# User Documentation for IDAS v3.0.0 (SUNDIALS v4.0.0)

Radu Serban, Cosmin Petra, and Alan C. Hindmarsh Center for Applied Scientific Computing Lawrence Livermore National Laboratory

December 7, 2018



#### DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

# Contents

Li	st of	Tables	ix		
Li	st of	Figures	xi		
1 Introduction					
	1.1	Changes from previous versions	2		
	1.2	Reading this User Guide	8		
	1.3	SUNDIALS Release License	9		
		1.3.1 Copyright Notices	9		
		1.3.1.1 SUNDIALS Copyright	9		
		1.3.1.2 ARKode Copyright	9		
		1.3.2 BSD License	10		
		1.0.2 DDD Electific	10		
<b>2</b>	Mat	thematical Considerations	11		
	2.1	IVP solution	11		
	2.2	Preconditioning	15		
	2.3	Rootfinding	16		
	2.4	Pure quadrature integration	17		
	2.5	Forward sensitivity analysis	17		
	2.0	2.5.1 Forward sensitivity methods	18		
		2.5.2 Selection of the absolute tolerances for sensitivity variables	19		
		2.5.2 Evaluation of the sensitivity right-hand side	19		
		<b>₽</b> 8	$\frac{19}{20}$		
	0.6	• • • • • • • • • • • • • • • • • • • •	$\frac{20}{20}$		
	2.6	Adjoint sensitivity analysis			
		2.6.1 Sensitivity of $G(p)$	20		
		2.6.2 Sensitivity of $g(T,p)$	21		
	~ <b>-</b>	2.6.3 Checkpointing scheme	22		
	2.7	Second-order sensitivity analysis	23		
0	<b>a</b>		0.5		
3		le Organization	<b>25</b>		
	3.1	SUNDIALS organization	25		
	3.2	IDAS organization	25		
4	Heir	ng IDAS for IVP Solution	31		
-	4.1	Access to library and header files	31		
	4.2	Data types	$\frac{31}{32}$		
	4.2	4.2.1 Floating point types	$\frac{32}{32}$		
		4.2.2 Integer types used for vector and matrix indices	$\frac{32}{32}$		
	4.9	~ · · ·			
	4.3	Header files	33		
	4.4	A skeleton of the user's main program	34		
	4.5	User-callable functions	37		
		4.5.1 IDAS initialization and deallocation functions	38		
		4.5.2 IDAS tolerance specification functions	38		

		4.5.3	Linear solver interface functions	40
		4.5.4	Nonlinear solver interface function	42
		4.5.5	Initial condition calculation function	42
		4.5.6	Rootfinding initialization function	43
		4.5.7	IDAS solver function	44
		4.5.8	Optional input functions	45
			4.5.8.1 Main solver optional input functions	45
			4.5.8.2 Linear solver interface optional input functions	51
			4.5.8.3 Initial condition calculation optional input functions	55
			4.5.8.4 Rootfinding optional input functions	57
		4.5.9	Interpolated output function	58
			Optional output functions	58
		1.0.10	4.5.10.1 SUNDIALS version information	58
			4.5.10.2 Main solver optional output functions	60
			4.5.10.3 Initial condition calculation optional output functions	66
			4.5.10.4 Rootfinding optional output functions	66
			4.5.10.5 IDALS linear solver interface optional output functions	67
		1511	IDAS reinitialization function	
	1.6			71
	4.6		supplied functions	72
		4.6.1	Residual function	72
		4.6.2	Error message handler function	73
		4.6.3	Error weight function	73
		4.6.4	Rootfinding function	74
		4.6.5	Jacobian construction (matrix-based linear solvers)	74
		4.6.6	Jacobian-vector product (matrix-free linear solvers)	76
		4.6.7	Jacobian-vector product setup (matrix-free linear solvers)	77
		4.6.8	Preconditioner solve (iterative linear solvers)	78
		4.6.9	Preconditioner setup (iterative linear solvers)	78
	4.7		ation of pure quadrature equations	79
		4.7.1	Quadrature initialization and deallocation functions	81
		4.7.2	IDAS solver function	82
		4.7.3	Quadrature extraction functions	82
		4.7.4	Optional inputs for quadrature integration	83
		4.7.5	Optional outputs for quadrature integration	84
		4.7.6	User-supplied function for quadrature integration	85
	4.8	A para	allel band-block-diagonal preconditioner module	86
5	$\mathbf{U}\mathbf{si}$	ng IDA	AS for Forward Sensitivity Analysis	93
	5.1	A skel	leton of the user's main program	93
	5.2	User-c	callable routines for forward sensitivity analysis	96
		5.2.1	Forward sensitivity initialization and deallocation functions	96
		5.2.2	Forward sensitivity tolerance specification functions	98
		5.2.3	Forward sensitivity nonlinear solver interface functions	100
		5.2.4	Forward sensitivity initial condition calculation function	101
		5.2.5	IDAS solver function	101
		5.2.6	Forward sensitivity extraction functions	101
		5.2.7	Optional inputs for forward sensitivity analysis	103
		5.2.8	Optional outputs for forward sensitivity analysis	105
		-	5.2.8.1 Main solver optional output functions	105
			5.2.8.2 Initial condition calculation optional output functions	108
	5.3	User-s	supplied routines for forward sensitivity analysis	108
	5.4		ation of quadrature equations depending on forward sensitivities	109
	J. 1	5.4.1	Sensitivity-dependent quadrature initialization and deallocation	111
		5.4.2	IDAS solver function	
		J. 1.		

	5.5	5.4.3 5.4.4 5.4.5 5.4.6 Note of	Sensitivity-dependent quadrature extraction functions	114 116 117
6	Usiı	ng IDA	AS for Adjoint Sensitivity Analysis	<b>121</b>
	6.1	_		121
	6.2		callable functions for adjoint sensitivity analysis	124
		6.2.1	Adjoint sensitivity allocation and deallocation functions	124
		6.2.2	Adjoint sensitivity optional input	
		6.2.3	Forward integration function	126
		6.2.4	Backward problem initialization functions	127
		6.2.5	Tolerance specification functions for backward problem	129
		6.2.6	Linear solver initialization functions for backward problem	
		6.2.7	Initial condition calculation functions for backward problem	
		6.2.8	Backward integration function	
		6.2.9	Optional input functions for the backward problem	
			6.2.9.1 Main solver optional input functions	
			6.2.9.2 Linear solver interface optional input functions	
		6.2.10	1 1	
			6.2.10.1 Main solver optional output functions	
		C O 11	6.2.10.2 Initial condition calculation optional output function	139
		0.2.11	Backward integration of quadrature equations	$\frac{140}{140}$
			6.2.11.2 Backward quadrature information functions	140 $141$
			6.2.11.3 Optional input/output functions for backward quadrature integration	141
	6.3	User-s	supplied functions for adjoint sensitivity analysis	142
	0.0	6.3.1	DAE residual for the backward problem	142
		6.3.2	DAE residual for the backward problem depending on the forward sensitivities	143
		6.3.3	Quadrature right-hand side for the backward problem	144
		6.3.4	Sensitivity-dependent quadrature right-hand side for the backward problem	145
		6.3.5	Jacobian construction for the backward problem (matrix-based linear solvers).	146
		6.3.6	Jacobian-vector product for the backward problem (matrix-free linear solvers).	148
		6.3.7	Jacobian-vector product setup for the backward problem (matrix-free linear solvers)	150
		6.3.8	Preconditioner solve for the backward problem (iterative linear solvers)	151
		6.3.9	Preconditioner setup for the backward problem (iterative linear solvers)	
	6.4	Using	the band-block-diagonal preconditioner for backward problems	
		6.4.1	Usage of IDABBDPRE for the backward problem	154
		6.4.2	User-supplied functions for IDABBDPRE	156
7	Des	criptio	on of the NVECTOR module	<b>159</b>
	7.1	_	CTOR functions used by IDAS	168
	7.2		VECTOR_SERIAL implementation	169
		7.2.1	NVECTOR_SERIAL accessor macros	169
		7.2.2	NVECTOR_SERIAL functions	170
		7.2.3	NVECTOR_SERIAL Fortran interfaces	173
	7.3	The N	VECTOR_PARALLEL implementation	173
		7.3.1	NVECTOR_PARALLEL accessor macros	174
		7.3.2	NVECTOR_PARALLEL functions	175
		7.3.3	NVECTOR_PARALLEL Fortran interfaces	178
	7.4		VECTOR_OPENMP implementation	
		7.4.1	NVECTOR_OPENMP accessor macros	178

		7.4.2 NVECTOR_OPENMP functions	<mark>79</mark>
		7.4.3 NVECTOR_OPENMP Fortran interfaces	82
	7.5	The NVECTOR_PTHREADS implementation	83
		7.5.1 NVECTOR_PTHREADS accessor macros	83
		7.5.2 NVECTOR_PTHREADS functions	84
		7.5.3 NVECTOR_PTHREADS Fortran interfaces	87
	7.6	The NVECTOR_PARHYP implementation	88
			88
	7.7		91
		7.7.1 NVECTOR_PETSC functions	
	7.8	The NVECTOR_CUDA implementation	
	•••	7.8.1 NVECTOR_CUDA functions	
	7.9		01
	1.0		$01 \\ 02$
	7 10		$02 \over 05$
	7.10	7.10.1 NVECTOR_OPENMPDEV accessor macros	
		7.10.1 NVECTOR_OPENMPDEV accessor macros	
	7 11		
	(.11	NVECTOR Examples	ΙŪ
8	Des	cription of the SUNMatrix module 21	15
Ü	8.1	SUNMatrix functions used by IDAS	
	8.2	The SUNMatrix_Dense implementation	
	0.2	8.2.1 SUNMatrix_Dense accessor macros	
		8.2.2 SUNMatrix_Dense functions	
		8.2.3 SUNMatrix_Dense Fortran interfaces	
	8.3	The SUNMatrix_Band implementation	
	0.5	8.3.1 SUNMatrix_Band accessor macros	
		8.3.2 SUNMatrix_Band functions	
	0.4	8.3.3 SUNMatrix_Band Fortran interfaces	
	8.4	The SUNMatrix_Sparse implementation	
		8.4.1 SUNMatrix_Sparse accessor macros	
		8.4.2 SUNMatrix_Sparse functions	
		8.4.3 SUNMatrix_Sparse Fortran interfaces	34
9	Des	cription of the SUNLinearSolver module 23	₹7
Ü	9.1	The SUNLinearSolver API	
	J.1	9.1.1 SUNLinearSolver core functions	
		9.1.2 SUNLinearSolver set functions	
		9.1.3 SUNLinearSolver get functions	
		9.1.4 Functions provided by SUNDIALS packages	
		9.1.5 SUNLinearSolver return codes	
	0.0		44 45
	9.2	1 0	45
	9.3	1 0	$\frac{45}{46}$
	0.4		46
	9.4		47
		00	48
			48
	9.5	r de la constant de l	49
		F	49
			49
			50
			51
	96	The SUNLinear Solver Band implementation 2:	52

	9.6.1	SUNLinearSolver_Band description	252
	9.6.2	SUNLinearSolver_Band functions	252
	9.6.3	SUNLinearSolver_Band Fortran interfaces	253
	9.6.4	SUNLinearSolver_Band content	254
9.7	The S	UNLinearSolver_LapackDense implementation	254
	9.7.1	SUNLinearSolver_LapackDense description	254
	9.7.2	SUNLinearSolver_LapackDense functions	
	9.7.3	SUNLinearSolver_LapackDense Fortran interfaces	255
	9.7.4	SUNLinearSolver_LapackDense content	
9.8	The S	UNLinearSolver_LapackBand implementation	
	9.8.1	SUNLinearSolver_LapackBand description	
	9.8.2	SUNLinearSolver_LapackBand functions	
	9.8.3	SUNLinearSolver_LapackBand Fortran interfaces	
	9.8.4	SUNLinearSolver_LapackBand content	
9.9		UNLinearSolver_KLU implementation	
0.0	9.9.1	-	259
	9.9.2	SUNLinear Solver_KLU functions	
	9.9.3	SUNLinear Solver_KLU Fortran interfaces	
	9.9.4	SUNLinear Solver-KLU content	
0.10		UNLinearSolver_SuperLUMT implementation	
3.10		SUNLinearSolver_SuperLUMT description	
		SUNLinear Solver-Super LUMT functions	
		SUNLinear Solver Super LUMT Fortran interfaces	
		SUNLinear Solver_SuperLUMT content	
0.11		UNLinearSolver_SPGMR implementation	
9.11		SUNLinearSolver_SPGMR description	
		SUNLinearSolver_SPGMR functions	
		SUNLinearSolver_SPGMR Fortran interfaces	
		SUNLinearSolver_SPGMR content	
0.10			
9.12		UNLinearSolver_SPFGMR implementation	
		SUNLinear Solver_SPFGMR description	
		SUNLinear Solver_SPFGMR functions	
		SUNLinearSolver_SPFGMR Fortran interfaces	
0.10		SUNLinearSolver_SPFGMR content	
9.13		UNLinearSolver_SPBCGS implementation	
		SUNLinearSolver_SPBCGS description	
			282
		SUNLinearSolver_SPBCGS Fortran interfaces	284
		SUNLinearSolver_SPBCGS content	286
9.14		UNLinearSolver_SPTFQMR implementation	287
		SUNLinearSolver_SPTFQMR description	287
		SUNLinearSolver_SPTFQMR functions	288
		SUNLinearSolver_SPTFQMR Fortran interfaces	289
		SUNLinearSolver_SPTFQMR content	291
9.15		UNLinearSolver_PCG implementation	292
		SUNLinearSolver_PCG description	293
		SUNLinearSolver_PCG functions	294
			295
	9.15.4	SUNLinearSolver_PCG content	297
9 16	SUNL	inearSolver Examples	298

10		cription of the SUNNonlinearSolver module	<b>301</b>				
	10.1	The SUNNonlinear Solver API	301				
		10.1.1 SUNNonlinearSolver core functions					
		10.1.2 SUNNonlinearSolver set functions					
		10.1.3 SUNNonlinearSolver get functions	304				
		10.1.4 Functions provided by SUNDIALS integrators	305				
		10.1.5 SUNNonlinearSolver return codes	307				
		10.1.6 The generic SUNNonlinearSolver module	307				
		10.1.7 Usage with sensitivity enabled integrators	308				
		10.1.8 Implementing a Custom SUNNonlinearSolver Module	310				
	10.2	The SUNNonlinearSolver_Newton implementation	310				
		10.2.1 SUNNonlinearSolver_Newton description	310				
		10.2.2 SUNNonlinearSolver_Newton functions	311				
		10.2.3 SUNNonlinearSolver_Newton Fortran interfaces	312				
		10.2.4 SUNNonlinearSolver_Newton content	313				
	10.3	The SUNNonlinearSolver_FixedPoint implementation	313				
		10.3.1 SUNNonlinearSolver_FixedPoint description	313				
		10.3.2 SUNNonlinearSolver_FixedPoint functions					
		10.3.3 SUNNonlinearSolver_FixedPoint Fortran interfaces					
		$10.3.4 \;\; SUNN on linear Solver\_Fixed Point \; content \;\; . \;\; . \;\; . \;\; . \;\; . \;\; . \;\; . \;\;$	316				
A	SUN	NDIALS Package Installation Procedure	319				
		CMake-based installation	320				
		A.1.1 Configuring, building, and installing on Unix-like systems	320				
		A.1.2 Configuration options (Unix/Linux)					
		A.1.3 Configuration examples					
		A.1.4 Working with external Libraries					
		A.1.5 Testing the build and installation					
	A.2	Building and Running Examples					
		Configuring, building, and installing on Windows					
		Installed libraries and exported header files					
В	IDA	AS Constants	339				
_		IDAS input constants					
		IDAS output constants					
Bi	343						
	dor	· ·	9 4 7				

# List of Tables

4.1 4.2 4.3	SUNDIALS linear solver interfaces and vector implementations that can be used for each.  Optional inputs for IDAS and IDALS	30 40 59
5.1 5.2	Forward sensitivity optional inputs	
7.1 7.2 7.3 7.4 7.5	Vector Identifications associated with vector kernels supplied with SUNDIALS.  Description of the NVECTOR operations.  Description of the NVECTOR fused operations.  Description of the NVECTOR vector array operations.  List of vector functions usage by IDAS code modules.	165 165 166
8.1 8.2 8.3 8.4	Identifiers associated with matrix kernels supplied with SUNDIALS	216
9.1 9.2	Description of the SUNLinearSolver error codes	244 245
9.3 10.1	List of linear solver function usage in the IDALS interface	30'
A.1	SUNDIALS libraries and header files	334

# List of Figures

2.1	Illustration of the checkpointing algorithm for generation of the forward solution during the integration of the adjoint system
	High-level diagram of the SUNDIALS suite
3.2	Organization of the SUNDIALS suite
3.3	Overall structure diagram of the IDA package
8.1	Diagram of the storage for a SUNMATRIX_BAND object
8.2	Diagram of the storage for a compressed-sparse-column matrix
A.1	Initial <i>ccmake</i> configuration screen
A.2	Changing the <i>instdir</i>

# Chapter 1

# Introduction

IDAS is part of a software family called SUNDIALS: SUite of Nonlinear and DIfferential/ALgebraic equation Solvers [26]. This suite consists of CVODE, ARKODE, KINSOL, and IDA, and variants of these with sensitivity analysis capabilities, CVODES and IDAS.

IDAS is a general purpose solver for the initial value problem (IVP) for systems of differential-algebraic equations (DAEs). The name IDAS stands for Implicit Differential-Algebraic solver with Sensitivity capabilities. IDAS is an extension of the IDA solver within SUNDIALS, itself based on DASPK [7, 8]; however, like all SUNDIALS solvers, IDAS is written in ANSI-standard C rather than FORTRAN77. Its most notable features are that, (1) in the solution of the underlying nonlinear system at each time step, it offers a choice of Newton/direct methods and a choice of Inexact Newton/Krylov (iterative) methods; (2) it is written in a data-independent manner in that it acts on generic vectors and matrices without any assumptions on the underlying organization of the data; and (3) it provides a flexible, extensible framework for sensitivity analysis, using either forward or adjoint methods. Thus IDAS shares significant modules previously written within CASC at LLNL to support the ordinary differential equation (ODE) solvers CVODE [27, 15] and PVODE [11, 12], the DAE solver IDA [30] on which IDAS is based, the sensitivity-enabled ODE solver CVODES [28, 42], and also the nonlinear system solver KINSOL [16].

At present, IDAS may utilize a variety of Krylov methods provided in SUNDIALS that can be used in conjuction with Newton iteration: these include the GMRES (Generalized Minimal RESidual) [41], FGMRES (Flexible Generalized Minimum RESidual) [40], Bi-CGStab (Bi-Conjugate Gradient Stabilized) [44], TFQMR (Transpose-Free Quasi-Minimal Residual) [23], and PCG (Preconditioned Conjugate Gradient) [24] linear iterative methods. As Krylov methods, these require little matrix storage for solving the Newton equations as compared to direct methods. However, the algorithms allow for a user-supplied preconditioner matrix, and, for most problems, preconditioning is essential for an efficient solution.

For very large DAE systems, the Krylov methods are preferable over direct linear solver methods, and are often the only feasible choice. Among the Krylov methods in SUNDIALS, we recommend GMRES as the best overall choice. However, users are encouraged to compare all options, especially if encountering convergence failures with GMRES. Bi-CGFStab and TFQMR have an advantage in storage requirements, in that the number of workspace vectors they require is fixed, while that number for GMRES depends on the desired Krylov subspace size. FGMRES has an advantage in that it is designed to support preconditioners that vary between iterations (e.g. iterative methods). PCG exhibits rapid convergence and minimal workspace vectors, but only works for symmetric linear systems.

IDAS is written with a functionality that is a superset of that of IDA. Sensitivity analysis capabilities, both forward and adjoint, have been added to the main integrator. Enabling forward sensitivity computations in IDAS will result in the code integrating the so-called *sensitivity equations* simultaneously with the original IVP, yielding both the solution and its sensitivity with respect to parameters in the model. Adjoint sensitivity analysis, most useful when the gradients of relatively few functionals of the solution with respect to many parameters are sought, involves integration of the original IVP

2 Introduction

forward in time followed by the integration of the so-called *adjoint equations* backward in time. IDAS provides the infrastructure needed to integrate any final-condition ODE dependent on the solution of the original IVP (in particular the adjoint system).

There are several motivations for choosing the C language for IDAS. First, a general movement away from FORTRAN and toward C in scientific computing was apparent. Second, the pointer, structure, and dynamic memory allocation features in C are extremely useful in software of this complexity, with the great variety of method options offered. Finally, we prefer C over C++ for IDAS because of the wider availability of C compilers, the potentially greater efficiency of C, and the greater ease of interfacing the solver to applications written in extended FORTRAN.

### 1.1 Changes from previous versions

#### Changes in v3.0.0

IDAS' previous direct and iterative linear solver interfaces, IDADLS and IDASPILS, have been merged into a single unified linear solver interface, IDALS, to support any valid SUNLINSOL module. This includes the "DIRECT" and "ITERATIVE" types as well as the new "MATRIX\_ITERATIVE" type. Details regarding how IDALS utilizes linear solvers of each type as well as discussion regarding intended use cases for user-supplied SUNLINSOL implementations are included in Chapter 9. All IDAS example programs and the standalone linear solver examples have been updated to use the unified linear solver interface.

The unified interface for the new IDALS module is very similar to the previous IDADLS and IDASPILS interfaces. To minimize challenges in user migration to the new names, the previous C routine names may still be used; these will be deprecated in future releases, so we recommend that users migrate to the new names soon.

The names of all constructor routines for SUNDIALS-provided SUNLINSOL implementations have been updated to follow the naming convention SUNLinSol\_\* where \* is the name of the linear solver. The new names are SUNLinSol\_Band, SUNLinSol\_Dense, SUNLinSol\_KLU, SUNLinSol\_LapackBand, SUNLinSol\_LapackDense, SUNLinSol\_PCG, SUNLinSol\_SPBCGS, SUNLinSol\_SPFGMR, SUNLinSol\_SPGMR, SUNLinSol\_SPFGMR, and SUNLinSol\_SuperLUMT. Solver-specific "set" routine names have been similarly standardized. To minimize challenges in user migration to the new names, the previous routine names may still be used; these will be deprecated in future releases, so we recommend that users migrate to the new names soon. All IDAS example programs and the standalone linear solver examples have been updated to use the new naming convention.

The SUNBandMatrix constructor has been simplified to remove the storage upper bandwidth argument.

SUNDIALS integrators have been updated to utilize generic nonlinear solver modules defined through the SUNNONLINSOL API. This API will ease the addition of new nonlinear solver options and allow for external or user-supplied nonlinear solvers. The SUNNONLINSOL API and SUNDIALS provided modules are described in Chapter 10 and follow the same object oriented design and implementation used by the NVECTOR, SUNMATRIX, and SUNLINSOL modules. Currently two SUNNONLINSOL implementations are provided, SUNNONLINSOL\_NEWTON and SUNNONLINSOL\_FIXEDPOINT. These replicate the previous integrator specific implementations of a Newton iteration and a fixed-point iteration (previously referred to as a functional iteration), respectively. Note the SUNNONLINSOL\_FIXEDPOINT module can optionally utilize Anderson's method to accelerate convergence. Example programs using each of these nonlinear solver modules in a standalone manner have been added and all IDAS example programs have been updated to use generic SUNNONLINSOL modules.

By default IDAS uses the SUNNONLINSOL\_NEWTON module. Since IDAS previously only used an internal implementation of a Newton iteration no changes are required to user programs and functions for setting the nonlinear solver options (e.g., IDASetMaxNonlinIters) or getting nonlinear solver statistics (e.g., IDAGetNumNonlinSolvIters) remain unchanged and internally call generic SUNNON-LINSOL functions as needed. While SUNDIALS includes a fixed-point nonlinear solver module, it is not currently supported in IDAS. For details on attaching a user-supplied nonlinear solver to IDAS see Chapter 4, 5, and 6.

Three fused vector operations and seven vector array operations have been added to the NVECTOR API. These *optional* operations are disabled by default and may be activated by calling vector specific routines after creating an NVECTOR (see Chapter 7 for more details). The new operations are intended to increase data reuse in vector operations, reduce parallel communication on distributed memory systems, and lower the number of kernel launches on systems with accelerators. The fused operations are N\_VLinearCombination, N\_VScaleAddMulti, and N\_VDotProdMulti and the vector array operations are N\_VLinearCombinationVectorArray, N\_VScaleVectorArray, N\_VConstVectorArray, N\_VWrmsNormVectorArray, N\_VWrmsNormMaskVectorArray, N\_VScaleAddMultiVectorArray, and N\_VLinearCombinationVectorArray. If an NVECTOR implementation defines any of these operations as NULL, then standard NVECTOR operations will automatically be called as necessary to complete the computation.

Multiple updates to NVECTOR\_CUDA were made:

- Changed N\_VGetLength\_Cuda to return the global vector length instead of the local vector length.
- Added N\_VGetLocalLength\_Cuda to return the local vector length.
- Added N\_VGetMPIComm\_Cuda to return the MPI communicator used.
- Removed the accessor functions in the namespace suncudavec.
- Changed the N\_VMake\_Cuda function to take a host data pointer and a device data pointer instead of an N\_VectorContent\_Cuda object.
- Added the ability to set the cudaStream\_t used for execution of the NVECTOR\_CUDA kernels. See the function N\_VSetCudaStreams\_Cuda.
- Added N\_VNewManaged\_Cuda, N\_VMakeManaged\_Cuda, and N\_VIsManagedMemory\_Cuda functions to accommodate using managed memory with the NVECTOR\_CUDA.

Multiple changes to NVECTOR\_RAJA were made:

- Changed N\_VGetLength\_Raja to return the global vector length instead of the local vector length.
- Added N\_VGetLocalLength\_Raja to return the local vector length.
- Added N\_VGetMPIComm\_Raja to return the MPI communicator used.
- Removed the accessor functions in the namespace suncudavec.

A new NVECTOR implementation for leveraging OpenMP 4.5+ device offloading has been added, NVECTOR\_OPENMPDEV. See §7.10 for more details.

#### Changes in v2.2.1

The changes in this minor release include the following:

- Fixed a bug in the CUDA NVECTOR where the N\_VInvTest operation could write beyond the allocated vector data.
- Fixed library installation path for multiarch systems. This fix changes the default library installation path to CMAKE\_INSTALL\_PREFIX/CMAKE\_INSTALL\_LIBDIR from CMAKE\_INSTALL\_PREFIX/lib. CMAKE\_INSTALL\_LIBDIR is automatically set, but is available as a CMake option that can modified.

4 Introduction

#### Changes in v2.2.0

Fixed a bug in IDAS where the saved residual value used in the nonlinear solve for consistent initial conditions was passed as temporary workspace and could be overwritten.

Fixed a thread-safety issue when using ajdoint sensitivity analysis.

Fixed a problem with setting sunindextype which would occur with some compilers (e.g. arm-clang) that did not define \_\_STDC\_VERSION\_\_.

Added hybrid MPI/CUDA and MPI/RAJA vectors to allow use of more than one MPI rank when using a GPU system. The vectors assume one GPU device per MPI rank.

Changed the name of the RAJA NVECTOR library to libsundials\_nveccudaraja.lib from libsundials\_nvecraja.lib to better reflect that we only support CUDA as a backend for RAJA currently.

Several changes were made to the build system:

- CMake 3.1.3 is now the minimum required CMake version.
- Deprecate the behavior of the SUNDIALS\_INDEX\_TYPE CMake option and added the SUNDIALS\_INDEX\_SIZE CMake option to select the sunindextype integer size.
- The native CMake FindMPI module is now used to locate an MPI installation.
- If MPI is enabled and MPI compiler wrappers are not set, the build system will check if CMAKE\_<language>\_COMPILER can compile MPI programs before trying to locate and use an MPI installation.
- The previous options for setting MPI compiler wrappers and the executable for running MPI programs have been have been depreated. The new options that align with those used in native CMake FindMPI module are MPI\_C\_COMPILER, MPI\_CXX\_COMPILER, MPI\_Fortran\_COMPILER, and MPIEXEC\_EXECUTABLE.
- When a Fortran name-mangling scheme is needed (e.g., LAPACK\_ENABLE is ON) the build system will infer the scheme from the Fortran compiler. If a Fortran compiler is not available or the inferred or default scheme needs to be overridden, the advanced options SUNDIALS\_F77\_FUNC\_CASE and SUNDIALS\_F77\_FUNC\_UNDERSCORES can be used to manually set the name-mangling scheme and bypass trying to infer the scheme.
- Parts of the main CMakeLists.txt file were moved to new files in the src and example directories to make the CMake configuration file structure more modular.

#### Changes in v2.1.2

The changes in this minor release include the following:

- Updated the minimum required version of CMake to 2.8.12 and enabled using rpath by default to locate shared libraries on OSX.
- Fixed Windows specific problem where sunindextype was not correctly defined when using 64-bit integers for the SUNDIALS index type. On Windows sunindextype is now defined as the MSVC basic type \_\_int64.
- Added sparse SUNMatrix "Reallocate" routine to allow specification of the nonzero storage.

- Updated the KLU sunlinsol module to set constants for the two reinitialization types, and fixed a bug in the full reinitialization approach where the sparse SUNMatrix pointer would go out of scope on some architectures.
- Updated the "ScaleAdd" and "ScaleAddI" implementations in the sparse SUNMatrix module to more optimally handle the case where the target matrix contained sufficient storage for the sum, but had the wrong sparsity pattern. The sum now occurs in-place, by performing the sum backwards in the existing storage. However, it is still more efficient if the user-supplied Jacobian routine allocates storage for the sum  $I + \gamma J$  manually (with zero entries if needed).
- Changed the LICENSE install path to instdir/include/sundials.

#### Changes in v2.1.1

The changes in this minor release include the following:

- Fixed a potential memory leak in the SPGMR and SPFGMR linear solvers: if "Initialize" was called multiple times then the solver memory was reallocated (without being freed).
- Updated KLU SUNLinearSolver module to use a typedef for the precision-specific solve function to be used (to avoid compiler warnings).
- Added missing typecasts for some (void\*) pointers (again, to avoid compiler warnings).
- Bugfix in sunmatrix\_sparse.c where we had used int instead of sunindextype in one location.
- Added missing #include <stdio.h> in NVECTOR and SUNMATRIX header files.
- Added missing prototype for IDASpilsGetNumJTSetupEvals.
- Fixed an indexing bug in the CUDA NVECTOR implementation of N\_VWrmsNormMask and revised the RAJA NVECTOR implementation of N\_VWrmsNormMask to work with mask arrays using values other than zero or one. Replaced double with realtype in the RAJA vector test functions.

In addition to the changes above, minor corrections were also made to the example programs, build system, and user documentation.

#### Changes in v2.1.0

Added NVECTOR print functions that write vector data to a specified file (e.g., N\_VPrintFile\_Serial). Added make test and make test\_install options to the build system for testing SUNDIALS after building with make and installing with make install respectively.

#### Changes in v2.0.0

All interfaces to matrix structures and linear solvers have been reworked, and all example programs have been updated. The goal of the redesign of these interfaces was to provide more encapsulation and to ease interfacing of custom linear solvers and interoperability with linear solver libraries. Specific changes include:

- Added generic Sunmatrix module with three provided implementations: dense, banded and sparse. These replicate previous Sundials Dls and Sls matrix structures in a single object-oriented API.
- Added example problems demonstrating use of generic Sunmatrix modules.
- Added generic SUNLinearSolver module with eleven provided implementations: SUNDIALS native dense, SUNDIALS native banded, LAPACK dense, LAPACK band, KLU, SuperLU\_MT, SPGMR, SPBCGS, SPTFQMR, SPFGMR, and PCG. These replicate previous SUNDIALS generic linear solvers in a single object-oriented API.

6 Introduction

- Added example problems demonstrating use of generic SUNLinearSolver modules.
- Expanded package-provided direct linear solver (Dls) interfaces and scaled, preconditioned, iterative linear solver (Spils) interfaces to utilize generic SUNMATRIX and SUNLinearSolver objects.
- Removed package-specific, linear solver-specific, solver modules (e.g. CVDENSE, KINBAND, IDAKLU, ARKSPGMR) since their functionality is entirely replicated by the generic Dls/Spils interfaces and SUNLinearSolver/SUNMATRIX modules. The exception is CVDIAG, a diagonal approximate Jacobian solver available to CVODE and CVODES.
- Converted all SUNDIALS example problems and files to utilize the new generic SUNMATRIX and SUNLinearSolver objects, along with updated Dls and Spils linear solver interfaces.
- Added Spils interface routines to ARKODE, CVODE, CVODES, IDA, and IDAS to allow specification of a user-provided "JTSetup" routine. This change supports users who wish to set up data structures for the user-provided Jacobian-times-vector ("JTimes") routine, and where the cost of one JTSetup setup per Newton iteration can be amortized between multiple JTimes calls.

Two additional NVECTOR implementations were added – one for CUDA and one for RAJA vectors. These vectors are supplied to provide very basic support for running on GPU architectures. Users are advised that these vectors both move all data to the GPU device upon construction, and speedup will only be realized if the user also conducts the right-hand-side function evaluation on the device. In addition, these vectors assume the problem fits on one GPU. Further information about RAJA, users are referred to the web site, https://software.llnl.gov/RAJA/. These additions are accompanied by additions to various interface functions and to user documentation.

All indices for data structures were updated to a new sunindextype that can be configured to be a 32- or 64-bit integer data index type. sunindextype is defined to be int32\_t or int64\_t when portable types are supported, otherwise it is defined as int or long int. The Fortran interfaces continue to use long int for indices, except for their sparse matrix interface that now uses the new sunindextype. This new flexible capability for index types includes interfaces to PETSc, hypre, SuperLU\_MT, and KLU with either 32-bit or 64-bit capabilities depending how the user configures SUNDIALS.

To avoid potential namespace conflicts, the macros defining booleantype values TRUE and FALSE have been changed to SUNTRUE and SUNFALSE respectively.

Temporary vectors were removed from preconditioner setup and solve routines for all packages. It is assumed that all necessary data for user-provided preconditioner operations will be allocated and stored in user-provided data structures.

The file include/sundials\_fconfig.h was added. This file contains SUNDIALS type information for use in Fortran programs.

The build system was expanded to support many of the xSDK-compliant keys. The xSDK is a movement in scientific software to provide a foundation for the rapid and efficient production of high-quality, sustainable extreme-scale scientific applications. More information can be found at, https://xsdk.info.

Added functions SUNDIALSGetVersion and SUNDIALSGetVersionNumber to get SUNDIALS release version information at runtime.

In addition, numerous changes were made to the build system. These include the addition of separate BLAS\_ENABLE and BLAS\_LIBRARIES CMake variables, additional error checking during CMake configuration, minor bug fixes, and renaming CMake options to enable/disable examples for greater clarity and an added option to enable/disable Fortran 77 examples. These changes included changing EXAMPLES\_ENABLE to EXAMPLES\_ENABLE\_CXX, changing F90\_ENABLE to EXAMPLES\_ENABLE\_F90, and adding an EXAMPLES\_ENABLE\_F77 option.

A bug fix was done to add a missing prototype for IDASetMaxBacksIC in ida.h.

Corrections and additions were made to the examples, to installation-related files, and to the user documentation.

#### Changes in v1.3.0

Two additional NVECTOR implementations were added – one for Hypre (parallel) ParVector vectors, and one for PETSc vectors. These additions are accompanied by additions to various interface functions and to user documentation.

Each NVECTOR module now includes a function, N\_VGetVectorID, that returns the NVECTOR module name.

An optional input function was added to set a maximum number of linesearch backtracks in the initial condition calculation, and four user-callable functions were added to support the use of LAPACK linear solvers in solving backward problems for adjoint sensitivity analysis.

For each linear solver, the various solver performance counters are now initialized to 0 in both the solver specification function and in solver linit function. This ensures that these solver counters are initialized upon linear solver instantiation as well as at the beginning of the problem solution.

A bug in for-loop indices was fixed in IDAAckpntAllocVectors. A bug was fixed in the interpolation functions used in solving backward problems.

A memory leak was fixed in the banded preconditioner interface. In addition, updates were done to return integers from linear solver and preconditioner 'free' functions.

In interpolation routines for backward problems, added logic to bypass sensitivity interpolation if input sensitivity argument is NULL.

The Krylov linear solver Bi-CGstab was enhanced by removing a redundant dot product. Various additions and corrections were made to the interfaces to the sparse solvers KLU and SuperLU\_MT, including support for CSR format when using KLU.

New examples were added for use of the OpenMP vector and for use of sparse direct solvers within sensitivity integrations.

Minor corrections and additions were made to the IDAS solver, to the examples, to installation-related files, and to the user documentation.

#### Changes in v1.2.0

Two major additions were made to the linear system solvers that are available for use with the IDAS solver. First, in the serial case, an interface to the sparse direct solver KLU was added. Second, an interface to SuperLU\_MT, the multi-threaded version of SuperLU, was added as a thread-parallel sparse direct solver option, to be used with the serial version of the NVECTOR module. As part of these additions, a sparse matrix (CSC format) structure was added to IDAS.

Otherwise, only relatively minor modifications were made to IDAS:

In IDARootfind, a minor bug was corrected, where the input array rootdir was ignored, and a line was added to break out of root-search loop if the initial interval size is below the tolerance ttol.

In IDALapackBand, the line smu = MIN(N-1,mu+ml) was changed to smu = mu + ml to correct an illegal input error for DGBTRF/DGBTRS.

An option was added in the case of Adjoint Sensitivity Analysis with dense or banded Jacobian: With a call to IDADlsSetDenseJacFnBS or IDADlsSetBandJacFnBS, the user can specify a user-supplied Jacobian function of type IDADls\*\*\*JacFnBS, for the case where the backward problem depends on the forward sensitivities.

A minor bug was fixed regarding the testing of the input tstop on the first call to IDASolve.

For the Adjoint Sensitivity Analysis case in which the backward problem depends on the forward sensitivities, options have been added to allow for user-supplied pset, psolve, and jtimes functions.

In order to avoid possible name conflicts, the mathematical macro and function names MIN, MAX, SQR, RAbs, RSqrt, RExp, RPowerI, and RPowerR were changed to SUNMIN, SUNMAX, SUNSQR, SUNRabs, SUNRsqrt, SUNRexp, SRpowerI, and SUNRpowerR, respectively. These names occur in both the solver and in various example programs.

In the User Guide, a paragraph was added in Section 6.2.1 on IDAAdjReInit, and a paragraph was added in Section 6.2.9 on IDAGetAdjY.

Two new NVECTOR modules have been added for thread-parallel computing environments — one for OpenMP, denoted NVECTOR\_OPENMP, and one for Pthreads, denoted NVECTOR\_PTHREADS.

8 Introduction

With this version of SUNDIALS, support and documentation of the Autotools mode of installation is being dropped, in favor of the CMake mode, which is considered more widely portable.

#### Changes in v1.1.0

One significant design change was made with this release: The problem size and its relatives, bandwidth parameters, related internal indices, pivot arrays, and the optional output lsflag have all been changed from type int to type long int, except for the problem size and bandwidths in user calls to routines specifying BLAS/LAPACK routines for the dense/band linear solvers. The function NewIntArray is replaced by a pair NewIntArray/NewLintArray, for int and long int arrays, respectively. In a minor change to the user interface, the type of the index which in IDAS was changed from long int to int.

Errors in the logic for the integration of backward problems were identified and fixed.

A large number of minor errors have been fixed. Among these are the following: A missing vector pointer setting was added in IDASensLineSrch. In IDACompleteStep, conditionals around lines loading a new column of three auxiliary divided difference arrays, for a possible order increase, were fixed. After the solver memory is created, it is set to zero before being filled. In each linear solver interface function, the linear solver memory is freed on an error return, and the \*\*Free function now includes a line setting to NULL the main memory pointer to the linear solver memory. A memory leak was fixed in two of the IDASp\*\*\*Free functions. In the rootfinding functions IDARcheck1/IDARcheck2, when an exact zero is found, the array glo of g values at the left endpoint is adjusted, instead of shifting the t location tlo slightly. In the installation files, we modified the treatment of the macro SUNDIALS\_USE\_GENERIC\_MATH, so that the parameter GENERIC\_MATH\_LIB is either defined (with no value) or not defined.

### 1.2 Reading this User Guide

The structure of this document is as follows:

- In Chapter 2, we give short descriptions of the numerical methods implemented by IDAS for the solution of initial value problems for systems of DAEs, continue with short descriptions of preconditioning (§2.2) and rootfinding (§2.3), and then give an overview of the mathematical aspects of sensitivity analysis, both forward (§2.5) and adjoint (§2.6).
- The following chapter describes the structure of the SUNDIALS suite of solvers (§3.1) and the software organization of the IDAS solver (§3.2).
- Chapter 4 is the main usage document for IDAS for simulation applications. It includes a complete description of the user interface for the integration of DAE initial value problems. Readers that are not interested in using IDAS for sensitivity analysis can then skip the next two chapters.
- Chapter 5 describes the usage of IDAS for forward sensitivity analysis as an extension of its IVP integration capabilities. We begin with a skeleton of the user main program, with emphasis on the steps that are required in addition to those already described in Chapter 4. Following that we provide detailed descriptions of the user-callable interface routines specific to forward sensitivity analysis and of the additional optional user-defined routines.
- Chapter 6 describes the usage of IDAS for adjoint sensitivity analysis. We begin by describing the IDAS checkpointing implementation for interpolation of the original IVP solution during integration of the adjoint system backward in time, and with an overview of a user's main program. Following that we provide complete descriptions of the user-callable interface routines for adjoint sensitivity analysis as well as descriptions of the required additional user-defined routines.
- Chapter 7 gives a brief overview of the generic NVECTOR module shared amongst the various components of SUNDIALS, as well as details on the NVECTOR implementations provided with SUNDIALS.

- Chapter 8 gives a brief overview of the generic SUNMATRIX module shared among the various components of SUNDIALS, and details on the SUNMATRIX implementations provided with SUNDIALS: a dense implementation (§8.2), a banded implementation (§8.3) and a sparse implementation (§8.4).
- Chapter 9 gives a brief overview of the generic Sunlinsol module shared among the various components of Sundials. This chapter contains details on the Sunlinsol implementations provided with Sundials. The chapter also contains details on the Sunlinsol implementations provided with Sundials that interface with external linear solver libraries.
- Chapter 10 describes the SUNNONLINSOL API and nonlinear solver implementations shared among the various components of SUNDIALS.
- Finally, in the appendices, we provide detailed instructions for the installation of IDAS, within the structure of SUNDIALS (Appendix A), as well as a list of all the constants used for input to and output from IDAS functions (Appendix B).

Finally, the reader should be aware of the following notational conventions in this user guide: program listings and identifiers (such as IDAInit) within textual explanations appear in typewriter type style; fields in C structures (such as content) appear in italics; and packages or modules, such as IDALS, are written in all capitals. Usage and installation instructions that constitute important warnings are marked with a triangular symbol in the margin.



#### 1.3 SUNDIALS Release License

The SUNDIALS packages are released open source, under a BSD license. The only requirements of the BSD license are preservation of copyright and a standard disclaimer of liability. Our Copyright notice is below along with the license.

\*\*PLEASE NOTE\*\* If you are using SUNDIALS with any third party libraries linked in (e.g., LaPACK, KLU, SuperLU\_MT, PETSC, or hypre), be sure to review the respective license of the package as that license may have more restrictive terms than the SUNDIALS license. For example, if someone builds SUNDIALS with a statically linked KLU, the build is subject to terms of the LGPL license (which is what KLU is released with) and \*not\* the SUNDIALS BSD license anymore.

### 1.3.1 Copyright Notices

All SUNDIALS packages except ARKode are subject to the following Copyright notice.

#### 1.3.1.1 SUNDIALS Copyright

Copyright (c) 2002-2016, Lawrence Livermore National Security. Produced at the Lawrence Livermore National Laboratory. Written by A.C. Hindmarsh, D.R. Reynolds, R. Serban, C.S. Woodward, S.D. Cohen, A.G. Taylor, S. Peles, L.E. Banks, and D. Shumaker.

UCRL-CODE-155951 (CVODE)

UCRL-CODE-155950 (CVODES)

UCRL-CODE-155952 (IDA)

UCRL-CODE-237203 (IDAS)

LLNL-CODE-665877 (KINSOL)

All rights reserved.

#### 1.3.1.2 ARKode Copyright

ARKode is subject to the following joint Copyright notice. Copyright (c) 2015-2016, Southern Methodist University and Lawrence Livermore National Security Written by D.R. Reynolds, D.J. Gardner, A.C. Hindmarsh, C.S. Woodward, and J.M. Sexton.



10 Introduction

LLNL-CODE-667205 (ARKODE) All rights reserved.

#### 1.3.2 BSD License

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the disclaimer below.
- 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the disclaimer (as noted below) in the documentation and/or other materials provided with the distribution
- 3. Neither the name of the LLNS/LLNL nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL LAWRENCE LIVERMORE NATIONAL SECURITY, LLC, THE U.S. DEPARTMENT OF ENERGY OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

#### Additional BSD Notice

- 1. This notice is required to be provided under our contract with the U.S. Department of Energy (DOE). This work was produced at Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344 with the DOE.
- 2. Neither the United States Government nor Lawrence Livermore National Security, LLC nor any of their employees, makes any warranty, express or implied, or assumes any liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights.
- 3. Also, reference herein to any specific commercial products, process, or services by trade name, trademark, manufacturer or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

# Chapter 2

# Mathematical Considerations

IDAS solves the initial-value problem (IVP) for a DAE system of the general form

$$F(t, y, \dot{y}) = 0, \quad y(t_0) = y_0, \ \dot{y}(t_0) = \dot{y}_0,$$
 (2.1)

where y,  $\dot{y}$ , and F are vectors in  $\mathbf{R}^N$ , t is the independent variable,  $\dot{y} = dy/dt$ , and initial values  $y_0$ ,  $\dot{y}_0$  are given. (Often t is time, but it certainly need not be.)

Additionally, if (2.1) depends on some parameters  $p \in \mathbf{R}^{N_p}$ , i.e.

$$F(t, y, \dot{y}, p) = 0$$
  

$$y(t_0) = y_0(p), \ \dot{y}(t_0) = \dot{y}_0(p),$$
(2.2)

IDAS can also compute first order derivative information, performing either forward sensitivity analysis or adjoint sensitivity analysis. In the first case, IDAS computes the sensitivities of the solution with respect to the parameters p, while in the second case, IDAS computes the gradient of a derived function with respect to the parameters p.

#### 2.1 IVP solution

Prior to integrating a DAE initial-value problem, an important requirement is that the pair of vectors  $y_0$  and  $\dot{y}_0$  are both initialized to satisfy the DAE residual  $F(t_0,y_0,\dot{y}_0)=0$ . For a class of problems that includes so-called semi-explicit index-one systems, IDAS provides a routine that computes consistent initial conditions from a user's initial guess [8]. For this, the user must identify sub-vectors of y (not necessarily contiguous), denoted  $y_d$  and  $y_a$ , which are its differential and algebraic parts, respectively, such that F depends on  $\dot{y}_d$  but not on any components of  $\dot{y}_a$ . The assumption that the system is "index one" means that for a given t and  $y_d$ , the system  $F(t,y,\dot{y})=0$  defines  $y_a$  uniquely. In this case, a solver within IDAS computes  $y_a$  and  $\dot{y}_d$  at  $t=t_0$ , given  $y_d$  and an initial guess for  $y_a$ . A second available option with this solver also computes all of  $y(t_0)$  given  $\dot{y}(t_0)$ ; this is intended mainly for quasi-steady-state problems, where  $\dot{y}(t_0)=0$  is given. In both cases, IDA solves the system  $F(t_0,y_0,\dot{y}_0)=0$  for the unknown components of  $y_0$  and  $\dot{y}_0$ , using Newton iteration augmented with a line search global strategy. In doing this, it makes use of the existing machinery that is to be used for solving the linear systems during the integration, in combination with certain tricks involving the step size (which is set artificially for this calculation). For problems that do not fall into either of these categories, the user is responsible for passing consistent values, or risks failure in the numerical integration.

The integration method used in IDAS is the variable-order, variable-coefficient BDF (Backward Differentiation Formula), in fixed-leading-coefficient form [4]. The method order ranges from 1 to 5, with the BDF of order q given by the multistep formula

$$\sum_{i=0}^{q} \alpha_{n,i} y_{n-i} = h_n \dot{y}_n \,, \tag{2.3}$$

where  $y_n$  and  $\dot{y}_n$  are the computed approximations to  $y(t_n)$  and  $\dot{y}(t_n)$ , respectively, and the step size is  $h_n = t_n - t_{n-1}$ . The coefficients  $\alpha_{n,i}$  are uniquely determined by the order q, and the history of the step sizes. The application of the BDF (2.3) to the DAE system (2.1) results in a nonlinear algebraic system to be solved at each step:

$$G(y_n) \equiv F\left(t_n, y_n, h_n^{-1} \sum_{i=0}^{q} \alpha_{n,i} y_{n-i}\right) = 0.$$
 (2.4)

By default IDAS solves (2.4) with a Newton iteration but IDAS also allows for user-defined nonlinear solvers (see Chapter 10). Each Newton iteration requires the solution of a linear system of the form

$$J[y_{n(m+1)} - y_{n(m)}] = -G(y_{n(m)}), (2.5)$$

where  $y_{n(m)}$  is the m-th approximation to  $y_n$ . Here J is some approximation to the system Jacobian

$$J = \frac{\partial G}{\partial y} = \frac{\partial F}{\partial y} + \alpha \frac{\partial F}{\partial \dot{y}}, \qquad (2.6)$$

where  $\alpha = \alpha_{n,0}/h_n$ . The scalar  $\alpha$  changes whenever the step size or method order changes.

For the solution of the linear systems within the Newton iteration, IDAS provides several choices, including the option of a user-supplied linear solver module (see Chapter 9). The linear solver modules distributed with SUNDIALS are organized in two families, a *direct* family comprising direct linear solvers for dense, banded, or sparse matrices and a *spils* family comprising scaled preconditioned iterative (Krylov) linear solvers. The methods offered through these modules are as follows:

- dense direct solvers, using either an internal implementation or a BLAS/LAPACK implementation (serial or threaded vector modules only),
- band direct solvers, using either an internal implementation or a BLAS/LAPACK implementation (serial or threaded vector modules only),
- sparse direct solver interfaces, using either the KLU sparse solver library [17, 1], or the threadenabled SuperLU\_MT sparse solver library [35, 19, 2] (serial or threaded vector modules only) [Note that users will need to download and install the KLU or SUPERLUMT packages independent of IDAS],
- SPGMR, a scaled preconditioned GMRES (Generalized Minimal Residual method) solver without restarts.
- SPFGMR, a scaled preconditioned FGMRES (Flexible Generalized Minimal Residual method) solver.
- SPBCGS, a scaled preconditioned Bi-CGStab (Bi-Conjugate Gradient Stable method) solver,
- SPTFQMR, a scaled preconditioned TFQMR (Transpose-Free Quasi-Minimal Residual method) solver, or
- PCG, a scaled preconditioned CG (Conjugate Gradient method) solver.

For large stiff systems, where direct methods are not feasible, the combination of a BDF integrator and a preconditioned Krylov method yields a powerful tool because it combines established methods for stiff integration, nonlinear iteration, and Krylov (linear) iteration with a problem-specific treatment of the dominant source of stiffness, in the form of the user-supplied preconditioner matrix [6]. For the *spils* linear solvers with IDAS, preconditioning is allowed only on the left (see §2.2). Note that the dense, band, and sparse direct linear solvers can only be used with serial and threaded vector representations.

2.1 IVP solution

In the process of controlling errors at various levels, IDAS uses a weighted root-mean-square norm, denoted  $\|\cdot\|_{WRMS}$ , for all error-like quantities. The multiplicative weights used are based on the current solution and on the relative and absolute tolerances input by the user, namely

$$W_i = 1/[\text{RTOL} \cdot |y_i| + \text{ATOL}_i]. \tag{2.7}$$

Because  $1/W_i$  represents a tolerance in the component  $y_i$ , a vector whose norm is 1 is regarded as "small." For brevity, we will usually drop the subscript WRMS on norms in what follows.

In the case of a matrix-based linear solver, the default Newton iteration is a Modified Newton iteration, in that the Jacobian J is fixed (and usually out of date) throughout the nonlinear iterations, with a coefficient  $\bar{\alpha}$  in place of  $\alpha$  in J. However, in the case that a matrix-free iterative linear solver is used, the default Newton iteration is an Inexact Newton iteration, in which J is applied in a matrix-free manner, with matrix-vector products Jv obtained by either difference quotients or a user-supplied routine. In this case, the linear residual  $J\Delta y + G$  is nonzero but controlled. With the default Newton iteration, the matrix J and preconditioner matrix P are updated as infrequently as possible to balance the high costs of matrix operations against other costs. Specifically, this matrix update occurs when:

- starting the problem,
- the value  $\bar{\alpha}$  at the last update is such that  $\alpha/\bar{\alpha} < 3/5$  or  $\alpha/\bar{\alpha} > 5/3$ , or
- $\bullet$  a non-fatal convergence failure occurred with an out-of-date J or P.

The above strategy balances the high cost of frequent matrix evaluations and preprocessing with the slow convergence due to infrequent updates. To reduce storage costs on an update, Jacobian information is always reevaluated from scratch.

The default stopping test for nonlinear solver iterations in IDAS ensures that the iteration error  $y_n - y_{n(m)}$  is small relative to y itself. For this, we estimate the linear convergence rate at all iterations m > 1 as

$$R = \left(\frac{\delta_m}{\delta_1}\right)^{\frac{1}{m-1}} ,$$

where the  $\delta_m = y_{n(m)} - y_{n(m-1)}$  is the correction at iteration  $m = 1, 2, \ldots$  The nonlinear solver iteration is halted if R > 0.9. The convergence test at the m-th iteration is then

$$S\|\delta_m\| < 0.33\,,\tag{2.8}$$

where S = R/(R-1) whenever m > 1 and  $R \le 0.9$ . The user has the option of changing the constant in the convergence test from its default value of 0.33. The quantity S is set to S = 20 initially and whenever J or P is updated, and it is reset to S = 100 on a step with  $\alpha \ne \bar{\alpha}$ . Note that at m = 1, the convergence test (2.8) uses an old value for S. Therefore, at the first nonlinear solver iteration, we make an additional test and stop the iteration if  $\|\delta_1\| < 0.33 \cdot 10^{-4}$  (since such a  $\delta_1$  is probably just noise and therefore not appropriate for use in evaluating R). We allow only a small number (default value 4) of nonlinear iterations. If convergence fails with J or P current, we are forced to reduce the step size  $h_n$ , and we replace  $h_n$  by  $h_n/4$ . The integration is halted after a preset number (default value 10) of convergence failures. Both the maximum number of allowable nonlinear iterations and the maximum number of nonlinear convergence failures can be changed by the user from their default values.

When an iterative method is used to solve the linear system, to minimize the effect of linear iteration errors on the nonlinear and local integration error controls, we require the preconditioned linear residual to be small relative to the allowed error in the nonlinear iteration, i.e.,  $||P^{-1}(Jx+G)|| < 0.05 \cdot 0.33$ . The safety factor 0.05 can be changed by the user.

When the Jacobian is stored using either dense or band Sunmatrix objects, the Jacobian J defined in (2.6) can be either supplied by the user or have IDAS compute one internally by difference quotients. In the latter case, we use the approximation

$$J_{ij} = [F_i(t, y + \sigma_j e_j, \dot{y} + \alpha \sigma_j e_j) - F_i(t, y, \dot{y})] / \sigma_j, \text{ with}$$
  
$$\sigma_j = \sqrt{U} \max\{|y_j|, |h\dot{y}_j|, 1/W_j\} \operatorname{sign}(h\dot{y}_j),$$

where U is the unit roundoff, h is the current step size, and  $W_j$  is the error weight for the component  $y_j$  defined by (2.7). We note that with sparse and user-supplied SUNMATRIX objects, the Jacobian must be supplied by a user routine.

In the case of an iterative linear solver, if a routine for Jv is not supplied, such products are approximated by

$$Jv = [F(t, y + \sigma v, \dot{y} + \alpha \sigma v) - F(t, y, \dot{y})]/\sigma \,,$$

where the increment  $\sigma = \sqrt{N}$ . As an option, the user can specify a constant factor that is inserted into this expression for  $\sigma$ .

During the course of integrating the system, IDAS computes an estimate of the local truncation error, LTE, at the *n*-th time step, and requires this to satisfy the inequality

$$\|LTE\|_{WRMS} \le 1$$
.

Asymptotically, LTE varies as  $h^{q+1}$  at step size h and order q, as does the predictor-corrector difference  $\Delta_n \equiv y_n - y_{n(0)}$ . Thus there is a constant C such that

$$LTE = C\Delta_n + O(h^{q+2}),$$

and so the norm of LTE is estimated as  $|C| \cdot ||\Delta_n||$ . In addition, IDAS requires that the error in the associated polynomial interpolant over the current step be bounded by 1 in norm. The leading term of the norm of this error is bounded by  $\bar{C}||\Delta_n||$  for another constant  $\bar{C}$ . Thus the local error test in IDAS is

$$\max\{|C|, \bar{C}\}\|\Delta_n\| \le 1. \tag{2.9}$$

A user option is available by which the algebraic components of the error vector are omitted from the test (2.9), if these have been so identified.

In IDAS, the local error test is tightly coupled with the logic for selecting the step size and order. First, there is an initial phase that is treated specially; for the first few steps, the step size is doubled and the order raised (from its initial value of 1) on every step, until (a) the local error test (2.9) fails, (b) the order is reduced (by the rules given below), or (c) the order reaches 5 (the maximum). For step and order selection on the general step, IDAS uses a different set of local error estimates, based on the asymptotic behavior of the local error in the case of fixed step sizes. At each of the orders q' equal to q, q-1 (if q>1), q-2 (if q>2), or q+1 (if q<5), there are constants C(q') such that the norm of the local truncation error at order q' satisfies

$$LTE(q') = C(q') \|\phi(q'+1)\| + O(h^{q'+2}),$$

where  $\phi(k)$  is a modified divided difference of order k that is retained by IDAS (and behaves asymptotically as  $h^k$ ). Thus the local truncation errors are estimated as  $\text{ELTE}(q') = C(q') \|\phi(q'+1)\|$  to select step sizes. But the choice of order in IDAS is based on the requirement that the scaled derivative norms,  $\|h^k y^{(k)}\|$ , are monotonically decreasing with k, for k near q. These norms are again estimated using the  $\phi(k)$ , and in fact

$$||h^{q'+1}y^{(q'+1)}|| \approx T(q') \equiv (q'+1)\text{ELTE}(q')$$
.

The step/order selection begins with a test for monotonicity that is made even before the local error test is performed. Namely, the order is reset to q' = q - 1 if (a) q = 2 and  $T(1) \le T(2)/2$ , or (b) q > 2 and  $\max\{T(q-1), T(q-2)\} \le T(q)$ ; otherwise q' = q. Next the local error test (2.9) is performed, and if it fails, the step is redone at order  $q \leftarrow q'$  and a new step size h'. The latter is based on the  $h^{q+1}$  asymptotic behavior of ELTE(q), and, with safety factors, is given by

$$\eta = h'/h = 0.9/[2\,\mathrm{ELTE}(q)]^{1/(q+1)}\,.$$

The value of  $\eta$  is adjusted so that  $0.25 \le \eta \le 0.9$  before setting  $h \leftarrow h' = \eta h$ . If the local error test fails a second time, IDAS uses  $\eta = 0.25$ , and on the third and subsequent failures it uses q = 1 and  $\eta = 0.25$ . After 10 failures, IDAS returns with a give-up message.

As soon as the local error test has passed, the step and order for the next step may be adjusted. No such change is made if q' = q - 1 from the prior test, if q = 5, or if q was increased on the previous step. Otherwise, if the last q + 1 steps were taken at a constant order q < 5 and a constant step size, IDAS considers raising the order to q + 1. The logic is as follows: (a) If q = 1, then reset q = 2 if T(2) < T(1)/2. (b) If q > 1 then

- reset  $q \leftarrow q 1$  if  $T(q 1) \le \min\{T(q), T(q + 1)\}$ ;
- else reset  $q \leftarrow q + 1$  if T(q + 1) < T(q);
- leave q unchanged otherwise [then  $T(q-1) > T(q) \le T(q+1)$ ].

In any case, the new step size h' is set much as before:

$$\eta = h'/h = 1/[2 \text{ELTE}(q)]^{1/(q+1)}$$
.

The value of  $\eta$  is adjusted such that (a) if  $\eta > 2$ ,  $\eta$  is reset to 2; (b) if  $\eta \le 1$ ,  $\eta$  is restricted to  $0.5 \le \eta \le 0.9$ ; and (c) if  $1 < \eta < 2$  we use  $\eta = 1$ . Finally h is reset to  $h' = \eta h$ . Thus we do not increase the step size unless it can be doubled. See [4] for details.

IDAS permits the user to impose optional inequality constraints on individual components of the solution vector y. Any of the following four constraints can be imposed:  $y_i > 0$ ,  $y_i < 0$ ,  $y_i \geq 0$ , or  $y_i \leq 0$ . The constraint satisfaction is tested after a successful nonlinear system solution. If any constraint fails, we declare a convergence failure of the nonlinear iteration and reduce the step size. Rather than cutting the step size by some arbitrary factor, IDAS estimates a new step size h' using a linear approximation of the components in y that failed the constraint test (including a safety factor of 0.9 to cover the strict inequality case). These additional constraints are also imposed during the calculation of consistent initial conditions.

Normally, IDAS takes steps until a user-defined output value  $t = t_{\text{out}}$  is overtaken, and then computes  $y(t_{\text{out}})$  by interpolation. However, a "one step" mode option is available, where control returns to the calling program after each step. There are also options to force IDAS not to integrate past a given stopping point  $t = t_{\text{stop}}$ .

# 2.2 Preconditioning

When using a nonlinear solver that requires the solution of a linear system of the form  $J\Delta y = -G$  (e.g., the default Newton iteration), IDAS makes repeated use of a linear solver. If this linear system solve is done with one of the scaled preconditioned iterative linear solvers supplied with SUNDIALS, these solvers are rarely successful if used without preconditioning; it is generally necessary to precondition the system in order to obtain acceptable efficiency. A system Ax = b can be preconditioned on the left, on the right, or on both sides. The Krylov method is then applied to a system with the matrix  $P^{-1}A$ , or  $AP^{-1}$ , or  $P_L^{-1}AP_R^{-1}$ , instead of A. However, within IDAS, preconditioning is allowed only on the left, so that the iterative method is applied to systems  $(P^{-1}J)\Delta y = -P^{-1}G$ . Left preconditioning is required to make the norm of the linear residual in the nonlinear iteration meaningful; in general,  $\|J\Delta y + G\|$  is meaningless, since the weights used in the WRMS-norm correspond to y.

In order to improve the convergence of the Krylov iteration, the preconditioner matrix P should in some sense approximate the system matrix A. Yet at the same time, in order to be cost-effective, the matrix P should be reasonably efficient to evaluate and solve. Finding a good point in this tradeoff between rapid convergence and low cost can be very difficult. Good choices are often problem-dependent (for example, see [6] for an extensive study of preconditioners for reaction-transport systems).

Typical preconditioners used with IDAS are based on approximations to the iteration matrix of the systems involved; in other words,  $P \approx \frac{\partial F}{\partial y} + \alpha \frac{\partial F}{\partial y}$ , where  $\alpha$  is a scalar inversely proportional to the integration step size h. Because the Krylov iteration occurs within a nonlinear solver iteration and further also within a time integration, and since each of these iterations has its own test for convergence, the preconditioner may use a very crude approximation, as long as it captures the dominant numerical feature(s) of the system. We have found that the combination of a preconditioner

with the Newton-Krylov iteration, using even a fairly poor approximation to the Jacobian, can be surprisingly superior to using the same matrix without Krylov acceleration (i.e., a modified Newton iteration), as well as to using the Newton-Krylov method with no preconditioning.

#### 2.3 Rootfinding

The IDAS solver has been augmented to include a rootfinding feature. This means that, while integrating the Initial Value Problem (2.1), IDAS can also find the roots of a set of user-defined functions  $g_i(t, y, \dot{y})$  that depend on t, the solution vector y = y(t), and its t-derivative  $\dot{y}(t)$ . The number of these root functions is arbitrary, and if more than one  $g_i$  is found to have a root in any given interval, the various root locations are found and reported in the order that they occur on the t axis, in the direction of integration.

Generally, this rootfinding feature finds only roots of odd multiplicity, corresponding to changes in sign of  $g_i(t, y(t), \dot{y}(t))$ , denoted  $g_i(t)$  for short. If a user root function has a root of even multiplicity (no sign change), it will probably be missed by IDAS. If such a root is desired, the user should reformulate the root function so that it changes sign at the desired root.

The basic scheme used is to check for sign changes of any  $g_i(t)$  over each time step taken, and then (when a sign change is found) to home in on the root (or roots) with a modified secant method [25]. In addition, each time g is computed, IDAS checks to see if  $g_i(t) = 0$  exactly, and if so it reports this as a root. However, if an exact zero of any  $g_i$  is found at a point t, IDAS computes g at  $t + \delta$  for a small increment  $\delta$ , slightly further in the direction of integration, and if any  $g_i(t + \delta) = 0$  also, IDAS stops and reports an error. This way, each time IDAS takes a time step, it is guaranteed that the values of all  $g_i$  are nonzero at some past value of t, beyond which a search for roots is to be done.

At any given time in the course of the time-stepping, after suitable checking and adjusting has been done, IDAS has an interval  $(t_{lo}, t_{hi}]$  in which roots of the  $g_i(t)$  are to be sought, such that  $t_{hi}$  is further ahead in the direction of integration, and all  $g_i(t_{lo}) \neq 0$ . The endpoint  $t_{hi}$  is either  $t_n$ , the end of the time step last taken, or the next requested output time  $t_{out}$  if this comes sooner. The endpoint  $t_{lo}$  is either  $t_{n-1}$ , or the last output time  $t_{out}$  (if this occurred within the last step), or the last root location (if a root was just located within this step), possibly adjusted slightly toward  $t_n$  if an exact zero was found. The algorithm checks g at  $t_{hi}$  for zeros and for sign changes in  $(t_{lo}, t_{hi})$ . If no sign changes are found, then either a root is reported (if some  $g_i(t_{hi}) = 0$ ) or we proceed to the next time interval (starting at  $t_{hi}$ ). If one or more sign changes were found, then a loop is entered to locate the root to within a rather tight tolerance, given by

$$\tau = 100 * U * (|t_n| + |h|)$$
 (U = unit roundoff).

Whenever sign changes are seen in two or more root functions, the one deemed most likely to have its root occur first is the one with the largest value of  $|g_i(t_{hi})|/|g_i(t_{hi}) - g_i(t_{lo})|$ , corresponding to the closest to  $t_{lo}$  of the secant method values. At each pass through the loop, a new value  $t_{mid}$  is set, strictly within the search interval, and the values of  $g_i(t_{mid})$  are checked. Then either  $t_{lo}$  or  $t_{hi}$  is reset to  $t_{mid}$  according to which subinterval is found to have the sign change. If there is none in  $(t_{lo}, t_{mid})$  but some  $g_i(t_{mid}) = 0$ , then that root is reported. The loop continues until  $|t_{hi} - t_{lo}| < \tau$ , and then the reported root location is  $t_{hi}$ .

In the loop to locate the root of  $g_i(t)$ , the formula for  $t_{mid}$  is

$$t_{mid} = t_{hi} - (t_{hi} - t_{lo})g_i(t_{hi})/[g_i(t_{hi}) - \alpha g_i(t_{lo})]$$
,

where  $\alpha$  a weight parameter. On the first two passes through the loop,  $\alpha$  is set to 1, making  $t_{mid}$  the secant method value. Thereafter,  $\alpha$  is reset according to the side of the subinterval (low vs high, i.e. toward  $t_{lo}$  vs toward  $t_{hi}$ ) in which the sign change was found in the previous two passes. If the two sides were opposite,  $\alpha$  is set to 1. If the two sides were the same,  $\alpha$  is halved (if on the low side) or doubled (if on the high side). The value of  $t_{mid}$  is closer to  $t_{lo}$  when  $\alpha < 1$  and closer to  $t_{hi}$  when  $\alpha > 1$ . If the above value of  $t_{mid}$  is within  $\tau/2$  of  $t_{lo}$  or  $t_{hi}$ , it is adjusted inward, such that its fractional distance from the endpoint (relative to the interval size) is between .1 and .5 (.5 being the midpoint), and the actual distance from the endpoint is at least  $\tau/2$ .

#### 2.4 Pure quadrature integration

In many applications, and most notably during the backward integration phase of an adjoint sensitivity analysis run (see §2.6) it is of interest to compute integral quantities of the form

$$z(t) = \int_{t_0}^{t} q(\tau, y(\tau), \dot{y}(\tau), p) d\tau.$$
 (2.10)

The most effective approach to compute z(t) is to extend the original problem with the additional ODEs (obtained by applying Leibnitz's differentiation rule):

$$\dot{z} = q(t, y, \dot{y}, p), \quad z(t_0) = 0.$$
 (2.11)

Note that this is equivalent to using a quadrature method based on the underlying linear multistep polynomial representation for y(t).

This can be done at the "user level" by simply exposing to IDAS the extended DAE system (2.2)+(2.10). However, in the context of an implicit integration solver, this approach is not desirable since the nonlinear solver module will require the Jacobian (or Jacobian-vector product) of this extended DAE. Moreover, since the additional states, z, do not enter the right-hand side of the ODE (2.10) and therefore the residual of the extended DAE system does not depend on z, it is much more efficient to treat the ODE system (2.10) separately from the original DAE system (2.2) by "taking out" the additional states z from the nonlinear system (2.4) that must be solved in the correction step of the LMM. Instead, "corrected" values  $z_n$  are computed explicitly as

$$z_n = \frac{1}{\alpha_{n,0}} \left( h_n q(t_n, y_n, \dot{y}_n, p) - \sum_{i=1}^q \alpha_{n,i} z_{n-i} \right),$$

once the new approximation  $y_n$  is available.

The quadrature variables z can be optionally included in the error test, in which case corresponding relative and absolute tolerances must be provided.

# 2.5 Forward sensitivity analysis

Typically, the governing equations of complex, large-scale models depend on various parameters, through the right-hand side vector and/or through the vector of initial conditions, as in (2.2). In addition to numerically solving the DAEs, it may be desirable to determine the sensitivity of the results with respect to the model parameters. Such sensitivity information can be used to estimate which parameters are most influential in affecting the behavior of the simulation or to evaluate optimization gradients (in the setting of dynamic optimization, parameter estimation, optimal control, etc.).

The solution sensitivity with respect to the model parameter  $p_i$  is defined as the vector  $s_i(t) = \partial y(t)/\partial p_i$  and satisfies the following forward sensitivity equations (or sensitivity equations for short):

$$\frac{\partial F}{\partial y}s_i + \frac{\partial F}{\partial \dot{y}}\dot{s}_i + \frac{\partial F}{\partial p_i} = 0$$

$$s_i(t_0) = \frac{\partial y_0(p)}{\partial p_i}, \ \dot{s}_i(t_0) = \frac{\partial \dot{y}_0(p)}{\partial p_i},$$
(2.12)

obtained by applying the chain rule of differentiation to the original DAEs (2.2).

When performing forward sensitivity analysis, IDAS carries out the time integration of the combined system, (2.2) and (2.12), by viewing it as a DAE system of size  $N(N_s + 1)$ , where  $N_s$  is the number of model parameters  $p_i$ , with respect to which sensitivities are desired  $(N_s \leq N_p)$ . However, major improvements in efficiency can be made by taking advantage of the special form of the sensitivity equations as linearizations of the original DAEs. In particular, the original DAE system and all sensitivity systems share the same Jacobian matrix J in (2.6).

The sensitivity equations are solved with the same linear multistep formula that was selected for the original DAEs and the same linear solver is used in the correction phase for both state and sensitivity variables. In addition, IDAS offers the option of including (full error control) or excluding (partial error control) the sensitivity variables from the local error test.

#### 2.5.1 Forward sensitivity methods

In what follows we briefly describe three methods that have been proposed for the solution of the combined DAE and sensitivity system for the vector  $\hat{y} = [y, s_1, \dots, s_{N_s}]$ .

- Staggered Direct In this approach [14], the nonlinear system (2.4) is first solved and, once an acceptable numerical solution is obtained, the sensitivity variables at the new step are found by directly solving (2.12) after the BDF discretization is used to eliminate  $\dot{s}_i$ . Although the system matrix of the above linear system is based on exactly the same information as the matrix J in (2.6), it must be updated and factored at every step of the integration, in contrast to an evaluation of J which is updated only occasionally. For problems with many parameters (relative to the problem size), the staggered direct method can outperform the methods described below [34]. However, the computational cost associated with matrix updates and factorizations makes this method unattractive for problems with many more states than parameters (such as those arising from semidiscretization of PDEs) and is therefore not implemented in IDAS.
- Simultaneous Corrector In this method [37], the discretization is applied simultaneously to both the original equations (2.2) and the sensitivity systems (2.12) resulting in an "extended" non-linear system  $\hat{G}(\hat{y}_n) = 0$  where  $\hat{y}_n = [y_n, \dots, s_i, \dots]$ . This combined nonlinear system can be solved using a modified Newton method as in (2.5) by solving the corrector equation

$$\hat{J}[\hat{y}_{n(m+1)} - \hat{y}_{n(m)}] = -\hat{G}(\hat{y}_{n(m)}) \tag{2.13}$$

at each iteration, where

$$\hat{J} = \begin{bmatrix} J & & & & \\ J_1 & J & & & \\ J_2 & 0 & J & & \\ \vdots & \vdots & \ddots & \ddots & \\ J_{N_s} & 0 & \dots & 0 & J \end{bmatrix} \,,$$

J is defined as in (2.6), and  $J_i = (\partial/\partial y) [F_y s_i + F_{\dot{y}} \dot{s}_i + F_{p_i}]$ . It can be shown that 2-step quadratic convergence can be retained by using only the block-diagonal portion of  $\hat{J}$  in the corrector equation (2.13). This results in a decoupling that allows the reuse of J without additional matrix factorizations. However, the sum  $F_y s_i + F_{\dot{y}} \dot{s}_i + F_{p_i}$  must still be reevaluated at each step of the iterative process (2.13) to update the sensitivity portions of the residual  $\hat{G}$ .

• Staggered corrector In this approach [22], as in the staggered direct method, the nonlinear system (2.4) is solved first using the Newton iteration (2.5). Then, for each sensitivity vector  $\xi \equiv s_i$ , a separate Newton iteration is used to solve the sensitivity system (2.12):

$$J[\xi_{n(m+1)} - \xi_{n(m)}] = -\left[F_{y}(t_{n}, y_{n}, \dot{y}_{n})\xi_{n(m)} + F_{\dot{y}}(t_{n}, y_{n}, \dot{y}_{n}) \cdot h_{n}^{-1} \left(\alpha_{n,0}\xi_{n(m)} + \sum_{i=1}^{q} \alpha_{n,i}\xi_{n-i}\right) + F_{p_{i}}(t_{n}, y_{n}, \dot{y}_{n})\right].$$
(2.14)

In other words, a modified Newton iteration is used to solve a linear system. In this approach, the matrices  $\partial F/\partial y$ ,  $\partial F/\partial \dot{y}$  and vectors  $\partial F/\partial p_i$  need be updated only once per integration step, after the state correction phase (2.5) has converged.

IDAS implements both the simultaneous corrector method and the staggered corrector method.

An important observation is that the staggered corrector method, combined with a Krylov linear solver, effectively results in a staggered direct method. Indeed, the Krylov solver requires only the action of the matrix J on a vector, and this can be provided with the current Jacobian information. Therefore, the modified Newton procedure (2.14) will theoretically converge after one iteration.

#### 2.5.2 Selection of the absolute tolerances for sensitivity variables

If the sensitivities are included in the error test, IDAS provides an automated estimation of absolute tolerances for the sensitivity variables based on the absolute tolerance for the corresponding state variable. The relative tolerance for sensitivity variables is set to be the same as for the state variables. The selection of absolute tolerances for the sensitivity variables is based on the observation that the sensitivity vector  $s_i$  will have units of  $[y]/[p_i]$ . With this, the absolute tolerance for the j-th component of the sensitivity vector  $s_i$  is set to  $\text{ATOL}_j/|\bar{p}_i|$ , where  $\text{ATOL}_j$  are the absolute tolerances for the state variables and  $\bar{p}$  is a vector of scaling factors that are dimensionally consistent with the model parameters p and give an indication of their order of magnitude. This choice of relative and absolute tolerances is equivalent to requiring that the weighted root-mean-square norm of the sensitivity vector  $s_i$  with weights based on  $s_i$  be the same as the weighted root-mean-square norm of the vector of scaled sensitivities  $\bar{s}_i = |\bar{p}_i|s_i$  with weights based on the state variables (the scaled sensitivities  $\bar{s}_i$  being dimensionally consistent with the state variables). However, this choice of tolerances for the  $s_i$  may be a poor one, and the user of IDAS can provide different values as an option.

#### 2.5.3 Evaluation of the sensitivity right-hand side

There are several methods for evaluating the residual functions in the sensitivity systems (2.12): analytic evaluation, automatic differentiation, complex-step approximation, and finite differences (or directional derivatives). IDAS provides all the software hooks for implementing interfaces to automatic differentiation (AD) or complex-step approximation; future versions will include a generic interface to AD-generated functions. At the present time, besides the option for analytical sensitivity right-hand sides (user-provided), IDAS can evaluate these quantities using various finite difference-based approximations to evaluate the terms  $(\partial F/\partial y)s_i + (\partial F/\partial y)\dot{s}_i$  and  $(\partial F/\partial p_i)$ , or using directional derivatives to evaluate  $[(\partial F/\partial y)s_i + (\partial F/\partial y)\dot{s}_i + (\partial F/\partial p_i)]$ . As is typical for finite differences, the proper choice of perturbations is a delicate matter. IDAS takes into account several problem-related features: the relative DAE error tolerance RTOL, the machine unit roundoff U, the scale factor  $\bar{p}_i$ , and the weighted root-mean-square norm of the sensitivity vector  $s_i$ .

Using central finite differences as an example, the two terms  $(\partial F/\partial y)s_i + (\partial F/\partial y)\dot{s}_i$  and  $\partial F/\partial p_i$  in (2.12) can be evaluated either separately:

$$\frac{\partial F}{\partial y}s_{i} + \frac{\partial F}{\partial \dot{y}}\dot{s}_{i} \approx \frac{F(t, y + \sigma_{y}s_{i}, \dot{y} + \sigma_{y}\dot{s}_{i}, p) - F(t, y - \sigma_{y}s_{i}, \dot{y} - \sigma_{y}\dot{s}_{i}, p)}{2\sigma_{y}}, \qquad (2.15)$$

$$\frac{\partial F}{\partial p_{i}} \approx \frac{F(t, y, \dot{y}, p + \sigma_{i}e_{i}) - F(t, y, \dot{y}, p - \sigma_{i}e_{i})}{2\sigma_{i}}, \qquad (2.15')$$

$$\sigma_{i} = |\bar{p}_{i}|\sqrt{\max(\text{RTOL}, U)}, \quad \sigma_{y} = \frac{1}{\max(1/\sigma_{i}, \|s_{i}\|_{\text{WRMS}}/|\bar{p}_{i}|)},$$

or simultaneously:

$$\frac{\partial F}{\partial y}s_i + \frac{\partial F}{\partial \dot{y}}\dot{s}_i + \frac{\partial F}{\partial p_i} \approx \frac{F(t, y + \sigma s_i, \dot{y} + \sigma \dot{s}_i, p + \sigma e_i) - F(t, y - \sigma s_i, \dot{y} - \sigma \dot{s}_i, p - \sigma e_i)}{2\sigma}, \qquad (2.16)$$

$$\sigma = \min(\sigma_i, \sigma_y),$$

or by adaptively switching between (2.15)+(2.15) and (2.16), depending on the relative size of the two finite difference increments  $\sigma_i$  and  $\sigma_y$ . In the adaptive scheme, if  $\rho = \max(\sigma_i/\sigma_y, \sigma_y/\sigma_i)$ , we use separate evaluations if  $\rho > \rho_{\rm max}$  (an input value), and simultaneous evaluations otherwise.

These procedures for choosing the perturbations  $(\sigma_i, \sigma_y, \sigma)$  and switching between derivative formulas have also been implemented for one-sided difference formulas. Forward finite differences can be applied to  $(\partial F/\partial y)s_i + (\partial F/\partial \dot{y})\dot{s}_i$  and  $\frac{\partial F}{\partial p_i}$  separately, or the single directional derivative formula

$$\frac{\partial F}{\partial y}s_i + \frac{\partial F}{\partial \dot{y}}\dot{s}_i + \frac{\partial F}{\partial p_i} \approx \frac{F(t,y+\sigma s_i,\dot{y}+\sigma \dot{s}_i,p+\sigma e_i) - F(t,y,\dot{y},p)}{\sigma}$$

can be used. In IDAS, the default value of  $\rho_{\text{max}} = 0$  indicates the use of the second-order centered directional derivative formula (2.16) exclusively. Otherwise, the magnitude of  $\rho_{\text{max}}$  and its sign (positive or negative) indicates whether this switching is done with regard to (centered or forward) finite differences, respectively.

#### 2.5.4 Quadratures depending on forward sensitivities

If pure quadrature variables are also included in the problem definition (see §2.4), IDAS does not carry their sensitivities automatically. Instead, we provide a more general feature through which integrals depending on both the states y of (2.2) and the state sensitivities  $s_i$  of (2.12) can be evaluated. In other words, IDAS provides support for computing integrals of the form:

$$\bar{z}(t) = \int_{t_0}^t \bar{q}(\tau, y(\tau), \dot{y}(\tau), s_1(\tau), \dots, s_{N_p}(\tau), p) d\tau.$$

If the sensitivities of the quadrature variables z of (2.10) are desired, these can then be computed by using:

$$\bar{q}_i = q_y s_i + q_{\dot{y}} \dot{s}_i + q_{p_i}, \quad i = 1, \dots, N_p$$

as integrands for  $\bar{z}$ , where  $q_y$ ,  $q_{\dot{y}}$ , and  $q_p$  are the partial derivatives of the integrand function q of (2.10).

As with the quadrature variables z, the new variables  $\bar{z}$  are also excluded from any nonlinear solver phase and "corrected" values  $\bar{z}_n$  are obtained through explicit formulas.

#### 2.6 Adjoint sensitivity analysis

In the forward sensitivity approach described in the previous section, obtaining sensitivities with respect to  $N_s$  parameters is roughly equivalent to solving an DAE system of size  $(1 + N_s)N$ . This can become prohibitively expensive, especially for large-scale problems, if sensitivities with respect to many parameters are desired. In this situation, the adjoint sensitivity method is a very attractive alternative, provided that we do not need the solution sensitivities  $s_i$ , but rather the gradients with respect to model parameters of a relatively few derived functionals of the solution. In other words, if y(t) is the solution of (2.2), we wish to evaluate the gradient dG/dp of

$$G(p) = \int_{t_0}^{T} g(t, y, p)dt,$$
 (2.17)

or, alternatively, the gradient dg/dp of the function g(t, y, p) at the final time t = T. The function g must be smooth enough that  $\partial g/\partial y$  and  $\partial g/\partial p$  exist and are bounded.

In what follows, we only sketch the analysis for the sensitivity problem for both G and g. For details on the derivation see [13].

#### **2.6.1** Sensitivity of G(p)

We focus first on solving the sensitivity problem for G(p) defined by (2.17). Introducing a Lagrange multiplier  $\lambda$ , we form the augmented objective function

$$I(p) = G(p) - \int_{t_0}^T \lambda^* F(t, y, \dot{y}, p) dt.$$

Since  $F(t, y, \dot{y}, p) = 0$ , the sensitivity of G with respect to p is

$$\frac{dG}{dp} = \frac{dI}{dp} = \int_{t_0}^{T} (g_p + g_y y_p) dt - \int_{t_0}^{T} \lambda^* (F_p + F_y y_p + F_{\dot{y}} \dot{y}_p) dt, \tag{2.18}$$

where subscripts on functions such as F or g are used to denote partial derivatives. By integration by parts, we have

$$\int_{t_0}^T \lambda^* F_{\dot{y}} \dot{y}_p dt = (\lambda^* F_{\dot{y}} y_p)|_{t_0}^T - \int_{t_0}^T (\lambda^* F_{\dot{y}})' y_p dt,$$

where  $(\cdots)'$  denotes the t-derivative. Thus equation (2.18) becomes

$$\frac{dG}{dp} = \int_{t_0}^T (g_p - \lambda^* F_p) dt - \int_{t_0}^T [-g_y + \lambda^* F_y - (\lambda^* F_{\dot{y}})'] y_p dt - (\lambda^* F_{\dot{y}} y_p)|_{t_0}^T.$$
 (2.19)

Now by requiring  $\lambda$  to satisfy

$$(\lambda^* F_{\dot{y}})' - \lambda^* F_y = -g_y, \tag{2.20}$$

we obtain

$$\frac{dG}{dp} = \int_{t_0}^{T} (g_p - \lambda^* F_p) dt - (\lambda^* F_{\dot{y}} y_p)|_{t_0}^{T}.$$
 (2.21)

Note that  $y_p$  at  $t = t_0$  is the sensitivity of the initial conditions with respect to p, which is easily obtained. To find the initial conditions (at t = T) for the adjoint system, we must take into consideration the structure of the DAE system.

For index-0 and index-1 DAE systems, we can simply take

$$\lambda^* F_{\dot{\eta}}|_{t=T} = 0, \tag{2.22}$$

yielding the sensitivity equation for dG/dp

$$\frac{dG}{dp} = \int_{t_0}^{T} (g_p - \lambda^* F_p) dt + (\lambda^* F_j y_p)|_{t=t_0}.$$
 (2.23)

This choice will not suffice for a Hessenberg index-2 DAE system. For a derivation of proper final conditions in such cases, see [13].

The first thing to notice about the adjoint system (2.20) is that there is no explicit specification of the parameters p; this implies that, once the solution  $\lambda$  is found, the formula (2.21) can then be used to find the gradient of G with respect to any of the parameters p. The second important remark is that the adjoint system (2.20) is a terminal value problem which depends on the solution y(t) of the original IVP (2.2). Therefore, a procedure is needed for providing the states y obtained during a forward integration phase of (2.2) to IDAS during the backward integration phase of (2.20). The approach adopted in IDAS, based on *checkpointing*, is described in §2.6.3 below.

#### **2.6.2** Sensitivity of q(T, p)

Now let us consider the computation of dg/dp(T). From dg/dp(T) = (d/dT)(dG/dp) and equation (2.21), we have

$$\frac{dg}{dp} = (g_p - \lambda^* F_p)(T) - \int_{t_0}^T \lambda_T^* F_p dt + (\lambda_T^* F_{\dot{y}} y_p)|_{t=t_0} - \frac{d(\lambda^* F_{\dot{y}} y_p)}{dT}$$
(2.24)

where  $\lambda_T$  denotes  $\partial \lambda/\partial T$ . For index-0 and index-1 DAEs, we obtain

$$\frac{d(\lambda^* F_i y_p)|_{t=T}}{dT} = 0,$$

while for a Hessenberg index-2 DAE system we have

$$\frac{d(\lambda^* F_{ij} y_p)|_{t=T}}{dT} = -\left. \frac{d(g_{y^a} (CB)^{-1} f_p^2)}{dt} \right|_{t=T}.$$

The corresponding adjoint equations are

$$(\lambda_T^* F_{\dot{\nu}})' - \lambda_T^* F_{\nu} = 0. \tag{2.25}$$

For index-0 and index-1 DAEs (as shown above, the index-2 case is different), to find the boundary condition for this equation we write  $\lambda$  as  $\lambda(t,T)$  because it depends on both t and T. Then

$$\lambda^*(T,T)F_{\dot{u}}|_{t=T} = 0.$$

Taking the total derivative, we obtain

$$(\lambda_t + \lambda_T)^*(T, T)F_{\dot{y}}|_{t=T} + \lambda^*(T, T)\frac{dF_{\dot{y}}}{dt}|_{t=T} = 0.$$

Since  $\lambda_t$  is just  $\dot{\lambda}$ , we have the boundary condition

$$(\lambda_T^* F_{\dot{y}})|_{t=T} = -\left[\lambda^*(T, T)\frac{dF_{\dot{y}}}{dt} + \dot{\lambda}^* F_{\dot{y}}\right]|_{t=T}.$$

For the index-one DAE case, the above relation and (2.20) yield

$$(\lambda_T^* F_{\dot{y}})|_{t=T} = [g_y - \lambda^* F_y]|_{t=T}. \tag{2.26}$$

For the regular implicit ODE case,  $F_{\dot{y}}$  is invertible; thus we have  $\lambda(T,T)=0$ , which leads to  $\lambda_T(T)=-\dot{\lambda}(T)$ . As with the final conditions for  $\lambda(T)$  in (2.20), the above selection for  $\lambda_T(T)$  is not sufficient for index-two Hessenberg DAEs (see [13] for details).

#### 2.6.3 Checkpointing scheme

During the backward integration, the evaluation of the right-hand side of the adjoint system requires, at the current time, the states y which were computed during the forward integration phase. Since IDAS implements variable-step integration formulas, it is unlikely that the states will be available at the desired time and so some form of interpolation is needed. The IDAS implementation being also variable-order, it is possible that during the forward integration phase the order may be reduced as low as first order, which means that there may be points in time where only y and  $\dot{y}$  are available. These requirements therefore limit the choices for possible interpolation schemes. IDAS implements two interpolation methods: a cubic Hermite interpolation algorithm and a variable-degree polynomial interpolation method which attempts to mimic the BDF interpolant for the forward integration.

However, especially for large-scale problems and long integration intervals, the number and size of the vectors y and  $\dot{y}$  that would need to be stored make this approach computationally intractable. Thus, IDAS settles for a compromise between storage space and execution time by implementing a socalled *checkpointing scheme*. At the cost of at most one additional forward integration, this approach offers the best possible estimate of memory requirements for adjoint sensitivity analysis. To begin with, based on the problem size N and the available memory, the user decides on the number  $N_d$ of data pairs  $(y, \dot{y})$  if cubic Hermite interpolation is selected, or on the number  $N_d$  of y vectors in the case of variable-degree polynomial interpolation, that can be kept in memory for the purpose of interpolation. Then, during the first forward integration stage, after every  $N_d$  integration steps a checkpoint is formed by saving enough information (either in memory or on disk) to allow for a hot restart, that is a restart which will exactly reproduce the forward integration. In order to avoid storing Jacobian-related data at each checkpoint, a reevaluation of the iteration matrix is forced before each checkpoint. At the end of this stage, we are left with  $N_c$  checkpoints, including one at  $t_0$ . During the backward integration stage, the adjoint variables are integrated backwards from T to  $t_0$ , going from one checkpoint to the previous one. The backward integration from checkpoint i+1 to checkpoint iis preceded by a forward integration from i to i+1 during which the  $N_d$  vectors y (and, if necessary  $\dot{y}$ ) are generated and stored in memory for interpolation<sup>1</sup>

<sup>&</sup>lt;sup>1</sup>The degree of the interpolation polynomial is always that of the current BDF order for the forward interpolation at

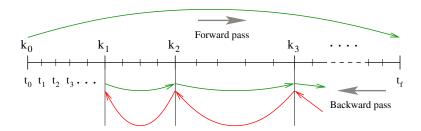


Figure 2.1: Illustration of the checkpointing algorithm for generation of the forward solution during the integration of the adjoint system.

This approach transfers the uncertainty in the number of integration steps in the forward integration phase to uncertainty in the final number of checkpoints. However,  $N_c$  is much smaller than the number of steps taken during the forward integration, and there is no major penalty for writing/reading the checkpoint data to/from a temporary file. Note that, at the end of the first forward integration stage, interpolation data are available from the last checkpoint to the end of the interval of integration. If no checkpoints are necessary ( $N_d$  is larger than the number of integration steps taken in the solution of (2.2)), the total cost of an adjoint sensitivity computation can be as low as one forward plus one backward integration. In addition, IDAS provides the capability of reusing a set of checkpoints for multiple backward integrations, thus allowing for efficient computation of gradients of several functionals (2.17).

Finally, we note that the adjoint sensitivity module in IDAS provides the necessary infrastructure to integrate backwards in time any DAE terminal value problem dependent on the solution of the IVP (2.2), including adjoint systems (2.20) or (2.25), as well as any other quadrature ODEs that may be needed in evaluating the integrals in (2.21). In particular, for DAE systems arising from semi-discretization of time-dependent PDEs, this feature allows for integration of either the discretized adjoint PDE system or the adjoint of the discretized PDE.

# 2.7 Second-order sensitivity analysis

In some applications (e.g., dynamically-constrained optimization) it may be desirable to compute second-order derivative information. Considering the DAE problem (2.2) and some model output functional<sup>2</sup> g(y), the Hessian  $d^2g/dp^2$  can be obtained in a forward sensitivity analysis setting as

$$\frac{d^2g}{dp^2} = \left(g_y \otimes I_{N_p}\right) y_{pp} + y_p^T g_{yy} y_p ,$$

where  $\otimes$  is the Kronecker product. The second-order sensitivities are solution of the matrix DAE system:

$$(F_{\dot{y}} \otimes I_{N_p}) \cdot \dot{y}_{pp} + (F_y \otimes I_{N_p}) \cdot y_{pp} + (I_N \otimes \dot{y}_p^T) \cdot (F_{\dot{y}\dot{y}}\dot{y}_p + F_{y\dot{y}}y_p) + (I_N \otimes y_p^T) \cdot (F_{y\dot{y}}\dot{y}_p + F_{yy}y_p) = 0$$

$$y_{pp}(t_0) = \frac{\partial^2 y_0}{\partial p^2} , \quad \dot{y}_{pp}(t_0) = \frac{\partial^2 \dot{y}_0}{\partial p^2} ,$$

the first point to the right of the time at which the interpolated value is sought (unless too close to the i-th checkpoint, in which case it uses the BDF order at the right-most relevant point). However, because of the FLC BDF implementation (see §2.1), the resulting interpolation polynomial is only an approximation to the underlying BDF interpolant.

The Hermite cubic interpolation option is present because it was implemented chronologically first and it is also used by other adjoint solvers (e.g. DASPKADJOINT). The variable-degree polynomial is more memory-efficient (it requires only half of the memory storage of the cubic Hermite interpolation) and is more accurate.

<sup>&</sup>lt;sup>2</sup>For the sake of simplifity in presentation, we do not include explicit dependencies of g on time t or parameters p. Moreover, we only consider the case in which the dependency of the original DAE (2.2) on the parameters p is through its initial conditions only. For details on the derivation in the general case, see [38].

where  $y_p$  denotes the first-order sensitivity matrix, the solution of  $N_p$  systems (2.12), and  $y_{pp}$  is a third-order tensor. It is easy to see that, except for situations in which the number of parameters  $N_p$  is very small, the computational cost of this so-called *forward-over-forward* approach is exorbitant as it requires the solution of  $N_p + N_p^2$  additional DAE systems of the same dimension as (2.2).

A much more efficient alternative is to compute Hessian-vector products using a so-called forward-over-adjoint approach. This method is based on using the same "trick" as the one used in computing gradients of pointwise functionals with the adjoint method, namely applying a formal directional forward derivation to the gradient of (2.21) (or the equivalent one for a pointwise functional g(T, y(T))). With that, the cost of computing a full Hessian is roughly equivalent to the cost of computing the gradient with forward sensitivity analysis. However, Hessian-vector products can be cheaply computed with one additional adjoint solve.

As an illustration<sup>3</sup>, consider the ODE problem

$$\dot{y} = f(t, y), \quad y(t_0) = y_0(p),$$

depending on some parameters p through the initial conditions only and consider the model functional output  $G(p) = \int_{t_0}^{t_f} g(t, y) dt$ . It can be shown that the product between the Hessian of G (with respect to the parameters p) and some vector u can be computed as

$$\frac{\partial^2 G}{\partial p^2} u = \left[ \left( \lambda^T \otimes I_{N_p} \right) y_{pp} u + y_p^T \mu \right]_{t=t_0},$$

where  $\lambda$  and  $\mu$  are solutions of

$$-\dot{\mu} = f_y^T \mu + (\lambda^T \otimes I_n) f_{yy} s; \quad \mu(t_f) = 0$$
  

$$-\dot{\lambda} = f_y^T \lambda + g_y^T; \quad \lambda(t_f) = 0$$
  

$$\dot{s} = f_y s; \quad s(t_0) = y_{0p} u.$$
(2.27)

In the above equation,  $s = y_p u$  is a linear combination of the columns of the sensitivity matrix  $y_p$ . The forward-over-adjoint approach hinges crucially on the fact that s can be computed at the cost of a forward sensitivity analysis with respect to a single parameter (the last ODE problem above) which is possible due to the linearity of the forward sensitivity equations (2.12).

Therefore (and this is also valid for the DAE case), the cost of computing the Hessian-vector product is roughly that of two forward and two backward integrations of a system of DAEs of size N. For more details, including the corresponding formulas for a pointwise model functional output, see the work by Ozyurt and Barton [38] who discuss this problem for ODE initial value problems. As far as we know, there is no published equivalent work on DAE problems. However, the derivations given in [38] for ODE problems can be extended to DAEs with some careful consideration given to the derivation of proper final conditions on the adjoint systems, following the ideas presented in [13].

To allow the foward-over-adjoint approach described above, IDAS provides support for:

- the integration of multiple backward problems depending on the same underlying forward problem (2.2), and
- the integration of backward problems and computation of backward quadratures depending on both the states y and forward sensitivities (for this particular application, s) of the original problem (2.2).

<sup>&</sup>lt;sup>3</sup>The derivation for the general DAE case is too involved for the purposes of this discussion.

# Chapter 3

# **Code Organization**

# 3.1 SUNDIALS organization

The family of solvers referred to as SUNDIALS consists of the solvers CVODE and ARKODE (for ODE systems), KINSOL (for nonlinear algebraic systems), and IDA (for differential-algebraic systems). In addition, SUNDIALS also includes variants of CVODE and IDA with sensitivity analysis capabilities (using either forward or adjoint methods), called CVODES and IDAS, respectively.

The various solvers of this family share many subordinate modules. For this reason, it is organized as a family, with a directory structure that exploits that sharing (see Figs. 3.1 and 3.2). The following is a list of the solver packages presently available, and the basic functionality of each:

- CVODE, a solver for stiff and nonstiff ODE systems dy/dt = f(t, y) based on Adams and BDF methods;
- CVODES, a solver for stiff and nonstiff ODE systems with sensitivity analysis capabilities;
- ARKODE, a solver for ODE systems  $Mdy/dt = f_E(t, y) + f_I(t, y)$  based on additive Runge-Kutta methods;
- IDA, a solver for differential-algebraic systems  $F(t, y, \dot{y}) = 0$  based on BDF methods;
- IDAS, a solver for differential-algebraic systems with sensitivity analysis capabilities;
- KINSOL, a solver for nonlinear algebraic systems F(u) = 0.

# 3.2 IDAS organization

The IDAS package is written in the ANSI C language. The following summarizes the basic structure of the package, although knowledge of this structure is not necessary for its use.

The overall organization of the IDAS package is shown in Figure 3.3. The central integration module, implemented in the files idas.h, idas\_impl.h, and idas.c, deals with the evaluation of integration coefficients, estimation of local error, selection of stepsize and order, and interpolation to user output points, among other issues.

IDAS utilizes generic linear and nonlinear solver modules defined by the SUNLINSOL API (see Chapter 9) and SUNNONLINSOL API (see Chapter 10) respectively. As such, IDAS has no knowledge of the method being used to solve the linear and nonlinear systems that arise in each time step. For any given user problem, there exists a single nonlinear solver interface and, if necessary, one of the linear system solver interfaces is specified, and invoked as needed during the integration. While SUNDIALS includes a fixed-point nonlinear solver module, it is not currently supported in IDAS (note the fixed-point module is listed in Figure 3.1 but not Figure 3.3).

In addition, if forward sensitivity analysis is turned on, the main module will integrate the forward sensitivity equations simultaneously with the original IVP. The sensitivity variables may be included 26 Code Organization

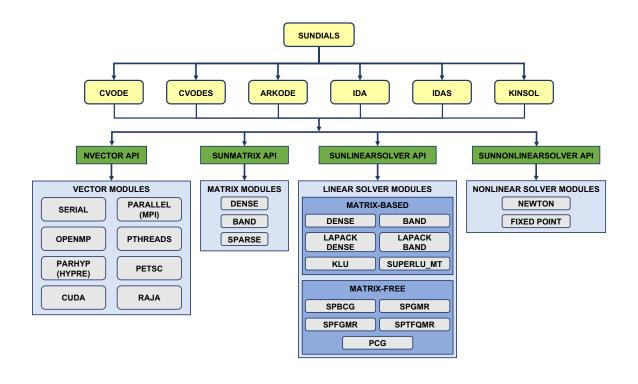


Figure 3.1: High-level diagram of the SUNDIALS suite

in the local error control mechanism of the main integrator. IDAS provides two different strategies for dealing with the correction stage for the sensitivity variables: IDA\_SIMULTANEOUS IDA\_STAGGERED (see §2.5). The IDAS package includes an algorithm for the approximation of the sensitivity equations residuals by difference quotients, but the user has the option of supplying these residual functions directly.

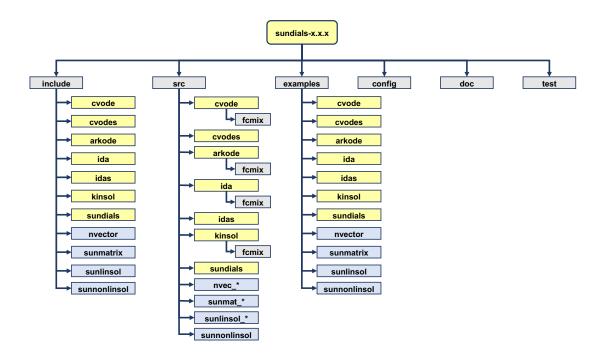
The adjoint sensitivity module (file idaa.c) provides the infrastructure needed for the backward integration of any system of DAEs which depends on the solution of the original IVP, in particular the adjoint system and any quadratures required in evaluating the gradient of the objective functional. This module deals with the setup of the checkpoints, the interpolation of the forward solution during the backward integration, and the backward integration of the adjoint equations.

IDAS now has a single unified linear solver interface, IDALS, supporting both direct and iterative linear solvers built using the generic SUNLINSOL API (see Chapter 9). These solvers may utilize a SUNMATRIX object (see Chapter 8) for storing Jacobian information, or they may be matrix-free. Since IDAS can operate on any valid SUNLINSOL implementation, the set of linear solver modules available to IDAS will expand as new SUNLINSOL modules are developed.

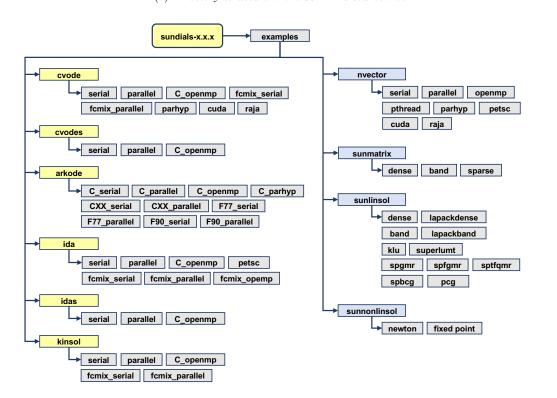
For users employing dense or banded Jacobian matrices, IDALS includes algorithms for their approximation through difference quotients, but the user also has the option of supplying the Jacobian (or an approximation to it) directly. This user-supplied routine is required when using sparse or user-supplied Jacobian matrices.

For users employing matrix-free iterative linear solvers, IDALS includes an algorithm for the approximation by difference quotients of the product between the Jacobian matrix and a vector, Jv. Again, the user has the option of providing routines for this operation, in two phases: setup (preprocessing of Jacobian data) and multiplication.

For preconditioned iterative methods, the preconditioning must be supplied by the user, again in two phases: setup and solve. While there is no default choice of preconditioner analogous to the difference-quotient approximation in the direct case, the references [6, 10], together with the example and demonstration programs included with IDAS, offer considerable assistance in building



(a) Directory structure of the SUNDIALS source tree



(b) Directory structure of the Sundials examples

Figure 3.2: Organization of the SUNDIALS suite

28 Code Organization

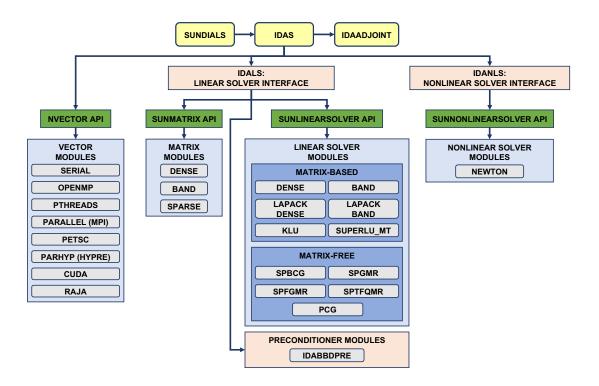


Figure 3.3: Overall structure diagram of the IDA package. Modules specific to IDA begin with "IDA" (IDALS, IDABBDPRE, and IDANLS), all other items correspond to generic solver and auxiliary modules. Note also that the LAPACK, KLU and SUPERLUMT support is through interfaces to external packages. Users will need to download and compile those packages independently.

preconditioners.

IDAS' linear solver interface consists of four primary routines, devoted to (1) memory allocation and initialization, (2) setup of the matrix data involved, (3) solution of the system, and (4) freeing of memory. The setup and solution phases are separate because the evaluation of Jacobians and preconditioners is done only periodically during the integration, as required to achieve convergence. The call list within the central IDAS module to each of the four associated functions is fixed, thus allowing the central module to be completely independent of the linear system method.

IDAS also provides a preconditioner module, IDABBDPRE, for use with any of the Krylov iterative linear solvers. It works in conjunction with NVECTOR\_PARALLEL and generates a preconditioner that is a block-diagonal matrix with each block being a banded matrix.

All state information used by IDAS to solve a given problem is saved in a structure, and a pointer to that structure is returned to the user. There is no global data in the IDAS package, and so, in this respect, it is reentrant. State information specific to the linear solver is saved in a separate structure, a pointer to which resides in the IDAS memory structure. The reentrancy of IDAS was motivated by the situation where two or more problems are solved by intermixed calls to the package from one user program.

# Chapter 4

# Using IDAS for IVP Solution

This chapter is concerned with the use of IDAS for the integration of DAEs in a C language setting. The following sections treat the header files, the layout of the user's main program, description of the IDAS user-callable functions, and description of user-supplied functions. This usage is essentially equivalent to using IDA [30].

The sample programs described in the companion document [43] may also be helpful. Those codes may be used as templates (with the removal of some lines involved in testing), and are included in the IDAS package.

The user should be aware that not all SUNLINSOL and SUNMATRIX modules are compatible with all NVECTOR implementations. Details on compatibility are given in the documentation for each SUNMATRIX module (Chapter 8) and each SUNLINSOL module (Chapter 9). For example, NVECTOR\_PARALLEL is not compatible with the dense, banded, or sparse SUNMATRIX types, or with the corresponding dense, banded, or sparse SUNLINSOL modules. Please check Chapters 8 and 9 to verify compatibility between these modules. In addition to that documentation, we note that the preconditioner module IDABBDPRE can only be used with NVECTOR\_PARALLEL. It is not recommended to use a threaded vector module with SuperLU\_MT unless it is the NVECTOR\_OPENMP module, and SuperLU\_MT is also compiled with OpenMP.

IDAS uses various constants for both input and output. These are defined as needed in this chapter, but for convenience are also listed separately in Appendix B.

# 4.1 Access to library and header files

At this point, it is assumed that the installation of IDAS, following the procedure described in Appendix A, has been completed successfully.

Regardless of where the user's application program resides, its associated compilation and load commands must make reference to the appropriate locations for the library and header files required by IDAS. The relevant library files are

- libdir/libsundials\_idas.lib,
- libdir/libsundials\_nvec\*.lib,

where the file extension .lib is typically .so for shared libraries and .a for static libraries. The relevant header files are located in the subdirectories

- incdir/include/idas
- incdir/include/sundials
- incdir/include/nvector
- incdir/include/sunmatrix

- incdir/include/sunlinsol
- incdir/include/sunnonlinsol

The directories *libdir* and *incdir* are the install library and include directories, respectively. For a default installation, these are *instdir*/lib and *instdir*/include, respectively, where *instdir* is the directory where SUNDIALS was installed (see Appendix A).

Note that an application cannot link to both the IDA and IDAS libraries because both contain user-callable functions with the same names (to ensure that IDAS is backward compatible with IDA). Therefore, applications that contain both DAE problems and DAEs with sensitivity analysis, should use IDAS.

# 4.2 Data types

The sundials\_types.h file contains the definition of the type realtype, which is used by the SUNDIALS solvers for all floating-point data, the definition of the integer type sunindextype, which is used for vector and matrix indices, and booleantype, which is used for certain logic operations within SUNDIALS.

# 4.2.1 Floating point types

The type realtype can be float, double, or long double, with the default being double. The user can change the precision of the SUNDIALS solvers arithmetic at the configuration stage (see §A.1.2).

Additionally, based on the current precision, sundials\_types.h defines BIG\_REAL to be the largest value representable as a realtype, SMALL\_REAL to be the smallest value representable as a realtype, and UNIT\_ROUNDOFF to be the difference between 1.0 and the minimum realtype greater than 1.0.

Within SUNDIALS, real constants are set by way of a macro called RCONST. It is this macro that needs the ability to branch on the definition realtype. In ANSI C, a floating-point constant with no suffix is stored as a double. Placing the suffix "F" at the end of a floating point constant makes it a float, whereas using the suffix "L" makes it a long double. For example,

```
#define A 1.0
#define B 1.0F
#define C 1.0L
```

defines A to be a double constant equal to 1.0, B to be a float constant equal to 1.0, and C to be a long double constant equal to 1.0. The macro call RCONST(1.0) automatically expands to 1.0 if realtype is double, to 1.0F if realtype is float, or to 1.0L if realtype is long double. SUNDIALS uses the RCONST macro internally to declare all of its floating-point constants.

A user program which uses the type realtype and the RCONST macro to handle floating-point constants is precision-independent except for any calls to precision-specific standard math library functions. (Our example programs use both realtype and RCONST.) Users can, however, use the type double, float, or long double in their code (assuming that this usage is consistent with the typedef for realtype). Thus, a previously existing piece of ANSI C code can use SUNDIALS without modifying the code to use realtype, so long as the SUNDIALS libraries use the correct precision (for details see  $\S A.1.2$ ).

# 4.2.2 Integer types used for vector and matrix indices

The type sunindextype can be either a 32- or 64-bit signed integer. The default is the portable int64\_t type, and the user can change it to int32\_t at the configuration stage. The configuration system will detect if the compiler does not support portable types, and will replace int32\_t and int64\_t with int and long int, respectively, to ensure use of the desired sizes on Linux, Mac OS X, and Windows platforms. SUNDIALS currently does not support unsigned integer types for vector and matrix indices, although these could be added in the future if there is sufficient demand.

4.3 Header files 33

A user program which uses sunindextype to handle vector and matrix indices will work with both index storage types except for any calls to index storage-specific external libraries. (Our C and C++ example programs use sunindextype.) Users can, however, use any one of int, long int, int32\_t, int64\_t or long long int in their code, assuming that this usage is consistent with the typedef for sunindextype on their architecture). Thus, a previously existing piece of ANSI C code can use SUNDIALS without modifying the code to use sunindextype, so long as the SUNDIALS libraries use the appropriate index storage type (for details see §A.1.2).

# 4.3 Header files

The calling program must include several header files so that various macros and data types can be used. The header file that is always required is:

• idas/idas.h, the header file for IDAS, which defines the several types and various constants, and includes function prototypes. This includes the header file for IDALS, ida/ida\_ls.h.

Note that idas.h includes sundials\_types.h, which defines the types realtype, sunindextype, and booleantype and the constants SUNFALSE and SUNTRUE.

The calling program must also include an NVECTOR implementation header file, of the form nvector/nvector\_\*\*\*.h. See Chapter 7 for the appropriate name. This file in turn includes the header file sundials\_nvector.h which defines the abstract N\_Vector data type.

If using a non-default nonlinear solver module, or when interacting with a SUNNONLINSOL module directly, the calling program must also include a SUNNONLINSOL implementation header file, of the form sunnonlinsol/sunnonlinsol\_\*\*\*.h where \*\*\* is the name of the nonlinear solver module (see Chapter 10 for more information). This file in turn includes the header file sundials\_nonlinearsolver.h which defines the abstract SUNNonlinearSolver data type.

If using a nonlinear solver that requires the solution of a linear system of the form (2.5) (e.g., the default Newton iteration), a linear solver module header file is also required. The header files corresponding to the various SUNDIALS-provided linear solver modules available for use with IDAS are:

#### • Direct linear solvers:

- sunlinsol/sunlinsol\_dense.h, which is used with the dense linear solver module, SUN-LINSOL\_DENSE;
- sunlinsol/sunlinsol\_band.h, which is used with the banded linear solver module, SUN-LINSOL\_BAND;
- sunlinsol/sunlinsol\_lapackdense.h, which is used with the LAPACK dense linear solver module, SUNLINSOL\_LAPACKDENSE;
- sunlinsol/sunlinsol\_lapackband.h, which is used with the LAPACK banded linear solver module, SUNLINSOL\_LAPACKBAND;
- sunlinsol/sunlinsol\_klu.h, which is used with the KLU sparse linear solver module, SUNLINSOL\_KLU;
- sunlinsol/sunlinsol\_superlumt.h, which is used with the SUPERLUMT sparse linear solver module, SUNLINSOL\_SUPERLUMT;

#### • Iterative linear solvers:

- sunlinsol/sunlinsol\_spgmr.h, which is used with the scaled, preconditioned GMRES Krylov linear solver module, SUNLINSOL\_SPGMR;
- sunlinsol/sunlinsol\_spfgmr.h, which is used with the scaled, preconditioned FGMRES Krylov linear solver module, SUNLINSOL\_SPFGMR;
- sunlinsol/sunlinsol\_spbcgs.h, which is used with the scaled, preconditioned Bi-CGStab Krylov linear solver module, SUNLINSOL\_SPBCGS;

- sunlinsol/sunlinsol\_sptfqmr.h, which is used with the scaled, preconditioned TFQMR
   Krylov linear solver module, SUNLINSOL\_SPTFQMR;
- sunlinsol/sunlinsol\_pcg.h, which is used with the scaled, preconditioned CG Krylov linear solver module, SUNLINSOL\_PCG;

The header files for the SUNLINSOL\_DENSE and SUNLINSOL\_LAPACKDENSE linear solver modules include the file sunmatrix/sunmatrix\_dense.h, which defines the SUNMATRIX\_DENSE matrix module, as as well as various functions and macros acting on such matrices.

The header files for the SUNLINSOL\_BAND and SUNLINSOL\_LAPACKBAND linear solver modules include the file sunmatrix/sunmatrix\_band.h, which defines the SUNMATRIX\_BAND matrix module, as as well as various functions and macros acting on such matrices.

The header files for the SUNLINSOL\_KLU and SUNLINSOL\_SUPERLUMT sparse linear solvers include the file sunmatrix\_sparse.h, which defines the SUNMATRIX\_SPARSE matrix module, as well as various functions and macros acting on such matrices.

The header files for the Krylov iterative solvers include the file sundials\_iterative.h, which enumerates the kind of preconditioning, and (for the SPGMR and SPFGMR solvers) the choices for the Gram-Schmidt process.

Other headers may be needed, according to the choice of preconditioner, etc. For example, in the idasFoodWeb\_kry\_p example (see [43]), preconditioning is done with a block-diagonal matrix. For this, even though the SUNLINSOL\_SPGMR linear solver is used, the header sundials/sundials\_dense.h is included for access to the underlying generic dense matrix arithmetic routines.

# 4.4 A skeleton of the user's main program

The following is a skeleton of the user's main program (or calling program) for the integration of a DAE IVP. Most of the steps are independent of the NVECTOR, SUNMATRIX, SUNLINSOL, and SUNNONLINSOL implementations used. For the steps that are not, refer to Chapter 7, 8, 9, and 10 for the specific name of the function to be called or macro to be referenced.

# 1. Initialize parallel or multi-threaded environment, if appropriate

For example, call MPI\_Init to initialize MPI if used, or set num\_threads, the number of threads to use within the threaded vector functions, if used.

# 2. Set problem dimensions etc.

This generally includes the problem size N, and may include the local vector length Nlocal.

Note: The variables N and Nlocal should be of type sunindextype.

#### 3. Set vectors of initial values

To set the vectors y0 and yp0 to initial values for y and  $\dot{y}$ , use the appropriate functions defined by the particular NVECTOR implementation.

For native SUNDIALS vector implementations (except the CUDA and RAJA-based ones), use a call of the form  $y0 = N_VMake_***(..., ydata)$  if the realtype array ydata containing the initial values of y already exists. Otherwise, create a new vector by making a call of the form  $y0 = N_VMew_***(...)$ , and then set its elements by accessing the underlying data with a call of the form ydata =  $N_VGetArrayPointer(y0)$ . See §7.2-7.5 for details.

For the hypre and PETSc vector wrappers, first create and initialize the underlying vector and then create an NVECTOR wrapper with a call of the form y0 = N\_VMake\_\*\*\*(yvec), where yvec is a hypre or PETSc vector. Note that calls like N\_VNew\_\*\*\*(...) and N\_VGetArrayPointer(...) are not available for these vector wrappers. See §7.6 and §7.7 for details.

If using either the CUDA- or RAJA-based vector implementations use a call of the form y0 = N\_VMake\_\*\*\*(..., c) where c is a pointer to a suncudavec or sunrajavec vector class if this class already exists. Otherwise, create a new vector by making a call of the form y0 = N\_VNew\_\*\*\*(...),

and then set its elements by accessing the underlying data where it is located with a call of the form N\_VGetDeviceArrayPointer\_\*\*\* or N\_VGetHostArrayPointer\_\*\*\*. Note that the vector class will allocate memory on both the host and device when instantiated. See §7.8-7.9 for details.

Set the vector yp0 of initial conditions for  $\dot{y}$  similarly.

#### 4. Create IDAS object

Call ida\_mem = IDACreate() to create the IDAS memory block. IDACreate returns a pointer to the IDAS memory structure. See §4.5.1 for details. This void \* pointer must then be passed as the first argument to all subsequent IDAS function calls.

#### 5. Initialize IDAS solver

Call IDAInit(...) to provide required problem specifications (residual function, initial time, and initial conditions), allocate internal memory for IDAS, and initialize IDAS. IDAInit returns an error flag to indicate success or an illegal argument value. See §4.5.1 for details.

#### 6. Specify integration tolerances

Call IDASStolerances(...) or IDASVtolerances(...) to specify, respectively, a scalar relative tolerance and scalar absolute tolerance, or a scalar relative tolerance and a vector of absolute tolerances. Alternatively, call IDAWFtolerances to specify a function which sets directly the weights used in evaluating WRMS vector norms. See §4.5.2 for details.

#### 7. Create matrix object

If a nonlinear solver requiring a linear solver will be used (e.g., the default Newton iteration) and the linear solver will be a matrix-based linear solver, then a template Jacobian matrix must be created by using the appropriate constructor function defined by the particular SUNMATRIX implementation.

For the SUNDIALS-supplied SUNMATRIX implementations, the matrix object may be created using a call of the form

```
SUNMatrix J = SUNBandMatrix(...);
or
SUNMatrix J = SUNDenseMatrix(...);
or
SUNMatrix J = SUNSparseMatrix(...);
```

NOTE: The dense, banded, and sparse matrix objects are usable only in a serial or threaded environment.

#### 8. Create linear solver object

If a nonlinear solver requiring a linear solver is chosen (e.g., the default Newton iteration), then the desired linear solver object must be created by calling the appropriate constructor function defined by the particular SUNLINSOL implementation.

For any of the SUNDIALS-supplied SUNLINSOL implementations, the linear solver object may be created using a call of the form

```
SUNLinearSolver LS = SUNLinSol_*(...);
```

where \* can be replaced with "Dense", "SPGMR", or other options, as discussed in §4.5.3 and Chapter 9.

#### 9. Set linear solver optional inputs

Call \*Set\* functions from the selected linear solver module to change optional inputs specific to that linear solver. See the documentation for each SUNLINSOL module in Chapter 9 for details.

#### 10. Attach linear solver module

If a nonlinear solver requiring a linear solver is chosen (e.g., the default Newton iteration), then initialize the IDALS linear solver interface by attaching the linear solver object (and matrix object, if applicable) with the following call (for details see §4.5.3):

```
ier = IDASetLinearSolver(...);
```

#### 11. Set optional inputs

Optionally, call IDASet\* functions to change from their default values any optional inputs that control the behavior of IDAS. See §4.5.8.1 and §4.5.8 for details.

# 12. Create nonlinear solver object (optional)

If using a non-default nonlinear solver (see §4.5.4), then create the desired nonlinear solver object by calling the appropriate constructor function defined by the particular SUNNONLINSOL implementation (e.g., NLS = SUNNonlinSol\_\*\*\*(...); where \*\*\* is the name of the nonlinear solver (see Chapter 10 for details).

# 13. Attach nonlinear solver module (optional)

If using a non-default nonlinear solver, then initialize the nonlinear solver interface by attaching the nonlinear solver object by calling ier = IDASetNonlinearSolver(ida\_mem, NLS); (see §4.5.4 for details).

#### 14. Set nonlinear solver optional inputs (optional)

Call the appropriate set functions for the selected nonlinear solver module to change optional inputs specific to that nonlinear solver. These *must* be called after IDAInit if using the default nonlinear solver or after attaching a new nonlinear solver to IDAS, otherwise the optional inputs will be overridden by IDAS defaults. See Chapter 10 for more information on optional inputs.

#### 15. Correct initial values

Optionally, call IDACalcIC to correct the initial values y0 and yp0 passed to IDAInit. See §4.5.5. Also see §4.5.8.3 for relevant optional input calls.

#### 16. Specify rootfinding problem

Optionally, call IDARootInit to initialize a rootfinding problem to be solved during the integration of the DAE system. See  $\S4.5.6$  for details, and see  $\S4.5.8.4$  for relevant optional input calls.

# 17. Advance solution in time

For each point at which output is desired, call flag = IDASolve(ida\_mem, tout, &tret, yret, yret, itask). Here itask specifies the return mode. The vector yret (which can be the same as the vector y0 above) will contain y(t), while the vector ypret (which can be the same as the vector yp0 above) will contain  $\dot{y}(t)$ . See §4.5.7 for details.

#### 18. Get optional outputs

Call IDA\*Get\* functions to obtain optional output. See §4.5.10 for details.

#### 19. Deallocate memory for solution vectors

Upon completion of the integration, deallocate memory for the vectors yret and ypret (or y and yp) by calling the appropriate destructor function defined by the NVECTOR implementation:

```
N_VDestroy(yret);
```

and similarly for ypret.

#### 20. Free solver memory

IDAFree (&ida\_mem) to free the memory allocated for IDAS.

#### 21. Free nonlinear solver memory (optional)

If a non-default nonlinear solver was used, then call SUNNonlinSolFree(NLS) to free any memory allocated for the SUNNONLINSOL object.

#### 22. Free linear solver and matrix memory

Call SUNLinSolFree and SUNMatDestroy to free any memory allocated for the linear solver and matrix objects created above.

#### 23. Finalize MPI, if used

Call MPI\_Finalize() to terminate MPI.

SUNDIALS provides some linear solvers only as a means for users to get problems running and not as highly efficient solvers. For example, if solving a dense system, we suggest using the LAPACK solvers if the size of the linear system is > 50,000. (Thanks to A. Nicolai for his testing and recommendation.) Table 4.1 shows the linear solver interfaces available as SUNLINSOL modules and the vector implementations required for use. As an example, one cannot use the dense direct solver interfaces with the MPI-based vector implementation. However, as discussed in Chapter 9 the SUNDIALS packages operate on generic SUNLINSOL objects, allowing a user to develop their own solvers should they so desire.

Linear Solver	Serial	Parallel (MPI)	OpenMP	pThreads	hypre	PETSC	CUDA	RAJA	User Supp.
Dense	<b>√</b>		<b>√</b>	<b>√</b>					✓
Band	<b>√</b>		<b>√</b>	<b>√</b>					✓
LapackDense	<b>√</b>		<b>√</b>	<b>√</b>					✓
LapackBand	<b>√</b>		<b>√</b>	<b>√</b>					<b>√</b>
KLU	<b>√</b>		<b>√</b>	<b>√</b>					✓
SUPERLUMT	<b>√</b>		<b>√</b>	<b>√</b>					<b>√</b>
SPGMR	<b>√</b>	✓	<b>√</b>						
SPFGMR	<b>√</b>	✓	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>✓</b>
SPBCGS	<b>√</b>	✓	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>✓</b>
SPTFQMR	<b>√</b>	✓	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	✓
PCG	<b>√</b>	✓	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	✓
User Supp.	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>	✓

Table 4.1: SUNDIALS linear solver interfaces and vector implementations that can be used for each.

# 4.5 User-callable functions

This section describes the IDAS functions that are called by the user to set up and solve a DAE. Some of these are required. However, starting with §4.5.8, the functions listed involve optional inputs/outputs or restarting, and those paragraphs can be skipped for a casual use of IDAS. In any case, refer to §4.4 for the correct order of these calls.

On an error, each user-callable function returns a negative value and sends an error message to the error handler routine, which prints the message on **stderr** by default. However, the user can set a file as error output or can provide his own error handler function (see §4.5.8.1).

# 4.5.1 IDAS initialization and deallocation functions

The following three functions must be called in the order listed. The last one is to be called only after the DAE solution is complete, as it frees the IDAS memory block created and allocated by the first two calls.

#### IDACreate

Call ida\_mem = IDACreate();

Description The function IDACreate instantiates an IDAS solver object.

Arguments IDACreate has no arguments.

Return value If successful, IDACreate returns a pointer to the newly created IDAS memory block (of type void \*). Otherwise it returns NULL.

#### IDAInit

Call flag = IDAInit(ida\_mem, res, t0, y0, yp0);

Description The function IDAInit provides required problem and solution specifications, allocates

internal memory, and initializes IDAS.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

res (IDAResFn) is the C function which computes the residual function F in the DAE. This function has the form res(t, yy, yp, resval, user\_data). For

full details see  $\S4.6.1$ .

to (realtype) is the initial value of t.

y0 (N\_Vector) is the initial value of y.

yp0 (N\_Vector) is the initial value of  $\dot{y}$ .

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAInit was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_MEM\_FAIL A memory allocation request has failed.

IDA\_ILL\_INPUT An input argument to IDAInit has an illegal value.

Notes If an error occurred, IDAInit also sends an error message to the error handler function.

#### IDAFree

Call IDAFree(&ida\_mem);

Description The function IDAFree frees the pointer allocated by a previous call to IDACreate.

Arguments The argument is the pointer to the IDAS memory block (of type void \*).

Return value The function IDAFree has no return value.

### 4.5.2 IDAS tolerance specification functions

One of the following three functions must be called to specify the integration tolerances (or directly specify the weights used in evaluating WRMS vector norms). Note that this call must be made after the call to IDAInit.

#### IDASStolerances

Call flag = IDASStolerances(ida\_mem, reltol, abstol);

Description The function IDASStolerances specifies scalar relative and absolute tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

reltol (realtype) is the scalar relative error tolerance. abstol (realtype) is the scalar absolute error tolerance.

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASStolerances was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to IDACreate.

IDA\_NO\_MALLOC The allocation function IDAInit has not been called.

IDA\_ILL\_INPUT One of the input tolerances was negative.

#### IDASVtolerances

Call flag = IDASVtolerances(ida\_mem, reltol, abstol);

 ${\bf Description} \quad {\bf The \ function \ IDASV tolerances \ specifies \ scalar \ relative \ tolerance \ and \ vector \ absolute}$ 

tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

reltol (realtype) is the scalar relative error tolerance.

abstol (N\_Vector) is the vector of absolute error tolerances.

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASVtolerances was successful.

 ${\tt IDA\_MEM\_NULL} \quad {\tt The\ IDAS\ memory\ block\ was\ not\ initialized\ through\ a\ previous\ call\ to}$ 

 ${\tt IDACreate}.$ 

IDA\_NO\_MALLOC The allocation function IDAInit has not been called.

 ${\tt IDA\_ILL\_INPUT \ The \ relative \ error \ tolerance \ was \ negative \ or \ the \ absolute \ tolerance \ had}$ 

a negative component.

Notes This choice of tolerances is important when the absolute error tolerance needs to be

different for each component of the state vector y.

# IDAWFtolerances

Call flag = IDAWFtolerances(ida\_mem, efun);

Description The function IDAWFtolerances specifies a user-supplied function efun that sets the

multiplicative error weights  $W_i$  for use in the weighted RMS norm, which are normally

defined by Eq. (2.7).

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

efun (IDAEwtFn) is the C function which defines the ewt vector (see §4.6.3).

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAWFtolerances was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_NO\_MALLOC The allocation function IDAInit has not been called.

General advice on choice of tolerances. For many users, the appropriate choices for tolerance values in reltol and abstol are a concern. The following pieces of advice are relevant.

(1) The scalar relative tolerance reltol is to be set to control relative errors. So reltol= $10^{-4}$  means that errors are controlled to .01%. We do not recommend using reltol larger than  $10^{-3}$ .

On the other hand, reltol should not be so small that it is comparable to the unit roundoff of the machine arithmetic (generally around  $10^{-15}$ ).

- (2) The absolute tolerances abstol (whether scalar or vector) need to be set to control absolute errors when any components of the solution vector y may be so small that pure relative error control is meaningless. For example, if y[i] starts at some nonzero value, but in time decays to zero, then pure relative error control on y[i] makes no sense (and is overly costly) after y[i] is below some noise level. Then abstol (if scalar) or abstol[i] (if a vector) needs to be set to that noise level. If the different components have different noise levels, then abstol should be a vector. See the example idasRoberts\_dns in the IDAS package, and the discussion of it in the IDAS Examples document [43]. In that problem, the three components vary between 0 and 1, and have different noise levels; hence the abstol vector. It is impossible to give any general advice on abstol values, because the appropriate noise levels are completely problem-dependent. The user or modeler hopefully has some idea as to what those noise levels are.
- (3) Finally, it is important to pick all the tolerance values conservatively, because they control the error committed on each individual time step. The final (global) errors are a sort of accumulation of those per-step errors. A good rule of thumb is to reduce the tolerances by a factor of .01 from the actual desired limits on errors. So if you want .01% accuracy (globally), a good choice is  $reltol=10^{-6}$ . But in any case, it is a good idea to do a few experiments with the tolerances to see how the computed solution values vary as tolerances are reduced.

Advice on controlling unphysical negative values. In many applications, some components in the true solution are always positive or non-negative, though at times very small. In the numerical solution, however, small negative (hence unphysical) values can then occur. In most cases, these values are harmless, and simply need to be controlled, not eliminated. The following pieces of advice are relevant.

- (1) The way to control the size of unwanted negative computed values is with tighter absolute tolerances. Again this requires some knowledge of the noise level of these components, which may or may not be different for different components. Some experimentation may be needed.
- (2) If output plots or tables are being generated, and it is important to avoid having negative numbers appear there (for the sake of avoiding a long explanation of them, if nothing else), then eliminate them, but only in the context of the output medium. Then the internal values carried by the solver are unaffected. Remember that a small negative value in yret returned by IDAS, with magnitude comparable to abstol or less, is equivalent to zero as far as the computation is concerned.
- (3) The user's residual routine **res** should never change a negative value in the solution vector yy to a non-negative value, as a "solution" to this problem. This can cause instability. If the **res** routine cannot tolerate a zero or negative value (e.g., because there is a square root or log of it), then the offending value should be changed to zero or a tiny positive number in a temporary variable (not in the input yy vector) for the purposes of computing  $F(t, y, \dot{y})$ .
- (4) IDAS provides the option of enforcing positivity or non-negativity on components. Also, such constraints can be enforced by use of the recoverable error return feature in the user-supplied residual function. However, because these options involve some extra overhead cost, they should only be exercised if the use of absolute tolerances to control the computed values is unsuccessful.

#### 4.5.3 Linear solver interface functions

As previously explained, if the nonlinear solver requires the solution of linear systems of the form (2.5) (e.g., the default Newton iteration, then solution of these linear systems is handled with the IDALS linear solver interface. This interface supports all valid SUNLINSOL modules. Here, matrix-based SUNLINSOL modules utilize SUNMATRIX objects to store the Jacobian matrix  $J = \partial F/\partial y + \alpha \partial F/\partial \dot{y}$  and factorizations used throughout the solution process. Conversely, matrix-free SUNLINSOL modules instead use iterative methods to solve the linear systems of equations, and only require the *action* of the Jacobian on a vector, Jv.

With most iterative linear solvers, preconditioning can be done on the left only, on the right only, on both the left and the right, or not at all. The exceptions to this rule are SPFGMR that supports

right preconditioning only and PCG that performs symmetric preconditioning. However, in IDAS only left preconditioning is supported. For the specification of a preconditioner, see the iterative linear solver sections in  $\S4.5.8$  and  $\S4.6$ . A preconditioner matrix P must approximate the Jacobian J, at least crudely.

To specify a generic linear solver to IDAS, after the call to IDACreate but before any calls to IDASolve, the user's program must create the appropriate SUNLINSOL object and call the function IDASetLinearSolver, as documented below. To create the SUNLinearSolver object, the user may call one of the SUNDIALS-packaged SUNLINSOL module constructor routines via a call of the form

```
SUNLinearSolver LS = SUNLinSol_*(...);
```

The current list of such constructor routines includes SUNLinSol\_Dense, SUNLinSol\_Band, SUNLinSol\_LapackDense, SUNLinSol\_LapackBand, SUNLinSol\_KLU, SUNLinSol\_SUNLinSol\_SPECGN, SUNLinSol\_SPECGN, SUNLinSol\_SPECGN

Alternately, a user-supplied SUNLinearSolver module may be created and used instead. The use of each of the generic linear solvers involves certain constants, functions and possibly some macros, that are likely to be needed in the user code. These are available in the corresponding header file associated with the specific SUNMATRIX or SUNLINSOL module in question, as described in Chapters 8 and 9.

Once this solver object has been constructed, the user should attach it to IDAS via a call to IDASetLinearSolver. The first argument passed to this function is the IDAS memory pointer returned by IDACreate; the second argument is the desired SUNLINSOL object to use for solving systems. The third argument is an optional SUNMATRIX object to accompany matrix-based SUNLINSOL inputs (for matrix-free linear solvers, the third argument should be NULL). A call to this function initializes the IDALS linear solver interface, linking it to the main IDAS integrator, and allows the user to specify additional parameters and routines pertinent to their choice of linear solver.

#### IDASetLinearSolver

Call flag = IDASetLinearSolver(ida\_mem, LS, J);

Description The function IDASetLinearSolver attaches a generic SUNLINSOL object LS and corresponding template Jacobian SUNMATRIX object J (if applicable) to IDAS, initializing the

IDALS linear solver interface.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

LS (SUNLinearSolver) SUNLINSOL object to use for solving linear systems of the form (2.5.

J (SUNMatrix) SUNMATRIX object for used as a template for the Jacobian (or NULL if not applicable).

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The IDALS initialization was successful.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_ILL\_INPUT The IDALS interface is not compatible with the LS or J input objects or is incompatible with the current NVECTOR module.

IDALS\_SUNLS\_FAIL A call to the LS object failed.

IDALS\_MEM\_FAIL A memory allocation request failed.

Notes

If LS is a matrix-based linear solver, then the template Jacobian matrix J will be used in the solve process, so if additional storage is required within the SUNMATRIX object (e.g., for factorization of a banded matrix), ensure that the input object is allocated with sufficient size (see the documentation of the particular SUNMATRIX type in Chapter 8 for further information).

The previous routines IDADlsSetLinearSolver and IDASpilsSetLinearSolver are now wrappers for this routine, and may still be used for backward-compatibility. However, these will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### 4.5.4 Nonlinear solver interface function

By default IDAS uses the SUNNONLINSOL implementation of Newton's method defined by the SUNNON-LINSOL\_NEWTON module (see §10.2). To specify a different nonlinear solver in IDAS, the user's program must create a SUNNONLINSOL object by calling the appropriate constructor routine. The user must then attach the SUNNONLINSOL object to IDAS by calling IDASetNonlinearSolver, as documented below.

When changing the nonlinear solver in IDAS, IDASetNonlinearSolver must be called after IDAInit. If any calls to IDASolve have been made, then IDAS will need to be reinitialized by calling IDAReInit to ensure that the nonlinear solver is initialized correctly before any subsequent calls to IDASolve.

The first argument passed to the routine IDASetNonlinearSolver is the IDAS memory pointer returned by IDACreate and the second argument is the SUNNONLINSOL object to use for solving the nonlinear system 2.4. A call to this function attaches the nonlinear solver to the main IDAS integrator. We note that at present, the SUNNONLINSOL object must be of type SUNNONLINEARSOLVER\_ROOTFIND.

# IDASetNonlinearSolver

Call flag = IDASetNonlinearSolver(ida\_mem, NLS);

Description The function IDASetNonLinearSolver attaches a SUNNONLINSOL object (NLS) to IDAS.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

NLS (SUNNonlinearSolver) SUNNONLINSOL object to use for solving nonlinear systems.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The nonlinear solver was successfully attached.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The SUNNONLINSOL object is NULL, does not implement the required nonlinear solver operations, is not of the correct type, or the residual function, convergence test function, or maximum number of nonlinear

iterations could not be set.

Notes When forward sensitivity analysis capabilities are enabled and the IDA\_STAGGERED corrector method is used this function sets the nonlinear solver method for correcting state

variables (see §5.2.3 for more details).

# 4.5.5 Initial condition calculation function

IDACalcIC calculates corrected initial conditions for the DAE system for certain index-one problems including a class of systems of semi-implicit form. (See §2.1 and Ref. [8].) It uses Newton iteration combined with a linesearch algorithm. Calling IDACalcIC is optional. It is only necessary when the initial conditions do not satisfy the given system. Thus if y0 and yp0 are known to satisfy  $F(t_0, y_0, \dot{y}_0) = 0$ , then a call to IDACalcIC is generally not necessary.

A call to the function IDACalcIC must be preceded by successful calls to IDACreate and IDAInit (or IDAReInit), and by a successful call to the linear system solver specification function. The call to IDACalcIC should precede the call(s) to IDASolve for the given problem.

#### IDACalcIC

Call flag = IDACalcIC(ida\_mem, icopt, tout1);

Description The function IDACalcIC corrects the initial values y0 and yp0 at time t0.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

icopt (int) is one of the following two options for the initial condition calculation.

icopt=IDA\_YA\_YDP\_INIT directs IDACalcIC to compute the algebraic components of y and differential components of  $\dot{y}$ , given the differential components

of y. This option requires that the N\_Vector id was set through IDASetId, specifying the differential and algebraic components.

 $icopt=IDA\_Y\_INIT$  directs IDACalcIC to compute all components of y, given  $\dot{y}$ . In this case,  $\dot{id}$  is not required.

tout1 (realtype) is the first value of t at which a solution will be requested (from IDASolve). This value is needed here only to determine the direction of integration and rough scale in the independent variable t.

Return value The return value flag (of type int) will be one of the following:

IDA_SUCCESS	IDASolve succeeded.
IDA_MEM_NULL	The argument ida_mem was NULL.
IDA_NO_MALLOC	The allocation function IDAInit has not been called.
IDA_ILL_INPUT	One of the input arguments was illegal.
IDA_LSETUP_FAIL	The linear solver's setup function failed in an unrecoverable manner. $$
IDA_LINIT_FAIL	The linear solver's initialization function failed.
IDA_LSOLVE_FAIL	The linear solver's solve function failed in an unrecoverable manner. $$
IDA_BAD_EWT	Some component of the error weight vector is zero (illegal), either for the input value of $y0$ or a corrected value.
IDA_FIRST_RES_FAIL	The user's residual function returned a recoverable error flag on the first call, but ${\tt IDACalcIC}$ was unable to recover.
IDA_RES_FAIL	The user's residual function returned a nonrecoverable error flag. $$
IDA_NO_RECOVERY	The user's residual function, or the linear solver's setup or solve function had a recoverable error, but ${\tt IDACalcIC}$ was unable to recover.
IDA_CONSTR_FAIL	${\tt IDACalcIC}$ was unable to find a solution satisfying the inequality constraints.
IDA_LINESEARCH_FAIL	The linesearch algorithm failed to find a solution with a step larger than steptol in weighted RMS norm, and within the allowed number of backtracks.
IDA_CONV_FAIL	IDACalcIC failed to get convergence of the Newton iterations.
A 11 C 11	

Notes

All failure return values are negative and therefore a test  ${\tt flag} < 0$  will trap all IDACalcIC failures.

Note that IDACalcIC will correct the values of  $y(t_0)$  and  $\dot{y}(t_0)$  which were specified in the previous call to IDAInit or IDAReInit. To obtain the corrected values, call IDAGetconsistentIC (see §4.5.10.3).

#### 4.5.6 Rootfinding initialization function

While integrating the IVP, IDAS has the capability of finding the roots of a set of user-defined functions. To activate the rootfinding algorithm, call the following function. This is normally called only once, prior to the first call to IDASolve, but if the rootfinding problem is to be changed during the solution, IDARootInit can also be called prior to a continuation call to IDASolve.

# IDARootInit

Call flag = IDARootInit(ida\_mem, nrtfn, g);

Description The function IDARootInit specifies that the roots of a set of functions  $g_i(t,y,\dot{y})$  are to

be found while the IVP is being solved.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

**nrtfn** (int) is the number of root functions  $g_i$ .

g (IDARootFn) is the C function which defines the nrtfn functions  $g_i(t, y, \dot{y})$  whose roots are sought. See §4.6.4 for details.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The call to IDARootInit was successful.

IDA\_MEM\_NULL The ida\_mem argument was NULL.

IDA\_MEM\_FAIL A memory allocation failed.

IDA\_ILL\_INPUT The function g is NULL, but nrtfn> 0.

Notes If a new IVP is to be solved with a call to IDAReInit, where the new IVP has no rootfinding problem but the prior one did, then call IDARootInit with nrtfn= 0.

#### 4.5.7 IDAS solver function

This is the central step in the solution process, the call to perform the integration of the DAE. One of the input arguments (itask) specifies one of two modes as to where IDAS is to return a solution. But these modes are modified if the user has set a stop time (with IDASetStopTime) or requested rootfinding.

#### IDASolve

Call flag = IDASolve(ida\_mem, tout, &tret, yret, ypret, itask);

Description The function IDASolve integrates the DAE over an interval in t.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

tout (realtype) the next time at which a computed solution is desired.

tret (realtype) the time reached by the solver (output).

yret (N\_Vector) the computed solution vector y.

ypret (N\_Vector) the computed solution vector  $\dot{y}$ .

itask (int) a flag indicating the job of the solver for the next user step. The IDA\_NORMAL task is to have the solver take internal steps until it has reached or just passed the user specified tout parameter. The solver then interpolates in order to return approximate values of y(tout) and  $\dot{y}(\texttt{tout})$ . The IDA\_ONE\_STEP option tells the solver to just take one internal step and return the solution at the point reached by that step.

Return value IDASolve returns vectors yret and ypret and a corresponding independent variable value t = tret, such that (yret, ypret) are the computed values of  $(y(t), \dot{y}(t))$ .

In IDA\_NORMAL mode with no errors, tret will be equal to tout and yret = y(tout), ypret =  $\dot{y}(tout)$ .

The return value flag (of type int) will be one of the following:

IDA\_SUCCESS IDASolve succeeded.

IDA\_TSTOP\_RETURN IDASolve succeeded by reaching the stop point specified through

the optional input function IDASetStopTime.

IDA\_ROOT\_RETURN IDASolve succeeded and found one or more roots. In this case,

tret is the location of the root. If nrtfn > 1, call IDAGetRootInfo to see which  $g_i$  were found to have a root. See §4.5.10.4 for more

information.

IDA\_MEM\_NULL The ida\_mem argument was NULL.

IDA\_ILL\_INPUT One of the inputs to IDASolve was illegal, or some other input

to the solver was either illegal or missing. The latter category includes the following situations: (a) The tolerances have not been set. (b) A component of the error weight vector became zero during

internal time-stepping. (c) The linear solver initialization function (called by the user after calling <code>IDACreate</code>) failed to set the linear solver-specific <code>lsolve</code> field in <code>ida\_mem</code>. (d) A root of one of the root functions was found both at a point t and also very near t. In any case, the user should see the printed error message for details.

 ${\tt IDA\_TOO\_MUCH\_WORK\ The\ solver\ took\ mxstep\ internal\ steps\ but\ could\ not\ reach\ tout.}$ 

The default value for mxstep is MXSTEP\_DEFAULT = 500.

IDA\_TOO\_MUCH\_ACC The solver could not satisfy the accuracy demanded by the user for

some internal step.

IDA\_ERR\_FAIL Error test failures occurred too many times (MXNEF = 10) during

one internal time step or occurred with  $|h| = h_{min}$ .

IDA\_CONV\_FAIL Convergence test failures occurred too many times (MXNCF = 10)

during one internal time step or occurred with  $|h| = h_{min}$ .

IDA\_LINIT\_FAIL The linear solver's initialization function failed.

IDALLSETUP\_FAIL The linear solver's setup function failed in an unrecoverable man-

ner.

IDA\_LSOLVE\_FAIL The linear solver's solve function failed in an unrecoverable manner.

IDA\_CONSTR\_FAIL The inequality constraints were violated and the solver was unable

to recover.

IDA\_REP\_RES\_ERR The user's residual function repeatedly returned a recoverable error

flag, but the solver was unable to recover.

IDA\_RES\_FAIL The user's residual function returned a nonrecoverable error flag.

IDA\_RTFUNC\_FAIL The rootfinding function failed.

The vector yret can occupy the same space as the vector y0 of initial conditions that was passed to IDAInit, and the vector ypret can occupy the same space as yp0.

In the IDA\_ONE\_STEP mode, tout is used on the first call only, and only to get the direction and rough scale of the independent variable.

All failure return values are negative and therefore a test  ${\tt flag} < 0$  will trap all IDASolve failures.

On any error return in which one or more internal steps were taken by IDASolve, the returned values of tret, yret, and ypret correspond to the farthest point reached in the integration. On all other error returns, these values are left unchanged from the previous IDASolve return.

# 4.5.8 Optional input functions

There are numerous optional input parameters that control the behavior of the IDAS solver. IDAS provides functions that can be used to change these optional input parameters from their default values. Table 4.2 lists all optional input functions in IDAS which are then described in detail in the remainder of this section. For the most casual use of IDAS, the reader can skip to §4.6.

We note that, on an error return, all these functions also send an error message to the error handler function. We also note that all error return values are negative, so a test  ${\tt flag} < 0$  will catch any error.

# 4.5.8.1 Main solver optional input functions

The calls listed here can be executed in any order. However, if the user's program calls either IDASetErrFile or IDASetErrHandlerFn, then that call should appear first, in order to take effect for any later error message.

Notes

Table 4.2: Optional inputs for IDAS and IDALS

Optional input	Function name	Default	
IDAS ma	in solver		
Pointer to an error file	IDASetErrFile	stderr	
Error handler function	IDASetErrHandlerFn	internal fn.	
User data	IDASetUserData	NULL	
Maximum order for BDF method	IDASetMaxOrd	5	
Maximum no. of internal steps before $t_{\text{out}}$	IDASetMaxNumSteps	500	
Initial step size	IDASetInitStep	estimated	
Maximum absolute step size	IDASetMaxStep	$\infty$	
Value of $t_{stop}$	IDASetStopTime	$\infty$	
Maximum no. of error test failures	IDASetMaxErrTestFails	10	
Maximum no. of nonlinear iterations	IDASetMaxNonlinIters	4	
Maximum no. of convergence failures	IDASetMaxConvFails	10	
Maximum no. of error test failures	IDASetMaxErrTestFails	7	
Coeff. in the nonlinear convergence test	IDASetNonlinConvCoef	0.33	
Suppress alg. vars. from error test	IDASetSuppressAlg	SUNFALSE	
Variable types (differential/algebraic)	IDASetId	NULL	
Inequality constraints on solution	IDASetConstraints	NULL	
Direction of zero-crossing	IDASetRootDirection	both	
Disable rootfinding warnings	IDASetNoInactiveRootWarn	none	
IDAS initial cond	itions calculation		
Coeff. in the nonlinear convergence test	IDASetNonlinConvCoefIC	0.0033	
Maximum no. of steps	IDASetMaxNumStepsIC	5	
Maximum no. of Jacobian/precond. evals.	IDASetMaxNumJacsIC	4	
Maximum no. of Newton iterations	IDASetMaxNumItersIC	10	
Max. linesearch backtracks per Newton iter.	IDASetMaxBacksIC	100	
Turn off linesearch	IDASetLineSearchOffIC	SUNFALSE	
Lower bound on Newton step	IDASetStepToleranceIC	$uround^{2/3}$	
IDALS linear solver interface			
Jacobian function	IDASetJacFn	DQ	
Jacobian-times-vector function	IDASetJacTimes	NULL, DQ	
Preconditioner functions	IDASetPreconditioner	NULL, NULL	
Ratio between linear and nonlinear tolerances	IDASetEpsLin	0.05	
Increment factor used in DQ $Jv$ approx.	IDASetIncrementFactor	1.0	

#### IDASetErrFile

Call flag = IDASetErrFile(ida\_mem, errfp);

Description The function IDASetErrFile specifies the pointer to the file where all IDAS messages

should be directed when the default IDAS error handler function is used.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

errfp (FILE \*) pointer to output file.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The default value for errfp is stderr.

Passing a value NULL disables all future error message output (except for the case in which the IDAS memory pointer is NULL). This use of IDASetErrFile is strongly discouraged.

If IDASetErrFile is to be called, it should be called before any other optional input functions, in order to take effect for any later error message.



# IDASetErrHandlerFn

Call flag = IDASetErrHandlerFn(ida\_mem, ehfun, eh\_data);

 $\label{prop:local_def} \textbf{Description} \quad \text{The function $\tt IDASetErrHandlerFn} \ \ \text{specifies the optional user-defined function to be}$ 

used in handling error messages.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

ehfun (IDAErrHandlerFn) is the user's C error handler function (see §4.6.2).

eh\_data (void \*) pointer to user data passed to ehfun every time it is called.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The function ehfun and data pointer eh\_data have been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes Error messages indicating that the IDAS solver memory is NULL will always be directed

to stderr.

#### IDASetUserData

Call flag = IDASetUserData(ida\_mem, user\_data);

Description The function IDASetUserData specifies the user data block user\_data and attaches it

to the main IDAS memory block.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

user\_data (void \*) pointer to the user data.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes If specified, the pointer to user\_data is passed to all user-supplied functions that have

it as an argument. Otherwise, a NULL pointer is passed.

If user\_data is needed in user linear solver or preconditioner functions, the call to IDASetUserData must be made *before* the call to specify the linear solver.



#### IDASetMaxOrd

Call flag = IDASetMaxOrd(ida\_mem, maxord);

Description The function IDASetMaxOrd specifies the maximum order of the linear multistep method.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxord (int) value of the maximum method order. This must be positive.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The input value maxord is  $\leq 0$ , or larger than its previous value.

Notes The default value is 5. If the input value exceeds 5, the value 5 will be used. Since maxord affects the memory requirements for the internal IDAS memory block, its value

cannot be increased past its previous value.

#### IDASetMaxNumSteps

Call flag = IDASetMaxNumSteps(ida\_mem, mxsteps);

Description The function IDASetMaxNumSteps specifies the maximum number of steps to be taken

by the solver in its attempt to reach the next output time.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

mxsteps (long int) maximum allowed number of steps.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes Passing mxsteps = 0 results in IDAS using the default value (500).

Passing mxsteps < 0 disables the test (not recommended).

# IDASetInitStep

Call flag = IDASetInitStep(ida\_mem, hin);

Description The function IDASetInitStep specifies the initial step size.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

hin (realtype) value of the initial step size to be attempted. Pass 0.0 to have

IDAS use the default value.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes By default, IDAS estimates the initial step as the solution of  $\|h\dot{y}\|_{WRMS} = 1/2$ , with an

added restriction that  $|h| \leq .001$  | tout - t0|.

#### IDASetMaxStep

Call flag = IDASetMaxStep(ida\_mem, hmax);

Description The function IDASetMaxStep specifies the maximum absolute value of the step size.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

hmax (realtype) maximum absolute value of the step size.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT Either hmax is not positive or it is smaller than the minimum allowable step.

Notes Pass hmax = 0 to obtain the default value  $\infty$ .

# IDASetStopTime

Call flag = IDASetStopTime(ida\_mem, tstop);

Description The function  ${\tt IDASetStopTime}$  specifies the value of the independent variable t past

which the solution is not to proceed.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

tstop (realtype) value of the independent variable past which the solution should not proceed.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The value of tstop is not beyond the current t value,  $t_n$ .

Notes The default, if this routine is not called, is that no stop time is imposed.

#### IDASetMaxErrTestFails

Call flag = IDASetMaxErrTestFails(ida\_mem, maxnef);

Description The function IDASetMaxErrTestFails specifies the maximum number of error test

failures in attempting one step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxnef (int) maximum number of error test failures allowed on one step (>0).

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The default value is 7.

#### IDASetMaxNonlinIters

Call flag = IDASetMaxNonlinIters(ida\_mem, maxcor);

Description The function IDASetMaxNonlinIters specifies the maximum number of nonlinear solver

iterations at one step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxcor (int) maximum number of nonlinear solver iterations allowed on one step

(>0).

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_MEM\_FAIL The SUNNONLINSOL module is NULL.

Notes The default value is 3.

# IDASetMaxConvFails

Call flag = IDASetMaxConvFails(ida\_mem, maxncf);

Description The function IDASetMaxConvFails specifies the maximum number of nonlinear solver

convergence failures at one step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxncf (int) maximum number of allowable nonlinear solver convergence failures on

one step (>0).

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The default value is 10.

#### IDASetNonlinConvCoef

Call flag = IDASetNonlinConvCoef(ida\_mem, nlscoef);

Description The function IDASetNonlinConvCoef specifies the safety factor in the nonlinear con-

vergence test; see Chapter 2, Eq. (2.8).

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nlscoef (realtype) coefficient in nonlinear convergence test (> 0.0).

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The value of nlscoef is  $\leq 0.0$ .

Notes The default value is 0.33.

# IDASetSuppressAlg

Call flag = IDASetSuppressAlg(ida\_mem, suppressalg);

Description The function IDASetSuppressAlg indicates whether or not to suppress algebraic vari-

ables in the local error test.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

suppressalg (booleantype) indicates whether to suppress (SUNTRUE) or not (SUNFALSE)

the algebraic variables in the local error test.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The default value is SUNFALSE.

If suppressalg=SUNTRUE is selected, then the id vector must be set (through IDASetId)

to specify the algebraic components.

In general, the use of this option (with suppressalg = SUNTRUE) is discouraged when solving DAE systems of index 1, whereas it is generally encouraged for systems of index 2 or more. See pp. 146-147 of Ref. [4] for more on this issue.

IDASetId

Call flag = IDASetId(ida\_mem, id);

Description The function IDASetId specifies algebraic/differential components in the y vector.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

> (N\_Vector) state vector. A value of 1.0 indicates a differential variable, while 0.0 indicates an algebraic variable.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The vector id is required if the algebraic variables are to be suppressed from the lo-

cal error test (see IDASetSuppressAlg) or if IDACalcIC is to be called with icopt =

 $IDA_YA_YDP_INIT$  (see §4.5.5).

#### IDASetConstraints

Call flag = IDASetConstraints(ida\_mem, constraints);

Description The function IDASetConstraints specifies a vector defining inequality constraints for

each component of the solution vector y.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

constraints (N\_Vector) vector of constraint flags. If constraints[i] is

0.0 then no constraint is imposed on  $y_i$ .

1.0 then  $y_i$  will be constrained to be  $y_i \ge 0.0$ .

-1.0 then  $y_i$  will be constrained to be  $y_i \leq 0.0$ .

2.0 then  $y_i$  will be constrained to be  $y_i > 0.0$ .

-2.0 then  $y_i$  will be constrained to be  $y_i < 0.0$ .

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The constraints vector contains illegal values or the simultaneous corrector option has been selected when doing forward sensitivity analysis.

Notes The presence of a non-NULL constraints vector that is not 0.0 in all components will cause constraint checking to be performed. However, a call with 0.0 in all components

of constraints will result in an illegal input return.

Constraint checking when doing forward sensitivity analysis with the simultaneous cor-

rector option is currently disallowed and will result in an illegal input return.

#### Linear solver interface optional input functions

The mathematical explanation of the linear solver methods available to IDAS is provided in §2.1. We group the user-callable routines into four categories: general routines concerning the overall IDALS linear solver interface, optional inputs for matrix-based linear solvers, optional inputs for matrix-free linear solvers, and optional inputs for iterative linear solvers. We note that the matrix-based and matrix-free groups are mutually exclusive, whereas the "iterative" tag can apply to either case.

When using matrix-based linear solver modules, the IDALS solver interface needs a function to compute an approximation to the Jacobian matrix  $J(t, y, \dot{y})$ . This function must be of type IDALsJacFn. The user can supply a Jacobian function, or if using a dense or banded matrix J can use the default internal difference quotient approximation that comes with the IDALS interface. To specify a user-supplied Jacobian function jac, IDALS provides the function IDASetJacFn. The IDALS interface

passes the pointer user\_data to the Jacobian function. This allows the user to create an arbitrary structure with relevant problem data and access it during the execution of the user-supplied Jacobian function, without using global data in the program. The pointer user\_data may be specified through IDASetUserData.

# IDASetJacFn

Call flag = IDASetJacFn(ida\_mem, jac);

Description The function IDASetJacFn specifies the Jacobian approximation function to be used for

a matrix-based solver within the IDALS interface.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

jac (IDALsJacFn) user-defined Jacobian approximation function.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver interface has not been initialized.

Notes This function must be called *after* the IDALS linear solver interface has been initialized through a call to IDASetLinearSolver.

By default, IDALS uses an internal difference quotient function for dense and band matrices. If NULL is passed to jac, this default function is used. An error will occur if no jac is supplied when using other matrix types.

The function type IDALsJacFn is described in §4.6.5.

The previous routine IDAD1sSetJacFn is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

When using matrix-free linear solver modules, the IDALS solver interface requires a function to compute an approximation to the product between the Jacobian matrix J(t,y) and a vector v. The user can supply a Jacobian-times-vector approximation function, or use the default internal difference quotient function that comes with the IDALS solver interface. A user-defined Jacobian-vector function must be of type IDALsJacTimesVecFn and can be specified through a call to IDASetJacTimes (see §4.6.6 for specification details). The evaluation and processing of any Jacobian-related data needed by the user's Jacobian-times-vector function may be done in the optional user-supplied function jtsetup (see §4.6.7 for specification details). The pointer user\_data received through IDASetUserData (or a pointer to NULL if user\_data was not specified) is passed to the Jacobian-times-vector setup and product functions, jtsetup and jtimes, each time they are called. This allows the user to create an arbitrary structure with relevant problem data and access it during the execution of the user-supplied preconditioner functions without using global data in the program.

#### IDASetJacTimes

Call flag = IDASetJacTimes(ida\_mem, jsetup, jtimes);

Description The function IDASetJacTimes specifies the Jacobian-vector setup and product func-

tions.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

jtsetup (IDALsJacTimesSetupFn) user-defined function to set up the Jacobian-vector product. Pass NULL if no setup is necessary.

jtimes (IDALsJacTimesVecFn) user-defined Jacobian-vector product function.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_SUNLS\_FAIL An error occurred when setting up the system matrix-times-vector routines in the SUNLINSOL object used by the IDALS interface.

Notes

The default is to use an internal finite difference quotient for jtimes and to omit jtsetup. If NULL is passed to jtimes, these defaults are used. A user may specify non-NULL jtimes and NULL jtsetup inputs.

This function must be called *after* the IDALS linear solver interface has been initialized through a call to IDASetLinearSolver.

The function type IDALsJacTimesSetupFn is described in §4.6.7.

The function type IDALsJacTimesVecFn is described in §4.6.6.

The previous routine IDASpilsSetJacTimes is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

Alternately, when using the default difference-quotient approximation to the Jacobian-vector product, the user may specify the factor to use in setting increments for the finite-difference approximation, via a call to IDASetIncrementFactor:

#### IDASetIncrementFactor

Call flag = IDASetIncrementFactor(ida\_mem, dqincfac);

Description

The function IDASetIncrementFactor specifies the increment factor to be used in the difference-quotient approximation to the product Jv. Specifically, Jv is approximated via the formula

$$Jv = \frac{1}{\sigma} \left[ F(t, \tilde{y}, \tilde{y}') - F(t, y, y') \right],$$

where  $\tilde{y} = y + \sigma v$ ,  $\tilde{y}' = y' + c_j \sigma v$ ,  $c_j$  is a BDF parameter proportional to the step size,  $\sigma = \sqrt{N}$  dqincfac, and N is the number of equations in the DAE system.

Arguments

ida\_mem (void \*) pointer to the IDAS memory block.

dqincfac (realtype) user-specified increment factor (positive).

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_ILL\_INPUT The specified value of dqincfac is  $\leq 0$ .

Notes

The default value is 1.0.

This function must be called *after* the IDALS linear solver interface has been initialized through a call to IDASetLinearSolver.

The previous routine IDASpilsSetIncrementFactor is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

When using an iterative linear solver, the user may supply a preconditioning operator to aid in solution of the system. This operator consists of two user-supplied functions, psetup and psolve, that are supplied to IDA using the function IDASetPreconditioner. The psetup function supplied to this routine should handle evaluation and preprocessing of any Jacobian data needed by the user's preconditioner solve function, psolve. Both of these functions are fully specified in §4.6. The user data pointer received through IDASetUserData (or a pointer to NULL if user data was not specified) is passed to the psetup and psolve functions. This allows the user to create an arbitrary structure with relevant problem data and access it during the execution of the user-supplied preconditioner functions without using global data in the program.

Also, as described in  $\S 2.1$ , the IDALS interface requires that iterative linear solvers stop when the norm of the preconditioned residual satisfies

$$||r|| \le \frac{\epsilon_L \epsilon}{10}$$

where  $\epsilon$  is the nonlinear solver tolerance, and the default  $\epsilon_L = 0.05$ ; this value may be modified by the user through the IDASetEpsLin function.

#### IDASetPreconditioner

Call flag = IDASetPreconditioner(ida\_mem, psetup, psolve);

Description The function IDASetPreconditioner specifies the preconditioner setup and solve func-

tions.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

psetup (IDALsPrecSetupFn) user-defined function to set up the preconditioner. Pass

NULL if no setup is necessary.

psolve (IDALsPrecSolveFn) user-defined preconditioner solve function.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional values have been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

 ${\tt IDALS\_SUNLS\_FAIL} \ \ {\tt An\ error\ occurred\ when\ setting\ up\ preconditioning\ in\ the\ SUNLINSOL}$ 

object used by the IDALS interface.

Notes The default is NULL for both arguments (i.e., no preconditioning).

This function must be called after the IDALS linear solver interface has been initialized

through a call to IDASetLinearSolver.

The function type IDALsPrecSolveFn is described in §4.6.8.

The function type IDALsPrecSetupFn is described in §4.6.9.

The previous routine IDASpilsSetPreconditioner is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

# IDASetEpsLin

Call flag = IDASetEpsLin(ida\_mem, eplifac);

Description The function IDASetEpsLin specifies the factor by which the Krylov linear solver's

convergence test constant is reduced from the nonlinear iteration test constant.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

eplifac (realtype) linear convergence safety factor ( $\geq 0.0$ ).

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_ILL\_INPUT The factor eplifac is negative.

Notes The default value is 0.05.

This function must be called *after* the IDALS linear solver interface has been initialized through a call to IDASetLinearSolver.

If eplifac= 0.0 is passed, the default value is used.

The previous routine IDASpilsSetEpsLin is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### 4.5.8.3 Initial condition calculation optional input functions

The following functions can be called just prior to calling IDACalcIC to set optional inputs controlling the initial condition calculation.

#### IDASetNonlinConvCoefIC

Call flag = IDASetNonlinConvCoefIC(ida\_mem, epiccon);

Description The function IDASetNonlinConvCoefIC specifies the positive constant in the Newton

iteration convergence test within the initial condition calculation.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

epiccon (realtype) coefficient in the Newton convergence test (>0).

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The epicon factor is <= 0.0.

Notes The default value is  $0.01 \cdot 0.33$ .

This test uses a weighted RMS norm (with weights defined by the tolerances). For new initial value vectors y and  $\dot{y}$  to be accepted, the norm of  $J^{-1}F(t_0, y, \dot{y})$  must be  $\leq$  epiccon, where J is the system Jacobian.

#### IDASetMaxNumStepsIC

Call flag = IDASetMaxNumStepsIC(ida\_mem, maxnh);

Description The function IDASetMaxNumStepsIC specifies the maximum number of steps allowed

when  $icopt=IDA\_YA\_YDP\_INIT$  in IDACalcIC, where h appears in the system Jacobian,

 $J = \partial F/\partial y + (1/h)\partial F/\partial \dot{y}.$ 

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxnh (int) maximum allowed number of values for h.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT maxnh is non-positive.

Notes The default value is 5.

#### IDASetMaxNumJacsIC

Call flag = IDASetMaxNumJacsIC(ida\_mem, maxnj);

Description The function IDASetMaxNumJacsIC specifies the maximum number of the approximate

Jacobian or preconditioner evaluations allowed when the Newton iteration appears to

be slowly converging.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxnj (int) maximum allowed number of Jacobian or preconditioner evaluations.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT maxnj is non-positive.

Notes The default value is 4.

#### IDASetMaxNumItersIC

Call flag = IDASetMaxNumItersIC(ida\_mem, maxnit);

Description The function IDASetMaxNumItersIC specifies the maximum number of Newton itera-

tions allowed in any one attempt to solve the initial conditions calculation problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxnit (int) maximum number of Newton iterations.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT maxnit is non-positive.

Notes The default value is 10.

# IDASetMaxBacksIC

Call flag = IDASetMaxBacksIC(ida\_mem, maxbacks);

Description The function IDASetMaxBacksIC specifies the maximum number of linesearch back-

tracks allowed in any Newton iteration, when solving the initial conditions calculation

problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxbacks (int) maximum number of linesearch backtracks per Newton step.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT maxbacks is non-positive.

Notes The default value is 100.

If IDASetMaxBacksIC is called in a Forward Sensitivity Analysis, the the limit maxbacks applies in the calculation of both the initial state values and the initial sensitivities.

#### | IDASetLineSearchOffIC

Call flag = IDASetLineSearchOffIC(ida\_mem, lsoff);

Description The function IDASetLineSearchOffIC specifies whether to turn on or off the linesearch

algorithm.

Arguments idamem (void \*) pointer to the IDAS memory block.

lsoff (booleantype) a flag to turn off (SUNTRUE) or keep (SUNFALSE) the linesearch

algorithm.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The default value is SUNFALSE.

#### IDASetStepToleranceIC

Call flag = IDASetStepToleranceIC(ida\_mem, steptol);

Description The function IDASetStepToleranceIC specifies a positive lower bound on the Newton

step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

steptol (int) Minimum allowed WRMS-norm of the Newton step (> 0.0).

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL. IDA\_ILL\_INPUT The steptol tolerance is  $\leq 0.0$ .

Notes The default value is (unit roundoff) $^{2/3}$ .

# 4.5.8.4 Rootfinding optional input functions

The following functions can be called to set optional inputs to control the rootfinding algorithm.

#### IDASetRootDirection

Call flag = IDASetRootDirection(ida\_mem, rootdir);

Description The function IDASetRootDirection specifies the direction of zero-crossings to be lo-

cated and returned to the user.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

rootdir (int \*) state array of length nrtfn, the number of root functions  $g_i$ , as specified in the call to the function IDARootInit. A value of 0 for rootdir[i] indicates that crossing in either direction should be reported for  $g_i$ . A value

of +1 or -1 indicates that the solver should report only zero-crossings where

 $g_i$  is increasing or decreasing, respectively.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT rootfinding has not been activated through a call to IDARootInit.

Notes The default behavior is to locate both zero-crossing directions.

# IDASetNoInactiveRootWarn

Call flag = IDASetNoInactiveRootWarn(ida\_mem);

Description The function IDASetNoInactiveRootWarn disables issuing a warning if some root func-

tion appears to be identically zero at the beginning of the integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes IDAS will not report the initial conditions as a possible zero-crossing (assuming that one or more components  $g_i$  are zero at the initial time). However, if it appears that some  $g_i$ 

is identically zero at the initial time (i.e.,  $g_i$  is zero at the initial time and after the first step), IDAS will issue a warning which can be disabled with this optional input function.

# 4.5.9 Interpolated output function

An optional function IDAGetDky is available to obtain additional output values. This function must be called after a successful return from IDASolve and provides interpolated values of y or its derivatives of order up to the last internal order used for any value of t in the last internal step taken by IDAS.

The call to the IDAGetDky function has the following form:

```
IDAGetDky
```

Call flag = IDAGetDky(ida\_mem, t, k, dky);

Description

The function IDAGetDky computes the interpolated values of the  $k^{th}$  derivative of y for any value of t in the last internal step taken by IDAS. The value of k must be nonnegative and smaller than the last internal order used. A value of 0 for k means that the y is interpolated. The value of t must satisfy  $t_n - h_u \le t \le t_n$ , where  $t_n$  denotes the current internal time reached, and  $h_u$  is the last internal step size used successfully.

Arguments

ida\_mem (void \*) pointer to the IDAS memory block.

t (realtype) time at which to interpolate.

 ${\tt k}$  (int) integer specifying the order of the derivative of y wanted.

dky (N\_Vector) vector containing the interpolated  $k^{th}$  derivative of y(t).

Return value The return value flag (of type int) is one of

IDA\_SUCCESS IDAGetDky succeeded.

 ${\tt IDA\_MEM\_NULL} \ \ {\tt The\ ida\_mem\ argument\ was\ NULL}.$ 

IDA\_BAD\_T t is not in the interval  $[t_n - h_u, t_n]$ .

IDA\_BAD\_K k is not one of  $\{0, 1, \dots, klast\}$ .

IDA\_BAD\_DKY dky is NULL.

Notes

It is only legal to call the function <code>IDAGetDky</code> after a successful return from <code>IDASolve</code>. Functions <code>IDAGetCurrentTime</code>, <code>IDAGetLastStep</code> and <code>IDAGetLastOrder</code> (see  $\S 4.5.10.2$ ) can be used to access  $t_n$ ,  $h_u$  and klast.

#### 4.5.10 Optional output functions

IDAS provides an extensive list of functions that can be used to obtain solver performance information. Table 4.3 lists all optional output functions in IDAS, which are then described in detail in the remainder of this section.

Some of the optional outputs, especially the various counters, can be very useful in determining how successful the IDAS solver is in doing its job. For example, the counters nsteps and nrevals provide a rough measure of the overall cost of a given run, and can be compared among runs with differing input options to suggest which set of options is most efficient. The ratio nniters/nsteps measures the performance of the nonlinear solver in solving the nonlinear systems at each time step; typical values for this range from 1.1 to 1.8. The ratio njevals/nniters (in the case of a matrix-based linear solver), and the ratio npevals/nniters (in the case of an iterative linear solver) measure the overall degree of nonlinearity in these systems, and also the quality of the approximate Jacobian or preconditioner being used. Thus, for example, njevals/nniters can indicate if a user-supplied Jacobian is inaccurate, if this ratio is larger than for the case of the corresponding internal Jacobian. The ratio nliters/nniters measures the performance of the Krylov iterative linear solver, and thus (indirectly) the quality of the preconditioner.

#### 4.5.10.1 SUNDIALS version information

The following functions provide a way to get SUNDIALS version information at runtime.

Table 4.3: Optional outputs from IDAS and IDALS

Optional output	Function name		
IDAS main solver			
Size of IDAS real and integer workspace	IDAGetWorkSpace		
Cumulative number of internal steps	IDAGetNumSteps		
No. of calls to residual function	IDAGetNumResEvals		
No. of calls to linear solver setup function	${\tt IDAGetNumLinSolvSetups}$		
No. of local error test failures that have occurred	${\tt IDAGetNumErrTestFails}$		
Order used during the last step	IDAGetLastOrder		
Order to be attempted on the next step	IDAGetCurrentOrder		
Order reductions due to stability limit detection	IDAGetNumStabLimOrderReds		
Actual initial step size used	IDAGetActualInitStep		
Step size used for the last step	IDAGetLastStep		
Step size to be attempted on the next step	IDAGetCurrentStep		
Current internal time reached by the solver	IDAGetCurrentTime		
Suggested factor for tolerance scaling	IDAGetTolScaleFactor		
Error weight vector for state variables	IDAGetErrWeights		
Estimated local errors	IDAGetEstLocalErrors		
No. of nonlinear solver iterations	IDAGetNumNonlinSolvIters		
No. of nonlinear convergence failures	IDAGetNumNonlinSolvConvFails		
Array showing roots found	IDAGetRootInfo		
No. of calls to user root function	IDAGetNumGEvals		
Name of constant associated with a return flag	IDAGetReturnFlagName		
IDAS initial conditions calc			
Number of backtrack operations	IDAGetNumBacktrackops		
Corrected initial conditions	IDAGetConsistentIC		
IDALS linear solver inter			
Size of real and integer workspace	${\tt IDAGetLinWorkSpace}$		
No. of Jacobian evaluations	IDAGetNumJacEvals		
No. of residual calls for finite diff. Jacobian[-vector] evals.	IDAGetNumLinResEvals		
No. of linear iterations	IDAGetNumLinIters		
No. of linear convergence failures	IDAGetNumLinConvFails		
No. of preconditioner evaluations	IDAGetNumPrecEvals		
No. of preconditioner solves	IDAGetNumPrecSolves		
No. of Jacobian-vector setup evaluations	${\tt IDAGetNumJTSetupEvals}$		
No. of Jacobian-vector product evaluations	IDAGetNumJtimesEvals		
Last return from a linear solver function	${\tt IDAGetLastLinFlag}$		
Name of constant associated with a return flag	IDAGetLinReturnFlagName		

#### SUNDIALSGetVersion

Call flag = SUNDIALSGetVersion(version, len);

Description The function SUNDIALSGetVersion fills a character array with SUNDIALS version infor-

mation.

Arguments version (char \*) character array to hold the SUNDIALS version information.

(int) allocated length of the version character array.

Return value If successful, SUNDIALSGetVersion returns 0 and version contains the SUNDIALS ver-

sion information. Otherwise, it returns -1 and version is not set (the input character

array is too short).

Notes A string of 25 characters should be sufficient to hold the version information. Any

trailing characters in the version array are removed.

#### SUNDIALSGetVersionNumber

Call flag = SUNDIALSGetVersionNumber(&major, &minor, &patch, label, len);

The function SUNDIALSGetVersionNumber set integers for the SUNDIALS major, minor, Description

and patch release numbers and fills a character array with the release label if applicable.

(int) SUNDIALS release major version number. Arguments

> (int) SUNDIALS release minor version number. patch (int) SUNDIALS release patch version number.

label (char \*) character array to hold the SUNDIALS release label.

(int) allocated length of the label character array. len

Return value If successful, SUNDIALSGetVersionNumber returns 0 and the major, minor, patch, and

label values are set. Otherwise, it returns -1 and the values are not set (the input

character array is too short).

Notes A string of 10 characters should be sufficient to hold the label information. If a label

is not used in the release version, no information is copied to label. Any trailing

characters in the label array are removed.

#### Main solver optional output functions 4.5.10.2

IDAS provides several user-callable functions that can be used to obtain different quantities that may be of interest to the user, such as solver workspace requirements, solver performance statistics, as well as additional data from the IDAS memory block (a suggested tolerance scaling factor, the error weight vector, and the vector of estimated local errors). Also provided are functions to extract statistics related to the performance of the SUNNONLINSOL nonlinear solver being used. As a convenience, additional extraction functions provide the optional outputs in groups. These optional output functions are described next.

# IDAGetWorkSpace

Call flag = IDAGetWorkSpace(ida\_mem, &lenrw, &leniw);

Description The function IDAGetWorkSpace returns the IDAS real and integer workspace sizes.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

> lenrw (long int) number of real values in the IDAS workspace.

(long int) number of integer values in the IDAS workspace. leniw

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes

In terms of the problem size N, the maximum method order maxord, and the number nrtfn of root functions (see §4.5.6), the actual size of the real workspace, in realtype words, is given by the following:

- base value: lenrw =  $55 + (m+6) * N_r + 3*nrtfn$ ;
- with IDASVtolerances: lenrw = lenrw  $+N_r$ ;
- with constraint checking (see IDASetConstraints): lenrw = lenrw  $+N_r$ ;
- with id specified (see IDASetId): lenrw = lenrw  $+N_r$ ;

where  $m = \max(\max, 3)$ , and  $N_r$  is the number of real words in one N\_Vector ( $\approx N$ ).

The size of the integer workspace (without distinction between int and long int words) is given by:

- base value: leniw =  $38 + (m+6) * N_i + \text{nrtfn}$ ;
- with IDASVtolerances: leniw = leniw  $+N_i$ ;
- with constraint checking:  $lenrw = lenrw + N_i$ ;
- with id specified: lenrw = lenrw  $+N_i$ ;

where  $N_i$  is the number of integer words in one N\_Vector (= 1 for NVECTOR\_SERIAL and 2\*npes for NVECTOR\_PARALLEL on npes processors).

For the default value of maxord, with no rootfinding, no id, no constraints, and with no call to IDASVtolerances, these lengths are given roughly by: lenrw = 55 + 11N, leniw = 49.

Note that additional memory is allocated if quadratures and/or forward sensitivity integration is enabled. See  $\S4.7.1$  and  $\S5.2.1$  for more details.

## IDAGetNumSteps

Call flag = IDAGetNumSteps(ida\_mem, &nsteps);

Description The function IDAGetNumSteps returns the cumulative number of internal steps taken by the solver (total so far).

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nsteps (long int) number of steps taken by IDAS.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### IDAGetNumResEvals

Call flag = IDAGetNumResEvals(ida\_mem, &nrevals);

Description The function IDAGetNumResEvals returns the number of calls to the user's residual evaluation function.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nrevals (long int) number of calls to the user's res function.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The nrevals value returned by IDAGetNumResEvals does not account for calls made to res from a linear solver or preconditioner module.

## IDAGetNumLinSolvSetups

Call flag = IDAGetNumLinSolvSetups(ida\_mem, &nlinsetups);

Description The function IDAGetNumLinSolvSetups returns the cumulative number of calls made

to the linear solver's setup function (total so far).

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nlinsetups (long int) number of calls made to the linear solver setup function.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### IDAGetNumErrTestFails

Call flag = IDAGetNumErrTestFails(ida\_mem, &netfails);

Description The function IDAGetNumErrTestFails returns the cumulative number of local error

test failures that have occurred (total so far).

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

netfails (long int) number of error test failures.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

## ${\tt IDAGetLastOrder}$

Call flag = IDAGetLastOrder(ida\_mem, &klast);

Description The function IDAGetLastOrder returns the integration method order used during the

last internal step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

klast (int) method order used on the last internal step.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

## IDAGetCurrentOrder

Call flag = IDAGetCurrentOrder(ida\_mem, &kcur);

Description The function IDAGetCurrentOrder returns the integration method order to be used on

the next internal step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

kcur (int) method order to be used on the next internal step.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

## IDAGetLastStep

Call flag = IDAGetLastStep(ida\_mem, &hlast);

Description The function IDAGetLastStep returns the integration step size taken on the last internal

step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

hlast (realtype) step size taken on the last internal step by IDAS, or last artificial

step size used in IDACalcIC, whichever was called last.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### IDAGetCurrentStep

Call flag = IDAGetCurrentStep(ida\_mem, &hcur);

Description The function IDAGetCurrentStep returns the integration step size to be attempted on

the next internal step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

hcur (realtype) step size to be attempted on the next internal step.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### IDAGetActualInitStep

Call flag = IDAGetActualInitStep(ida\_mem, &hinused);

Description The function IDAGetActualInitStep returns the value of the integration step size used

on the first step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

hinused (realtype) actual value of initial step size.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes Even if the value of the initial integration step size was specified by the user through a call to IDASetInitStep, this value might have been changed by IDAS to ensure that

a call to IDASetInitStep, this value might have been changed by IDAS to ensure that the step size is within the prescribed bounds  $(h_{\min} \leq h_0 \leq h_{\max})$ , or to meet the local

error test.

## IDAGetCurrentTime

Call flag = IDAGetCurrentTime(ida\_mem, &tcur);

Description The function IDAGetCurrentTime returns the current internal time reached by the

solver.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

tcur (realtype) current internal time reached.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### IDAGetTolScaleFactor

Call flag = IDAGetTolScaleFactor(ida\_mem, &tolsfac);

Description The function IDAGetTolScaleFactor returns a suggested factor by which the user's

tolerances should be scaled when too much accuracy has been requested for some internal

step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

tolsfac (realtype) suggested scaling factor for user tolerances.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### IDAGetErrWeights

Call flag = IDAGetErrWeights(ida\_mem, eweight);

Description The function IDAGetErrWeights returns the solution error weights at the current time.

These are the  $W_i$  given by Eq. (2.7) (or by the user's IDAEwtFn).

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

eweight (N\_Vector) solution error weights at the current time.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The user must allocate space for eweight.

## IDAGetEstLocalErrors

Call flag = IDAGetEstLocalErrors(ida\_mem, ele);

Description The function IDAGetEstLocalErrors returns the estimated local errors.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

ele (N\_Vector) estimated local errors at the current time.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes The user must allocate space for ele.

The values returned in ele are only valid if IDASolve returned a non-negative value.

The ele vector, together with the eweight vector from IDAGetErrWeights, can be used to determine how the various components of the system contributed to the estimated local error test. Specifically, that error test uses the RMS norm of a vector whose components are the products of the components of these two vectors. Thus, for example, if there were recent error test failures, the components causing the failures are those with largest values for the products, denoted loosely as eweight[i]\*ele[i].

#### IDAGetIntegratorStats

Call flag = IDAGetIntegratorStats(ida\_mem, &nsteps, &nrevals, &nlinsetups, &netfails, &klast, &kcur, &hinused, &hlast, &hcur, &tcur);

Description The function IDAGetIntegratorStats returns the IDAS integrator statistics as a group.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.



nsteps (long int) cumulative number of steps taken by IDAS.

nrevals (long int) cumulative number of calls to the user's res function.

nlinsetups (long int) cumulative number of calls made to the linear solver setup

function.

netfails (long int) cumulative number of error test failures.

klast (int) method order used on the last internal step.

kcur (int) method order to be used on the next internal step.

hinused (realtype) actual value of initial step size.

hlast (realtype) step size taken on the last internal step.

hcur (realtype) step size to be attempted on the next internal step.

tcur (realtype) current internal time reached.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS the optional output values have been successfully set.

IDA\_MEM\_NULL the ida\_mem pointer is NULL.

#### IDAGetNumNonlinSolvIters

Call flag = IDAGetNumNonlinSolvIters(ida\_mem, &nniters);

Description The function IDAGetNumNonlinSolvIters returns the cumulative number of nonlinear

iterations performed.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nniters (long int) number of nonlinear iterations performed.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_MEM\_FAIL The SUNNONLINSOL module is NULL.

## IDAGetNumNonlinSolvConvFails

Call flag = IDAGetNumNonlinSolvConvFails(ida\_mem, &nncfails);

Description The function IDAGetNumNonlinSolvConvFails returns the cumulative number of non-

linear convergence failures that have occurred.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nncfails (long int) number of nonlinear convergence failures.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

## IDAGetNonlinSolvStats

Call flag = IDAGetNonlinSolvStats(ida\_mem, &nniters, &nncfails);

Description The function IDAGetNonlinSolvStats returns the IDAS nonlinear solver statistics as a

group.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nniters (long int) cumulative number of nonlinear iterations performed.

nncfails (long int) cumulative number of nonlinear convergence failures.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_MEM\_FAIL The SUNNONLINSOL module is NULL.

## IDAGetReturnFlagName

Call name = IDAGetReturnFlagName(flag);

Description The function IDAGetReturnFlagName returns the name of the IDAS constant correspond-

ing to flag.

Arguments The only argument, of type int, is a return flag from an IDAS function.

Return value The return value is a string containing the name of the corresponding constant.

#### 4.5.10.3 Initial condition calculation optional output functions

#### IDAGetNumBcktrackOps

Call flag = IDAGetNumBacktrackOps(ida\_mem, &nbacktr);

Description The function IDAGetNumBacktrackOps returns the number of backtrack operations done

in the linesearch algorithm in IDACalcIC.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nbacktr (long int) the cumulative number of backtrack operations.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### IDAGetConsistentIC

Call flag = IDAGetConsistentIC(ida\_mem, yy0\_mod, yp0\_mod);

Description The function IDAGetConsistentIC returns the corrected initial conditions calculated

by IDACalcIC.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

yy0\_mod (N\_Vector) consistent solution vector. yp0\_mod (N\_Vector) consistent derivative vector.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_ILL\_INPUT The function was not called before the first call to IDASolve.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes If the consistent solution vector or consistent derivative vector is not desired, pass NULL

for the corresponding argument.

The user must allocate space for yyo mod and ypo mod (if not NULL).

#### 4.5.10.4 Rootfinding optional output functions

There are two optional output functions associated with rootfinding.



#### IDAGetRootInfo

Call flag = IDAGetRootInfo(ida\_mem, rootsfound);

Description The function IDAGetRootInfo returns an array showing which functions were found to

have a root.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

rootsfound (int \*) array of length nrtfn with the indices of the user functions  $g_i$  found to have a root. For i = 0, ..., nrtfn -1, rootsfound $[i] \neq 0$  if  $g_i$  has a

root, and = 0 if not.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output values have been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes Note that, for the components  $g_i$  for which a root was found, the sign of rootsfound[i]

indicates the direction of zero-crossing. A value of +1 indicates that  $g_i$  is increasing,

while a value of -1 indicates a decreasing  $g_i$ .

The user must allocate memory for the vector rootsfound.

# <u>!</u>

#### IDAGetNumGEvals

Call flag = IDAGetNumGEvals(ida\_mem, &ngevals);

Description The function IDAGetNumGEvals returns the cumulative number of calls to the user root

function g.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

ngevals (long int) number of calls to the user's function g so far.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

#### 4.5.10.5 IDALS linear solver interface optional output functions

The following optional outputs are available from the IDALS modules: workspace requirements, number of calls to the Jacobian routine, number of calls to the residual routine for finite-difference Jacobian or Jacobian-vector product approximation, number of linear iterations, number of linear convergence failures, number of calls to the preconditioner setup and solve routines, number of calls to the Jacobian-vector setup and product routines, and last return value from an IDALS function. Note that, where the name of an output would otherwise conflict with the name of an optional output from the main solver, a suffix LS (for Linear Solver) has been added (e.g., lenrwLS).

## IDAGetLinWorkSpace

Call flag = IDAGetLinWorkSpace(ida\_mem, &lenrwLS, &leniwLS);

Description The function IDAGetLinWorkSpace returns the sizes of the real and integer workspaces

used by the IDALS linear solver interface.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

lenrwLS (long int) the number of real values in the IDALS workspace.

leniwLS (long int) the number of integer values in the IDALS workspace.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes

The workspace requirements reported by this routine correspond only to memory allocated within this interface and to memory allocated by the SUNLINSOL object attached to it. The template Jacobian matrix allocated by the user outside of IDALS is not included in this report.

The previous routines IDADlsGetWorkspace and IDASpilsGetWorkspace are now wrappers for this routine, and may still be used for backward-compatibility. However, these will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDAGetNumJacEvals

Call flag = IDAGetNumJacEvals(ida\_mem, &njevals);

Description The function IDAGetNumJacEvals returns the cumulative number of calls to the IDALS

Jacobian approximation function.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

njevals (long int) the cumulative number of calls to the Jacobian function (total so

far).

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes

The previous routine IDADlsGetNumJacEvals is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

## IDAGetNumLinResEvals

Call flag = IDAGetNumLinResEvals(ida\_mem, &nrevalsLS);

Description The function IDAGetNumLinResEvals returns the cumulative number of calls to the user

residual function due to the finite difference Jacobian approximation or finite difference

Jacobian-vector product approximation.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nrevalsLS (long int) the cumulative number of calls to the user residual function.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes The value nrevalsLS is incremented only if one of the default internal difference quotient

functions is used.

The previous routines IDADlsGetNumRhsEvals and IDASpilsGetNumRhsEvals are now wrappers for this routine, and may still be used for backward-compatibility. However, these will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDAGetNumLinIters

Call flag = IDAGetNumLinIters(ida\_mem, &nliters);

Description The function IDAGetNumLinIters returns the cumulative number of linear iterations.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nliters (long int) the current number of linear iterations.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes

The previous routine IDASpilsGetNumLinIters is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

## IDAGetNumLinConvFails

Call flag = IDAGetNumLinConvFails(ida\_mem, &nlcfails);

Description The function IDAGetNumLinConvFails returns the cumulative number of linear conver-

gence failures.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nlcfails (long int) the current number of linear convergence failures.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes

The previous routine IDASpilsGetNumConvFails is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

## IDAGetNumPrecEvals

Call flag = IDAGetNumPrecEvals(ida\_mem, &npevals);

Description The function IDAGetNumPrecEvals returns the cumulative number of preconditioner

evaluations, i.e., the number of calls made to psetup.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

npevals (long int) the cumulative number of calls to psetup.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes The previous routine IDASpilsGetNumPrecEvals is now a wrapper for this routine, and

may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

## IDAGetNumPrecSolves

Call flag = IDAGetNumPrecSolves(ida\_mem, &npsolves);

Description The function IDAGetNumPrecSolves returns the cumulative number of calls made to

the preconditioner solve function, psolve.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

npsolves (long int) the cumulative number of calls to psolve.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes

The previous routine IDASpilsGetNumPrecSolves is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDAGetNumJTSetupEvals

Call flag = IDAGetNumJTSetupEvals(ida\_mem, &njtsetup);

Description The function IDAGetNumJTSetupEvals returns the cumulative number of calls made to

the Jacobian-vector setup function jtsetup.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

njtsetup (long int) the current number of calls to jtsetup.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_LMEM\_NULL The IDA linear solver has not been initialized.

Notes

The previous routine <code>IDASpilsGetNumJTSetupEvals</code> is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDAGetNumJtimesEvals

Call flag = IDAGetNumJtimesEvals(ida\_mem, &njvevals);

Description The function IDAGetNumJtimesEvals returns the cumulative number of calls made to

the Jacobian-vector function, jtimes.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

njvevals (long int) the cumulative number of calls to jtimes.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_LMEM\_NULL The IDA linear solver has not been initialized.

Notes

The previous routine <code>IDASpilsGetNumJtimesEvals</code> is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDAGetLastLinFlag

Call flag = IDAGetLastLinFlag(ida.mem, &lsflag);

Description The function IDAGetLastLinFlag returns the last return value from an IDALS routine.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

lsflag (long int) the value of the last return flag from an IDALS function.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

Notes

If the IDALS setup function failed (i.e., IDASolve returned IDALSETUP\_FAIL) when using the SUNLINSOL\_DENSE or SUNLINSOL\_BAND modules, then the value of lsflag is equal to the column index (numbered from one) at which a zero diagonal element was encountered during the LU factorization of the (dense or banded) Jacobian matrix.

If the IDALS setup function failed when using another SUNLINSOL module, then lsflag will be SUNLS\_PSET\_FAIL\_UNREC, SUNLS\_ASET\_FAIL\_UNREC, or SUNLS\_PACKAGE\_FAIL\_UNREC.

If the IDALS solve function failed (IDASolve returned IDALSOLVE\_FAIL), lsflag contains the error return flag from the SUNLINSOL object, which will be one of: SUNLS\_MEM\_NULL, indicating that the SUNLINSOL memory is NULL;

SUNLS\_ATIMES\_FAIL\_UNREC, indicating an unrecoverable failure in the J\*v function; SUNLS\_PSOLVE\_FAIL\_UNREC, indicating that the preconditioner solve function psolve failed unrecoverably; SUNLS\_GS\_FAIL, indicating a failure in the Gram-Schmidt procedure (generated only in SPGMR or SPFGMR); SUNLS\_QRSOL\_FAIL, indicating that the matrix R was found to be singular during the QR solve phase (SPGMR and SPFGMR only); or SUNLS\_PACKAGE\_FAIL\_UNREC, indicating an unrecoverable failure in an external iterative linear solver package.

The previous routines IDADlsGetLastFlag and IDASpilsGetLastFlag are now wrappers for this routine, and may still be used for backward-compatibility. However, these will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDAGetLinReturnFlagName

Call name = IDAGetLinReturnFlagName(lsflag);

Description The function IDAGetLinReturnFlagName returns the name of the IDA constant corre-

sponding to lsflag.

Arguments The only argument, of type long int, is a return flag from an IDALS function.

Return value The return value is a string containing the name of the corresponding constant.

If  $1 \leq lsflag \leq N$  (LU factorization failed), this function returns "NONE".

Notes

The previous routines IDADlsGetReturnFlagName and IDASpilsGetReturnFlagName are now wrappers for this routine, and may still be used for backward-compatibility. However, these will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

## 4.5.11 IDAS reinitialization function

The function IDAReInit reinitializes the main IDAS solver for the solution of a new problem, where a prior call to IDAInit has been made. The new problem must have the same size as the previous one. IDAReInit performs the same input checking and initializations that IDAInit does, but does no memory allocation, as it assumes that the existing internal memory is sufficient for the new problem. A call to IDAReInit deletes the solution history that was stored internally during the previous integration. Following a successful call to IDAReInit, call IDASolve again for the solution of the new problem.

The use of IDAReInit requires that the maximum method order, maxord, is no larger for the new problem than for the problem specified in the last call to IDAInit. In addition, the same NVECTOR module set for the previous problem will be reused for the new problem.

If there are changes to the linear solver specifications, make the appropriate calls to either the linear solver objects themselves, or to the IDALS interface routines, as described in §4.5.3.

If there are changes to any optional inputs, make the appropriate IDASet\*\*\* calls, as described in §4.5.8. Otherwise, all solver inputs set previously remain in effect.

One important use of the IDAReInit function is in the treating of jump discontinuities in the residual function. Except in cases of fairly small jumps, it is usually more efficient to stop at each point of discontinuity and restart the integrator with a readjusted DAE model, using a call to IDAReInit. To stop when the location of the discontinuity is known, simply make that location a value of tout. To stop when the location of the discontinuity is determined by the solution, use the rootfinding feature. In either case, it is critical that the residual function not incorporate the discontinuity, but rather have a smooth extention over the discontinuity, so that the step across it (and subsequent rootfinding, if used) can be done efficiently. Then use a switch within the residual function (communicated through user\_data) that can be flipped between the stopping of the integration and the restart, so that the restarted problem uses the new values (which have jumped). Similar comments apply if there is to be a jump in the dependent variable vector.

```
IDAReInit
Call
             flag = IDAReInit(ida_mem, t0, y0, yp0);
Description
             The function IDAReInit provides required problem specifications and reinitializes IDAS.
Arguments
             ida_mem (void *) pointer to the IDAS memory block.
                       (realtype) is the initial value of t.
             t0
                       (N_Vector) is the initial value of y.
             y0
                       (N_Vector) is the initial value of \dot{y}.
             ур0
Return value The return value flag (of type int) will be one of the following:
             TDA SUCCESS
                             The call to IDAReInit was successful.
             IDA_MEM_NULL
                             The IDAS memory block was not initialized through a previous call to
                              IDACreate.
             IDA_NO_MALLOC Memory space for the IDAS memory block was not allocated through a
                             previous call to IDAInit.
             IDA_ILL_INPUT An input argument to IDAReInit has an illegal value.
Notes
             If an error occurred, IDAReInit also sends an error message to the error handler func-
```

# 4.6 User-supplied functions

The user-supplied functions consist of one function defining the DAE residual, (optionally) a function that handles error and warning messages, (optionally) a function that provides the error weight vector, (optionally) one or two functions that provide Jacobian-related information for the linear solver, and (optionally) one or two functions that define the preconditioner for use in any of the Krylov iteration algorithms.

## 4.6.1 Residual function

tion.

The user must provide a function of type IDAResFn defined as follows:

```
Definition typedef int (*IDAResFn) (realtype tt, N_Vector yy, N_Vector yp, N_Vector rr, void *user_data);

Purpose This function computes the problem residual for given values of the independent variable t, state vector y, and derivative \dot{y}.

Arguments tt is the current value of the independent variable. yy is the current value of the dependent variable vector, y(t). yp is the current value of \dot{y}(t).
```

is the output residual vector  $F(t, y, \dot{y})$ . rr

user\_data is a pointer to user data, the same as the user\_data parameter passed to IDASetUserData.

Return value An IDAResFn function type should return a value of 0 if successful, a positive value if a recoverable error occurred (e.g., yy has an illegal value), or a negative value if a nonrecoverable error occurred. In the last case, the integrator halts. If a recoverable error occurred, the integrator will attempt to correct and retry.

Notes

A recoverable failure error return from the IDAResFn is typically used to flag a value of the dependent variable y that is "illegal" in some way (e.g., negative where only a non-negative value is physically meaningful). If such a return is made, IDAS will attempt to recover (possibly repeating the nonlinear solve, or reducing the step size) in order to avoid this recoverable error return.

For efficiency reasons, the DAE residual function is not evaluated at the converged solution of the nonlinear solver. Therefore, in general, a recoverable error in that converged value cannot be corrected. (It may be detected when the right-hand side function is called the first time during the following integration step, but a successful step cannot be undone.) However, if the user program also includes quadrature integration, the state variables can be checked for legality in the call to IDAQuadRhsFn, which is called at the converged solution of the nonlinear system, and therefore IDAS can be flagged to attempt to recover from such a situation. Also, if sensitivity analysis is performed with the staggered method, the DAE residual function is called at the converged solution of the nonlinear system, and a recoverable error at that point can be flagged, and IDAS will then try to correct it.

Allocation of memory for yp is handled within IDAS.

#### 4.6.2 Error message handler function

As an alternative to the default behavior of directing error and warning messages to the file pointed to by errfp (see IDASetErrFile), the user may provide a function of type IDAErrHandlerFn to process any such messages. The function type IDAErrHandlerFn is defined as follows:

### IDAErrHandlerFn

Definition typedef void (\*IDAErrHandlerFn)(int error\_code, const char \*module, const char \*function, char \*msg, void \*eh\_data);

Purpose This function processes error and warning messages from IDAS and its sub-modules.

Arguments error\_code is the error code.

> module is the name of the IDAS module reporting the error. function is the name of the function in which the error occurred.

is the error message. msg

 $eh_data$ is a pointer to user data, the same as the eh\_data parameter passed to

IDASetErrHandlerFn.

Return value A IDAErrHandlerFn function has no return value.

Notes error\_code is negative for errors and positive (IDA\_WARNING) for warnings. If a function that returns a pointer to memory encounters an error, it sets error\_code to 0.

#### 4.6.3 Error weight function

As an alternative to providing the relative and absolute tolerances, the user may provide a function of type IDAEwtFn to compute a vector ewt containing the multiplicative weights  $W_i$  used in the WRMS

norm  $||v||_{\text{WRMS}} = \sqrt{(1/N)\sum_{1}^{N}(W_i \cdot v_i)^2}$ . These weights will used in place of those defined by Eq. (2.7). The function type IDAEwtFn is defined as follows:

#### IDAEwtFn

Definition typedef int (\*IDAEwtFn)(N\_Vector y, N\_Vector ewt, void \*user\_data);

Purpose This function computes the WRMS error weights for the vector y.

Arguments y is the value of the dependent variable vector at which the weight vector is

to be computed.

ewt is the output vector containing the error weights.

user\_data is a pointer to user data, the same as the user\_data parameter passed to IDASetUserData.

Return value An IDAEwtFn function type must return 0 if it successfully set the error weights and -1

otherwise.

Notes Allocation of memory for ewt is handled within IDAS.

The error weight vector must have all components positive. It is the user's responsibility to perform this test and return -1 if it is not satisfied.

## 4.6.4 Rootfinding function

If a rootfinding problem is to be solved during the integration of the DAE system, the user must supply a C function of type IDARootFn, defined as follows:

#### IDARootFn

Definition typedef int (\*IDARootFn)(realtype t, N\_Vector y, N\_Vector yp, realtype \*gout, void \*user\_data);

Purpose This function computes a vector-valued function  $q(t, y, \dot{y})$  such that the roots of the

**nrtfn** components  $g_i(t, y, \dot{y})$  are to be found during the integration.

Arguments t is the current value of the independent variable.

y is the current value of the dependent variable vector, y(t).

yp is the current value of  $\dot{y}(t)$ , the t-derivative of y.

gout is the output array, of length nrtfn, with components  $g_i(t, y, \dot{y})$ .

user\_data is a pointer to user data, the same as the user\_data parameter passed to

IDASetUserData.

Return value An IDARootFn should return 0 if successful or a non-zero value if an error occurred (in

which case the integration is halted and IDASolve returns IDA\_RTFUNC\_FAIL).

Notes Allocation of memory for gout is handled within IDAS.

## 4.6.5 Jacobian construction (matrix-based linear solvers)

If a matrix-based linear solver module is used (i.e. a non-NULL SUNMATRIX object was supplied to IDASetLinearSolver), the user may provide a function of type IDALsJacFn defined as follows:

## IDALsJacFn



Purpose This function computes the Jacobian matrix J of the DAE system (or an approximation to it), defined by Eq. (2.6).

Arguments tt is the current value of the independent variable t.

cj is the scalar in the system Jacobian, proportional to the inverse of the step size ( $\alpha$  in Eq. (2.6)).

yy is the current value of the dependent variable vector, y(t).

yp is the current value of  $\dot{y}(t)$ .

rr is the current value of the residual vector  $F(t, y, \dot{y})$ .

Jac is the output (approximate) Jacobian matrix (of type SUNMatrix),  $J = \partial F/\partial y + cj \ \partial F/\partial \dot{y}$ .

user\_data is a pointer to user data, the same as the user\_data parameter passed to IDASetUserData.

tmp1

tmp3 are pointers to memory allocated for variables of type N\_Vector which can be used by IDALsJacFn function as temporary storage or work space.

Return value An IDALsJacFn should return 0 if successful, a positive value if a recoverable error occurred, or a negative value if a nonrecoverable error occurred.

In the case of a recoverable error return, the integrator will attempt to recover by reducing the stepsize, and hence changing  $\alpha$  in (2.6).

Notes

Information regarding the structure of the specific SUNMATRIX structure (e.g., number of rows, upper/lower bandwidth, sparsity type) may be obtained through using the implementation-specific SUNMATRIX interface functions (see Chapter 8 for details).

Prior to calling the user-supplied Jacobian function, the Jacobian matrix J(t, y) is zeroed out, so only nonzero elements need to be loaded into Jac.

If the user's IDALsJacFn function uses difference quotient approximations, it may need to access quantities not in the call list. These quantities may include the current stepsize, the error weights, etc. To obtain these, the user will need to add a pointer to ida\_mem to user\_data and then use the IDAGet\* functions described in §4.5.10.2. The unit roundoff can be accessed as UNIT\_ROUNDOFF defined in sundials\_types.h.

#### dansa

A user-supplied dense Jacobian function must load the Neq  $\times$  Neq dense matrix Jac with an approximation to the Jacobian matrix  $J(t,y,\dot{y})$  at the point (tt, yy, yp). The accessor macros SM\_ELEMENT\_D and SM\_COLUMN\_D allow the user to read and write dense matrix elements without making explicit references to the underlying representation of the SUNMATRIX\_DENSE type. SM\_ELEMENT\_D(J, i, j) references the (i, j)-th element of the dense matrix Jac (with i, j = 0...N - 1). This macro is meant for small problems for which efficiency of access is not a major concern. Thus, in terms of the indices m and n ranging from 1 to N, the Jacobian element  $J_{m,n}$  can be set using the statement SM\_ELEMENT\_D(J, m-1, n-1) =  $J_{m,n}$ . Alternatively, SM\_COLUMN\_D(J, j) returns a pointer to the first element of the j-th column of Jac (with j = 0...N - 1), and the elements of the j-th column can then be accessed using ordinary array indexing. Consequently,  $J_{m,n}$  can be loaded using the statements col\_n = SM\_COLUMN\_D(J, n-1); col\_n[m-1] =  $J_{m,n}$ . For large problems, it is more efficient to use SM\_COLUMN\_D than to use SM\_ELEMENT\_D. Note that both of these macros number rows and columns starting from 0. The SUNMATRIX\_DENSE type and accessor macros are documented in §8.2.

## banded:

A user-supplied banded Jacobian function must load the Neq  $\times$  Neq banded matrix Jac with an approximation to the Jacobian matrix  $J(t, y, \dot{y})$  at the point (tt, yy, yp). The accessor macros SM\_ELEMENT\_B, SM\_COLUMN\_B, and SM\_COLUMN\_ELEMENT\_B allow the

user to read and write banded matrix elements without making specific references to the underlying representation of the SUNMATRIX\_BAND type. SM\_ELEMENT\_B(J, i, j) references the (i, j)-th element of the banded matrix Jac, counting from 0. This macro is meant for use in small problems for which efficiency of access is not a major concern. Thus, in terms of the indices m and n ranging from 1 to N with (m,n)within the band defined by mupper and mlower, the Jacobian element  $J_{m,n}$  can be loaded using the statement SM\_ELEMENT\_B(J, m-1, n-1) =  $J_{m,n}$ . The elements within the band are those with  $-mupper \le m-n \le mlower$ . Alternatively, SM\_COLUMN\_B(J, j) returns a pointer to the diagonal element of the j-th column of Jac, and if we assign this address to realtype \*col\_j, then the i-th element of the j-th column is given by SM\_COLUMN\_ELEMENT\_B(col\_j, i, j), counting from 0. Thus, for (m, n)within the band,  $J_{m,n}$  can be loaded by setting col\_n = SM\_COLUMN\_B(J, n-1); and SM\_COLUMN\_ELEMENT\_B(col\_n, m-1, n-1) =  $J_{m,n}$ . The elements of the j-th column can also be accessed via ordinary array indexing, but this approach requires knowledge of the underlying storage for a band matrix of type SUNMATRIX\_BAND. The array col\_n can be indexed from -mupper to mlower. For large problems, it is more efficient to use SM\_COLUMN\_B and SM\_COLUMN\_ELEMENT\_B than to use the SM\_ELEMENT\_B macro. As in the dense case, these macros all number rows and columns starting from 0. The SUNMATRIX\_BAND type and accessor macros are documented in §8.3.

#### sparse:

A user-supplied sparse Jacobian function must load the Neq  $\times$  Neq compressed-sparse-column or compressed-sparse-row matrix Jac with an approximation to the Jacobian matrix  $J(t,y,\dot{y})$  at the point (tt, yy, yp). Storage for Jac already exists on entry to this function, although the user should ensure that sufficient space is allocated in Jac to hold the nonzero values to be set; if the existing space is insufficient the user may reallocate the data and index arrays as needed. The amount of allocated space in a SUNMATRIX\_SPARSE object may be accessed using the macro SM\_NNZ\_S or the routine SUNSparseMatrix\_NNZ. The SUNMATRIX\_SPARSE type and accessor macros are documented in §8.4.

The previous function type IDADlsJacFn is identical to IDALsJacFn, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

## 4.6.6 Jacobian-vector product (matrix-free linear solvers)

If a matrix-free linear solver is to be used (i.e., a NULL-valued SUNMATRIX was supplied to IDASetLinearSolver), the user may provide a function of type IDALsJacTimesVecFn in the following form, to compute matrix-vector products Jv. If such a function is not supplied, the default is a difference quotient approximation to these products.

```
IDALsJacTimesVecFn
Definition
              typedef int (*IDALsJacTimesVecFn)(realtype tt, N_Vector yy,
                                                      N_Vector yp, N_Vector rr,
                                                      N_Vector v, N_Vector Jv,
                                                      realtype cj, void *user_data,
                                                      N_Vector tmp1, N_Vector tmp2);
Purpose
              This function computes the product Jv of the DAE system Jacobian J (or an approxi-
              mation to it) and a given vector \mathbf{v}, where J is defined by Eq. (2.6).
Arguments
                         is the current value of the independent variable.
              tt
                         is the current value of the dependent variable vector, y(t).
              уу
                         is the current value of \dot{y}(t).
              ур
                         is the current value of the residual vector F(t, y, \dot{y}).
              rr
```

v is the vector by which the Jacobian must be multiplied to the right.

Jv is the computed output vector.

cj is the scalar in the system Jacobian, proportional to the inverse of the step size ( $\alpha$  in Eq. (2.6)).

user\_data is a pointer to user data, the same as the user\_data parameter passed to IDASetUserData.

tmp1

are pointers to memory allocated for variables of type N\_Vector which can be used by IDALsJacTimesVecFn as temporary storage or work space.

Return value The value returned by the Jacobian-times-vector function should be 0 if successful. A nonzero value indicates that a nonrecoverable error occurred.

Notes This function must return a value of J \* v that uses the *current* value of J, i.e. as evaluated at the current  $(t, y, \dot{y})$ .

If the user's IDALsJacTimesVecFn function uses difference quotient approximations, it may need to access quantities not in the call list. These include the current stepsize, the error weights, etc. To obtain these, the user will need to add a pointer to ida\_mem to user\_data and then use the IDAGet\* functions described in §4.5.10.2. The unit roundoff can be accessed as UNIT\_ROUNDOFF defined in sundials\_types.h.

The previous function type IDASpilsJacTimesVecFn is identical to IDALsJacTimesVecFn, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

## 4.6.7 Jacobian-vector product setup (matrix-free linear solvers)

If the user's Jacobian-times-vector requires that any Jacobian-related data be preprocessed or evaluated, then this needs to be done in a user-supplied function of type IDALsJacTimesSetupFn, defined as follows:

## IDAJacTimesSetupFn

Notes

```
Definition typedef int (*IDAJacTimesSetupFn)(realtype tt, N_Vector yy, N_Vector rr, realtype cj, void *user_data);
```

Purpose This function preprocesses and/or evaluates Jacobian data needed by the Jacobian-times-vector routine.

Arguments tt is the current value of the independent variable.

yy is the current value of the dependent variable vector, y(t).

yp is the current value of  $\dot{y}(t)$ .

rr is the current value of the residual vector  $F(t, y, \dot{y})$ .

cj is the scalar in the system Jacobian, proportional to the inverse of the step size ( $\alpha$  in Eq. (2.6)).

user\_data is a pointer to user data, the same as the user\_data parameter passed to IDASetUserData.

Return value The value returned by the Jacobian-vector setup function should be 0 if successful, positive for a recoverable error (in which case the step will be retried), or negative for an unrecoverable error (in which case the integration is halted).

Each call to the Jacobian-vector setup function is preceded by a call to the IDAResFn user function with the same (t,y, yp) arguments. Thus, the setup function can use any auxiliary data that is computed and saved during the evaluation of the DAE residual.

Notes

soon.

If the user's IDALsJacTimesVecFn function uses difference quotient approximations, it may need to access quantities not in the call list. These include the current stepsize, the error weights, etc. To obtain these, the user will need to add a pointer to ida\_mem to user\_data and then use the IDAGet\* functions described in §4.5.10.2. The unit roundoff can be accessed as UNIT\_ROUNDOFF defined in sundials\_types.h.

The previous function type IDASpilsJacTimesSetupFn is identical to IDALsJacTimesSetupFn, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

## 4.6.8 Preconditioner solve (iterative linear solvers)

If a user-supplied preconditioner is to be used with a SUNLINSOL solver module, then the user must provide a function to solve the linear system Pz = r where P is a left preconditioner matrix which approximates (at least crudely) the Jacobian matrix  $J = \partial F/\partial y + cj \partial F/\partial \dot{y}$ . This function must be of type IDALsPrecSolveFn, defined as follows:

```
IDALsPrecSolveFn
Definition
              typedef int (*IDALsPrecSolveFn)(realtype tt, N_Vector yy,
                                                     N_Vector yp, N_Vector rr,
                                                     N_Vector rvec, N_Vector zvec,
                                                     realtype cj, realtype delta,
                                                     void *user_data);
              This function solves the preconditioning system Pz = r.
Purpose
                          is the current value of the independent variable.
Arguments
              tt
                          is the current value of the dependent variable vector, y(t).
              уу
                          is the current value of \dot{y}(t).
              ур
                          is the current value of the residual vector F(t, y, \dot{y}).
              rr
                          is the right-hand side vector r of the linear system to be solved.
              rvec
              zvec
                          is the computed output vector.
                          is the scalar in the system Jacobian, proportional to the inverse of the step
              сj
                          size (\alpha in Eq. (2.6)).
              delta
                          is an input tolerance to be used if an iterative method is employed in the
                          solution. In that case, the residual vector Res = r - Pz of the system should
                          be made less than delta in weighted l_2 norm, i.e., \sqrt{\sum_i (Res_i \cdot ewt_i)^2}
                          delta. To obtain the N_Vector ewt, call IDAGetErrWeights (see §4.5.10.2).
              user_data is a pointer to user data, the same as the user_data parameter passed to
                          the function IDASetUserData.
Return value The value to be returned by the preconditioner solve function is a flag indicating whether
              it was successful. This value should be 0 if successful, positive for a recoverable error
              (in which case the step will be retried), negative for an unrecoverable error (in which
              case the integration is halted).
```

## 4.6.9 Preconditioner setup (iterative linear solvers)

If the user's preconditioner requires that any Jacobian-related data be evaluated or preprocessed, then this needs to be done in a user-supplied function of type IDALsPrecSetupFn, defined as follows:

The previous function type IDASpilsPrecSolveFn is identical to IDALsPrecSolveFn, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name

## IDALsPrecSetupFn

Notes

Definition typedef int (\*IDALsPrecSetupFn)(realtype tt, N\_Vector yy, N\_Vector rr, realtype cj, void \*user\_data);

Purpose This function evaluates and/or preprocesses Jacobian-related data needed by the preconditioner.

Arguments tt is the current value of the independent variable.

yy is the current value of the dependent variable vector, y(t).

yp is the current value of  $\dot{y}(t)$ .

rr is the current value of the residual vector  $F(t, y, \dot{y})$ .

cj is the scalar in the system Jacobian, proportional to the inverse of the step size ( $\alpha$  in Eq. (2.6)).

user\_data is a pointer to user data, the same as the user\_data parameter passed to the function IDASetUserData.

Return value The value returned by the preconditioner setup function is a flag indicating whether it was successful. This value should be 0 if successful, positive for a recoverable error (in which case the step will be retried), negative for an unrecoverable error (in which case the integration is halted).

The operations performed by this function might include forming a crude approximate Jacobian, and performing an LU factorization on the resulting approximation.

Each call to the preconditioner setup function is preceded by a call to the IDAResFn user function with the same (tt, yy, yp) arguments. Thus the preconditioner setup function can use any auxiliary data that is computed and saved during the evaluation of the DAE residual.

This function is not called in advance of every call to the preconditioner solve function, but rather is called only as often as needed to achieve convergence in the nonlinear solver

If the user's IDALsPrecSetupFn function uses difference quotient approximations, it may need to access quantities not in the call list. These include the current stepsize, the error weights, etc. To obtain these, the user will need to add a pointer to ida\_mem to user\_data and then use the IDAGet\* functions described in §4.5.10.2. The unit roundoff can be accessed as UNIT\_ROUNDOFF defined in sundials\_types.h.

The previous function type IDASpilsPrecSetupFn is identical to IDALsPrecSetupFn, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

# 4.7 Integration of pure quadrature equations

IDAS allows the DAE system to include *pure quadratures*. In this case, it is more efficient to treat the quadratures separately by excluding them from the nonlinear solution stage. To do this, begin by excluding the quadrature variables from the vectors yy and yp and the quadrature equations from within res. Thus a separate vector yQ of quadrature variables is to satisfy (d/dt)yQ =  $f_Q(t, y, \dot{y})$ . The following is an overview of the sequence of calls in a user's main program in this situation. Steps that are unchanged from the skeleton program presented in §4.4 are grayed out.

- 1. Initialize parallel or multi-threaded environment, if appropriate
- 2. Set problem dimensions, etc.

This generally includes N, the problem size N (excluding quadrature variables), Nq, the number of quadrature variables, and may include the local vector length Nlocal (excluding quadrature variables), and local number of quadrature variables Nqlocal.

- 3. Set vectors of initial values
- 4. Create IDAS object
- 5. Initialize IDAS solver
- 6. Specify integration tolerances
- 7. Create matrix object
- 8. Create linear solver object
- 9. Set linear solver optional inputs
- 10. Attach linear solver module
- 11. Set optional inputs
- 12. Create nonlinear solver object
- 13. Attach nonlinear solver module
- 14. Set nonlinear solver optional inputs
- 15. Correct initial values

#### 16. Set vector of initial values for quadrature variables

Typically, the quadrature variables should be initialized to 0.

#### 17. Initialize quadrature integration

Call IDAQuadInit to specify the quadrature equation right-hand side function and to allocate internal memory related to quadrature integration. See §4.7.1 for details.

#### 18. Set optional inputs for quadrature integration

Call IDASetQuadErrCon to indicate whether or not quadrature variables should be used in the step size control mechanism. If so, one of the IDAQuad\*tolerances functions must be called to specify the integration tolerances for quadrature variables. See §4.7.4 for details.

19. Advance solution in time

## 20. Extract quadrature variables

Call IDAGetQuad or IDAGetQuadDky to obtain the values of the quadrature variables or their derivatives at the current time. See §4.7.3 for details.

21. Get optional outputs

## 22. Get quadrature optional outputs

Call IDAGetQuad\* functions to obtain optional output related to the integration of quadratures. See §4.7.5 for details.

- 23. Deallocate memory for solution vectors and for the vector of quadrature variables
- 24. Free solver memory

- 25. Free nonlinear solver memory
- 26. Free linear solver and matrix memory
- 27. Finalize MPI, if used

IDAQuadInit can be called and quadrature-related optional inputs (step 18 above) can be set, anywhere between steps 4 and 19.

## 4.7.1 Quadrature initialization and deallocation functions

The function IDAQuadInit activates integration of quadrature equations and allocates internal memory related to these calculations. The form of the call to this function is as follows:

#### IDAQuadInit

Call flag = IDAQuadInit(ida\_mem, rhsQ, yQ0);

Description The function IDAQuadInit provides required problem specifications, allocates internal memory, and initializes quadrature integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

rhsQ (IDAQuadRhsFn) is the C function which computes  $f_Q$ , the right-hand side of the quadrature equations. This function has the form fQ(t, yy, yp, rhsQ, user\_data) (for full details see §4.7.6).

yQ0 (N\_Vector) is the initial value of  $y_Q$ .

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAQuadInit was successful.

IDA\_MEM\_NULL The IDAS memory was not initialized by a prior call to IDACreate.

IDA\_MEM\_FAIL A memory allocation request failed.

Notes If an error occurred, IDAQuadInit also sends an error message to the error handler function.

In terms of the number of quadrature variables  $N_q$  and maximum method order maxord, the size of the real workspace is increased as follows:

- Base value: lenrw = lenrw + (maxord+5) $N_a$
- If IDAQuadSVtolerances is called: lenrw = lenrw  $+N_q$

and the size of the integer workspace is increased as follows:

- Base value: leniw = leniw + (maxord+5) $N_q$
- If IDAQuadSVtolerances is called: leniw = leniw  $+N_a$

The function IDAQuadReInit, useful during the solution of a sequence of problems of same size, reinitializes the quadrature-related internal memory and must follow a call to IDAQuadInit (and maybe a call to IDAQuadInit). The number Nq of quadratures is assumed to be unchanged from the prior call to IDAQuadInit. The call to the IDAQuadReInit function has the following form:

#### IDAQuadReInit

```
Call flag = IDAQuadReInit(ida_mem, yQ0);
```

Description The function IDAQuadReInit provides required problem specifications and reinitializes the quadrature integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

yQ0 (N\_Vector) is the initial value of  $y_Q$ .

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAReInit was successful.

IDA\_MEM\_NULL The IDAS memory was not initialized by a prior call to IDACreate.

<code>IDA\_NO\_QUAD</code> Memory space for the quadrature integration was not allocated by a prior

call to IDAQuadInit.

Notes If an error occurred, IDAQuadReInit also sends an error message to the error handler

function.

#### IDAQuadFree

Call IDAQuadFree(ida\_mem);

Description The function IDAQuadFree frees the memory allocated for quadrature integration.

Arguments The argument is the pointer to the IDAS memory block (of type void \*).

Return value The function IDAQuadFree has no return value.

Notes In general, IDAQuadFree need not be called by the user as it is invoked automatically

by IDAFree.

## 4.7.2 IDAS solver function

Even if quadrature integration was enabled, the call to the main solver function IDASolve is exactly the same as in §4.5.7. However, in this case the return value flag can also be one of the following:

IDA\_QRHS\_FAIL The quadrature right-hand side function failed in an unrecoverable man-

ner.

IDA\_FIRST\_QRHS\_ERR The quadrature right-hand side function failed at the first call.

IDA\_REP\_QRHS\_ERR Convergence test failures occurred too many times due to repeated recov-

erable errors in the quadrature right-hand side function. This value will also be returned if the quadrature right-hand side function had repeated recoverable errors during the estimation of an initial step size (assuming

the quadrature variables are included in the error tests).

## 4.7.3 Quadrature extraction functions

If quadrature integration has been initialized by a call to  ${\tt IDAQuadInit}$ , or reinitialized by a call to  ${\tt IDAQuadReInit}$ , then  ${\tt IDAS}$  computes both a solution and quadratures at time  ${\tt t}$ . However,  ${\tt IDASolve}$  will still return only the solution y in y. Solution quadratures can be obtained using the following function:

#### IDAGetQuad

Call flag = IDAGetQuad(ida\_mem, &tret, yQ);

Description The function IDAGetQuad returns the quadrature solution vector after a successful return

from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

tret (realtype) the time reached by the solver (output).

yQ (N\_Vector) the computed quadrature vector.

Return value The return value flag of IDAGetQuad is one of:

IDA\_SUCCESS IDAGetQuad was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_QUAD Quadrature integration was not initialized.

IDA\_BAD\_DKY yQ is NULL.

The function IDAGetQuadDky computes the k-th derivatives of the interpolating polynomials for the quadrature variables at time t. This function is called by IDAGetQuad with k = 0 and with the current time at which IDASolve has returned, but may also be called directly by the user.

## ${\tt IDAGetQuadDky}$

Call flag = IDAGetQuadDky(ida\_mem, t, k, dkyQ);

Description The function IDAGetQuadDky returns derivatives of the quadrature solution vector after

a successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

t (realtype) the time at which quadrature information is requested. The time t must fall within the interval defined by the last successful step taken by IDAS.

k (int) order of the requested derivative. This must be  $\leq klast$ .

 ${\tt dkyQ} \qquad ({\tt N\_Vector}) \ {\tt the} \ {\tt vector} \ {\tt containing} \ {\tt the} \ {\tt derivative}. \ {\tt This} \ {\tt vector} \ {\tt must} \ {\tt be} \ {\tt allocated}$ 

by the user.

Return value The return value flag of IDAGetQuadDky is one of:

IDA\_SUCCESS IDAGetQuadDky succeeded.

IDA\_MEM\_NULL The pointer to ida\_mem was NULL.

IDA\_NO\_QUAD Quadrature integration was not initialized.

IDA\_BAD\_DKY The vector dkyQ is NULL.

IDA\_BAD\_K k is not in the range 0, 1, ..., klast.

IDA\_BAD\_T The time t is not in the allowed range.

## 4.7.4 Optional inputs for quadrature integration

IDAS provides the following optional input functions to control the integration of quadrature equations.

#### IDASetQuadErrCon

Call flag = IDASetQuadErrCon(ida\_mem, errconQ);

Description The function IDASetQuadErrCon specifies whether or not the quadrature variables are

to be used in the step size control mechanism within IDAS. If they are, the user must call either  ${\tt IDAQuadSStolerances}$  or  ${\tt IDAQuadSVtolerances}$  to specify the integration

tolerances for the quadrature variables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

 $\verb|errconQ| (booleantype) specifies whether quadrature variables are included (\verb|SUNTRUE|)|$ 

or not (SUNFALSE) in the error control mechanism.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL

IDA\_NO\_QUAD Quadrature integration has not been initialized.

Notes By default, errconQ is set to SUNFALSE.

It is illegal to call IDASetQuadErrCon before a call to IDAQuadInit.

If the quadrature variables are part of the step size control mechanism, one of the following functions must be called to specify the integration tolerances for quadrature variables.



## IDAQuadSStolerances

Call flag = IDAQuadSVtolerances(ida\_mem, reltolQ, abstolQ);

Description The function IDAQuadSStolerances specifies scalar relative and absolute tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

reltolQ (realtype) is the scalar relative error tolerance. abstolQ (realtype) is the scalar absolute error tolerance.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set. IDA\_NO\_QUAD Quadrature integration was not initialized.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT One of the input tolerances was negative.

#### IDAQuadSVtolerances

Call flag = IDAQuadSVtolerances(ida\_mem, reltolQ, abstolQ);

Description The function IDAQuadSVtolerances specifies scalar relative and vector absolute toler-

ances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

reltolQ (realtype) is the scalar relative error tolerance. abstolQ (N\_Vector) is the vector absolute error tolerance.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_NO\_QUAD Quadrature integration was not initialized.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT One of the input tolerances was negative.

## 4.7.5 Optional outputs for quadrature integration

IDAS provides the following functions that can be used to obtain solver performance information related to quadrature integration.

#### IDAGetQuadNumRhsEvals

Call flag = IDAGetQuadNumRhsEvals(ida\_mem, &nrhsQevals);

Description The function IDAGetQuadNumRhsEvals returns the number of calls made to the user's

quadrature right-hand side function.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nrhsQevals (long int) number of calls made to the user's rhsQ function.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_QUAD Quadrature integration has not been initialized.

## IDAGetQuadNumErrTestFails

Call flag = IDAGetQuadNumErrTestFails(ida\_mem, &nQetfails);

Description The function IDAGetQuadNumErrTestFails returns the number of local error test fail-

ures due to quadrature variables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nQetfails (long int) number of error test failures due to quadrature variables.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_QUAD Quadrature integration has not been initialized.

## IDAGetQuadErrWeights

Call flag = IDAGetQuadErrWeights(ida\_mem, eQweight);

Description The function IDAGetQuadErrWeights returns the quadrature error weights at the cur-

rent time.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

eQweight (N\_Vector) quadrature error weights at the current time.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_QUAD Quadrature integration has not been initialized.

Notes The user must allocate memory for eQweight.

If quadratures were not included in the error control mechanism (through a call to  ${\tt IDASetQuadErrCon}$  with  ${\tt errconQ}$  =  ${\tt SUNTRUE}$ ),  ${\tt IDAGetQuadErrWeights}$  does not set

the eQweight vector.

#### ${\tt IDAGetQuadStats}$

Call flag = IDAGetQuadStats(ida.mem, &nrhsQevals, &nQetfails);

Description The function IDAGetQuadStats returns the IDAS integrator statistics as a group.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nrhsQevals (long int) number of calls to the user's rhsQ function.

nQetfails (long int) number of error test failures due to quadrature variables.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS the optional output values have been successfully set.

IDA\_MEM\_NULL the ida\_mem pointer is NULL.

IDA\_NO\_QUAD Quadrature integration has not been initialized.

## 4.7.6 User-supplied function for quadrature integration

For integration of quadrature equations, the user must provide a function that defines the right-hand side of the quadrature equations (in other words, the integrand function of the integral that must be evaluated). This function must be of type IDAQuadRhsFn defined as follows:



#### IDAQuadRhsFn

Definition typedef int (\*IDAQuadRhsFn)(realtype t, N\_Vector yy, N\_Vector yp, N\_Vector rhsQ, void \*user\_data);

Purpose This function computes the quadrature equation right-hand side for a given value of the

independent variable t and state vectors y and  $\dot{y}$ .

Arguments t is the current value of the independent variable.

yy is the current value of the dependent variable vector, y(t).

yp is the current value of the dependent variable derivative vector,  $\dot{y}(t)$ .

rhsQ is the output vector  $f_Q(t, y, \dot{y})$ .

user\_data is the user\_data pointer passed to IDASetUserData.

Return value A IDAQuadRhsFn should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if it failed unrecoverably (in which case the integration is halted and IDA\_QRHS\_FAIL is returned).

Notes Allocation of memory for rhsQ is automatically handled within IDAS.

Both y and rhsQ are of type N\_Vector, but they typically have different internal representations. It is the user's responsibility to access the vector data consistently (including the use of the correct accessor macros from each NVECTOR implementation). For the sake of computational efficiency, the vector functions in the two NVECTOR implementations provided with IDAS do not perform any consistency checks with respect to their N\_Vector arguments (see §7.2 and §7.3).

There is one situation in which recovery is not possible even if IDAQuadRhsFn function returns a recoverable error flag. This is when this occurs at the very first call to the IDAQuadRhsFn (in which case IDAS returns IDA\_FIRST\_QRHS\_ERR).

## 4.8 A parallel band-block-diagonal preconditioner module

A principal reason for using a parallel DAE solver such as IDAS lies in the solution of partial differential equations (PDEs). Moreover, the use of a Krylov iterative method for the solution of many such problems is motivated by the nature of the underlying linear system of equations (2.5) that must be solved at each time step. The linear algebraic system is large, sparse, and structured. However, if a Krylov iterative method is to be effective in this setting, then a nontrivial preconditioner needs to be used. Otherwise, the rate of convergence of the Krylov iterative method is usually unacceptably slow. Unfortunately, an effective preconditioner tends to be problem-specific.

However, we have developed one type of preconditioner that treats a rather broad class of PDE-based problems. It has been successfully used for several realistic, large-scale problems [32] and is included in a software module within the IDAS package. This module works with the parallel vector module NVECTOR\_PARALLEL and generates a preconditioner that is a block-diagonal matrix with each block being a band matrix. The blocks need not have the same number of super- and sub-diagonals, and these numbers may vary from block to block. This Band-Block-Diagonal Preconditioner module is called IDABBDPRE.

One way to envision these preconditioners is to think of the domain of the computational PDE problem as being subdivided into M non-overlapping sub-domains. Each of these sub-domains is then assigned to one of the M processors to be used to solve the DAE system. The basic idea is to isolate the preconditioning so that it is local to each processor, and also to use a (possibly cheaper) approximate residual function. This requires the definition of a new function  $G(t, y, \dot{y})$  which approximates the function  $F(t, y, \dot{y})$  in the definition of the DAE system (2.1). However, the user may set G = F. Corresponding to the domain decomposition, there is a decomposition of the solution vectors y and  $\dot{y}$  into M disjoint blocks  $y_m$  and  $\dot{y}_m$ , and a decomposition of G into blocks  $G_m$ . The block  $G_m$  depends on  $y_m$  and  $\dot{y}_m$ , and also on components of  $y_{m'}$  and  $\dot{y}_{m'}$  associated with neighboring sub-domains

(so-called ghost-cell data). Let  $\bar{y}_m$  and  $\bar{y}_m$  denote  $y_m$  and  $y_m$  (respectively) augmented with those other components on which  $G_m$  depends. Then we have

$$G(t, y, \dot{y}) = [G_1(t, \bar{y}_1, \bar{y}_1), G_2(t, \bar{y}_2, \bar{y}_2), \dots, G_M(t, \bar{y}_M, \bar{y}_M)]^T,$$
(4.1)

and each of the blocks  $G_m(t, \bar{y}_m, \bar{\dot{y}}_m)$  is uncoupled from the others.

The preconditioner associated with this decomposition has the form

$$P = diag[P_1, P_2, \dots, P_M] \tag{4.2}$$

where

$$P_m \approx \partial G_m / \partial y_m + \alpha \partial G_m / \partial \dot{y}_m \tag{4.3}$$

This matrix is taken to be banded, with upper and lower half-bandwidths mudq and mldq defined as the number of non-zero diagonals above and below the main diagonal, respectively. The difference quotient approximation is computed using  $\operatorname{mudq} + \operatorname{mldq} + 2$  evaluations of  $G_m$ , but only a matrix of bandwidth  $\operatorname{mukeep} + \operatorname{mlkeep} + 1$  is retained.

Neither pair of parameters need be the true half-bandwidths of the Jacobians of the local block of G, if smaller values provide a more efficient preconditioner. Such an efficiency gain may occur if the couplings in the DAE system outside a certain bandwidth are considerably weaker than those within the band. Reducing mukeep and mlkeep while keeping mudq and mldq at their true values, discards the elements outside the narrower band. Reducing both pairs has the additional effect of lumping the outer Jacobian elements into the computed elements within the band, and requires more caution and experimentation.

The solution of the complete linear system

$$Px = b (4.4)$$

reduces to solving each of the equations

$$P_m x_m = b_m (4.5)$$

and this is done by banded LU factorization of  $P_m$  followed by a banded backsolve.

Similar block-diagonal preconditioners could be considered with different treatment of the blocks  $P_m$ . For example, incomplete LU factorization or an iterative method could be used instead of banded LU factorization

The IDABBDPRE module calls two user-provided functions to construct P: a required function  $\operatorname{Gres}$  (of type IDABBDLocalFn) which approximates the residual function  $G(t,y,\dot{y})\approx F(t,y,\dot{y})$  and which is computed locally, and an optional function  $\operatorname{Gcomm}$  (of type IDABBDCommFn) which performs all inter-process communication necessary to evaluate the approximate residual G. These are in addition to the user-supplied residual function res. Both functions take as input the same pointer user\_data as passed by the user to IDASetUserData and passed to the user's function res. The user is responsible for providing space (presumably within user\_data) for components of yy and yp that are communicated by Gcomm from the other processors, and that are then used by Gres, which should not do any communication.

## IDABBDLocalFn

Definition typedef int (\*IDABBDLocalFn)(sunindextype Nlocal, realtype tt, N\_Vector yy, N\_Vector yp, N\_Vector gval, void \*user\_data);

Purpose This Gres function computes  $G(t, y, \dot{y})$ . It loads the vector gval as a function of tt, yy, and yp.

Arguments Nlocal is the local vector length.

tt is the value of the independent variable.

yy is the dependent variable.

yp is the derivative of the dependent variable.

gval is the output vector.

user\_data is a pointer to user data, the same as the user\_data parameter passed to IDASetUserData.

Return value An IDABBDLocalFn function type should return 0 to indicate success, 1 for a recoverable error, or -1 for a non-recoverable error.

Notes

This function must assume that all inter-processor communication of data needed to calculate gval has already been done, and this data is accessible within user\_data.

The case where G is mathematically identical to F is allowed.

#### IDABBDCommFn

Definition typedef int (\*IDABBDCommFn)(sunindextype Nlocal, realtype tt, N\_Vector yy, N\_Vector yp, void \*user\_data);

Purpose This Gcomm function performs all inter-processor communications necessary for the execution of the Gres function above, using the input vectors yy and yp.

Arguments Nlocal is the local vector length.

tt is the value of the independent variable.

yy is the dependent variable.

yp is the derivative of the dependent variable.

user\_data is a pointer to user data, the same as the user\_data parameter passed to IDASetUserData.

Return value An IDABBDCommFn function type should return 0 to indicate success, 1 for a recoverable error, or -1 for a non-recoverable error.

Notes

The Gcomm function is expected to save communicated data in space defined within the structure user\_data.

Each call to the Gcomm function is preceded by a call to the residual function res with the same (tt, yy, yp) arguments. Thus Gcomm can omit any communications done by res if relevant to the evaluation of Gres. If all necessary communication was done in res, then Gcomm = NULL can be passed in the call to IDABBDPrecInit (see below).

Besides the header files required for the integration of the DAE problem (see §4.3), to use the IDABBDPRE module, the main program must include the header file idas\_bbdpre.h which declares the needed function prototypes.

The following is a summary of the usage of this module and describes the sequence of calls in the user main program. Steps that are unchanged from the user main program presented in  $\S4.4$  are grayed-out.

- 1. Initialize MPI
- 2. Set problem dimensions etc.
- 3. Set vectors of initial values
- 4. Create IDAS object
- 5. Initialize IDAS solver
- 6. Specify integration tolerances

#### 7. Create linear solver object

When creating the iterative linear solver object, specify the use of left preconditioning (PREC\_LEFT) as IDAS only supports left preconditioning.

8. Set linear solver optional inputs

#### 9. Attach linear solver module

10. Set optional inputs

Note that the user should not overwrite the preconditioner setup function or solve function through calls to idIDASetPreconditioner optional input function.

#### 11. Initialize the IDABBDPRE preconditioner module

Specify the upper and lower bandwidths mudq, mldq and mukeep, mlkeep and call

to allocate memory and initialize the internal preconditioner data. The last two arguments of <code>IDABBDPrecInit</code> are the two user-supplied functions described above.

- 12. Create nonlinear solver object
- 13. Attach nonlinear solver module
- 14. Set nonlinear solver optional inputs
- 15. Correct initial values
- 16. Specify rootfinding problem
- 17. Advance solution in time

#### 18. Get optional outputs

Additional optional outputs associated with IDABBDPRE are available by way of two routines described below, IDABBDPrecGetWorkSpace and IDABBDPrecGetNumGfnEvals.

- 19. Deallocate memory for solution vectors
- 20. Free solver memory
- 21. Free nonlinear solver memory
- 22. Free linear solver memory

mukeep

23. Finalize MPI

The user-callable functions that initialize (step 11 above) or re-initialize the IDABBDPRE preconditioner module are described next.

## IDABBDPrecInit

Description The function IDABBDPrecInit initializes and allocates (internal) memory for the ID-ABBDPRE preconditioner.

 $\label{eq:conditional} \mbox{Arguments} \quad \mbox{ida\_mem} \quad \mbox{(void *) pointer to the IDAS memory block.}$ 

Nlocal (sunindextype) local vector dimension.

mudq (sunindextype) upper half-bandwidth to be used in the difference-quotient Jacobian approximation.

mldq (sunindextype) lower half-bandwidth to be used in the difference-quotient Jacobian approximation.

(sunindextype) upper half-bandwidth of the retained banded approximate

Jacobian block.

mlkeep (sunindextype) lower half-bandwidth of the retained banded approximate Jacobian block.

dq\_rel\_yy (realtype) the relative increment in components of y used in the difference quotient approximations. The default is dq\_rel\_yy=  $\sqrt{\text{unit roundoff}}$ , which can be specified by passing dq\_rel\_yy= 0.0.

Gres (IDABBDLocalFn) the C function which computes the local residual approximation  $G(t, y, \dot{y})$ .

Gcomm (IDABBDCommFn) the optional C function which performs all inter-process communication required for the computation of  $G(t, y, \dot{y})$ .

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The call to IDABBDPrecInit was successful.

IDALS\_MEM\_NULL The ida\_mem pointer was NULL.

IDALS\_MEM\_FAIL A memory allocation request has failed.

IDALS\_LMEM\_NULL An IDALS linear solver memory was not attached.

IDALS\_ILL\_INPUT The supplied vector implementation was not compatible with the block band preconditioner.

Notes

If one of the half-bandwidths mudq or mldq to be used in the difference-quotient calculation of the approximate Jacobian is negative or exceeds the value Nlocal-1, it is replaced by 0 or Nlocal-1 accordingly.

The half-bandwidths mudq and mldq need not be the true half-bandwidths of the Jacobian of the local block of G, when smaller values may provide a greater efficiency.

Also, the half-bandwidths mukeep and mlkeep of the retained banded approximate Jacobian block may be even smaller, to reduce storage and computation costs further.

For all four half-bandwidths, the values need not be the same on every processor.

The IDABBDPRE module also provides a reinitialization function to allow for a sequence of problems of the same size, with the same linear solver choice, provided there is no change in local\_N, mukeep, or mlkeep. After solving one problem, and after calling IDAReInit to re-initialize IDAS for a subsequent problem, a call to IDABBDPrecReInit can be made to change any of the following: the half-bandwidths mudq and mldq used in the difference-quotient Jacobian approximations, the relative increment dq\_rel\_yy, or one of the user-supplied functions Gres and Gcomm. If there is a change in any of the linear solver inputs, an additional call to the "Set" routines provided by the SUNLINSOL module, and/or one or more of the corresponding IDASet\*\*\* functions, must also be made (in the proper order).

#### IDABBDPrecReInit

Call flag = IDABBDPrecReInit(ida\_mem, mudq, mldq, dq\_rel\_yy);

Description The function IDABBDPrecReInit reinitializes the IDABBDPRE preconditioner.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

mudq (sunindextype) upper half-bandwidth to be used in the difference-quotient

Jacobian approximation.

mldq (sunindextype) lower half-bandwidth to be used in the difference-quotient Jacobian approximation.

dq\_rel\_yy (realtype) the relative increment in components of y used in the difference quotient approximations. The default is  $dq_rel_yy = \sqrt{unit roundoff}$ , which can be specified by passing  $dq_rel_yy = 0.0$ .

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The call to IDABBDPrecReInit was successful.

IDALS\_MEM\_NULL The ida\_mem pointer was NULL.

IDALS\_LMEM\_NULL An IDALS linear solver memory was not attached.

IDALS\_PMEM\_NULL The function IDABBDPrecInit was not previously called.

Notes If one of the half-bandwidths mudq or mldq is negative or exceeds the value Nlocal-1,

it is replaced by 0 or Nlocal-1, accordingly.

The following two optional output functions are available for use with the IDABBDPRE module:

## IDABBDPrecGetWorkSpace

Call flag = IDABBDPrecGetWorkSpace(ida\_mem, &lenrwBBDP, &leniwBBDP);

Description The function IDABBDPrecGetWorkSpace returns the local sizes of the IDABBDPRE real

and integer workspaces.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

lenrwBBDP (long int) local number of real values in the IDABBDPRE workspace.

leniwBBDP (long int) local number of integer values in the IDABBDPRE workspace.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer was NULL.

IDALS\_PMEM\_NULL The IDABBDPRE preconditioner has not been initialized.

Notes The workspace requirements reported by this routine correspond only to memory allo-

cated within the IDABBDPRE module (the banded matrix approximation, banded SUN-LINSOL object, temporary vectors). These values are local to each process.

The workspaces referred to here exist in addition to those given by the corresponding

function IDAGetLinWorkSpace.

## IDABBDPrecGetNumGfnEvals

Call flag = IDABBDPrecGetNumGfnEvals(ida\_mem, &ngevalsBBDP);

Description The function IDABBDPrecGetNumGfnEvals returns the cumulative number of calls to

the user Gres function due to the finite difference approximation of the Jacobian blocks

used within IDABBDPRE's preconditioner setup function.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

ngevalsBBDP (long int) the cumulative number of calls to the user Gres function.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional output value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer was NULL.

IDALS\_PMEM\_NULL The IDABBDPRE preconditioner has not been initialized.

In addition to the ngevalsBBDP Gres evaluations, the costs associated with IDABBDPRE also include nlinsetups LU factorizations, nlinsetups calls to Gcomm, npsolves banded backsolve calls, and nrevalsLS residual function evaluations, where nlinsetups is an optional IDAS output (see  $\S4.5.10.2$ ), and npsolves and nrevalsLS are linear solver optional outputs (see  $\S4.5.10.5$ ).

# Chapter 5

# Using IDAS for Forward Sensitivity Analysis

This chapter describes the use of IDAS to compute solution sensitivities using forward sensitivity analysis. One of our main guiding principles was to design the IDAS user interface for forward sensitivity analysis as an extension of that for IVP integration. Assuming a user main program and user-defined support routines for IVP integration have already been defined, in order to perform forward sensitivity analysis the user only has to insert a few more calls into the main program and (optionally) define an additional routine which computes the residuals for sensitivity systems (2.12). The only departure from this philosophy is due to the IDAResFn type definition (§4.6.1). Without changing the definition of this type, the only way to pass values of the problem parameters to the DAE residual function is to require the user data structure user\_data to contain a pointer to the array of real parameters p.

IDAS uses various constants for both input and output. These are defined as needed in this chapter, but for convenience are also listed separately in Appendix B.

We begin with a brief overview, in the form of a skeleton user program. Following that are detailed descriptions of the interface to the various user-callable routines and of the user-supplied routines that were not already described in Chapter 4.

## 5.1 A skeleton of the user's main program

The following is a skeleton of the user's main program (or calling program) as an application of IDAS. The user program is to have these steps in the order indicated, unless otherwise noted. For the sake of brevity, we defer many of the details to the later sections. As in §4.4, most steps are independent of the NVECTOR, SUNMATRIX, SUNLINSOL, and SUNNONLINSOL implementations used. For the steps that are not, refer to Chapter 7, 8, 9, and 10 for the specific name of the function to be called or macro to be referenced.

Differences between the user main program in  $\S4.4$  and the one below start only at step (16). Steps that are unchanged from the skeleton program presented in  $\S4.4$  are grayed out.

First, note that no additional header files need be included for forward sensitivity analysis beyond those for IVP solution (§4.4).

- 1. Initialize parallel or multi-threaded environment, if appropriate
- 2. Set problem dimensions etc.
- 3. Set vectors of initial values
- 4. Create IDAS object
- 5. Initialize IDAS solver

- 6. Specify integration tolerances
- 7. Create matrix object
- 8. Create linear solver object
- 9. Set linear solver optional inputs
- 10. Attach linear solver module
- 11. Set optional inputs
- 12. Create nonlinear solver object
- 13. Attach nonlinear solver module
- 14. Set nonlinear solver optional inputs
- 15. Initialize quadrature problem, if not sensitivity-dependent

## 16. Define the sensitivity problem

•Number of sensitivities (required)

Set  $Ns = N_s$ , the number of parameters with respect to which sensitivities are to be computed.

•Problem parameters (optional)

If IDAS is to evaluate the residuals of the sensitivity systems, set p, an array of Np real parameters upon which the IVP depends. Only parameters with respect to which sensitivities are (potentially) desired need to be included. Attach p to the user data structure user\_data. For example, user\_data->p = p;

If the user provides a function to evaluate the sensitivity residuals, p need not be specified.

•Parameter list (optional)

If IDAS is to evaluate the sensitivity residuals, set plist, an array of Ns integers to specify the parameters p with respect to which solution sensitivities are to be computed. If sensitivities with respect to the j-th parameter p[j]  $(0 \le j < Np)$  are desired, set plist<sub>i</sub> = j, for some  $i = 0, ..., N_s - 1$ .

If plist is not specified, IDAS will compute sensitivities with respect to the first Ns parameters; i.e., plist<sub>i</sub> = i ( $i = 0, ..., N_s - 1$ ).

If the user provides a function to evaluate the sensitivity residuals, plist need not be specified.

•Parameter scaling factors (optional)

If IDAS is to estimate tolerances for the sensitivity solution vectors (based on tolerances for the state solution vector) or if IDAS is to evaluate the residuals of the sensitivity systems using the internal difference-quotient function, the results will be more accurate if order of magnitude information is provided.

Set pbar, an array of Ns positive scaling factors. Typically, if  $p_i \neq 0$ , the value  $\bar{p}_i = |p_{\text{plist}_i}|$  can be used.

If pbar is not specified, IDAS will use  $\bar{p}_i = 1.0$ .

If the user provides a function to evaluate the sensitivity residual and specifies tolerances for the sensitivity variables, **pbar** need not be specified.

Note that the names for p, pbar, plist, as well as the field p of user\_data are arbitrary, but they must agree with the arguments passed to IDASetSensParams below.

#### 17. Set sensitivity initial conditions

Set the Ns vectors yS0[i] and ypS0[i] of initial values for sensitivities (for i = 0, ..., Ns - 1), using the appropriate functions defined by the particular NVECTOR implementation chosen.

First, create an array of Ns vectors by making the appropriate call

```
yS0 = N_VCloneVectorArray_***(Ns, y0);
or
yS0 = N_VCloneVectorArrayEmpty_***(Ns, y0);
```

Here the argument y0 serves only to provide the N\_Vector type for cloning.

Then, for each i = 0, ..., Ns - 1, load initial values for the i-th sensitivity vector ySO[i].

Set the initial conditions for the Ns sensitivity derivative vectors ypS0 of  $\dot{y}$  similarly.

#### 18. Activate sensitivity calculations

Call flag = IDASensInit(...); to activate forward sensitivity computations and allocate internal memory for IDAS related to sensitivity calculations (see §5.2.1).

#### 19. Set sensitivity tolerances

Call IDASensStolerances, IDASensSVtolerances, or IDASensEEtolerances. See §5.2.2.

#### 20. Set sensitivity analysis optional inputs

Call IDASetSens\* routines to change from their default values any optional inputs that control the behavior of IDAS in computing forward sensitivities. See §5.2.7.

#### 21. Create sensitivity nonlinear solver object (optional)

If using a non-default nonlinear solver (see §5.2.3), then create the desired nonlinear solver object by calling the appropriate constructor function defined by the particular SUNNONLINSOL implementation e.g.,

```
NLSSens = SUNNonlinSol_***Sens(...);
```

where \*\*\* is the name of the nonlinear solver and ... are constructor specific arguments (see Chapter 10 for details).

### 22. Attach the sensitvity nonlinear solver module (optional)

If using a non-default nonlinear solver, then initialize the nonlinear solver interface by attaching the nonlinear solver object by calling

```
ier = IDASetNonlinearSolverSensSim(ida_mem, NLSSens);
when using the IDA_SIMULTANEOUS corrector method or
```

```
ier = IDASetNonlinearSolverSensStg(ida_mem, NLSSens);
```

when using the IDA\_STAGGERED corrector method (see §5.2.3 for details).

## 23. Set sensitivity nonlinear solver optional inputs (optional)

Call the appropriate set functions for the selected nonlinear solver module to change optional inputs specific to that nonlinear solver. These *must* be called after <code>IDASensInit</code> if using the default nonlinear solver or after attaching a new nonlinear solver to <code>IDAS</code>, otherwise the optional inputs will be overridden by <code>IDAS</code> defaults. See Chapter 10 for more information on optional inputs.

- 24. Correct initial values
- 25. Specify rootfinding problem
- 26. Advance solution in time

#### 27. Extract sensitivity solution

After each successful return from IDASolve, the solution of the original IVP is available in the y argument of IDASolve, while the sensitivity solution can be extracted into yS and ypS (which can be the same as ySO and ypSO, respectively) by calling one of the following routines: IDAGetSens, IDAGetSensDky or IDAGetSensDky1 (see §5.2.6).

- 28. Get optional outputs
- 29. Deallocate memory for solution vector

## 30. Deallocate memory for sensitivity vectors

Upon completion of the integration, deallocate memory for the vectors contained in ySO and ypSO:

```
N_VDestroyVectorArray_***(yS0, Ns);
```

If yS was created from realtype arrays yS\_i, it is the user's responsibility to also free the space for the arrays yS\_i, and likewise for ypS.

- 31. Free user data structure
- 32. Free solver memory
- 33. Free nonlinear solver memory
- 34. Free vector specification memory
- 35. Free linear solver and matrix memory
- 36. Finalize MPI, if used

## 5.2 User-callable routines for forward sensitivity analysis

This section describes the IDAS functions, in addition to those presented in §4.5, that are called by the user to set up and solve a forward sensitivity problem.

## 5.2.1 Forward sensitivity initialization and deallocation functions

Activation of forward sensitivity computation is done by calling IDASensInit. The form of the call to this routine is as follows:

## IDASensInit

```
Call flag = IDASensInit(ida_mem, Ns, ism, resS, yS0, ypS0);
```

Description The routine IDASensInit activates forward sensitivity computations and allocates internal memory related to sensitivity calculations.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

Ns (int) the number of sensitivities to be computed.

ism (int) a flag used to select the sensitivity solution method. Its value can be either IDA\_SIMULTANEOUS or IDA\_STAGGERED:

- In the IDA\_SIMULTANEOUS approach, the state and sensitivity variables are corrected at the same time. If the default Newton nonlinear solver is used, this amounts to performing a modified Newton iteration on the combined nonlinear system;
- In the IDA\_STAGGERED approach, the correction step for the sensitivity variables takes place at the same time for all sensitivity equations, but only after the correction of the state variables has converged and the state variables have passed the local error test;

resS (IDASensResFn) is the C function which computes the residual of the sensitivity DAE. For full details see §5.3.

yS0 (N\_Vector \*) a pointer to an array of Ns vectors containing the initial values of the sensitivities of y.

ypS0 (N\_Vector \*) a pointer to an array of Ns vectors containing the initial values of the sensitivities of  $\dot{y}$ .

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASensInit was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to IDACreate.

IDA\_MEM\_FAIL A memory allocation request has failed.

IDA\_ILL\_INPUT An input argument to IDASensInit has an illegal value.

Notes Passing resS=NULL indicates using the default internal difference quotient sensitivity residual routine.

If an error occurred, IDASensInit also prints an error message to the file specified by the optional input errfp.

In terms of the problem size N, number of sensitivity vectors  $N_s$ , and maximum method order maxord, the size of the real workspace is increased as follows:

- Base value: lenrw = lenrw + (maxord+5) $N_sN$
- With IDASensSVtolerances: lenrw = lenrw  $+N_sN$

the size of the integer workspace is increased as follows:

- Base value: leniw = leniw +  $(maxord+5)N_sN_i$
- With IDASensSVtolerances: leniw = leniw  $+N_sN_i$ ,

where  $N_i$  is the number of integer words in one N\_Vector.

The routine IDASensReInit, useful during the solution of a sequence of problems of same size, reinitializes the sensitivity-related internal memory and must follow a call to IDASensInit (and maybe a call to IDAReInit). The number Ns of sensitivities is assumed to be unchanged since the call to IDASensInit. The call to the IDASensReInit function has the form:

### IDASensReInit

Call flag = IDASensReInit(ida\_mem, ism, yS0, ypS0);

Description The routine IDASensReInit reinitializes forward sensitivity computations.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

ism (int) a flag used to select the sensitivity solution method. Its value can be either IDA\_SIMULTANEOUS or IDA\_STAGGERED.

ySO (N\_Vector \*) a pointer to an array of Ns variables of type N\_Vector containing the initial values of the sensitivities of y.

ypS0 (N\_Vector \*) a pointer to an array of Ns variables of type N\_Vector containing the initial values of the sensitivities of  $\dot{y}$ .

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAReInit was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_NO\_SENS Memory space for sensitivity integration was not allocated through a

previous call to IDASensInit.

IDA\_ILL\_INPUT An input argument to IDASensReInit has an illegal value.

IDA\_MEM\_FAIL A memory allocation request has failed.

Notes All arguments of IDASensReInit are the same as those of IDASensInit.

If an error occurred, IDASensReInit also prints an error message to the file specified

by the optional input errfp.

To deallocate all forward sensitivity-related memory (allocated in a prior call to IDASensInit), the user must call

#### IDASensFree

Call IDASensFree(ida\_mem);

Description The function IDASensFree frees the memory allocated for forward sensitivity compu-

tations by a previous call to IDASensInit.

Arguments The argument is the pointer to the IDAS memory block (of type void \*).

Return value The function IDASensFree has no return value.

Notes In general, IDASensFree need not be called by the user as it is invoked automatically

by IDAFree.

After a call to IDASensFree, forward sensitivity computations can be reactivated only

by calling IDASensInit again.

To activate and deactivate forward sensitivity calculations for successive IDAS runs, without having to allocate and deallocate memory, the following function is provided:

#### IDASensToggleOff

Call IDASensToggleOff(ida\_mem);

Description The function IDASensToggleOff deactivates forward sensitivity calculations. It does

not deallocate sensitivity-related memory.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

Return value The return value flag of IDASensToggle is one of:

IDA\_SUCCESS IDASensToggleOff was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

Notes Since sensitivity-related memory is not deallocated, sensitivities can be reactivated at

a later time (using IDASensReInit).

## 5.2.2 Forward sensitivity tolerance specification functions

One of the following three functions must be called to specify the integration tolerances for sensitivities. Note that this call must be made after the call to IDASensInit.

#### IDASensSStolerances

Call flag = IDASensSStolerances(ida\_mem, reltolS, abstolS);

Description The function IDASensSStolerances specifies scalar relative and absolute tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

reltolS (realtype) is the scalar relative error tolerance.

abstolS (realtype\*) is a pointer to an array of length Ns containing the scalar absolute error tolerances.

Return value The return flag flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASStolerances was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_NO\_SENS The sensitivity allocation function IDASensInit has not been called.

IDA\_ILL\_INPUT One of the input tolerances was negative.

#### IDASensSVtolerances

Call flag = IDASensSVtolerances(ida\_mem, reltolS, abstolS);

Description The function IDASensSVtolerances specifies scalar relative tolerance and vector abso-

lute tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

reltolS (realtype) is the scalar relative error tolerance.

abstolS (N\_Vector\*) is an array of Ns variables of type N\_Vector. The N\_Vector from

abstolS[is] specifies the vector tolerances for is-th sensitivity.

Return value The return flag flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASVtolerances was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_NO\_SENS The sensitivity allocation function IDASensInit has not been called.

 ${\tt IDA\_ILL\_INPUT \ The \ relative \ error \ tolerance \ was \ negative \ or \ one \ of \ the \ absolute \ tolerance}$ 

vectors had a negative component.

Notes This choice of tolerances is important when the absolute error tolerance needs to be

different for each component of any vector yS[i].

#### IDASensEEtolerances

Call flag = IDASensEEtolerances(ida\_mem);

Description When IDASensEEtolerances is called, IDAS will estimate tolerances for sensitivity vari-

ables based on the tolerances supplied for states variables and the scaling factors  $\bar{p}$ .

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

Return value The return flag flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASensEEtolerances was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_NO\_SENS The sensitivity allocation function IDASensInit has not been called.

#### 5.2.3 Forward sensitivity nonlinear solver interface functions

As in the pure DAE case, when computing solution sensitivities using forward sensitivitiy analysis IDAS uses the SUNNONLINSOL implementation of Newton's method defined by the SUNNONLINSOL\_NEWTON module (see §10.2) by default. To specify a different nonlinear solver in IDAS, the user's program must create a SUNNONLINSOL object by calling the appropriate constructor routine. The user must then attach the SUNNONLINSOL object to IDAS by calling either IDASetNonlinearSolverSensSim when using the IDA\_SIMULTANEOUS corrector option, or IDASetNonlinearSolver (see §4.5.4) and IDASetNonlinearSolverSensStg when using the IDA\_STAGGERED corrector option, as documented below.

When changing the nonlinear solver in IDAS, IDASetNonlinearSolver must be called after IDAInit; similarly IDASetNonlinearSolverSensSim and IDASetNonlinearSolverStg must be called after IDASensInit. If any calls to IDASolve have been made, then IDAS will need to be reinitialized by calling IDAReInit to ensure that the nonlinear solver is initialized correctly before any subsequent calls to IDASolve.

The first argument passed to the routines IDASetNonlinearSolverSensSim and IDASetNonlinearSolverSensStg is the IDAS memory pointer returned by IDACreate and the second argument is the SUNNONLINSOL object to use for solving the nonlinear system 2.4. A call to this function attaches the nonlinear solver to the main IDAS integrator. We note that at present, the SUNNONLINSOL object must be of type SUNNONLINEARSOLVER\_ROOTFIND.

#### IDASetNonlinearSolverSensSim

Call flag = IDASetNonlinearSolverSensSim(ida\_mem, NLS);

Description The function IDASetNonLinearSolverSensSim attaches a SUNNONLINSOL object (NLS)

to IDAS when using the IDA\_SIMULTANEOUS approach to correct the state and sensitivity

variables at the same time.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

NLS (SUNNonlinearSolver) SUNNONLINSOL object to use for solving nonlinear sys-

tems.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The nonlinear solver was successfully attached.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The SUNNONLINSOL object is NULL, does not implement the required

nonlinear solver operations, is not of the correct type, or the residual function, convergence test function, or maximum number of nonlinear

iterations could not be set.

#### IDASetNonlinearSolverSensStg

Call flag = IDASetNonlinearSolverSensStg(ida\_mem, NLS);

Description The function IDASetNonLinearSolverSensStg attaches a SUNNONLINSOL object (NLS)

to IDAS when using the IDA\_STAGGERED approach to correct the sensitivity variables

after the correction of the state variables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

NLS (SUNNonlinearSolver) SUNNONLINSOL object to use for solving nonlinear sys-

tems.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The nonlinear solver was successfully attached.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT The SUNNONLINSOL object is NULL, does not implement the required nonlinear solver operations, is not of the correct type, or the residual function, convergence test function, or maximum number of nonlinear iterations could not be set.

Notes

This function only attaches the SUNNONLINSOL object for correcting the sensitivity variables. To attach a SUNNONLINSOL object for the state variable correction use IDASetNonlinearSolver (see  $\S4.5.4$ ).

#### 5.2.4 Forward sensitivity initial condition calculation function

IDACalcIC also calculates corrected initial conditions for sensitivity variables of a DAE system. When used for initial conditions calculation of the forward sensitivities, IDACalcIC must be preceded by successful calls to IDASensInit (or IDASensReInit) and should precede the call(s) to IDASolve. For restrictions that apply for initial conditions calculation of the state variables, see §4.5.5.

Calling IDACalcIC is optional. It is only necessary when the initial conditions do not satisfy the sensitivity systems. Even if forward sensitivity analysis was enabled, the call to the initial conditions calculation function IDACalcIC is exactly the same as for state variables.

```
flag = IDACalcIC(ida_mem, icopt, tout1);
```

See §4.5.5 for a list of possible return values.

#### 5.2.5 **IDAS** solver function

Even if forward sensitivity analysis was enabled, the call to the main solver function IDASolve is exactly the same as in §4.5.7. However, in this case the return value flag can also be one of the following:

IDA\_SRES\_FAIL The sensitivity residual function failed in an unrecoverable manner.

IDA\_REP\_SRES\_ERR The user's residual function repeatedly returned a recoverable error flag, but the solver was unable to recover.

#### 5.2.6 Forward sensitivity extraction functions

If forward sensitivity computations have been initialized by a call to IDASensInit, or reinitialized by a call to IDASensReInit, then IDAS computes both a solution and sensitivities at time t. However, IDASolve will still return only the solutions y and  $\dot{y}$  in yret and ypret, respectively. Solution sensitivities can be obtained through one of the following functions:

#### IDAGetSens

Call flag = IDAGetSens(ida\_mem, &tret, yS);

Description The function IDAGetSens returns the sensitivity solution vectors after a successful return

from IDASolve.

ida\_mem (void \*) pointer to the memory previously allocated by IDAInit. Arguments

(realtype) the time reached by the solver (output).

(N\_Vector \*) the array of Ns computed forward sensitivity vectors. vS

Return value The return value flag of IDAGetSens is one of:

IDA\_SUCCESS IDAGetSens was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_BAD\_DKY vS is NULL.

Notes Note that the argument tret is an output for this function. Its value will be the same

as that returned at the last IDASolve call.

The function IDAGetSensDky computes the k-th derivatives of the interpolating polynomials for the sensitivity variables at time t. This function is called by IDAGetSens with k=0, but may also be called directly by the user.

#### IDAGetSensDky

Call flag = IDAGetSensDky(ida\_mem, t, k, dkyS);

Description The function IDAGetSensDky returns derivatives of the sensitivity solution vectors after a successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

t (realtype) specifies the time at which sensitivity information is requested. The time t must fall within the interval defined by the last successful step taken by IDAS.

k (int) order of derivatives.

dkyS (N\_Vector \*) array of Ns vectors containing the derivatives on output. The space for dkyS must be allocated by the user.

Return value The return value flag of IDAGetSensDky is one of:

IDA\_SUCCESS IDAGetSensDky succeeded.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_BAD\_DKY dkyS or one of the vectors dkyS[i] is NULL.

 ${\tt IDA\_BAD\_K} \qquad {\tt k} \text{ is not in the range } 0,1,...,klast.$ 

IDA\_BAD\_T The time t is not in the allowed range.

Forward sensitivity solution vectors can also be extracted separately for each parameter in turn through the functions IDAGetSens1 and IDAGetSensDky1, defined as follows:

#### IDAGetSens1

Call flag = IDAGetSens1(ida\_mem, &tret, is, yS);

Description The function IDAGetSens1 returns the is-th sensitivity solution vector after a successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

tret (realtype \*) the time reached by the solver (output).

is (int) specifies which sensitivity vector is to be returned  $(0 \le is < N_s)$ .

yS (N\_Vector) the computed forward sensitivity vector.

Return value The return value flag of IDAGetSens1 is one of:

IDA\_SUCCESS IDAGetSens1 was successful.

 ${\tt IDA\_MEM\_NULL} \ \, {\tt ida\_mem} \ \, {\tt was} \ \, {\tt NULL}.$ 

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_BAD\_IS The index is is not in the allowed range.

IDA\_BAD\_DKY yS is NULL.

IDA\_BAD\_T The time t is not in the allowed range.

Notes Note that the argument tret is an output for this function. Its value will be the same as that returned at the last IDASolve call.

#### IDAGetSensDky1

Call flag = IDAGetSensDky1(ida\_mem, t, k, is, dkyS);

Description The function IDAGetSensDky1 returns the k-th derivative of the is-th sensitivity solution vector after a successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

t (realtype) specifies the time at which sensitivity information is requested.

The time t must fall within the interval defined by the last successful step taken by IDAS.

k (int) order of derivative.

is (int) specifies the sensitivity derivative vector to be returned  $(0 \le is < N_s)$ .

dkyS (N\_Vector) the vector containing the derivative on output. The space for dkyS must be allocated by the user.

Return value The return value flag of IDAGetSensDky1 is one of:

IDA\_SUCCESS IDAGetQuadDky1 succeeded.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_BAD\_DKY dkyS is NULL.

IDA\_BAD\_IS The index is is not in the allowed range.

IDA\_BAD\_K k is not in the range 0, 1, ..., klast.

IDA\_BAD\_T The time t is not in the allowed range.

## 5.2.7 Optional inputs for forward sensitivity analysis

Optional input variables that control the computation of sensitivities can be changed from their default values through calls to IDASetSens\* functions. Table 5.1 lists all forward sensitivity optional input functions in IDAS which are described in detail in the remainder of this section.

#### IDASetSensParams

Call flag = IDASetSensParams(ida\_mem, p, pbar, plist);

Description The function IDASetSensParams specifies problem parameter information for sensitivity calculations.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

p (realtype \*) a pointer to the array of real problem parameters used to evaluate  $F(t, y, \dot{y}, p)$ . If non-NULL, p must point to a field in the user's data structure user\_data passed to the user's residual function. (See §5.1).

pbar (realtype \*) an array of Ns positive scaling factors. If non-NULL, pbar must have all its components > 0.0. (See §5.1).

plist (int \*) an array of Ns non-negative indices to specify which components of p to use in estimating the sensitivity equations. If non-NULL, plist must have all components  $\geq 0$ . (See §5.1).

Table 5.1: Forward sensitivity optional inputs

Optional input	Routine name	Default
Sensitivity scaling factors	IDASetSensParams	NULL
DQ approximation method	IDASetSensDQMethod	$_{\rm centered,0.0}$
Error control strategy	IDASetSensErrCon	SUNFALSE
Maximum no. of nonlinear iterations	${\tt IDASetSensMaxNonlinIters}$	3

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_ILL\_INPUT An argument has an illegal value.

Notes This function must be preceded by a call to IDASensInit.

## IDASetSensDQMethod

Call flag = IDASetSensDQMethod(ida\_mem, DQtype, DQrhomax);

Description The function IDASetSensDQMethod specifies the difference quotient strategy in the case

in which the residual of the sensitivity equations are to be computed by IDAS.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

 $\mathtt{DQtype}$  (int) specifies the difference quotient type and can be either  $\mathtt{IDA\_CENTERED}$  or

IDA\_FORWARD.

DQrhomax (realtype) positive value of the selection parameter used in deciding switching between a simultaneous or separate approximation of the two terms in the

sensitivity residual.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_ILL\_INPUT An argument has an illegal value.

Notes If DQrhomax = 0.0, then no switching is performed. The approximation is done simultaneously using either centered or forward finite differences, depending on the value of DQtype. For values of DQrhomax \geq 1.0, the simultaneous approximation is used whenever the estimated finite difference porturbations for states and parameters are within

ever the estimated finite difference perturbations for states and parameters are within a factor of DQrhomax, and the separate approximation is used otherwise. Note that a value DQrhomax < 1.0 will effectively disable switching. See §2.5 for more details.

The default value are DQtype=IDA\_CENTERED and DQrhomax= 0.0.

#### IDASetSensErrCon

Call flag = IDASetSensErrCon(ida\_mem, errconS);

Description The function IDASetSensErrCon specifies the error control strategy for sensitivity vari-

ables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

errconS (booleantype) specifies whether sensitivity variables are included (SUNTRUE)

or not (SUNFALSE) in the error control mechanism.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

Notes By default, errconS is set to SUNFALSE. If errconS=SUNTRUE then both state variables and sensitivity variables are included in the error tests. If errconS=SUNFALSE then the sensitivity variables are excluded from the error tests. Note that, in any event, all

variables are considered in the convergence tests.

#### IDASetSensMaxNonlinIters

Call flag = IDASetSensMaxNonlinIters(ida\_mem, maxcors);

Description The function IDASetSensMaxNonlinIters specifies the maximum number of nonlinear

solver iterations for sensitivity variables per step.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

maxcorS (int) maximum number of nonlinear solver iterations allowed per step (>0).

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_MEM\_FAIL The ida\_mem SUNNONLINSOL module is NULL.

Notes The default value is 3.

#### 5.2.8 Optional outputs for forward sensitivity analysis

#### 5.2.8.1 Main solver optional output functions

Optional output functions that return statistics and solver performance information related to forward sensitivity computations are listed in Table 5.2 and described in detail in the remainder of this section.

#### IDAGetSensNumResEvals

Call flag = IDAGetSensNumResEvals(ida\_mem, &nfSevals);

Description The function IDAGetSensNumResEvals returns the number of calls to the sensitivity

residual function.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nfSevals (long int) number of calls to the sensitivity residual function.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

#### IDAGetNumResEvalsSens

Call flag = IDAGetNumResEvalsSens(ida\_mem, &nfevalsS);

 $\label{prop:local_decomposition} \textbf{Description} \quad \text{The function $\tt IDAGetNumResEvalsSEns} \ \ \text{returns the number of calls to the user's residual} \\$ 

function due to the internal finite difference approximation of the sensitivity residuals.

Table 5.2: Forward sensitivity optional outputs

Optional output	Routine name
No. of calls to sensitivity residual function	IDAGetSensNumResEvals
No. of calls to residual function for sensitivity	IDAGetNumResEvalsSens
No. of sensitivity local error test failures	IDAGetSensNumErrTestFails
No. of calls to lin. solv. setup routine for sens.	IDAGetSensNumLinSolvSetups
Sensitivity-related statistics as a group	IDAGetSensStats
Error weight vector for sensitivity variables	IDAGetSensErrWeights
No. of sens. nonlinear solver iterations	IDAGetSensNumNonlinSolvIters
No. of sens. convergence failures	IDAGetSensNumNonlinSolvConvFails
Sens. nonlinear solver statistics as a group	IDAGetSensNonlinSolvStats

Arguments idamem (void \*) pointer to the IDAS memory block.

nfevalsS (long int) number of calls to the user residual function for sensitivity residuals.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

Notes This counter is incremented only if the internal finite difference approximation routines are used for the evaluation of the sensitivity residuals.

#### IDAGetSensNumErrTestFails

Call flag = IDAGetSensNumErrTestFails(ida\_mem, &nSetfails);

Description The function IDAGetSensNumErrTestFails returns the number of local error test fail-

ures for the sensitivity variables that have occurred.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nSetfails (long int) number of error test failures.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

Notes This counter is incremented only if the sensitivity variables have been included in the

error test (see IDASetSensErrCon in §5.2.7). Even in that case, this counter is not incremented if the ism=IDA\_SIMULTANEOUS sensitivity solution method has been used.

#### IDAGetSensNumLinSolvSetups

Call flag = IDAGetSensNumLinSolvSetups(ida\_mem, &nlinsetupsS);

Description The function IDAGetSensNumLinSolvSetups returns the number of calls to the linear

solver setup function due to forward sensitivity calculations.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nlinsetups (long int) number of calls to the linear solver setup function.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

Notes This counter is incremented only if a nonlinear solver requiring linear solves has been

 ${\tt used \ and \ staggered \ sensitivity \ solution \ method \ (ism=IDA\_STAGGERED) \ was \ specified \ in}$ 

the call to IDASensInit (see  $\S5.2.1$ ).

#### IDAGetSensStats

Call flag = IDAGetSensStats(ida.mem, &nfSevals, &nfevalsS, &nSetfails,

&nlinsetupsS);

Description The function IDAGetSensStats returns all of the above sensitivity-related solver statis-

tics as a group.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nfSevals (long int) number of calls to the sensitivity residual function.

nfevalsS (long int) number of calls to the user-supplied residual function.

nSetfails (long int) number of error test failures.

nlinsetupsS (long int) number of calls to the linear solver setup function.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output values have been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

#### IDAGetSensErrWeights

Call flag = IDAGetSensErrWeights(ida\_mem, eSweight);

Description The function IDAGetSensErrWeights returns the sensitivity error weight vectors at the

current time. These are the reciprocals of the  $W_i$  of (2.7) for the sensitivity variables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

eSweight (N\_Vector\_S) pointer to the array of error weight vectors.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

Notes The user must allocate memory for eweights.

#### IDAGetSensNumNonlinSolvIters

Call flag = IDAGetSensNumNonlinSolvIters(ida\_mem, &nSniters);

Description The function IDAGetSensNumNonlinSolvIters returns the number of nonlinear itera-

tions performed for sensitivity calculations.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nSniters (long int) number of nonlinear iterations performed.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_MEM\_FAIL The SUNNONLINSOL module is NULL.

Notes This counter is incremented only if ism was IDA\_STAGGERED in the call to IDASensInit

(see  $\S 5.2.1$ ).

## IDAGetSensNumNonlinSolvConvFails

Call flag = IDAGetSensNumNonlinSolvConvFails(ida\_mem, &nSncfails);

Description The function IDAGetSensNumNonlinSolvConvFails returns the number of nonlinear

convergence failures that have occurred for sensitivity calculations.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nSncfails (long int) number of nonlinear convergence failures.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

Notes This counter is incremented only if ism was IDA\_STAGGERED in the call to IDASensInit

(see  $\S 5.2.1$ ).

#### IDAGetSensNonlinSolvStats

Call flag = IDAGetSensNonlinSolvStats(ida\_mem, &nSniters, &nSncfails);

Description The function IDAGetSensNonlinSolvStats returns the sensitivity-related nonlinear

solver statistics as a group.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nSniters (long int) number of nonlinear iterations performed. nSncfails (long int) number of nonlinear convergence failures.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output values have been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_MEM\_FAIL The SUNNONLINSOL module is NULL.

#### 5.2.8.2 Initial condition calculation optional output functions

The sensitivity consistent initial conditions found by IDAS (after a successful call to IDACalcIC) can be obtained by calling the following function:

#### IDAGetSensConsistentIC

Call flag = IDAGetSensConsistentIC(ida\_mem, yyS0\_mod, ypS0\_mod);

Description The function IDAGetSensConsistentIC returns the corrected initial conditions calcu-

lated by IDACalcIC for sensitivities variables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

 $\verb|yyS0_mod| (\verb|N_Vector|*) a pointer to an array of \verb|Ns| vectors containing consistent sensitive for the sensitive formula and the sensitive formula of the sensitive fo$ 

tivity vectors.

ypSO\_mod (N\_Vector \*) a pointer to an array of Ns vectors containing consistent sensi-

tivity derivative vectors.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS IDAGetSensConsistentIC succeeded.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS The function IDASensInit has not been previously called.

IDA\_ILL\_INPUT IDASolve has been already called.

Notes If the consistent sensitivity vectors or consistent derivative vectors are not desired, pass

 ${\tt NULL}$  for the corresponding argument.

The user must allocate space for yyS0\_mod and ypS0\_mod (if not NULL).

# <u>!</u>

## 5.3 User-supplied routines for forward sensitivity analysis

In addition to the required and optional user-supplied routines described in §4.6, when using IDAS for forward sensitivity analysis, the user has the option of providing a routine that calculates the residual of the sensitivity equations (2.12).

By default, IDAS uses difference quotient approximation routines for the residual of the sensitivity equations. However, IDAS allows the option for user-defined sensitivity residual routines (which also provides a mechanism for interfacing IDAS to routines generated by automatic differentiation).

The user may provide the residuals of the sensitivity equations (2.12), for all sensitivity parameters at once, through a function of type IDASensResFn defined by:

```
IDASensResFn
Definition
              typedef int (*IDASensResFn)(int Ns, realtype t,
                                                  N_Vector yy, N_Vector yp, N_Vector resval,
                                                  N_Vector *yS, N_Vector *ypS,
                                                  N_Vector *resvalS, void *user_data,
                                                  N_Vector tmp1, N_Vector tmp2, N_Vector tmp3);
              This function computes the sensitivity residual for all sensitivity equations. It must com-
Purpose
              pute the vectors (\partial F/\partial y)s_i(t) + (\partial F/\partial \dot{y})\dot{s}_i(t) + (\partial F/\partial p_i) and store them in resvalS[i].
Arguments
                           is the number of sensitivities.
              Ns
              t
                           is the current value of the independent variable.
                           is the current value of the state vector, y(t).
              уу
                           is the current value of \dot{y}(t).
              ур
                           contains the current value F of the original DAE residual.
              resval
                           contains the current values of the sensitivities s_i.
              yS
                           contains the current values of the sensitivity derivatives \dot{s}_i.
              ypS
                           contains the output sensitivity residual vectors.
              resvalS
              user_data is a pointer to user data.
              tmp1
              tmp2
              tmp3
                           are N_{\text{-}}Vectors of length N which can be used as temporary storage.
Return value An IDASensResFn should return 0 if successful, a positive value if a recoverable error
              occurred (in which case IDAS will attempt to correct), or a negative value if it failed
              unrecoverably (in which case the integration is halted and IDA_SRES_FAIL is returned).
Notes
              There is one situation in which recovery is not possible even if IDASensResFn function
              returns a recoverable error flag. That is when this occurs at the very first call to the
```

# 5.4 Integration of quadrature equations depending on forward sensitivities

IDASensResFn, in which case IDAS returns IDA\_FIRST\_RES\_FAIL.

IDAS provides support for integration of quadrature equations that depends not only on the state variables but also on forward sensitivities.

The following is an overview of the sequence of calls in a user's main program in this situation. Steps that are unchanged from the skeleton program presented in §5.1 are grayed out. See also §4.7.

- 1. Initialize parallel or multi-threaded environment
- 2. Set problem dimensions, etc.
- 3. Set vectors of initial values
- 4. Create IDAS object
- 5. Initialize IDAS solver
- 6. Specify integration tolerances
- 7. Create matrix object
- 8. Create linear solver object
- 9. Set linear solver optional inputs

- 10. Attach linear solver module
- 11. Set optional inputs
- 12. Create nonlinear solver object
- 13. Attach nonlinear solver module
- 14. Set nonlinear solver optional inputs
- 15. Initialize sensitivity-independent quadrature problem
- 16. Define the sensitivity problem
- 17. Set sensitivity initial conditions
- 18. Activate sensitivity calculations
- 19. Set sensitivity tolerances
- 20. Set sensitivity analysis optional inputs
- 21. Create sensitivity nonlinear solver object
- 22. Attach the sensitvity nonlinear solver module
- 23. Set sensitivity nonlinear solver optional inputs

#### 24. Set vector of initial values for quadrature variables

Typically, the quadrature variables should be initialized to 0.

#### 25. Initialize sensitivity-dependent quadrature integration

Call IDAQuadSensInit to specify the quadrature equation right-hand side function and to allocate internal memory related to quadrature integration. See §5.4.1 for details.

#### 26. Set optional inputs for sensitivity-dependent quadrature integration

Call IDASetQuadSensErrCon to indicate whether or not quadrature variables should be used in the step size control mechanism. If so, one of the IDAQuadSens\*tolerances functions must be called to specify the integration tolerances for quadrature variables. See §5.4.4 for details.

27. Advance solution in time

#### 28. Extract sensitivity-dependent quadrature variables

Call IDAGetQuadSens, IDAGetQuadSens1, IDAGetQuadSensDky or IDAGetQuadSensDky1 to obtain the values of the quadrature variables or their derivatives at the current time. See §5.4.3 for details.

- 29. Get optional outputs
- 30. Extract sensitivity solution

#### 31. Get sensitivity-dependent quadrature optional outputs

Call IDAGetQuadSens\* functions to obtain optional output related to the integration of sensitivity-dependent quadratures. See §5.4.5 for details.

- 32. Deallocate memory for solutions vector
- 33. Deallocate memory for sensitivity vectors

- 34. Deallocate memory for sensitivity-dependent quadrature variables
- 35. Free solver memory
- 36. Free nonlinear solver memory
- 37. Free vector specification memory
- 38. Free linear solver and matrix memory
- 39. Finalize MPI, if used

Note: IDAQuadSensInit (step 25 above) can be called and quadrature-related optional inputs (step 26 above) can be set, anywhere between steps 16 and 27.

### 5.4.1 Sensitivity-dependent quadrature initialization and deallocation

The function IDAQuadSensInit activates integration of quadrature equations depending on sensitivities and allocates internal memory related to these calculations. If rhsQS is input as NULL, then IDAS uses an internal function that computes difference quotient approximations to the functions  $\bar{q}_i = (\partial q/\partial y)s_i + (\partial q/\partial y)\dot{s}_i + \partial q/\partial p_i$ , in the notation of (2.10). The form of the call to this function is as follows:

#### IDAQuadSensInit

Call flag = IDAQuadSensInit(ida\_mem, rhsQS, yQSO);

Description The function IDAQuadSensInit provides required problem specifications, allocates internal memory, and initializes quadrature integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

rhsQS (IDAQuadSensRhsFn) is the C function which computes  $f_{QS}$ , the right-hand side of the sensitivity-dependent quadrature equations (for full details see §5.4.6).

yQSO (N\_Vector \*) contains the initial values of sensitivity-dependent quadratures.

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAQuadSensInit was successful.

IDA\_MEM\_NULL The IDAS memory was not initialized by a prior call to IDACreate.

IDA\_MEM\_FAIL A memory allocation request failed.

IDA\_NO\_SENS The sensitivities were not initialized by a prior call to IDASensInit.

IDA\_ILL\_INPUT The parameter yQSO is NULL.

Notes Before calling IDAQuadSensInit, the user must enable the sensitivites by calling IDASensInit.

If an error occurred,  ${\tt IDAQuadSensInit}$  also sends an error message to the error handler function

In terms of the number of quadrature variables  $N_q$  and maximum method order maxord, the size of the real workspace is increased as follows:

- Base value: lenrw = lenrw + (maxord+5) $N_q$
- ullet If IDAQuadSensSVtolerances is called: lenrw = lenrw  $+N_qN_s$

and the size of the integer workspace is increased as follows:

- $\bullet \ \, {\rm Base \ value:} \ \, {\rm leniw} = {\rm leniw} + ({\rm maxord+5})N_q \\$
- ullet If IDAQuadSensSVtolerances is called: leniw = leniw  $+N_qN_s$



The function IDAQuadSensReInit, useful during the solution of a sequence of problems of same size, reinitializes the quadrature related internal memory and must follow a call to IDAQuadSensInit. The number Nq of quadratures as well as the number Ns of sensitivities are assumed to be unchanged from the prior call to IDAQuadSensInit. The call to the IDAQuadSensReInit function has the form:

#### ${\tt IDAQuadSensReInit}$

Call flag = IDAQuadSensReInit(ida\_mem, yQS0);

Description The function IDAQuadSensReInit provides required problem specifications and reini-

tializes the sensitivity-dependent quadrature integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

yQSO (N\_Vector \*) contains the initial values of sensitivity-dependent quadratures.

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAQuadSensReInit was successful.

IDA\_MEM\_NULL The IDAS memory was not initialized by a prior call to IDACreate.

prior call to IDASensInit.

IDA\_NO\_QUADSENS Memory space for the sensitivity quadratures integration was not

allocated by a prior call to IDAQuadSensInit.

IDA\_ILL\_INPUT The parameter yQSO is NULL.

Notes If an error occurred, IDAQuadSensReInit also sends an error message to the error

handler function.

#### IDAQuadSensFree

Call IDAQuadSensFree(ida\_mem):

Description The function IDAQuadSensFree frees the memory allocated for sensitivity quadrature

integration.

Arguments The argument is the pointer to the IDAS memory block (of type void \*).

Return value The function IDAQuadSensFree has no return value.

Notes In general, IDAQuadSensFree need not be called by the user as it is called automatically

by IDAFree.

#### 5.4.2 IDAS solver function

Even if quadrature integration was enabled, the call to the main solver function IDASolve is exactly the same as in §4.5.7. However, in this case the return value flag can also be one of the following:

IDA\_QSRHS\_FAIL The sensitivity quadrature right-hand side function failed in an unrecoverable

manner.

IDA\_FIRST\_QSRHS\_ERR The sensitivity quadrature right-hand side function failed at the first call.

IDA\_REP\_QSRHS\_ERR Convergence test failures occurred too many times due to repeated recover-

able errors in the quadrature right-hand side function. The IDA\_REP\_RES\_ERR will also be returned if the quadrature right-hand side function had repeated recoverable errors during the estimation of an initial step size (assuming the sensitivity quadrature variables are included in the error tests).

#### 5.4.3 Sensitivity-dependent quadrature extraction functions

If sensitivity-dependent quadratures have been initialized by a call to IDAQuadSensInit, or reinitialized by a call to IDAQuadSensReInit, then IDAS computes a solution, sensitivities, and quadratures depending on sensitivities at time t. However, IDASolve will still return only the solutions y and  $\dot{y}$ . Sensitivity-dependent quadratures can be obtained using one of the following functions:

#### IDAGetQuadSens

Call flag = IDAGetQuadSens(ida\_mem, &tret, yQS);

Description The function IDAGetQuadSens returns the quadrature sensitivity solution vectors after

a successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

tret (realtype) the time reached by the solver (output).

yQS (N\_Vector \*) array of Ns computed sensitivity-dependent quadrature vectors.

Return value The return value flag of IDAGetQuadSens is one of:

IDA\_SUCCESS IDAGetQuadSens was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_SENS Sensitivities were not activated.

IDA\_NO\_QUADSENS Quadratures depending on the sensitivities were not activated.

IDA\_BAD\_DKY
yQS or one of the yQS[i] is NULL.

The function IDAGetQuadSensDky computes the k-th derivatives of the interpolating polynomials for the sensitivity-dependent quadrature variables at time t. This function is called by IDAGetQuadSens with k = 0, but may also be called directly by the user.

#### IDAGetQuadSensDky

Call flag = IDAGetQuadSensDky(ida\_mem, t, k, dkyQS);

 $\label{thm:decomposition} \textbf{Description} \quad \textbf{The function $\mathtt{IDAGetQuadSensDky}$ returns derivatives of the quadrature sensitivities}$ 

solution vectors after a successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

t (realtype) the time at which information is requested. The time t must fall within the interval defined by the last successful step taken by IDAS.

k (int) order of the requested derivative.

 ${\tt dkyQS} \quad \hbox{($\tt N\_Vector *) array of $\tt Ns$ vectors containing the derivatives. This vector array}$ 

must be allocated by the user.

Return value The return value flag of IDAGetQuadSensDky is one of:

IDA\_SUCCESS IDAGetQuadSensDky succeeded.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_SENS Sensitivities were not activated.

IDA\_NO\_QUADSENS Quadratures depending on the sensitivities were not activated.

IDA\_BAD\_DKY dkyQS or one of the vectors dkyQS[i] is NULL.

IDA\_BAD\_K k is not in the range 0, 1, ..., klast.

IDA\_BAD\_T The time t is not in the allowed range.

Quadrature sensitivity solution vectors can also be extracted separately for each parameter in turn through the functions IDAGetQuadSens1 and IDAGetQuadSensDky1, defined as follows:

#### IDAGetQuadSens1

Call flag = IDAGetQuadSens1(ida\_mem, &tret, is, yQS);

 $\hbox{Description} \quad \hbox{The function $\tt IDAGetQuadSens1$ returns the $\tt is$-th sensitivity of quadratures after a } \\$ 

successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

tret (realtype) the time reached by the solver (output).

is (int) specifies which sensitivity vector is to be returned  $(0 \le is < N_s)$ .

yQS (N\_Vector) the computed sensitivity-dependent quadrature vector.

Return value The return value flag of IDAGetQuadSens1 is one of:

IDA\_SUCCESS IDAGetQuadSens1 was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

 ${\tt IDA\_NO\_QUADSENS}\ \ {\tt Quadratures}\ \ {\tt depending}\ \ {\tt on}\ \ {\tt the}\ \ {\tt sensitivities}\ \ {\tt were}\ \ {\tt not}\ \ {\tt activated}.$ 

IDA\_BAD\_IS The index is is not in the allowed range.

IDA\_BAD\_DKY yQS is NULL.

#### IDAGetQuadSensDky1

Call flag = IDAGetQuadSensDky1(ida\_mem, t, k, is, dkyQS);

Description The function IDAGetQuadSensDky1 returns the k-th derivative of the is-th sensitivity

solution vector after a successful return from IDASolve.

Arguments ida\_mem (void \*) pointer to the memory previously allocated by IDAInit.

t (realtype) specifies the time at which sensitivity information is requested.

The time t must fall within the interval defined by the last successful step

taken by IDAS.

k (int) order of derivative.

is (int) specifies the sensitivity derivative vector to be returned ( $0 \le is < N_s$ ).

dkyQS (N\_Vector) the vector containing the derivative. The space for dkyQS must be

allocated by the user.

Return value The return value flag of IDAGetQuadSensDky1 is one of:

IDA\_SUCCESS IDAGetQuadDky1 succeeded.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_SENS Forward sensitivity analysis was not initialized.

IDA\_NO\_QUADSENS Quadratures depending on the sensitivities were not activated.

IDA\_BAD\_DKY dkyQS is NULL.

IDA\_BAD\_IS The index is is not in the allowed range.

IDA\_BAD\_K k is not in the range 0, 1, ..., klast.

 $\label{eq:total_bad_tau} \textbf{IDA\_BAD\_T} \qquad \qquad \textbf{The time t is not in the allowed range.}$ 

#### 5.4.4 Optional inputs for sensitivity-dependent quadrature integration

IDAS provides the following optional input functions to control the integration of sensitivity-dependent quadrature equations.

#### IDASetQuadSensErrCon

Call flag = IDASetQuadSensErrCon(ida\_mem, errconQS)

Description The function IDASetQuadSensErrCon specifies whether or not the quadrature variables

are to be used in the local error control mechanism. If they are, the user must specify the error tolerances for the quadrature variables by calling IDAQuadSensSStolerances,

 ${\tt IDAQuadSensSV tolerances}, \ {\tt or} \ {\tt IDAQuadSensEE tolerances}.$ 

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

errconQS (booleantype) specifies whether sensitivity quadrature variables are included

(SUNTRUE) or not (SUNFALSE) in the error control mechanism.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Sensitivities were not activated.

IDA\_NO\_QUADSENS Quadratures depending on the sensitivities were not activated.

Notes By default, errconQS is set to SUNFALSE.

It is illegal to call IDASetQuadSensErrCon before a call to IDAQuadSensInit.

If the quadrature variables are part of the step size control mechanism, one of the following functions must be called to specify the integration tolerances for quadrature variables.

#### IDAQuadSensSStolerances

Call flag = IDAQuadSensSVtolerances(ida\_mem, reltolQS, abstolQS);

Description The function IDAQuadSensSStolerances specifies scalar relative and absolute toler-

ances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

reltolQS (realtype) is the scalar relative error tolerance.

abstolQS (realtype\*) is a pointer to an array containing the Ns scalar absolute error

tolerances.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Sensitivities were not activated.

IDA\_NO\_QUADSENS Quadratures depending on the sensitivities were not activated.

IDA\_ILL\_INPUT One of the input tolerances was negative.

#### IDAQuadSensSVtolerances

Call flag = IDAQuadSensSVtolerances(ida\_mem, reltolQS, abstolQS);

Description The function IDAQuadSensSVtolerances specifies scalar relative and vector absolute

tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

reltolQS (realtype) is the scalar relative error tolerance.

abstolQS (N\_Vector\*) is an array of Ns variables of type N\_Vector. The N\_Vector from

abstolS[is] specifies the vector tolerances for is-th quadrature sensitivity.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional value has been successfully set.

IDA\_NO\_QUAD Quadrature integration was not initialized.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Sensitivities were not activated.

IDA\_NO\_QUADSENS Quadratures depending on the sensitivities were not activated.

IDA\_ILL\_INPUT One of the input tolerances was negative.

#### IDAQuadSensEEtolerances

Call flag = IDAQuadSensEEtolerances(ida\_mem);

Description The function IDAQuadSensEEtolerances specifies that the tolerances for the sensitivity-

dependent quadratures should be estimated from those provided for the pure quadrature

variables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

Return value The return value flag (of type int) is one of:



IDA\_SUCCESS The optional value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_SENS Sensitivities were not activated.

IDA\_NO\_QUADSENS Quadratures depending on the sensitivities were not activated.

Notes

When IDAQuadSensEEtolerances is used, before calling IDASolve, integration of pure quadratures must be initialized (see 4.7.1) and tolerances for pure quadratures must be also specified (see 4.7.4).

#### 5.4.5 Optional outputs for sensitivity-dependent quadrature integration

IDAS provides the following functions that can be used to obtain solver performance information related to quadrature integration.

#### ${\tt IDAGetQuadSensNumRhsEvals}$

Call flag = IDAGetQuadSensNumRhsEvals(ida\_mem, &nrhsQSevals);

 $Description \quad The \ function \ \ IDAGetQuadSensNumRhsEvals \ returns \ the \ number \ of \ calls \ made \ to \ the$ 

user's quadrature right-hand side function.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nrhsQSevals (long int) number of calls made to the user's rhsQS function.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_QUADSENS Sensitivity-dependent quadrature integration has not been initialized.

#### IDAGetQuadSensNumErrTestFails

Call flag = IDAGetQuadSensNumErrTestFails(ida\_mem, &nQSetfails);

Description The function IDAGetQuadSensNumErrTestFails returns the number of local error test

failures due to quadrature variables.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nQSetfails (long int) number of error test failures due to quadrature variables.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

 ${\tt IDA\_NO\_QUADSENS} \ \ {\tt Sensitivity-dependent} \ \ {\tt quadrature} \ \ {\tt integration} \ \ {\tt has} \ \ {\tt not} \ \ {\tt been} \ \ {\tt initialized}.$ 

#### IDAGetQuadSensErrWeights

Call flag = IDAGetQuadSensErrWeights(ida.mem, eQSweight);

 $\label{lem:decomposition} \textbf{Description} \quad \textbf{The function IDAGetQuadSensErrWeights} \ \ \textbf{returns the quadrature error weights} \ \ \textbf{at the} \\$ 

current time.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

eQSweight (N\_Vector \*) array of quadrature error weight vectors at the current time.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_QUADSENS Sensitivity-dependent quadrature integration has not been initialized.



Notes The user must allocate memory for eQSweight.

If quadratures were not included in the error control mechanism (through a call to IDASetQuadSensErrCon with errconQS=SUNTRUE), IDAGetQuadSensErrWeights does not set the eQSweight vector.

#### IDAGetQuadSensStats

Call flag = IDAGetQuadSensStats(ida\_mem, &nrhsQSevals, &nQSetfails);

Description The function IDAGetQuadSensStats returns the IDAS integrator statistics as a group.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

nrhsQSevals (long int) number of calls to the user's rhsQS function.

nQSetfails (long int) number of error test failures due to quadrature variables.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS the optional output values have been successfully set.

the optional output values have been successionly set.

IDA\_MEM\_NULL the ida\_mem pointer is NULL.

IDA\_NO\_QUADSENS Sensitivity-dependent quadrature integration has not been initialized.

# 5.4.6 User-supplied function for sensitivity-dependent quadrature integration

For the integration of sensitivity-dependent quadrature equations, the user must provide a function that defines the right-hand side of the sensitivity quadrature equations. For sensitivities of quadratures (2.10) with integrands q, the appropriate right-hand side functions are given by  $\bar{q}_i = (\partial q/\partial y)s_i + (\partial q/\partial y)\dot{s}_i + \partial q/\partial p_i$ . This user function must be of type IDAQuadSensRhsFn, defined as follows:

#### IDAQuadSensRhsFn

tmp3

Definition	typedef i	int (*IDAQuadSensRhsFn)(int Ns, realtype t, N_Vector yy,				
Purpose		ction computes the sensitivity quadrature equation right-hand side for a given the independent variable $t$ and state vector $y$ .				
Arguments	Ns	is the number of sensitivity vectors.				
	t	is the current value of the independent variable.				
	уу	is the current value of the dependent variable vector, $y(t)$ .				
	yp is the current value of the dependent variable vector, $\dot{y}(t)$ .					
	ууЅ	is an array of Ns variables of type N_Vector containing the dependent sensitivity vectors $\boldsymbol{s}_i.$				
	урS	is an array of Ns variables of type N_Vector containing the dependent sensitivity derivatives $\dot{s}_i.$				
	rrQ	is the current value of the quadrature right-hand side $q$ .				
	rhsvalQS	contains the Ns output vectors.				
	user_data is the user_data pointer passed to IDASetUserData.					
	tmp1					
	tmp2					

are N\_Vectors which can be used as temporary storage.

Return value An IDAQuadSensRhsFn should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if it failed unrecoverably (in which case the integration is halted and IDA\_QRHS\_FAIL is returned).

Notes Allocation of memory for rhsvalQS is automatically handled within IDAS.

Both yy and yp are of type N\_Vector and both yyS and ypS are pointers to an array containing Ns vectors of type N\_Vector. It is the user's responsibility to access the vector data consistently (including the use of the correct accessor macros from each NVECTOR implementation). For the sake of computational efficiency, the vector functions in the two NVECTOR implementations provided with IDAS do not perform any consistency checks with respect to their N\_Vector arguments (see §7.2 and §7.3).

There is one situation in which recovery is not possible even if IDAQuadSensRhsFn function returns a recoverable error flag. That is when this occurs at the very first call to the IDAQuadSensRhsFn, in which case IDAS returns IDA\_FIRST\_QSRHS\_ERR).

## 5.5 Note on using partial error control

For some problems, when sensitivities are excluded from the error control test, the behavior of IDAS may appear at first glance to be erroneous. One would expect that, in such cases, the sensitivity variables would not influence in any way the step size selection.

The short explanation of this behavior is that the step size selection implemented by the error control mechanism in IDAS is based on the magnitude of the correction calculated by the nonlinear solver. As mentioned in §5.2.1, even with partial error control selected in the call to IDASensInit, the sensitivity variables are included in the convergence tests of the nonlinear solver.

When using the simultaneous corrector method (§2.5), the nonlinear system that is solved at each step involves both the state and sensitivity equations. In this case, it is easy to see how the sensitivity variables may affect the convergence rate of the nonlinear solver and therefore the step size selection. The case of the staggered corrector approach is more subtle. The sensitivity variables at a given step are computed only once the solver for the nonlinear state equations has converged. However, if the nonlinear system corresponding to the sensitivity equations has convergence problems, IDAS will attempt to improve the initial guess by reducing the step size in order to provide a better prediction of the sensitivity variables. Moreover, even if there are no convergence failures in the solution of the sensitivity system, IDAS may trigger a call to the linear solver's setup routine which typically involves reevaluation of Jacobian information (Jacobian approximation in the case of IDADENSE and IDABAND, or preconditioner data in the case of the Krylov solvers). The new Jacobian information will be used by subsequent calls to the nonlinear solver for the state equations and, in this way, potentially affect the step size selection.

When using the simultaneous corrector method it is not possible to decide whether nonlinear solver convergence failures or calls to the linear solver setup routine have been triggered by convergence problems due to the state or the sensitivity equations. When using one of the staggered corrector methods, however, these situations can be identified by carefully monitoring the diagnostic information provided through optional outputs. If there are no convergence failures in the sensitivity nonlinear solver, and none of the calls to the linear solver setup routine were made by the sensitivity nonlinear solver, then the step size selection is not affected by the sensitivity variables.

Finally, the user must be warned that the effect of appending sensitivity equations to a given system of DAEs on the step size selection (through the mechanisms described above) is problem-dependent and can therefore lead to either an increase or decrease of the total number of steps that IDAS takes to complete the simulation. At first glance, one would expect that the impact of the sensitivity variables, if any, would be in the direction of increasing the step size and therefore reducing the total number of steps. The argument for this is that the presence of the sensitivity variables in the convergence test of the nonlinear solver can only lead to additional iterations (and therefore a smaller iteration error), or to additional calls to the linear solver setup routine (and therefore more up-to-date Jacobian information), both of which will lead to larger steps being taken by IDAS. However, this is true only

locally. Overall, a larger integration step taken at a given time may lead to step size reductions at later times, due to either nonlinear solver convergence failures or error test failures.

# Chapter 6

# Using IDAS for Adjoint Sensitivity Analysis

This chapter describes the use of IDAS to compute sensitivities of derived functions using adjoint sensitivity analysis. As mentioned before, the adjoint sensitivity module of IDAS provides the infrastructure for integrating backward in time any system of DAEs that depends on the solution of the original IVP, by providing various interfaces to the main IDAS integrator, as well as several supporting user-callable functions. For this reason, in the following sections we refer to the *backward problem* and not to the *adjoint problem* when discussing details relevant to the DAEs that are integrated backward in time. The backward problem can be the adjoint problem (2.20) or (2.25), and can be augmented with some quadrature differential equations.

IDAS uses various constants for both input and output. These are defined as needed in this chapter, but for convenience are also listed separately in Appendix B.

We begin with a brief overview, in the form of a skeleton user program. Following that are detailed descriptions of the interface to the various user-callable functions and of the user-supplied functions that were not already described in Chapter 4.

# 6.1 A skeleton of the user's main program

The following is a skeleton of the user's main program as an application of IDAS. The user program is to have these steps in the order indicated, unless otherwise noted. For the sake of brevity, we defer many of the details to the later sections. As in §4.4, most steps are independent of the NVECTOR, SUNMATRIX, SUNLINSOL, and SUNNONLINSOL implementations used. For the steps that are not, refer to Chapters 7, 8, 9, and 10 for the specific name of the function to be called or macro to be referenced.

Steps that are unchanged from the skeleton programs presented in §4.4, §5.1, and §5.4, are grayed out

1. Include necessary header files

The idas.h header file also defines additional types, constants, and function prototypes for the adjoint sensitivity module user-callable functions. In addition, the main program should include an NVECTOR implementation header file (for the particular implementation used) and, if a nonlinear solver requiring a linear solver (e.g., the default Newton iteration) will be used, the header file of the desired linear solver module.

2. Initialize parallel or multi-threaded environment

#### Forward problem

3. Set problem dimensions etc. for the forward problem

- 4. Set initial conditions for the forward problem
- 5. Create IDAS object for the forward problem
- 6. Initialize IDAS solver for the forward problem
- 7. Specify integration tolerances for forward problem
- 8. Set optional inputs for the forward problem
- 9. Create matrix object for the forward problem
- 10. Create linear solver object for the forward problem
- 11. Set linear solver optional inputs for the forward problem
- 12. Attach linear solver module for the forward problem
- 13. Create nonlinear solver module for the forward problem
- 14. Attach nonlinear solver module for the forward problem
- 15. Set nonlinear solver optional inputs for the forward problem
- 16. Initialize quadrature problem or problems for forward problems, using IDAQuadInit and/or IDAQuadSensInit.
- 17. Initialize forward sensitivity problem
- 18. Specify rootfinding

#### 19. Allocate space for the adjoint computation

Call IDAAdjInit() to allocate memory for the combined forward-backward problem (see §6.2.1 for details). This call requires Nd, the number of steps between two consecutive checkpoints. IDAAdjInit also specifies the type of interpolation used (see §2.6.3).

#### 20. Integrate forward problem

Call IDASolveF, a wrapper for the IDAS main integration function IDASolve, either in IDA\_NORMAL mode to the time tout or in IDA\_ONE\_STEP mode inside a loop (if intermediate solutions of the forward problem are desired (see  $\S 6.2.3$ )). The final value of tret is then the maximum allowable value for the endpoint T of the backward problem.

#### Backward problem(s)

#### 21. Set problem dimensions etc. for the backward problem

This generally includes NB, the number of variables in the backward problem and possibly the local vector length NBlocal.

#### 22. Set initial values for the backward problem

Set the endpoint time  $\mathtt{tB0} = T$ , and set the corresponding vectors  $\mathtt{yB0}$  and  $\mathtt{ypB0}$  at which the backward problem starts.

#### 23. Create the backward problem

Call IDACreateB, a wrapper for IDACreate, to create the IDAS memory block for the new backward problem. Unlike IDACreate, the function IDACreateB does not return a pointer to the newly created memory block (see §6.2.4). Instead, this pointer is attached to the internal adjoint memory block (created by IDAAdjInit) and returns an identifier called which that the user must later specify in any actions on the newly created backward problem.

#### 24. Allocate memory for the backward problem

Call IDAInitB (or IDAInitBS, when the backward problem depends on the forward sensitivities). The two functions are actually wrappers for IDAInit and allocate internal memory, specify problem data, and initialize IDAS at tBO for the backward problem (see §6.2.4).

#### 25. Specify integration tolerances for backward problem

Call IDASStolerancesB(...) or IDASVtolerancesB(...) to specify a scalar relative tolerance and scalar absolute tolerance, or a scalar relative tolerance and a vector of absolute tolerances, respectively. The functions are wrappers for IDASStolerances(...) and IDASVtolerances(...) but they require an extra argument which, the identifier of the backward problem returned by IDACreateB. See §6.2.5 for more information.

#### 26. Set optional inputs for the backward problem

Call IDASet\*B functions to change from their default values any optional inputs that control the behavior of IDAS. Unlike their counterparts for the forward problem, these functions take an extra argument which, the identifier of the backward problem returned by IDACreateB (see §6.2.9).

#### 27. Create matrix object for the backward problem

If a nonlinear solver requiring a linear solve will be used (e.g., the the default Newton iteration) and the linear solver will be a direct linear solver, then a template Jacobian matrix must be created by calling the appropriate constructor function defined by the particular SUNMATRIX implementation.

NOTE: The dense, banded, and sparse matrix objects are usable only in a serial or threaded environment.

Note also that it is not required to use the same matrix type for both the forward and the backward problems.

#### 28. Create linear solver object for the backward problem

If a nonlinear solver requiring a linear solver is chosen (e.g., the default Newton iteration), then the desired linear solver object for the backward problem must be created by calling the appropriate constructor function defined by the particular SUNLINSOL implementation.

Note that it is not required to use the same linear solver module for both the forward and the backward problems; for example, the forward problem could be solved with the SUNLINSOL\_DENSE linear solver module and the backward problem with SUNLINSOL\_SPGMR linear solver module.

#### 29. Set linear solver interface optional inputs for the backward problem

Call IDASet\*B functions to change optional inputs specific to the linear solver interface. See §6.2.9 for details.

#### 30. Attach linear solver module for the backward problem

If a nonlinear solver requiring a linear solver is chosen for the backward problem (e.g., the default Newton iteration), then initialize the IDALS linear solver interface by attaching the linear solver object (and matrix object, if applicable) with the following call (for details see §4.5.3):

```
ier = IDASetLinearSolverB(...);
```

#### 31. Create nonlinear solver object for the backward problem (optional)

If using a non-default nonlinear solver for the backward problem, then create the desired nonlinear solver object by calling the appropriate constructor function defined by the particular SUNNON-LINSOL implementation e.g., NLSB = SUNNonlinSol\_\*\*\*(...); where \*\*\* is the name of the nonlinear solver (see Chapter 10 for details).

#### 32. Attach nonlinear solver module for the backward problem (optional)

If using a non-default nonlinear solver for the backward problem, then initialize the nonlinear solver interface by attaching the nonlinear solver object by calling

ier = IDASetNonlinearSolverB(idaode\_mem, NLSB); (see §4.5.4 for details).

#### 33. Initialize quadrature calculation

If additional quadrature equations must be evaluated, call <code>IDAQuadInitB</code> or <code>IDAQuadInitBS</code> (if quadrature depends also on the forward sensitivities) as shown in §6.2.11.1. These functions are wrappers around <code>IDAQuadInit</code> and can be used to initialize and allocate memory for quadrature integration. Optionally, call <code>IDASetQuad\*B</code> functions to change from their default values optional inputs that control the integration of quadratures during the backward phase.

#### 34. Integrate backward problem

Call IDASolveB, a second wrapper around the IDAS main integration function IDASolve, to integrate the backward problem from tB0 (see  $\S6.2.8$ ). This function can be called either in IDA\_NORMAL or IDA\_ONE\_STEP mode. Typically, IDASolveB will be called in IDA\_NORMAL mode with an end time equal to the initial time  $t_0$  of the forward problem.

#### 35. Extract quadrature variables

If applicable, call IDAGetQuadB, a wrapper around IDAGetQuad, to extract the values of the quadrature variables at the time returned by the last call to IDASolveB. See §6.2.11.2.

#### 36. Deallocate memory

Upon completion of the backward integration, call all necessary deallocation functions. These include appropriate destructors for the vectors y and yB, a call to IDAFree to free the IDAS memory block for the forward problem. If one or more additional adjoint sensitivity analyses are to be done for this problem, a call to IDAAdjFree (see §6.2.1) may be made to free and deallocate the memory allocated for the backward problems, followed by a call to IDAAdjInit.

- 37. Free the nonlinear solver memory for the forward and backward problems
- 38. Free linear solver and matrix memory for the forward and backward problems
- 39. Finalize MPI, if used

The above user interface to the adjoint sensitivity module in IDAS was motivated by the desire to keep it as close as possible in look and feel to the one for DAE IVP integration. Note that if steps (21)-(35) are not present, a program with the above structure will have the same functionality as one described in §4.4 for integration of DAEs, albeit with some overhead due to the checkpointing scheme.

If there are multiple backward problems associated with the same forward problem, repeat steps (21)-(35) above for each successive backward problem. In the process, each call to IDACreateB creates a new value of the identifier which.

## 6.2 User-callable functions for adjoint sensitivity analysis

#### 6.2.1 Adjoint sensitivity allocation and deallocation functions

After the setup phase for the forward problem, but before the call to <code>IDASolveF</code>, memory for the combined forward-backward problem must be allocated by a call to the function <code>IDAAdjInit</code>. The form of the call to this function is

```
IDAAdjInit
```

Description The function IDAAdjInit updates IDAS memory block by allocating the internal memory

needed for backward integration. Space is allocated for the  $Nd = N_d$  interpolation data

points, and a linked list of checkpoints is initialized.

Arguments ida\_mem (void \*) is the pointer to the IDAS memory block returned by a previous

call to IDACreate.

Nd (long int) is the number of integration steps between two consecutive

checkpoints.

interpType (int) specifies the type of interpolation used and can be IDA\_POLYNOMIAL

or IDA\_HERMITE, indicating variable-degree polynomial and cubic Hermite

interpolation, respectively (see §2.6.3).

Return value The return value flag (of type int) is one of:

 ${\tt IDA\_SUCCESS} \quad {\tt IDAAdjInit} \ {\rm was} \ {\rm successful}.$ 

IDA\_MEM\_FAIL A memory allocation request has failed.

IDA\_MEM\_NULL ida\_mem was NULL.

 ${\tt IDA\_ILL\_INPUT} \ \ {\tt One} \ \ {\tt of} \ \ {\tt the} \ \ {\tt parameters} \ \ {\tt was} \ \ {\tt invalid:} \ \ {\tt Nd} \ \ {\tt was} \ \ {\tt not} \ \ {\tt positive} \ \ {\tt or} \ \ {\tt interpType}$ 

is not one of the IDA\_POLYNOMIAL or IDA\_HERMITE.

Notes The user must set Nd so that all data needed for interpolation of the forward problem solution between two checkpoints fits in memory. IDAAdjInit attempts to allocate

space for (2Nd+3) variables of type N\_Vector.

If an error occurred, IDAAdjInit also sends a message to the error handler function.

#### IDAAdjReInit

Call flag = IDAAdjReInit(ida\_mem);

Description The function IDAAdjReInit reinitializes the IDAS memory block for ASA, assuming

that the number of steps between check points and the type of interpolation remain

unchanged.

Arguments ida\_mem (void \*) is the pointer to the IDAS memory block returned by a previous call

to IDACreate.

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS IDAAdjReInit was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit was not previously called.

Notes The list of check points (and associated memory) is deleted.

The list of backward problems is kept. However, new backward problems can be added to this list by calling IDACreateB. If a new list of backward problems is also needed, then free the adjoint memory (by calling IDAAdjFree) and reinitialize ASA with IDAAdjInit.

The IDAS memory for the forward and backward problems can be reinitialized separately by calling IDAReInit and IDAReInitB, respectively.

#### IDAAdjFree

Call IDAAdjFree(ida\_mem);

Description The function IDAAdjFree frees the memory related to backward integration allocated

by a previous call to IDAAdjInit.

Arguments The only argument is the IDAS memory block pointer returned by a previous call to

 ${\tt IDACreate}.$ 

Return value The function IDAAdjFree has no return value.

Notes

This function frees all memory allocated by IDAAdjInit. This includes workspace memory, the linked list of checkpoints, memory for the interpolation data, as well as the IDAS memory for the backward integration phase.

Unless one or more further calls to IDAAdjInit are to be made, IDAAdjFree should not be called by the user, as it is invoked automatically by IDAFree.

#### 6.2.2 Adjoint sensitivity optional input

At any time during the integration of the forward problem, the user can disable the checkpointing of the forward sensitivities by calling the following function:

#### IDAAdjSetNoSensi

Call flag = IDAAdjSetNoSensi(ida\_mem);

Description The function IDAAdjSetNoSensi instructs IDASolveF not to save checkpointing data

for forward sensitivities any more.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

Return value The return flag (of type int) is one of:

IDA\_SUCCESS The call to IDACreateB was successful.

IDA MEM NULL The ida mem was NULL.

The function IDAAdjInit has not been previously called. IDA\_NO\_ADJ

#### 6.2.3 Forward integration function

The function IDASolve is very similar to the IDAS function IDASolve (see §4.5.7) in that it integrates the solution of the forward problem and returns the solution  $(y, \dot{y})$ . At the same time, however, IDASolveF stores checkpoint data every Nd integration steps. IDASolveF can be called repeatedly by the user. Note that IDASolveF is used only for the forward integration pass within an Adjoint Sensitivity Analysis. It is not for use in Forward Sensitivity Analysis; for that, see Chapter 5. The call to this function has the form

#### IDASolveF

Call flag = IDASolveF(ida\_mem, tout, &tret, yret, ypret, itask, &ncheck);

Description The function IDASolveF integrates the forward problem over an interval in t and saves

checkpointing data.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

> tout (realtype) the next time at which a computed solution is desired.

(realtype) the time reached by the solver (output). tret

( $N_{\text{-}}$ Vector) the computed solution vector y. yret

(N\_Vector) the computed solution vector  $\dot{y}$ . ypret

(int) a flag indicating the job of the solver for the next step. The IDA\_NORMAL itask task is to have the solver take internal steps until it has reached or just passed the user-specified tout parameter. The solver then interpolates in order to return an approximate value of y(tout) and  $\dot{y}(tout)$ . The IDA\_ONE\_STEP option tells the solver to take just one internal step and return the solution at the point reached by that step.

ncheck (int) the number of (internal) checkpoints stored so far.

Return value On return, IDASolveF returns vectors yret, ypret and a corresponding independent variable value t = tret, such that yret is the computed value of y(t) and ypret the value of  $\dot{y}(t)$ . Additionally, it returns in ncheck the number of internal checkpoints saved; the total number of checkpoint intervals is ncheck+1. The return value flag (of type int) will be one of the following. For more details see §4.5.7.

IDA\_SUCCESS IDASolveF succeeded.

IDA\_TSTOP\_RETURN IDASolveF succeeded by reaching the optional stopping point.

IDA\_ROOT\_RETURN IDASolveF succeeded and found one or more roots. In this case,

tret is the location of the root. If nrtfn > 1, call IDAGetRootInfo

to see which  $g_i$  were found to have a root.

IDA\_NO\_MALLOC The function IDAInit has not been previously called.

IDA\_ILL\_INPUT One of the inputs to IDASolveF is illegal.

IDA\_TOO\_MUCH\_WORK The solver took mxstep internal steps but could not reach tout.

IDA\_TOO\_MUCH\_ACC The solver could not satisfy the accuracy demanded by the user for

some internal step.

time step or occurred with  $|h| = h_{min}$ .

IDA\_CONV\_FAILURE Convergence test failures occurred too many times during one in-

ternal time step or occurred with  $|h| = h_{min}$ .

IDA\_LSETUP\_FAIL The linear solver's setup function failed in an unrecoverable man-

ner.

IDA\_LSOLVE\_FAIL The linear solver's solve function failed in an unrecoverable manner.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_MEM\_FAIL A memory allocation request has failed (in an attempt to allocate

space for a new checkpoint).

Notes All failure return values are negative and therefore a test flag< 0 will trap all IDASolveF failures.

At this time, IDASolveF stores checkpoint information in memory only. Future versions will provide for a safeguard option of dumping checkpoint data into a temporary file as needed. The data stored at each checkpoint is basically a snapshot of the IDAS internal memory block and contains enough information to restart the integration from that time and to proceed with the same step size and method order sequence as during the forward integration.

In addition, IDASolveF also stores interpolation data between consecutive checkpoints so that, at the end of this first forward integration phase, interpolation information is already available from the last checkpoint forward. In particular, if no checkpoints were necessary, there is no need for the second forward integration phase.

It is illegal to change the integration tolerances between consecutive calls to IDASolveF, as this information is not captured in the checkpoint data.

#### 6.2.4 Backward problem initialization functions

The functions IDACreateB and IDAInitB (or IDAInitBS) must be called in the order listed. They instantiate an IDAS solver object, provide problem and solution specifications, and allocate internal memory for the backward problem.

#### IDACreateB

Call flag = IDACreateB(ida\_mem, &which);

Description The function IDACreateB instantiates an IDAS solver object for the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

which (int) contains the identifier assigned by IDAS for the newly created backward problem. Any call to IDA\*B functions requires such an identifier.



Return value The return flag (of type int) is one of:

IDA\_SUCCESS The call to IDACreateB was successful.

IDA\_MEM\_NULL The ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_MEM\_FAIL A memory allocation request has failed.

There are two initialization functions for the backward problem – one for the case when the backward problem does not depend on the forward sensitivities, and one for the case when it does. These two functions are described next.

The function IDAInitB initializes the backward problem when it does not depend on the forward sensitivities. It is essentially wrapper for IDAInit with some particularization for backward integration, as described below.

#### IDAInitB

Call flag = IDAInitB(ida\_mem, which, resB, tB0, yB0, ypB0);

Description The function IDAInitB provides problem specification, allocates internal memory, and initializes the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

which (int) represents the identifier of the backward problem.

resB (IDAResFnB) is the C function which computes fB, the residual of the backward DAE problem. This function has the form resB(t, y, yp, yB, ypB, resvalB, user\_dataB) (for full details see  $\S6.3.1$ ).

tB0 (realtype) specifies the endpoint T where final conditions are provided for the backward problem, normally equal to the endpoint of the forward integration.

yB0 (N\_Vector) is the initial value (at t = tB0) of the backward solution.

ypB0 (N\_Vector) is the initial derivative value (at t = tB0) of the backward solution.

Return value The return flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAInitB was successful.

IDA\_NO\_MALLOC The function IDAInit has not been previously called.

IDA\_MEM\_NULL The ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_BAD\_TBO The final time tBO was outside the interval over which the forward problem was solved.

IDA\_ILL\_INPUT The parameter which represented an invalid identifier, or one of yBO, ypBO, resB was NULL.

Notes The memory allocated by IDAInitB is deallocated by the function IDAAdjFree.

For the case when backward problem also depends on the forward sensitivities, user must call IDAInitBS instead of IDAInitB. Only the third argument of each function differs between these functions.

#### IDAInitBS

Call flag = IDAInitBS(ida\_mem, which, resBS, tB0, yB0, ypB0);

Description The function IDAInitBS provides problem specification, allocates internal memory, and initializes the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

which (int) represents the identifier of the backward problem.

resBS (IDAResFnBS) is the C function which computes fB, the residual or the backward DAE problem. This function has the form resBS(t, y, yp, yS, yps, yps, yps, resvalb, user\_datab) (for full details see  $\S 6.3.2$ ).

tB0 (realtype) specifies the endpoint T where final conditions are provided for the backward problem.

yBO (N\_Vector) is the initial value (at t = tBO) of the backward solution.

 $\mathtt{ypB0}$  (N\_Vector) is the initial derivative value (at  $t = \mathtt{tB0}$ ) of the backward solution.

Return value The return flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAInitB was successful.

IDA\_NO\_MALLOC The function IDAInit has not been previously called.

IDA\_MEM\_NULL The ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_BAD\_TBO The final time tBO was outside the interval over which the forward problem was solved.

IDA\_ILL\_INPUT The parameter which represented an invalid identifier, or one of yBO, ypBO, resB was NULL, or sensitivities were not active during the forward integration.

Notes The memory allocated by IDAInitBS is deallocated by the function IDAAdjFree.

The function IDAReInitB reinitializes IDAS for the solution of a series of backward problems, each identified by a value of the parameter which. IDAReInitB is essentially a wrapper for IDAReInit, and so all details given for IDAReInit in §4.5.11 apply here. Also, IDAReInitB can be called to reinitialize a backward problem even if it has been initialized with the sensitivity-dependent version IDAInitBS. Before calling IDAReInitB for a new backward problem, call any desired solution extraction functions IDAGet\*\* associated with the previous backward problem. The call to the IDAReInitB function has the form

#### IDAReInitB

Call flag = IDAReInitB(ida\_mem, which, tB0, yB0, ypB0)

Description The function IDAReInitB reinitializes an IDAS backward problem.

Arguments ida\_mem (void \*) pointer to IDAS memory block returned by IDACreate.

which (int) represents the identifier of the backward problem.

tBO (realtype) specifies the endpoint T where final conditions are provided for the backward problem.

yB0 (N\_Vector) is the initial value (at t = tB0) of the backward solution.

ypB0 (N\_Vector) is the initial derivative value (at t = tB0) of the backward solution.

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAReInitB was successful.

IDA\_NO\_MALLOC The function IDAInit has not been previously called.

IDA\_MEM\_NULL The ida\_mem memory block pointer was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_BAD\_TB0 The final time tB0 is outside the interval over which the forward problem was solved.

IDA\_ILL\_INPUT The parameter which represented an invalid identifier, or one of yBO, ypBO was NULL.

#### 6.2.5 Tolerance specification functions for backward problem

One of the following two functions must be called to specify the integration tolerances for the backward problem. Note that this call must be made after the call to IDAInitB or IDAInitBS.

#### IDASStolerancesB

Call flag = IDASStolerances(ida\_mem, which, reltolB, abstolB);

Description The function IDASStolerancesB specifies scalar relative and absolute tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

which (int) represents the identifier of the backward problem.

reltolB (realtype) is the scalar relative error tolerance.

abstolB (realtype) is the scalar absolute error tolerance.

Return value The return flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASStolerancesB was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_NO\_MALLOC The allocation function IDAInit has not been called.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_ILL\_INPUT One of the input tolerances was negative.

#### IDASVtolerancesB

Call flag = IDASVtolerancesB(ida\_mem, which, reltolB, abstolB);

Description The function IDASVtolerancesB specifies scalar relative tolerance and vector absolute

tolerances.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

which (int) represents the identifier of the backward problem.

reltol (realtype) is the scalar relative error tolerance.

abstol (N\_Vector) is the vector of absolute error tolerances.

Return value The return flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDASVtolerancesB was successful.

IDA\_MEM\_NULL The IDAS memory block was not initialized through a previous call to

IDACreate.

IDA\_NO\_MALLOC The allocation function IDAInit has not been called.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

 ${\tt IDA\_ILL\_INPUT \ The \ relative \ error \ tolerance \ was \ negative \ or \ the \ absolute \ tolerance \ had}$ 

a negative component.

Notes This choice of tolerances is important when the absolute error tolerance needs to be

different for each component of the DAE state vector y.

#### 6.2.6 Linear solver initialization functions for backward problem

All IDAS linear solver modules available for forward problems are available for the backward problem. They should be created as for the forward problem then attached to the memory structure for the backward problem using the following function.

#### IDASetLinearSolverB

Call flag = IDASetLinearSolverB(ida\_mem, which, LS, A);

Description The function IDASetLinearSolverB attaches a generic SUNLINSOL object LS and cor-

responding template Jacobian SUNMATRIX object A (if applicable) to IDAS, initializing

the IDALS linear solver interface for solution of the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) represents the identifier of the backward problem returned by IDACreateB.

LS (SUNLinearSolver) SUNLINSOL object to use for solving linear systems for the backward problem.

A (SUNMatrix) SUNMATRIX object for used as a template for the Jacobian for the backward problem (or NULL if not applicable).

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The IDALS initialization was successful.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_ILL\_INPUT The IDALS interface is not compatible with the LS or A input objects or is incompatible with the current NVECTOR module.

IDALS\_MEM\_FAIL A memory allocation request failed.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes

If LS is a matrix-based linear solver, then the template Jacobian matrix A will be used in the solve process, so if additional storage is required within the SUNMATRIX object (e.g. for factorization of a banded matrix), ensure that the input object is allocated with sufficient size (see the documentation of the particular SUNMATRIX type in Chapter 8 for further information).

The previous routines IDADlsSetLinearSolverB and IDASpilsSetLinearSolverB are now wrappers for this routine, and may still be used for backward-compatibility. However, these will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

### 6.2.7 Initial condition calculation functions for backward problem

IDAS provides support for calculation of consistent initial conditions for certain backward index-one problems of semi-implicit form through the functions IDACalcICB and IDACalcICBS. Calling them is optional. It is only necessary when the initial conditions do not satisfy the adjoint system.

The above functions provide the same functionality for backward problems as IDACalcIC with parameter icopt = IDA\_YA\_YDP\_INIT provides for forward problems (see §4.5.5): compute the algebraic components of yB and differential components of yB, given the differential components of yB. They require that the IDASetIdB was previously called to specify the differential and algebraic components.

Both functions require forward solutions at the final time tBO. IDACalcICBS also needs forward sensitivities at the final time tBO.

#### IDACalcICB

Call flag = IDACalcICB(ida\_mem, which, tBout1, N\_Vector yfin, N\_Vector ypfin);

Description The function IDACalcICB corrects the initial values yB0 and ypB0 at time tB0 for the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) is the identifier of the backward problem.

tBout1 (realtype) is the first value of t at which a solution will be requested (from IDASolveB). This value is needed here only to determine the direction of integration and rough scale in the independent variable t.

yfin (N\_Vector) the forward solution at the final time tBO.

ypfin (N\_Vector) the forward solution derivative at the final time tB0.

Return value The return value flag (of type int) can be any that is returned by IDACalcIC (see §4.5.5). However IDACalcICB can also return one of the following:

IDA\_NO\_ADJ IDAAdjInit has not been previously called.

IDA\_ILL\_INPUT Parameter which represented an invalid identifier.

Notes

All failure return values are negative and therefore a test flag < 0 will trap all IDACalcICB failures.

Note that IDACalcICB will correct the values of  $yB(tB_0)$  and  $\dot{y}B(tB_0)$  which were specified in the previous call to IDAInitB or IDAReInitB. To obtain the corrected values, call IDAGetconsistentICB (see §6.2.10.2).

In the case where the backward problem also depends on the forward sensitivities, user must call the following function to correct the initial conditions:

#### IDACalcICBS

Call flag = IDACalcICBS(ida\_mem, which, tBout1, N\_Vector yfin, N\_Vector ypfin, N\_Vector ypfin, N\_Vector ypfin);

Description The function IDACalcICBS corrects the initial values yB0 and ypB0 at time tB0 for the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) is the identifier of the backward problem.

tBout1 (realtype) is the first value of t at which a solution will be requested (from IDASolveB). This value is needed here only to determine the direction of integration and rough scale in the independent variable t.

yfin (N\_Vector) the forward solution at the final time tBO.

ypfin (N\_Vector) the forward solution derivative at the final time tB0.

ySfin (N\_Vector \*) a pointer to an array of Ns vectors containing the sensitivities of the forward solution at the final time tBO.

ypSfin (N\_Vector \*) a pointer to an array of Ns vectors containing the derivatives of
the forward solution sensitivities at the final time tBO.

Return value The return value flag (of type int) can be any that is returned by IDACalcIC (see §4.5.5). However IDACalcICBS can also return one of the following:

IDA\_NO\_ADJ IDAAdjInit has not been previously called.

IDA\_ILL\_INPUT Parameter which represented an invalid identifier, sensitivities were not active during forward integration, or IDAInitBS (or IDAReInitBS) has not been previously called.

Notes

All failure return values are negative and therefore a test flag < 0 will trap all IDACalcICBS failures.

Note that IDACalcICBS will correct the values of  $yB(tB_0)$  and  $\dot{y}B(tB_0)$  which were specified in the previous call to IDAInitBS or IDAReInitBS. To obtain the corrected values, call IDAGetConsistentICB (see §6.2.10.2).

#### 6.2.8 Backward integration function

The function IDASolveB performs the integration of the backward problem. It is essentially a wrapper for the IDAS main integration function IDASolve and, in the case in which checkpoints were needed, it evolves the solution of the backward problem through a sequence of forward-backward integration pairs between consecutive checkpoints. In each pair, the first run integrates the original IVP forward in time and stores interpolation data; the second run integrates the backward problem backward in time and performs the required interpolation to provide the solution of the IVP to the backward problem.

The function IDASolveB does not return the solution yB itself. To obtain that, call the function IDAGetB, which is also described below.

The IDASolveB function does not support rootfinding, unlike IDASoveF, which supports the finding of roots of functions of  $(t, y, \dot{y})$ . If rootfinding was performed by IDASolveF, then for the sake of efficiency, it should be disabled for IDASolveB by first calling IDARootInit with nrtfn = 0.

The call to IDASolveB has the form

IDASolveB

Call flag = IDASolveB(ida\_mem, tBout, itaskB);

Description The function IDASolveB integrates the backward DAE problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory returned by IDACreate.

tBout (realtype) the next time at which a computed solution is desired.

itaskB (int) a flag indicating the job of the solver for the next step. The IDA\_NORMAL task is to have the solver take internal steps until it has reached or just passed the user-specified value tBout. The solver then interpolates in order to return an approximate value of yB(tBout). The IDA\_ONE\_STEP option tells the solver to take just one internal step in the direction of tBout and return.

Return value The return value flag (of type int) will be one of the following. For more details see §4.5.7.

IDA\_SUCCESS IDASolveB succeeded.

IDA\_MEM\_NULL The ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_NO\_BCK No backward problem has been added to the list of backward prob-

lems by a call to IDACreateB

IDA\_NO\_FWD The function IDASolveF has not been previously called.

IDA\_ILL\_INPUT One of the inputs to IDASolveB is illegal.

IDA\_BAD\_ITASK The itaskB argument has an illegal value.

 ${\tt IDA\_TOO\_MUCH\_WORK} \ \ {\tt The \ solver \ took \ mxstep \ internal \ steps \ but \ could \ not \ reach \ tBout.}$ 

IDA\_TOO\_MUCH\_ACC The solver could not satisfy the accuracy demanded by the user for

some internal step.

IDA\_ERR\_FAILURE Error test failures occurred too many times during one internal

time step.

IDA\_CONV\_FAILURE Convergence test failures occurred too many times during one in-

ternal time step.

IDA\_LSETUP\_FAIL The linear solver's setup function failed in an unrecoverable man-

ner.

IDA\_SOLVE\_FAIL The linear solver's solve function failed in an unrecoverable manner.

IDA\_BCKMEM\_NULL The IDAS memory for the backward problem was not created with

a call to IDACreateB.

IDA\_BAD\_TBOUT The desired output time tBout is outside the interval over which

the forward problem was solved.

IDA\_REIFWD\_FAIL Reinitialization of the forward problem failed at the first checkpoint

(corresponding to the initial time of the forward problem).

IDA\_FWD\_FAIL An error occurred during the integration of the forward problem.

All failure return values are negative and therefore a test flag< 0 will trap all IDASolveB failures.

In the case of multiple checkpoints and multiple backward problems, a given call to IDASolveB in IDA\_ONE\_STEP mode may not advance every problem one step, depending on the relative locations of the current times reached. But repeated calls will eventually advance all problems to tBout.

To obtain the solution yB to the backward problem, call the function IDAGetB as follows:

#### IDAGetB

Notes

Call flag = IDAGetB(ida\_mem, which, &tret, yB, ypB);

Description The function IDAGetB provides the solution yB of the backward DAE problem.

```
ida_mem (void *) pointer to the IDAS memory returned by IDACreate.
Arguments
             which
                      (int) the identifier of the backward problem.
             tret
                      (realtype) the time reached by the solver (output).
                      (N_Vector) the backward solution at time tret.
             vΒ
             урВ
                      (N_Vector) the backward solution derivative at time tret.
Return value The return value flag (of type int) will be one of the following.
             IDA_SUCCESS
                             IDAGetB was successful.
             IDA_MEM_NULL
                            ida_mem is NULL.
             IDA NO ADJ
                             The function IDAAdjInit has not been previously called.
             IDA_ILL_INPUT The parameter which is an invalid identifier.
Notes
              The user must allocate space for yB and ypB.
```



#### 6.2.9 Optional input functions for the backward problem

#### 6.2.9.1 Main solver optional input functions

The adjoint module in IDAS provides wrappers for most of the optional input functions defined in §4.5.8.1. The only difference is that the user must specify the identifier which of the backward problem within the list managed by IDAS.

The optional input functions defined for the backward problem are:

```
flag = IDASetNonlinearSolverB(ida_mem, which, NLSB);
flag = IDASetUserDataB(ida_mem, which, user_dataB);
flag = IDASetMaxOrdB(ida_mem, which, maxordB);
flag = IDASetMaxNumStepsB(ida_mem, which, mxstepsB);
flag = IDASetInitStepB(ida_mem, which, hinB)
flag = IDASetMaxStepB(ida_mem, which, hmaxB);
flag = IDASetSuppressAlgB(ida_mem, which, suppressalgB);
flag = IDASetIdB(ida_mem, which, idB);
flag = IDASetConstraintsB(ida_mem, which, constraintsB);
```

Their return value flag (of type int) can have any of the return values of their counterparts, but it can also be IDA\_NO\_ADJ if IDAAdjInit has not been called, or IDA\_ILL\_INPUT if which was an invalid identifier.

#### 6.2.9.2 Linear solver interface optional input functions

When using matrix-based linear solver modules for the backward problem, i.e., a non-NULL SUNMATRIX object A was passed to IDASetLinearSolverB, the IDALS linear solver interface needs a function to compute an approximation to the Jacobian matrix. This can be attached through a call to either IDASetJacFnB or IDAIDASetJacFnBS, with the second used when the backward problem depends on the forward sensitivities.

```
Call flag = IDASetJacFnB(ida_mem, which, jacB);

Description The function IDASetJacFnB specifies the Jacobian approximation function to be used for the backward problem.

Arguments ida_mem (void *) pointer to the IDAS memory block.

which (int) represents the identifier of the backward problem.

jacB (IDALsJacFnB) user-defined Jacobian approximation function.

Return value The return value flag (of type int) is one of IDALS_SUCCESS IDASetJacFnB succeeded.
```

IDALS\_MEM\_NULL The ida\_mem was NULL.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDALS\_LMEM\_NULL The linear solver has not been initialized with a call to IDASetLinearSolverB.

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes

The function type IDALsJacFnB is described in §6.3.5.

The previous routine IDADlsSetJacFnB is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDASetJacFnBS

Call flag = IDASetJacFnBS(ida\_mem, which, jacBS);

Description The function IDASetJacFnBS specifies the Jacobian approximation function to be used

for the backward problem in the case where the backward problem depends on the

forward sensitivities.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) represents the identifier of the backward problem.

jacBS (IDALJacFnBS) user-defined Jacobian approximation function.

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS IDASetJacFnBS succeeded.

IDALS\_MEM\_NULL The ida\_mem was NULL.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

 ${\tt IDALS\_LMEM\_NULL}$  The linear solver has not been initialized with a call to

 ${\tt IDASetLinearSolverBS}.$ 

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes

The function type IDALsJacFnBS is described in §6.3.5.

The previous routine IDAD1sSetJacFnBS is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

When using a matrix-free linear solver module for the backward problem, the IDALS linear solver interface requires a function to compute an approximation to the product between the Jacobian matrix J(t,y) and a vector v. This may be performed internally using a difference-quotient approximation, or it may be supplied by the user by calling one of the following two functions:

#### IDASetJacTimesB

Call flag = IDASetJacTimesB(ida\_mem, which, jsetupB, jtimesB);

Description The function IDASetJacTimesB specifies the Jacobian-vector setup and product func-

tions to be used.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

jtsetupB (IDALsJacTimesSetupFnB) user-defined function to set up the Jacobian-vector product. Pass NULL if no setup is necessary.

jtimesB (IDALsJacTimesVecFnB) user-defined Jacobian-vector product function.

Return value The return value flag (of type int) is one of:

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem memory block pointer was NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes

The function types IDALsJacTimesVecFnB and IDALsJacTimesSetupFnB are described in §6.3.6.

The previous routine IDASpilsSetJacTimesB is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDASetJacTimesBS

Call flag = IDASetJacTimesBS(ida\_mem, which, jsetupBS, jtimesBS);

Description The function IDASetJacTimesBS specifies the Jacobian-vector product setup and evaluation functions to be used, in the case where the backward problem depends on the

forward sensitivities.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

jtsetupBS (IDALsJacTimesSetupFnBS) user-defined function to set up the Jacobian-vector product. Pass NULL if no setup is necessary.

jtimesBS (IDALsJacTimesVecFnBS) user-defined Jacobian-vector product function.

Return value The return value flag (of type int) is one of:

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem memory block pointer was NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

 ${\tt IDALS\_NO\_ADJ} \qquad {\tt The \ function \ IDAAdjInit \ has \ not \ been \ previously \ called}.$ 

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes The function types IDALsJacTimesVecFnBS and IDALsJacTimesSetupFnBS are described in §6.3.6.

The previous routine IDASpilsSetJacTimesBS is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

Alternately, when using the default difference-quotient approximation to the Jacobian-vector product for the backward problem, the user may specify the factor to use in setting increments for the finite-difference approximation, via a call to IDASetIncrementFactorB:

#### IDASetIncrementFactorB

Call flag = IDASetIncrementFactorB(ida\_mem, which, dqincfacB);

Description The function IDASetIncrementFactorB specifies the factor in the increments used in the difference quotient approximations to matrix-vector products for the backward problem.

This routine can be used in both the cases wherethe backward problem does and does not depend on the forward sensitivities.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

dqincfacB (realtype) difference quotient approximation factor.

Return value The return value flag (of type int) is one of

 ${\tt IDALS\_SUCCESS} \quad \text{ The optional value has been successfully set.}$ 

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDALS\_ILL\_INPUT The value of eplifacB is negative.

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes The default value is 1.0.

The previous routine IDASpilsSetIncrementFactorB is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

When using an iterative linear solver for the backward problem, the user may supply a preconditioning operator to aid in solution of the system, or she/he may adjust the convergence tolerance factor for the iterative linear solver. These may be accomplished through calling the following functions:

#### IDASetPreconditionerB

Call flag = IDASetPreconditionerB(ida\_mem, which, psetupB, psolveB);

Description The function IDASetPrecSolveFnB specifies the preconditioner setup and solve func-

tions for the backward integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

psetupB (IDALsPrecSetupFnB) user-defined preconditioner setup function.

psolveB (IDALsPrecSolveFnB) user-defined preconditioner solve function.

Return value The return value flag (of type int) is one of:

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem memory block pointer was NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes The function types IDALsPrecSolveFnB and IDALsPrecSetupFnB are described in §6.3.8

and §6.3.9, respectively. The psetupB argument may be NULL if no setup operation is

involved in the preconditioner.

The previous routine IDASpilsSetPreconditionerB is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### IDASetPreconditionerBS

Call flag = IDASetPreconditionerBS(ida\_mem, which, psetupBS, psolveBS);

Description The function IDASetPrecSolveFnBS specifies the preconditioner setup and solve func-

tions for the backward integration, in the case where the backward problem depends on

the forward sensitivities.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

psetupBS (IDALsPrecSetupFnBS) user-defined preconditioner setup function.

psolveBS (IDALsPrecSolveFnBS) user-defined preconditioner solve function.

Return value The return value flag (of type int) is one of:

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem memory block pointer was NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes

The function types IDALsPrecSolveFnBS and IDALsPrecSetupFnBS are described in §6.3.8 and §6.3.9, respectively. The psetupBS argument may be NULL if no setup operation is involved in the preconditioner.

The previous routine IDASpilsSetPreconditionerBS is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### ${\tt IDASetEpsLinB}$

Call flag = IDASetEpsLinB(ida\_mem, which, eplifacB);

Description

The function IDASetEpsLinB specifies the factor by which the Krylov linear solver's convergence test constant is reduced from the nonlinear iteration test constant. (See §2.1). This routine can be used in both the cases wherethe backward problem does and does not depend on the forward sensitivities.

Arguments

ida\_mem (void \*) pointer to the IDAS memory block.
which (int) the identifier of the backward problem.

eplifacB (realtype) linear convergence safety factor (>= 0.0).

Return value The return value flag (of type int) is one of

IDALS\_SUCCESS The optional value has been successfully set.

IDALS\_MEM\_NULL The ida\_mem pointer is NULL.

IDALS\_LMEM\_NULL The IDALS linear solver has not been initialized.

IDALS\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDALS\_ILL\_INPUT The value of eplifacB is negative.

IDALS\_ILL\_INPUT The parameter which represented an invalid identifier.

Notes

The default value is 0.05.

Passing a value eplifacB= 0.0 also indicates using the default value.

The previous routine IDASpilsSetEpsLinB is now a wrapper for this routine, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new routine name soon.

#### 6.2.10 Optional output functions for the backward problem

#### 6.2.10.1 Main solver optional output functions

The user of the adjoint module in IDAS has access to any of the optional output functions described in §4.5.10, both for the main solver and for the linear solver modules. The first argument of these IDAGet\* and IDA\*Get\* functions is the pointer to the IDAS memory block for the backward problem. In order to call any of these functions, the user must first call the following function to obtain this pointer:

#### IDAGetAdjIDABmem

Call ida\_memB = IDAGetAdjIDABmem(ida\_mem, which);

Description The function IDAGetAdjIDABmem returns a pointer to the IDAS memory block for the

backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block created by IDACreate.

which (int) the identifier of the backward problem.

Return value The return value, ida\_memB (of type void \*), is a pointer to the IDAS memory for the backward problem.

Notes

The user should not modify ida\_memB in any way.

Optional output calls should pass ida\_memB as the first argument; thus, for example, to get the number of integration steps: flag = IDAGetNumSteps(idas\_memB,&nsteps).

To get values of the *forward* solution during a backward integration, use the following function. The input value of t would typically be equal to that at which the backward solution has just been obtained with IDAGetB. In any case, it must be within the last checkpoint interval used by IDASolveB.

#### IDAGetAdjY

Call flag = IDAGetAdjY(ida\_mem, t, y, yp);

Description The function  $\mathtt{IDAGetAdjY}$  returns the interpolated value of the forward solution y and its derivative during a backward integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block created by IDACreate.

t (realtype) value of the independent variable at which y is desired (input).

y (N\_Vector) forward solution y(t).

yp (N\_Vector) forward solution derivative  $\dot{y}(t)$ .

Return value The return value flag (of type int) is one of:

IDA\_SUCCESS IDAGetAdjY was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_GETY\_BADT The value of t was outside the current checkpoint interval.

Notes The user must allocate space for y and yp.

#### IDAGetAdjCheckPointsInfo

Call flag = IDAGetAdjCheckPointsInfo(ida\_mem, IDAadjCheckPointRec \*ckpnt);

Description The function IDAGetAdjCheckPointsInfo loads an array of ncheck+1 records of type IDAadjCheckPointRec. The user must allocate space for the array ckpnt.

Arguments ida\_mem (void \*) pointer to the IDAS memory block created by IDACreate.

Return value The return value is IDA\_SUCCESS if successful, or IDA\_MEM\_NULL if ida\_mem is NULL, or IDA\_NO\_ADJ if ASA was not initialized.

Notes The members of each record ckpnt[i] are:

- ckpnt[i].my\_addr (void \*) address of current checkpoint in ida\_mem->ida\_adj\_mem
- ckpnt[i].next\_addr (void \*) address of next checkpoint
- ckpnt[i].t0 (realtype) start of checkpoint interval
- ckpnt[i].t1 (realtype) end of checkpoint interval
- ckpnt[i].nstep (long int) step counter at ckeckpoint t0
- ckpnt[i].order (int) method order at checkpoint t0
- ckpnt[i].step (realtype) step size at checkpoint t0

## 6.2.10.2 Initial condition calculation optional output function





#### IDAGetConsistentICB

Call flag = IDAGetConsistentICB(ida\_mem, which, yB0\_mod, ypB0\_mod);

Description The function IDAGetConsistentICB returns the corrected initial conditions for back-

ward problem calculated by IDACalcICB.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which is the identifier of the backward problem.

yBO\_mod (N\_Vector) consistent initial vector.

ypBO\_mod (N\_Vector) consistent initial derivative vector.

Return value The return value flag (of type int) is one of

IDA\_SUCCESS The optional output value has been successfully set.

IDA\_MEM\_NULL The ida\_mem pointer is NULL.

IDA\_NO\_ADJ IDAAdjInit has not been previously called.

IDA\_ILL\_INPUT Parameter which did not refer a valid backward problem identifier.

Notes If the consistent solution vector or consistent derivative vector is not desired, pass NULL

for the corresponding argument.

The user must allocate space for yB0\_mod and ypB0\_mod (if not NULL).

#### 6.2.11 Backward integration of quadrature equations

Not only the backward problem but also the backward quadrature equations may or may not depend on the forward sensitivities. Accordingly, one of the IDAQuadInitB or IDAQuadInitBS should be used to allocate internal memory and to initialize backward quadratures. For any other operation (extraction, optional input/output, reinitialization, deallocation), the same function is called regardless of whether or not the quadratures are sensitivity-dependent.

#### 6.2.11.1 Backward quadrature initialization functions

The function IDAQuadInitB initializes and allocates memory for the backward integration of quadrature equations that do not depend on forward sensitivities. It has the following form:

#### IDAQuadInitB

Call flag = IDAQuadInitB(ida\_mem, which, rhsQB, yQBO);

Description The function IDAQuadInitB provides required problem specifications, allocates internal

memory, and initializes backward quadrature integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

 ${\tt rhsQB}$  (IDAQuadRhsFnB) is the C function which computes fQB, the residual of the

backward quadrature equations. This function has the form rhsQB(t, y, yp,

yB, ypB, rhsvalBQ, user\_dataB) (see §6.3.3). (N\_Vector) is the value of the quadrature variables at tB0.

yQBO (N\_Vector) is the value of the quadrature variables at tBo

Return value The return value flag (of type int) will be one of the following:

 ${\tt IDA\_SUCCESS} \quad \text{ The call to } {\tt IDAQuadInitB} \text{ was successful.}$ 

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_MEM\_FAIL A memory allocation request has failed.

IDA\_ILL\_INPUT The parameter which is an invalid identifier.

The function IDAQuadInitBS initializes and allocates memory for the backward integration of quadrature equations that depend on the forward sensitivities.



#### IDAQuadInitBS

Call flag = IDAQuadInitBS(ida\_mem, which, rhsQBS, yQBS0);

Description The function IDAQuadInitBS provides required problem specifications, allocates internal

memory, and initializes backward quadrature integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

rhsQBS (IDAQuadRhsFnBS) is the C function which computes fQBS, the residual of the backward quadrature equations. This function has the form rhsQBS(t, y, yp, yS, ypS, yB, ypB, rhsvalBQS, user\_dataB) (see §6.3.4).

yQBSO (N\_Vector) is the value of the sensitivity-dependent quadrature variables at tBO.

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAQuadInitBS was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_MEM\_FAIL A memory allocation request has failed.

IDA\_ILL\_INPUT The parameter which is an invalid identifier.

The integration of quadrature equations during the backward phase can be re-initialized by calling the following function. Before calling <code>IDAQuadReInitB</code> for a new backward problem, call any desired solution extraction functions <code>IDAGet\*\*</code> associated with the previous backward problem.

#### IDAQuadReInitB

Call flag = IDAQuadReInitB(ida\_mem, which, yQB0);

Description The function IDAQuadReInitB re-initializes the backward quadrature integration.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

yQBO (N\_Vector) is the value of the quadrature variables at tBO.

Return value The return value flag (of type int) will be one of the following:

IDA\_SUCCESS The call to IDAQuadReInitB was successful.

IDA\_MEM\_NULL ida\_mem was NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_MEM\_FAIL A memory allocation request has failed.

IDA\_NO\_QUAD Quadrature integration was not activated through a previous call to

IDAQuadInitB.

IDA\_ILL\_INPUT The parameter which is an invalid identifier.

Notes IDAQuadReInitB can be used after a call to either IDAQuadInitB or IDAQuadInitBS.

#### 6.2.11.2 Backward quadrature extraction function

To extract the values of the quadrature variables at the last return time of IDASolveB, IDAS provides a wrapper for the function IDAGetQuad (see §4.7.3). The call to this function has the form

#### ${\tt IDAGetQuadB}$

Call flag = IDAGetQuadB(ida\_mem, which, &tret, yQB);

 $\label{prop:local_prop} \textbf{Description} \quad \textbf{The function } \textbf{IDAGetQuadB} \ \ \textbf{returns the quadrature solution vector after a successful}$ 

return from IDASolveB.

Arguments ida\_mem (void \*) pointer to the IDAS memory.



```
tret (realtype) the time reached by the solver (output).
yQB (N_Vector) the computed quadrature vector.
```

Return value

Notes T

he user must allocate space for yQB. The return value flag of IDAGetQuadB is one of:

IDA\_SUCCESS IDAGetQuadB was successful.

IDA\_MEM\_NULL ida\_mem is NULL.

IDA\_NO\_ADJ The function IDAAdjInit has not been previously called.

IDA\_NO\_QUAD Quadrature integration was not initialized.

IDA\_BAD\_DKY yQB was NULL.

IDA\_ILL\_INPUT The parameter which is an invalid identifier.

#### 6.2.11.3 Optional input/output functions for backward quadrature integration

Optional values controlling the backward integration of quadrature equations can be changed from their default values through calls to one of the following functions which are wrappers for the corresponding optional input functions defined in §4.7.4. The user must specify the identifier which of the backward problem for which the optional values are specified.

```
flag = IDASetQuadErrConB(ida_mem, which, errconQ);
flag = IDAQuadSStolerancesB(ida_mem, which, reltolQ, abstolQ);
flag = IDAQuadSVtolerancesB(ida_mem, which, reltolQ, abstolQ);
```

Their return value flag (of type int) can have any of the return values of its counterparts, but it can also be IDA\_NO\_ADJ if the function IDAAdjInit has not been previously called or IDA\_ILL\_INPUT if the parameter which was an invalid identifier.

Access to optional outputs related to backward quadrature integration can be obtained by calling the corresponding IDAGetQuad\* functions (see §4.7.5). A pointer ida\_memB to the IDAS memory block for the backward problem, required as the first argument of these functions, can be obtained through a call to the functions IDAGetAdjIDABmem (see §6.2.10).

### 6.3 User-supplied functions for adjoint sensitivity analysis

In addition to the required DAE residual function and any optional functions for the forward problem, when using the adjoint sensitivity module in IDAS, the user must supply one function defining the backward problem DAE and, optionally, functions to supply Jacobian-related information and one or two functions that define the preconditioner (if applicable for the choice of SUNLINSOL object) for the backward problem. Type definitions for all these user-supplied functions are given below.

#### 6.3.1 DAE residual for the backward problem

The user must provide a resB function of type IDAResFnB defined as follows:

```
IDAResFnB
```

```
Definition typedef int (*IDAResFnB)(realtype t, N_Vector y, N_Vector yp, N_Vector yB, N_Vector ypB,
```

N\_Vector resvalB, void \*user\_dataB);

Purpose This function evaluates the residual of the backward problem DAE system. This could

be (2.20) or (2.25).

Arguments t is the current value of the independent variable.

у	is the current value of the forward solution vector.
ур	is the current value of the forward solution derivative vector.
уВ	is the current value of the backward dependent variable vector.
урВ	is the current value of the backward dependent derivative vector.

is the output vector containing the residual for the backward DAE problem. resvalB

user\_dataB is a pointer to user data, same as passed to IDASetUserDataB.

Return value An IDAResFnB should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if an unrecoverabl failure occurred (in which case the integration stops and IDASolveB returns IDA\_RESFUNC\_FAIL).

Notes

Allocation of memory for resvalB is handled within IDAS.

The y, yp, yB, ypB, and resvalB arguments are all of type N\_Vector, but yB, ypB, and resvalB typically have different internal representations from y and yp. It is the user's responsibility to access the vector data consistently (including the use of the correct accessor macros from each NVECTOR implementation). For the sake of computational efficiency, the vector functions in the two NVECTOR implementations provided with IDAS do not perform any consistency checks with respect to their N\_Vector arguments (see  $\S 7.2 \text{ and } \S 7.3$ ).

The user\_dataB pointer is passed to the user's resB function every time it is called and can be the same as the user\_data pointer used for the forward problem.

Before calling the user's resB function, IDAS needs to evaluate (through interpolation) the values of the states from the forward integration. If an error occurs in the interpolation, IDAS triggers an unrecoverable failure in the residual function which will halt the integration and IDASolveB will return IDA\_RESFUNC\_FAIL.

#### 6.3.2DAE residual for the backward problem depending on the forward sensitivities

The user must provide a resBS function of type IDAResFnBS defined as follows:

#### IDAResFnBS Definition typedef int (\*IDAResFnBS)(realtype t, N\_Vector y, N\_Vector yp, N\_Vector \*yS, N\_Vector \*ypS, N\_Vector yB, N\_Vector ypB, N\_Vector resvalB, void \*user\_dataB); Purpose This function evaluates the residual of the backward problem DAE system. This could be (2.20) or (2.25). is the current value of the independent variable. Arguments t is the current value of the forward solution vector. is the current value of the forward solution derivative vector. ур a pointer to an array of Ns vectors containing the sensitivities of the forward yS a pointer to an array of Ns vectors containing the derivatives of the forward ypS sensitivities. is the current value of the backward dependent variable vector. yВ is the current value of the backward dependent derivative vector. урВ is the output vector containing the residual for the backward DAE problem. resvalB user\_dataB is a pointer to user data, same as passed to IDASetUserDataB.

Return value An IDAResFnBS should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if an unrecoverable error occurred (in which case the integration stops and IDASolveB returns IDA\_RESFUNC\_FAIL).

Notes

Allocation of memory for resvalB is handled within IDAS.

The y, yp, yB, ypB, and resvalB arguments are all of type N\_Vector, but yB, ypB, and resvalB typically have different internal representations from y and yp. Likewise for each yS[i] and ypS[i]. It is the user's responsibility to access the vector data consistently (including the use of the correct accessor macros from each NVECTOR implementation). For the sake of computational efficiency, the vector functions in the two NVECTOR implementations provided with IDAS do not perform any consistency checks with respect to their N\_Vector arguments (see  $\S7.2$  and  $\S7.3$ ).

The user\_dataB pointer is passed to the user's resBS function every time it is called and can be the same as the user\_data pointer used for the forward problem.

Before calling the user's resBS function, IDAS needs to evaluate (through interpolation) the values of the states from the forward integration. If an error occurs in the interpolation, IDAS triggers an unrecoverable failure in the residual function which will halt the integration and IDASolveB will return IDA\_RESFUNC\_FAIL.

#### 6.3.3 Quadrature right-hand side for the backward problem

The user must provide an fQB function of type IDAQuadRhsFnB defined by

#### IDAQuadRhsFnB

Definition typedef int (\*IDAQuadRhsFnB) (realtype t, N\_Vector y, N\_Vector yp, N\_Vector yB, N\_Vector ypB, N\_Vector rhsvalBQ, void \*user\_dataB);

Purpose

This function computes the quadrature equation right-hand side for the backward prob-

Arguments

t is the current value of the independent variable. is the current value of the forward solution vector. У

is the current value of the forward solution derivative vector. ур yВ is the current value of the backward dependent variable vector. is the current value of the backward dependent derivative vector. урВ

is the output vector containing the residual for the backward quadrature rhsvalBQ

equations.

user\_dataB is a pointer to user data, same as passed to IDASetUserDataB.

Return value An IDAQuadRhsFnB should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if it failed unrecoverably (in which case the integration is halted and IDASolveB returns IDA\_QRHSFUNC\_FAIL).

Notes

Allocation of memory for rhsvalBQ is handled within IDAS.

The y, yp, yB, ypB, and rhsvalBQ arguments are all of type N\_Vector, but they typically all have different internal representations. It is the user's responsibility to access the vector data consistently (including the use of the correct accessor macros from each NVECTOR implementation). For the sake of computational efficiency, the vector functions in the two NVECTOR implementations provided with IDAS do not perform any consistency checks with repsect to their N\_Vector arguments (see §7.2 and §7.3).

The user\_dataB pointer is passed to the user's fQB function every time it is called and can be the same as the user\_data pointer used for the forward problem.



Before calling the user's fQB function, IDAS needs to evaluate (through interpolation) the values of the states from the forward integration. If an error occurs in the interpolation, IDAS triggers an unrecoverable failure in the quadrature right-hand side function which will halt the integration and IDASolveB will return IDA\_QRHSFUNC\_FAIL.



#### 6.3.4 Sensitivity-dependent quadrature right-hand side for the backward problem

The user must provide an fQBS function of type IDAQuadRhsFnBS defined by

#### IDAQuadRhsFnBS

Definition	typedef i	nt (*IDAQuadRhsFnBS)(realtype t, N_Vector y, N_Vector yp,
		N_Vector *yS, N_Vector *ypS,
		N_Vector yB, N_Vector ypB,
		$N_{Vector}$ rhsvalBQS, void *user_dataB);
Purpose	This functi	on computes the quadrature equation residual for the backward problem.
Arguments	t	is the current value of the independent variable.
		is the company value of the farmend solution reactor

is the current value of the forward solution vector. У is the current value of the forward solution derivative vector. ур

a pointer to an array of Ns vectors containing the sensitivities of the forward yS

solution.

a pointer to an array of Ns vectors containing the derivatives of the forward ypS

sensitivities.

is the current value of the backward dependent variable vector. yВ

is the current value of the backward dependent derivative vector. урВ

is the output vector containing the residual for the backward quadrature rhsvalBQS

equations.

user\_dataB is a pointer to user data, same as passed to IDASetUserDataB.

Return value An IDAQuadRhsFnBS should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if it failed unrecoverably (in which case the integration is halted and IDASolveB returns IDA\_QRHSFUNC\_FAIL).

Notes

Allocation of memory for rhsvalBQS is handled within IDAS.

The y, yp, yB, ypB, and rhsvalBQS arguments are all of type N\_Vector, but they typically do not all have the same internal representations. Likewise for each yS[i] and ypS[i]. It is the user's responsibility to access the vector data consistently (including the use of the correct accessor macros from each NVECTOR implementation). For the sake of computational efficiency, the vector functions in the two NVECTOR implementations provided with IDAS do not perform any consistency checks with repsect to their N\_Vector arguments (see  $\S7.2$  and  $\S7.3$ ).

The user\_dataB pointer is passed to the user's fQBS function every time it is called and can be the same as the user\_data pointer used for the forward problem.

Before calling the user's fQBS function, IDAS needs to evaluate (through interpolation) the values of the states from the forward integration. If an error occurs in the interpolation, IDAS triggers an unrecoverable failure in the quadrature right-hand side function which will halt the integration and IDASolveB will return IDA\_QRHSFUNC\_FAIL.



## 6.3.5 Jacobian construction for the backward problem (matrix-based linear solvers)

If a matrix-based linear solver module is is used for the backward problem (i.e., IDASetLinearSolverB is called with non-NULL SUNMATRIX argument in the step described in §6.1), the user may provide a function of type IDALsJacFnB or IDALsJacFnBS (see §6.2.9), defined as follows:

#### IDALsJacFnB

Definition typedef int (\*IDALsJacFnB) (realtype tt, realtype cjB, N\_Vector yy, N\_Vector yp, N\_Vector yB, N\_Vector ypB, N\_Vector resvalB, SUNMatrix JacB, void \*user\_dataB, N\_Vector tmp1B, N\_Vector tmp2B, N\_Vector tmp3B); This function computes the Jacobian of the backward problem (or an approximation to Purpose it). Arguments is the current value of the independent variable. tt is the scalar in the system Jacobian, proportional to the inverse of the step cjB size ( $\alpha$  in Eq. (2.6)). is the current value of the forward solution vector. уу is the current value of the forward solution derivative vector. ур is the current value of the backward dependent variable vector. yВ is the current value of the backward dependent derivative vector. урВ resvalB is the current value of the residual for the backward problem. is the output approximate Jacobian matrix. JacB

user\_dataB is a pointer to user data — the parameter passed to IDASetUserDataB.

tmp1B tmp2B

tmp3B are pointers to memory allocated for variables of type N\_Vector which can be used by the IDALsJacFnB function as temporary storage or work space.

Return value An IDALsJacFnB should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct, while IDALS sets last\_flag to IDALS\_JACFUNC\_RECVR), or a negative value if it failed unrecoverably (in which case the integration is halted, IDASolveB returns IDA\_LSETUP\_FAIL and IDALS sets last\_flag to IDALS\_JACFUNC\_UNRECVR).

A user-supplied Jacobian function must load the matrix JacB with an approximation to the Jacobian matrix at the point (tt,yy,yB), where yy is the solution of the original IVP at time tt, and yB is the solution of the backward problem at the same time. Information regarding the structure of the specific SUNMATRIX structure (e.g. number of rows, upper/lower bandwidth, sparsity type) may be obtained through using the implementation-specific SUNMATRIX interface functions (see Chapter 8 for details). Only nonzero elements need to be loaded into JacB as this matrix is set to zero before the call to the Jacobian function.

Before calling the user's IDALsJacFnB, IDAS needs to evaluate (through interpolation) the values of the states from the forward integration. If an error occurs in the interpolation, IDAS triggers an unrecoverable failure in the Jacobian function which will halt the integration (IDASolveB returns IDA\_LSETUP\_FAIL and IDALS sets last\_flag to IDALS\_JACFUNC\_UNRECVR).



Notes

The previous function type IDADlsJacFnB is identical to IDALsJacFnB, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

#### IDALsJacFnBS

Definition typedef int (\*IDALsJacFnBS)(realtype tt, realtype cjB, N\_Vector yy, N\_Vector yp, N\_Vector \*yS, N\_Vector \*ypS, N\_Vector yB, N\_Vector ypB, N\_Vector resvalB. SUNMatrix JacB, void \*user\_dataB, N\_Vector tmp1B, N\_Vector tmp2B, N\_Vector tmp3B); Purpose This function computes the Jacobian of the backward problem (or an approximation to it), in the case where the backward problem depends on the forward sensitivities. is the current value of the independent variable. Arguments tt сјВ is the scalar in the system Jacobian, proportional to the inverse of the step size ( $\alpha$  in Eq. (2.6)). is the current value of the forward solution vector. уу is the current value of the forward solution derivative vector. ур yS a pointer to an array of Ns vectors containing the sensitivities of the forward solution. a pointer to an array of Ns vectors containing the derivatives of the forward ypS solution sensitivities. yВ is the current value of the backward dependent variable vector. is the current value of the backward dependent derivative vector. урВ is the current value of the residual for the backward problem. resvalB JacB is the output approximate Jacobian matrix. user\_dataB is a pointer to user data — the parameter passed to IDASetUserDataB. tmp1B tmp2B tmp3B are pointers to memory allocated for variables of type N\_Vector which can

Return value An IDALsJacFnBS should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct, while IDALS sets last\_flag to IDALS\_JACFUNC\_RECVR), or a negative value if it failed unrecoverably (in which case the integration is halted, IDASolveB returns IDA\_LSETUP\_FAIL and IDALS sets last\_flag to IDALS\_JACFUNC\_UNRECVR).

be used by IDALsJacFnBS as temporary storage or work space.

Notes

A user-supplied dense Jacobian function must load the matrix JacB with an approximation to the Jacobian matrix at the point (tt,yy,yS,yB), where yy is the solution of the original IVP at time tt, yS is the array of forward sensitivities at time tt, and yB is the solution of the backward problem at the same time. Information regarding the structure of the specific SUNMATRIX structure (e.g. number of rows, upper/lower bandwidth, sparsity type) may be obtained through using the implementation-specific SUNMATRIX interface functions (see Chapter 8 for details). Only nonzero elements need to be loaded into JacB as this matrix is set to zero before the call to the Jacobian function.

Before calling the user's IDALsJacFnBS, IDAS needs to evaluate (through interpolation) the values of the states from the forward integration. If an error occurs in the interpolation, IDAS triggers an unrecoverable failure in the Jacobian function which will



halt the integration (IDASolveB returns IDA\_LSETUP\_FAIL and IDALS sets last\_flag to IDALS\_JACFUNC\_UNRECVR).

The previous function type IDADlsJacFnBS is identical to IDALsJacFnBS, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

## 6.3.6 Jacobian-vector product for the backward problem (matrix-free linear solvers)

If a matrix-free linear solver is selected for the backward problem (i.e., IDASetLinearSolverB is called with NULL-valued SUNMATRIX argument in the steps described in §6.1), the user may provide a function of type

IDALsJacTimesVecFnBs or IDALsJacTimesVecFnBs in the following form, to compute matrix-vector products Jv. If such a function is not supplied, the default is a difference quotient approximation to these products.

#### IDALsJacTimesVecFnB

```
Definition
             typedef int (*IDALsJacTimesVecFnB)(realtype t,
                                                      N_Vector yy, N_Vector yp,
                                                      N_Vector yB, N_Vector ypB,
                                                      N_Vector resvalB,
                                                      N_Vector vB, N_Vector JvB,
                                                      realtype cjB, void *user_dataB,
                                                      N_Vector tmp1B, N_Vector tmp2B);
Purpose
             This function computes the action of the backward problem Jacobian JB on a given
             vector vB.
Arguments
                          is the current value of the independent variable.
                          is the current value of the forward solution vector.
             уу
                          is the current value of the forward solution derivative vector.
             ур
                          is the current value of the backward dependent variable vector.
             yВ
                          is the current value of the backward dependent derivative vector.
             ypB
                          is the current value of the residual for the backward problem.
             resvalB
                          is the vector by which the Jacobian must be multiplied.
             vΒ
             JvB
                          is the computed output vector, JB*vB.
                          is the scalar in the system Jacobian, proportional to the inverse of the step
             cjB
                          size (\alpha in Eq. (2.6)).
             user_dataB is a pointer to user data — the same as the user_dataB parameter passed
                          to IDASetUserDataB.
             tmp1B
             tmp2B
                          are pointers to memory allocated for variables of type N_Vector which can
                          be used by IDALsJacTimesVecFnB as temporary storage or work space.
```

Return value The return value of a function of type IDALsJtimesVecFnB should be 0 if successful or nonzero if an error was encountered, in which case the integration is halted.

Notes

A user-supplied Jacobian-vector product function must load the vector JvB with the product of the Jacobian of the backward problem at the point (t, y, yB) and the vector vB. Here, y is the solution of the original IVP at time t and yB is the solution of the backward problem at the same time. The rest of the arguments are equivalent to those passed to a function of type IDALsJacTimesVecFn (see §4.6.6). If the backward problem is the adjoint of  $\dot{y} = f(t, y)$ , then this function is to compute  $-(\partial f/\partial y)^T v_B$ .

The previous function type IDASpilsJacTimesVecFnB is identical to IDALsJacTimesVecFnB, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

#### IDALsJacTimesVecFnBS

Definition typedef int (\*IDALsJacTimesVecFnBS)(realtype t, N\_Vector yy, N\_Vector yp, N\_Vector \*yyS, N\_Vector \*ypS, N\_Vector yB, N\_Vector ypB, N\_Vector resvalB, N\_Vector vB, N\_Vector JvB, realtype cjB, void \*user\_dataB, N\_Vector tmp1B, N\_Vector tmp2B); This function computes the action of the backward problem Jacobian JB on a given Purpose vector vB, in the case where the backward problem depends on the forward sensitivities. Arguments t is the current value of the independent variable. is the current value of the forward solution vector. уу is the current value of the forward solution derivative vector. ур a pointer to an array of Ns vectors containing the sensitivities of the forward yyS solution. a pointer to an array of Ns vectors containing the derivatives of the forward ypS sensitivities. is the current value of the backward dependent variable vector. yВ is the current value of the backward dependent derivative vector. урВ resvalB is the current value of the residual for the backward problem. is the vector by which the Jacobian must be multiplied. vΒ JvB is the computed output vector, JB\*vB. is the scalar in the system Jacobian, proportional to the inverse of the step cjB size ( $\alpha$  in Eq. (2.6)). user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed to IDASetUserDataB.

Return value The return value of a function of type IDALsJtimesVecFnBS should be 0 if successful or nonzero if an error was encountered, in which case the integration is halted.

Notes

tmp1B

tmp2B

A user-supplied Jacobian-vector product function must load the vector JvB with the product of the Jacobian of the backward problem at the point (t, y, yB) and the vector vB. Here, y is the solution of the original IVP at time t and yB is the solution of the backward problem at the same time. The rest of the arguments are equivalent to those passed to a function of type IDALsJacTimesVecFn (see §4.6.6).

are pointers to memory allocated for variables of type N\_Vector which can

be used by IDALsJacTimesVecFnBS as temporary storage or work space.

The previous function type IDASpilsJacTimesVecFnBS is identical to IDALsJacTimesVecFnBS, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

## 6.3.7 Jacobian-vector product setup for the backward problem (matrix-free linear solvers)

If the user's Jacobian-times-vector requires that any Jacobian-related data be preprocessed or evaluated, then this needs to be done in a user-supplied function of type IDALsJacTimesSetupFnB or IDALsJacTimesSetupFnBS, defined as follows:

#### IDALsJacTimesSetupFnB

Purpose This function preprocesses and/or evaluates Jacobian data needed by the Jacobian-times-vector routine for the backward problem.

Arguments tt is the current value of the independent variable.

yy is the current value of the dependent variable vector, y(t).

yp is the current value of  $\dot{y}(t)$ .

yB is the current value of the backward dependent variable vector. ypB is the current value of the backward dependent derivative vector.

resvalB is the current value of the residual for the backward problem.

cjB is the scalar in the system Jacobian, proportional to the inverse of the step

size ( $\alpha$  in Eq. (2.6)).

Return value The value returned by the Jacobian-vector setup function should be 0 if successful, positive for a recoverable error (in which case the step will be retried), or negative for an unrecoverable error (in which case the integration is halted).

Notes

Each call to the Jacobian-vector setup function is preceded by a call to the backward problem residual user function with the same (t,y, yp, yB, ypB) arguments. Thus, the setup function can use any auxiliary data that is computed and saved during the evaluation of the DAE residual.

If the user's IDALsJacTimesVecFnB function uses difference quotient approximations, it may need to access quantities not in the call list. These include the current stepsize, the error weights, etc. To obtain these, the user will need to add a pointer to ida\_mem to user\_dataB and then use the IDAGet\* functions described in §4.5.10.2. The unit roundoff can be accessed as UNIT\_ROUNDOFF defined in sundials\_types.h.

The previous function type IDASpilsJacTimesSetupFnB is identical to IDALsJacTimesSetupFnB, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

#### IDALsJacTimesSetupFnBS

Purpose This function preprocesses and/or evaluates Jacobian data needed by the Jacobian-times-vector routine for the backward problem, in the case that the backward problem depends on the forward sensitivities.

Arguments tt is the current value of the independent variable.

yy is the current value of the dependent variable vector, y(t).

yp is the current value of  $\dot{y}(t)$ .

yyS a pointer to an array of Ns vectors containing the sensitivities of the forward

solution.

ypS a pointer to an array of Ns vectors containing the derivatives of the forward

sensitivities.

yB is the current value of the backward dependent variable vector.

ypB is the current value of the backward dependent derivative vector.

resvalB is the current value of the residual for the backward problem.

cjB is the scalar in the system Jacobian, proportional to the inverse of the step

size ( $\alpha$  in Eq. (2.6)).

user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed to IDASetUserDataB.

Return value The value returned by the Jacobian-vector setup function should be 0 if successful, positive for a recoverable error (in which case the step will be retried), or negative for an unrecoverable error (in which case the integration is halted).

Notes

Each call to the Jacobian-vector setup function is preceded by a call to the backward problem residual user function with the same (t,y, yp, yyS, ypS, ypB, ypB) arguments. Thus, the setup function can use any auxiliary data that is computed and saved during the evaluation of the DAE residual.

If the user's IDALsJacTimesVecFnB function uses difference quotient approximations, it may need to access quantities not in the call list. These include the current stepsize, the error weights, etc. To obtain these, the user will need to add a pointer to ida\_mem to user\_dataB and then use the IDAGet\* functions described in §4.5.10.2. The unit roundoff can be accessed as UNIT\_ROUNDOFF defined in sundials\_types.h.

The previous function type IDASpilsJacTimesSetupFnBS is identical to IDALsJacTimesSetupFnBS, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

#### 6.3.8 Preconditioner solve for the backward problem (iterative linear solvers)

If preconditioning is used during integration of the backward problem, then the user must provide a function to solve the linear system Pz = r, where P is a left preconditioner matrix. This function must have one of the following two forms:

#### IDALsPrecSolveFnB

Purpose This function solves the preconditioning system Pz = r for the backward problem.

Arguments t is the current value of the independent variable.

yy is the current value of the forward solution vector.

yp is the current value of the forward solution derivative vector. yB is the current value of the backward dependent variable vector. ypB is the current value of the backward dependent derivative vector. resvalB is the current value of the residual for the backward problem. rvecB is the right-hand side vector r of the linear system to be solved.

zvecB is the computed output vector.

cjB is the scalar in the system Jacobian, proportional to the inverse of the step

size ( $\alpha$  in Eq. (2.6)).

deltaB is an input tolerance to be used if an iterative method is employed in the

solution.

user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed

to the function IDASetUserDataB.

Return value The return value of a preconditioner solve function for the backward problem should be 0 if successful, positive for a recoverable error (in which case the step will be retried),

or negative for an unrecoverable error (in which case the integration is halted).

or negative for an unrecoverable error (in which case the integration is natted).

Notes The previous function type IDASpilsPrecSolveFnB is identical to IDALsPrecSolveFnB, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name

soon.

#### IDALsPrecSolveFnBS

Definition typedef int (\*IDALsPrecSolveFnBS)(realtype t,

N\_Vector yy, N\_Vector yp,
N\_Vector \*yyS, N\_Vector \*ypS,
N\_Vector yB, N\_Vector ypB,

 $N_{-}$ Vector resvalB,

N\_Vector rvecB, N\_Vector zvecB, realtype cjB, realtype deltaB,

void \*user\_dataB);

Purpose This function solves the preconditioning system Pz = r for the backward problem, for

the case in which the backward problem depends on the forward sensitivities.

Arguments t is the current value of the independent variable.

yy is the current value of the forward solution vector.

yp is the current value of the forward solution derivative vector.

yyS a pointer to an array of Ns vectors containing the sensitivities of the forward

solution.

ypS a pointer to an array of Ns vectors containing the derivatives of the forward

sensitivities.

yB is the current value of the backward dependent variable vector.

ypB is the current value of the backward dependent derivative vector.

resvalB is the current value of the residual for the backward problem.

rvecB is the right-hand side vector r of the linear system to be solved.

zvecB is the computed output vector.

cjB is the scalar in the system Jacobian, proportional to the inverse of the step

size ( $\alpha$  in Eq. (2.6)).

deltaB is an input tolerance to be used if an iterative method is employed in the

solution.

user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed to the function IDASetUserDataB.

Return value The return value of a preconditioner solve function for the backward problem should be

0 if successful, positive for a recoverable error (in which case the step will be retried),

or negative for an unrecoverable error (in which case the integration is halted).

The previous function type IDASpilsPrecSolveFnBS is identical to IDALsPrecSolveFnBS, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

#### 6.3.9 Preconditioner setup for the backward problem (iterative linear solvers)

If the user's preconditioner requires that any Jacobian-related data be preprocessed or evaluated, then this needs to be done in a user-supplied function of one of the following two types:

#### IDALsPrecSetupFnB

Notes

```
Definition
            typedef int (*IDALsPrecSetupFnB)(realtype t,
                                               N_Vector yy, N_Vector yp,
                                               N_Vector yB, N_Vector ypB,
                                               N_Vector resvalB,
                                              realtype cjB, void *user_dataB);
```

This function preprocesses and/or evaluates Jacobian-related data needed by the pre-Purpose

conditioner for the backward problem.

The arguments of an IDALsPrecSetupFnB are as follows: Arguments

> is the current value of the independent variable. is the current value of the forward solution vector. уу is the current value of the forward solution vector. ур

is the current value of the backward dependent variable vector. yВ is the current value of the backward dependent derivative vector. ypB is the current value of the residual for the backward problem. resvalB

cjB is the scalar in the system Jacobian, proportional to the inverse of the step

size ( $\alpha$  in Eq. (2.6)).

user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed to the function IDASetUserDataB.

Return value The return value of a preconditioner setup function for the backward problem should be 0 if successful, positive for a recoverable error (in which case the step will be retried),

or negative for an unrecoverable error (in which case the integration is halted).

Notes

The previous function type IDASpilsPrecSetupFnB is identical to IDALsPrecSetupFnB, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

#### IDALsPrecSetupFnBS

```
Definition
            typedef int (*IDALsPrecSetupFnBS)(realtype t,
                                               N_Vector yy, N_Vector yp,
                                               N_Vector *yyS, N_Vector *ypS,
                                               N_Vector yB, N_Vector ypB,
                                               N_Vector resvalB,
                                               realtype cjB, void *user_dataB);
```

Purpose This function preprocesses and/or evaluates Jacobian-related data needed by the preconditioner for the backward problem, in the case where the backward problem depends on the forward sensitivities.

Arguments The arguments of an IDALsPrecSetupFnBS are as follows:

t is the current value of the independent variable.

yy is the current value of the forward solution vector.

yp is the current value of the forward solution vector.

yyS a pointer to an array of Ns vectors containing the sensitivities of the forward

solution.

 ${\tt ypS}$  a pointer to an array of  ${\tt Ns}$  vectors containing the derivatives of the forward

sensitivities.

yB is the current value of the backward dependent variable vector.

ypB is the current value of the backward dependent derivative vector.

resvalB is the current value of the residual for the backward problem.

cjB is the scalar in the system Jacobian, proportional to the inverse of the step

size ( $\alpha$  in Eq. (2.6)).

user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed
to the function IDASetUserDataB.

Return value The return value of a preconditioner setup function for the backward problem should be 0 if successful, positive for a recoverable error (in which case the step will be retried), or negative for an unrecoverable error (in which case the integration is halted).

Notes The previous function type IDASpilsPrecSetupFnBS is identical to IDALsPrecSetupFnBS, and may still be used for backward-compatibility. However, this will be deprecated in future releases, so we recommend that users transition to the new function type name soon.

# 6.4 Using the band-block-diagonal preconditioner for backward problems

As on the forward integration phase, the efficiency of Krylov iterative methods for the solution of linear systems can be greatly enhanced through preconditioning. The band-block-diagonal preconditioner module IDABBDPRE, provides interface functions through which it can be used on the backward integration phase.

The adjoint module in IDAS offers an interface to the band-block-diagonal preconditioner module IDABBDPRE described in section §4.8. This generates a preconditioner that is a block-diagonal matrix with each block being a band matrix and can be used with one of the Krylov linear solvers and with the MPI-parallel vector module NVECTOR\_PARALLEL.

In order to use the IDABBDPRE module in the solution of the backward problem, the user must define one or two additional functions, described at the end of this section.

#### 6.4.1 Usage of IDABBDPRE for the backward problem

The IDABBDPRE module is initialized by calling the following function, *after* an iterative linear solver for the backward problem has been attached to IDAS by calling IDASetLinearSolverB (see §6.2.6).

#### IDABBDPrecInitB

```
Call flag = IDABBDPrecInitB(ida_mem, which, NlocalB, mudqB, mldqB, mukeepB, mlkeepB, dqrelyB, GresB, GcommB);
```

Description The function IDABBDPrecInitB initializes and allocates memory for the IDABBDPRE preconditioner for the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block.

which (int) the identifier of the backward problem.

NlocalB (sunindextype) local vector dimension for the backward problem.

mudqB (sunindextype) upper half-bandwidth to be used in the difference-quotient Jacobian approximation.

mldqB (sunindextype) lower half-bandwidth to be used in the difference-quotient Jacobian approximation.

mukeepB (sunindextype) upper half-bandwidth of the retained banded approximate Jacobian block.

mlkeepB (sunindextype) lower half-bandwidth of the retained banded approximate Jacobian block.

dqrelyB (realtype) the relative increment in components of yB used in the difference quotient approximations. The default is dqrelyB=  $\sqrt{\text{unit roundoff}}$ , which can be specified by passing dqrely= 0.0.

GresB (IDABBDLocalFnB) the C function which computes  $G_B(t, y, \dot{y}, y_B, \dot{y}_B)$ , the function approximating the residual of the backward problem.

GcommB (IDABBDCommFnB) the optional C function which performs all interprocess communication required for the computation of  $G_B$ .

Return value If successful, IDABBDPrecInitB creates, allocates, and stores (internally in the IDAS solver block) a pointer to the newly created IDABBDPRE memory block. The return value flag (of type int) is one of:

IDALS\_SUCCESS The call to IDABBDPrecInitB was successful.

IDALS\_MEM\_FAIL A memory allocation request has failed.

IDALS\_MEM\_NULL The ida\_mem argument was NULL.

IDALS\_LMEM\_NULL No linear solver has been attached.

IDALS\_ILL\_INPUT An invalid parameter has been passed.

To reinitialize the IDABBDPRE preconditioner module for the backward problem, possibly with a change in mudqB, mldqB, or dqrelyB, call the following function:

#### IDABBDPrecReInitB

Call flag = IDABBDPrecReInitB(ida\_mem, which, mudqB, mldqB, dqrelyB);

Description The function IDABBDPrecReInitB reinitializes the IDABBDPRE preconditioner for the backward problem.

Arguments ida\_mem (void \*) pointer to the IDAS memory block returned by IDACreate.

which (int) the identifier of the backward problem.

mudqB (sunindextype) upper half-bandwidth to be used in the difference-quotient Jacobian approximation.

mldqB (sunindextype) lower half-bandwidth to be used in the difference-quotient Jacobian approximation.

dqrelyB (realtype) the relative increment in components of yB used in the difference quotient approximations.

Return value The return value flag (of type int) is one of:

IDALS\_SUCCESS The call to IDABBDPrecReInitB was successful.

IDALS\_MEM\_FAIL A memory allocation request has failed.

IDALS\_MEM\_NULL The ida\_mem argument was NULL.

IDALS\_PMEM\_NULL The IDABBDPrecInitB has not been previously called.

IDALS\_LMEM\_NULL No linear solver has been attached.

IDALS\_ILL\_INPUT An invalid parameter has been passed.

For more details on IDABBDPRE see §4.8.

#### 6.4.2 User-supplied functions for IDABBDPRE

To use the IDABBDPRE module, the user must supply one or two functions which the module calls to construct the preconditioner: a required function GresB (of type IDABBDLocalFnB) which approximates the residual of the backward problem and which is computed locally, and an optional function GcommB (of type IDABBDCommFnB) which performs all interprocess communication necessary to evaluate this approximate residual (see §4.8). The prototypes for these two functions are described below.

#### IDABBDLocalFnB

typedef int (\*IDABBDLocalFnB)(sunindextype NlocalB, realtype t, Definition N\_Vector y, N\_Vector yp, N\_Vector yB, N\_Vector ypB, N\_Vector gB, void \*user\_dataB); Purpose This GresB function loads the vector gB, an approximation to the residual of the backward problem, as a function of t, y, yp, and yB and ypB. Arguments NlocalB is the local vector length for the backward problem. is the value of the independent variable. t is the current value of the forward solution vector. у is the current value of the forward solution derivative vector. ур уВ is the current value of the backward dependent variable vector. is the current value of the backward dependent derivative vector. ypB gΒ is the output vector,  $G_B(t, y, \dot{y}, y_B, \dot{y}_B)$ . user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed

Return value An IDABBDLocalFnB should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if it failed unrecoverably (in which case the integration is halted and IDASolveB returns IDA\_LSETUP\_FAIL).

to IDASetUserDataB.

This routine must assume that all interprocess communication of data needed to calculate gB has already been done, and this data is accessible within user\_dataB.

Before calling the user's IDABBDLocalFnB, IDAS needs to evaluate (through interpolation) the values of the states from the forward integration. If an error occurs in the interpolation, IDAS triggers an unrecoverable failure in the preconditioner setup function which will halt the integration (IDASolveB returns IDA\_LSETUP\_FAIL).

#### IDABBDCommFnB

Notes

Definition	<pre>typedef int (*IDABBDCommFnB)(sunindextype NlocalB, realtype t,</pre>
	N_Vector y, N_Vector yp,
	$N_{-}$ Vector yB, $N_{-}$ Vector ypB,
	<pre>void *user_dataB);</pre>
Purpose	This $GcommB$ function performs all interprocess communications necessary for the execution of the $GresB$ function above, using the input vectors $y$ , $yp$ , $yB$ and $ypB$ .

Arguments NlocalB is the local vector length.

t is the value of the independent variable.

 ${\tt y}$  is the current value of the forward solution vector.

yp is the current value of the forward solution derivative vector.
yB is the current value of the backward dependent variable vector.
ypB is the current value of the backward dependent derivative vector.



user\_dataB is a pointer to user data — the same as the user\_dataB parameter passed to IDASetUserDataB.

Return value An IDABBDCommFnB should return 0 if successful, a positive value if a recoverable error occurred (in which case IDAS will attempt to correct), or a negative value if it failed unrecoverably (in which case the integration is halted and IDASolveB returns IDA\_LSETUP\_FAIL).

Notes

The GcommB function is expected to save communicated data in space defined within the structure user\_dataB.

Each call to the GcommB function is preceded by a call to the function that evaluates the residual of the backward problem with the same t, y, yp, yB and ypB arguments. If there is no additional communication needed, then pass GcommB = NULL to IDABBDPrecInitB.

## Chapter 7

# Description of the NVECTOR module

The SUNDIALS solvers are written in a data-independent manner. They all operate on generic vectors (of type N\_Vector) through a set of operations defined by the particular NVECTOR implementation. Users can provide their own specific implementation of the NVECTOR module, or use one of the implementations provided with SUNDIALS. The generic operations are described below and the implementations provided with SUNDIALS are described in the following sections.

The generic N\_Vector type is a pointer to a structure that has an implementation-dependent content field containing the description and actual data of the vector, and an ops field pointing to a structure with generic vector operations. The type N\_Vector is defined as

```
typedef struct _generic_N_Vector *N_Vector;
struct _generic_N_Vector {
    void *content;
    struct _generic_N_Vector_Ops *ops;
};
```

The \_generic\_N\_Vector\_Ops structure is essentially a list of pointers to the various actual vector operations, and is defined as

```
struct _generic_N_Vector_Ops {
  N_Vector_ID (*nvgetvectorid)(N_Vector);
  N_Vector
              (*nvclone)(N_Vector);
  N_Vector
              (*nvcloneempty)(N_Vector);
  void
              (*nvdestroy)(N_Vector);
              (*nvspace)(N_Vector, sunindextype *, sunindextype *);
  void
              (*nvgetarraypointer)(N_Vector);
  realtype*
              (*nvsetarraypointer)(realtype *, N_Vector);
  void
              (*nvlinearsum)(realtype, N_Vector, realtype, N_Vector, N_Vector);
  void
              (*nvconst)(realtype, N_Vector);
  void
  void
              (*nvprod)(N_Vector, N_Vector, N_Vector);
              (*nvdiv)(N_Vector, N_Vector, N_Vector);
  void
              (*nvscale)(realtype, N_Vector, N_Vector);
  void
  void
              (*nvabs)(N_Vector, N_Vector);
              (*nvinv)(N_Vector, N_Vector);
  void
  void
              (*nvaddconst)(N_Vector, realtype, N_Vector);
  realtype
              (*nvdotprod)(N_Vector, N_Vector);
  realtype
              (*nvmaxnorm)(N_Vector);
  realtype
              (*nvwrmsnorm)(N_Vector, N_Vector);
```

```
(*nvwrmsnormmask)(N_Vector, N_Vector, N_Vector);
  realtype
              (*nvmin)(N_Vector);
  realtype
              (*nvwl2norm)(N_Vector, N_Vector);
  realtype
  realtype
              (*nvl1norm)(N_Vector);
  void
              (*nvcompare)(realtype, N_Vector, N_Vector);
  booleantype (*nvinvtest)(N_Vector, N_Vector);
  booleantype (*nvconstrmask)(N_Vector, N_Vector, N_Vector);
              (*nvminquotient)(N_Vector, N_Vector);
  realtype
              (*nvlinearcombination)(int, realtype*, N_Vector*, N_Vector);
  int
  int
              (*nvscaleaddmulti)(int, realtype*, N_Vector, N_Vector*, N_Vector*);
              (*nvdotprodmulti)(int, N_Vector, N_Vector*, realtype*);
  int
              (*nvlinearsumvectorarray)(int, realtype, N_Vector*, realtype,
  int
                                         N_Vector*, N_Vector*);
  int
              (*nvscalevectorarray)(int, realtype*, N_Vector*, N_Vector*);
  int
              (*nvconstvectorarray)(int, realtype, N_Vector*);
              (*nvwrmsnomrvectorarray)(int, N_Vector*, N_Vector*, realtype*);
  int
              (*nvwrmsnomrmaskvectorarray)(int, N_Vector*, N_Vector*, N_Vector,
  int
                                            realtype*);
              (*nvscaleaddmultivectorarray)(int, int, realtype*, N_Vector*,
  int
                                             N_Vector**, N_Vector**);
              (*nvlinearcombinationvectorarray)(int, int, realtype*, N_Vector**,
  int
                                                 N Vector*):
};
```

These routines are nothing but wrappers for the vector operations defined by a particular NVECTOR implementation, which are accessed through the *ops* field of the N\_Vector structure. To illustrate this point we show below the implementation of a typical vector operation from the generic NVECTOR module, namely N\_VScale, which performs the scaling of a vector x by a scalar c:

```
void N_VScale(realtype c, N_Vector x, N_Vector z)
{
   z->ops->nvscale(c, x, z);
}
```

Table 7.2 contains a complete list of all standard vector operations defined by the generic NVECTOR module. Tables 7.3 and 7.4 list *optional* fused and vector array operations respectively.

Fused and vector array operations are intended to increase data reuse, reduce parallel communication on distributed memory systems, and lower the number of kernel launches on systems with accelerators. If a particular NVECTOR implementation defines a fused or vector array operation as NULL, the generic NVECTOR module will automatically call standard vector operations as necessary to complete the desired operation. Currently, all fused and vector array operations are disabled by default however, SUNDIALS provided NVECTOR implementations define additional user-callable functions to enable/disable any or all of the fused and vector array operations. See the following sections for the implementation specific functions to enable/disable operations.

Finally, note that the generic NVECTOR module defines the functions N\_VCloneVectorArray and N\_VCloneVectorArrayEmpty. Both functions create (by cloning) an array of count variables of type N\_Vector, each of the same type as an existing N\_Vector. Their prototypes are

```
N_Vector *N_VCloneVectorArray(int count, N_Vector w);
N_Vector *N_VCloneVectorArrayEmpty(int count, N_Vector w);
```

and their definitions are based on the implementation-specific N\_VClone and N\_VCloneEmpty operations, respectively.

An array of variables of type  $N_{\text{-}}$ Vector can be destroyed by calling  $N_{\text{-}}$ VDestroyVectorArray, whose prototype is

Table 7.1: Vector Identifications associated	with vector	r kernels supplied with SUNDIALS	3.
--	-------------	----------------------------------	----

Vector ID	Vector type	ID Value
SUNDIALS_NVEC_SERIAL	Serial	0
SUNDIALS_NVEC_PARALLEL	Distributed memory parallel (MPI)	1
SUNDIALS_NVEC_OPENMP	OpenMP shared memory parallel	2
SUNDIALS_NVEC_PTHREADS	PThreads shared memory parallel	3
SUNDIALS_NVEC_PARHYP	hypre ParHyp parallel vector	4
SUNDIALS_NVEC_PETSC	PETSc parallel vector	5
SUNDIALS_NVEC_OPENMPDEV	OpenMP shared memory parallel with device offloading	6
SUNDIALS_NVEC_CUSTOM	User-provided custom vector	7

#### void N\_VDestroyVectorArray(N\_Vector \*vs, int count);

and whose definition is based on the implementation-specific N\_VDestroy operation.

A particular implementation of the NVECTOR module must:

- Specify the *content* field of N\_Vector.
- Define and implement the vector operations. Note that the names of these routines should be unique to that implementation in order to permit using more than one NVECTOR module (each with different N\_Vector internal data representations) in the same code.
- Define and implement user-callable constructor and destructor routines to create and free an N\_Vector with the new *content* field and with *ops* pointing to the new vector operations.
- Optionally, define and implement additional user-callable routines acting on the newly defined N\_Vector (e.g., a routine to print the content for debugging purposes).
- Optionally, provide accessor macros as needed for that particular implementation to be used to access different parts in the *content* field of the newly defined N\_Vector.

Each NVECTOR implementation included in SUNDIALS has a unique identifier specified in enumeration and shown in Table 7.1. It is recommended that a user-supplied NVECTOR implementation use the SUNDIALS\_NVEC\_CUSTOM identifier.

Table 7.2: Description of the NVECTOR operations

Name	Usage and Description
N_VGetVectorID	id = N_VGetVectorID(w); Returns the vector type identifier for the vector w. It is used to determine the vector implementation type (e.g. serial, parallel,) from the abstract N_Vector interface. Returned values are given in Table 7.1.
N_VClone	<pre>v = N_VClone(w); Creates a new N_Vector of the same type as an existing vector w and sets the ops field. It does not copy the vector, but rather allocates storage for the new vector.</pre>
N_VCloneEmpty	<pre>v = N_VCloneEmpty(w); Creates a new N_Vector of the same type as an existing vector w and sets the ops field. It does not allocate storage for data.</pre>
${ t NVDestroy}$	N_VDestroy(v); Destroys the N_Vector v and frees memory allocated for its internal data.
N_VSpace	N_VSpace(nvSpec, &lrw, &liw); Returns storage requirements for one N_Vector. lrw contains the number of realtype words and liw contains the number of integer words. This function is advisory only, for use in determining a user's total space requirements; it could be a dummy function in a user-supplied NVECTOR module if that information is not of interest.
N_VGetArrayPointer	vdata = N_VGetArrayPointer(v); Returns a pointer to a realtype array from the N_Vector v. Note that this assumes that the internal data in N_Vector is a contiguous array of realtype. This routine is only used in the solver-specific interfaces to the dense and banded (serial) linear solvers, the sparse linear solvers (serial and threaded), and in the interfaces to the banded (serial) and band-block- diagonal (parallel) preconditioner modules provided with SUNDIALS.
N_VSetArrayPointer	N_VSetArrayPointer(vdata, v); Overwrites the data in an N_Vector with a given array of realtype. Note that this assumes that the internal data in N_Vector is a contiguous array of realtype. This routine is only used in the interfaces to the dense (serial) linear solver, hence need not exist in a user-supplied NVECTOR module for a parallel environment.
	continued on next page

continued from last page		
Name	Usage and Description	
N_VLinearSum	N_VLinearSum(a, x, b, y, z); Performs the operation $z = ax + by$ , where a and b are realtype scalars and x and y are of type N_Vector: $z_i = ax_i + by_i$ , $i = 0, \ldots, n-1$ .	
N_VConst	N_VConst(c, z); Sets all components of the N_Vector z to realtype c: $z_i=c, i=0,\ldots,n-1.$	
N_VProd	N_VProd(x, y, z); Sets the N_Vector z to be the component-wise product of the N_Vector inputs x and y: $z_i = x_i y_i$ , $i = 0, \ldots, n-1$ .	
N_VDiv	N_VDiv(x, y, z); Sets the N_Vector z to be the component-wise ratio of the N_Vector inputs x and y: $z_i = x_i/y_i$ , $i = 0, \ldots, n-1$ . The $y_i$ may not be tested for 0 values. It should only be called with a y that is guaranteed to have all nonzero components.	
N_VScale	N_VScale(c, x, z); Scales the N_Vector x by the realtype scalar c and returns the result in z: $z_i = cx_i$ , $i = 0,, n-1$ .	
N_VAbs	N_VAbs(x, z); Sets the components of the N_Vector z to be the absolute values of the components of the N_Vector x: $y_i =  x_i , i = 0, \ldots, n-1$ .	
N_VInv	N_VInv(x, z); Sets the components of the N_Vector z to be the inverses of the components of the N_Vector x: $z_i = 1.0/x_i$ , $i = 0, \ldots, n-1$ . This routine may not check for division by 0. It should be called only with an x which is guaranteed to have all nonzero components.	
$N_{-}VAddConst$	N_VAddConst(x, b, z); Adds the realtype scalar b to all components of x and returns the result in the N_Vector z: $z_i = x_i + b, i = 0, \ldots, n-1$ .	
N_VDotProd	d = N_VDotProd(x, y); Returns the value of the ordinary dot product of x and y: $d = \sum_{i=0}^{n-1} x_i y_i$ .	
N_VMaxNorm	m = N_VMaxNorm(x); Returns the maximum norm of the N_Vector x: $m = \max_i  x_i $ .	
	continued on next page	

continued from last page		
Name	Usage and Description	
N_VWrmsNorm	m = N_VWrmsNorm(x, w) Returns the weighted root-mean-square norm of the N_Vector x with realtype weight vector w: $m = \sqrt{\left(\sum_{i=0}^{n-1} (x_i w_i)^2\right)/n}$ .	
N_VWrmsNormMask	$ \begin{aligned} \mathbf{m} &= \mathbf{N}_{-}\mathbf{V}\mathbf{W}\mathbf{r}\mathbf{m}\mathbf{s}\mathbf{N}\mathbf{o}\mathbf{r}\mathbf{m}\mathbf{M}\mathbf{a}\mathbf{s}\mathbf{k}(\mathbf{x},\ \mathbf{w},\ \mathbf{id}); \\ \text{Returns the weighted root mean square norm of the } \mathbf{N}_{-}\mathbf{V}\mathbf{e}\mathbf{c}\mathbf{t}\mathbf{o}\mathbf{r}\ \mathbf{x}\ \mathbf{w}\mathbf{i}\mathbf{t}\mathbf{h} \\ \mathbf{realtype}\ \mathbf{weight}\ \mathbf{vector}\ \mathbf{w}\ \mathbf{built}\ \mathbf{u}\mathbf{s}\mathbf{i}\mathbf{n}\mathbf{g}\ \mathbf{o}\mathbf{n}\mathbf{l}\mathbf{y}\ \mathbf{t}\mathbf{h}\ \mathbf{e}\ \mathbf{e}\mathbf{l}\mathbf{e}\mathbf{m}\mathbf{e}\mathbf{n}\mathbf{t}\mathbf{s}\ \mathbf{o}\mathbf{f}\ \mathbf{x}\ \mathbf{c}\mathbf{o}\mathbf{r}\mathbf{r}\mathbf{s}\mathbf{p}\mathbf{o}\mathbf{n}\mathbf{d}\mathbf{i}\mathbf{n}\mathbf{g} \\ \mathbf{m} &= \sqrt{\left(\sum_{i=0}^{n-1}(x_iw_iH(id_i))^2\right)/n},\ \mathbf{w}\mathbf{h}\mathbf{e}\mathbf{r}\mathbf{e}\ \mathbf{H}(\alpha) = \begin{cases} 1 & \alpha > 0 \\ 0 & \alpha \leq 0 \end{cases} \end{aligned} $	
$N_{-}VMin$	$m = N_{\text{VMin}}(x);$ Returns the smallest element of the N_Vector x: $m = \min_{i} x_{i}.$	
N_VWL2Norm	m = N_VWL2Norm(x, w); Returns the weighted Euclidean $\ell_2$ norm of the N_Vector x with realtype weight vector w: $m = \sqrt{\sum_{i=0}^{n-1} (x_i w_i)^2}$ .	
N_VL1Norm	m = N_VL1Norm(x); Returns the $\ell_1$ norm of the N_Vector x: $m = \sum_{i=0}^{n-1}  x_i $ .	
N_VCompare	N_VCompare(c, x, z); Compares the components of the N_Vector x to the realtype scalar c and returns an N_Vector z such that: $z_i = 1.0$ if $ x_i  \ge c$ and $z_i = 0.0$ otherwise.	
N_VInvTest	t = N_VInvTest(x, z); Sets the components of the N_Vector z to be the inverses of the components of the N_Vector x, with prior testing for zero values: $z_i = 1.0/x_i$ , $i = 0, \ldots, n-1$ . This routine returns a boolean assigned to SUNTRUE if all components of x are nonzero (successful inversion) and returns SUNFALSE otherwise.	
N_VConstrMask	t = N_VConstrMask(c, x, m); Performs the following constraint tests: $x_i > 0$ if $c_i = 2$ , $x_i \ge 0$ if $c_i = 1$ , $x_i \le 0$ if $c_i = -1$ , $x_i < 0$ if $c_i = -2$ . There is no constraint on $x_i$ if $c_i = 0$ . This routine returns a boolean assigned to SUNFALSE if any element failed the constraint test and assigned to SUNTRUE if all passed. It also sets a mask vector m, with elements equal to 1.0 where the constraint test failed, and 0.0 where the test passed. This routine is used only for constraint checking.	
	continued on next page	

continued from last page		
Name	Usage and Description	
N_VMinQuotient	minq = $NVMinQuotient(num, denom)$ ; This routine returns the minimum of the quotients obtained by term-wise dividing $num_i$ by $denom_i$ . A zero element in $denom$ will be skipped. If no such quotients are found, then the large value $BIG_REAL$ (defined in the header file $sundials\_types.h$ ) is returned.	

Table 7.3: Description of the NVECTOR fused operations

Name	Usage and Description
N_VLinearCombination	ier = N_VLinearCombination(nv, c, X, z); This routine computes the linear combination of $n_v$ vectors with $n$ elements: $z_i = \sum_{j=0}^{n_v-1} c_j x_{j,i},  i=0,\ldots,n-1,$
	where $c$ is an array of $n_v$ scalars (type realtype*), $X$ is an array of $n_v$ vectors (type N_Vector*), and $z$ is the output vector (type N_Vector). If the output vector $z$ is one of the vectors in $X$ , then it $must$ be the first vector in the vector array. The operation returns 0 for success and a non-zero value otherwise.
N_VScaleAddMulti	ier = N_VScaleAddMulti(nv, c, x, Y, Z); This routine scales and adds one vector to $n_v$ vectors with $n$ elements: $z_{j,i} = c_j x_i + y_{j,i},  j = 0, \dots, n_v - 1  i = 0, \dots, n-1,$ where $c$ is an array of $n_v$ scalars (type realtype*), $x$ is the vector (type N_Vector) to be scaled and added to each vector in the vector array of $n_v$ vectors $Y$ (type N_Vector*), and $Z$ (type N_Vector*) is a vector array of $n_v$ output vectors. The operation returns 0 for success and a non-zero value otherwise.

continued from last page	
Name	Usage and Description
N_VDotProdMulti	ier = $N_VDotProdMulti(nv, x, Y, d)$ ; This routine computes the dot product of a vector with $n_v$ other vectors:
	$d_j = \sum_{i=0}^{n-1} x_i y_{j,i},  j = 0, \dots, n_v - 1,$
	where $d$ (type realtype*) is an array of $n_v$ scalars containing the dot products of the vector $x$ (type N_Vector) with each of the $n_v$ vectors in the vector array $Y$ (type N_Vector*). The operation returns 0 for
	success and a non-zero value otherwise.

Table 7.4: Description of the NVECTOR vector array operations

Name	Usage and Description
N_VLinearSumVectorArray  N_VScaleVectorArray	ier = N_VLinearSumVectorArray(nv, a, X, b, Y, Z); This routine comuptes the linear sum of two vector arrays containing $n_v$ vectors of $n$ elements: $z_{j,i} = ax_{j,i} + by_{j,i},  i = 0, \dots, n-1  j = 0, \dots, n_v - 1,$ where $a$ and $b$ are realtype scalars and $b$ , $b$ , and $b$ are arrays of $b$ , vectors (type N_Vector*). The operation returns 0 for success and a non-zero value otherwise.  ier = N_VScaleVectorArray(nv, c, X, Z); This routine scales each vector of $b$ elements in a vector array of $b$ array of $b$ vectors by a potentially different constant: $z_{j,i} = c_j x_{j,i},  i = 0, \dots, n-1  j = 0, \dots, n_v - 1,$ where $b$ is an array of $b$ vectors (type realtype*) and $b$ and $b$ are arrays of $b$ vectors (type N_Vector*). The operation returns 0 for success and a non-zero value otherwise.
	continued on next page

continued from last page	
Name	Usage and Description
N_VConstVectorArray	ier = N_VConstVectorArray(nv, c, X); This routine sets each element in a vector of $n$ elements in a vector array of $n_v$ vectors to the same value:
	$z_{j,i} = c,  i = 0, \dots, n-1  j = 0, \dots, n_v - 1,$
	where $c$ is a realtype scalar and $X$ is an array of $n_v$ vectors (type N_Vector*). The operation returns 0 for success and a non-zero value otherwise.
${\tt N\_VWrmsNormVectorArray}$	ier = N_VWrmsNormVectorArray(nv, X, W, m); This routine computes the weighted root mean square norm of $n_v$ vectors with $n$ elements:
	$m_j = \left(\frac{1}{n}\sum_{i=0}^{n-1} (x_{j,i}w_{j,i})^2\right)^{1/2},  j = 0, \dots, n_v - 1,$
	where $m$ (type realtype*) contains the $n_v$ norms of the vectors in the vector array $X$ (type N_Vector*) with corresponding weight vectors $W$ (type N_Vector*). The operation returns 0 for success and a non-zero value otherwise.
N_VWrmsNormMaskVectorArray	ier = N_VWrmsNormMaskVectorArray(nv, X, W, id, m); This routine computes the masked weighted root mean square norm of $n_v$ vectors with $n$ elements:
	$m_j = \left(\frac{1}{n}\sum_{i=0}^{n-1} (x_{j,i}w_{j,i}H(id_i))^2\right)^{1/2},  j = 0, \dots, n_v - 1,$
	$H(id_i) = 1$ for $id_i > 0$ and is zero otherwise, $m$ (type realtype*) contains the $n_v$ norms of the vectors in the vector array $X$ (type N_Vector*) with corresponding weight vectors $W$ (type N_Vector*) and mask vector $id_v$ (type N_Vector). The operation returns 0 for success and a non-zero value otherwise.
	continued on next pag

continued from last page	
Name	Usage and Description
N_VScaleAddMultiVectorArray	ier = N_VScaleAddMultiVectorArray(nv, ns, c, X, YY, ZZ); This routine scales and adds a vector in a vector array of $n_v$ vectors to the corresponding vector in $n_s$ vector arrays: $z_{j,i} = \sum_{k=0}^{n_s-1} c_k x_{k,j,i},  i=0,\ldots,n-1  j=0,\ldots,n_v-1,$ where $c$ is an array of $n_s$ scalars (type realtype*), $X$ is a vector array of $n_v$ vectors (type idN_Vector*) to be scaled and added to the corresponding vector in each of the $n_s$ vector arrays in the array of vector arrays $YY$ (type N_Vector**) and stored in the output array of vector arrays $ZZ$ (type N_Vector**). The operation returns 0 for success and a non-zero value otherwise.
$N_{-}VLinear Combination Vector Array$	ier = N_VLinearCombinationVectorArray(nv, ns, c, XX, Z); This routine computes the linear combination of $n_s$ vector arrays containing $n_v$ vectors with $n$ elements: $z_{j,i} = \sum_{k=0}^{n_s-1} c_k x_{k,j,i},  i=0,\ldots,n-1  j=0,\ldots,n_v-1,$ where $c$ is an array of $n_s$ scalars (type realtype*), $XX$ (type N_Vector**) is an array of $n_s$ vector arrays each containing $n_v$ vectors to be summed into the output vector array of $n_v$ vectors $Z$ (type N_Vector*). If the output
	vector array $Z$ is one of the vector arrays in $XX$ , then it $must$ be the first vector array in $XX$ . The operation returns 0 for success and a non-zero value otherwise.

### 7.1 NVECTOR functions used by IDAS

In Table 7.5 below, we list the vector functions used in the NVECTOR module used by the IDAS package. The table also shows, for each function, which of the code modules uses the function. The IDAS column shows function usage within the main integrator module, while the remaining columns show function usage within the IDAS linear solvers interface, the IDABBDPRE preconditioner module, and the IDAA module.

At this point, we should emphasize that the IDAS user does not need to know anything about the usage of vector functions by the IDAS code modules in order to use IDAS. The information is presented as an implementation detail for the interested reader.

Special cases (numbers match markings in table):

- 1. These routines are only required if an internal difference-quotient routine for constructing dense or band Jacobian matrices is used.
- 2. This routine is optional, and is only used in estimating space requirements for IDAS modules for user feedback.
- 3. The optional function N\_VDotProdMulti is only used when Classical Gram-Schmidt is enabled with SPGMR or SPFGMR. The remaining operations from Tables 7.3 and 7.4 not listed above are

unused and a user-supplied NVECTOR module for IDAS could omit these operations.

Of the functions listed in Table 7.2, N\_VWL2Norm, N\_VL1Norm, and N\_VInvTest are *not* used by IDAS. Therefore a user-supplied NVECTOR module for IDAS could omit these functions.

# 7.2 The NVECTOR\_SERIAL implementation

The serial implementation of the NVECTOR module provided with SUNDIALS, NVECTOR\_SERIAL, defines the *content* field of N\_Vector to be a structure containing the length of the vector, a pointer to the beginning of a contiguous data array, and a boolean flag *own\_data* which specifies the ownership of *data*.

```
struct _N_VectorContent_Serial {
   sunindextype length;
   booleantype own_data;
   realtype *data;
};
```

The header file to include when using this module is nvector\_serial.h. The installed module library to link to is libsundials\_nvecserial.lib where .lib is typically .so for shared libraries and .a for static libraries.

# 7.2.1 NVECTOR\_SERIAL accessor macros

The following macros are provided to access the content of an NVECTOR\_SERIAL vector. The suffix \_S in the names denotes the serial version.

#### • NV\_CONTENT\_S

This routine gives access to the contents of the serial vector N\_Vector.

The assignment  $v_{cont} = NV_{content_s(v)}$  sets  $v_{cont}$  to be a pointer to the serial  $N_{content}$  content structure.

Implementation:

```
#define NV_CONTENT_S(v) ( (N_VectorContent_Serial)(v->content) )
```

• NV\_OWN\_DATA\_S, NV\_DATA\_S, NV\_LENGTH\_S

These macros give individual access to the parts of the content of a serial N\_Vector.

The assignment  $v_{data} = NV_DATA_S(v)$  sets  $v_{data}$  to be a pointer to the first component of the data for the  $N_Vector v$ . The assignment  $NV_DATA_S(v) = v_{data}$  sets the component array of v to be  $v_{data}$  by storing the pointer  $v_{data}$ .

The assignment  $v_len = NV_LENGTH_S(v)$  sets  $v_len$  to be the length of v. On the other hand, the call  $NV_LENGTH_S(v) = len_v$  sets the length of v to be  $len_v$ .

Implementation:

```
#define NV_OWN_DATA_S(v) ( NV_CONTENT_S(v)->own_data )
#define NV_DATA_S(v) ( NV_CONTENT_S(v)->data )
#define NV_LENGTH_S(v) ( NV_CONTENT_S(v)->length )
```

#### • NV\_Ith\_S

This macro gives access to the individual components of the data array of an N-Vector.

The assignment  $r = NV_{i,i}$  sets r to be the value of the i-th component of v. The assignment  $NV_{i,i} = r$  sets the value of the i-th component of v to be r.

Here i ranges from 0 to n-1 for a vector of length n.

Implementation:

```
#define NV_Ith_S(v,i) ( NV_DATA_S(v)[i] )
```

# 7.2.2 NVECTOR\_SERIAL functions

The NVECTOR\_SERIAL module defines serial implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4. Their names are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix \_Serial (e.g. N\_VDestroy\_Serial). All the standard vector operations listed in 7.2 with the suffix \_Serial appended are callable via the FORTRAN 2003 interface by prepending an 'F' (e.g. FN\_VDestroy\_Serial).

The module NVECTOR\_SERIAL provides the following additional user-callable routines:

# N\_VNew\_Serial

Prototype N\_Vector N\_VNew\_Serial(sunindextype vec\_length);

Description This function creates and allocates memory for a serial N\_Vector. Its only argument is the vector length.

F2003 Name This function is callable as FN\_VNew\_Serial when using the Fortran 2003 interface module.

# N\_VNewEmpty\_Serial

Prototype N\_Vector N\_VNewEmpty\_Serial(sunindextype vec\_length);

Description This function creates a new serial N\_Vector with an empty (NULL) data array.

F2003 Name This function is callable as FN\_VNewEmpty\_Serial when using the Fortran 2003 interface module.

# N\_VMake\_Serial

Prototype N\_Vector N\_VMake\_Serial(sunindextype vec\_length, realtype \*v\_data);

Description This function creates and allocates memory for a serial vector with user-provided data

array.

(This function does *not* allocate memory for v\_data itself.)

F2003 Name This function is callable as FN\_VMake\_Serial when using the Fortran 2003 interface module.

# N\_VCloneVectorArray\_Serial

Prototype N\_Vector \*N\_VCloneVectorArray\_Serial(int count, N\_Vector w);

Description This function creates (by cloning) an array of count serial vectors.

# N\_VCloneVectorArrayEmpty\_Serial

Prototype N\_Vector \*N\_VCloneVectorArrayEmpty\_Serial(int count, N\_Vector w);

Description This function creates (by cloning) an array of count serial vectors, each with an empty (NULL) data array.

# N\_VDestroyVectorArray\_Serial

Prototype void N\_VDestroyVectorArray\_Serial(N\_Vector \*vs, int count);

Description This function frees memory allocated for the array of count variables of type N\_Vector

created with N\_VCloneVectorArray\_Serial or with

N\_VCloneVectorArrayEmpty\_Serial.

# N\_VGetLength\_Serial

Prototype sunindextype N\_VGetLength\_Serial(N\_Vector v);

Description This function returns the number of vector elements.

 $F2003\ Name\ This\ function\ is\ callable\ as\ {\tt FN\_VGetLength\_Serial}\ when\ using\ the\ Fortran\ 2003\ interface$ 

module.

# N\_VPrint\_Serial

Prototype void N\_VPrint\_Serial(N\_Vector v);

Description This function prints the content of a serial vector to stdout.

F2003 Name This function is callable as FN\_VPrint\_Serial when using the Fortran 2003 interface

module.

### N\_VPrintFile\_Serial

Prototype void N\_VPrintFile\_Serial(N\_Vector v, FILE \*outfile);

Description This function prints the content of a serial vector to outfile.

By default all fused and vector array operations are disabled in the NVECTOR\_SERIAL module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VNew\_Serial, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VNew\_Serial will have the default settings for the NVECTOR\_SERIAL module.

#### N\_VEnableFusedOps\_Serial

Prototype int N\_VEnableFusedOps\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array op-

erations in the serial vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

#### N\_VEnableLinearCombination\_Serial

Prototype int N\_VEnableLinearCombination\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused

operation in the serial vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

#### N\_VEnableScaleAddMulti\_Serial

Prototype int N\_VEnableScaleAddMulti\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to

multiple vectors fused operation in the serial vector. The return value is 0 for success

and -1 if the input vector or its ops structure are NULL.

# N\_VEnableDotProdMulti\_Serial

Prototype int N\_VEnableDotProdMulti\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused operation in the serial vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableLinearSumVectorArray\_Serial

Prototype int N\_VEnableLinearSumVectorArray\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for vector arrays in the serial vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableScaleVectorArray\_Serial

Prototype int N\_VEnableScaleVectorArray\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector arrays in the serial vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

#### N\_VEnableConstVectorArray\_Serial

Prototype int N\_VEnableConstVectorArray\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector arrays in the serial vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

### N\_VEnableWrmsNormVectorArray\_Serial

Prototype int N\_VEnableWrmsNormVectorArray\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for vector arrays in the serial vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

### N\_VEnableWrmsNormMaskVectorArray\_Serial

Prototype int N\_VEnableWrmsNormMaskVectorArray\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm operation for vector arrays in the serial vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableScaleAddMultiVectorArray\_Serial

Prototype int N\_VEnableScaleAddMultiVectorArray\_Serial(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array to multiple vector arrays operation in the serial vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableLinearCombinationVectorArray\_Serial

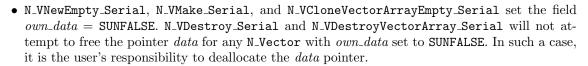
Prototype int N\_VEnableLinearCombinationVectorArray\_Serial(N\_Vector v, booleantype tf);

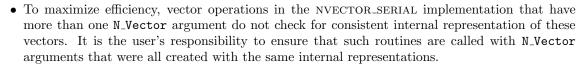
Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation

for vector arrays in the serial vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

### Notes

• When looping over the components of an N\_Vector v, it is more efficient to first obtain the component array via v\_data = NV\_DATA\_S(v) and then access v\_data[i] within the loop than it is to use NV\_Ith\_S(v,i) within the loop.





### 7.2.3 NVECTOR\_SERIAL Fortran interfaces

The NVECTOR\_SERIAL module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

#### FORTRAN 2003 interface module

The fnvector\_serial\_mod FORTRAN module defines interfaces to all NVECTOR\_SERIAL C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function N\_VNew\_Serial is interfaced as FN\_VNew\_Serial.

The FORTRAN 2003 NVECTOR\_SERIAL interface module can be accessed with the use statement, i.e. use fnvector\_serial\_mod, and linking to the library libsundials\_fnvectorserial\_mod.lib in addition to the C library. For details on where the library and module file fnvector\_serial\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fnvectorserial\_mod library.

### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the NVECTOR\_SERIAL module also includes a FORTRAN-callable function FNVINITS(code, NEQ, IER), to initialize this NVECTOR\_SERIAL module. Here code is an input solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, 4 for ARKODE); NEQ is the problem size (declared so as to match C type long int); and IER is an error return flag equal 0 for success and -1 for failure.

# 7.3 The NVECTOR\_PARALLEL implementation

The NVECTOR\_PARALLEL implementation of the NVECTOR module provided with SUNDIALS is based on MPI. It defines the *content* field of N\_Vector to be a structure containing the global and local lengths of the vector, a pointer to the beginning of a contiguous local data array, an MPI communicator, and a boolean flag own\_data indicating ownership of the data array data.





```
struct _N_VectorContent_Parallel {
   sunindextype local_length;
   sunindextype global_length;
   booleantype own_data;
   realtype *data;
   MPI_Comm comm;
};
```

The header file to include when using this module is nvector\_parallel.h. The installed module library to link to is libsundials\_nvecparallel.lib where .lib is typically .so for shared libraries and .a for static libraries.

### 7.3.1 NVECTOR\_PARALLEL accessor macros

The following macros are provided to access the content of a NVECTOR\_PARALLEL vector. The suffix \_P in the names denotes the distributed memory parallel version.

#### • NV CONTENT P

This macro gives access to the contents of the parallel vector N\_Vector.

The assignment  $v\_cont = NV\_CONTENT\_P(v)$  sets  $v\_cont$  to be a pointer to the  $N\_Vector$  content structure of type struct  $\_N\_VectorContent\_Parallel$ .

Implementation:

```
#define NV_CONTENT_P(v) ( (N_VectorContent_Parallel)(v->content) )
```

• NV\_OWN\_DATA\_P, NV\_DATA\_P, NV\_LOCLENGTH\_P, NV\_GLOBLENGTH\_P

These macros give individual access to the parts of the content of a parallel N\_Vector.

The assignment  $v_{data} = NV_DATA_P(v)$  sets  $v_{data}$  to be a pointer to the first component of the local data for the  $N_Vector v$ . The assignment  $NV_DATA_P(v) = v_{data}$  sets the component array of v to be  $v_{data}$  by storing the pointer  $v_{data}$ .

The assignment v\_llen = NV\_LOCLENGTH\_P(v) sets v\_llen to be the length of the local part of v. The call NV\_LENGTH\_P(v) =  $llen_v$  sets the local length of v to be  $llen_v$ .

The assignment  $v\_glen = NV\_GLOBLENGTH\_P(v)$  sets  $v\_glen$  to be the global length of the vector v. The call  $NV\_GLOBLENGTH\_P(v) = glen\_v$  sets the global length of v to be  $glen\_v$ .

Implementation:

```
#define NV_OWN_DATA_P(v) ( NV_CONTENT_P(v)->own_data )
#define NV_DATA_P(v) ( NV_CONTENT_P(v)->data )
#define NV_LOCLENGTH_P(v) ( NV_CONTENT_P(v)->local_length )
#define NV_GLOBLENGTH_P(v) ( NV_CONTENT_P(v)->global_length )
```

#### NV\_COMM\_P

This macro provides access to the MPI communicator used by the NVECTOR\_PARALLEL vectors. Implementation:

```
#define NV_COMM_P(v) ( NV_CONTENT_P(v)->comm )
```

# • NV\_Ith\_P

This macro gives access to the individual components of the local data array of an N-Vector.

The assignment  $r = NV_{in}(v,i)$  sets r to be the value of the i-th component of the local part of v. The assignment  $NV_{in}(v,i) = r$  sets the value of the i-th component of the local part of v to be r.

Here i ranges from 0 to n-1, where n is the local length.

Implementation:

```
#define NV_Ith_P(v,i) ( NV_DATA_P(v)[i] )
```

# 7.3.2 NVECTOR\_PARALLEL functions

The NVECTOR\_PARALLEL module defines parallel implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4. Their names are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix \_Parallel (e.g. N\_VDestroy\_Parallel). The module NVECTOR\_PARALLEL provides the following additional user-callable routines:

#### N\_VNew\_Parallel

Prototype N\_Vector N\_VNew\_Parallel(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length);

Description This function creates and allocates memory for a parallel vector.

# N\_VNewEmpty\_Parallel

Prototype N\_Vector N\_VNewEmpty\_Parallel(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length);

Description This function creates a new parallel N\_Vector with an empty (NULL) data array.

# N\_VMake\_Parallel

Prototype N\_Vector N\_VMake\_Parallel(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length, realtype \*v\_data);

Description This function creates and allocates memory for a parallel vector with user-provided data array. This function does *not* allocate memory for v\_data itself.

# N\_VCloneVectorArray\_Parallel

Prototype N\_Vector \*N\_VCloneVectorArray\_Parallel(int count, N\_Vector w);

Description This function creates (by cloning) an array of count parallel vectors.

### N\_VCloneVectorArrayEmpty\_Parallel

Prototype N\_Vector \*N\_VCloneVectorArrayEmpty\_Parallel(int count, N\_Vector w);

Description This function creates (by cloning) an array of count parallel vectors, each with an empty (NULL) data array.

# N\_VDestroyVectorArray\_Parallel

Prototype void N\_VDestroyVectorArray\_Parallel(N\_Vector \*vs, int count);

Description This function frees memory allocated for the array of count variables of type N\_Vector

created with  ${\tt N\_VCloneVectorArray\_Parallel}$  or with

 ${\tt N\_VCloneVectorArrayEmpty\_Parallel}.$ 

#### N\_VGetLength\_Parallel

Prototype sunindextype N\_VGetLength\_Parallel(N\_Vector v);

Description This function returns the number of vector elements (global vector length).

# N\_VGetLocalLength\_Parallel

Prototype sunindextype N\_VGetLocalLength\_Parallel(N\_Vector v);

Description This function returns the local vector length.

# N\_VPrint\_Parallel

Prototype void N\_VPrint\_Parallel(N\_Vector v);

Description This function prints the local content of a parallel vector to stdout.

#### N\_VPrintFile\_Parallel

Prototype void N\_VPrintFile\_Parallel(N\_Vector v, FILE \*outfile);

Description This function prints the local content of a parallel vector to outfile.

By default all fused and vector array operations are disabled in the NVECTOR\_PARALLEL module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VNew\_Parallel, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone with that vector. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VNew\_Parallel will have the default settings for the NVECTOR\_PARALLEL module.

### N\_VEnableFusedOps\_Parallel

Prototype int N\_VEnableFusedOps\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array operations in the parallel vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

#### N\_VEnableLinearCombination\_Parallel

Prototype int N\_VEnableLinearCombination\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused operation in the parallel vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableScaleAddMulti\_Parallel

Prototype int N\_VEnableScaleAddMulti\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to

multiple vectors fused operation in the parallel vector. The return value is  $\boldsymbol{0}$  for success

and -1 if the input vector or its ops structure are NULL.

# N\_VEnableDotProdMulti\_Parallel

Prototype int N\_VEnableDotProdMulti\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused

operation in the parallel vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

### N\_VEnableLinearSumVectorArray\_Parallel

Prototype int N\_VEnableLinearSumVectorArray\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for vector arrays in the parallel vector. The return value is 0 for success and -1 if the input

vector arrays in the parallel vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableScaleVectorArray\_Parallel

Prototype int N\_VEnableScaleVectorArray\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector arrays in the parallel vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableConstVectorArray\_Parallel

Prototype int N\_VEnableConstVectorArray\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector

arrays in the parallel vector. The return value is 0 for success and -1 if the input vector

or its  $\ensuremath{\mathsf{ops}}$  structure are NULL.

# N\_VEnableWrmsNormVectorArray\_Parallel

Prototype int N\_VEnableWrmsNormVectorArray\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for

vector arrays in the parallel vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableWrmsNormMaskVectorArray\_Parallel

Prototype int N\_VEnableWrmsNormMaskVectorArray\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm op-

eration for vector arrays in the parallel vector. The return value is 0 for success and -1

if the input vector or its ops structure are NULL.

#### N\_VEnableScaleAddMultiVectorArray\_Parallel

Prototype int N\_VEnableScaleAddMultiVectorArray\_Parallel(N\_Vector v,

booleantype tf);

 $\hbox{Description This function enables ($\tt SUNTRUE)$ or disables ($\tt SUNFALSE)$ the scale and add a vector $\tt SUNFALSE. The scale and add a vector {\tt SUNFALSE}$ is a scale and add a vector {\tt SUNFALSE}$ and {\tt SUNFALSE}$ is a scale and add a vector {\tt SUNFALSE}$. The scale and add a vector {\tt SUNFALSE}$ is a scale and add a vector {\tt SUNFALSE}$ is a scale and add a vector {\tt SUNFALSE}$. The scale and add a vector {\tt SUNFALSE}$ is a scale and add a vector {\tt SUNFALSE}$ is a scale and add a vector {\tt SUNFALSE}$. The scale and {\tt SUNFALSE}$ is a scale and {\tt SUNFA$ 

array to multiple vector arrays operation in the parallel vector. The return value is 0

for success and -1 if the input vector or its ops structure are NULL.

### N\_VEnableLinearCombinationVectorArray\_Parallel

Prototype int N\_VEnableLinearCombinationVectorArray\_Parallel(N\_Vector v, booleantype tf);

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation

for vector arrays in the parallel vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

Notes

- When looping over the components of an N\_Vector v, it is more efficient to first obtain the local component array via v\_data = NV\_DATA\_P(v) and then access v\_data[i] within the loop than it is to use NV\_Ith\_P(v,i) within the loop.
- N\_VNewEmpty\_Parallel, N\_VMake\_Parallel, and N\_VCloneVectorArrayEmpty\_Parallel set the field own\_data = SUNFALSE. N\_VDestroy\_Parallel and N\_VDestroyVectorArray\_Parallel will not attempt to free the pointer data for any N\_Vector with own\_data set to SUNFALSE. In such a case, it is the user's responsibility to deallocate the data pointer.





• To maximize efficiency, vector operations in the NVECTOR\_PARALLEL implementation that have more than one N\_Vector argument do not check for consistent internal representation of these vectors. It is the user's responsibility to ensure that such routines are called with N\_Vector arguments that were all created with the same internal representations.

#### 7.3.3 NVECTOR\_PARALLEL Fortran interfaces

For solvers that include a FORTRAN 77 interface module, the NVECTOR\_PARALLEL module also includes a FORTRAN-callable function FNVINITP(COMM, code, NLOCAL, NGLOBAL, IER), to initialize this NVECTOR\_PARALLEL module. Here COMM is the MPI communicator, code is an input solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, 4 for ARKODE); NLOCAL and NGLOBAL are the local and global vector sizes, respectively (declared so as to match C type long int); and IER is an error return flag equal 0 for success and -1 for failure. NOTE: If the header file sundials\_config.h defines SUNDIALS\_MPI\_COMM\_F2C to be 1 (meaning the MPI implementation used to build SUNDIALS includes the MPI\_Comm\_f2c function), then COMM can be any valid MPI communicator. Otherwise, MPI\_COMM\_WORLD will be used, so just pass an integer value as a placeholder.

# 7.4 The NVECTOR\_OPENMP implementation

In situations where a user has a multi-core processing unit capable of running multiple parallel threads with shared memory, SUNDIALS provides an implementation of NVECTOR using OpenMP, called NVECTOR\_OPENMP, and an implementation using Pthreads, called NVECTOR\_PTHREADS. Testing has shown that vectors should be of length at least 100,000 before the overhead associated with creating and using the threads is made up by the parallelism in the vector calculations.

The OpenMP NVECTOR implementation provided with SUNDIALS, NVECTOR\_OPENMP, defines the content field of N\_Vector to be a structure containing the length of the vector, a pointer to the beginning of a contiguous data array, a boolean flag own\_data which specifies the ownership of data, and the number of threads. Operations on the vector are threaded using OpenMP.

```
struct _N_VectorContent_OpenMP {
   sunindextype length;
   booleantype own_data;
   realtype *data;
   int num_threads;
};
```

The header file to include when using this module is nvector\_openmp.h. The installed module library to link to is libsundials\_nvecopenmp.lib where .lib is typically .so for shared libraries and .a for static libraries. The FORTRAN module file to use when using the FORTRAN 2003 interface to this module is fnvector\_openmp\_mod.mod.

# 7.4.1 NVECTOR\_OPENMP accessor macros

The following macros are provided to access the content of an NVECTOR\_OPENMP vector. The suffix \_OMP in the names denotes the OpenMP version.

# • NV\_CONTENT\_OMP

This routine gives access to the contents of the OpenMP vector N\_Vector.

The assignment  $v\_cont = NV\_CONTENT\_OMP(v)$  sets  $v\_cont$  to be a pointer to the OpenMP  $N\_Vector$  content structure.

Implementation:

```
#define NV_CONTENT_OMP(v) ( (N_VectorContent_OpenMP)(v->content) )
```



• NV\_OWN\_DATA\_OMP, NV\_DATA\_OMP, NV\_LENGTH\_OMP, NV\_NUM\_THREADS\_OMP

These macros give individual access to the parts of the content of a OpenMP N\_Vector.

The assignment  $v_{data} = NV_DATA_OMP(v)$  sets  $v_{data}$  to be a pointer to the first component of the data for the  $N_Vector v$ . The assignment  $NV_DATA_OMP(v) = v_{data}$  sets the component array of v to be  $v_{data}$  by storing the pointer  $v_{data}$ .

The assignment  $v_len = NV_length_OMP(v)$  sets  $v_len$  to be the length of v. On the other hand, the call  $NV_length_OMP(v) = len_v$  sets the length of v to be  $len_v$ .

The assignment v\_num\_threads = NV\_NUM\_THREADS\_OMP(v) sets v\_num\_threads to be the number of threads from v. On the other hand, the call NV\_NUM\_THREADS\_OMP(v) = num\_threads\_v sets the number of threads for v to be num\_threads\_v.

Implementation:

```
#define NV_OWN_DATA_OMP(v) ( NV_CONTENT_OMP(v)->own_data )
#define NV_DATA_OMP(v) ( NV_CONTENT_OMP(v)->data )
#define NV_LENGTH_OMP(v) ( NV_CONTENT_OMP(v)->length )
#define NV_NUM_THREADS_OMP(v) ( NV_CONTENT_OMP(v)->num_threads )
```

• NV\_Ith\_OMP

This macro gives access to the individual components of the data array of an N\_Vector.

The assignment  $r = NV_{int}(v,i)$  sets r to be the value of the i-th component of v. The assignment  $NV_{int}(v,i) = r$  sets the value of the i-th component of v to be r.

Here i ranges from 0 to n-1 for a vector of length n.

Implementation:

```
#define NV_Ith_OMP(v,i) ( NV_DATA_OMP(v)[i] )
```

# 7.4.2 NVECTOR OPENMP functions

The NVECTOR\_OPENMP module defines OpenMP implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4. Their names are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix \_OpenMP (e.g. N\_VDestroy\_OpenMP). All the standard vector operations listed in 7.2 with the suffix \_OpenMP appended are callable via the FORTRAN 2003 interface by prepending an 'F' (e.g. FN\_VDestroy\_OpenMP).

The module NVECTOR\_OPENMP provides the following additional user-callable routines:

# N\_VNew\_OpenMP

Prototype N\_Vector N\_VNew\_OpenMP(sunindextype vec\_length, int num\_threads)

Description This function creates and allocates memory for a OpenMP N\_Vector. Arguments are the vector length and number of threads.

F2003 Name This function is callable as FN\_VNew\_OpenMP when using the Fortran 2003 interface module.

# N\_VNewEmpty\_OpenMP

Prototype N\_Vector N\_VNewEmpty\_OpenMP(sunindextype vec\_length, int num\_threads)

Description This function creates a new OpenMP N\_Vector with an empty (NULL) data array.

F2003 Name This function is callable as FN\_VNewEmpty\_OpenMP when using the Fortran 2003 interface module.

# N\_VMake\_OpenMP

Prototype N\_Vector N\_VMake\_OpenMP(sunindextype vec\_length, realtype \*v\_data, int num\_threads);

Description This function creates and allocates memory for a OpenMP vector with user-provided

data array. This function does not allocate memory for v\_data itself.

 ${\tt F2003~Name~This~function~is~callable~as~{\tt FN\_VMake\_OpenMP~when~using~the~Fortran~2003~interface}}$ 

module.

# N\_VCloneVectorArray\_OpenMP

Prototype N\_Vector \*N\_VCloneVectorArray\_OpenMP(int count, N\_Vector w)

Description This function creates (by cloning) an array of count OpenMP vectors.

# N\_VCloneVectorArrayEmpty\_OpenMP

Prototype N\_Vector \*N\_VCloneVectorArrayEmpty\_OpenMP(int count, N\_Vector w)

Description This function creates (by cloning) an array of count OpenMP vectors, each with an

empty (NULL) data array.

# N\_VDestroyVectorArray\_OpenMP

Prototype void N\_VDestroyVectorArray\_OpenMP(N\_Vector \*vs, int count)

Description This function frees memory allocated for the array of count variables of type N\_Vector

created with N\_VCloneVectorArray\_OpenMP or with N\_VCloneVectorArrayEmpty\_OpenMP.

### N\_VGetLength\_OpenMP

Prototype sunindextype N\_VGetLength\_OpenMP(N\_Vector v)

Description This function returns number of vector elements.

F2003 Name This function is callable as FN\_VGetLength\_OpenMP when using the Fortran 2003 interface

module.

#### N\_VPrint\_OpenMP

Prototype void N\_VPrint\_OpenMP(N\_Vector v)

Description This function prints the content of an OpenMP vector to stdout.

F2003 Name This function is callable as FN\_VPrint\_OpenMP when using the Fortran 2003 interface

module.

# N\_VPrintFile\_OpenMP

Prototype void N\_VPrintFile\_OpenMP(N\_Vector v, FILE \*outfile)

Description This function prints the content of an OpenMP vector to outfile.

By default all fused and vector array operations are disabled in the NVECTOR\_OPENMP module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VNew\_OpenMP, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VNew\_OpenMP will have the default settings for the NVECTOR\_OPENMP module.

# N\_VEnableFusedOps\_OpenMP

Prototype int N\_VEnableFusedOps\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array op-

erations in the OpenMP vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableLinearCombination\_OpenMP

Prototype int N\_VEnableLinearCombination\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused

operation in the OpenMP vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

#### N\_VEnableScaleAddMulti\_OpenMP

Prototype int N\_VEnableScaleAddMulti\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to

multiple vectors fused operation in the OpenMP vector. The return value is  $\boldsymbol{0}$  for success

and -1 if the input vector or its ops structure are NULL.

### N\_VEnableDotProdMulti\_OpenMP

Prototype int N\_VEnableDotProdMulti\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused

operation in the OpenMP vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

### ${\tt N\_VEnableLinearSumVectorArray\_OpenMP}$

Prototype int N\_VEnableLinearSumVectorArray\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for

vector arrays in the OpenMP vector. The return value is 0 for success and -1 if the

input vector or its ops structure are NULL.

# N\_VEnableScaleVectorArray\_OpenMP

Prototype int N\_VEnableScaleVectorArray\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector

arrays in the OpenMP vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableConstVectorArray\_OpenMP

Prototype int N\_VEnableConstVectorArray\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector

arrays in the OpenMP vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableWrmsNormVectorArray\_OpenMP

Prototype int N\_VEnableWrmsNormVectorArray\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for vector arrays in the OpenMP vector. The return value is 0 for success and -1 if the

input vector or its ops structure are NULL.

# N\_VEnableWrmsNormMaskVectorArray\_OpenMP

Prototype int N\_VEnableWrmsNormMaskVectorArray\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm operation for vector arrays in the OpenMP vector. The return value is 0 for success and

-1 if the input vector or its ops structure are NULL.

# ${\tt N\_VEnableScaleAddMultiVectorArray\_OpenMP}$

Prototype int N\_VEnableScaleAddMultiVectorArray\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array to multiple vector arrays operation in the OpenMP vector. The return value is 0 for

success and -1 if the input vector or its ops structure are NULL.

#### N\_VEnableLinearCombinationVectorArray\_OpenMP

Prototype int N\_VEnableLinearCombinationVectorArray\_OpenMP(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation

for vector arrays in the OpenMP vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

Notes

• When looping over the components of an N\_Vector v, it is more efficient to first obtain the component array via v\_data = NV\_DATA\_OMP(v) and then access v\_data[i] within the loop than it is to use NV\_Ith\_OMP(v,i) within the loop.



• N\_VNewEmpty\_OpenMP, N\_VMake\_OpenMP, and N\_VCloneVectorArrayEmpty\_OpenMP set the field  $own\_data = SUNFALSE$ . N\_VDestroy\_OpenMP and N\_VDestroyVectorArray\_OpenMP will not attempt to free the pointer data for any N\_Vector with  $own\_data$  set to SUNFALSE. In such a case, it is the user's responsibility to deallocate the data pointer.



• To maximize efficiency, vector operations in the NVECTOR\_OPENMP implementation that have more than one N\_Vector argument do not check for consistent internal representation of these vectors. It is the user's responsibility to ensure that such routines are called with N\_Vector arguments that were all created with the same internal representations.

# 7.4.3 NVECTOR\_OPENMP Fortran interfaces

The NVECTOR\_OPENMP module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

#### FORTRAN 2003 interface module

The nvector\_openmp\_mod FORTRAN module defines interfaces to most NVECTOR\_OPENMP C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function N\_VNew\_OpenMP is interfaced as FN\_VNew\_OpenMP.

The FORTRAN 2003 NVECTOR\_OPENMP interface module can be accessed with the use statement, i.e. use fnvector\_openmp\_mod, and linking to the library libsundials\_fnvectoropenmp\_mod.lib in addition to the C library. For details on where the library and module file fnvector\_openmp\_mod.mod are installed see Appendix A.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the NVECTOR\_OPENMP module also includes a FORTRAN-callable function FNVINITOMP(code, NEQ, NUMTHREADS, IER), to initialize this module. Here code is an input solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, 4 for ARKODE); NEQ is the problem size (declared so as to match C type long int); NUMTHREADS is the number of threads; and IER is an error return flag equal 0 for success and -1 for failure.

# 7.5 The NVECTOR\_PTHREADS implementation

In situations where a user has a multi-core processing unit capable of running multiple parallel threads with shared memory, SUNDIALS provides an implementation of NVECTOR using OpenMP, called NVECTOR\_OPENMP, and an implementation using Pthreads, called NVECTOR\_PTHREADS. Testing has shown that vectors should be of length at least 100,000 before the overhead associated with creating and using the threads is made up by the parallelism in the vector calculations.

The Pthreads NVECTOR implementation provided with SUNDIALS, denoted NVECTOR\_PTHREADS, defines the *content* field of N\_Vector to be a structure containing the length of the vector, a pointer to the beginning of a contiguous data array, a boolean flag  $own\_data$  which specifies the ownership of data, and the number of threads. Operations on the vector are threaded using POSIX threads (Pthreads).

```
struct _N_VectorContent_Pthreads {
   sunindextype length;
   booleantype own_data;
   realtype *data;
   int num_threads;
};
```

The header file to include when using this module is nvector\_pthreads.h. The installed module library to link to is libsundials\_nvecpthreads.lib where .lib is typically .so for shared libraries and .a for static libraries.

# 7.5.1 NVECTOR\_PTHREADS accessor macros

The following macros are provided to access the content of an NVECTOR\_PTHREADS vector. The suffix \_PT in the names denotes the Pthreads version.

# • NV\_CONTENT\_PT

This routine gives access to the contents of the Pthreads vector N\_Vector.

The assignment  $v\_cont = NV\_CONTENT\_PT(v)$  sets  $v\_cont$  to be a pointer to the Pthreads  $N\_Vector$  content structure.

Implementation:

```
#define NV_CONTENT_PT(v) ( (N_VectorContent_Pthreads)(v->content) )
```

### • NV\_OWN\_DATA\_PT, NV\_DATA\_PT, NV\_LENGTH\_PT, NV\_NUM\_THREADS\_PT

These macros give individual access to the parts of the content of a Pthreads N\_Vector.

The assignment v\_data = NV\_DATA\_PT(v) sets v\_data to be a pointer to the first component of the data for the N\_Vector v. The assignment NV\_DATA\_PT(v) = v\_data sets the component array of v to be v\_data by storing the pointer v\_data.

The assignment  $v_len = NV_LENGTH_PT(v)$  sets  $v_len$  to be the length of v. On the other hand, the call  $NV_LENGTH_PT(v) = len_v$  sets the length of v to be  $len_v$ .

The assignment v\_num\_threads = NV\_NUM\_THREADS\_PT(v) sets v\_num\_threads to be the number of threads from v. On the other hand, the call NV\_NUM\_THREADS\_PT(v) = num\_threads\_v sets the number of threads for v to be num\_threads\_v.

Implementation:

```
#define NV_OWN_DATA_PT(v) ( NV_CONTENT_PT(v)->own_data )
#define NV_DATA_PT(v) ( NV_CONTENT_PT(v)->data )
#define NV_LENGTH_PT(v) ( NV_CONTENT_PT(v)->length )
#define NV_NUM_THREADS_PT(v) ( NV_CONTENT_PT(v)->num_threads )
```

#### • NV\_Ith\_PT

This macro gives access to the individual components of the data array of an N\_Vector.

The assignment  $r = NV_{int}PT(v,i)$  sets r to be the value of the i-th component of v. The assignment  $NV_{int}PT(v,i) = r$  sets the value of the i-th component of v to be r.

Here i ranges from 0 to n-1 for a vector of length n.

Implementation:

```
#define NV_Ith_PT(v,i) ( NV_DATA_PT(v)[i] )
```

### 7.5.2 NVECTOR\_PTHREADS functions

The NVECTOR\_PTHREADS module defines Pthreads implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4. Their names are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix \_Pthreads (e.g. N\_VDestroy\_Pthreads). All the standard vector operations listed in 7.2 are callable via the FORTRAN 2003 interface by prepending an 'F' (e.g. FN\_VDestroy\_Pthreads). The module NVECTOR\_PTHREADS provides the following additional user-callable routines:

# N\_VNew\_Pthreads

Prototype N\_Vector N\_VNew\_Pthreads(sunindextype vec\_length, int num\_threads)

Description This function creates and allocates memory for a Pthreads N\_Vector. Arguments are the vector length and number of threads.

F2003 Name This function is callable as FN\_VNew\_Pthreads when using the Fortran 2003 interface module.

#### N\_VNewEmpty\_Pthreads

Prototype N\_Vector N\_VNewEmpty\_Pthreads(sunindextype vec\_length, int num\_threads)

Description This function creates a new Pthreads N\_Vector with an empty (NULL) data array.

F2003 Name This function is callable as FN\_VNewEmpty\_Pthreads when using the Fortran 2003 interface module.

### N\_VMake\_Pthreads

Prototype N\_Vector N\_VMake\_Pthreads(sunindextype vec\_length, realtype \*v\_data, int num\_threads);

Description This function creates and allocates memory for a Pthreads vector with user-provided data array. This function does *not* allocate memory for v\_data itself.

F2003 Name This function is callable as FN\_VMake\_Pthreads when using the Fortran 2003 interface

# N\_VCloneVectorArray\_Pthreads

Prototype N\_Vector \*N\_VCloneVectorArray\_Pthreads(int count, N\_Vector w)
Description This function creates (by cloning) an array of count Pthreads vectors.

# N\_VCloneVectorArrayEmpty\_Pthreads

Prototype N\_Vector \*N\_VCloneVectorArrayEmpty\_Pthreads(int count, N\_Vector w)

Description This function creates (by cloning) an array of count Pthreads vectors, each with an empty (NULL) data array.

# N\_VDestroyVectorArray\_Pthreads

Prototype void N\_VDestroyVectorArray\_Pthreads(N\_Vector \*vs, int count)

Description This function frees memory allocated for the array of count variables of type N\_Vector

created with N\_VCloneVectorArray\_Pthreads or with

 ${\tt N\_VCloneVectorArrayEmpty\_Pthreads}.$ 

#### N\_VGetLength\_Pthreads

Prototype sunindextype N\_VGetLength\_Pthreads(N\_Vector v)
Description This function returns the number of vector elements.

F2003 Name This function is callable as FN\_VGetLength\_Pthreads when using the Fortran 2003 in-

terface module.

# N\_VPrint\_Pthreads

Prototype void N\_VPrint\_Pthreads(N\_Vector v)

Description This function prints the content of a Pthreads vector to stdout.

F2003 Name This function is callable as FN\_VPrint\_Pthreads when using the Fortran 2003 interface

module.

# N\_VPrintFile\_Pthreads

Prototype void N\_VPrintFile\_Pthreads(N\_Vector v, FILE \*outfile)

Description This function prints the content of a Pthreads vector to outfile.

By default all fused and vector array operations are disabled in the NVECTOR\_PTHREADS module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VNew\_Pthreads, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VNew\_Pthreads will have the default settings for the NVECTOR\_PTHREADS module.

# N\_VEnableFusedOps\_Pthreads

Prototype int N\_VEnableFusedOps\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array op-

erations in the Pthreads vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableLinearCombination\_Pthreads

Prototype int N\_VEnableLinearCombination\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused

operation in the Pthreads vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

#### N\_VEnableScaleAddMulti\_Pthreads

Prototype int N\_VEnableScaleAddMulti\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to

multiple vectors fused operation in the Pthreads vector. The return value is  $\boldsymbol{0}$  for success

and -1 if the input vector or its ops structure are NULL.

#### N\_VEnableDotProdMulti\_Pthreads

Prototype int N\_VEnableDotProdMulti\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused

operation in the Pthreads vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

### N\_VEnableLinearSumVectorArray\_Pthreads

Prototype int N\_VEnableLinearSumVectorArray\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for

vector arrays in the Pthreads vector. The return value is 0 for success and -1 if the

input vector or its ops structure are NULL.

### N\_VEnableScaleVectorArray\_Pthreads

Prototype int N\_VEnableScaleVectorArray\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector

arrays in the Pthreads vector. The return value is  ${\tt 0}$  for success and  ${\tt -1}$  if the input

vector or its ops structure are NULL.

#### N\_VEnableConstVectorArray\_Pthreads

Prototype int N\_VEnableConstVectorArray\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector

arrays in the Pthreads vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableWrmsNormVectorArray\_Pthreads

Prototype int N\_VEnableWrmsNormVectorArray\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for vector arrays in the Pthreads vector. The return value is 0 for success and -1 if the

input vector or its ops structure are NULL.

# ${\tt N\_VEnableWrmsNormMaskVectorArray\_Pthreads}$

Prototype int N\_VEnableWrmsNormMaskVectorArray\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm operation for vector arrays in the Pthreads vector. The return value is 0 for success and

-1 if the input vector or its ops structure are NULL.

# N\_VEnableScaleAddMultiVectorArray\_Pthreads

Prototype int N\_VEnableScaleAddMultiVectorArray\_Pthreads(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array

to multiple vector arrays operation in the Pthreads vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

N\_VEnableLinearCombinationVectorArray\_Pthreads

 $\label{prototype} Prototype \quad \text{ int $N_V$EnableLinearCombinationVectorArray\_Pthreads($N_V$ector $v$, }$ 

booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation for vector arrays in the Pthreads vector. The return value is 0 for success and -1 if the

input vector or its ops structure are NULL.

#### Notes

- When looping over the components of an N\_Vector v, it is more efficient to first obtain the component array via v\_data = NV\_DATA\_PT(v) and then access v\_data[i] within the loop than it is to use NV\_Ith\_PT(v,i) within the loop.
- N\_VNewEmpty\_Pthreads, N\_VMake\_Pthreads, and N\_VCloneVectorArrayEmpty\_Pthreads set the field own\_data = SUNFALSE. N\_VDestroy\_Pthreads and N\_VDestroyVectorArray\_Pthreads will not attempt to free the pointer data for any N\_Vector with own\_data set to SUNFALSE. In such a case, it is the user's responsibility to deallocate the data pointer.
- To maximize efficiency, vector operations in the NVECTOR\_PTHREADS implementation that have more than one N\_Vector argument do not check for consistent internal representation of these vectors. It is the user's responsibility to ensure that such routines are called with N\_Vector arguments that were all created with the same internal representations.

# 7.5.3 NVECTOR\_PTHREADS Fortran interfaces

The NVECTOR\_PTHREADS module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.





#### FORTRAN 2003 interface module

The nvector\_pthreads\_mod FORTRAN module defines interfaces to most NVECTOR\_PTHREADS C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function N\_VNew\_Pthreads is interfaced as FN\_VNew\_Pthreads.

The FORTRAN 2003 NVECTOR\_PTHREADS interface module can be accessed with the use statement, i.e. use fnvector\_pthreads\_mod, and linking to the library libsundials\_fnvectorpthreads\_mod.lib in addition to the C library. For details on where the library and module file fnvector\_pthreads\_mod.mod are installed see Appendix A.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN interface module, the NVECTOR\_PTHREADS module also includes a FORTRAN-callable function FNVINITPTS(code, NEQ, NUMTHREADS, IER), to initialize this module. Here code is an input solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, 4 for ARKODE); NEQ is the problem size (declared so as to match C type long int); NUMTHREADS is the number of threads; and IER is an error return flag equal 0 for success and -1 for failure.

# 7.6 The NVECTOR\_PARHYP implementation

The NVECTOR\_PARHYP implementation of the NVECTOR module provided with SUNDIALS is a wrapper around *hypre*'s ParVector class. Most of the vector kernels simply call *hypre* vector operations. The implementation defines the *content* field of N\_Vector to be a structure containing the global and local lengths of the vector, a pointer to an object of type HYPRE\_ParVector, an MPI communicator, and a boolean flag *own\_parvector* indicating ownership of the *hypre* parallel vector object x.

```
struct _N_VectorContent_ParHyp {
   sunindextype local_length;
   sunindextype global_length;
   booleantype own_parvector;
   MPI_Comm comm;
   HYPRE_ParVector x;
};
```

The header file to include when using this module is nvector\_parhyp.h. The installed module library to link to is libsundials\_nvecparhyp.lib where .lib is typically .so for shared libraries and .a for static libraries.

Unlike native SUNDIALS vector types, NVECTOR\_PARHYP does not provide macros to access its member variables. Note that NVECTOR\_PARHYP requires SUNDIALS to be built with MPI support.

# 7.6.1 NVECTOR\_PARHYP functions

The NVECTOR\_PARHYP module defines implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4, except for N\_VSetArrayPointer and N\_VGetArrayPointer, because accessing raw vector data is handled by low-level hypre functions. As such, this vector is not available for use with SUNDIALS Fortran interfaces. When access to raw vector data is needed, one should extract the hypre vector first, and then use hypre methods to access the data. Usage examples of NVECTOR\_PARHYP are provided in the cvAdvDiff\_non\_ph.c example program for CVODE [31] and the ark\_diurnal\_kry\_ph.c example program for ARKODE [39].

The names of parhyp methods are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix Parhyp (e.g. N\_VDestroy\_Parhyp). The module NVECTOR\_PARHYP provides the following additional user-callable routines:

# N\_VNewEmpty\_ParHyp

Prototype N\_Vector N\_VNewEmpty\_ParHyp(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length)

Description This function creates a new parhyp  $N\_Vector$  with the pointer to the hypre vector set to NULL.

# N\_VMake\_ParHyp

Prototype N\_Vector N\_VMake\_ParHyp(HYPRE\_ParVector x)

Description This function creates an N\_Vector wrapper around an existing hypre parallel vector. It does not allocate memory for x itself.

# N\_VGetVector\_ParHyp

 $\label{eq:prototype} Prototype \qquad \texttt{HYPRE\_ParVector} \ \ \texttt{N\_VGetVector\_ParHyp(N\_Vector} \ \ \texttt{v)}$ 

Description This function returns the underlying hypre vector.

# N\_VCloneVectorArray\_ParHyp

 $\label{eq:prototype} Prototype \quad \mbox{N\_Vector *N\_VCloneVectorArray\_ParHyp(int count, N\_Vector w)}$ 

Description This function creates (by cloning) an array of count parallel vectors.

# N\_VCloneVectorArrayEmpty\_ParHyp

Prototype N\_Vector \*N\_VCloneVectorArrayEmpty\_ParHyp(int count, N\_Vector w)

Description This function creates (by cloning) an array of count parallel vectors, each with an empty

(NULL) data array.

#### N\_VDestroyVectorArray\_ParHyp

Prototype void N\_VDestroyVectorArray\_ParHyp(N\_Vector \*vs, int count)

Description This function frees memory allocated for the array of count variables of type N\_Vector created with N\_VCloneVectorArray\_ParHyp or with N\_VCloneVectorArrayEmpty\_ParHyp.

# $|N_VPrint_ParHyp|$

Prototype void N\_VPrint\_ParHyp(N\_Vector v)

Description This function prints the local content of a parhyp vector to stdout.

### $N_VPrintFile_ParHyp$

Prototype void N\_VPrintFile\_ParHyp(N\_Vector v, FILE \*outfile)

Description This function prints the local content of a parhyp vector to outfile.

By default all fused and vector array operations are disabled in the NVECTOR\_PARHYP module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VMake\_ParHyp, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VMake\_ParHyp will have the default settings for the NVECTOR\_PARHYP module.

# N\_VEnableFusedOps\_ParHyp

Prototype int N\_VEnableFusedOps\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array oper-

ations in the parhyp vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableLinearCombination\_ParHyp

Prototype int N\_VEnableLinearCombination\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused

operation in the parhyp vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableScaleAddMulti\_ParHyp

Prototype int N\_VEnableScaleAddMulti\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to

multiple vectors fused operation in the parhyp vector. The return value is 0 for success

and -1 if the input vector or its ops structure are NULL.

# N\_VEnableDotProdMulti\_ParHyp

Prototype int N\_VEnableDotProdMulti\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused

operation in the parhyp vector. The return value is  ${\tt 0}$  for success and  ${\tt -1}$  if the input

vector or its ops structure are NULL.

### N\_VEnableLinearSumVectorArray\_ParHyp

Prototype int N\_VEnableLinearSumVectorArray\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for

vector arrays in the parhyp vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

### N\_VEnableScaleVectorArray\_ParHyp

Prototype int N\_VEnableScaleVectorArray\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector

arrays in the parhyp vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableConstVectorArray\_ParHyp

Prototype int N\_VEnableConstVectorArray\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector

arrays in the parhyp vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableWrmsNormVectorArray\_ParHyp

Prototype int N\_VEnableWrmsNormVectorArray\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for vector arrays in the parhyp vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

N\_VEnableWrmsNormMaskVectorArray\_ParHyp

Prototype int N\_VEnableWrmsNormMaskVectorArray\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm operation for vector arrays in the parhyp vector. The return value is 0 for success and -1

if the input vector or its ops structure are NULL.

# ${\tt N\_VEnableScaleAddMultiVectorArray\_ParHyp}$

Prototype int N\_VEnableScaleAddMultiVectorArray\_ParHyp(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array

to multiple vector arrays operation in the parhyp vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

N\_VEnableLinearCombinationVectorArray\_ParHyp

Prototype int N\_VEnableLinearCombinationVectorArray\_ParHyp(N\_Vector v,

booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation for vector arrays in the parhyp vector. The return value is 0 for success and -1 if the

input vector or its ops structure are NULL.

#### Notes

- When there is a need to access components of an N\_Vector\_ParHyp, v, it is recommended to extract the hypre vector via x\_vec = N\_VGetVector\_ParHyp(v) and then access components using appropriate hypre functions.
- N\_VNewEmpty\_ParHyp, N\_VMake\_ParHyp, and N\_VCloneVectorArrayEmpty\_ParHyp set the field own\_parvector to SUNFALSE. N\_VDestroy\_ParHyp and N\_VDestroyVectorArray\_ParHyp will not attempt to delete an underlying hypre vector for any N\_Vector with own\_parvector set to SUNFALSE. In such a case, it is the user's responsibility to delete the underlying vector.
- To maximize efficiency, vector operations in the NVECTOR\_PARHYP implementation that have more than one N\_Vector argument do not check for consistent internal representations of these vectors. It is the user's responsibility to ensure that such routines are called with N\_Vector arguments that were all created with the same internal representations.

# 7.7 The NVECTOR\_PETSC implementation

The NVECTOR\_PETSC module is an NVECTOR wrapper around the PETSc vector. It defines the *content* field of a N\_Vector to be a structure containing the global and local lengths of the vector, a pointer to the PETSc vector, an MPI communicator, and a boolean flag own\_data indicating ownership of the wrapped PETSc vector.





```
struct _N_VectorContent_Petsc {
   sunindextype local_length;
   sunindextype global_length;
   booleantype own_data;
   Vec *pvec;
   MPI_Comm comm;
};
```

The header file to include when using this module is nvector\_petsc.h. The installed module library to link to is libsundials\_nvecpetsc.lib where .lib is typically .so for shared libraries and .a for static libraries.

Unlike native SUNDIALS vector types, NVECTOR\_PETSC does not provide macros to access its member variables. Note that NVECTOR\_PETSC requires SUNDIALS to be built with MPI support.

# 7.7.1 NVECTOR\_PETSC functions

The NVECTOR\_PETSC module defines implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4, except for N\_VGetArrayPointer and N\_VSetArrayPointer. As such, this vector cannot be used with SUNDIALS Fortran interfaces. When access to raw vector data is needed, it is recommended to extract the PETSc vector first, and then use PETSc methods to access the data. Usage examples of NVECTOR\_PETSC are provided in example programs for IDA [29].

The names of vector operations are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix \_Petsc (e.g. N\_VDestroy\_Petsc). The module NVECTOR\_PETSC provides the following additional user-callable routines:

### N\_VNewEmpty\_Petsc

Prototype N\_Vector N\_VNewEmpty\_Petsc(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length)

Description This function creates a new NVECTOR wrapper with the pointer to the wrapped PETSc vector set to (NULL). It is used by the N\_VMake\_Petsc and N\_VClone\_Petsc implementations.

### N\_VMake\_Petsc

Prototype N\_Vector N\_VMake\_Petsc(Vec \*pvec)

Description This function creates and allocates memory for an NVECTOR\_PETSC wrapper around a user-provided PETSc vector. It does *not* allocate memory for the vector **pvec** itself.

# N\_VGetVector\_Petsc

Prototype Vec \*N\_VGetVector\_Petsc(N\_Vector v)

Description This function returns a pointer to the underlying PETSc vector.

# N\_VCloneVectorArray\_Petsc

Prototype N\_Vector \*N\_VCloneVectorArray\_Petsc(int count, N\_Vector w)

Description This function creates (by cloning) an array of count NVECTOR\_PETSC vectors.

# N\_VCloneVectorArrayEmpty\_Petsc

Prototype N\_Vector \*N\_VCloneVectorArrayEmpty\_Petsc(int count, N\_Vector w)

Description This function creates (by cloning) an array of count NVECTOR\_PETSC vectors, each with pointers to PETSc vectors set to (NULL).

# N\_VDestroyVectorArray\_Petsc

Prototype void N\_VDestroyVectorArray\_Petsc(N\_Vector \*vs, int count)

Description This function frees memory allocated for the array of count variables of type N\_Vector created with N\_VCloneVectorArray\_Petsc or with N\_VCloneVectorArrayEmpty\_Petsc.

#### N\_VPrint\_Petsc

Prototype void N\_VPrint\_Petsc(N\_Vector v)

Description This function prints the global content of a wrapped PETSc vector to stdout.

# N\_VPrintFile\_Petsc

Prototype void N\_VPrintFile\_Petsc(N\_Vector v, const char fname[])

Description This function prints the global content of a wrapped PETSc vector to fname.

By default all fused and vector array operations are disabled in the NVECTOR\_PETSC module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VMake\_Petsc, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VMake\_Petsc will have the default settings for the NVECTOR\_PETSC module.

#### N\_VEnableFusedOps\_Petsc

Prototype int N\_VEnableFusedOps\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array operations in the PETSc vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableLinearCombination\_Petsc

Prototype int N\_VEnableLinearCombination\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused operation in the PETSc vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableScaleAddMulti\_Petsc

Prototype int N\_VEnableScaleAddMulti\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to multiple vectors fused operation in the PETSc vector. The return value is 0 for success

and -1 if the input vector or its ops structure are NULL.

#### N\_VEnableDotProdMulti\_Petsc

Prototype int N\_VEnableDotProdMulti\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused operation in the PETSc vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableLinearSumVectorArray\_Petsc

Prototype int N\_VEnableLinearSumVectorArray\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for vector arrays in the PETSC vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableScaleVectorArray\_Petsc

Prototype int N\_VEnableScaleVectorArray\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector

arrays in the PETSc vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

### N\_VEnableConstVectorArray\_Petsc

Prototype int N\_VEnableConstVectorArray\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector

arrays in the PETSc vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableWrmsNormVectorArray\_Petsc

Prototype int N\_VEnableWrmsNormVectorArray\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for

vector arrays in the PETSc vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

### N\_VEnableWrmsNormMaskVectorArray\_Petsc

Prototype int N\_VEnableWrmsNormMaskVectorArray\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm op-

eration for vector arrays in the PETSc vector. The return value is 0 for success and -1

if the input vector or its ops structure are NULL.

#### N\_VEnableScaleAddMultiVectorArray\_Petsc

Prototype int N\_VEnableScaleAddMultiVectorArray\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array

to multiple vector arrays operation in the PETSc vector. The return value is  $\boldsymbol{0}$  for success

and  ${\tt -1}$  if the input vector or its  ${\tt ops}$  structure are NULL.

# ${\tt N\_VEnableLinearCombinationVectorArray\_Petsc}$

Prototype int N\_VEnableLinearCombinationVectorArray\_Petsc(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation

for vector arrays in the PETSc vector. The return value is  ${\tt 0}$  for success and  ${\tt -1}$  if the

input vector or its ops structure are NULL.

#### Notes

- When there is a need to access components of an N\_Vector\_Petsc, v, it is recommeded to extract the PETSc vector via x\_vec = N\_VGetVector\_Petsc(v) and then access components using appropriate PETSc functions.
- The functions N\_VNewEmpty\_Petsc, N\_VMake\_Petsc, and N\_VCloneVectorArrayEmpty\_Petsc set the field own\_data to SUNFALSE. N\_VDestroy\_Petsc and N\_VDestroyVectorArray\_Petsc will not attempt to free the pointer pvec for any N\_Vector with own\_data set to SUNFALSE. In such a case, it is the user's responsibility to deallocate the pvec pointer.
- To maximize efficiency, vector operations in the NVECTOR\_PETSC implementation that have more than one N\_Vector argument do not check for consistent internal representations of these vectors. It is the user's responsibility to ensure that such routines are called with N\_Vector arguments that were all created with the same internal representations.





# 7.8 The NVECTOR\_CUDA implementation

The NVECTOR\_CUDA module is an experimental NVECTOR implementation in the CUDA language. The module allows for SUNDIALS vector kernels to run on GPU devices. It is intended for users who are already familiar with CUDA and GPU programming. Building this vector module requires a CUDA compiler and, by extension, a C++ compiler. The class **Vector** in the namespace **suncudavec** manages the vector data layout:

```
template <class T, class I>
class Vector {
    I size_;
    I mem_size_;
    I global_size_;
    T* h_vec_;
    T* d_vec_;
    ThreadPartitioning<T, I>* partStream_;
    ThreadPartitioning<T, I>* partReduce_;
    bool ownPartitioning_;
    bool ownData_;
    bool managed_mem_;
    SUNMPI_Comm comm_;
    ...
};
```

The class members are vector size (length), size of the vector data memory block, pointers to vector data on the host and the device, pointers to ThreadPartitioning implementations that handle thread partitioning for streaming and reduction vector kernels, a boolean flag that signals if the vector owns the thread partitioning, a boolean flag that signals if the vector owns the data, a boolean flag that signals if managed memory is used for the data arrays, and the MPI communicator. The class Vector inherits from the empty structure

```
struct _N_VectorContent_Cuda {};
```

to interface the C++ class with the NVECTOR C code. Due to the rapid progress of CUDA development, we expect that the suncudavec::Vector class will change frequently in future SUNDIALS releases. The code is structured so that it can tolerate significant changes in the suncudavec::Vector class without requiring changes to the user API.

When instantiated with N\_VNew\_Cuda, the class Vector will allocate memory on both the host and the device. Alternatively, a user can provide host and device data arrays by using the N\_VMake\_Cuda

constructor. To use CUDA managed memory, the constructors N\_VNewManaged\_Cuda and N\_VMakeManaged\_Cuda are provided. Details on each of these constructors are provided below.

The NVECTOR\_CUDA module can be utilized for single-node parallelism or in a distributed context with MPI. In the single-node case the header file to include nvector\_cuda.h and the library to link to is libsundials\_nveccuda.lib. In the a distributed setting the header file to include is nvector\_mpicuda.h and the library to link to is libsundials\_nvecmpicuda.lib. The extension, .lib, is typically .so for shared libraries and .a for static libraries. Only one of these libraries may be linked to when creating an executable or library. SUNDIALS must be built with MPI support if the distributed library is desired.

### 7.8.1 NVECTOR\_CUDA functions

Unlike other native SUNDIALS vector types, NVECTOR\_CUDA does not provide macros to access its member variables. Instead, user should use the accessor functions:

### N\_VGetLength\_Cuda

Prototype sunindextype N\_VGetLength\_Cuda(N\_Vector v)

Description This function returns the global length of the vector.

# N\_VGetLocalLength\_Cuda

Prototype sunindextype N\_VGetLocalLength\_Cuda(N\_Vector v)

Description This function returns the local length of the vector.

Note: This function is for use in a *distributed* context and is defined in the header nvector\_mpicuda.h and the library to link to is libsundials\_nvecmpicuda.lib.

### N\_VGetHostArrayPointer\_Cuda

Prototype realtype \*N\_VGetHostArrayPointer\_Cuda(N\_Vector v)

Description This function returns a pointer to the vector data on the host.

# ${\tt N\_VGetDeviceArrayPointer\_Cuda}$

Prototype realtype \*N\_VGetDeviceArrayPointer\_Cuda(N\_Vector v)

Description This function returns a pointer to the vector data on the device.

# N\_VGetMPIComm\_Cuda

Prototype MPI\_Comm N\_VGetMPIComm\_Cuda(N\_Vector v)

Description This function returns the MPI communicator for the vector.

Note: This function is for use in a *distributed* context and is defined in the header nvector\_mpicuda.h and the library to link to is libsundials\_nvecmpicuda.lib.

# N\_VIsManagedMemory\_Cuda

Prototype booleantype \*N\_VIsManagedMemory\_Cuda(N\_Vector v)

Description This function returns a boolean flag indicating if the vector data is allocated in managed

memory or not.

The NVECTOR\_CUDA module defines implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4, except for N\_VGetArrayPointer and N\_VSetArrayPointer. As such, this vector cannot be used with the SUNDIALS Fortran interfaces, nor with the SUNDIALS direct solvers and preconditioners. Instead, the NVECTOR\_CUDA module provides separate functions to access data on the host and on the device. It also provides methods for copying from the host to the device and vice versa. Usage examples of NVECTOR\_CUDA are provided in some example programs for CVODE [31].

The names of vector operations are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix \_Cuda (e.g. N\_VDestroy\_Cuda). The module NVECTOR\_CUDA provides the following functions:

### N\_VNew\_Cuda

Single-node usage

Prototype N\_Vector N\_VNew\_Cuda(sunindextype length)

Description This function creates and allocates memory for a CUDA N\_Vector. The vector data array is allocated on both the host and device. In the *single-node* setting, the only input is

the vector length. This constructor is defined in the header nvector\_cuda.h and the

library to link to is libsundials\_nveccuda.lib.

Distributed-memory parallel usage

Prototype N\_Vector N\_VNew\_Cuda(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length)

Description This function creates and allocates memory for a CUDA N\_Vector. The vector data array is allocated on both the host and device. When used in a distributed context with MPI, the arguments are the MPI communicator, the local vector length, and the global vector length. This constructor is defined in the header nvector\_mpicuda.h and the library to link to is libsundials\_nvecmpicuda.lib.

# N\_VNewManaged\_Cuda

Single-node usage

Prototype N\_Vector N\_VNewManaged\_Cuda(sunindextype length)

Description This function creates and allocates memory for a CUDA N\_Vector on a single node. The vector data array is allocated in managed memory. In the *single-node* setting, the only input is the vector length. This constructor is defined in the header nvector\_cuda.h

and the library to link to is libsundials\_nveccuda.lib.

Distributed-memory parallel usage

Prototype N\_Vector N\_VNewManaged\_Cuda(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length)

sunlindextype global\_length

Description This function creates and allocates memory for a CUDA N\_Vector on a single node. The vector data array is allocated in managed memory. When used in a distributed context with MPI, the arguments are the MPI communicator, the local vector length, and the global vector length. This constructor is defined in the header nvector\_mpicuda.h and

the library to link to is libsundials\_nvecmpicuda.lib.

### N\_VNewEmpty\_Cuda

Prototype N\_Vector N\_VNewEmpty\_Cuda()

Description This function creates a new NVECTOR wrapper with the pointer to the wrapped CUDA vector set to NULL. It is used by the N VNew Cuda, N VMake Cuda, and N VClone Cuda

vector set to NULL. It is used by the N\_VNew\_Cuda, N\_VMake\_Cuda, and N\_VClone\_Cuda

implementations.

# N\_VMake\_Cuda

 $Single-node\ usage$ 

Prototype N\_Vector N\_VMake\_Cuda(sunindextype length, realtype \*h\_vdata, realtype \*d\_vdata)

Description

This function creates an NVECTOR\_CUDA with user-supplied vector data arrays h\_vdata and d\_vdata. This function does not allocate memory for data itself. In the *single-node* setting, the inputs are the vector length, the host data array, and the device data. This constructor is defined in the header nvector\_cuda.h and the library to link to is libsundials\_nveccuda.lib.

Distributed-memory parallel usage

Prototype N\_Vector N\_VMake\_Cuda(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length, realtype \*h\_vdata, realtype \*d\_vdata)

Description

This function creates an NVECTOR\_CUDA with user-supplied vector data arrays h\_vdata and d\_vdata. This function does not allocate memory for data itself. When used in a distributed context with MPI, the arguments are the MPI communicator, the local vector length, the global vector length, the host data array, and the device data array. This constructor is defined in the header nvector\_mpicuda.h and the library to link to is libsundials\_nvecmpicuda.lib.

### N\_VMakeManaged\_Cuda

Single-node usage

Prototype N\_Vector N\_VMakeManaged\_Cuda(sunindextype length, realtype \*vdata)

Description

This function creates an NVECTOR\_CUDA with a user-supplied managed memory data array. This function does not allocate memory for data itself. In the *single-node* setting, the inputs are the vector length and the managed data array. This constructor is defined in the header nvector\_cuda.h and the library to link to is libsundials\_nveccuda.lib.

Distributed-memory parallel usage

Prototype N\_Vector N\_VMakeManaged\_Cuda(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length, realtype \*vdata)

Description This function creates an NVECTOR\_CUDA with a user-supplied managed memory data array. This function does not allocate memory for data itself. When used in a *distributed* context with MPI, the arguments are the MPI communicator, the local vector length, the global vector length, the managed data array. This constructor is defined in the header nvector\_mpicuda.h and the library to link to is libsundials\_nvecmpicuda.lib.

The module NVECTOR\_CUDA also provides the following user-callable routines:

### N\_VSetCudaStream\_Cuda

Prototype void N\_VSetCudaStream\_Cuda(N\_Vector v, cudaStream\_t \*stream)

Description This function sets the CUDA stream that all vector kernels will be launched on. By default an NVECTOR\_CUDA uses the default CUDA stream.

Note: All vectors used in a single instance of a SUNDIALS solver must use the same CUDA stream, and the CUDA stream must be set prior to solver initialization. Additionally, if manually instantiating the stream and reduce ThreadPartitioning of a suncudavec::Vector, ensure that they use the same CUDA stream.

# N\_VCopyToDevice\_Cuda

Prototype realtype \*N\_VCopyToDevice\_Cuda(N\_Vector v)

Description This function copies host vector data to the device.

# N\_VCopyFromDevice\_Cuda

Prototype realtype \*N\_VCopyFromDevice\_Cuda(N\_Vector v)

Description This function copies vector data from the device to the host.

### N\_VPrint\_Cuda

Prototype void N\_VPrint\_Cuda(N\_Vector v)

Description This function prints the content of a CUDA vector to stdout.

### N\_VPrintFile\_Cuda

Prototype void N\_VPrintFile\_Cuda(N\_Vector v, FILE \*outfile)

Description This function prints the content of a CUDA vector to outfile.

By default all fused and vector array operations are disabled in the NVECTOR\_CUDA module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VNew\_Cuda, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VNew\_Cuda will have the default settings for the NVECTOR\_CUDA module.

#### N\_VEnableFusedOps\_Cuda

Prototype int N\_VEnableFusedOps\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array operations in the CUDA vector. The return value is 0 for success and -1 if the input vector

or its  $\ensuremath{\mathsf{ops}}$  structure are NULL.

#### N\_VEnableLinearCombination\_Cuda

Prototype int N\_VEnableLinearCombination\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused operation in the CUDA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableScaleAddMulti\_Cuda

Prototype int N\_VEnableScaleAddMulti\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to multiple vectors fused operation in the CUDA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableDotProdMulti\_Cuda

Prototype int N\_VEnableDotProdMulti\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused operation in the CUDA vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableLinearSumVectorArray\_Cuda

Prototype int N\_VEnableLinearSumVectorArray\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for vector arrays in the CUDA vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableScaleVectorArray\_Cuda

Prototype int N\_VEnableScaleVectorArray\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector

arrays in the CUDA vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

#### N\_VEnableConstVectorArray\_Cuda

Prototype int N\_VEnableConstVectorArray\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector

arrays in the CUDA vector. The return value is 0 for success and -1 if the input vector

or its ops structure are NULL.

# N\_VEnableWrmsNormVectorArray\_Cuda

Prototype int N\_VEnableWrmsNormVectorArray\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for

vector arrays in the CUDA vector. The return value is 0 for success and -1 if the input

vector or its ops structure are NULL.

# N\_VEnableWrmsNormMaskVectorArray\_Cuda

Prototype int N\_VEnableWrmsNormMaskVectorArray\_Cuda(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm op-

eration for vector arrays in the CUDA vector. The return value is 0 for success and -1 if

the input vector or its ops structure are NULL.

# ${\tt N\_VEnableScaleAddMultiVectorArray\_Cuda}$

 $Prototype \quad \text{ int $N_V$EnableScaleAddMultiVectorArray\_Cuda($N_V$ector $v$, booleantype tf)} \\$ 

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array to multiple vector arrays operation in the CUDA vector. The return value is 0 for success

and -1 if the input vector or its ops structure are NULL.

# N\_VEnableLinearCombinationVectorArray\_Cuda

```
Prototype int N_VEnableLinearCombinationVectorArray_Cuda(N_Vector v, booleantype tf)
```

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation for vector arrays in the CUDA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

#### Notes

- When there is a need to access components of an N\_Vector\_Cuda, v, it is recommeded to use functions N\_VGetDeviceArrayPointer\_Cuda or N\_VGetHostArrayPointer\_Cuda.
- To maximize efficiency, vector operations in the NVECTOR\_CUDA implementation that have more than one N\_Vector argument do not check for consistent internal representations of these vectors. It is the user's responsibility to ensure that such routines are called with N\_Vector arguments that were all created with the same internal representations.



# 7.9 The NVECTOR\_RAJA implementation

The NVECTOR\_RAJA module is an experimental NVECTOR implementation using the RAJA hardware abstraction layer. In this implementation, RAJA allows for SUNDIALS vector kernels to run on GPU devices. The module is intended for users who are already familiar with RAJA and GPU programming. Building this vector module requires a C++11 compliant compiler and a CUDA software development toolkit. Besides the CUDA backend, RAJA has other backends such as serial, OpenMP, and OpenACC. These backends are not used in this SUNDIALS release. Class Vector in namespace sunrajavec manages the vector data layout:

```
template <class T, class I>
class Vector {
    I size_;
    I mem_size_;
    I global_size_;
    T* h_vec_;
    T* d_vec_;
    SUNMPI_Comm comm_;
    ...
};
```

The class members are: vector size (length), size of the vector data memory block, the global vector size (length), pointers to vector data on the host and on the device, and the MPI communicator. The class **Vector** inherits from an empty structure

```
struct _N_VectorContent_Raja {
};
```

to interface the C++ class with the NVECTOR C code. When instantiated, the class Vector will allocate memory on both the host and the device. Due to the rapid progress of RAJA development, we expect that the sunrajavec::Vector class will change frequently in future SUNDIALS releases. The code is structured so that it can tolerate significant changes in the sunrajavec::Vector class without requiring changes to the user API.

The NVECTOR\_RAJA module can be utilized for single-node parallelism or in a distributed context with MPI. The header file to include when using this module for single-node parallelism is nvector\_raja.h. The header file to include when using this module in the distributed case is nvector\_mpiraja.h. The installed module libraries to link to are libsundials\_nvecraja.lib in the single-node case, or libsundials\_nvecmpicudaraja.lib in the distributed case. Only one one

of these libraries may be linked to when creating an executable or library. SUNDIALS must be built with MPI support if the distributed library is desired.

### 7.9.1 NVECTOR\_RAJA functions

Unlike other native SUNDIALS vector types, NVECTOR\_RAJA does not provide macros to access its member variables. Instead, user should use the accessor functions:

# N\_VGetLength\_Raja

Prototype sunindextype N\_VGetLength\_Raja(N\_Vector v)

Description This function returns the global length of the vector.

# N\_VGetLocalLength\_Raja

Prototype sunindextype N\_VGetLocalLength\_Raja(N\_Vector v)

Description This function returns the local length of the vector.

Note: This function is for use in a *distributed context* and is defined in the header nvector\_mpiraja.h and the library to link to is libsundials\_nvecmpicudaraja.lib.

#### N\_VGetHostArrayPointer\_Raja

Prototype realtype \*N\_VGetHostArrayPointer\_Raja(N\_Vector v)

Description This function returns a pointer to the vector data on the host.

# N\_VGetDeviceArrayPointer\_Raja

Prototype realtype \*N\_VGetDeviceArrayPointer\_Raja(N\_Vector v)

Description This function returns a pointer to the vector data on the device.

### N\_VGetMPIComm\_Raja

Prototype MPI\_Comm N\_VGetMPIComm\_Raja(N\_Vector v)

Description This function returns the MPI communicator for the vector.

Note: This function is for use in a *distributed context* and is defined in the header nvector\_mpiraja.h and the library to link to is libsundials\_nvecmpicudaraja.lib.

The NVECTOR\_RAJA module defines the implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4, except for N\_VDotProdMulti, N\_VWrmsNormVectorArray, and

N\_VWrmsNormMaskVectorArray as support for arrays of reduction vectors is not yet supported in RAJA. These function will be added to the NVECTOR\_RAJA implementation in the future. Additionally the vector operations N\_VGetArrayPointer and N\_VSetArrayPointer are not implemented by the RAJA vector. As such, this vector cannot be used with the SUNDIALS Fortran interfaces, nor with the SUNDIALS direct solvers and preconditioners. The NVECTOR\_RAJA module provides separate functions to access data on the host and on the device. It also provides methods for copying data from the host to the device and vice versa. Usage examples of NVECTOR\_RAJA are provided in some example programs for CVODE [31].

The names of vector operations are obtained from those in Tables 7.2, 7.3, and 7.4, by appending the suffix \_Raja (e.g. N\_VDestroy\_Raja). The module NVECTOR\_RAJA provides the following additional user-callable routines:

# N\_VNew\_Raja

 $Single-node\ usage$ 

Prototype N\_Vector N\_VNew\_Raja(sunindextype length)

Description This function creates and allocates memory for a CUDA N\_Vector. The vector data array is allocated on both the host and device. In the *single-node* setting, the only input is the vector length. This constructor is defined in the header nvector\_raja.h and the

library to link to is libsundials\_nveccudaraja.lib.

Distributed-memory parallel usage

Prototype N\_Vector N\_VNew\_Raja(MPI\_Comm comm, sunindextype local\_length, sunindextype global\_length)

Description This function creates and allocates memory for a CUDA N\_Vector. The vector data array is allocated on both the host and device. When used in a *distributed* context with MPI, the arguments are the MPI communicator, the local vector length, and the

 ${\it global\ vector\ length.\ This\ constructor\ is\ defined\ in\ the\ header\ {\it nvector\_mpiraja.h}\ and}$ 

the library to link to is libsundials\_nvecmpicudaraja.lib.

# $N_VNewEmpty_Raja$

Prototype N\_Vector N\_VNewEmpty\_Raja()

Description This function creates a new NVECTOR wrapper with the pointer to the wrapped RAJA

 $vector\ set\ to\ {\tt NULL}.\ It\ is\ used\ by\ the\ {\tt N\_VNew\_Raja},\ {\tt N\_VMake\_Raja},\ {\tt and}\ {\tt N\_VClone\_Raja}$ 

implementations.

#### N\_VMake\_Raja

Prototype N\_Vector N\_VMake\_Raja(N\_VectorContent\_Raja c)

Description This function creates and allocates memory for an NVECTOR\_RAJA wrapper around a

user-provided sunrajavec::Vector class. Its only argument is of type

N\_VectorContent\_Raja, which is the pointer to the class.

#### N\_VCopyToDevice\_Raja

Prototype realtype \*N\_VCopyToDevice\_Raja(N\_Vector v)

Description This function copies host vector data to the device.

# $N_VCopyFromDevice_Raja$

Prototype realtype \*N\_VCopyFromDevice\_Raja(N\_Vector v)

Description This function copies vector data from the device to the host.

# $N_VPrint_Raja$

Prototype void N\_VPrint\_Raja(N\_Vector v)

Description This function prints the content of a RAJA vector to stdout.

# N\_VPrintFile\_Raja

Prototype void N\_VPrintFile\_Raja(N\_Vector v, FILE \*outfile)

Description This function prints the content of a RAJA vector to outfile.

By default all fused and vector array operations are disabled in the NVECTOR\_RAJA module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VNew\_Raja, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VNew\_Raja will have the default settings for the NVECTOR\_RAJA module.

# N\_VEnableFusedOps\_Raja

Prototype int N\_VEnableFusedOps\_Raja(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array operations in the RAJA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableLinearCombination\_Raja

Prototype int N\_VEnableLinearCombination\_Raja(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused operation in the RAJA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableScaleAddMulti\_Raja

Prototype int N\_VEnableScaleAddMulti\_Raja(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to multiple vectors fused operation in the RAJA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

#### N\_VEnableLinearSumVectorArray\_Raja

Prototype int N\_VEnableLinearSumVectorArray\_Raja(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for vector arrays in the RAJA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableScaleVectorArray\_Raja

Prototype int N\_VEnableScaleVectorArray\_Raja(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale operation for vector arrays in the RAJA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

#### N\_VEnableConstVectorArray\_Raja

Prototype int N\_VEnableConstVectorArray\_Raja(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector arrays in the RAJA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

#### N\_VEnableScaleAddMultiVectorArray\_Raja

Prototype int N\_VEnableScaleAddMultiVectorArray\_Raja(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array to multiple vector arrays operation in the RAJA vector. The return value is 0 for success

and -1 if the input vector or its ops structure are NULL.

# ${\tt N\_VEnableLinearCombinationVectorArray\_Raja}$

```
Prototype int N_VEnableLinearCombinationVectorArray_Raja(N_Vector v, booleantype tf)
```

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation

for vector arrays in the RAJA vector. The return value is 0 for success and -1 if the input vector or its ops structure are NULL.

#### Notes

• When there is a need to access components of an N\_Vector\_Raja, v, it is recommeded to use functions N\_VGetDeviceArrayPointer\_Raja or N\_VGetHostArrayPointer\_Raja.



• To maximize efficiency, vector operations in the NVECTOR\_RAJA implementation that have more than one N\_Vector argument do not check for consistent internal representations of these vectors. It is the user's responsibility to ensure that such routines are called with N\_Vector arguments that were all created with the same internal representations.

# 7.10 The NVECTOR\_OPENMPDEV implementation

In situations where a user has access to a device such as a GPU for offloading computation, SUNDIALS provides an NVECTOR implementation using OpenMP device offloading, called NVECTOR\_OPENMPDEV.

The NVECTOR\_OPENMPDEV implementation defines the *content* field of the N\_Vector to be a structure containing the length of the vector, a pointer to the beginning of a contiguous data array on the host, a pointer to the beginning of a contiguous data array on the device, and a boolean flag own\_data which specifies the ownership of host and device data arrays.

```
struct _N_VectorContent_OpenMPDEV {
   sunindextype length;
   booleantype own_data;
   realtype *host_data;
   realtype *dev_data;
};
```

The header file to include when using this module is nvector\_openmpdev.h. The installed module library to link to is libsundials\_nvecopenmpdev.lib where .lib is typically .so for shared libraries and .a for static libraries.

#### 7.10.1 NVECTOR\_OPENMPDEV accessor macros

The following macros are provided to access the content of an NVECTOR\_OPENMPDEV vector.

#### NV\_CONTENT\_OMPDEV

This routine gives access to the contents of the NVECTOR\_OPENMPDEV vector N\_Vector.

The assignment  $v\_cont = NV\_CONTENT\_OMPDEV(v)$  sets  $v\_cont$  to be a pointer to the NVECTOR\_OPENMPDEV  $N\_Vector$  content structure.

Implementation:

```
#define NV_CONTENT_OMPDEV(v) ( (N_VectorContent_OpenMPDEV)(v->content) )
```

NV\_OWN\_DATA\_OMPDEV, NV\_DATA\_HOST\_OMPDEV, NV\_DATA\_DEV\_OMPDEV, NV\_LENGTH\_OMPDEV

These macros give individual access to the parts of the content of an NVECTOR\_OPENMPDEV N\_Vector.

The assignment v\_data = NV\_DATA\_HOST\_OMPDEV(v) sets v\_data to be a pointer to the first component of the data on the host for the N\_Vector v. The assignment NV\_DATA\_HOST\_OMPDEV(v) = v\_data sets the host component array of v to be v\_data by storing the pointer v\_data.

The assignment v\_dev\_data = NV\_DATA\_DEV\_OMPDEV(v) sets v\_dev\_data to be a pointer to the first component of the data on the device for the N\_Vector v. The assignment NV\_DATA\_DEV\_OMPDEV(v) = v\_dev\_data sets the device component array of v to be v\_dev\_data by storing the pointer v\_dev\_data.

The assignment  $v_len = NV_length_ompdev(v)$  sets  $v_len$  to be the length of v. On the other hand, the call  $NV_length_ompdev(v) = len_v$  sets the length of v to be  $len_v$ .

Implementation:

```
#define NV_OWN_DATA_OMPDEV(v) ( NV_CONTENT_OMPDEV(v)->own_data )
#define NV_DATA_HOST_OMPDEV(v) ( NV_CONTENT_OMPDEV(v)->host_data )
#define NV_DATA_DEV_OMPDEV(v) ( NV_CONTENT_OMPDEV(v)->dev_data )
#define NV_LENGTH_OMPDEV(v) ( NV_CONTENT_OMPDEV(v)->length )
```

# 7.10.2 NVECTOR\_OPENMPDEV functions

The NVECTOR\_OPENMPDEV module defines OpenMP device offloading implementations of all vector operations listed in Tables 7.2, 7.3, and 7.4, except for N\_VGetArrayPointer and N\_VSetArrayPointer. As such, this vector cannot be used with the SUNDIALS Fortran interfaces, nor with the SUNDIALS direct solvers and preconditioners. It also provides methods for copying from the host to the device and vice versa.

The names of vector operations are obtained from those in Tables 7.2, 7.3, and 7.4 by appending the suffix \_OpenMPDEV (e.g. N\_VDestroy\_OpenMPDEV). The module NVECTOR\_OPENMPDEV provides the following additional user-callable routines:

#### N\_VNew\_OpenMPDEV

Prototype N\_Vector N\_VNew\_OpenMPDEV(sunindextype vec\_length)

Description This function creates and allocates memory for an NVECTOR\_OPENMPDEV N\_Vector.

# N\_VNewEmpty\_OpenMPDEV

Prototype N\_Vector N\_VNewEmpty\_OpenMPDEV(sunindextype vec\_length)

Description This function creates a new NVECTOR\_OPENMPDEV N\_Vector with an empty (NULL) host and device data arrays.

# N\_VMake\_OpenMPDEV

Prototype N\_Vector N\_VMake\_OpenMPDEV(sunindextype vec\_length, realtype \*h\_vdata, realtype \*d\_vdata)

Description This function creates an NVECTOR\_OPENMPDEV vector with user-supplied vector data arrays h\_vdata and d\_vdata. This function does not allocate memory for data itself.

#### N\_VCloneVectorArray\_OpenMPDEV

Prototype N\_Vector \*N\_VCloneVectorArray\_OpenMPDEV(int count, N\_Vector w)

Description This function creates (by cloning) an array of count NVECTOR\_OPENMPDEV vectors.

# N\_VCloneVectorArrayEmpty\_OpenMPDEV

Prototype N\_Vector \*N\_VCloneVectorArrayEmpty\_OpenMPDEV(int count, N\_Vector w)

Description This function creates (by cloning) an array of count NVECTOR\_OPENMPDEV vectors,

each with an empty (NULL) data array.

# N\_VDestroyVectorArray\_OpenMPDEV

Prototype void N\_VDestroyVectorArray\_OpenMPDEV(N\_Vector \*vs, int count)

Description This function frees memory allocated for the array of count variables of type N\_Vector

created with N\_VCloneVectorArray\_OpenMPDEV or with

 ${\tt N\_VCloneVectorArrayEmpty\_OpenMPDEV}.$ 

#### N\_VGetLength\_OpenMPDEV

Prototype sunindextype N\_VGetLength\_OpenMPDEV(N\_Vector v)

Description This function returns the number of vector elements.

#### $N_VGetHostArrayPointer_OpenMPDEV$

Prototype realtype \*N\_VGetHostArrayPointer\_OpenMPDEV(N\_Vector v)

Description This function returns a pointer to the host data array.

# | N\_VGetDeviceArrayPointer\_OpenMPDEV

Prototype realtype \*N\_VGetDeviceArrayPointer\_OpenMPDEV(N\_Vector v)

Description This function returns a pointer to the device data array.

#### N\_VPrint\_OpenMPDEV

Prototype void N\_VPrint\_OpenMPDEV(N\_Vector v)

Description This function prints the content of an NVECTOR\_OPENMPDEV vector to stdout.

#### $N_VPrintFile_OpenMPDEV$

Prototype void N\_VPrintFile\_OpenMPDEV(N\_Vector v, FILE \*outfile)

Description This function prints the content of an NVECTOR\_OPENMPDEV vector to outfile.

#### N\_VCopyToDevice\_OpenMPDEV

Prototype void N\_VCopyToDevice\_OpenMPDEV(N\_Vector v)

Description This function copies the content of an NVECTOR\_OPENMPDEV vector's host data array

to the device data array.

#### N\_VCopyFromDevice\_OpenMPDEV

Prototype void N\_VCopyFromDevice\_OpenMPDEV(N\_Vector v)

Description This function copies the content of an NVECTOR\_OPENMPDEV vector's device data array

to the host data array.

By default all fused and vector array operations are disabled in the NVECTOR\_OPENMPDEV module. The following additional user-callable routines are provided to enable or disable fused and vector array operations for a specific vector. To ensure consistency across vectors it is recommended to first create a vector with N\_VNew\_OpenMPDEV, enable/disable the desired operations for that vector with the functions below, and create any additional vectors from that vector using N\_VClone. This guarantees the new vectors will have the same operations enabled/disabled as cloned vectors inherit the same enable/disable options as the vector they are cloned from while vectors created with N\_VNew\_OpenMPDEV will have the default settings for the NVECTOR\_OPENMPDEV module.

#### N\_VEnableFusedOps\_OpenMPDEV

Prototype int N\_VEnableFusedOps\_OpenMPDEV(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) all fused and vector array op-

erations in the NVECTOR\_OPENMPDEV vector. The return value is 0 for success and -1

if the input vector or its ops structure are NULL.

# ${\tt N\_VEnableLinearCombination\_OpenMPDEV}$

Prototype int N\_VEnableLinearCombination\_OpenMPDEV(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination fused

operation in the NVECTOR\_OPENMPDEV vector. The return value is 0 for success and

-1 if the input vector or its ops structure are NULL.

#### N\_VEnableScaleAddMulti\_OpenMPDEV

Prototype int N\_VEnableScaleAddMulti\_OpenMPDEV(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector to

multiple vectors fused operation in the NVECTOR\_OPENMPDEV vector. The return value

is 0 for success and -1 if the input vector or its ops structure are NULL.

# N\_VEnableDotProdMulti\_OpenMPDEV

Prototype int N\_VEnableDotProdMulti\_OpenMPDEV(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the multiple dot products fused

operation in the  ${\tt NVECTOR\_OPENMPDEV}$  vector. The return value is 0 for success and

-1 if the input vector or its ops structure are NULL.

#### N\_VEnableLinearSumVectorArray\_OpenMPDEV

Prototype int N\_VEnableLinearSumVectorArray\_OpenMPDEV(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear sum operation for

vector arrays in the NVECTOR\_OPENMPDEV vector. The return value is  ${\tt 0}$  for success

and -1 if the input vector or its ops structure are NULL.

N\_VEnableScaleVectorArray\_OpenMPDEV

Prototype int N\_VEnableScaleVectorArray\_OpenMPDEV(N\_Vector v, booleantype tf)

 $\label{thm:condition} Description \quad This \ function \ enables \ ({\tt SUNTRUE}) \ or \ disables \ ({\tt SUNFALSE}) \ the \ scale \ operation \ for \ vector$ 

arrays in the NVECTOR\_OPENMPDEV vector. The return value is 0 for success and -1 if

the input vector or its ops structure are NULL.

N\_VEnableConstVectorArray\_OpenMPDEV

Prototype int N\_VEnableConstVectorArray\_OpenMPDEV(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the const operation for vector

arrays in the  ${\tt NVECTOR\_OPENMPDEV}$  vector. The return value is 0 for success and -1 if

the input vector or its ops structure are NULL.

| N\_VEnableWrmsNormVectorArray\_OpenMPDEV

Prototype int N\_VEnableWrmsNormVectorArray\_OpenMPDEV(N\_Vector v, booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the WRMS norm operation for

vector arrays in the NVECTOR\_OPENMPDEV vector. The return value is 0 for success

and -1 if the input vector or its ops structure are NULL.

N\_VEnableWrmsNormMaskVectorArray\_OpenMPDEV

Prototype int N\_VEnableWrmsNormMaskVectorArray\_OpenMPDEV(N\_Vector v,

booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the masked WRMS norm op-

eration for vector arrays in the  $NVECTOR\_OPENMPDEV$  vector. The return value is 0 for

success and -1 if the input vector or its ops structure are NULL.

N\_VEnableScaleAddMultiVectorArray\_OpenMPDEV

 $\label{lem:lemma$ 

booleantype tf)

Description This function enables (SUNTRUE) or disables (SUNFALSE) the scale and add a vector array

to multiple vector arrays operation in the  ${\tt NVECTOR\_OPENMPDEV}$  vector. The return

value is 0 for success and -1 if the input vector or its ops structure are NULL.

 ${\tt N\_VEnableLinearCombinationVectorArray\_OpenMPDEV}$ 

 $\label{thm:continuous} Prototype \quad \text{ int $\tt N\_VEnableLinearCombinationVectorArray\_OpenMPDEV($\tt N\_Vector \ v$, and the prototype of the prot$ 

booleantype tf)

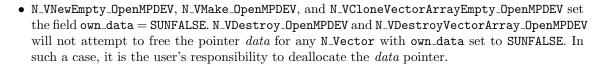
Description This function enables (SUNTRUE) or disables (SUNFALSE) the linear combination operation

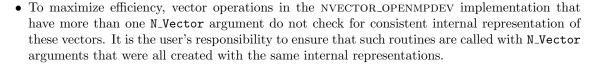
for vector arrays in the NVECTOR\_OPENMPDEV vector. The return value is 0 for success

and -1 if the input vector or its ops structure are NULL.

#### Notes

- When looping over the components of an N\_Vector v, it is most efficient to first obtain the component array via h\_data = NV\_DATA\_HOST\_OMPDEV(v) for the host array or d\_data = NV\_DATA\_DEV\_OMPDEV(v) for the device array and then access h\_data[i] or d\_data[i] within the loop.
- When accessing individual components of an N\_Vector v on the host remember to first copy the
  array back from the device with N\_VCopyFromDevice\_OpenMPDEV(v) to ensure the array is up
  to date.





# 7.11 NVECTOR Examples

There are NVector examples that may be installed for the implementations provided with SUNDIALS. Each implementation makes use of the functions in test\_nvector.c. These example functions show simple usage of the NVector family of functions. The input to the examples are the vector length, number of threads (if threaded implementation), and a print timing flag.

The following is a list of the example functions in test\_nvector.c:

- Test\_N\_VClone: Creates clone of vector and checks validity of clone.
- Test\_N\_VCloneEmpty: Creates clone of empty vector and checks validity of clone.
- Test\_N\_VCloneVectorArray: Creates clone of vector array and checks validity of cloned array.
- Test\_N\_VCloneVectorArray: Creates clone of empty vector array and checks validity of cloned array.
- Test\_N\_VGetArrayPointer: Get array pointer.
- Test\_N\_VSetArrayPointer: Allocate new vector, set pointer to new vector array, and check values.
- Test\_N\_VLinearSum Case 1a: Test y = x + y
- Test\_N\_VLinearSum Case 1b: Test y = -x + y
- Test\_N\_VLinearSum Case 1c: Test y = ax + y
- Test\_N\_VLinearSum Case 2a: Test x = x + y
- Test\_N\_VLinearSum Case 2b: Test x = x y
- Test\_N\_VLinearSum Case 2c: Test x = x + by
- Test\_N\_VLinearSum Case 3: Test z = x + y
- Test\_N\_VLinearSum Case 4a: Test z = x y
- Test\_N\_VLinearSum Case 4b: Test z = -x + y





- Test\_N\_VLinearSum Case 5a: Test z = x + by
- Test\_N\_VLinearSum Case 5b: Test z = ax + y
- Test\_N\_VLinearSum Case 6a: Test z = -x + by
- Test\_N\_VLinearSum Case 6b: Test z = ax y
- Test\_N\_VLinearSum Case 7: Test z = a(x + y)
- Test\_N\_VLinearSum Case 8: Test z = a(x y)
- Test\_N\_VLinearSum Case 9: Test z = ax + by
- Test\_N\_VConst: Fill vector with constant and check result.
- Test\_N\_VProd: Test vector multiply: z = x \* y
- Test\_N\_VDiv: Test vector division: z = x / y
- Test\_N\_VScale: Case 1: scale: x = cx
- Test\_N\_VScale: Case 2: copy: z = x
- Test\_N\_VScale: Case 3: negate: z = -x
- Test\_N\_VScale: Case 4: combination: z = cx
- Test\_N\_VAbs: Create absolute value of vector.
- Test\_N\_VAddConst: add constant vector: z = c + x
- Test\_N\_VDotProd: Calculate dot product of two vectors.
- Test\_N\_VMaxNorm: Create vector with known values, find and validate the max norm.
- Test\_N\_VWrmsNorm: Create vector of known values, find and validate the weighted root mean square.
- Test\_N\_VWrmsNormMask: Create vector of known values, find and validate the weighted root mean square using all elements except one.
- Test\_N\_VMin: Create vector, find and validate the min.
- Test\_N\_VWL2Norm: Create vector, find and validate the weighted Euclidean L2 norm.
- Test\_N\_VL1Norm: Create vector, find and validate the L1 norm.
- Test\_N\_VCompare: Compare vector with constant returning and validating comparison vector.
- Test\_N\_VConstrMask: Test mask of vector x with vector c.
- Test\_N\_VMinQuotient: Fill two vectors with known values. Calculate and validate minimum quotient.
- Test\_N\_VLinearCombination Case 1a: Test x = a x
- $\bullet$  Test\_N\_VLinearCombination Case 1b: Test  $z=a\ x$
- Test\_N\_VLinearCombination Case 2b: Test z = a x + b y

- Test\_N\_VLinearCombination Case 3a: Test x = x + a y + b z
- Test\_N\_VLinearCombination Case 3b: Test x = a x + b y + c z
- Test\_N\_VLinearCombination Case 3c: Test w = a x + b y + c z
- Test\_N\_VScaleAddMulti Case 1a: y = a x + y
- Test\_N\_VScaleAddMulti Case 1b: z = a x + y
- Test\_N\_VScaleAddMulti Case 2a: Y[i] = c[i] x + Y[i], i = 1,2,3
- Test\_N\_VDotProdMulti Case 1: Calculate the dot product of two vectors
- Test\_N\_VDotProdMulti Case 2: Calculate the dot product of one vector with three other vectors in a vector array.
- Test\_N\_VLinearSumVectorArray Case 1: z = a x + b y
- Test\_N\_VLinearSumVectorArray Case 2a: Z[i] = a X[i] + b Y[i]
- Test\_N\_VLinearSumVectorArray Case 2b: X[i] = a X[i] + b Y[i]
- Test\_N\_VLinearSumVectorArray Case 2c: Y[i] = a X[i] + b Y[i]
- Test\_N\_VScaleVectorArray Case 1a: y = c y
- Test\_N\_VScaleVectorArray Case 1b: z = c y
- Test\_N\_VScaleVectorArray Case 2a: Y[i] = c[i] Y[i]
- Test\_N\_VScaleVectorArray Case 2b: Z[i] = c[i] Y[i]
- Test\_N\_VScaleVectorArray Case 1a: z = c
- Test\_N\_VScaleVectorArray Case 1b: Z[i] = c
- Test\_N\_VWrmsNormVectorArray Case 1a: Create a vector of know values, find and validate the weighted root mean square norm.
- Test\_N\_VWrmsNormVectorArray Case 1b: Create a vector array of three vectors of know values, find and validate the weighted root mean square norm of each.
- Test\_N\_VWrmsNormMaskVectorArray Case 1a: Create a vector of know values, find and validate the weighted root mean square norm using all elements except one.
- Test\_N\_VWrmsNormMaskVectorArray Case 1b: Create a vector array of three vectors of know values, find and validate the weighted root mean square norm of each using all elements except one.
- Test\_N\_VScaleAddMultiVectorArray Case 1a: y = a x + y
- Test\_N\_VScaleAddMultiVectorArray Case 1b: z = a x + y
- Test\_N\_VScaleAddMultiVectorArray Case 2a: Y[j][0] = a[j] X[0] + Y[j][0]
- Test\_N\_VScaleAddMultiVectorArray Case 3a: Y[0][i] = a[0] X[i] + Y[0][i]
- Test\_N\_VScaleAddMultiVectorArray Case 3b: Z[0][i] = a[0] X[i] + Y[0][i]

- Test\_N\_VScaleAddMultiVectorArray Case 4a: Y[j][i] = a[j] X[i] + Y[j][i]
- ullet Test\_N\_VLinearCombinationVectorArray Case 1a:  $x=a\ x$
- ullet Test\_N\_VLinearCombinationVectorArray Case 1b:  $z=a\ x$
- Test\_N\_VLinearCombinationVectorArray Case 2a: x = a x + b y
- Test\_N\_VLinearCombinationVectorArray Case 2b: z = a x + b y
- Test\_N\_VLinearCombinationVectorArray Case 3b: w = a x + b y + c z
- Test\_N\_VLinearCombinationVectorArray Case 4a: X[0][i] = c[0] X[0][i]
- Test\_N\_VLinearCombinationVectorArray Case 4b: Z[i] = c[0] X[0][i]
- $\bullet \ \, \mathsf{Test\_N\_VLinearCombinationVectorArray} \ \, \mathsf{Case} \ \, \mathsf{5a:} \ \, \mathsf{X}[0][i] = \mathsf{c}[0] \ \, \mathsf{X}[0][i] + \mathsf{c}[1] \ \, \mathsf{X}[1][i]$
- Test\_N\_VLinearCombinationVectorArray Case 5b: Z[i] = c[0] X[0][i] + c[1] X[1][i]
- $\bullet \ \, \mathsf{Test\_N\_VLinearCombinationVectorArray} \ \, \mathsf{Case} \ \, 6a: \ \, \mathsf{X}[0][i] = \mathsf{X}[0][i] + c[1] \ \, \mathsf{X}[1][i] + c[2] \ \, \mathsf{X}[2][i]$
- $\bullet \ \, \mathsf{Test\_N\_VLinearCombinationVectorArray} \ \, \mathsf{Case} \ \, 6c: \ \, \mathsf{Z[i]} = \mathsf{c[0]} \ \, \mathsf{X[0][i]} + \mathsf{c[1]} \ \, \mathsf{X[1][i]} + \mathsf{c[2]} \ \, \mathsf{X[2][i]}$

Table 7.5: List of vector functions usage by IDAS code modules

	IDAS	IDALS	IDABBDPRE	IDAA
N_VGetVectorID				
N_VClone	<b>√</b>	<b>√</b>	<b>√</b>	<b>V</b>
N_VCloneEmpty		1		
N_VDestroy	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>
N_VCloneVectorArray	<b>√</b>			<b>V</b>
N_VDestroyVectorArray	<b>√</b>			<b>1</b>
N_VSpace	<b>√</b>	2		
N_VGetArrayPointer		1	<b>√</b>	
N_VSetArrayPointer		1		
$N_{-}VLinearSum$	<b>√</b>	<b>√</b>		<b></b>
N_VConst	<b>√</b>	<b>√</b>		<b>√</b>
N_VProd	<b>√</b>			
N_VDiv	<b>√</b>			
N_VScale	<b>√</b>	<b>√</b>	<b>√</b>	<b>√</b>
$N_{-}VAbs$	<b>√</b>			
$N_{-}VInv$	<b>√</b>			
${\tt N\_VAddConst}$	<b>√</b>			
$N_{V}$ DotProd		<b>√</b>		
$N_{ m L}VMaxNorm$	<b>√</b>			
$N_{ m L}VWrmsNorm$	<b>√</b>	<b>√</b>		
N_VMin	<b>√</b>			
${ t NVMinQuotient}$	<b>√</b>			
N_VConstrMask	<b>√</b>			
N_VWrmsNormMask	<b>√</b>			
$N_{-}VCompare$	<b>√</b>			
N_VLinearCombination	<b>_</b>			
N_VScaleAddMulti	<b>√</b>			
$N_{-}VDotProdMulti$		3		П
N_VLinearSumVectorArray	<b>\</b>		İ	Ħ
N_VScaleVectorArray	<b>√</b>			-
N_VConstVectorArray	<b>√</b>			
N_VWrmsNormVectorArray	<b>√</b>			
N_VWrmsNormMaskVectorArray	<b>√</b>			П
N_VScaleAddMultiVectorArray	<b>√</b>			П
N_VLinearCombinationVectorArray	<b>V</b>			

# Chapter 8

# Description of the SUNMatrix module

For problems that involve direct methods for solving linear systems, the SUNDIALS solvers not only operate on generic vectors, but also on generic matrices (of type SUNMatrix), through a set of operations defined by the particular SUNMATRIX implementation. Users can provide their own specific implementation of the SUNMATRIX module, particularly in cases where they provide their own NVECTOR and/or linear solver modules, and require matrices that are compatible with those implementations. Alternately, we provide three SUNMATRIX implementations: dense, banded, and sparse. The generic operations are described below, and descriptions of the implementations provided with SUNDIALS follow.

The generic SUNMatrix type has been modeled after the object-oriented style of the generic N\_Vector type. Specifically, a generic SUNMatrix is a pointer to a structure that has an implementation-dependent *content* field containing the description and actual data of the matrix, and an *ops* field pointing to a structure with generic matrix operations. The type SUNMatrix is defined as

```
typedef struct _generic_SUNMatrix *SUNMatrix;
struct _generic_SUNMatrix {
    void *content;
    struct _generic_SUNMatrix_Ops *ops;
};
```

The \_generic\_SUNMatrix\_Ops structure is essentially a list of pointers to the various actual matrix operations, and is defined as

```
struct _generic_SUNMatrix_Ops {
  SUNMatrix_ID (*getid)(SUNMatrix);
  SUNMatrix
               (*clone)(SUNMatrix);
  void
               (*destroy)(SUNMatrix);
  int
               (*zero)(SUNMatrix);
  int
               (*copy)(SUNMatrix, SUNMatrix);
               (*scaleadd)(realtype, SUNMatrix, SUNMatrix);
  int
  int
               (*scaleaddi)(realtype, SUNMatrix);
  int
               (*matvec)(SUNMatrix, N_Vector, N_Vector);
  int
               (*space)(SUNMatrix, long int*, long int*);
};
```

The generic SUNMATRIX module defines and implements the matrix operations acting on SUNMatrix objects. These routines are nothing but wrappers for the matrix operations defined by a particular SUNMATRIX implementation, which are accessed through the *ops* field of the SUNMatrix structure. To

Matrix ID	Matrix type	ID Value
SUNMATRIX_DENSE	Dense $M \times N$ matrix	0
SUNMATRIX_BAND	Band $M \times M$ matrix	1
SUNMATRIX_SPARSE	Sparse (CSR or CSC) $M \times N$ matrix	2
SUNMATRIX_CUSTOM	User-provided custom matrix	3

Table 8.1: Identifiers associated with matrix kernels supplied with SUNDIALS.

illustrate this point we show below the implementation of a typical matrix operation from the generic SUNMATRIX module, namely SUNMatZero, which sets all values of a matrix A to zero, returning a flag denoting a successful/failed operation:

```
int SUNMatZero(SUNMatrix A)
{
  return((int) A->ops->zero(A));
}
```

Table 8.2 contains a complete list of all matrix operations defined by the generic Sunmatrix module. A particular implementation of the Sunmatrix module must:

- Specify the *content* field of the SUNMatrix object.
- Define and implement a minimal subset of the matrix operations. See the documentation for each SUNDIALS solver to determine which SUNMATRIX operations they require.
  - Note that the names of these routines should be unique to that implementation in order to permit using more than one SUNMATRIX module (each with different SUNMatrix internal data representations) in the same code.
- Define and implement user-callable constructor and destructor routines to create and free a SUNMatrix with the new *content* field and with *ops* pointing to the new matrix operations.
- Optionally, define and implement additional user-callable routines acting on the newly defined SUNMatrix (e.g., a routine to print the content for debugging purposes).
- Optionally, provide accessor macros or functions as needed for that particular implementation to access different parts of the *content* field of the newly defined SUNMatrix.

Each SUNMATRIX implementation included in SUNDIALS has a unique identifier specified in enumeration and shown in Table 8.1. It is recommended that a user-supplied SUNMATRIX implementation use the SUNMATRIX\_CUSTOM identifier.

Name	Usage and Description
SUNMatGetID	id = SUNMatGetID(A); Returns the type identifier for the matrix A. It is used to determine the matrix implementation type (e.g. dense, banded, sparse,) from the abstract SUNMatrix interface. This is used to assess compatibility with SUNDIALS-provided linear solver implementations. Returned values are given in the Table 8.1.
	continued on next page

Table 8.2: Description of the SUNMatrix operations

Name	Usage and Description
SUNMatClone	B = SUNMatClone(A); Creates a new SUNMatrix of the same type as an existing matrix A and sets the <i>ops</i> field. It does not copy the matrix, but rather allocates storage for the new matrix.
SUNMatDestroy	SUNMatDestroy(A); Destroys the SUNMatrix A and frees memory allocated for its internal data.
SUNMatSpace	ier = SUNMatSpace(A, &lrw, &liw); Returns the storage requirements for the matrix A. lrw is a long int containing the number of realtype words and liw is a long int containing the number of integer words. The return value is an integer flag denoting success/failure of the operation.  This function is advisory only, for use in determining a user's total space requirements; it could be a dummy function in a user-supplied SUNMATRIX module if that information is not of interest.
SUNMatZero	ier = SUNMatZero(A); Performs the operation $A_{ij} = 0$ for all entries of the matrix $A$ . The return value is an integer flag denoting success/failure of the operation.
SUNMatCopy	ier = SUNMatCopy(A,B); Performs the operation $B_{ij} = A_{i,j}$ for all entries of the matrices $A$ and $B$ . The return value is an integer flag denoting success/failure of the operation.
SUNMatScaleAdd	ier = SUNMatScaleAdd(c, A, B); Performs the operation $A = cA + B$ . The return value is an integer flag denoting success/failure of the operation.
SUNMatScaleAddI	ier = SUNMatScaleAddI(c, A); Performs the operation $A = cA + I$ . The return value is an integer flag denoting success/failure of the operation.
SUNMatMatvec	ier = SUNMatMatvec(A, x, y); Performs the matrix-vector product operation, $y = Ax$ . It should only be called with vectors x and y that are compatible with the matrix A – both in storage type and dimensions. The return value is an integer flag denoting success/failure of the operation.

We note that not all SUNMATRIX types are compatible with all NVECTOR types provided with SUNDIALS. This is primarily due to the need for compatibility within the SUNMatMatvec routine; however, compatibility between SUNMATRIX and NVECTOR implementations is more crucial when considering their interaction within SUNLINSOL objects, as will be described in more detail in Chapter 9. More specifically, in Table 8.3 we show the matrix interfaces available as SUNMATRIX modules, and the compatible vector implementations.

Table 8.3: SUNDIALS matrix interfaces and vector implementations that can be used for each.

Matrix Interface	Serial	Parallel (MPI)	OpenMP	pThreads	hypre Vec.	PETSC Vec.	CUDA	RAJA	User Suppl.
Dense	✓		✓	✓					✓
continued on next page									

Matrix Interface	Serial	Parallel (MPI)	OpenMP	pThreads	hypre Vec.	PETSC Vec.	CUDA	RAJA	User Suppl.
		(11111)			vec.	vec.			Suppi.
Band	✓		✓	√					✓
Sparse	✓		✓	✓					✓
User supplied	✓	$\checkmark$	✓	✓	$\checkmark$	✓	✓	✓	✓

# 8.1 SUNMatrix functions used by IDAS

In Table 8.4, we list the matrix functions in the SUNMATRIX module used within the IDAS package. The table also shows, for each function, which of the code modules uses the function. The main IDAS integrator does not call any SUNMATRIX functions directly, so the table columns are specific to the IDALS interface and the IDABBDPRE preconditioner module. We further note that the IDALS interface only utilizes these routines when supplied with a *matrix-based* linear solver, i.e., the SUNMATRIX object passed to IDASetLinearSolver was not NULL.

At this point, we should emphasize that the IDAS user does not need to know anything about the usage of matrix functions by the IDAS code modules in order to use IDAS. The information is presented as an implementation detail for the interested reader.

	IDALS	IDABBDPRE
SUNMatGetID	<b>√</b>	
SUNMatDestroy		<b>√</b>
SUNMatZero	<b>√</b>	<b>√</b>

Table 8.4: List of matrix functions usage by IDAS code modules

The matrix functions listed in Table 8.2 with a † symbol are optionally used, in that these are only called if they are implemented in the SUNMATRIX module that is being used (i.e. their function pointers are non-NULL). The matrix functions listed in Table 8.2 that are *not* used by IDAS are: SUNMatCopy, SUNMatClone, SUNMatScaleAdd, SUNMatScaleAddI and SUNMatMatvec. Therefore a user-supplied SUNMATRIX module for IDAS could omit these functions.

# 8.2 The SUNMatrix\_Dense implementation

The dense implementation of the SUNMATRIX module provided with SUNDIALS, SUNMATRIX\_DENSE, defines the *content* field of SUNMatrix to be the following structure:

```
struct _SUNMatrixContent_Dense {
   sunindextype M;
   sunindextype N;
   realtype *data;
   sunindextype ldata;
   realtype **cols;
};
```

These entries of the *content* field contain the following information:

- M number of rows
- N number of columns

data - pointer to a contiguous block of realtype variables. The elements of the dense matrix are stored columnwise, i.e. the (i,j)-th element of a dense SUNMATRIX A (with  $0 \le i < M$  and  $0 \le j < N$ ) may be accessed via data[j\*M+i].

ldata - length of the data array (=  $M \cdot N$ ).

cols - array of pointers. cols[j] points to the first element of the j-th column of the matrix in the array data. The (i,j)-th element of a dense SUNMATRIX A (with  $0 \le i < M$  and  $0 \le j < N$ ) may be accessed via cols[j][i].

The header file to include when using this module is sunmatrix\_dense.h. The SUNMATRIX\_DENSE module is accessible from all SUNDIALS solvers without linking to the libsundials\_sunmatrixdense module library.

#### 8.2.1 SUNMatrix\_Dense accessor macros

The following macros are provided to access the content of a SUNMATRIX\_DENSE matrix. The prefix  $SM_{-}$  in the names denotes that these macros are for SUNMatrix implementations, and the suffix  $_{-}D$  denotes that these are specific to the dense version.

#### • SM\_CONTENT\_D

This macro gives access to the contents of the dense SUNMatrix.

The assignment  $A\_cont = SM\_CONTENT\_D(A)$  sets  $A\_cont$  to be a pointer to the dense SUNMatrix content structure.

Implementation:

```
#define SM_CONTENT_D(A) ( (SUNMatrixContent_Dense)(A->content) )
```

• SM\_ROWS\_D, SM\_COLUMNS\_D, and SM\_LDATA\_D

These macros give individual access to various lengths relevant to the content of a dense SUNMatrix.

These may be used either to retrieve or to set these values. For example, the assignment A\_rows = SM\_ROWS\_D(A) sets A\_rows to be the number of rows in the matrix A. Similarly, the assignment SM\_COLUMNS\_D(A) = A\_cols sets the number of columns in A to equal A\_cols.

Implementation:

```
#define SM_ROWS_D(A) ( SM_CONTENT_D(A)->M )
#define SM_COLUMNS_D(A) ( SM_CONTENT_D(A)->N )
#define SM_LDATA_D(A) ( SM_CONTENT_D(A)->1data )
```

#### • SM\_DATA\_D and SM\_COLS\_D

These macros give access to the data and cols pointers for the matrix entries.

The assignment A\_data = SM\_DATA\_D(A) sets A\_data to be a pointer to the first component of the data array for the dense SUNMatrix A. The assignment SM\_DATA\_D(A) = A\_data sets the data array of A to be A\_data by storing the pointer A\_data.

Similarly, the assignment A\_cols = SM\_COLS\_D(A) sets A\_cols to be a pointer to the array of column pointers for the dense SUNMatrix A. The assignment SM\_COLS\_D(A) = A\_cols sets the column pointer array of A to be A\_cols by storing the pointer A\_cols.

Implementation:

```
#define SM_DATA_D(A) ( SM_CONTENT_D(A)->data )
#define SM_COLS_D(A) ( SM_CONTENT_D(A)->cols )
```

#### • SM\_COLUMN\_D and SM\_ELEMENT\_D

These macros give access to the individual columns and entries of the data array of a dense SUNMatrix.

The assignment col\_j = SM\_COLUMN\_D(A,j) sets col\_j to be a pointer to the first entry of the j-th column of the M  $\times$  N dense matrix A (with  $0 \le j < N$ ). The type of the expression SM\_COLUMN\_D(A,j) is realtype \*. The pointer returned by the call SM\_COLUMN\_D(A,j) can be treated as an array which is indexed from 0 to M - 1.

The assignments SM\_ELEMENT\_D(A,i,j) = a\_ij and a\_ij = SM\_ELEMENT\_D(A,i,j) reference the (i,j)-th element of the M × N dense matrix A (with  $0 \le i < M$  and  $0 \le j < N$ ).

Implementation:

```
#define SM_COLUMN_D(A,j) ( (SM_CONTENT_D(A)->cols)[j] )
#define SM_ELEMENT_D(A,i,j) ( (SM_CONTENT_D(A)->cols)[j][i] )
```

#### 8.2.2 SUNMatrix\_Dense functions

The SUNMATRIX\_DENSE module defines dense implementations of all matrix operations listed in Table 8.2. Their names are obtained from those in Table 8.2 by appending the suffix \_Dense (e.g. SUNMatCopy\_Dense). All the standard matrix operations listed in 8.2 with the suffix \_Dense appended are callable via the FORTRAN 2003 interface by prepending an 'F' (e.g. FSUNMatCopy\_Dense).

The module SUNMATRIX\_DENSE provides the following additional user-callable routines:

#### SUNDenseMatrix

Prototype SUNMatrix SUNDenseMatrix(sunindextype M, sunindextype N)

Description This constructor function creates and allocates memory for a dense SUNMatrix. Its arguments are the number of rows, M, and columns, N, for the dense matrix.

F2003 Name This function is callable as FSUNDenseMatrix when using the Fortran 2003 interface module.

#### SUNDenseMatrix\_Print

Prototype void SUNDenseMatrix\_Print(SUNMatrix A, FILE\* outfile)

Description This function prints the content of a dense SUNMatrix to the output stream specified by outfile. Note: stdout or stderr may be used as arguments for outfile to print directly to standard output or standard error, respectively.

#### SUNDenseMatrix\_Rows

Prototype sunindextype SUNDenseMatrix\_Rows(SUNMatrix A)

Description This function returns the number of rows in the dense SUNMatrix.

F2003 Name This function is callable as FSUNDenseMatrix\_Rows when using the Fortran 2003 interface module.

## SUNDenseMatrix\_Columns

Prototype sunindextype SUNDenseMatrix\_Columns(SUNMatrix A)

Description This function returns the number of columns in the dense SUNMatrix.

F2003 Name This function is callable as FSUNDenseMatrix\_Columns when using the Fortran 2003 interface module.

#### SUNDenseMatrix\_LData

Prototype sunindextype SUNDenseMatrix\_LData(SUNMatrix A)

Description This function returns the length of the data array for the dense SUNMatrix.

F2003 Name This function is callable as FSUNDenseMatrix\_LData when using the Fortran 2003 inter-

face module.

#### SUNDenseMatrix\_Data

Prototype realtype\* SUNDenseMatrix\_Data(SUNMatrix A)

Description This function returns a pointer to the data array for the dense SUNMatrix.

F2003 Name This function is callable as FSUNDenseMatrix\_Data when using the Fortran 2003 inter-

face module.

#### SUNDenseMatrix\_Cols

Prototype realtype\*\* SUNDenseMatrix\_Cols(SUNMatrix A)

Description This function returns a pointer to the cols array for the dense SUNMatrix.

#### SUNDenseMatrix\_Column

Prototype realtype\* SUNDenseMatrix\_Column(SUNMatrix A, sunindextype j)

 $\label{prop:control} Description \quad This function \ returns \ a \ pointer \ to \ the \ first \ entry \ of \ the \ jth \ column \ of \ the \ dense \ {\tt SUNMatrix}.$ 

The resulting pointer should be indexed over the range 0 to M-1.

F2003 Name This function is callable as FSUNDenseMatrix\_Column when using the Fortran 2003 in-

terface module.

#### Notes

- When looping over the components of a dense SUNMatrix A, the most efficient approaches are to:
  - First obtain the component array via A\_data = SM\_DATA\_D(A) or A\_data = SUNDenseMatrix\_Data(A) and then access A\_data[i] within the loop.
  - First obtain the array of column pointers via A\_cols = SM\_COLS\_D(A) or
     A\_cols = SUNDenseMatrix\_Cols(A), and then access A\_cols[j][i] within the loop.
  - Within a loop over the columns, access the column pointer via
     A\_colj = SUNDenseMatrix\_Column(A,j) and then to access the entries within that column using A\_colj[i] within the loop.

All three of these are more efficient than using SM\_ELEMENT\_D(A,i,j) within a double loop.

• Within the SUNMatMatvec\_Dense routine, internal consistency checks are performed to ensure that the matrix is called with consistent NVECTOR implementations. These are currently limited to: NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS. As additional compatible vector implementations are added to SUNDIALS, these will be included within this compatibility check.

# 8.2.3 SUNMatrix\_Dense Fortran interfaces

The SUNMATRIX\_DENSE module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.



#### FORTRAN 2003 interface module

The fsunmatrix\_dense\_mod FORTRAN module defines interfaces to most SUNMATRIX\_DENSE C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNDenseMatrix is interfaced as FSUNDenseMatrix.

The Fortran 2003 sunmatrix\_dense\_mod, and linking to the library libsundials\_fsunmatrix\_dense\_mod. lib in addition to the C library. For details on where the library and module file fsunmatrix\_dense\_mod.mod are installed see Appendix A. We note that the module is accessible from the Fortran 2003 sundials integrators without separately linking to the libsundials\_fsunmatrixdense\_mod library.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN interface module, the SUNMATRIX\_DENSE module also includes the FORTRAN-callable function FSUNDenseMatInit(code, M, N, ier) to initialize this SUNMATRIX\_DENSE module for a given SUNDIALS solver. Here code is an integer input solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, 4 for ARKODE); M and N are the corresponding dense matrix construction arguments (declared to match C type long int); and ier is an error return flag equal to 0 for success and -1 for failure. Both code and ier are declared to match C type int. Additionally, when using ARKODE with a non-identity mass matrix, the FORTRAN-callable function FSUNDenseMassMatInit(M, N, ier) initializes this SUNMATRIX\_DENSE module for storing the mass matrix.

# 8.3 The SUNMatrix\_Band implementation

The banded implementation of the SUNMATRIX module provided with SUNDIALS, SUNMATRIX\_BAND, defines the *content* field of SUNMatrix to be the following structure:

```
struct _SUNMatrixContent_Band {
   sunindextype M;
   sunindextype N;
   sunindextype mu;
   sunindextype ml;
   sunindextype s_mu;
   sunindextype ldim;
   realtype *data;
   sunindextype ldata;
   realtype **cols;
};
```

A diagram of the underlying data representation in a banded matrix is shown in Figure 8.1. A more complete description of the parts of this *content* field is given below:

```
complete description of the parts of this content field is given below:
M - number of rows
N - number of columns (N = M)
mu - upper half-bandwidth, 0 ≤ mu < N</li>
ml - lower half-bandwidth, 0 ≤ ml < N</li>
s_mu - storage upper bandwidth, mu ≤ s_mu < N. The LU decomposition routines in the associated SUNLINSOL_BAND and SUNLINSOL_LAPACKBAND modules write the LU factors into the storage for A. The upper triangular factor U, however, may have an upper bandwidth as big as min(N-1,mu+ml) because of partial pivoting. The s_mu field holds the upper half-bandwidth allocated for A.</li>
ldim - leading dimension (ldim ≥ s_mu+ml+1)
```

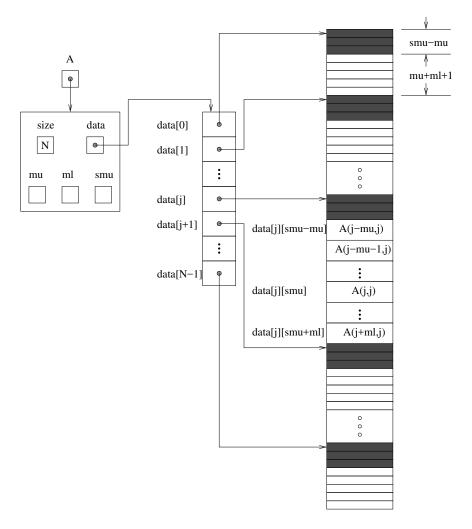


Figure 8.1: Diagram of the storage for the SUNMATRIX\_BAND module. Here A is an  $\mathbb{N} \times \mathbb{N}$  band matrix with upper and lower half-bandwidths mu and ml, respectively. The rows and columns of A are numbered from 0 to  $\mathbb{N}-1$  and the (i,j)-th element of A is denoted A(i,j). The greyed out areas of the underlying component storage are used by the associated SUNLINSOL\_BAND linear solver.

data - pointer to a contiguous block of realtype variables. The elements of the banded matrix are stored columnwise (i.e. columns are stored one on top of the other in memory). Only elements within the specified half-bandwidths are stored. data is a pointer to ldata contiguous locations which hold the elements within the band of A.

ldata - length of the data array (=  $ldim \cdot N$ )

cols - array of pointers. cols[j] is a pointer to the uppermost element within the band in the j-th column. This pointer may be treated as an array indexed from s\_mu-mu (to access the uppermost element within the band in the j-th column) to s\_mu+ml (to access the lowest element within the band in the j-th column). Indices from 0 to s\_mu-mu-1 give access to extra storage elements required by the LU decomposition function. Finally, cols[j][i-j+s\_mu] is the (i,j)-th element with j-mu  $\leq i \leq j$ +ml.

The header file to include when using this module is sunmatrix/sunmatrix\_band.h. The SUNMATRIX\_BAND module is accessible from all SUNDIALS solvers without linking to the libsundials\_sunmatrixband module library.

#### 8.3.1 SUNMatrix Band accessor macros

The following macros are provided to access the content of a SUNMATRIX\_BAND matrix. The prefix SM\_ in the names denotes that these macros are for *SUNMatrix* implementations, and the suffix \_B denotes that these are specific to the *banded* version.

#### • SM\_CONTENT\_B

This routine gives access to the contents of the banded SUNMatrix.

The assignment A\_cont = SM\_CONTENT\_B(A) sets A\_cont to be a pointer to the banded SUNMatrix content structure.

Implementation:

```
#define SM_CONTENT_B(A) ((SUNMatrixContent_Band)(A->content) )
```

• SM\_ROWS\_B, SM\_COLUMNS\_B, SM\_UBAND\_B, SM\_LBAND\_B, SM\_SUBAND\_B, SM\_LDIM\_B, and SM\_LDATA\_B

These macros give individual access to various lengths relevant to the content of a banded SUNMatrix.

These may be used either to retrieve or to set these values. For example, the assignment A\_rows = SM\_ROWS\_B(A) sets A\_rows to be the number of rows in the matrix A. Similarly, the assignment SM\_COLUMNS\_B(A) = A\_cols sets the number of columns in A to equal A\_cols.

Implementation:

```
#define SM_ROWS_B(A) ( SM_CONTENT_B(A)->M )
#define SM_COLUMNS_B(A) ( SM_CONTENT_B(A)->N )
#define SM_UBAND_B(A) ( SM_CONTENT_B(A)->mu )
#define SM_LBAND_B(A) ( SM_CONTENT_B(A)->ml )
#define SM_SUBAND_B(A) ( SM_CONTENT_B(A)->s_mu )
#define SM_LDIM_B(A) ( SM_CONTENT_B(A)->ldim )
#define SM_LDATA_B(A) ( SM_CONTENT_B(A)->ldata )
```

#### • SM\_DATA\_B and SM\_COLS\_B

These macros give access to the data and cols pointers for the matrix entries.

The assignment A\_data = SM\_DATA\_B(A) sets A\_data to be a pointer to the first component of the data array for the banded SUNMatrix A. The assignment SM\_DATA\_B(A) = A\_data sets the data array of A to be A\_data by storing the pointer A\_data.

Similarly, the assignment  $A\_cols = SM\_COLS\_B(A)$  sets  $A\_cols$  to be a pointer to the array of column pointers for the banded SUNMatrix A. The assignment  $SM\_COLS\_B(A) = A\_cols$  sets the column pointer array of A to be  $A\_cols$  by storing the pointer  $A\_cols$ .

Implementation:

```
#define SM_DATA_B(A) ( SM_CONTENT_B(A)->data )
#define SM_COLS_B(A) ( SM_CONTENT_B(A)->cols )
```

• SM\_COLUMN\_B, SM\_COLUMN\_ELEMENT\_B, and SM\_ELEMENT\_B

These macros give access to the individual columns and entries of the data array of a banded SUNMatrix.

The assignments SM\_ELEMENT\_B(A,i,j) = a\_ij and a\_ij = SM\_ELEMENT\_B(A,i,j) reference the (i,j)-th element of the N × N band matrix A, where  $0 \le i, j \le N-1$ . The location (i,j) should further satisfy  $j-mu \le i \le j+ml$ .

The assignment col\_j = SM\_COLUMN\_B(A,j) sets col\_j to be a pointer to the diagonal element of the j-th column of the N  $\times$  N band matrix A,  $0 \le j \le N - 1$ . The type of the expression

 $SM\_COLUMN\_B(A,j)$  is realtype \*. The pointer returned by the call  $SM\_COLUMN\_B(A,j)$  can be treated as an array which is indexed from -mu to ml.

The assignments SM\_COLUMN\_ELEMENT\_B(col\_j,i,j) = a\_ij and a\_ij = SM\_COLUMN\_ELEMENT\_B(col\_j,i,j) reference the (i,j)-th entry of the band matrix A when used in conjunction with SM\_COLUMN\_B to reference the j-th column through col\_j. The index (i,j) should satisfy  $j-mu \le i \le j+ml$ .

Implementation:

# 8.3.2 SUNMatrix\_Band functions

The SUNMATRIX\_BAND module defines banded implementations of all matrix operations listed in Table 8.2. Their names are obtained from those in Table 8.2 by appending the suffix \_Band (e.g. SUNMatCopy\_Band). All the standard matrix operations listed in 8.2 with the suffix \_Band appended are callable via the FORTRAN 2003 interface by prepending an 'F' (e.g. FSUNMatCopy\_Band).

The module SUNMATRIX\_BAND provides the following additional user-callable routines:

#### SUNBandMatrix

Prototype SUNMatrix SUNBandMatrix(sunindextype N, sunindextype mu, sunindextype ml)

Description This constructor function creates and allocates memory for a banded SUNMatrix. Its arguments are the matrix size, N, and the upper and lower half-bandwidths of the matrix, mu and ml. The stored upper bandwidth is set to mu+ml to accommodate subsequent factorization in the SUNLINSOL\_BAND and SUNLINSOL\_LAPACKBAND modules.

F2003 Name This function is callable as FSUNBandMatrix when using the Fortran 2003 interface module.

#### SUNBandMatrixStorage

Prototype SUNMatrix SUNBandMatrixStorage(sunindextype N, sunindextype mu, sunindextype ml, sunindextype smu)

Description This constructor function creates and allocates memory for a banded SUNMatrix. Its arguments are the matrix size, N, the upper and lower half-bandwidths of the matrix, mu and ml, and the stored upper bandwidth, smu. When creating a band SUNMatrix, this value should be

- at least min(N-1,mu+ml) if the matrix will be used by the SUNLINSOL\_BAND module;
- exactly equal to mu+ml if the matrix will be used by the SUNLINSOL\_LAPACKBAND module:
- at least mu if used in some other manner.

Note: it is strongly recommended that users call the default constructor, SUNBandMatrix, in all standard use cases. This advanced constructor is used internally within SUNDIALS solvers, and is provided to users who require banded matrices for non-default purposes.

# SUNBandMatrix\_Print

Prototype void SUNBandMatrix\_Print(SUNMatrix A, FILE\* outfile)

Description This function prints the content of a banded SUNMatrix to the output stream specified

by outfile. Note: stdout or stderr may be used as arguments for outfile to print

directly to standard output or standard error, respectively.

#### SUNBandMatrix\_Rows

Prototype sunindextype SUNBandMatrix\_Rows(SUNMatrix A)

Description This function returns the number of rows in the banded SUNMatrix.

F2003 Name This function is callable as FSUNBandMatrix\_Rows when using the Fortran 2003 interface

module.

#### SUNBandMatrix\_Columns

Prototype sunindextype SUNBandMatrix\_Columns(SUNMatrix A)

Description This function returns the number of columns in the banded SUNMatrix.

F2003 Name This function is callable as FSUNBandMatrix\_Columns when using the Fortran 2003 in-

terface module.

# SUNBandMatrix\_LowerBandwidth

Prototype sunindextype SUNBandMatrix\_LowerBandwidth(SUNMatrix A)

Description This function returns the lower half-bandwidth of the banded SUNMatrix.

F2003 Name This function is callable as FSUNBandMatrix\_LowerBandwidth when using the Fortran

2003 interface module.

# ${\tt SUNBandMatrix\_UpperBandwidth}$

Prototype sunindextype SUNBandMatrix\_UpperBandwidth(SUNMatrix A)

Description This function returns the upper half-bandwidth of the banded SUNMatrix.

F2003 Name This function is callable as FSUNBandMatrix\_UpperBandwidth when using the Fortran

2003 interface module.

#### SUNBandMatrix\_StoredUpperBandwidth

Prototype sunindextype SUNBandMatrix\_StoredUpperBandwidth(SUNMatrix A)

Description This function returns the stored upper half-bandwidth of the banded SUNMatrix.

F2003 Name This function is callable as FSUNBandMatrix\_StoredUpperBandwidth when using the

Fortran 2003 interface module.

#### SUNBandMatrix\_LDim

Prototype sunindextype SUNBandMatrix\_LDim(SUNMatrix A)

Description This function returns the length of the leading dimension of the banded SUNMatrix.

F2003 Name This function is callable as FSUNBandMatrix\_LDim when using the Fortran 2003 interface

module.

#### SUNBandMatrix\_Data

Prototype realtype\* SUNBandMatrix\_Data(SUNMatrix A)

Description This function returns a pointer to the data array for the banded SUNMatrix.

F2003 Name This function is callable as FSUNBandMatrix\_Data when using the Fortran 2003 interface module.

# SUNBandMatrix\_Cols

Prototype realtype\*\* SUNBandMatrix\_Cols(SUNMatrix A)

Description This function returns a pointer to the cols array for the banded SUNMatrix.

#### SUNBandMatrix\_Column

Prototype realtype\* SUNBandMatrix\_Column(SUNMatrix A, sunindextype j)

Description This function returns a pointer to the diagonal entry of the j-th column of the banded SUNMatrix. The resulting pointer should be indexed over the range -mu to ml.

F2003 Name This function is callable as FSUNBandMatrix\_Column when using the Fortran 2003 interface module.

#### Notes

- When looping over the components of a banded SUNMatrix A, the most efficient approaches are to:
  - First obtain the component array via A\_data = SM\_DATA\_B(A) or
     A\_data = SUNBandMatrix\_Data(A) and then access A\_data[i] within the loop.
  - First obtain the array of column pointers via A\_cols = SM\_COLS\_B(A) or
     A\_cols = SUNBandMatrix\_Cols(A), and then access A\_cols[j][i] within the loop.
  - Within a loop over the columns, access the column pointer via
     A\_colj = SUNBandMatrix\_Column(A,j) and then to access the entries within that column using SM\_COLUMN\_ELEMENT\_B(A\_colj,i,j).

All three of these are more efficient than using SM\_ELEMENT\_B(A,i,j) within a double loop.

• Within the SUNMatMatvec\_Band routine, internal consistency checks are performed to ensure that the matrix is called with consistent NVECTOR implementations. These are currently limited to: NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS. As additional compatible vector implementations are added to SUNDIALS, these will be included within this compatibility check.

#### 8.3.3 SUNMatrix\_Band Fortran interfaces

The SUNMATRIX\_BAND module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

#### FORTRAN 2003 interface module

The fsunmatrix\_band\_mod FORTRAN module defines interfaces to most SUNMATRIX\_BAND C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNBandMatrix is interfaced as FSUNBandMatrix.

The FORTRAN 2003 SUNMATRIX\_BAND interface module can be accessed with the use statement, i.e. use fsunmatrix\_band\_mod, and linking to the library libsundials\_fsunmatrixband\_mod.lib in



addition to the C library. For details on where the library and module file fsunmatrix\_band\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunmatrixband\_mod library.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN interface module, the SUNMATRIX\_BAND module also includes the FORTRAN-callable function FSUNBandMatInit(code, N, mu, ml, ier) to initialize this SUNMATRIX\_BAND module for a given SUNDIALS solver. Here code is an integer input solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, 4 for ARKODE); N, mu, and ml are the corresponding band matrix construction arguments (declared to match C type long int); and ier is an error return flag equal to 0 for success and -1 for failure. Both code and ier are declared to match C type int. Additionally, when using ARKODE with a non-identity mass matrix, the FORTRAN-callable function FSUNBandMassMatInit(N, mu, ml, ier) initializes this SUNMATRIX\_BAND module for storing the mass matrix.

# 8.4 The SUNMatrix\_Sparse implementation

The sparse implementation of the Sunmatrix module provided with Sundials, Sunmatrix\_sparse, is designed to work with either *compressed-sparse-column* (CSC) or *compressed-sparse-row* (CSR) sparse matrix formats. To this end, it defines the *content* field of Sunmatrix to be the following structure:

```
struct _SUNMatrixContent_Sparse {
   sunindextype M;
   sunindextype N;
   sunindextype NNZ;
   sunindextype NP;
   realtype *data;
   int sparsetype;
   sunindextype *indexvals;
   sunindextype *indexptrs;
   /* CSC indices */
   sunindextype **rowvals;
   sunindextype **colptrs;
   /* CSR indices */
   sunindextype **colvals;
   sunindextype **rowptrs;
};
```

A diagram of the underlying data representation for a CSC matrix is shown in Figure 8.2 (the CSR format is similar). A more complete description of the parts of this *content* field is given below:

```
- number of rows
            - number of columns
N
            - maximum number of nonzero entries in the matrix (allocated length of data and
NNZ
            indexvals arrays)
            - number of index pointers (e.g. number of column pointers for CSC matrix). For CSC
NP
            matrices NP = N, and for CSR matrices NP = M. This value is set automatically based
            the input for sparsetype.
data
            - pointer to a contiguous block of realtype variables (of length NNZ), containing the
            values of the nonzero entries in the matrix
sparsetype - type of the sparse matrix (CSC_MAT or CSR_MAT)
indexvals - pointer to a contiguous block of int variables (of length NNZ), containing the row indices
            (if CSC) or column indices (if CSR) of each nonzero matrix entry held in data
```

- pointer to a contiguous block of int variables (of length NP+1). For CSC matrices each entry provides the index of the first column entry into the data and indexvals arrays, e.g. if indexptr[3]=7, then the first nonzero entry in the fourth column of the matrix is located in data[7], and is located in row indexvals[7] of the matrix. The last entry contains the total number of nonzero values in the matrix and hence points one past the end of the active data in the data and indexvals arrays. For CSR matrices, each entry provides the index of the first row entry into the data and indexvals arrays.

The following pointers are added to the SlsMat type for user convenience, to provide a more intuitive interface to the CSC and CSR sparse matrix data structures. They are set automatically when creating a sparse SUNMATRIX, based on the sparse matrix storage type.

```
rowvals - pointer to indexvals when sparsetype is CSC_MAT, otherwise set to NULL.
```

colptrs - pointer to indexptrs when sparsetype is CSC\_MAT, otherwise set to NULL.

colvals - pointer to indexvals when sparsetype is CSR\_MAT, otherwise set to NULL.

rowptrs - pointer to indexptrs when sparsetype is CSR\_MAT, otherwise set to NULL. For example, the  $5 \times 4$  CSC matrix

 $\left[\begin{array}{ccccc}
0 & 3 & 1 & 0 \\
3 & 0 & 0 & 2 \\
0 & 7 & 0 & 0 \\
1 & 0 & 0 & 9 \\
0 & 0 & 0 & 5
\end{array}\right]$ 

could be stored in this structure as either

```
M = 5:
  N = 4;
  NNZ = 8;
  NP = N;
  data = \{3.0, 1.0, 3.0, 7.0, 1.0, 2.0, 9.0, 5.0\};
  sparsetype = CSC_MAT;
  indexvals = \{1, 3, 0, 2, 0, 1, 3, 4\};
  indexptrs = \{0, 2, 4, 5, 8\};
or
 M = 5;
  N = 4;
  NNZ = 10;
  NP = N;
  data = \{3.0, 1.0, 3.0, 7.0, 1.0, 2.0, 9.0, 5.0, *, *\};
  sparsetype = CSC_MAT;
  indexvals = \{1, 3, 0, 2, 0, 1, 3, 4, *, *\};
  indexptrs = \{0, 2, 4, 5, 8\};
```

where the first has no unused space, and the second has additional storage (the entries marked with \* may contain any values). Note in both cases that the final value in indexptrs is 8, indicating the total number of nonzero entries in the matrix.

Similarly, in CSR format, the same matrix could be stored as

```
M = 5;
N = 4;
NNZ = 8;
NP = N;
data = {3.0, 1.0, 3.0, 2.0, 7.0, 1.0, 9.0, 5.0};
sparsetype = CSR_MAT;
indexvals = {1, 2, 0, 3, 1, 0, 3, 3};
indexptrs = {0, 2, 4, 5, 7, 8};
```

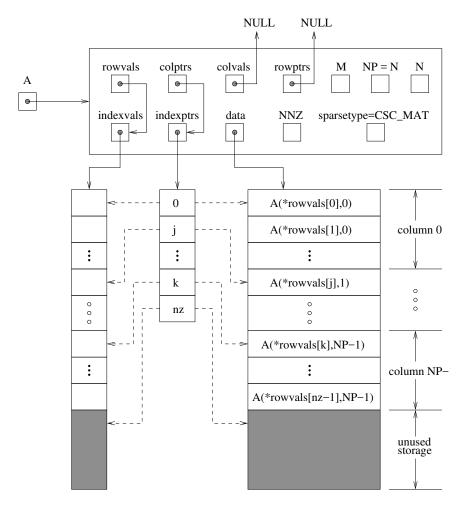


Figure 8.2: Diagram of the storage for a compressed-sparse-column matrix. Here A is an  $M \times N$  sparse matrix with storage for up to NNZ nonzero entries (the allocated length of both data and indexvals). The entries in indexvals may assume values from 0 to M-1, corresponding to the row index (zero-based) of each nonzero value. The entries in data contain the values of the nonzero entries, with the row i, column j entry of A (again, zero-based) denoted as A(i,j). The indexptrs array contains N+1 entries; the first N denote the starting index of each column within the indexvals and data arrays, while the final entry points one past the final nonzero entry. Here, although NNZ values are allocated, only nz are actually filled in; the greyed-out portions of data and indexvals indicate extra allocated space.

The header file to include when using this module is sunmatrix/sunmatrix\_sparse.h. The SUNMATRIX\_SPARSE module is accessible from all SUNDIALS solvers without linking to the libsundials\_sunmatrixsparse module library.

# 8.4.1 SUNMatrix\_Sparse accessor macros

The following macros are provided to access the content of a SUNMATRIX\_SPARSE matrix. The prefix SM\_ in the names denotes that these macros are for *SUNMatrix* implementations, and the suffix \_S denotes that these are specific to the *sparse* version.

# • SM\_CONTENT\_S

This routine gives access to the contents of the sparse SUNMatrix.

The assignment  $A_{cont} = SM_{cont} = SM_{cont} = SM_{cont}$  to be a pointer to the sparse SUNMatrix content structure.

Implementation:

```
#define SM_CONTENT_S(A) ( (SUNMatrixContent_Sparse)(A->content) )
```

• SM\_ROWS\_S, SM\_COLUMNS\_S, SM\_NNZ\_S, SM\_NP\_S, and SM\_SPARSETYPE\_S

These macros give individual access to various lengths relevant to the content of a sparse SUNMatrix.

These may be used either to retrieve or to set these values. For example, the assignment A\_rows = SM\_ROWS\_S(A) sets A\_rows to be the number of rows in the matrix A. Similarly, the assignment SM\_COLUMNS\_S(A) = A\_cols sets the number of columns in A to equal A\_cols.

Implementation:

```
#define SM_ROWS_S(A) ( SM_CONTENT_S(A)->M )
#define SM_COLUMNS_S(A) ( SM_CONTENT_S(A)->N )
#define SM_NNZ_S(A) ( SM_CONTENT_S(A)->NNZ )
#define SM_NP_S(A) ( SM_CONTENT_S(A)->NP )
#define SM_SPARSETYPE_S(A) ( SM_CONTENT_S(A)->sparsetype )
```

#### • SM\_DATA\_S, SM\_INDEXVALS\_S, and SM\_INDEXPTRS\_S

These macros give access to the data and index arrays for the matrix entries.

The assignment A\_data = SM\_DATA\_S(A) sets A\_data to be a pointer to the first component of the data array for the sparse SUNMatrix A. The assignment SM\_DATA\_S(A) = A\_data sets the data array of A to be A\_data by storing the pointer A\_data.

Similarly, the assignment A\_indexvals = SM\_INDEXVALS\_S(A) sets A\_indexvals to be a pointer to the array of index values (i.e. row indices for a CSC matrix, or column indices for a CSR matrix) for the sparse SUNMatrix A. The assignment A\_indexptrs = SM\_INDEXPTRS\_S(A) sets A\_indexptrs to be a pointer to the array of index pointers (i.e. the starting indices in the data/indexvals arrays for each row or column in CSR or CSC formats, respectively).

Implementation:

```
#define SM_DATA_S(A) ( SM_CONTENT_S(A)->data )
#define SM_INDEXVALS_S(A) ( SM_CONTENT_S(A)->indexvals )
#define SM_INDEXPTRS_S(A) ( SM_CONTENT_S(A)->indexptrs )
```

# 8.4.2 SUNMatrix\_Sparse functions

The SUNMATRIX\_SPARSE module defines sparse implementations of all matrix operations listed in Table 8.2. Their names are obtained from those in Table 8.2 by appending the suffix \_Sparse (e.g. SUNMatCopy\_Sparse). All the standard matrix operations listed in 8.2 with the suffix \_Sparse appended are callable via the FORTRAN 2003 interface by prepending an 'F' (e.g. FSUNMatCopy\_Sparse).

The module SUNMATRIX\_SPARSE provides the following additional user-callable routines:

# ${\tt SUNSparseMatrix}$

Prototype SUNMatrix SUNSparseMatrix(sunindextype M, sunindextype N, sunindextype NNZ, int sparsetype)

Description This function creates and allocates memory for a sparse SUNMatrix. Its arguments are the number of rows and columns of the matrix, M and N, the maximum number of nonzeros to be stored in the matrix, NNZ, and a flag sparsetype indicating whether to use CSR or CSC format (valid arguments are CSR\_MAT or CSC\_MAT).

F2003 Name This function is callable as FSUNSparseMatrix when using the Fortran 2003 interface module.

#### SUNSparseFromDenseMatrix

Prototype SUNMatrix SUNSparseFromDenseMatrix(SUNMatrix A, realtype droptol, int sparsetype);

Description This function creates a new sparse matrix from an existing dense matrix by copying all values with magnitude larger than droptol into the sparse matrix structure.

Requirements:

- A must have type SUNMATRIX\_DENSE;
- droptol must be non-negative;
- sparsetype must be either CSC\_MAT or CSR\_MAT.

The function returns NULL if any requirements are violated, or if the matrix storage request cannot be satisfied.

F2003 Name This function is callable as FSUNSparseFromDenseMatrix when using the Fortran 2003 interface module.

#### SUNSparseFromBandMatrix

Prototype SUNMatrix SUNSparseFromBandMatrix(SUNMatrix A, realtype droptol, int sparsetype);

Description This function creates a new sparse matrix from an existing band matrix by copying all values with magnitude larger than droptol into the sparse matrix structure.

Requirements:

- A must have type SUNMATRIX\_BAND;
- droptol must be non-negative;
- sparsetype must be either CSC\_MAT or CSR\_MAT.

The function returns NULL if any requirements are violated, or if the matrix storage request cannot be satisfied.

F2003 Name This function is callable as FSUNSparseFromBandMatrix when using the Fortran 2003 interface module.

# SUNSparseMatrix\_Realloc

Prototype int SUNSparseMatrix\_Realloc(SUNMatrix A)

Description This function reallocates internal storage arrays in a sparse matrix so that the resulting

sparse matrix has no wasted space (i.e. the space allocated for nonzero entries equals the actual number of nonzeros, indexptrs[NP]). Returns 0 on success and 1 on failure

(e.g. if the input matrix is not sparse).

F2003 Name This function is callable as FSUNSparseMatrix\_Realloc when using the Fortran 2003

interface module.

#### ${\tt SUNSparseMatrix\_Reallocate}$

Prototype int SUNSparseMatrix\_Reallocate(SUNMatrix A, sunindextype NNZ)

Description This function reallocates internal storage arrays in a sparse matrix so that the resulting

sparse matrix has storage for a specified number of nonzeros. Returns 0 on success and

1 on failure (e.g. if the input matrix is not sparse or if NNZ is negative).

F2003 Name This function is callable as FSUNSparseMatrix\_Reallocate when using the Fortran 2003

interface module.

#### SUNSparseMatrix\_Print

Prototype void SUNSparseMatrix\_Print(SUNMatrix A, FILE\* outfile)

Description This function prints the content of a sparse SUNMatrix to the output stream specified

by outfile. Note: stdout or stderr may be used as arguments for outfile to print

directly to standard output or standard error, respectively.

#### SUNSparseMatrix\_Rows

Prototype sunindextype SUNSparseMatrix\_Rows(SUNMatrix A)

Description This function returns the number of rows in the sparse SUNMatrix.

F2003 Name This function is callable as FSUNSparseMatrix\_Rows when using the Fortran 2003 inter-

face module.

#### SUNSparseMatrix\_Columns

Prototype sunindextype SUNSparseMatrix\_Columns(SUNMatrix A)

Description This function returns the number of columns in the sparse SUNMatrix.

F2003 Name This function is callable as FSUNSparseMatrix\_Columns when using the Fortran 2003

interface module.

#### SUNSparseMatrix\_NNZ

Prototype sunindextype SUNSparseMatrix\_NNZ(SUNMatrix A)

Description This function returns the number of entries allocated for nonzero storage for the sparse

matrix SUNMatrix.

F2003 Name This function is callable as FSUNSparseMatrix\_NNZ when using the Fortran 2003 inter-

face module.

# |SUNSparseMatrix\_NP

Prototype sunindextype SUNSparseMatrix\_NP(SUNMatrix A)

Description This function returns the number of columns/rows for the sparse SUNMatrix, depending

on whether the matrix uses CSC/CSR format, respectively. The indexptrs array has

NP+1 entries.

F2003 Name This function is callable as FSUNSparseMatrix\_NP when using the Fortran 2003 interface

module.

## SUNSparseMatrix\_SparseType

Prototype int SUNSparseMatrix\_SparseType(SUNMatrix A)

Description This function returns the storage type (CSR\_MAT or CSC\_MAT) for the sparse SUNMatrix.

F2003 Name This function is callable as FSUNSparseMatrix\_SparseType when using the Fortran 2003

interface module.

#### SUNSparseMatrix\_Data

Prototype realtype\* SUNSparseMatrix\_Data(SUNMatrix A)

Description This function returns a pointer to the data array for the sparse SUNMatrix.

F2003 Name This function is callable as FSUNSparseMatrix\_Data when using the Fortran 2003 inter-

face module.

# SUNSparseMatrix\_IndexValues

Prototype sunindextype\* SUNSparseMatrix\_IndexValues(SUNMatrix A)

Description This function returns a pointer to index value array for the sparse SUNMatrix: for CSR

format this is the column index for each nonzero entry, for CSC format this is the row

index for each nonzero entry.

F2003 Name This function is callable as FSUNSparseMatrix\_IndexValues when using the Fortran

2003 interface module.

# SUNSparseMatrix\_IndexPointers

Prototype sunindextype\* SUNSparseMatrix\_IndexPointers(SUNMatrix A)

Description This function returns a pointer to the index pointer array for the sparse SUNMatrix:

for CSR format this is the location of the first entry of each row in the data and indexvalues arrays, for CSC format this is the location of the first entry of each column.

F2003 Name This function is callable as FSUNSparseMatrix\_IndexPointers when using the Fortran

2003 interface module.

Within the SUNMatMatvec\_Sparse routine, internal consistency checks are performed to ensure that the matrix is called with consistent NVECTOR implementations. These are currently limited to: NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS. As additional compatible vector implementations are added to SUNDIALS, these will be included within this compatibility check.

# 8.4.3 SUNMatrix\_Sparse Fortran interfaces

The SUNMATRIX\_SPARSE module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.



#### FORTRAN 2003 interface module

The fsunmatrix\_sparse\_mod FORTRAN module defines interfaces to most SUNMATRIX\_SPARSE C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNSparseMatrix is interfaced as FSUNSparseMatrix.

The Fortran 2003 Sunmatrix\_sparse interface module can be accessed with the use statement, i.e. use fsunmatrix\_sparse\_mod, and linking to the library libsundials\_fsunmatrixsparse\_mod.lib in addition to the C library. For details on where the library and module file fsunmatrix\_sparse\_mod.mod are installed see Appendix A. We note that the module is accessible from the Fortran 2003 Sundials integrators without separately linking to the libsundials\_fsunmatrixsparse\_mod library.

#### FORTRAN 77 interface functions

For solvers that include a Fortran interface module, the SUNMATRIX\_SPARSE module also includes the Fortran-callable function FSUNSparseMatInit(code, M, N, NNZ, sparsetype, ier) to initialize this SUNMATRIX\_SPARSE module for a given SUNDIALS solver. Here code is an integer input for the solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, 4 for ARKODE); M, N and NNZ are the corresponding sparse matrix construction arguments (declared to match C type long int); sparsetype is an integer flag indicating the sparse storage type (0 for CSC, 1 for CSR); and ier is an error return flag equal to 0 for success and -1 for failure. Each of code, sparsetype and ier are declared so as to match C type int. Additionally, when using ARKODE with a non-identity mass matrix, the Fortran-callable function FSUNSparseMassMatInit(M, N, NNZ, sparsetype, ier) initializes this SUNMATRIX\_SPARSE module for storing the mass matrix.

# Chapter 9

# Description of the SUNLinearSolver module

For problems that involve the solution of linear systems of equations, the SUNDIALS packages operate using generic linear solver modules defined through the SUNLINSOL API. This allows SUNDIALS packages to utilize any valid SUNLINSOL implementation that provides a set of required functions. These functions can be divided into three categories. The first are the core linear solver functions. The second group consists of "set" routines to supply the linear solver object with functions provided by the SUNDIALS package, or for modification of solver parameters. The last group consists of "get" routines for retrieving artifacts (statistics, residual vectors, etc.) from the linear solver. All of these functions are defined in the header file sundials\_linearsolver.h.

The implementations provided with SUNDIALS work in coordination with the SUNDIALS generic NVECTOR and SUNMATRIX modules to provide a set of compatible data structures and solvers for the solution of linear systems using direct or iterative (matrix-based or matrix-free) methods. Moreover, advanced users can provide a customized SUNLinearSolver implementation to any SUNDIALS package, particularly in cases where they provide their own NVECTOR and/or SUNMATRIX modules.

Historically, the SUNDIALS packages have been designed to specifically leverage the use of either direct linear solvers or matrix-free, scaled, preconditioned, iterative linear solvers. However, matrix-based iterative linear solvers are also supported.

The iterative linear solvers packaged with SUNDIALS leverage scaling and preconditioning, as applicable, to balance error between solution components and to accelerate convergence of the linear solver. To this end, instead of solving the linear system Ax = b directly, these apply the underlying iterative algorithm to the transformed system

$$\tilde{A}\tilde{x} = \tilde{b} \tag{9.1}$$

where

$$\tilde{A} = S_1 P_1^{-1} A P_2^{-1} S_2^{-1},$$

$$\tilde{b} = S_1 P_1^{-1} b,$$

$$\tilde{x} = S_2 P_2 x,$$
(9.2)

and where

- $P_1$  is the left preconditioner,
- $P_2$  is the right preconditioner,
- $S_1$  is a diagonal matrix of scale factors for  $P_1^{-1}b$ ,
- $S_2$  is a diagonal matrix of scale factors for  $P_2x$ .

SUNDIALS packages request that iterative linear solvers stop based on the 2-norm of the scaled preconditioned residual meeting a prescribed tolerance

$$\left\| \tilde{b} - \tilde{A}\tilde{x} \right\|_2 < \text{tol.}$$

When provided an iterative SUNLINSOL implementation that does not support the scaling matrices  $S_1$  and  $S_2$ , SUNDIALS' packages will adjust the value of tol accordingly (see §9.4.2 for more details). In this case, they instead request that iterative linear solvers stop based on the criteria

$$||P_1^{-1}b - P_1^{-1}Ax||_2 < \text{tol.}$$

We note that the corresponding adjustments to tol in this case are non-optimal, in that they cannot balance error between specific entries of the solution x, only the aggregate error in the overall solution vector.

We further note that not all of the SUNDIALS-provided iterative linear solvers support the full range of the above options (e.g., separate left/right preconditioning), and that some of the SUNDIALS packages only utilize a subset of these options. Further details on these exceptions are described in the documentation for each SUNDIALS package.

For users interested in providing their own SUNLINSOL module, the following section presents the SUNLINSOL API and its implementation beginning with the definition of SUNLINSOL functions in sections 9.1.1 – 9.1.3. This is followed by the definition of functions supplied to a linear solver implementation in section 9.1.4. A table of linear solver return codes is given in section 9.1.5. The SUNLinearSolver type and the generic SUNLINSOL module are defined in section 9.1.6. The section 9.2 discusses compatibility between the SUNDIALS-provided SUNLINSOL modules and SUNMATRIX modules. Section 9.3 lists the requirements for supplying a custom SUNLINSOL module and discusses some intended use cases. Users wishing to supply their own SUNLINSOL module are encouraged to use the SUNLINSOL implementations provided with SUNDIALS as a template for supplying custom linear solver modules. The SUNLINSOL functions required by this SUNDIALS package as well as other package specific details are given in section 9.4. The remaining sections of this chapter present the SUNLINSOL modules provided with SUNDIALS.

# 9.1 The SUNLinear Solver API

The SUNLINSOL API defines several linear solver operations that enable SUNDIALS packages to utilize any SUNLINSOL implementation that provides the required functions. These functions can be divided into three categories. The first are the core linear solver functions. The second group of functions consists of set routines to supply the linear solver with functions provided by the SUNDIALS time integrators and to modify solver parameters. The final group consists of get routines for retrieving linear solver statistics. All of these functions are defined in the header file sundials/sundials\_linearsolver.h.

#### 9.1.1 SUNLinearSolver core functions

The core linear solver functions consist of four required routines to get the linear solver type (SUNLinSolGetType), initialize the linear solver object once all solver-specific options have been set (SUNLinSolInitialize), set up the linear solver object to utilize an updated matrix A (SUNLinSolSetup), and solve the linear system Ax = b (SUNLinSolSolve). The remaining routine for destruction of the linear solver object (SUNLinSolFree) is optional.

```
SUNLinSolGetType
```

Call type = SUNLinSolGetType(LS);

Description The required function SUNLinSolGetType returns the type identifier for the linear solver LS. It is used to determine the solver type (direct, iterative, or matrix-iterative) from the abstract SUNLinearSolver interface.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

Return value The return value type (of type int) will be one of the following:

- SUNLINEARSOLVER\_DIRECT 0, the SUNLINSOL module requires a matrix, and computes an 'exact' solution to the linear system defined by that matrix.
- SUNLINEARSOLVER\_ITERATIVE 1, the SUNLINSOL module does not require a matrix (though one may be provided), and computes an inexact solution to the linear system using a matrix-free iterative algorithm. That is it solves the linear system defined by the package-supplied ATimes routine (see SUNLinSolSetATimes below), even if that linear system differs from the one encoded in the matrix object (if one is provided). As the solver computes the solution only inexactly (or may diverge), the linear solver should check for solution convergence/accuracy as appropriate.
- SUNLINEARSOLVER\_MATRIX\_ITERATIVE 2, the SUNLINSOL module requires a matrix, and computes an inexact solution to the linear system defined by that matrix using an iterative algorithm. That is it solves the linear system defined by the matrix object even if that linear system differs from that encoded by the package-supplied ATimes routine. As the solver computes the solution only inexactly (or may diverge), the linear solver should check for solution convergence/accuracy as appropriate.

Notes See section 9.3.1 for more information on intended use cases corresponding to the linear solver type.

#### SUNLinSolInitialize

Call retval = SUNLinSolInitialize(LS);

Description The required function SUNLinSolInitialize performs linear solver initialization (assuming that all solver-specific options have been set).

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

Return value This should return zero for a successful call, and a negative value for a failure, ideally returning one of the generic error codes listed in Table 9.1.

#### SUNLinSolSetup

Call retval = SUNLinSolSetup(LS, A);

Description The required function SUNLinSolSetup performs any linear solver setup needed, based on an updated system SUNMATRIX A. This may be called frequently (e.g., with a full Newton method) or infrequently (for a modified Newton method), based on the type of integrator and/or nonlinear solver requesting the solves.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

A (SUNMatrix) a SUNMATRIX object.

Return value This should return zero for a successful call, a positive value for a recoverable failure and a negative value for an unrecoverable failure, ideally returning one of the generic error codes listed in Table 9.1.

# SUNLinSolSolve

Call retval = SUNLinSolSolve(LS, A, x, b, tol);

Description The required function SUNLinSolSolve solves a linear system Ax = b.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

A (SUNMatrix) a SUNMATRIX object.

- x (N\_Vector) a NVECTOR object containing the initial guess for the solution of the linear system, and the solution to the linear system upon return.
- b (N\_Vector) a NVECTOR object containing the linear system right-hand side.

tol (realtype) the desired linear solver tolerance.

Return value This should return zero for a successful call, a positive value for a recoverable failure and a negative value for an unrecoverable failure, ideally returning one of the generic error codes listed in Table 9.1.

Notes **Direct solvers:** can ignore the tol argument.

Matrix-free solvers: (those that identify as SUNLINEARSOLVER\_ITERATIVE) can ignore the SUNMATRIX input A, and should instead rely on the matrix-vector product function supplied through the routine SUNLinSolSetATimes.

Iterative solvers: (those that identify as SUNLINEARSOLVER\_ITERATIVE or SUNLINEARSOLVER\_MATRIX\_ITERATIVE) should attempt to solve to the specified tolerance tol in a weighted 2-norm. If the solver does not support scaling then it should just use a 2-norm.

#### SUNLinSolFree

Call retval = SUNLinSolFree(LS);

Description The optional function SUNLinSolFree frees memory allocated by the linear solver.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

Return value This should return zero for a successful call and a negative value for a failure.

#### 9.1.2 SUNLinearSolver set functions

The following set functions are used to supply linear solver modules with functions defined by the SUNDIALS packages and to modify solver parameters. Only the routine for setting the matrix-vector product routine is required, and that is only for matrix-free linear solver modules. Otherwise, all other set functions are optional. SUNLINSOL implementations that do not provide the functionality for any optional routine should leave the corresponding function pointer NULL instead of supplying a dummy routine.

# SUNLinSolSetATimes

Call retval = SUNLinSolSetATimes(LS, A\_data, ATimes);

Description The function SUNLinSolSetATimes is required for matrix-free linear solvers; otherwise it is optional.

This routine provides an ATimesFn function pointer, as well as a void\* pointer to a data structure used by this routine, to a linear solver object. SUNDIALS packages will call this function to set the matrix-vector product function to either a solver-provided difference-quotient via vector operations or a user-supplied solver-specific routine.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

A\_data (void\*) data structure passed to ATimes.

ATimes (ATimesFn) function pointer implementing the matrix-vector product routine.

Return value This routine should return zero for a successful call, and a negative value for a failure, ideally returning one of the generic error codes listed in Table 9.1.

### SUNLinSolSetPreconditioner

Call retval = SUNLinSolSetPreconditioner(LS, Pdata, Pset, Psol);

 $\label{thm:condition} \textbf{Description} \quad \textbf{The } \textit{optional } \textbf{function } \textbf{SUNLinSolSetPreconditioner } \textbf{provides } \textbf{PSetupFn } \textbf{and } \textbf{PSolveFn} \\ \textbf{SolveFn} \quad \textbf{Suppose } \textbf{Suppo$ 

function pointers that implement the preconditioner solves  $P_1^{-1}$  and  $P_2^{-1}$  from equations (9.1)-(9.2). This routine will be called by a SUNDIALS package, which will provide translation between the generic Pset and Psol calls and the package- or user-supplied

routines.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

Pdata (void\*) data structure passed to both Pset and Psol.

Pset (PSetupFn) function pointer implementing the preconditioner setup.

Psol (PSolveFn) function pointer implementing the preconditioner solve.

Return value This routine should return zero for a successful call, and a negative value for a failure,

ideally returning one of the generic error codes listed in Table 9.1.

### SUNLinSolSetScalingVectors

Call retval = SUNLinSolSetScalingVectors(LS, s1, s2);

Description The optional function SUNLinSolSetScalingVectors provides left/right scaling vectors

for the linear system solve. Here, s1 and s2 are NVECTOR of positive scale factors containing the diagonal of the matrices  $S_1$  and  $S_2$  from equations (9.1)-(9.2), respectively. Neither of these vectors need to be tested for positivity, and a NULL argument for either

indicates that the corresponding scaling matrix is the identity.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

s1 (N\_Vector) diagonal of the matrix  $S_1$ 

s2 (N\_Vector) diagonal of the matrix  $S_2$ 

Return value This routine should return zero for a successful call, and a negative value for a failure,

ideally returning one of the generic error codes listed in Table 9.1.

### 9.1.3 SUNLinearSolver get functions

The following get functions allow SUNDIALS packages to retrieve results from a linear solve. All routines are optional.

### SUNLinSolNumIters

Call its = SUNLinSolNumIters(LS);

Description The optional function SUNLinSolNumIters should return the number of linear iterations

performed in the last 'solve' call.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

Return value int containing the number of iterations

### SUNLinSolResNorm

Call rnorm = SUNLinSolResNorm(LS);

Description The optional function SUNLinSolResNorm should return the final residual norm from

the last 'solve' call.

Arguments LS (SUNLinearSolver) a SUNLINSOL object.

Return value realtype containing the final residual norm

### SUNLinSolResid

Call rvec = SUNLinSolResid(LS);

If an iterative method computes the preconditioned initial residual and returns with Description

> a successful solve without performing any iterations (i.e., either the initial guess or the preconditioner is sufficiently accurate), then this optional routine may be called by the SUNDIALS package. This routine should return the NVECTOR containing the

preconditioned initial residual vector.

LS (SUNLinearSolver) a SUNLINSOL object. Arguments

Return value N\_Vector containing the final residual vector

Since N\_Vector is actually a pointer, and the results are not modified, this routine Notes

> should not require additional memory allocation. If the SUNLINSOL object does not retain a vector for this purpose, then this function pointer should be set to NULL in the

implementation.

### SUNLinSolLastFlag

lflag = SUNLinSolLastFlag(LS); Call

Description The optional function SUNLinSollastFlag should return the last error flag encountered

within the linear solver. This is not called by the SUNDIALS packages directly; it allows

the user to investigate linear solver issues after a failed solve.

LS (SUNLinearSolver) a SUNLINSOL object. Arguments

Return value long int containing the most recent error flag

### ${\tt SUNLinSolSpace}$

Call retval = SUNLinSolSpace(LS, &lrw, &liw);

Description The optional function SUNLinSolSpace should return the storage requirements for the

linear solver LS.

LS (SUNLinearSolver) a SUNLINSOL object. Arguments

1rw (long int\*) the number of realtype words stored by the linear solver.

liw (long int\*) the number of integer words stored by the linear solver.

Return value This should return zero for a successful call, and a negative value for a failure, ideally

returning one of the generic error codes listed in Table 9.1.

Notes This function is advisory only, for use in determining a user's total space requirements.

### 9.1.4 Functions provided by SUNDIALS packages

To interface with the SUNLINSOL modules, the SUNDIALS packages supply a variety of routines for evaluating the matrix-vector product, and setting up and applying the preconditioner. These packageprovided routines translate between the user-supplied ODE, DAE, or nonlinear systems and the generic interfaces to the linear systems of equations that result in their solution. The types for functions provided to a SUNLINSOL module are defined in the header file sundials/sundials\_iterative.h, and are described below.

### ATimesFn

Definition typedef int (\*ATimesFn)(void \*A\_data, N\_Vector v, N\_Vector z);

Purpose These functions compute the action of a matrix on a vector, performing the operation

z = Av. Memory for z should already be allocted prior to calling this function. The

vector v should be left unchanged.

Arguments A\_data is a pointer to client data, the same as that supplied to SUNLinSolSetATimes.

v is the input vector to multiply.

z is the output vector computed.

Return value This routine should return 0 if successful and a non-zero value if unsuccessful.

PSetupFn

Definition typedef int (\*PSetupFn) (void \*P\_data)

Purpose These functions set up any requisite problem data in preparation for calls to the corre-

sponding PSolveFn.

Arguments P\_data is a pointer to client data, the same pointer as that supplied to the routine SUNLinSolSetPreconditioner.

Return value This routine should return 0 if successful and a non-zero value if unsuccessful.

PSolveFn

Definition typedef int (\*PSolveFn)(void \*P\_data, N\_Vector r, N\_Vector z,

Purpose These functions solve the preconditioner equation Pz = r for the vector z. Memory for z should already be allocted prior to calling this function. The parameter  $P_{-}$ data is a point of the parameter  $p_{-}$ data is a parameter  $p_{-}$ data is a point of the parameter  $p_$ 

z should already be allocted prior to calling this function. The parameter  $P_{-}$ data is a pointer to any information about P which the function needs in order to do its job (set up by the corresponding PSetupFn). The parameter lr is input, and indicates whether P is to be taken as the left preconditioner or the right preconditioner: lr = 1 for left and lr = 2 for right. If preconditioning is on one side only, lr can be ignored. If the preconditioner is iterative, then it should strive to solve the preconditioner equation so that

$$||Pz - r||_{\text{wrms}} < tol$$

where the weight vector for the WRMS norm may be accessed from the main package memory structure. The vector **r** should not be modified by the PSolveFn.

Arguments P\_data is a pointer to client data, the same pointer as that supplied to the routine SUNLinSolSetPreconditioner.

r is the right-hand side vector for the preconditioner system.

**z** is the solution vector for the preconditioner system.

tol is the desired tolerance for an iterative preconditioner.

is flag indicating whether the routine should perform left (1) or right (2) preconditioning.

Return value This routine should return 0 if successful and a non-zero value if unsuccessful. On a failure, a negative return value indicates an unrecoverable condition, while a positive value indicates a recoverable one, in which the calling routine may reattempt the solution after updating preconditioner data.

### 9.1.5 SUNLinearSolver return codes

The functions provided to SUNLINSOL modules by each SUNDIALS package, and functions within the SUNDIALS-provided SUNLINSOL implementations utilize a common set of return codes, shown in Table 9.1. These adhere to a common pattern: 0 indicates success, a postitive value corresponds to a recoverable failure, and a negative value indicates a non-recoverable failure. Aside from this pattern, the actual values of each error code are primarily to provide additional information to the user in case of a linear solver failure.

Table 9.1: Description of the SUNLinearSolver error codes				
Name	Value	Description		
SUNLS_SUCCESS	0	successful call or converged solve		
SUNLS_MEM_NULL	-1	the memory argument to the function is NULL		
SUNLS_ILL_INPUT	-2	an illegal input has been provided to the function		
SUNLS_MEM_FAIL	-3	failed memory access or allocation		
SUNLS_ATIMES_FAIL_UNREC	-4	an unrecoverable failure occurred in the ATimes routine		
SUNLS_PSET_FAIL_UNREC	-5	an unrecoverable failure occurred in the Pset routine		
SUNLS_PSOLVE_FAIL_UNREC	-6	an unrecoverable failure occurred in the Psolve routine		
SUNLS_PACKAGE_FAIL_UNREC	-7	an unrecoverable failure occurred in an external linear solver package		
SUNLS_GS_FAIL	-8	a failure occurred during Gram-Schmidt orthogonalization (SUNLINSOL_SPGMR/SUNLINSOL_SPFGMR)		
SUNLS_QRSOL_FAIL	-9	a singular $R$ matrix was encountered in a QR factorization (SUNLINSOL_SPGMR/SUNLINSOL_SPFGMR)		
SUNLS_RES_REDUCED	1	an iterative solver reduced the residual, but did not converge to the desired tolerance		
SUNLS_CONV_FAIL	2	an iterative solver did not converge (and the residual was not reduced)		
SUNLS_ATIMES_FAIL_REC	3	a recoverable failure occurred in the ATimes routine		
SUNLS_PSET_FAIL_REC	4	a recoverable failure occurred in the Pset routine		
SUNLS_PSOLVE_FAIL_REC	5	a recoverable failure occurred in the Psolve routine		
SUNLS_PACKAGE_FAIL_REC	6	a recoverable failure occurred in an external linear solver package		
SUNLS_QRFACT_FAIL	7	a singular matrix was encountered during a QR factorization (SUNLINSOL_SPGMR/SUNLINSOL_SPFGMR)		
SUNLS_LUFACT_FAIL	8	a singular matrix was encountered during a LU factorization		

### 9.1.6The generic SUNLinearSolver module

SUNDIALS packages interact with specific SUNLINSOL implementations through the generic SUNLINSOL module on which all other SUNLINSOL iplementations are built. The SUNLinearSolver type is a pointer to a structure containing an implementation-dependent content field, and an ops field. The type SUNLinearSolver is defined as

(SUNLINSOL\_DENSE/SUNLINSOL\_BAND)

```
typedef struct _generic_SUNLinearSolver *SUNLinearSolver;
struct _generic_SUNLinearSolver {
  void *content;
  struct _generic_SUNLinearSolver_Ops *ops;
};
```

where the \_generic\_SUNLinearSolver\_Ops structure is a list of pointers to the various actual linear solver operations provided by a specific implementation. The \_generic\_SUNLinearSolver\_Ops structure is defined as

```
struct _generic_SUNLinearSolver_Ops {
  SUNLinearSolver_Type (*gettype)(SUNLinearSolver);
```

```
(*setatimes)(SUNLinearSolver, void*, ATimesFn);
  int
  int
                        (*setpreconditioner)(SUNLinearSolver, void*,
                                              PSetupFn, PSolveFn);
  int
                        (*setscalingvectors)(SUNLinearSolver,
                                              N_Vector, N_Vector);
  int
                        (*initialize)(SUNLinearSolver);
  int
                        (*setup)(SUNLinearSolver, SUNMatrix);
                        (*solve)(SUNLinearSolver, SUNMatrix, N_Vector,
  int
                                 N_Vector, realtype);
  int
                        (*numiters)(SUNLinearSolver);
                        (*resnorm)(SUNLinearSolver);
  realtype
                        (*lastflag)(SUNLinearSolver);
  long int
                        (*space)(SUNLinearSolver, long int*, long int*);
  int
  N_Vector
                        (*resid)(SUNLinearSolver);
                        (*free)(SUNLinearSolver);
  int.
};
```

The generic SUNLINSOL module defines and implements the linear solver operations defined in Sections 9.1.1-9.1.3. These routines are in fact only wrappers to the linear solver operations defined by a particular SUNLINSOL implementation, which are accessed through the *ops* field of the SUNLinearSolver structure. To illustrate this point we show below the implementation of a typical linear solver operation from the generic SUNLINSOL module, namely SUNLinSolInitialize, which initializes a SUNLINSOL object for use after it has been created and configured, and returns a flag denoting a successful/failed operation:

```
int SUNLinSolInitialize(SUNLinearSolver S)
{
  return ((int) S->ops->initialize(S));
}
```

# 9.2 Compatibility of SUNLinearSolver modules

We note that not all SUNLINSOL types are compatible with all SUNMATRIX and NVECTOR types provided with SUNDIALS. In Table 9.2 we show the matrix-based linear solvers available as SUNLINSOL modules, and the compatible matrix implementations. Recall that Table 4.1 shows the compatibility between all SUNLINSOL modules and vector implementations.

Table 9.2:	SUNDIALS	matrix-based	linear se	olvers a	and m	atrix	implem	entations	that	can	be ı	ısed	for
each.													

Linear Solver	Dense	Banded	Sparse	User
Interface	Matrix	Matrix	Matrix	Supplied
Dense	✓			✓
Band		<b>√</b>		✓
LapackDense	✓			✓
LapackBand		✓		✓
KLU			✓	✓
SUPERLUMT			<b>√</b>	✓
User supplied	<b>√</b>	<b>√</b>	<b>√</b>	✓

# 9.3 Implementing a custom SUNLinearSolver module

A particular implementation of the Sunlinsol module must:

- Specify the *content* field of the SUNLinearSolver object.
- Define and implement a minimal subset of the linear solver operations. See the section 9.4 to determine which SUNLINSOL operations are required for this SUNDIALS package.
  - Note that the names of these routines should be unique to that implementation in order to permit using more than one SUNLINSOL module (each with different SUNLinearSolver internal data representations) in the same code.
- Define and implement user-callable constructor and destructor routines to create and free a SUNLinearSolver with the new content field and with ops pointing to the new linear solver operations.

We note that the function pointers for all unsupported optional routines should be set to NULL in the *ops* structure. This allows the SUNDIALS package that is using the SUNLINSOL object to know that the associated functionality is not supported.

Additionally, a SUNLINSOL implementation may do the following:

- Define and implement additional user-callable "set" routines acting on the SUNLinearSolver, e.g., for setting various configuration options to tune the linear solver to a particular problem.
- Provide additional user-callable "get" routines acting on the SUNLinearSolver object, e.g., for returning various solve statistics.

### 9.3.1 Intended use cases

The SUNLINSOL (and SUNMATRIX) APIs are designed to require a minimal set of routines to ease interfacing with custom or third-party linear solver libraries. External solvers provide similar routines with the necessary functionality and thus will require minimal effort to wrap within custom SUNMATRIX and SUNLINSOL implementations. Sections 8.1 and 9.4 include a list of the required set of routines that compatible SUNMATRIX and SUNLINSOL implementations must provide. As SUNDIALS packages utilize generic SUNLINSOL modules allowing for user-supplied SUNLinearSolver implementations, there exists a wide range of possible linear solver combinations. Some intended use cases for both the SUNDIALS-provided and user-supplied SUNLINSOL modules are discussed in the following sections.

### Direct linear solvers

Direct linear solver modules require a matrix and compute an 'exact' solution to the linear system defined by the matrix. Multiple matrix formats and associated direct linear solvers are supplied with SUNDIALS through different SUNMATRIX and SUNLINSOL implementations. SUNDIALS packages strive to amortize the high cost of matrix construction by reusing matrix information for multiple nonlinear iterations. As a result, each package's linear solver interface recomputes Jacobian information as infrequently as possible.

Alternative matrix storage formats and compatible linear solvers that are not currently provided by, or interfaced with, SUNDIALS can leverage this infrastructure with minimal effort. To do so, a user must implement custom SUNMATRIX and SUNLINSOL wrappers for the desired matrix format and/or linear solver following the APIs described in Chapters 8 and 9. This user-supplied SUNLINSOL module must then self-identify as having SUNLINEARSOLVER\_DIRECT type.

### Matrix-free iterative linear solvers

Matrix-free iterative linear solver modules do not require a matrix and compute an inexact solution to the linear system defined by the package-supplied ATimes routine. SUNDIALS supplies multiple scaled, preconditioned iterative linear solver (spils) SUNLINSOL modules that support scaling to allow users to handle non-dimensionalization (as best as possible) within each SUNDIALS package and retain variables and define equations as desired in their applications. For linear solvers that do not support left/right scaling, the tolerance supplied to the linear solver is adjusted to compensate (see section 9.4.2 for

more details); however, this use case may be non-optimal and cannot handle situations where the magnitudes of different solution components or equations vary dramatically within a single problem.

To utilize alternative linear solvers that are not currently provided by, or interfaced with, SUNDIALS a user must implement a custom SUNLINSOL wrapper for the linear solver following the API described in Chapter 9. This user-supplied SUNLINSOL module must then self-identify as having SUNLINEARSOLVER\_ITERATIVE type.

### Matrix-based iterative linear solvers (reusing A)

Matrix-based iterative linear solver modules require a matrix and compute an inexact solution to the linear system *defined by the matrix*. This matrix will be updated infrequently and resued across multiple solves to amortize cost of matrix construction. As in the direct linear solver case, only wrappers for the matrix and linear solver in Sunmatrix and sunlinsol implementations need to be created to utilize a new linear solver. *This user-supplied* Sunlinsol *module must then self-identify as having* Sunlinear Solver Matrix\_Iterative type.

At present, SUNDIALS has one example problem that uses this approach for wrapping a structured-grid matrix, linear solver, and preconditioner from the *hypre* library that may be used as a template for other customized implementations (see examples/arkode/CXX\_parhyp/ark\_heat2D\_hypre.cpp).

### Matrix-based iterative linear solvers (current A)

For users who wish to utilize a matrix-based iterative linear solver module where the matrix is *purely* for preconditioning and the linear system is defined by the package-supplied ATimes routine, we envision two current possibilities.

The preferred approach is for users to employ one of the SUNDIALS spils SUNLINSOL implementations (SUNLINSOL\_SPGMR, SUNLINSOL\_SPFGMR, SUNLINSOL\_SPFGMR, SUNLINSOL\_SPTFQMR, or SUNLINSOL\_PCG) as the outer solver. The creation and storage of the preconditioner matrix, and interfacing with the corresponding linear solver, can be handled through a package's preconditioner 'setup' and 'solve' functionality (see §4.5.8.2) without creating SUNMATRIX and SUNLINSOL implementations. This usage mode is recommended primarily because the SUNDIALS-provided spils modules support the scaling as described above.

A second approach supported by the linear solver APIs is as follows. If the SUNLINSOL implementation is matrix-based, self-identifies as having SUNLINEARSOLVER\_ITERATIVE type, and also provides a non-NULL SUNLinSolSetATimes routine, then each SUNDIALS package will call that routine to attach its package-specific matrix-vector product routine to the SUNLINSOL object. The SUNDIALS package will then call the SUNLINSOL-provided SUNLinSolSetup routine (infrequently) to update matrix information, but will provide current matrix-vector products to the SUNLINSOL implementation through the package-supplied ATimesFn routine.

### 9.4 IDAS SUNLinearSolver interface

Table 9.3 below lists the SUNLINSOL module linear solver functions used within the IDALS interface. As with the SUNMATRIX module, we emphasize that the IDA user does not need to know detailed usage of linear solver functions by the IDA code modules in order to use IDA. The information is presented as an implementation detail for the interested reader.

The linear solver functions listed below are marked with  $\checkmark$  to indicate that they are required, or with  $\dagger$  to indicate that they are only called if they are non-NULL in the SUNLINSOL implementation that is being used. Note:

- 1. Although IDALS does not call SUNLinSollastFlag directly, this routine is available for users to query linear solver issues directly.
- 2. Although IDALS does not call SUNLinSolFree directly, this routine should be available for users to call when cleaning up from a simulation.

	DIRECT	ITERATIVE	MATRIX_ITERATIVE
SUNLinSolGetType	<b>√</b>	<b>√</b>	<b>√</b>
SUNLinSolSetATimes	†	<b>√</b>	†
SUNLinSolSetPreconditioner	†	†	†
SUNLinSolSetScalingVectors	†	†	†
SUNLinSolInitialize	<b>√</b>	<b>√</b>	<b>√</b>
SUNLinSolSetup	<b>√</b>	<b>√</b>	<b>√</b>
SUNLinSolSolve	<b>√</b>	<b>√</b>	<b>✓</b>
SUNLinSolNumIters		<b>√</b>	<b>√</b>
SUNLinSolResid		<b>√</b>	<b>√</b>
$^1$ SUNLinSolLastFlag			
$^2$ SUNLinSolFree			
SUNLinSolSpace	†	+	+

Table 9.3: List of linear solver function usage in the IDALS interface

Since there are a wide range of potential SUNLINSOL use cases, the following subsections describe some details of the IDALS interface, in the case that interested users wish to develop custom SUNLINSOL modules.

### 9.4.1 Lagged matrix information

If the SUNLINSOL object self-identifies as having type SUNLINEARSOLVER\_DIRECT or SUNLINEARSOLVER\_MATRIX\_ITERATIVE, then the SUNLINSOL object solves a linear system defined by a SUNMATRIX object. CVLS will update the matrix information infrequently according to the strategies outlined in §2.1. When solving a linear system  $J\bar{x}=b$ , it is likely that the value  $\bar{\alpha}$  used to construct J differs from the current value of  $\alpha$  in the BDF method, since J is updated infrequently. Therefore, after calling the SUNLINSOL-provided SUNLinSolve routine, we test whether  $\alpha/\bar{\alpha}\neq 1$ , and if this is the case we scale the solution  $\bar{x}$  to obtain the desired linear system solution x via

$$x = \frac{2}{1 + \alpha/\bar{\alpha}}\bar{x}.\tag{9.3}$$

For values of  $\alpha/\bar{\alpha}$  that are "close" to 1, this rescaling approximately solves the original linear system.

### 9.4.2 Iterative linear solver tolerance

If the SUNLINSOL object self-identifies as having type SUNLINEARSOLVER\_ITERATIVE or SUNLINEARSOLVER\_MATRIX\_ITERATIVE then IDALS will set the input tolerance delta as described in §2.1. However, if the iterative linear solver does not support scaling matrices (i.e., the SUNLinSolSetScalingVectors routine is NULL), then IDALS will attempt to adjust the linear solver tolerance to account for this lack of functionality. To this end, the following assumptions are made:

1. All solution components have similar magnitude; hence the error weight vector W used in the WRMS norm (see §2.1) should satisfy the assumption

$$W_i \approx W_{mean}$$
, for  $i = 0, \dots, n-1$ .

2. The SUNLINSOL object uses a standard 2-norm to measure convergence.

Since IDA uses identical left and right scaling matrices,  $S_1 = S_2 = S = \text{diag}(W)$ , then the linear solver convergence requirement is converted as follows (using the notation from equations (9.1)-(9.2)):

$$\begin{split} & \left\| \tilde{b} - \tilde{A} \tilde{x} \right\|_2 < \text{tol} \\ \Leftrightarrow & \left\| S P_1^{-1} b - S P_1^{-1} A x \right\|_2 < \text{tol} \\ \Leftrightarrow & \sum_{i=0}^{n-1} \left[ W_i \left( P_1^{-1} (b - A x) \right)_i \right]^2 < \text{tol}^2 \\ \Leftrightarrow & W_{mean}^2 \sum_{i=0}^{n-1} \left[ \left( P_1^{-1} (b - A x) \right)_i \right]^2 < \text{tol}^2 \\ \Leftrightarrow & \sum_{i=0}^{n-1} \left[ \left( P_1^{-1} (b - A x) \right)_i \right]^2 < \left( \frac{\text{tol}}{W_{mean}} \right)^2 \\ \Leftrightarrow & \left\| P_1^{-1} (b - A x) \right\|_2 < \frac{\text{tol}}{W_{mean}} \end{split}$$

Therefore the tolerance scaling factor

$$W_{mean} = ||W||_2/\sqrt{n}$$

is computed and the scaled tolerance  $delta = tol/W_{mean}$  is supplied to the SUNLINSOL object.

## 9.5 The SUNLinearSolver\_Dense implementation

This section describes the SUNLINSOL implementation for solving dense linear systems. The SUNLINSOL\_DENSE module is designed to be used with the corresponding SUNMATRIX\_DENSE matrix type, and one of the serial or shared-memory NVECTOR implementations (NVECTOR\_SERIAL, NVECTOR\_OPENMP, or NVECTOR\_PTHREADS).

To access the SUNLINSOL\_DENSE module, include the header file sunlinsol/sunlinsol\_dense.h. We note that the SUNLINSOL\_DENSE module is accessible from SUNDIALS packages without separately linking to the libsundials\_sunlinsoldense module library.

### 9.5.1 SUNLinearSolver\_Dense description

This solver is constructed to perform the following operations:

- The "setup" call performs a LU factorization with partial (row) pivoting  $(\mathcal{O}(N^3) \cos t)$ , PA = LU, where P is a permutation matrix, L is a lower triangular matrix with 1's on the diagonal, and U is an upper triangular matrix. This factorization is stored in-place on the input SUNMATRIX\_DENSE object A, with pivoting information encoding P stored in the pivots array.
- The "solve" call performs pivoting and forward and backward substitution using the stored pivots array and the LU factors held in the SUNMATRIX\_DENSE object ( $\mathcal{O}(N^2)$  cost).

### 9.5.2 SUNLinearSolver\_Dense functions

The SUNLINSOL\_DENSE module provides the following user-callable constructor for creating a SUNLinearSolver object.

### SUNLinSol\_Dense

Call LS = SUNLinSol\_Dense(y, A);

Description The function SUNLinSol\_Dense creates and allocates memory for a dense

SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

A (SUNMatrix) a  ${\tt SUNMATRIX\_DENSE}$  matrix template for cloning matrices needed

within the solver

Return value This returns a SUNLinearSolver object. If either A or y are incompatible then this

routine will return NULL.

Notes This routine will perform consistency checks to ensure that it is called with con-

sistent NVECTOR and SUNMATRIX implementations. These are currently limited to the SUNMATRIX\_DENSE matrix type and the NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS vector types. As additional compatible matrix and vector implementations are added to SUNDIALS, these will be included within this

compatibility check.

Deprecated Name For backward compatibility, the wrapper function SUNDenseLinearSolver with

idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_Dense when using the Fortran 2003 inter-

face module.

The SUNLINSOL\_DENSE module defines implementations of all "direct" linear solver operations listed in Sections 9.1.1 - 9.1.3:

SUNLinSolGetType\_Dense

- SUNLinSolInitialize\_Dense this does nothing, since all consistency checks are performed at solver creation.
- SUNLinSolSetup\_Dense this performs the LU factorization.
- SUNLinSolSolve\_Dense this uses the LU factors and pivots array to perform the solve.
- SUNLinSolLastFlag\_Dense
- SUNLinSolSpace\_Dense this only returns information for the storage within the solver object, i.e. storage for N, last\_flag, and pivots.
- SUNLinSolFree\_Dense

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

## 9.5.3 SUNLinearSolver\_Dense Fortran interfaces

The SUNLINSOL\_DENSE module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

### FORTRAN 2003 interface module

The fsunlinsol\_dense\_mod FORTRAN module defines interfaces to all SUNLINSOL\_DENSE C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_Dense is interfaced as FSUNLinSol\_Dense.

The FORTRAN 2003 SUNLINSOL\_DENSE interface module can be accessed with the use statement, i.e. use fsunlinsol\_dense\_mod, and linking to the library libsundials\_fsunlinsoldense\_mod. lib in

addition to the C library. For details on where the library and module file fsunlinsol\_dense\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunlinsoldense\_mod library.

### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_DENSE module also includes a Fortran-callable function for creating a SUNLinearSolver object.

### FSUNDENSELINSOLINIT

Call FSUNDENSELINSOLINIT(code, ier)

Description The function FSUNDENSELINSOLINIT can be called for Fortran programs to create a

dense SUNLinearSolver object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* both the NVECTOR and SUNMATRIX objects have been

initialized.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_DENSE module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

### FSUNMASSDENSELINSOLINIT

Call FSUNMASSDENSELINSOLINIT(ier)

Description The function FSUNMASSDENSELINSOLINIT can be called for Fortran programs to create

a dense SUNLinearSolver object for mass matrix linear systems.

Arguments None

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called after both the NVECTOR and SUNMATRIX mass-matrix

objects have been initialized.

### 9.5.4 SUNLinearSolver\_Dense content

The  $SUNLINSOL_DENSE$  module defines the content field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_Dense {
   sunindextype N;
   sunindextype *pivots;
   long int last_flag;
};
```

These entries of the *content* field contain the following information:

N - size of the linear system,

pivots - index array for partial pivoting in LU factorization,

last\_flag - last error return flag from internal function evaluations.

## 9.6 The SUNLinearSolver\_Band implementation

This section describes the SUNLINSOL implementation for solving banded linear systems. The SUNLINSOL\_BAND module is designed to be used with the corresponding SUNMATRIX\_BAND matrix type, and one of the serial or shared-memory NVECTOR implementations (NVECTOR\_SERIAL, NVECTOR\_OPENMP, or NVECTOR\_PTHREADS).

To access the SUNLINSOL\_BAND module, include the header file sunlinsol/sunlinsol\_band.h. We note that the SUNLINSOL\_BAND module is accessible from SUNDIALS packages without separately linking to the libsundials\_sunlinsolband module library.

### 9.6.1 SUNLinearSolver\_Band description

This solver is constructed to perform the following operations:

- The "setup" call performs a LU factorization with partial (row) pivoting, PA = LU, where P is a permutation matrix, L is a lower triangular matrix with 1's on the diagonal, and U is an upper triangular matrix. This factorization is stored in-place on the input SUNMATRIX\_BAND object A, with pivoting information encoding P stored in the pivots array.
- The "solve" call performs pivoting and forward and backward substitution using the stored pivots array and the *LU* factors held in the SUNMATRIX\_BAND object.
- A must be allocated to accommodate the increase in upper bandwidth that occurs during factorization. More precisely, if A is a band matrix with upper bandwidth  $\mathtt{mu}$  and lower bandwidth  $\mathtt{ml}$ , then the upper triangular factor U can have upper bandwidth as big as  $\mathtt{smu} = \mathtt{MIN(N-1,mu+ml)}$ . The lower triangular factor L has lower bandwidth  $\mathtt{ml}$ .

### 9.6.2 SUNLinearSolver\_Band functions

The SUNLINSOL\_BAND module provides the following user-callable constructor for creating a SUNLinearSolver object.

SUNLi	nSol	_Band
-------	------	-------

Call LS = SUNLinSol\_Band(y, A);

Description The function SUNLinSol\_Band creates and allocates memory for a band

SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

A (SUNMatrix) a  $\operatorname{SUNMATRIX\_BAND}$  matrix template for cloning matrices needed

within the solver

Return value This returns a SUNLinearSolver object. If either A or y are incompatible then this

routine will return NULL.

Notes This routine will perform consistency checks to ensure that it is called with con-

sistent NVECTOR and SUNMATRIX implementations. These are currently limited to the SUNMATRIX\_BAND matrix type and the NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS vector types. As additional compatible matrix and vector implementations are added to SUNDIALS, these will be included within this

compatibility check.

Additionally, this routine will verify that the input matrix A is allocated with appropriate upper bandwidth storage for the LU factorization.

Deprecated Name For backward compatibility, the wrapper function SUNBandLinearSolver with idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_Band when using the Fortran 2003 interface

module.



The SUNLINSOL\_BAND module defines band implementations of all "direct" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_Band
- SUNLinSolInitialize\_Band this does nothing, since all consistency checks are performed at solver creation.
- SUNLinSolSetup\_Band this performs the LU factorization.
- SUNLinSolSolve\_Band this uses the LU factors and pivots array to perform the solve.
- SUNLinSolLastFlag\_Band
- SUNLinSolSpace\_Band this only returns information for the storage within the solver object, i.e. storage for N, last\_flag, and pivots.
- SUNLinSolFree\_Band

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

### 9.6.3 SUNLinearSolver Band Fortran interfaces

The SUNLINSOL\_BAND module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

### FORTRAN 2003 interface module

The fsunlinsol\_band\_mod FORTRAN module defines interfaces to all SUNLINSOL\_BAND C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_Band is interfaced as FSUNLinSol\_Band.

The Fortran 2003 sunlinsol\_band interface module can be accessed with the use statement, i.e. use fsunlinsol\_band\_mod, and linking to the library libsundials\_fsunlinsolband\_mod.lib in addition to the C library. For details on where the library and module file fsunlinsol\_band\_mod.mod are installed see Appendix A. We note that the module is accessible from the Fortran 2003 sundials integrators without separately linking to the libsundials\_fsunlinsolband\_mod library.

### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_BAND module also includes a Fortran-callable function for creating a SUNLinearSolver object.

### FSUNBANDLINSOLINIT

Call FSUNBANDLINSOLINIT(code, ier)

Description The function FSUNBANDLINSOLINIT can be called for Fortran programs to create a band

SUNLinearSolver object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* both the NVECTOR and SUNMATRIX objects have been

initialized.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_BAND module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

### FSUNMASSBANDLINSOLINIT

Call FSUNMASSBANDLINSOLINIT(ier)

Description The function FSUNMASSBANDLINSOLINIT can be called for Fortran programs to create a

band SUNLinearSolver object for mass matrix linear systems.

Arguments None

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called after both the NVECTOR and SUNMATRIX mass-matrix

objects have been initialized.

### 9.6.4 SUNLinearSolver\_Band content

The SUNLINSOL\_BAND module defines the *content* field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_Band {
   sunindextype N;
   sunindextype *pivots;
   long int last_flag;
};
```

These entries of the *content* field contain the following information:

N - size of the linear system,

pivots - index array for partial pivoting in LU factorization,

last\_flag - last error return flag from internal function evaluations.

## 9.7 The SUNLinearSolver\_LapackDense implementation

This section describes the SUNLINSOL implementation for solving dense linear systems with LA-PACK. The SUNLINSOL\_LAPACKDENSE module is designed to be used with the corresponding SUNMA-TRIX\_DENSE matrix type, and one of the serial or shared-memory NVECTOR implementations (NVECTOR\_SERIAL, NVECTOR\_OPENMP, or NVECTOR\_PTHREADS).

To access the SUNLINSOL\_LAPACKDENSE module, include the header file sunlinsol/sunlinsol\_lapackdense.h. The installed module library to link to is libsundials\_sunlinsollapackdense.lib where .lib is typically .so for shared libraries and .a for static libraries.

The SUNLINSOL\_LAPACKDENSE module is a SUNLINSOL wrapper for the LAPACK dense matrix factorization and solve routines, \*GETRF and \*GETRS, where \* is either D or S, depending on whether SUNDIALS was configured to have realtype set to double or single, respectively (see Section 4.2). In order to use the SUNLINSOL\_LAPACKDENSE module it is assumed that LAPACK has been installed on the system prior to installation of SUNDIALS, and that SUNDIALS has been configured appropriately to link with LAPACK (see Appendix A for details). We note that since there do not exist 128-bit floating-point factorization and solve routines in LAPACK, this interface cannot be compiled when using extended precision for realtype. Similarly, since there do not exist 64-bit integer LAPACK routines, the SUNLINSOL\_LAPACKDENSE module also cannot be compiled when using 64-bit integers for the sunindextype.



### 9.7.1 SUNLinearSolver\_LapackDense description

This solver is constructed to perform the following operations:

- The "setup" call performs a LU factorization with partial (row) pivoting  $(\mathcal{O}(N^3) \cos t)$ , PA = LU, where P is a permutation matrix, L is a lower triangular matrix with 1's on the diagonal, and U is an upper triangular matrix. This factorization is stored in-place on the input SUNMATRIX\_DENSE object A, with pivoting information encoding P stored in the pivots array.
- The "solve" call performs pivoting and forward and backward substitution using the stored pivots array and the LU factors held in the SUNMATRIX\_DENSE object  $(\mathcal{O}(N^2) \text{ cost})$ .

## 9.7.2 SUNLinearSolver\_LapackDense functions

The SUNLINSOL\_LAPACKDENSE module provides the following user-callable constructor for creating a SUNLinearSolver object.

### SUNLinSol\_LapackDense

Call LS = SUNLinSol\_LapackDense(y, A);

Description The function SUNLinSol\_LapackDense creates and allocates memory for a LAPACK-

based, dense SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

A (SUNMatrix) a  ${\tt SUNMATRIX\_DENSE}$  matrix template for cloning matrices needed

within the solver

Return value This returns a SUNLinearSolver object. If either A or y are incompatible then this

routine will return NULL.

Notes This routine will perform consistency checks to ensure that it is called with con-

sistent NVECTOR and SUNMATRIX implementations. These are currently limited to the SUNMATRIX\_DENSE matrix type and the NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS vector types. As additional compatible matrix and vector implementations are added to SUNDIALS, these will be included within this

compatibility check.

 $\label{thm:local_problem} Deprecated \ Name \ For \ backward \ compatibility, \ the \ wrapper \ function \ {\tt SUNLapackDense} \ with \ idential$ 

input and output arguments is also provided.

The SUNLINSOL\_LAPACKDENSE module defines dense implementations of all "direct" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_LapackDense
- SUNLinSolInitialize\_LapackDense this does nothing, since all consistency checks are performed at solver creation.
- SUNLinSolSetup\_LapackDense this calls either DGETRF or SGETRF to perform the LU factorization.
- SUNLinSolSolve\_LapackDense this calls either DGETRS or SGETRS to use the *LU* factors and pivots array to perform the solve.
- $\bullet \ {\tt SUNLinSolLastFlag\_LapackDense} \\$
- SUNLinSolSpace\_LapackDense this only returns information for the storage within the solver object, i.e. storage for N, last\_flag, and pivots.
- SUNLinSolFree\_LapackDense

### 9.7.3 SUNLinearSolver\_LapackDense Fortran interfaces

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_LAPACKDENSE module also includes a Fortran-callable function for creating a SUNLinearSolver object.

### FSUNLAPACKDENSEINIT

Call FSUNLAPACKDENSEINIT(code, ier)

Description The function FSUNLAPACKDENSEINIT can be called for Fortran programs to create a

LAPACK-based dense SUNLinearSolver object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* both the NVECTOR and SUNMATRIX objects have been

initialized.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_LAPACKDENSE module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

### FSUNMASSLAPACKDENSEINIT

Call FSUNMASSLAPACKDENSEINIT(ier)

Description The function FSUNMASSLAPACKDENSEINIT can be called for Fortran programs to create

a LAPACK-based, dense SUNLinearSolver object for mass matrix linear systems.

Arguments None

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called after both the NVECTOR and SUNMATRIX mass-matrix

objects have been initialized.

### 9.7.4 SUNLinearSolver\_LapackDense content

The SUNLINSOL\_LAPACKDENSE module defines the *content* field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_Dense {
   sunindextype N;
   sunindextype *pivots;
   long int last_flag;
};
```

These entries of the *content* field contain the following information:

N - size of the linear system,

pivots - index array for partial pivoting in LU factorization,

last\_flag - last error return flag from internal function evaluations.

# 9.8 The SUNLinearSolver\_LapackBand implementation

This section describes the SUNLINSOL implementation for solving banded linear systems with LA-PACK. The SUNLINSOL\_LAPACKBAND module is designed to be used with the corresponding SUNMA-TRIX\_BAND matrix type, and one of the serial or shared-memory NVECTOR implementations (NVECTOR\_SERIAL, NVECTOR\_OPENMP, or NVECTOR\_PTHREADS).

To access the SUNLINSOL\_LAPACKBAND module, include the header file sunlinsol/sunlinsol\_lapackband.h. The installed module library to link to is libsundials\_sunlinsollapackband.lib where .lib is typically .so for shared libraries and .a for static libraries.

The SUNLINSOL\_LAPACKBAND module is a SUNLINSOL wrapper for the LAPACK band matrix factorization and solve routines, \*GBTRF and \*GBTRS, where \* is either D or S, depending on whether SUNDIALS was configured to have realtype set to double or single, respectively (see Section 4.2). In order to use the SUNLINSOL\_LAPACKBAND module it is assumed that LAPACK has been installed on the system prior to installation of SUNDIALS, and that SUNDIALS has been configured appropriately to link with LAPACK (see Appendix A for details). We note that since there do not exist 128-bit floating-point factorization and solve routines in LAPACK, this interface cannot be compiled when using extended precision for realtype. Similarly, since there do not exist 64-bit integer LAPACK routines, the SUNLINSOL\_LAPACKBAND module also cannot be compiled when using 64-bit integers for the sunindextype.



### 9.8.1 SUNLinearSolver\_LapackBand description

This solver is constructed to perform the following operations:

- The "setup" call performs a LU factorization with partial (row) pivoting, PA = LU, where P is a permutation matrix, L is a lower triangular matrix with 1's on the diagonal, and U is an upper triangular matrix. This factorization is stored in-place on the input SUNMATRIX\_BAND object A, with pivoting information encoding P stored in the pivots array.
- The "solve" call performs pivoting and forward and backward substitution using the stored pivots array and the *LU* factors held in the SUNMATRIX\_BAND object.
- A must be allocated to accommodate the increase in upper bandwidth that occurs during factorization. More precisely, if A is a band matrix with upper bandwidth  $\mathtt{mu}$  and lower bandwidth  $\mathtt{ml}$ , then the upper triangular factor U can have upper bandwidth as big as  $\mathtt{smu} = \mathtt{MIN(N-1,mu+ml)}$ . The lower triangular factor L has lower bandwidth  $\mathtt{ml}$ .



### 9.8.2 SUNLinearSolver\_LapackBand functions

The SUNLINSOL\_LAPACKBAND module provides the following user-callable constructor for creating a SUNLinearSolver object.

SUNLinSol_LapackBand

Call LS = SUNLinSol\_LapackBand(y, A);

Description The function SUNLinSol\_LapackBand creates and allocates memory for a LAPACK-

based, band SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

A (SUNMatrix) a SUNMATRIX\_BAND matrix template for cloning matrices needed

within the solver

Return value This returns a SUNLinearSolver object. If either A or y are incompatible then this

routine will return NULL.

Notes This routine will perform consistency checks to ensure that it is called with con-

sistent NVECTOR and SUNMATRIX implementations. These are currently limited to the SUNMATRIX\_BAND matrix type and the NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS vector types. As additional compatible matrix and vector implementations are added to SUNDIALS, these will be included within this

compatibility check.

Additionally, this routine will verify that the input matrix A is allocated with

appropriate upper bandwidth storage for the LU factorization.

Deprecated Name For backward compatibility, the wrapper function SUNLapackBand with idential

input and output arguments is also provided.

The SUNLINSOL\_LAPACKBAND module defines band implementations of all "direct" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_LapackBand
- SUNLinSolInitialize\_LapackBand this does nothing, since all consistency checks are performed at solver creation.
- SUNLinSolSetup\_LapackBand this calls either DGBTRF or SGBTRF to perform the LU factorization.
- SUNLinSolSolve\_LapackBand this calls either DGBTRS or SGBTRS to use the *LU* factors and pivots array to perform the solve.
- SUNLinSolLastFlag\_LapackBand
- SUNLinSolSpace\_LapackBand this only returns information for the storage within the solver object, i.e. storage for N, last\_flag, and pivots.
- SUNLinSolFree\_LapackBand

### 9.8.3 SUNLinearSolver\_LapackBand Fortran interfaces

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_LAPACKBAND module also includes a Fortran-callable function for creating a SUNLinearSolver object.

### FSUNLAPACKDENSEINIT

Call FSUNLAPACKBANDINIT(code, ier)

Description The function FSUNLAPACKBANDINIT can be called for Fortran programs to create a LAPACK-based band SUNLinearSolver object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for KINSOL, and 4 for ARKODE).

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See printed message for details in case of failure.

Notes This routine must be called *after* both the NVECTOR and SUNMATRIX objects have been initialized.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_LAPACKBAND module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

### FSUNMASSLAPACKBANDINIT

Call FSUNMASSLAPACKBANDINIT(ier)

Description The function FSUNMASSLAPACKBANDINIT can be called for Fortran programs to create a LAPACK-based, band SUNLinearSolver object for mass matrix linear systems.

Arguments None

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise. See printed message for details in case of failure.

Notes This routine must be called *after* both the NVECTOR and SUNMATRIX mass-matrix objects have been initialized.

### 9.8.4 SUNLinearSolver\_LapackBand content

The SUNLINSOL\_LAPACKBAND module defines the *content* field of a SUNLinearSolver as the following structure:

## 9.9 The SUNLinearSolver\_KLU implementation

This section describes the SUNLINSOL implementation for solving sparse linear systems with KLU. The SUNLINSOL\_KLU module is designed to be used with the corresponding SUNMATRIX\_SPARSE matrix type, and one of the serial or shared-memory NVECTOR implementations (NVECTOR\_SERIAL, NVECTOR\_OPENMP, or NVECTOR\_PTHREADS).

The header file to include when using this module is sunlinsol/sunlinsol\_klu.h. The installed module library to link to is libsundials\_sunlinsolklu.lib where .lib is typically .so for shared libraries and .a for static libraries.

The SUNLINSOL\_KLU module is a SUNLINSOL wrapper for the KLU sparse matrix factorization and solver library written by Tim Davis [1, 17]. In order to use the SUNLINSOL\_KLU interface to KLU, it is assumed that KLU has been installed on the system prior to installation of SUNDIALS, and that SUNDIALS has been configured appropriately to link with KLU (see Appendix A for details). Additionally, this wrapper only supports double-precision calculations, and therefore cannot be compiled if SUNDIALS is configured to have realtype set to either extended or single (see Section 4.2). Since the KLU library supports both 32-bit and 64-bit integers, this interface will be compiled for either of the available sunindextype options.

# <u>!</u>

### 9.9.1 SUNLinearSolver\_KLU description

The KLU library has a symbolic factorization routine that computes the permutation of the linear system matrix to block triangular form and the permutations that will pre-order the diagonal blocks (the only ones that need to be factored) to reduce fill-in (using AMD, COLAMD, CHOLAMD, natural, or an ordering given by the user). Of these ordering choices, the default value in the SUNLINSOL\_KLU module is the COLAMD ordering.

KLU breaks the factorization into two separate parts. The first is a symbolic factorization and the second is a numeric factorization that returns the factored matrix along with final pivot information. KLU also has a refactor routine that can be called instead of the numeric factorization. This routine will reuse the pivot information. This routine also returns diagnostic information that a user can examine to determine if numerical stability is being lost and a full numerical factorization should be done instead of the refactor.

Since the linear systems that arise within the context of SUNDIALS calculations will typically have identical sparsity patterns, the SUNLINSOL\_KLU module is constructed to perform the following operations:

• The first time that the "setup" routine is called, it performs the symbolic factorization, followed by an initial numerical factorization.

- On subsequent calls to the "setup" routine, it calls the appropriate KLU "refactor" routine, followed by estimates of the numerical conditioning using the relevant "rcond", and if necessary "condest", routine(s). If these estimates of the condition number are larger than  $\varepsilon^{-2/3}$  (where  $\varepsilon$  is the double-precision unit roundoff), then a new factorization is performed.
- The module includes the routine SUNKLUReInit, that can be called by the user to force a full or partial refactorization at the next "setup" call.
- The "solve" call performs pivoting and forward and backward substitution using the stored KLU data structures. We note that in this solve KLU operates on the native data arrays for the right-hand side and solution vectors, without requiring costly data copies.

### 9.9.2 SUNLinearSolver\_KLU functions

The SUNLINSOL\_KLU module provides the following user-callable constructor for creating a SUNLinearSolver object.

CIIMIT	inSol	וו זע
DUNL	insor	_KLU

Call LS = SUNLinSol\_KLU(y, A);

Description The function SUNLinSol\_KLU creates and allocates memory for a KLU-based

SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

A (SUNMatrix) a SUNMATRIX\_SPARSE matrix template for cloning matrices needed

within the solver

Return value This returns a SUNLinearSolver object. If either A or y are incompatible then this

routine will return NULL.

Notes This routine will perform consistency checks to ensure that it is called with con-

sistent NVECTOR and SUNMATRIX implementations. These are currently limited to the SUNMATRIX\_SPARSE matrix type (using either CSR or CSC storage formats) and the NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS vector types. As additional compatible matrix and vector implementations are added to

SUNDIALS, these will be included within this compatibility check.

Deprecated Name For backward compatibility, the wrapper function SUNKLU with idential input and

output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_KLU when using the Fortran 2003 interface

module.

The SUNLINSOL\_KLU module defines implementations of all "direct" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_KLU
- SUNLinSolInitialize\_KLU this sets the first\_factorize flag to 1, forcing both symbolic and numerical factorizations on the subsequent "setup" call.
- SUNLinSolSetup\_KLU this performs either a *LU* factorization or refactorization of the input matrix.
- SUNLinSolSolve\_KLU this calls the appropriate KLU solve routine to utilize the *LU* factors to solve the linear system.
- SUNLinSolLastFlag\_KLU
- SUNLinSolSpace\_KLU this only returns information for the storage within the solver *interface*, i.e. storage for the integers last\_flag and first\_factorize. For additional space requirements, see the KLU documentation.

### • SUNLinSolFree\_KLU

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

The SUNLINSOL\_KLU module also defines the following additional user-callable functions.

### SUNLinSol\_KLUReInit

Call retval = SUNLinSol\_KLUReInit(LS, A, nnz, reinit\_type);

Description The function SUNLinSol\_KLUReInit reinitializes memory and flags for a new fac-

torization (symbolic and numeric) to be conducted at the next solver setup call. This routine is useful in the cases where the number of nonzeroes has changed or if the structure of the linear system has changed which would require a new symbolic

(and numeric factorization).

Arguments LS (SUNLinearSolver) a template for cloning vectors needed within the

solver

A (SUNMatrix) a SUNMATRIX\_SPARSE matrix template for cloning ma-

trices needed within the solver

nnz (sunindextype) the new number of nonzeros in the matrix

reinit\_type (int) flag governing the level of reinitialization. The allowed values

- SUNKLU\_REINIT\_FULL The Jacobian matrix will be destroyed and a new one will be allocated based on the nnz value passed to this call. New symbolic and numeric factorizations will be completed at the next solver setup.
- SUNKLU\_REINIT\_PARTIAL Only symbolic and numeric factorizations will be completed. It is assumed that the Jacobian size has not exceeded the size of nnz given in the sparse matrix provided to the original constructor routine (or the previous SUNLinSol\_KLUReInit call).

Return value The return values from this function are SUNLS\_MEM\_NULL (either S or A are NULL),

SUNLS\_ILL\_INPUT (A does not have type SUNMATRIX\_SPARSE or reinit\_type is invalid), SUNLS\_MEM\_FAIL (reallocation of the sparse matrix failed) or SUNLS\_SUCCESS.

Notes This routine will perform consistency checks to ensure that it is called with con-

sistent NVECTOR and SUNMATRIX implementations. These are currently limited to the SUNMATRIX\_SPARSE matrix type (using either CSR or CSC storage formats) and the NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS vector types. As additional compatible matrix and vector implementations are added to

SUNDIALS, these will be included within this compatibility check.

This routine assumes no other changes to solver use are necessary.

Deprecated Name For backward compatibility, the wrapper function SUNKLUReInit with idential in-

put and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_KLUReInit when using the Fortran 2003

interface module.

### SUNLinSol\_KLUSetOrdering

Call retval = SUNLinSol\_KLUSetOrdering(LS, ordering);

Description This function sets the ordering used by KLU for reducing fill in the linear solve.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_KLU object

ordering (int) flag indicating the reordering algorithm to use, the options are:

0 AMD,

1 COLAMD, and

2 the natural ordering.

The default is 1 for COLAMD.

Return value The return values from this function are SUNLS\_MEM\_NULL (S is NULL),

SUNLS\_ILL\_INPUT (invalid ordering choice), or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNKLUSetOrdering with iden-

tial input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_KLUSetOrdering when using the Fortran

2003 interface module.

### 9.9.3 SUNLinearSolver\_KLU Fortran interfaces

The SUNLINSOL\_KLU module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

### FORTRAN 2003 interface module

The fsunlinsol\_klu\_mod FORTRAN module defines interfaces to all SUNLINSOL\_KLU C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_klu is interfaced as FSUNLinSol\_klu.

The FORTRAN 2003 SUNLINSOL\_KLU interface module can be accessed with the use statement, i.e. use fsunlinsol\_klu\_mod, and linking to the library libsundials\_fsunlinsolklu\_mod.lib in addition to the C library. For details on where the library and module file fsunlinsol\_klu\_mod.mod are installed see Appendix A.

### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_KLU module also includes a Fortran-callable function for creating a SUNLinearSolver object.

### FSUNKLUINIT

Call FSUNKLUINIT(code, ier)

Description The function FSUNKLUINIT can be called for Fortran programs to create a SUNLIN-

SOL\_KLU object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* both the NVECTOR and SUNMATRIX objects have been

initialized.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_KLU module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

### FSUNMASSKLUINIT

Call FSUNMASSKLUINIT(ier)

Description The function FSUNMASSKLUINIT can be called for Fortran programs to create a KLU-

based SUNLinearSolver object for mass matrix linear systems.

Arguments None

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called after both the NVECTOR and SUNMATRIX mass-matrix

objects have been initialized.

The SUNLinSol\_KLUReInit and SUNLinSol\_KLUSetOrdering routines also support FORTRAN interfaces for the system and mass matrix solvers:

### FSUNKLUREINIT

Call FSUNKLUREINIT(code, nnz, reinit\_type, ier)

Description The function FSUNKLUREINIT can be called for Fortran programs to re-initialize a SUN-

LINSOL\_KLU object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA,

3 for KINSOL, and 4 for ARKODE).

nnz (sunindextype\*) the new number of nonzeros in the matrix

reinit\_type (int\*) flag governing the level of reinitialization. The allowed values are:

1 – The Jacobian matrix will be destroyed and a new one will be allocated based on the nnz value passed to this call. New symbolic and numeric factorizations will be completed at the next solver setup.

2 - Only symbolic and numeric factorizations will be completed. It is assumed that the Jacobian size has not exceeded the size of nnz given in the sparse matrix provided to the original constructor routine (or the previous SUNLinSol\_KLUReInit call).

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise. See printed message for details in case of failure.

See SUNLinSol\_KLUReInit for complete further documentation of this routine.

occ printed message for details in ease of failure.

FSUNMASSKLUREINIT

Notes

Call FSUNMASSKLUREINIT(nnz, reinit\_type, ier)

Description The function FSUNMASSKLUREINIT can be called for Fortran programs to re-initialize a

SUNLINSOL\_KLU object for mass matrix linear systems.

Arguments The arguments are identical to FSUNKLUREINIT above, except that code is not needed

since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_KLUReInit for complete further documentation of this routine.

### FSUNKLUSETORDERING

Call FSUNKLUSETORDERING(code, ordering, ier)

Description The function FSUNKLUSETORDERING can be called for Fortran programs to change the

reordering algorithm used by KLU.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

ordering (int\*) flag indication the reordering algorithm to use. Options include:

0 AMD,

1 COLAMD, and

2 the natural ordering.

The default is 1 for COLAMD.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_KLUSetOrdering for complete further documentation of this routine.

### FSUNMASSKLUSETORDERING

Call FSUNMASSKLUSETORDERING(ier)

Description The function FSUNMASSKLUSETORDERING can be called for Fortran programs to change

the reordering algorithm used by KLU for mass matrix linear systems.

Arguments The arguments are identical to FSUNKLUSETORDERING above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_KLUSetOrdering for complete further documentation of this routine.

### 9.9.4 SUNLinearSolver\_KLU content

The SUNLINSOL\_KLU module defines the content field of a SUNLinearSolver as the following structure:

These entries of the *content* field contain the following information:

last\_flag - last error return flag from internal function evaluations,

first\_factorize - flag indicating whether the factorization has ever been performed,

symbolic - KLU storage structure for symbolic factorization components,

numeric - KLU storage structure for numeric factorization components,

- storage structure for common KLU solver components,

klu\_solver — pointer to the appropriate KLU solver function (depending on whether it is using

a CSR or CSC sparse matrix).

## 9.10 The SUNLinearSolver\_SuperLUMT implementation

This section describes the SUNLINSOL implementation for solving sparse linear systems with SuperLU\_MT. The SUPERLUMT module is designed to be used with the corresponding SUNMATRIX\_SPARSE matrix type, and one of the serial or shared-memory NVECTOR implementations (NVECTOR\_SERIAL, NVECTOR\_OPENMP, or NVECTOR\_PTHREADS). While these are compatible, it is not recommended to use a threaded vector module with SUNLINSOL\_SUPERLUMT unless it is the NVECTOR\_OPENMP module and the SUPERLUMT library has also been compiled with OpenMP.

The header file to include when using this module is sunlinsol/sunlinsol\_superlumt.h. The installed module library to link to is libsundials\_sunlinsolsuperlumt. lib where .lib is typically .so for shared libraries and .a for static libraries.

The SUNLINSOL\_SUPERLUMT module is a SUNLINSOL wrapper for the SUPERLUMT sparse matrix factorization and solver library written by X. Sherry Li [2, 35, 19]. The package performs matrix factorization using threads to enhance efficiency in shared memory parallel environments. It should be noted that threads are only used in the factorization step. In order to use the SUNLINSOL\_SUPERLUMT interface to SUPERLUMT, it is assumed that SUPERLUMT has been installed on the system prior to installation of SUNDIALS, and that SUNDIALS has been configured appropriately to link with SUPERLUMT (see Appendix A for details). Additionally, this wrapper only supports single- and double-precision calculations, and therefore cannot be compiled if SUNDIALS is configured to have realtype set to extended (see Section 4.2). Moreover, since the SUPERLUMT library may be installed to support either 32-bit or 64-bit integers, it is assumed that the SUPERLUMT library is installed using the same integer precision as the SUNDIALS sunindextype option.



## 9.10.1 SUNLinearSolver\_SuperLUMT description

The SUPERLUMT library has a symbolic factorization routine that computes the permutation of the linear system matrix to reduce fill-in on subsequent LU factorizations (using COLAMD, minimal degree ordering on  $A^T * A$ , minimal degree ordering on  $A^T * A$ , or natural ordering). Of these ordering choices, the default value in the SUNLINSOL\_SUPERLUMT module is the COLAMD ordering.

Since the linear systems that arise within the context of SUNDIALS calculations will typically have identical sparsity patterns, the SUNLINSOL\_SUPERLUMT module is constructed to perform the following operations:

- The first time that the "setup" routine is called, it performs the symbolic factorization, followed by an initial numerical factorization.
- On subsequent calls to the "setup" routine, it skips the symbolic factorization, and only refactors the input matrix.
- The "solve" call performs pivoting and forward and backward substitution using the stored SUPERLUMT data structures. We note that in this solve SUPERLUMT operates on the native data arrays for the right-hand side and solution vectors, without requiring costly data copies.

## 9.10.2 SUNLinearSolver\_SuperLUMT functions

The module SUNLINSOL\_SUPERLUMT provides the following user-callable constructor for creating a SUNLinearSolver object.

SUNLinSol_Super	rLUMT				
Call	LS = SUNLin	= SUNLinSol_SuperLUMT(y, A, num_threads);			
Description		The function SUNLinSol_SuperLUMT creates and allocates memory for a SuperLU_MT-based SUNLinearSolver object.			
Arguments	y A	(N_Vector) a template for cloning vectors needed within the solver (SUNMatrix) a SUNMATRIX_SPARSE matrix template for cloning matrices needed within the solver			
	num_threads	(int) desired number of threads (OpenMP or Pthreads, depending on how SUPERLUMT was installed) to use during the factorization steps			
Return value	This returns a SUNLinearSolver object. If either A or y are incompatible then t routine will return NULL.				

Notes

This routine analyzes the input matrix and vector to determine the linear system size and to assess compatibility with the SUPERLUMT library.

This routine will perform consistency checks to ensure that it is called with consistent NVECTOR and SUNMATRIX implementations. These are currently limited to the SUNMATRIX\_SPARSE matrix type (using either CSR or CSC storage formats) and the NVECTOR\_SERIAL, NVECTOR\_OPENMP, and NVECTOR\_PTHREADS vector types. As additional compatible matrix and vector implementations are added to SUNDIALS, these will be included within this compatibility check.

The num\_threads argument is not checked and is passed directly to SUPERLUMT routines.

Deprecated Name For backward compatibility, the wrapper function SUNSuperLUMT with idential input and output arguments is also provided.

The SUNLINSOL\_SUPERLUMT module defines implementations of all "direct" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_SuperLUMT
- SUNLinSolInitialize\_SuperLUMT this sets the first\_factorize flag to 1 and resets the internal SUPERLUMT statistics variables.
- SUNLinSolSetup\_SuperLUMT this performs either a *LU* factorization or refactorization of the input matrix.
- SUNLinSolSolve\_SuperLUMT this calls the appropriate SUPERLUMT solve routine to utilize the *LU* factors to solve the linear system.
- SUNLinSolLastFlag\_SuperLUMT
- SUNLinSolSpace\_SuperLUMT this only returns information for the storage within the solver *interface*, i.e. storage for the integers last\_flag and first\_factorize. For additional space requirements, see the SUPERLUMT documentation.
- SUNLinSolFree\_SuperLUMT

The SUNLINSOL\_SUPERLUMT module also defines the following additional user-callable function.

## ${\tt SUNLinSol\_SuperLUMTSetOrdering}$

Call retval = SUNLinSol\_SuperLUMTSetOrdering(LS, ordering);

Description This function sets the ordering used by SUPERLUMT for reducing fill in the linear

solve.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SUPERLUMT object

ordering (int) a flag indicating the ordering algorithm to use, the options are:

0 natural ordering

1 minimal degree ordering on  $A^TA$ 

2 minimal degree ordering on  $A^T + A$ 

3 COLAMD ordering for unsymmetric matrices

The default is 3 for COLAMD.

 ${\it Return\ values\ from\ this\ function\ are\ {\tt SUNLS\_MEM\_NULL\ (S\ is\ NULL)}},$ 

SUNLS\_ILL\_INPUT (invalid ordering choice), or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSuperLUMTSetOrdering with idential input and output arguments is also provided.

### 9.10.3 SUNLinearSolver\_SuperLUMT Fortran interfaces

For solvers that include a Fortran interface module, the SUNLINSOL\_SUPERLUMT module also includes a Fortran-callable function for creating a SUNLinearSolver object.

### FSUNSUPERLUMTINIT

Call FSUNSUPERLUMTINIT(code, num\_threads, ier)

Description The function FSUNSUPERLUMTINIT can be called for Fortran programs to create a SUN-

LINSOL\_KLU object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA,

3 for KINSOL, and 4 for ARKODE).

num\_threads (int\*) desired number of threads (OpenMP or Pthreads, depending on

how superlumt was installed) to use during the factorization steps

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* both the NVECTOR and SUNMATRIX objects have been

initialized.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_SUPERLUMT module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

### FSUNMASSSUPERLUMTINIT

Call FSUNMASSSUPERLUMTINIT(num\_threads, ier)

Description The function FSUNMASSSUPERLUMTINIT can be called for Fortran programs to create a

SuperLU\_MT-based SUNLinearSolver object for mass matrix linear systems.

Arguments num\_threads (int\*) desired number of threads (OpenMP or Pthreads, depending on

how superlumt was installed) to use during the factorization steps.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called after both the NVECTOR and SUNMATRIX mass-matrix

objects have been initialized.

The SUNLinSol\_SuperLUMTSetOrdering routine also supports Fortran interfaces for the system and mass matrix solvers:

### FSUNSUPERLUMTSETORDERING

Call FSUNSUPERLUMTSETORDERING(code, ordering, ier)

Description The function FSUNSUPERLUMTSETORDERING can be called for Fortran programs to update

the ordering algorithm in a SUNLINSOL\_SUPERLUMT object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

ordering (int\*) a flag indicating the ordering algorithm, options are:

0 natural ordering

1 minimal degree ordering on  $A^TA$ 

2 minimal degree ordering on  $A^T + A$ 

3 COLAMD ordering for unsymmetric matrices

The default is 3 for COLAMD.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise. See printed message for details in case of failure.

Notes See SUNLinSol\_SuperLUMTSetOrdering for complete further documentation of this rou-

tine.

### FSUNMASSUPERLUMTSETORDERING

Call FSUNMASSUPERLUMTSETORDERING(ordering, ier)

Description The function FSUNMASSUPERLUMTSETORDERING can be called for Fortran programs to

 $update\ the\ ordering\ algorithm\ in\ a\ {\tt SUNLINSOL\_SUPERLUMT}\ object\ for\ mass\ matrix\ linear$ 

systems.

Arguments ordering (int\*) a flag indicating the ordering algorithm, options are:

0 natural ordering

1 minimal degree ordering on  $A^TA$ 

2 minimal degree ordering on  $A^T + A$ 

3 COLAMD ordering for unsymmetric matrices

The default is 3 for COLAMD.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SuperLUMTSetOrdering for complete further documentation of this rou-

tine.

## 9.10.4 SUNLinearSolver\_SuperLUMT content

The SUNLINSOL\_SUPERLUMT module defines the *content* field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_SuperLUMT {
               last_flag;
  long int
               first_factorize;
  int
  SuperMatrix *A, *AC, *L, *U, *B;
  Gstat_t
               *Gstat;
  sunindextype *perm_r, *perm_c;
  sunindextype N;
  int
               num_threads;
               diag_pivot_thresh;
  realtype
  int
               ordering;
  superlumt_options_t *options;
};
```

These entries of the *content* field contain the following information:

last\_flag - last error return flag from internal function evaluations,

first\_factorize - flag indicating whether the factorization has ever been performed,

A, AC, L, U, B - SuperMatrix pointers used in solve,

 ${\tt Gstat\_t\ object\ used\ in\ solve},$ 

perm\_r, perm\_c - permutation arrays used in solve,

N - size of the linear system,

num\_threads - number of OpenMP/Pthreads threads to use,

diag\_pivot\_thresh - threshold on diagonal pivoting,

ordering - flag for which reordering algorithm to use,
options - pointer to SUPERLUMT options structure.

## 9.11 The SUNLinearSolver\_SPGMR implementation

This section describes the SUNLINSOL implementation of the SPGMR (Scaled, Preconditioned, Generalized Minimum Residual [41]) iterative linear solver. The SUNLINSOL\_SPGMR module is designed to be compatible with any NVECTOR implementation that supports a minimal subset of operations (N\_VClone, N\_VDotProd, N\_VScale, N\_VLinearSum, N\_VProd, N\_VConst, N\_VDiv, and N\_VDestroy). When using Classical Gram-Schmidt, the optional function N\_VDotProdMulti may be supplied for increased efficiency.

To access the SUNLINSOL\_SPGMR module, include the header file sunlinsol/sunlinsol\_spgmr.h. We note that the SUNLINSOL\_SPGMR module is accessible from SUNDIALS packages without separately linking to the libsundials\_sunlinsolspgmr module library.

### 9.11.1 SUNLinearSolver\_SPGMR description

This solver is constructed to perform the following operations:

- During construction, the xcor and vtemp arrays are cloned from a template NVECTOR that is input, and default solver parameters are set.
- User-facing "set" routines may be called to modify default solver parameters.
- Additional "set" routines are called by the SUNDIALS solver that interfaces with SUNLINSOL\_SPGMR to supply the ATimes, PSetup, and Psolve function pointers and s1 and s2 scaling vectors.
- In the "initialize" call, the remaining solver data is allocated (V, Hes, givens, and yg)
- In the "setup" call, any non-NULL PSetup function is called. Typically, this is provided by the SUNDIALS solver itself, that translates between the generic PSetup function and the solver-specific routine (solver-supplied or user-supplied).
- In the "solve" call, the GMRES iteration is performed. This will include scaling, preconditioning, and restarts if those options have been supplied.

### 9.11.2 SUNLinearSolver\_SPGMR functions

The SUNLINSOL\_SPGMR module provides the following user-callable constructor for creating a SUNLinearSolver object.

### SUNLinSol\_SPGMR

Call LS = SUNLinSol\_SPGMR(y, pretype, maxl);

Description The function SUNLinSol\_SPGMR creates and allocates memory for a SPGMR

SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

- PREC\_NONE (0)
- PREC\_LEFT (1)
- PREC\_RIGHT (2)
- PREC\_BOTH (3)

Any other integer input will result in the default (no preconditioning).

maxl (int) the number of Krylov basis vectors to use. Values  $\leq 0$  will result in the default value (5).

Return value This returns a SUNLinearSolver object. If either y is incompatible then this routine will return NULL.

Notes

This routine will perform consistency checks to ensure that it is called with a consistent NVECTOR implementation (i.e. that it supplies the requisite vector operations). If y is incompatible, then this routine will return NULL.

We note that some SUNDIALS solvers are designed to only work with left preconditioning (IDA and IDAS) and others with only right preconditioning (KINSOL). While it is possible to configure a SUNLINSOL\_SPGMR object to use any of the preconditioning options with these solvers, this use mode is not supported and may result in inferior performance.

Deprecated Name For backward compatibility, the wrapper function SUNSPGMR with idential input and output arguments is also provided.

F2003 Name

This function is callable as FSUNLinSol\_SPGMR when using the Fortran 2003 interface module.

The SUNLINSOL\_SPGMR module defines implementations of all "iterative" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_SPGMR
- SUNLinSolInitialize\_SPGMR
- SUNLinSolSetATimes\_SPGMR
- SUNLinSolSetPreconditioner\_SPGMR
- SUNLinSolSetScalingVectors\_SPGMR
- SUNLinSolSetup\_SPGMR
- SUNLinSolSolve\_SPGMR
- SUNLinSolNumIters\_SPGMR
- SUNLinSolResNorm\_SPGMR
- SUNLinSolResid\_SPGMR
- SUNLinSolLastFlag\_SPGMR
- SUNLinSolSpace\_SPGMR
- SUNLinSolFree\_SPGMR

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

The SUNLINSOL\_SPGMR module also defines the following additional user-callable functions.

### SUNLinSol\_SPGMRSetPrecType

Call retval = SUNLinSol\_SPGMRSetPrecType(LS, pretype);

The function SUNLinSol\_SPGMRSetPrecType updates the type of preconditioning Description

to use in the SUNLINSOL\_SPGMR object.

Arguments (SUNLinearSolver) the SUNLINSOL\_SPGMR object to update

pretype (int) flag indicating the desired type of preconditioning, allowed values

match those discussed in SUNLinSol\_SPGMR.

Return value This routine will return with one of the error codes SUNLS\_ILL\_INPUT (illegal

pretype), SUNLS\_MEM\_NULL (S is NULL) or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPGMRSetPrecType with

idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPGMRSetPrecType when using the Fortran

2003 interface module.

### SUNLinSol\_SPGMRSetGSType

Call retval = SUNLinSol\_SPGMRSetGSType(LS, gstype);

Description The function SUNLinSol\_SPGMRSetPrecType sets the type of Gram-Schmidt or-

thogonalization to use in the SUNLINSOL\_SPGMR object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPGMR object to update

gstype (int) flag indicating the desired orthogonalization algorithm; allowed val-

ues are:

• MODIFIED\_GS (1)

• CLASSICAL\_GS (2)

Any other integer input will result in a failure, returning error code SUNLS\_ILL\_INPUT.

Return value This routine will return with one of the error codes SUNLS\_ILL\_INPUT (illegal

pretype), SUNLS\_MEM\_NULL (S is NULL) or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPGMRSetGSType with iden-

tial input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPGMRSetGSType when using the Fortran

2003 interface module.

### SUNLinSol\_SPGMRSetMaxRestarts

Call retval = SUNLinSol\_SPGMRSetMaxRestarts(LS, maxrs);

Description The function SUNLinSol\_SPGMRSetMaxRestarts sets the number of GMRES restarts

to allow in the SUNLINSOL\_SPGMR object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPGMR object to update

maxrs (int) integer indicating number of restarts to allow. A negative input will

result in the default of 0.

Return value This routine will return with one of the error codes SUNLS\_MEM\_NULL (S is NULL) or

SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPGMRSetMaxRestarts with

idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPGMRSetMaxRestarts when using the

Fortran 2003 interface module.

### 9.11.3 SUNLinearSolver\_SPGMR Fortran interfaces

The SUNLINSOL\_SPGMR module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

### FORTRAN 2003 interface module

The fsunlinsol\_spgmr\_mod Fortran module defines interfaces to all Sunlinsol\_spgmr C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_SPGMR is interfaced as FSUNLinSol\_SPGMR.

The Fortran 2003 sunlinsol\_spgmr\_mod, and linking to the library libsundials\_fsunlinsolspgmr\_mod. lib in addition to the C library. For details on where the library and module file fsunlinsol\_spgmr\_mod.mod are installed see Appendix A. We note that the module is accessible from the Fortran 2003 sundials integrators without separately linking to the libsundials\_fsunlinsolspgmr\_mod library.

### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_SPGMR module also includes a Fortran-callable function for creating a SUNLinearSolver object.

### FSUNSPGMRINIT

Call FSUNSPGMRINIT(code, pretype, maxl, ier)

Description The function FSUNSPGMRINIT can be called for Fortran programs to create a SUNLIN-

SOL\_SPGMR object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating desired preconditioning type

maxl (int\*) flag indicating Krylov subspace size

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

SUNLinSol\_SPGMR.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_SPGMR module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

### FSUNMASSSPGMRINIT

Call FSUNMASSSPGMRINIT(pretype, maxl, ier)

Description The function FSUNMASSSPGMRINIT can be called for Fortran programs to create a SUN-

LINSOL\_SPGMR object for mass matrix linear systems.

Arguments pretype (int\*) flag indicating desired preconditioning type

maxl (int\*) flag indicating Krylov subspace size

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

 ${\tt SUNLinSol\_SPGMR}.$ 

The SUNLinSol\_SPGMRSetPrecType, SUNLinSol\_SPGMRSetGSType and

SUNLinSol\_SPGMRSetMaxRestarts routines also support Fortran interfaces for the system and mass matrix solvers.

### FSUNSPGMRSETGSTYPE

Call FSUNSPGMRSETGSTYPE(code, gstype, ier)

Description The function FSUNSPGMRSETGSTYPE can be called for Fortran programs to change the

Gram-Schmidt orthogonaliation algorithm.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

gstype (int\*) flag indicating the desired orthogonalization algorithm.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPGMRSetGSType for complete further documentation of this routine.

### FSUNMASSSPGMRSETGSTYPE

Call FSUNMASSSPGMRSETGSTYPE(gstype, ier)

Description The function FSUNMASSSPGMRSETGSTYPE can be called for Fortran programs to change

the Gram-Schmidt orthogonaliation algorithm for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPGMRSETGSTYPE above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPGMRSetGSType for complete further documentation of this routine.

### FSUNSPGMRSETPRECTYPE

Call FSUNSPGMRSETPRECTYPE(code, pretype, ier)

Description The function FSUNSPGMRSETPRECTYPE can be called for Fortran programs to change the

type of preconditioning to use.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating the type of preconditioning to use.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPGMRSetPrecType for complete further documentation of this routine.

### FSUNMASSSPGMRSETPRECTYPE

Call FSUNMASSSPGMRSETPRECTYPE(pretype, ier)

Description The function FSUNMASSSPGMRSETPRECTYPE can be called for Fortran programs to change

the type of preconditioning for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPGMRSETPRECTYPE above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPGMRSetPrecType for complete further documentation of this routine.

### FSUNSPGMRSETMAXRS

Call FSUNSPGMRSETMAXRS(code, maxrs, ier)

Description The function FSUNSPGMRSETMAXRS can be called for Fortran programs to change the

maximum number of restarts allowed for SPGMR.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

maxrs (int\*) maximum allowed number of restarts.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPGMRSetMaxRestarts for complete further documentation of this rou-

tine.

Psetup

### FSUNMASSSPGMRSETMAXRS

Call FSUNMASSSPGMRSETMAXRS(maxrs, ier)

Description The function FSUNMASSSPGMRSETMAXRS can be called for Fortran programs to change

the maximum number of restarts allowed for SPGMR for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPGMRSETMAXRS above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPGMRSetMaxRestarts for complete further documentation of this rou-

tine.

### 9.11.4 SUNLinearSolver\_SPGMR content

The  $\mathtt{SUNLINSOL\_SPGMR}$  module defines the content field of a  $\mathtt{SUNLinearSolver}$  as the following structure:

```
struct _SUNLinearSolverContent_SPGMR {
  int maxl;
  int pretype;
  int gstype;
  int max_restarts;
  int numiters;
  realtype resnorm;
  long int last_flag;
  ATimesFn ATimes;
  void* ATData;
  PSetupFn Psetup;
  PSolveFn Psolve;
  void* PData;
  N_Vector s1;
  N_Vector s2;
  N_Vector *V;
  realtype **Hes;
  realtype *givens;
  N_Vector xcor;
  realtype *yg;
  N_Vector vtemp;
};
These entries of the content field contain the following information:
              - number of GMRES basis vectors to use (default is 5),
maxl
              - flag for type of preconditioning to employ (default is none),
pretype
              - flag for type of Gram-Schmidt orthogonalization (default is modified Gram-Schmidt),
gstype
max_restarts - number of GMRES restarts to allow (default is 0),
numiters
              - number of iterations from the most-recent solve,
resnorm
              - final linear residual norm from the most-recent solve,
              - last error return flag from an internal function,
last_flag
              - function pointer to perform Av product,
ATimes
ATData
              - pointer to structure for ATimes,
```

- function pointer to preconditioner setup routine,

Psolve - function pointer to preconditioner solve routine,
PData - pointer to structure for Psetup and Psolve,
s1, s2 - vector pointers for supplied scaling matrices (default is NULL),
V - the array of Krylov basis vectors v<sub>1</sub>,..., v<sub>maxl+1</sub>, stored in V[0],..., V[maxl]. Each v<sub>i</sub> is a vector of type NVECTOR.,
Hes - the (maxl + 1) × maxl Hessenberg matrix. It is stored row-wise so that the (i.j)th

- the  $(\max 1 + 1) \times \max 1$  Hessenberg matrix. It is stored row-wise so that the (i,j)th element is given by Hes[i][j].,

- a length 2\*max1 array which represents the Givens rotation matrices that arise in the GMRES algorithm. These matrices are  $F_0, F_1, \ldots, F_j$ , where

$$F_i = \begin{bmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & \\ & & & c_i & -s_i & & & \\ & & & s_i & c_i & & & \\ & & & & 1 & & \\ & & & & \ddots & & \\ & & & & & 1 \end{bmatrix},$$

are represented in the givens vector as givens [0] =  $c_0$ , givens [1] =  $s_0$ , givens [2] =  $c_1$ , givens [3] =  $s_1$ , ... givens [2j] =  $c_j$ , givens [2j+1] =  $s_j$ .

xcor - a vector which holds the scaled, preconditioned correction to the initial guess,

yg - a length (maxl+1) array of realtype values used to hold "short" vectors (e.g. y and g),

vtemp - temporary vector storage.

## 9.12 The SUNLinearSolver\_SPFGMR implementation

This section describes the SUNLINSOL implementation of the SPFGMR (Scaled, Preconditioned, Flexible, Generalized Minimum Residual [40]) iterative linear solver. The SUNLINSOL\_SPFGMR module is designed to be compatible with any NVECTOR implementation that supports a minimal subset of operations (N\_VClone, N\_VDotProd, N\_VScale, N\_VLinearSum, N\_VProd, N\_VConst, N\_VDiv, and N\_VDestroy). When using Classical Gram-Schmidt, the optional function N\_VDotProdMulti may be supplied for increased efficiency. Unlike the other Krylov iterative linear solvers supplied with SUNDIALS, SPFGMR is specifically designed to work with a changing preconditioner (e.g. from an iterative method).

To access the SUNLINSOL\_SPFGMR module, include the header file sunlinsol/sunlinsol\_spfgmr.h. We note that the SUNLINSOL\_SPFGMR module is accessible from SUNDIALS packages without separately linking to the libsundials\_sunlinsolspfgmr module library.

### 9.12.1 SUNLinearSolver\_SPFGMR description

This solver is constructed to perform the following operations:

- During construction, the xcor and vtemp arrays are cloned from a template NVECTOR that is input, and default solver parameters are set.
- User-facing "set" routines may be called to modify default solver parameters.
- Additional "set" routines are called by the SUNDIALS solver that interfaces with SUNLINSOL\_SPFGMR to supply the ATimes, PSetup, and Psolve function pointers and s1 and s2 scaling vectors.

- In the "initialize" call, the remaining solver data is allocated (V, Hes, givens, and yg)
- In the "setup" call, any non-NULL PSetup function is called. Typically, this is provided by the SUNDIALS solver itself, that translates between the generic PSetup function and the solver-specific routine (solver-supplied or user-supplied).
- In the "solve" call, the FGMRES iteration is performed. This will include scaling, preconditioning, and restarts if those options have been supplied.

### 9.12.2 SUNLinearSolver\_SPFGMR functions

The SUNLINSOL\_SPFGMR module provides the following user-callable constructor for creating a SUNLinearSolver object.

### SUNLinSol\_SPFGMR

Call LS = SUNLinSol\_SPFGMR(y, pretype, maxl);

Description The function SUNLinSol\_SPFGMR creates and allocates memory for a SPFGMR

SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver pretype (int) flag indicating the desired type of preconditioning, allowed values are:

- PREC\_NONE (0)
- PREC\_LEFT (1)
- PREC\_RIGHT (2)
- PREC\_BOTH (3)

Any other integer input will result in the default (no preconditioning).

maxl (int) the number of Krylov basis vectors to use. Values  $\leq 0$  will result in the default value (5).

Return value This returns a SUNLinearSolver object. If either y is incompatible then this routine will return NULL.

Notes This routine will perform consistency checks to ensure that it is called with a consistent NVECTOR implementation (i.e. that it supplies the requisite vector operations). If y is incompatible, then this routine will return NULL.

We note that some SUNDIALS solvers are designed to only work with left preconditioning (IDA and IDAS) and others with only right preconditioning (KINSOL). While it is possible to configure a SUNLINSOL\_SPFGMR object to use any of the preconditioning options with these solvers, this use mode is not supported and may result in inferior performance.

F2003 Name This function is callable as FSUNLinSol\_SPFGMR when using the Fortran 2003 interface module.

SUNSPFGMR The SUNLINSOL\_SPFGMR module defines implementations of all "iterative" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_SPFGMR
- SUNLinSolInitialize\_SPFGMR
- SUNLinSolSetATimes\_SPFGMR
- SUNLinSolSetPreconditioner\_SPFGMR
- SUNLinSolSetScalingVectors\_SPFGMR
- SUNLinSolSetup\_SPFGMR

- SUNLinSolSolve\_SPFGMR
- SUNLinSolNumIters\_SPFGMR
- SUNLinSolResNorm\_SPFGMR
- SUNLinSolResid\_SPFGMR
- SUNLinSolLastFlag\_SPFGMR
- SUNLinSolSpace\_SPFGMR
- SUNLinSolFree\_SPFGMR

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

The SUNLINSOL\_SPFGMR module also defines the following additional user-callable functions.

# SUNLinSol\_SPFGMRSetPrecType

Call retval = SUNLinSol\_SPFGMRSetPrecType(LS, pretype);

Description The function SUNLinSol\_SPFGMRSetPrecType updates the type of preconditioning

to use in the SUNLINSOL\_SPFGMR object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPFGMR object to update

pretype (int) flag indicating the desired type of preconditioning, allowed values

match those discussed in SUNLinSol\_SPFGMR.

Return value This routine will return with one of the error codes SUNLS\_ILL\_INPUT (illegal

pretype), SUNLS\_MEM\_NULL (S is NULL) or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPFGMRSetPrecType with

idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPFGMRSetPrecType when using the For-

tran 2003 interface module.

# ${\tt SUNLinSol\_SPFGMRSetGSType}$

Call retval = SUNLinSol\_SPFGMRSetGSType(LS, gstype);

Description The function SUNLinSol\_SPFGMRSetPrecType sets the type of Gram-Schmidt or-

thogonalization to use in the SUNLINSOL\_SPFGMR object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPFGMR object to update

gstype (int) flag indicating the desired orthogonalization algorithm; allowed val-

ues are:

• MODIFIED\_GS (1)

• CLASSICAL\_GS (2)

Any other integer input will result in a failure, returning error code SUNLS\_ILL\_INPUT.

Return value This routine will return with one of the error codes SUNLS\_ILL\_INPUT (illegal

pretype), SUNLS\_MEM\_NULL (S is NULL) or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPFGMRSetGSType with iden-

tial input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPFGMRSetGSType when using the Fortran

2003 interface module.

#### SUNLinSol\_SPFGMRSetMaxRestarts

Call retval = SUNLinSol\_SPFGMRSetMaxRestarts(LS, maxrs);

Description The function SUNLinSol\_SPFGMRSetMaxRestarts sets the number of GMRES

restarts to allow in the SUNLINSOL\_SPFGMR object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPFGMR object to update

maxrs (int) integer indicating number of restarts to allow. A negative input will

result in the default of 0.

Return value This routine will return with one of the error codes SUNLS\_MEM\_NULL (S is NULL) or

SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPFGMRSetMaxRestarts with

idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPFGMRSetMaxRestarts when using the

Fortran 2003 interface module.

#### 9.12.3 SUNLinearSolver\_SPFGMR Fortran interfaces

The SUNLINSOL\_SPFGMR module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

# FORTRAN 2003 interface module

The fsunlinsol\_spfgmr\_mod FORTRAN module defines interfaces to all SUNLINSOL\_SPFGMR C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_SPFGMR is interfaced as FSUNLinSol\_SPFGMR.

The FORTRAN 2003 SUNLINSOL\_SPFGMR interface module can be accessed with the use statement, i.e. use fsunlinsol\_spfgmr\_mod, and linking to the library libsundials\_fsunlinsolspfgmr\_mod.lib in addition to the C library. For details on where the library and module file

fsunlinsol\_spfgmr\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunlinsolspfgmr\_mod library.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_SPFGMR module also includes a Fortran-callable function for creating a SUNLinearSolver object.

#### FSUNSPFGMRINIT

Call FSUNSPFGMRINIT(code, pretype, maxl, ier)

Description The function FSUNSPFGMRINIT can be called for Fortran programs to create a SUNLIN-

SOL\_SPFGMR object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating desired preconditioning type

maxl (int\*) flag indicating Krylov subspace size

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

SUNLinSol\_SPFGMR.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_SPFGMR module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

#### FSUNMASSSPFGMRINIT

Call FSUNMASSSPFGMRINIT(pretype, maxl, ier)

Description The function FSUNMASSSPFGMRINIT can be called for Fortran programs to create a SUN-

LINSOL\_SPFGMR object for mass matrix linear systems.

Arguments pretype (int\*) flag indicating desired preconditioning type

maxl (int\*) flag indicating Krylov subspace size

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

SUNLinSol\_SPFGMR.

The SUNLinSol\_SPFGMRSetPrecType, SUNLinSol\_SPFGMRSetGSType and

SUNLinSol\_SPFGMRSetMaxRestarts routines also support Fortran interfaces for the system and mass matrix solvers.

#### FSUNSPFGMRSETGSTYPE

Call FSUNSPFGMRSETGSTYPE(code, gstype, ier)

Description The function FSUNSPFGMRSETGSTYPE can be called for Fortran programs to change the

Gram-Schmidt orthogonaliation algorithm.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

gstype (int\*) flag indicating the desired orthogonalization algorithm.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPFGMRSetGSType for complete further documentation of this routine.

#### FSUNMASSSPFGMRSETGSTYPE

Call FSUNMASSSPFGMRSETGSTYPE(gstype, ier)

Description The function FSUNMASSSPFGMRSETGSTYPE can be called for Fortran programs to change

the Gram-Schmidt orthogonaliation algorithm for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPFGMRSETGSTYPE above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPFGMRSetGSType for complete further documentation of this routine.

#### FSUNSPFGMRSETPRECTYPE

Call FSUNSPFGMRSETPRECTYPE(code, pretype, ier)

Description The function FSUNSPFGMRSETPRECTYPE can be called for Fortran programs to change

the type of preconditioning to use.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating the type of preconditioning to use.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise. See printed message for details in case of failure.

Notes See SUNLinSol\_SPFGMRSetPrecType for complete further documentation of this routine.

#### FSUNMASSSPFGMRSETPRECTYPE

Call FSUNMASSSPFGMRSETPRECTYPE(pretype, ier)

Description The function FSUNMASSSPFGMRSETPRECTYPE can be called for Fortran programs to change

the type of preconditioning for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPFGMRSETPRECTYPE above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPFGMRSetPrecType for complete further documentation of this routine.

#### FSUNSPFGMRSETMAXRS

Call FSUNSPFGMRSETMAXRS(code, maxrs, ier)

Description The function FSUNSPFGMRSETMAXRS can be called for Fortran programs to change the

maximum number of restarts allowed for SPFGMR.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

maxrs (int\*) maximum allowed number of restarts.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPFGMRSetMaxRestarts for complete further documentation of this rou-

tine.

#### FSUNMASSSPFGMRSETMAXRS

Call FSUNMASSSPFGMRSETMAXRS(maxrs, ier)

 $\label{thm:constraint} \textbf{Description} \quad \textbf{The function } \textbf{FSUNMASSSPFGMRSETMAXRS} \ \ \textbf{can be called for Fortran programs to change}$ 

the maximum number of restarts allowed for SPFGMR for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPFGMRSETMAXRS above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPFGMRSetMaxRestarts for complete further documentation of this rou-

tine.

#### 9.12.4 SUNLinearSolver\_SPFGMR content

The SUNLINSOL\_SPFGMR module defines the *content* field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_SPFGMR {
  int maxl;
  int pretype;
  int gstype;
  int max_restarts;
  int numiters;
```

```
realtype resnorm;
  long int last_flag;
  ATimesFn ATimes;
  void* ATData;
  PSetupFn Psetup;
  PSolveFn Psolve;
  void* PData;
  N_Vector s1;
  N_Vector s2;
  N_Vector *V;
  N_Vector *Z;
  realtype **Hes;
  realtype *givens;
  N_Vector xcor;
  realtype *yg;
  N_Vector vtemp;
};
These entries of the content field contain the following information:
               - number of FGMRES basis vectors to use (default is 5),
maxl
               - flag for type of preconditioning to employ (default is none),
pretype
               - flag for type of Gram-Schmidt orthogonalization (default is modified Gram-Schmidt),
gstype
max_restarts - number of FGMRES restarts to allow (default is 0),
               - number of iterations from the most-recent solve,
numiters
               - final linear residual norm from the most-recent solve,
resnorm
last_flag
               - last error return flag from an internal function,
ATimes
               - function pointer to perform Av product,
ATData
               - pointer to structure for ATimes,
               - function pointer to preconditioner setup routine,
Psetup
Psolve
               - function pointer to preconditioner solve routine,
               - pointer to structure for Psetup and Psolve,
PData
               - vector pointers for supplied scaling matrices (default is NULL),
s1, s2
V
               - the array of Krylov basis vectors v_1, \ldots, v_{\mathtt{max1}+1}, stored in V[0], \ldots, V[\mathtt{max1}]. Each
               v_i is a vector of type NVECTOR.,
Z
               - the array of preconditioned Krylov basis vectors z_1, \ldots, z_{\max 1+1}, stored in Z[0], \ldots
               Z[max1]. Each z_i is a vector of type NVECTOR.,
               - the (\max 1 + 1) \times \max 1 Hessenberg matrix. It is stored row-wise so that the (i,j)th
Hes
               element is given by Hes[i][j].,
               - a length 2*maxl array which represents the Givens rotation matrices that arise in the
givens
               FGMRES algorithm. These matrices are F_0, F_1, \ldots, F_i, where
```

$$F_i = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & c_i & -s_i & & & & \\ & & & c_i & & & & & \\ & & & s_i & c_i & & & & \\ & & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 \end{bmatrix},$$

```
are represented in the givens vector as givens [0] = c_0, givens [1] = s_0, givens [2] = c_1, givens [3] = s_1, ... givens [2j] = c_j, givens [2j+1] = s_j.,

xcor

- a vector which holds the scaled, preconditioned correction to the initial guess,

yg

- a length (maxl+1) array of realtype values used to hold "short" vectors (e.g. y and g),

vtemp

- temporary vector storage.
```

# 9.13 The SUNLinearSolver\_SPBCGS implementation

This section describes the SUNLINSOL implementation of the SPBCGS (Scaled, Preconditioned, Bi-Conjugate Gradient, Stabilized [44]) iterative linear solver. The SUNLINSOL\_SPBCGS module is designed to be compatible with any NVECTOR implementation that supports a minimal subset of operations (N\_VClone, N\_VDotProd, N\_VScale, N\_VLinearSum, N\_VProd, N\_VDiv, and N\_VDestroy). Unlike the SPGMR and SPFGMR algorithms, SPBCGS requires a fixed amount of memory that does not increase with the number of allowed iterations.

To access the SUNLINSOL\_SPBCGS module, include the header file sunlinsol/sunlinsol\_spbcgs.h. We note that the SUNLINSOL\_SPBCGS module is accessible from SUNDIALS packages without separately linking to the libsundials\_sunlinsolspbcgs module library.

# 9.13.1 SUNLinearSolver\_SPBCGS description

This solver is constructed to perform the following operations:

- During construction all NVECTOR solver data is allocated, with vectors cloned from a template NVECTOR that is input, and default solver parameters are set.
- User-facing "set" routines may be called to modify default solver parameters.
- Additional "set" routines are called by the SUNDIALS solver that interfaces with SUNLINSOL\_SPBCGS to supply the ATimes, PSetup, and Psolve function pointers and s1 and s2 scaling vectors.
- In the "initialize" call, the solver parameters are checked for validity.
- In the "setup" call, any non-NULL PSetup function is called. Typically, this is provided by the SUNDIALS solver itself, that translates between the generic PSetup function and the solver-specific routine (solver-supplied or user-supplied).
- In the "solve" call the SPBCGS iteration is performed. This will include scaling and preconditioning if those options have been supplied.

#### 9.13.2 SUNLinearSolver\_SPBCGS functions

The SUNLINSOL\_SPBCGS module provides the following user-callable constructor for creating a SUNLinearSolver object.

```
Call LS = SUNLinSol_SPBCGS(y, pretype, maxl);

Description The function SUNLinSol_SPBCGS creates and allocates memory for a SPBCGS SUNLinearSolver object.

Arguments y (N_Vector) a template for cloning vectors needed within the solver pretype (int) flag indicating the desired type of preconditioning, allowed values are:
```

• PREC\_NONE (0)

- PREC\_LEFT (1)
- PREC\_RIGHT (2)
- PREC\_BOTH (3)

Any other integer input will result in the default (no preconditioning).

maxl

(int) the number of linear iterations to allow. Values  $\leq 0$  will result in the default value (5).

Return value

This returns a SUNLinearSolver object. If either y is incompatible then this routine will return NULL.

Notes

This routine will perform consistency checks to ensure that it is called with a consistent NVECTOR implementation (i.e. that it supplies the requisite vector operations). If y is incompatible, then this routine will return NULL.

We note that some SUNDIALS solvers are designed to only work with left preconditioning (IDA and IDAS) and others with only right preconditioning (KINSOL). While it is possible to configure a SUNLINSOL\_SPBCGS object to use any of the preconditioning options with these solvers, this use mode is not supported and may result in inferior performance.

Deprecated Name For backward compatibility, the wrapper function SUNSPBCGS with idential input and output arguments is also provided.

F2003 Name

This function is callable as FSUNLinSol\_SPBCGS when using the Fortran 2003 interface module.

The SUNLINSOL\_SPBCGS module defines implementations of all "iterative" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_SPBCGS
- SUNLinSolInitialize\_SPBCGS
- SUNLinSolSetATimes\_SPBCGS
- SUNLinSolSetPreconditioner\_SPBCGS
- SUNLinSolSetScalingVectors\_SPBCGS
- SUNLinSolSetup\_SPBCGS
- SUNLinSolSolve SPBCGS
- SUNLinSolNumIters\_SPBCGS
- SUNLinSolResNorm\_SPBCGS
- SUNLinSolResid\_SPBCGS
- SUNLinSolLastFlag\_SPBCGS
- SUNLinSolSpace\_SPBCGS
- SUNLinSolFree\_SPBCGS

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

The SUNLINSOL\_SPBCGS module also defines the following additional user-callable functions.

# SUNLinSol\_SPBCGSSetPrecType

Call retval = SUNLinSol\_SPBCGSSetPrecType(LS, pretype);

Description The function SUNLinSol\_SPBCGSSetPrecType updates the type of preconditioning

to use in the SUNLINSOL\_SPBCGS object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPBCGS object to update

pretype (int) flag indicating the desired type of preconditioning, allowed values

match those discussed in SUNLinSol\_SPBCGS.

Return value This routine will return with one of the error codes SUNLS\_ILL\_INPUT (illegal

pretype), SUNLS\_MEM\_NULL (S is NULL) or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPBCGSSetPrecType with

idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPBCGSSetPrecType when using the For-

tran 2003 interface module.

#### SUNLinSol\_SPBCGSSetMax1

Call retval = SUNLinSol\_SPBCGSSetMax1(LS, max1);

Description The function SUNLinSol\_SPBCGSSetMaxl updates the number of linear solver iter-

ations to allow.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPBCGS object to update

max1 (int) flag indicating the number of iterations to allow. Values  $\leq 0$  will result

in the default value (5).

Return value This routine will return with one of the error codes SUNLS\_MEM\_NULL (S is NULL) or

SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPBCGSSetMaxl with idential

input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPBCGSSetMaxl when using the Fortran

2003 interface module.

### 9.13.3 SUNLinearSolver\_SPBCGS Fortran interfaces

The SUNLINSOL\_SPBCGS module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

#### FORTRAN 2003 interface module

The fsunlinsol\_spbcgs\_mod FORTRAN module defines interfaces to all SUNLINSOL\_SPBCGS C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_SPBCGS is interfaced as FSUNLinSol\_SPBCGS.

The FORTRAN 2003 SUNLINSOL\_SPBCGS interface module can be accessed with the use statement, i.e. use fsunlinsol\_spbcgs\_mod, and linking to the library libsundials\_fsunlinsolspbcgs\_mod. lib in addition to the C library. For details on where the library and module file

fsunlinsol\_spbcgs\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunlinsolspbcgs\_mod library.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_SPBCGS module also includes a Fortran-callable function for creating a SUNLinearSolver object.

#### FSUNSPBCGSINIT

Call FSUNSPBCGSINIT(code, pretype, maxl, ier)

Description The function FSUNSPBCGSINIT can be called for Fortran programs to create a SUNLIN-

SOL\_SPBCGS object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating desired preconditioning type
maxl (int\*) flag indicating number of iterations to allow

Return value  $\mathtt{ier}$  is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

SUNLinSol\_SPBCGS.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_SPBCGS module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

#### FSUNMASSSPBCGSINIT

Call FSUNMASSSPBCGSINIT(pretype, maxl, ier)

Description The function FSUNMASSSPBCGSINIT can be called for Fortran programs to create a SUN-

LINSOL\_SPBCGS object for mass matrix linear systems.

Arguments pretype (int\*) flag indicating desired preconditioning type

maxl (int\*) flag indicating number of iterations to allow

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

SUNLinSol\_SPBCGS.

The SUNLinSol\_SPBCGSSetPrecType and SUNLinSol\_SPBCGSSetMaxl routines also support Fortran interfaces for the system and mass matrix solvers.

#### FSUNSPBCGSSETPRECTYPE

Call FSUNSPBCGSSETPRECTYPE(code, pretype, ier)

Description The function FSUNSPBCGSSETPRECTYPE can be called for Fortran programs to change

the type of preconditioning to use.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating the type of preconditioning to use.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPBCGSSetPrecType for complete further documentation of this routine.

#### FSUNMASSSPBCGSSETPRECTYPE

Call FSUNMASSSPBCGSSETPRECTYPE(pretype, ier)

Description The function FSUNMASSSPBCGSSETPRECTYPE can be called for Fortran programs to change

the type of preconditioning for mass matrix linear systems.

 $\label{lem:arguments} \textbf{Arguments} \quad \text{The arguments are identical to } \textbf{FSUNSPBCGSSETPRECTYPE} \ above, \ except \ that \ \textbf{code} \ is \ not$ 

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPBCGSSetPrecType for complete further documentation of this routine.

#### FSUNSPBCGSSETMAXL

Call FSUNSPBCGSSETMAXL(code, maxl, ier)

Description The function FSUNSPBCGSSETMAXL can be called for Fortran programs to change the

maximum number of iterations to allow.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

maxl (int\*) the number of iterations to allow.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPBCGSSetMaxl for complete further documentation of this routine.

#### FSUNMASSSPBCGSSETMAXL

Call FSUNMASSSPBCGSSETMAXL(maxl, ier)

Description The function FSUNMASSSPBCGSSETMAXL can be called for Fortran programs to change

the type of preconditioning for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPBCGSSETMAXL above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPBCGSSetMaxl for complete further documentation of this routine.

#### 9.13.4 SUNLinearSolver\_SPBCGS content

The SUNLINSOL\_SPBCGS module defines the *content* field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_SPBCGS {
  int maxl;
  int pretype;
  int numiters;
  realtype resnorm;
  long int last_flag;
  ATimesFn ATimes;
  void* ATData;
  PSetupFn Psetup;
  PSolveFn Psolve;
  void* PData;
  N_Vector s1;
  N_Vector s2;
  N_Vector r;
  N_Vector r_star;
  N_Vector p;
  N_Vector q;
  N_Vector u;
```

```
N_Vector Ap;
  N_Vector vtemp;
These entries of the content field contain the following information:
           - number of SPBCGS iterations to allow (default is 5),
maxl
           - flag for type of preconditioning to employ (default is none),
pretype
numiters - number of iterations from the most-recent solve,
resnorm
           - final linear residual norm from the most-recent solve.
last_flag - last error return flag from an internal function,
           - function pointer to perform Av product,
ATimes
ATData
           - pointer to structure for ATimes,
Psetup
           - function pointer to preconditioner setup routine,
Psolve
           - function pointer to preconditioner solve routine,
PData
           - pointer to structure for Psetup and Psolve,
           - vector pointers for supplied scaling matrices (default is NULL),
s1, s2
           - a NVECTOR which holds the current scaled, preconditioned linear system residual,
           - a NVECTOR which holds the initial scaled, preconditioned linear system residual,
r_star
p, q, u, Ap, vtemp - NVECTORS used for workspace by the SPBCGS algorithm.
```

# 9.14 The SUNLinearSolver\_SPTFQMR implementation

This section describes the SUNLINSOL implementation of the SPTFQMR (Scaled, Preconditioned, Transpose-Free Quasi-Minimum Residual [23]) iterative linear solver. The SUNLINSOL\_SPTFQMR module is designed to be compatible with any NVECTOR implementation that supports a minimal subset of operations (N\_VClone, N\_VDotProd, N\_VScale, N\_VLinearSum, N\_VProd, N\_VConst, N\_VDiv, and N\_VDestroy). Unlike the SPGMR and SPFGMR algorithms, SPTFQMR requires a fixed amount of memory that does not increase with the number of allowed iterations.

To access the SUNLINSOL\_SPTFQMR module, include the header file sunlinsol/sunlinsol\_sptfqmr.h. We note that the SUNLINSOL\_SPTFQMR module is accessible from SUNDIALS packages without separately linking to the libsundials\_sunlinsolsptfqmr module library.

# 9.14.1 SUNLinearSolver\_SPTFQMR description

This solver is constructed to perform the following operations:

- During construction all NVECTOR solver data is allocated, with vectors cloned from a template NVECTOR that is input, and default solver parameters are set.
- User-facing "set" routines may be called to modify default solver parameters.
- Additional "set" routines are called by the SUNDIALS solver that interfaces with SUNLINSOL\_SPTFQMR to supply the ATimes, PSetup, and Psolve function pointers and s1 and s2 scaling vectors.
- In the "initialize" call, the solver parameters are checked for validity.
- In the "setup" call, any non-NULL PSetup function is called. Typically, this is provided by the SUNDIALS solver itself, that translates between the generic PSetup function and the solver-specific routine (solver-supplied or user-supplied).
- In the "solve" call the TFQMR iteration is performed. This will include scaling and preconditioning if those options have been supplied.

Notes

# 9.14.2 SUNLinearSolver\_SPTFQMR functions

The SUNLINSOL\_SPTFQMR module provides the following user-callable constructor for creating a SUNLinearSolver object.

#### SUNLinSol\_SPTFQMR

Call LS = SUNLinSol\_SPTFQMR(y, pretype, maxl);

Description The function SUNLinSol\_SPTFQMR creates and allocates memory for a SPTFQMR

SUNLinearSolver object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

 ${\tt pretype}$  (int) flag indicating the desired type of preconditioning, allowed values

• PREC\_NONE (0)

• PREC\_LEFT (1)

• PREC\_RIGHT (2)

• PREC\_BOTH (3)

Any other integer input will result in the default (no preconditioning).

maxl (int) the number of linear iterations to allow. Values  $\leq 0$  will result in the default value (5).

Return value This returns a SUNLinearSolver object. If either y is incompatible then this

routine will return NULL.

This routine will perform consistency checks to ensure that it is called with a consistent NVECTOR implementation (i.e. that it supplies the requisite vector operations). If y is incompatible, then this routine will return NULL.

We note that some SUNDIALS solvers are designed to only work with left preconditioning (IDA and IDAS) and others with only right preconditioning (KINSOL). While it is possible to configure a SUNLINSOL\_SPTFQMR object to use any of the preconditioning options with these solvers, this use mode is not supported and may result in inferior performance.

Deprecated Name For backward compatibility, the wrapper function SUNSPTFQMR with idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPTFQMR when using the Fortran 2003 interface module.

The SUNLINSOL\_SPTFQMR module defines implementations of all "iterative" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_SPTFQMR
- SUNLinSolInitialize\_SPTFQMR
- SUNLinSolSetATimes\_SPTFQMR
- SUNLinSolSetPreconditioner\_SPTFQMR
- SUNLinSolSetScalingVectors\_SPTFQMR
- SUNLinSolSetup\_SPTFQMR
- SUNLinSolSolve\_SPTFQMR
- SUNLinSolNumIters\_SPTFQMR
- SUNLinSolResNorm\_SPTFQMR

- SUNLinSolResid\_SPTFQMR
- SUNLinSolLastFlag\_SPTFQMR
- SUNLinSolSpace\_SPTFQMR
- SUNLinSolFree\_SPTFQMR

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

The SUNLINSOL\_SPTFQMR module also defines the following additional user-callable functions.

# SUNLinSol\_SPTFQMRSetPrecType

Call retval = SUNLinSol\_SPTFQMRSetPrecType(LS, pretype);

Description The function SUNLinSol\_SPTFQMRSetPrecType updates the type of preconditioning

to use in the SUNLINSOL\_SPTFQMR object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPTFQMR object to update

pretype (int) flag indicating the desired type of preconditioning, allowed values

match those discussed in SUNLinSol\_SPTFQMR.

Return value This routine will return with one of the error codes SUNLS\_ILL\_INPUT (illegal

pretype), SUNLS\_MEM\_NULL (S is NULL) or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNSPTFQMRSetPrecType with

idential input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_SPTFQMRSetPrecType when using the For-

tran 2003 interface module.

# ${\tt SUNLinSol\_SPTFQMRSetMaxl}$

Call retval = SUNLinSol\_SPTFQMRSetMaxl(LS, maxl);

Description The function SUNLinSol\_SPTFQMRSetMaxl updates the number of linear solver iterations

to allow.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_SPTFQMR object to update

max1 (int) flag indicating the number of iterations to allow; values < 0 will result in

the default value (5)

Return value This routine will return with one of the error codes SUNLS\_MEM\_NULL (S is NULL) or

SUNLS\_SUCCESS.

F2003 Name This function is callable as FSUNLinSol\_SPTFQMRSetMaxl when using the Fortran 2003

interface module.

 ${\bf SUNSPTFQMRSetMaxl}$ 

#### 9.14.3 SUNLinearSolver\_SPTFQMR Fortran interfaces

The SUNLINSOL\_SPFGMR module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

#### FORTRAN 2003 interface module

The fsunlinsol\_sptfqmr\_mod FORTRAN module defines interfaces to all SUNLINSOL\_SPFGMR C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_SPTFQMR is interfaced as FSUNLinSol\_SPTFQMR.

The Fortran 2003 sunlinsol\_spfgmr interface module can be accessed with the use statement, i.e. use fsunlinsol\_sptfqmr\_mod, and linking to the library libsundials\_fsunlinsolsptfqmr\_mod. lib in addition to the C library. For details on where the library and module file

fsunlinsol\_sptfqmr\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunlinsolsptfqmr\_mod library.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_SPTFQMR module also includes a Fortran-callable function for creating a SUNLinearSolver object.

#### FSUNSPTFQMRINIT

Call FSUNSPTFQMRINIT(code, pretype, maxl, ier)

Description The function FSUNSPTFQMRINIT can be called for Fortran programs to create a SUNLIN-

SOL\_SPTFQMR object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating desired preconditioning type
maxl (int\*) flag indicating number of iterations to allow

Return value  $\mathtt{ier}$  is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

SUNLinSol\_SPTFQMR.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_SPTFQMR module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

# FSUNMASSSPTFQMRINIT

Call FSUNMASSSPTFQMRINIT(pretype, maxl, ier)

Description The function FSUNMASSSPTFQMRINIT can be called for Fortran programs to create a

SUNLINSOL\_SPTFQMR object for mass matrix linear systems.

Arguments pretype (int\*) flag indicating desired preconditioning type

maxl (int\*) flag indicating number of iterations to allow

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function

SUNLinSol\_SPTFQMR.

The SUNLinSol\_SPTFQMRSetPrecType and SUNLinSol\_SPTFQMRSetMaxl routines also support Fortran interfaces for the system and mass matrix solvers.

#### FSUNSPTFQMRSETPRECTYPE

Call FSUNSPTFQMRSETPRECTYPE(code, pretype, ier)

Description The function FSUNSPTFQMRSETPRECTYPE can be called for Fortran programs to change

the type of preconditioning to use.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating the type of preconditioning to use.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPTFQMRSetPrecType for complete further documentation of this rou-

tine.

#### FSUNMASSSPTFQMRSETPRECTYPE

Call FSUNMASSSPTFQMRSETPRECTYPE(pretype, ier)

Description The function FSUNMASSSPTFQMRSETPRECTYPE can be called for Fortran programs to

change the type of preconditioning for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPTFQMRSETPRECTYPE above, except that code is

not needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPTFQMRSetPrecType for complete further documentation of this rou-

tine.

#### FSUNSPTFQMRSETMAXL

Call FSUNSPTFQMRSETMAXL(code, maxl, ier)

Description The function FSUNSPTFQMRSETMAXL can be called for Fortran programs to change the

maximum number of iterations to allow.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

maxl (int\*) the number of iterations to allow.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPTFQMRSetMaxl for complete further documentation of this routine.

#### FSUNMASSSPTFQMRSETMAXL

Call FSUNMASSSPTFQMRSETMAXL(maxl, ier)

Description The function FSUNMASSSPTFQMRSETMAXL can be called for Fortran programs to change

the type of preconditioning for mass matrix linear systems.

Arguments The arguments are identical to FSUNSPTFQMRSETMAXL above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_SPTFQMRSetMax1 for complete further documentation of this routine.

# 9.14.4 SUNLinearSolver\_SPTFQMR content

The SUNLINSOL\_SPTFQMR module defines the *content* field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_SPTFQMR {
  int maxl;
  int pretype;
  int numiters;
  realtype resnorm;
```

```
long int last_flag;
  ATimesFn ATimes;
  void* ATData;
  PSetupFn Psetup;
  PSolveFn Psolve;
  void* PData;
  N_Vector s1;
  N_Vector s2;
  N_Vector r_star;
  N_Vector q;
  N_Vector d;
  N_Vector v;
  N_Vector p;
  N_Vector *r;
  N_Vector u;
  N_Vector vtemp1;
  N_Vector vtemp2;
  N_Vector vtemp3;
};
These entries of the content field contain the following information:
           - number of TFQMR iterations to allow (default is 5),
           - flag for type of preconditioning to employ (default is none),
pretype
numiters - number of iterations from the most-recent solve,
resnorm
          - final linear residual norm from the most-recent solve,
last_flag - last error return flag from an internal function,
           - function pointer to perform Av product,
ATimes
ATData
           - pointer to structure for ATimes,
           - function pointer to preconditioner setup routine,
Psetup
          - function pointer to preconditioner solve routine,
Psolve
PData
          - pointer to structure for Psetup and Psolve,
s1, s2
          - vector pointers for supplied scaling matrices (default is NULL),
          - a NVECTOR which holds the initial scaled, preconditioned linear system residual,
r star
q, d, v, p, u - NVECTORS used for workspace by the SPTFQMR algorithm,
           - array of two NVECTORS used for workspace within the SPTFQMR algorithm,
vtemp1, vtemp2, vtemp3 - temporary vector storage.
```

# 9.15 The SUNLinearSolver\_PCG implementation

This section describes the SUNLINSOL implementation of the PCG (Preconditioned Conjugate Gradient [24]) iterative linear solver. The SUNLINSOL\_PCG module is designed to be compatible with any NVECTOR implementation that supports a minimal subset of operations (N\_VClone, N\_VDotProd, N\_VScale, N\_VLinearSum, N\_VProd, and N\_VDestroy). Unlike the SPGMR and SPFGMR algorithms, PCG requires a fixed amount of memory that does not increase with the number of allowed iterations.

To access the SUNLINSOL\_PCG module, include the header file sunlinsol/sunlinsol\_pcg.h. We note that the SUNLINSOL\_PCG module is accessible from SUNDIALS packages without separately linking to the libsundials\_sunlinsolpcg module library.

# 9.15.1 SUNLinearSolver\_PCG description

Unlike all of the other iterative linear solvers supplied with SUNDIALS, PCG should only be used on symmetric linear systems (e.g. mass matrix linear systems encountered in ARKODE). As a result, the explanation of the role of scaling and preconditioning matrices given in general must be modified in this scenario. The PCG algorithm solves a linear system Ax = b where A is a symmetric ( $A^T = A$ ), real-valued matrix. Preconditioning is allowed, and is applied in a symmetric fashion on both the right and left. Scaling is also allowed and is applied symmetrically. We denote the preconditioner and scaling matrices as follows:

- P is the preconditioner (assumed symmetric),
- S is a diagonal matrix of scale factors.

The matrices A and P are not required explicitly; only routines that provide A and  $P^{-1}$  as operators are required. The diagonal of the matrix S is held in a single NVECTOR, supplied by the user.

In this notation, PCG applies the underlying CG algorithm to the equivalent transformed system

$$\tilde{A}\tilde{x} = \tilde{b} \tag{9.4}$$

where

$$\tilde{A} = SP^{-1}AP^{-1}S,$$

$$\tilde{b} = SP^{-1}b,$$

$$\tilde{x} = S^{-1}Px.$$

$$(9.5)$$

The scaling matrix must be chosen so that the vectors  $SP^{-1}b$  and  $S^{-1}Px$  have dimensionless components.

The stopping test for the PCG iterations is on the L2 norm of the scaled preconditioned residual:

$$\begin{split} & \|\tilde{b} - \tilde{A}\tilde{x}\|_2 < \delta \\ \Leftrightarrow & \\ & \|SP^{-1}b - SP^{-1}Ax\|_2 < \delta \\ \Leftrightarrow & \\ & \|P^{-1}b - P^{-1}Ax\|_S < \delta \end{split}$$

where  $||v||_S = \sqrt{v^T S^T S v}$ , with an input tolerance  $\delta$ .

This solver is constructed to perform the following operations:

- During construction all NVECTOR solver data is allocated, with vectors cloned from a template NVECTOR that is input, and default solver parameters are set.
- User-facing "set" routines may be called to modify default solver parameters.
- Additional "set" routines are called by the SUNDIALS solver that interfaces with SUNLINSOL\_PCG to supply the ATimes, PSetup, and Psolve function pointers and s scaling vector.
- In the "initialize" call, the solver parameters are checked for validity.
- In the "setup" call, any non-NULL PSetup function is called. Typically, this is provided by the SUNDIALS solver itself, that translates between the generic PSetup function and the solver-specific routine (solver-supplied or user-supplied).
- In the "solve" call the PCG iteration is performed. This will include scaling and preconditioning if those options have been supplied.

#### 9.15.2 SUNLinearSolver\_PCG functions

The SUNLINSOL\_PCG module provides the following user-callable constructor for creating a SUNLinearSolver object.

SUNLinSol\_PCG

Call LS = SUNLinSol\_PCG(y, pretype, maxl);

Description The function SUNLinSol\_PCG creates and allocates memory for a PCG SUNLinearSolver

object.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver

pretype (int) flag indicating whether to use preconditioning. Since the PCG al-

gorithm is designed to only support symmetric preconditioning, then any of the pretype inputs PREC\_LEFT (1), PREC\_RIGHT (2), or PREC\_BOTH (3) will result in use of the symmetric preconditioner; any other integer input

will result in the default (no preconditioning).

maxl (int) the number of linear iterations to allow; values  $\leq 0$  will result in

the default value (5).

Return value This returns a SUNLinearSolver object. If either y is incompatible then this

routine will return NULL.

Notes This routine will perform consistency checks to ensure that it is called with a consistent NVECTOR implementation (i.e. that it supplies the requisite vector operations).

If y is incompatible, then this routine will return NULL.

Although some SUNDIALS solvers are designed to only work with left preconditioning (IDA and IDAS) and others with only right preconditioning (KINSOL), PCG should *only* be used with these packages when the linear systems are known to be *symmetric*. Since the scaling of matrix rows and columns must be identical in a symmetric matrix, symmetric preconditioning should work appropriately even for

packages designed with one-sided preconditioning in mind.

output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_PCG when using the Fortran 2003 interface

module

The SUNLINSOL\_PCG module defines implementations of all "iterative" linear solver operations listed in Sections 9.1.1 - 9.1.3:

- SUNLinSolGetType\_PCG
- SUNLinSolInitialize\_PCG
- SUNLinSolSetATimes\_PCG
- SUNLinSolSetPreconditioner\_PCG
- SUNLinSolSetScalingVectors\_PCG since PCG only supports symmetric scaling, the second NVECTOR argument to this function is ignored
- SUNLinSolSetup\_PCG
- SUNLinSolSolve\_PCG
- SUNLinSolNumIters\_PCG
- SUNLinSolResNorm\_PCG
- SUNLinSolResid\_PCG

- SUNLinSolLastFlag\_PCG
- SUNLinSolSpace\_PCG
- SUNLinSolFree\_PCG

All of the listed operations are callable via the FORTRAN 2003 interface module by prepending an 'F' to the function name.

The SUNLINSOL\_PCG module also defines the following additional user-callable functions.

### SUNLinSol\_PCGSetPrecType

Call retval = SUNLinSol\_PCGSetPrecType(LS, pretype);

Description The function SUNLinSol\_PCGSetPrecType updates the flag indicating use of pre-

conditioning in the SUNLINSOL\_PCG object.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_PCG object to update

pretype (int) flag indicating use of preconditioning, allowed values match those

discussed in SUNLinSol\_PCG.

Return value This routine will return with one of the error codes SUNLS\_ILL\_INPUT (illegal

pretype), SUNLS\_MEM\_NULL (S is NULL) or SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNPCGSetPrecType with iden-

tial input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_PCGSetPrecType when using the Fortran

2003 interface module.

# ${\tt SUNLinSol\_PCGSetMaxl}$

Call retval = SUNLinSol\_PCGSetMaxl(LS, maxl);

Description The function SUNLinSol\_PCGSetMaxl updates the number of linear solver iterations

to allow.

Arguments LS (SUNLinearSolver) the SUNLINSOL\_PCG object to update

maxl (int) flag indicating the number of iterations to allow; values  $\leq 0$  will result

in the default value (5)

Return value This routine will return with one of the error codes SUNLS\_MEM\_NULL (S is NULL) or

SUNLS\_SUCCESS.

Deprecated Name For backward compatibility, the wrapper function SUNPCGSetMaxl with idential

input and output arguments is also provided.

F2003 Name This function is callable as FSUNLinSol\_PCGSetMaxl when using the Fortran 2003

interface module.

#### 9.15.3 SUNLinearSolver\_PCG Fortran interfaces

The SUNLINSOL\_PCG module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

#### FORTRAN 2003 interface module

The fsunlinsol\_pcg\_mod FORTRAN module defines interfaces to all SUNLINSOL\_PCG C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNLinSol\_PCG is interfaced as FSUNLinSol\_PCG.

The FORTRAN 2003 SUNLINSOL\_PCG interface module can be accessed with the use statement, i.e. use fsunlinsol\_pcg\_mod, and linking to the library libsundials\_fsunlinsolpcg\_mod.lib in addition to the C library. For details on where the library and module file fsunlinsol\_pcg\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunlinsolpcg\_mod library.

#### FORTRAN 77 interface functions

For solvers that include a FORTRAN 77 interface module, the SUNLINSOL\_PCG module also includes a Fortran-callable function for creating a SUNLinearSolver object.

#### FSUNPCGINIT

Call FSUNPCGINIT(code, pretype, maxl, ier)

Description The function FSUNPCGINIT can be called for Fortran programs to create a SUNLIN-

SOL\_PCG object.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating desired preconditioning type
maxl (int\*) flag indicating number of iterations to allow

Return value  $\mathtt{ier}$  is a return completion flag equal to 0 for a success return and -1 otherwise. See

printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and max1 are the same as for the C function SUNLinSol\_PCG.

Additionally, when using ARKODE with a non-identity mass matrix, the SUNLINSOL\_PCG module includes a Fortran-callable function for creating a SUNLinearSolver mass matrix solver object.

#### FSUNMASSPCGINIT

Call FSUNMASSPCGINIT(pretype, maxl, ier)

Description The function FSUNMASSPCGINIT can be called for Fortran programs to create a SUNLIN-

SOL\_PCG object for mass matrix linear systems.

Arguments pretype (int\*) flag indicating desired preconditioning type

maxl (int\*) flag indicating number of iterations to allow

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes This routine must be called *after* the NVECTOR object has been initialized.

Allowable values for pretype and maxl are the same as for the C function SUNLinSol\_PCG.

The SUNLinSol\_PCGSetPrecType and SUNLinSol\_PCGSetMaxl routines also support Fortran interfaces for the system and mass matrix solvers.

#### FSUNPCGSETPRECTYPE

Call FSUNPCGSETPRECTYPE(code, pretype, ier)

Description The function FSUNPCGSETPRECTYPE can be called for Fortran programs to change the

type of preconditioning to use.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3

for KINSOL, and 4 for ARKODE).

pretype (int\*) flag indicating the type of preconditioning to use.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_PCGSetPrecType for complete further documentation of this routine.

#### FSUNMASSPCGSETPRECTYPE

Call FSUNMASSPCGSETPRECTYPE(pretype, ier)

Description The function FSUNMASSPCGSETPRECTYPE can be called for Fortran programs to change

the type of preconditioning for mass matrix linear systems.

Arguments The arguments are identical to FSUNPCGSETPRECTYPE above, except that code is not

needed since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_PCGSetPrecType for complete further documentation of this routine.

#### FSUNPCGSETMAXL

Call FSUNPCGSETMAXL(code, maxl, ier)

Description The function FSUNPCGSETMAXL can be called for Fortran programs to change the maxi-

mum number of iterations to allow.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, 3 for

KINSOL, and 4 for ARKODE).

maxl (int\*) the number of iterations to allow.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_PCGSetMaxl for complete further documentation of this routine.

#### FSUNMASSPCGSETMAXL

Call FSUNMASSPCGSETMAXL(maxl, ier)

Description The function FSUNMASSPCGSETMAXL can be called for Fortran programs to change the

type of preconditioning for mass matrix linear systems.

Arguments The arguments are identical to FSUNPCGSETMAXL above, except that code is not needed

since mass matrix linear systems only arise in ARKODE.

Return value ier is a int return completion flag equal to 0 for a success return and -1 otherwise.

See printed message for details in case of failure.

Notes See SUNLinSol\_PCGSetMaxl for complete further documentation of this routine.

#### 9.15.4 SUNLinearSolver\_PCG content

The SUNLINSOL\_PCG module defines the *content* field of a SUNLinearSolver as the following structure:

```
struct _SUNLinearSolverContent_PCG {
  int maxl;
  int pretype;
  int numiters;
  realtype resnorm;
  long int last_flag;
  ATimesFn ATimes;
  void* ATData;
  PSetupFn Psetup;
  PSolveFn Psolve;
  void* PData;
```

```
N_Vector s;
  N_Vector r;
  N_Vector p;
  N_Vector z;
  N_Vector Ap;
};
These entries of the content field contain the following information:
           - number of PCG iterations to allow (default is 5),
maxl
           - flag for use of preconditioning (default is none),
pretype
numiters - number of iterations from the most-recent solve,
resnorm
           - final linear residual norm from the most-recent solve,
last_flag - last error return flag from an internal function,
           - function pointer to perform Av product,
ATimes
ATData
           - pointer to structure for ATimes,
           - function pointer to preconditioner setup routine,
Psetup
Psolve
           - function pointer to preconditioner solve routine,
           - pointer to structure for Psetup and Psolve,
PData
           - vector pointer for supplied scaling matrix (default is NULL),
           - a NVECTOR which holds the preconditioned linear system residual,
p, z, Ap - NVECTORS used for workspace by the PCG algorithm.
```

# 9.16 SUNLinearSolver Examples

There are SUNLinearSolver examples that may be installed for each implementation; these make use of the functions in test\_sunlinsol.c. These example functions show simple usage of the SUNLinearSolver family of functions. The inputs to the examples depend on the linear solver type, and are output to stdout if the example is run without the appropriate number of command-line arguments.

The following is a list of the example functions in test\_sunlinsol.c:

- Test\_SUNLinSolGetType: Verifies the returned solver type against the value that should be returned.
- Test\_SUNLinSolInitialize: Verifies that SUNLinSolInitialize can be called and returns successfully.
- Test\_SUNLinSolSetup: Verifies that SUNLinSolSetup can be called and returns successfully.
- Test\_SUNLinSolSolve: Given a SUNMATRIX object A, NVECTOR objects x and b (where Ax = b) and a desired solution tolerance tol, this routine clones x into a new vector y, calls SUNLinSolSolve to fill y as the solution to Ay = b (to the input tolerance), verifies that each entry in x and y match to within 10\*tol, and overwrites x with y prior to returning (in case the calling routine would like to investigate further).
- Test\_SUNLinSolSetATimes (iterative solvers only): Verifies that SUNLinSolSetATimes can be called and returns successfully.
- Test\_SUNLinSolSetPreconditioner (iterative solvers only): Verifies that SUNLinSolSetPreconditioner can be called and returns successfully.
- Test\_SUNLinSolSetScalingVectors (iterative solvers only): Verifies that SUNLinSolSetScalingVectors can be called and returns successfully.

- Test\_SUNLinSolLastFlag: Verifies that SUNLinSolLastFlag can be called, and outputs the result to stdout.
- Test\_SUNLinSolNumIters (iterative solvers only): Verifies that SUNLinSolNumIters can be called, and outputs the result to stdout.
- Test\_SUNLinSolResNorm (iterative solvers only): Verifies that SUNLinSolResNorm can be called, and that the result is non-negative.
- Test\_SUNLinSolResid (iterative solvers only): Verifies that SUNLinSolResid can be called.
- Test\_SUNLinSolSpace verifies that SUNLinSolSpace can be called, and outputs the results to stdout.

We'll note that these tests should be performed in a particular order. For either direct or iterative linear solvers, Test\_SUNLinSolInitialize must be called before Test\_SUNLinSolSolve, which must be called before Test\_SUNLinSolSolve. Additionally, for iterative linear solvers

 ${\tt Test\_SUNLinSolSetATimes}, {\tt Test\_SUNLinSolSetPreconditioner} \ {\rm and} \ \\$ 

Test\_SUNLinSolSetScalingVectors should be called before Test\_SUNLinSolInitialize; similarly Test\_SUNLinSolNumIters, Test\_SUNLinSolResNorm and Test\_SUNLinSolResid should be called after Test\_SUNLinSolSolve. These are called in the appropriate order in all of the example problems.

# Chapter 10

# Description of the SUNNonlinearSolver module

SUNDIALS time integration packages are written in terms of generic nonlinear solver operations defined by the SUNNONLINSOL API and implemented by a particular SUNNONLINSOL module of type SUNNonlinearSolver. Users can supply their own SUNNONLINSOL module, or use one of the modules provided with SUNDIALS.

The time integrators in SUNDIALS specify a default nonlinear solver module and as such this chapter is intended for users that wish to use a non-default nonlinear solver module or would like to provide their own nonlinear solver implementation. Users interested in using a non-default solver module may skip the description of the SUNNONLINSOL API in section 10.1 and proceeded to the subsequent sections in this chapter that describe the SUNNONLINSOL modules provided with SUNDIALS.

For users interested in providing their own Sunnonlinsol module, the following section presents the sunnonlinsol API and its implementation beginning with the definition of sunnonlinsol functions in sections 10.1.1 - 10.1.3. This is followed by the definition of functions supplied to a nonlinear solver implementation in section 10.1.4. A table of nonlinear solver return codes is given in section 10.1.5. The SunnonlinearSolver type and the generic sunnonlinsol module are defined in section 10.1.6. Section 10.1.7 describes how sunnonlinsol models interface with sundials integrators providing sensitivity analysis capabilities (CVODES and IDAS). Finally, section 10.1.8 lists the requirements for supplying a custom sunnonlinsol module. Users wishing to supply their own sunnonlinsol module are encouraged to use the sunnonlinsol implementations provided with sundials as a template for supplying custom nonlinear solver modules.

# 10.1 The SUNNonlinear Solver API

The SUNNONLINSOL API defines several nonlinear solver operations that enable SUNDIALS integrators to utilize any SUNNONLINSOL implementation that provides the required functions. These functions can be divided into three categories. The first are the core nonlinear solver functions. The second group of functions consists of set routines to supply the nonlinear solver with functions provided by the SUNDIALS time integrators and to modify solver parameters. The final group consists of get routines for retrieving nonlinear solver statistics. All of these functions are defined in the header file sundials/sundials\_nonlinearsolver.h.

#### 10.1.1 SUNNonlinearSolver core functions

The core nonlinear solver functions consist of two required functions to get the nonlinear solver type (SUNNonlinsSolGetType) and solve the nonlinear system (SUNNonlinSolSolve). The remaining three functions for nonlinear solver initialization (SUNNonlinSolInitialization), setup (SUNNonlinSolSetup), and destruction (SUNNonlinSolFree) are optional.

# SUNNonlinSolGetType

Call type = SUNNonlinSolGetType(NLS);

Description The required function SUNNonlinSolGetType returns nonlinear solver type.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

Return value The return value type (of type int) will be one of the following:

SUNNONLINEARSOLVER\_ROOTFIND 0, the SUNNONLINSOL module solves F(y) = 0. SUNNONLINEARSOLVER\_FIXEDPOINT 1, the SUNNONLINSOL module solves G(y) = y.

#### SUNNonlinSolInitialize

Call retval = SUNNonlinSolInitialize(NLS);

Description The optional function SUNNonlinSolInitialize performs nonlinear solver initialization

and may perform any necessary memory allocations.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

Return value The return value retval (of type int) is zero for a successful call and a negative value

for a failure.

Notes It is assumed all solver-specific options have been set prior to calling

SUNNonlinSolInitialize. SUNNONLINSOL implementations that do not require initial-

ization may set this operation to NULL.

#### SUNNonlinSolSetup

Call retval = SUNNonlinSolSetup(NLS, y, mem);

Description The optional function SUNNonlinSolSetup performs any solver setup needed for a non-

linear solve.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

y (N\_Vector) the initial iteration passed to the nonlinear solver.

mem (void \*) the SUNDIALS integrator memory structure.

Return value The return value retval (of type int) is zero for a successful call and a negative value

for a failure.

Notes SUNDIALS integrators call SUNonlinSolSetup before each step attempt. SUNNONLINSOL

implementations that do not require setup may set this operation to NULL.

#### SUNNonlinSolSolve

Call retval = SUNNonlinSolSolve(NLS, y0, y, w, tol, callLSetup, mem);

Description The required function SUNNonlinSolSolve solves the nonlinear system F(y) = 0 or

G(y) = y.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

y0 (N\_Vector) the initial iterate for the nonlinear solve. This must remain

unchanged throughout the solution process.

y (N\_Vector) the solution to the nonlinear system.

(N\_Vector) the solution error weight vector used for computing weighted

error norms.

tol (realtype) the requested solution tolerance in the weighted root-mean-

squared norm.

callLSetup (booleantype) a flag indicating that the integrator recommends for the

linear solver setup function to be called.

(void \*) the SUNDIALS integrator memory structure. mem

Return value The return value retval (of type int) is zero for a successul solve, a positive value for a recoverable error, and a negative value for an unrecoverable error.

#### SUNNonlinSolFree

Call retval = SUNNonlinSolFree(NLS);

Description The optional function SUNNonlinSolFree frees any memory allocated by the nonlinear

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure. SUNNONLINSOL implementations that do not allocate data may set

this operation to NULL.

#### 10.1.2SUNNonlinear Solver set functions

The following set functions are used to supply nonlinear solver modules with functions defined by the SUNDIALS integrators and to modify solver parameters. Only the routine for setting the nonlinear system defining function (SUNNonlinSolSetSysFn is required. All other set functions are optional.

#### SUNNonlinSolSetSysFn

Call retval = SUNNonlinSolSetSysFn(NLS, SysFn);

Description The required function SUNNonlinSolSetSysFn is used to provide the nonlinear solver

with the function defining the nonlinear system. This is the function F(y) in F(y) = 0

for SUNNONLINEARSOLVER\_ROOTFIND modules or G(y) in G(y) = y for

SUNNONLINEARSOLVER\_FIXEDPOINT modules.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

SysFn (SUNNonlinSolSysFn) the function defining the nonlinear system. See section

10.1.4 for the definition of SUNNonlinSolSysFn.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

#### SUNNonlinSolSetLSetupFn

Call retval = SUNNonlinSolSetLSetupFn(NLS, LSetupFn);

Description The optional function SUNNonlinSollSetupFn is called by SUNDIALS integrators to

provide the nonlinear solver with access to its linear solver setup function.

Arguments (SUNNonlinearSolver) a SUNNONLINSOL object.

LSetupFn (SUNNonlinSolLSetupFn) a wrapper function to the SUNDIALS integrator's

linear solver setup function. See section 10.1.4 for the definition of

SUNNonlinLSetupFn.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

Notes The SUNNonlinLSetupFn function sets up the linear system Ax = b where  $A = \frac{\partial F}{\partial y}$  is

the linearization of the nonlinear residual function F(y) = 0 (when using SUNLINSOL direct linear solvers) or calls the user-defined preconditioner setup function (when using SUNLINSOL iterative linear solvers). SUNNONLINSOL implementations that do not require solving this system, do not utilize SUNLINSOL linear solvers, or use SUNLINSOL linear solvers that do not require setup may set this operation to NULL.

#### SUNNonlinSolSetLSolveFn

Call retval = SUNNonlinSolSetLSolveFn(NLS, LSolveFn);

Description The optional function SUNNonlinSolSetLSolveFn is called by SUNDIALS integrators to

provide the nonlinear solver with access to its linear solver solve function.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object

LSolveFn (SUNNonlinSolLSolveFn) a wrapper function to the SUNDIALS integrator's

linear solver solve function. See section 10.1.4 for the definition of

SUNNonlinSolLSolveFn.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

Notes The SUNNonlinLSolveFn function solves the linear system Ax = b where  $A = \frac{\partial F}{\partial y}$  is the

linearization of the nonlinear residual function F(y) = 0. SUNNONLINSOL implementations that do not require solving this system or do not use SUNLINSOL linear solvers may

set this operation to NULL.

# SUNNonlinSolSetConvTestFn

Call retval = SUNNonlinSolSetConvTestFn(NLS, CTestFn);

 $\label{thm:conv} \textbf{Description} \quad \textbf{The } \textit{optional } \textbf{function } \textbf{SUNNonlinSolSetConvTestFn} \ \text{is used to provide the nonlinear}$ 

solver with a function for determining if the nonlinear solver iteration has converged. This is typically called by SUNDIALS integrators to define their nonlinear convergence

criteria, but may be replaced by the user.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

CTestFn (SUNNonlineSolConvTestFn) a SUNDIALS integrator's nonlinear solver conver-

gence test function. See section 10.1.4 for the definition of

SUNNonlinSolConvTestFn.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

Notes Sunnonlinsol implementations utilizing their own convergence test criteria may set

this function to NULL.

#### SUNNonlinSolSetMaxIters

Call retval = SUNNonlinSolSetMaxIters(NLS, maxiters);

Description The optional function SUNNonlinSolSetMaxIters sets the maximum number of non-

linear solver iterations. This is typically called by SUNDIALS integrators to define their

default iteration limit, but may be adjusted by the user.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object.

maxiters (int) the maximum number of nonlinear iterations.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure (e.g., maxiters < 1).

### 10.1.3 SUNNonlinearSolver get functions

The following get functions allow SUNDIALS integrators to retrieve nonlinear solver statistics. The routines to get the current total number of iterations (SUNNonlinSolGetNumIters) and number of convergence failures (SUNNonlinSolGetNumConvFails) are optional. The routine to get the current nonlinear solver iteration (SUNNonlinSolGetCurIter) is required when using the convergence test provided by the SUNDIALS integrator or by the ARKODE and CVODE linear solver interfaces. Otherwise, SUNNonlinSolGetCurIter is optional.

#### SUNNonlinSolGetNumIters

Call retval = SUNNonlinSolGetNumIters(NLS, numiters);

Description The optional function SUNNonlinSolGetNumIters returns the total number of nonlin-

ear solver iterations. This is typically called by the SUNDIALS integrator to store the

nonlinear solver statistics, but may also be called by the user.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object

numiters (long int\*) the total number of nonlinear solver iterations.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

#### SUNNonlinSolGetCurIter

Call retval = SUNNonlinSolGetCurIter(NLS, iter);

Description The function SUNNonlinSolGetCurIter returns the iteration index of the current non-

linear solve. This function is *required* when using SUNDIALS integrator-provided convergence tests or when using a SUNLINSOL spils linear solver; otherwise it is *optional*.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object

iter (int\*) the nonlinear solver iteration in the current solve starting from zero.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

#### SUNNonlinSolGetNumConvFails

Call retval = SUNNonlinSolGetNumConvFails(NLS, nconvfails);

Description The optional function SUNNonlinSolGetNumConvFails returns the total number of non-

linear solver convergence failures. This may be called by the SUNDIALS integrator to

store the nonlinear solver statistics, but may also be called by the user.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object

nconvfails (long int\*) the total number of nonlinear solver convergence failures.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

#### 10.1.4 Functions provided by SUNDIALS integrators

To interface with SUNNONLINSOL modules, the SUNDIALS integrators supply a variety of routines for evaluating the nonlinear system, calling the SUNLINSOL setup and solve functions, and testing the nonlinear iteration for convergence. These integrator-provided routines translate between the user-supplied ODE or DAE systems and the generic interfaces to the nonlinear or linear systems of equations that result in their solution. The types for functions provided to a SUNNONLINSOL module are defined in the header file sundials/sundials\_nonlinearsolver.h, and are described below.

# SUNNonlinSolSysFn

Definition typedef int (\*SUNNonlinSolSysFn)(N\_Vector y, N\_Vector F, void\* mem);

Purpose These functions evaluate the nonlinear system F(y) for SUNNONLINEARSOLVER\_ROOTFIND

type modules or G(y) for SUNNONLINEARSOLVER\_FIXEDPOINT type modules. Memory for F must by be allocated prior to calling this function. The vector y must be left

unchanged.

Arguments y is the state vector at which the nonlinear system should be evaluated.

F is the output vector containing F(y) or G(y), depending on the solver type.

mem is the SUNDIALS integrator memory structure.

Return value The return value retval (of type int) is zero for a successul solve, a positive value for a recoverable error, and a negative value for an unrecoverable error.

#### SUNNonlinSolLSetupFn

Definition typedef int (\*SUNNonlinSolLSetupFn)(N\_Vector y, N\_Vector F,

booleantype jbad,

booleantype\* jcur, void\* mem);

Purpose These functions are wrappers to the Sundials integrator's function for setting up linear solves with Sunlinsol modules.

Arguments y is the state vector at which the linear system should be setup.

F is the value of the nonlinear system function at y.

jbad is an input indicating whether the nonlinear solver believes that A has gone stale (SUNTRUE) or not (SUNFALSE).

jcur is an output indicating whether the routine has updated the Jacobian A (SUNTRUE) or not (SUNFALSE).

mem is the SUNDIALS integrator memory structure.

Return value The return value retval (of type int) is zero for a successul solve, a positive value for a recoverable error, and a negative value for an unrecoverable error.

Notes

The SUNNonlinLSetupFn function sets up the linear system Ax = b where  $A = \frac{\partial F}{\partial y}$  is the linearization of the nonlinear residual function F(y) = 0 (when using SUNLINSOL direct linear solvers) or calls the user-defined preconditioner setup function (when using SUNLINSOL iterative linear solvers). SUNNONLINSOL implementations that do not require solving this system, do not utilize SUNLINSOL linear solvers, or use SUNLINSOL linear solvers that do not require setup may ignore these functions.

#### SUNNonlinSolLSolveFn

Definition typedef int (\*SUNNonlinSolLSolveFn)(N\_Vector y, N\_Vector b, void\* mem);

Purpose These functions are wrappers to the SUNDIALS integrator's function for solving linear systems with SUNLINSOL modules.

Arguments y is the input vector containing the current nonlinear iteration.

b contains the right-hand side vector for the linear solve on input and the solution to the linear system on output.

mem is the SUNDIALS integrator memory structure.

Return value The return value retval (of type int) is zero for a successul solve, a positive value for a recoverable error, and a negative value for an unrecoverable error.

Notes The SUNNonlinLSolveFn function solves the linear system Ax = b where  $A = \frac{\partial F}{\partial y}$  is the linearization of the nonlinear residual function F(y) = 0. SUNNONLINSOL implementations that do not require solving this system or do not use SUNLINSOL linear solvers may ignore these functions.

#### SUNNonlinSolConvTestFn

Definition typedef int (\*SUNNonlinSolConvTestFn)(SUNNonlinearSolver NLS, N\_Vector y, N\_Vector del, realtype tol, N\_Vector ewt, void\* mem);

Purpose These functions are SUNDIALS integrator-specific convergence tests for nonlinear solvers and are typically supplied by each SUNDIALS integrator, but users may supply custom problem-specific versions as desired.

y is the current nonlinear iterate.

del is the difference between the current and prior nonlinear iterates.

tol is the nonlinear solver tolerance.

ewt is the weight vector used in computing weighted norms.

mem is the SUNDIALS integrator memory structure.

Return value The return value of this routine will be a negative value if an unrecoverable error occurred or one of the following:

SUN\_NLS\_SUCCESS the iteration is converged.

SUN\_NLS\_CONTINUE the iteration has not converged, keep iterating.

SUN\_NLS\_CONV\_RECVR the iteration appears to be diverging, try to recover.

Notes

The tolerance passed to this routine by SUNDIALS integrators is the tolerance in a weighted root-mean-squared norm with error weight vector ewt. SUNNONLINSOL modules utilizing their own convergence criteria may ignore these functions.

#### 10.1.5 SUNNonlinearSolver return codes

The functions provided to SUNNONLINSOL modules by each SUNDIALS integrator, and functions within the SUNDIALS-provided SUNNONLINSOL implementations utilize a common set of return codes, shown below in Table 10.1. Here, negative values correspond to non-recoverable failures, positive values to recoverable failures, and zero to a successful call.

Name	Value	Description
SUN_NLS_SUCCESS	0	successful call or converged solve
SUN_NLS_CONTINUE	1	the nonlinear solver is not converged, keep iterating
SUN_NLS_CONV_RECVR	2	the nonlinear solver appears to be diverging, try to recover
SUN_NLS_MEM_NULL	-1	a memory argument is NULL
SUN_NLS_MEM_FAIL	-2	a memory access or allocation failed
SUN_NLS_ILL_INPUT	-3	an illegal input option was provided

Table 10.1: Description of the SUNNonlinearSolver return codes

# 10.1.6 The generic SUNNonlinear Solver module

SUNDIALS integrators interact with specific SUNNONLINSOL implementations through the generic SUNNONLINSOL module on which all other SUNNONLINSOL implementations are built. The SUNNonlinearSolver type is a pointer to a structure containing an implementation-dependent *content* field and an *ops* field. The type SUNNonlinearSolver is defined as follows:

```
typedef struct _generic_SUNNonlinearSolver *SUNNonlinearSolver;
struct _generic_SUNNonlinearSolver {
  void *content;
  struct _generic_SUNNonlinearSolver_Ops *ops;
};
```

where the <code>\_generic\_SUNNonlinearSolver\_Ops</code> structure is a list of pointers to the various actual non-linear solver operations provided by a specific implementation. The <code>\_generic\_SUNNonlinearSolver\_Ops</code> structure is defined as

```
struct _generic_SUNNonlinearSolver_Ops {
  SUNNonlinearSolver_Type (*gettype)(SUNNonlinearSolver);
  int
                           (*initialize)(SUNNonlinearSolver);
  int
                           (*setup)(SUNNonlinearSolver, N_Vector, void*);
  int
                           (*solve)(SUNNonlinearSolver, N_Vector, N_Vector,
                                    N_Vector, realtype, booleantype, void*);
  int
                           (*free)(SUNNonlinearSolver);
  int
                           (*setsysfn)(SUNNonlinearSolver, SUNNonlinSolSysFn);
                           (*setlsetupfn)(SUNNonlinearSolver, SUNNonlinSolLSetupFn);
  int
  int
                           (*setlsolvefn)(SUNNonlinearSolver, SUNNonlinSolLSolveFn);
  int
                           (*setctestfn)(SUNNonlinearSolver, SUNNonlinSolConvTestFn);
                           (*setmaxiters)(SUNNonlinearSolver, int);
  int
  int
                           (*getnumiters)(SUNNonlinearSolver, long int*);
                           (*getcuriter)(SUNNonlinearSolver, int*);
  int
  int
                           (*getnumconvfails)(SUNNonlinearSolver, long int*);
};
```

The generic SUNNONLINSOL module defines and implements the nonlinear solver operations defined in Sections 10.1.1 – 10.1.3. These routines are in fact only wrappers to the nonlinear solver operations provided by a particular SUNNONLINSOL implementation, which are accessed through the ops field of the SUNNonlinearSolver structure. To illustrate this point we show below the implementation of a typical nonlinear solver operation from the generic SUNNONLINSOL module, namely SUNNonlinSolSolve, which solves the nonlinear system and returns a flag denoting a successful or failed solve:

# 10.1.7 Usage with sensitivity enabled integrators

When used with SUNDIALS packages that support sensitivity analysis capabilities (e.g., CVODES and IDAS) a special NVECTOR module is used to interface with SUNNONLINSOL modules for solves involving sensitivity vectors stored in an NVECTOR array. As described below, the NVECTOR\_SENSWRAPPER module is an NVECTOR implementation where the vector content is an NVECTOR array. This wrapper vector allows SUNNONLINSOL modules to operate on data stored as a collection of vectors.

For all SUNDIALS-provided SUNNONLINSOL modules a special constructor wrapper is provided so users do not need to interact directly with the NVECTOR\_SENSWRAPPER module. These constructors follow the naming convention SUNNonlinSol\_\*\*\*Sens(count,...) where \*\*\* is the name of the SUNNONLINSOL module, count is the size of the vector wrapper, and ... are the module-specific constructor arguments.

#### The NVECTOR\_SENSWRAPPER module

This section describes the NVECTOR\_SENSWRAPPER implementation of an NVECTOR. To access the NVECTOR\_SENSWRAPPER module, include the header file sundials\_nvector\_senswrapper.h.

The NVECTOR\_SENSWRAPPER module defines an N\_Vector implementing all of the standard vectors operations defined in Table 7.2 but with some changes to how operations are computed in order to accommodate operating on a collection of vectors.

1. Element-wise vector operations are computed on a vector-by-vector basis. For example, the linear sum of two wrappers containing  $n_v$  vectors of length n, N\_VLinearSum(a,x,b,y,z), is computed as

$$z_{j,i} = ax_{j,i} + by_{j,i}, \quad i = 0, \dots, n-1, \quad j = 0, \dots, n_v - 1.$$

2. The dot product of two wrappers containing  $n_v$  vectors of length n is computed as if it were the dot product of two vectors of length  $nn_v$ . Thus  $d = N_v Dot Prod(x, y)$  is

$$d = \sum_{j=0}^{n_v - 1} \sum_{i=0}^{n-1} x_{j,i} y_{j,i}.$$

3. All norms are computed as the maximum of the individual norms of the  $n_v$  vectors in the wrapper. For example, the weighted root mean square norm  $\mathbf{m} = \mathbf{N}_{\text{-}}\mathbf{V}\mathbf{W}\mathbf{r}\mathbf{m}\mathbf{s}\mathbf{N}\mathbf{o}\mathbf{r}\mathbf{m}(\mathbf{x}, \mathbf{w})$  is

$$m = \max_{j} \sqrt{\left(\frac{1}{n} \sum_{i=0}^{n-1} (x_{j,i} w_{j,i})^2\right)}$$

To enable usage alongside other NVECTOR modules the NVECTOR\_SENSWRAPPER functions implementing vector operations have \_SensWrapper appended to the generic vector operation name.

The NVECTOR\_SENSWRAPPER module provides the following constructors for creating an NVECTOR\_SENSWRAPPER:

#### N\_VNewEmpty\_SensWrapper

Call w = N\_VNewEmpty\_SensWrapper(count);

Description The function N\_VNewEmpty\_SensWrapper creates an empty NVECTOR\_SENSWRAPPER wrapper with space for count vectors.

Arguments count (int) the number of vectors the wrapper will contain.

Return value The return value w (of type N\_Vector) will be a NVECTOR object if the constructor exits successfully, otherwise w will be NULL.

#### N\_VNew\_SensWrapper

Call w = N\_VNew\_SensWrapper(count, y);

Description The function N\_VNew\_SensWrapper creates an NVECTOR\_SENSWRAPPER wrapper containing count vectors cloned from y.

Arguments count (int) the number of vectors the wrapper will contain.

y (N\_Vector) the template vectors to use in creating the vector wrapper.

Return value The return value w (of type N\_Vector) will be a NVECTOR object if the constructor exits successfully, otherwise w will be NULL.

The NVECTOR\_SENSWRAPPER implementation of the NVECTOR module defines the *content* field of the N\_Vector to be a structure containing an N\_Vector array, the number of vectors in the vector array, and a boolean flag indicating ownership of the vectors in the vector array.

```
struct _N_VectorContent_SensWrapper {
   N_Vector* vecs;
   int nvecs;
   booleantype own_vecs;
};
```

The following macros are provided to access the content of an NVECTOR\_SENSWRAPPER vector.

- NV\_CONTENT\_SW(v) provides access to the content structure
- NV\_VECS\_SW(v) provides access to the vector array
- NV\_NVECS\_SW(v) provides access to the number of vectors
- NV\_OWN\_VECS\_SW(v) provides access to the ownership flag
- NV\_VEC\_SW(v,i) provides access to the i-th vector in the vector array

# 10.1.8 Implementing a Custom SUNNonlinear Solver Module

A SUNNONLINSOL implementation must do the following:

- 1. Specify the content of the SUNNONLINSOL module.
- 2. Define and implement the required nonlinear solver operations defined in Sections 10.1.1 10.1.3. Note that the names of the module routines should be unique to that implementation in order to permit using more than one SUNNONLINSOL module (each with different SUNNonlinearSolver internal data representations) in the same code.
- 3. Define and implement a user-callable constructor to create a SUNNonlinearSolver object.

Additionally, a SUNNonlinearSolver implementation may do the following:

- 1. Define and implement additional user-callable "set" routines acting on the SUNNonlinearSolver object, e.g., for setting various configuration options to tune the performance of the nonlinear solve algorithm.
- 2. Provide additional user-callable "get" routines acting on the SUNNonlinearSolver object, e.g., for returning various solve statistics.

# 10.2 The SUNNonlinearSolver\_Newton implementation

This section describes the SUNNONLINSOL implementation of Newton's method. To access the SUNNON-LINSOL\_NEWTON module, include the header file sunnonlinsol/sunnonlinsol\_newton.h. We note that the SUNNONLINSOL\_NEWTON module is accessible from SUNDIALS integrators without separately linking to the libsundials\_sunnonlinsolnewton module library.

# 10.2.1 SUNNonlinearSolver\_Newton description

To find the solution to

$$F(y) = 0 \tag{10.1}$$

given an initial guess  $y^{(0)}$ , Newton's method computes a series of approximate solutions

$$y^{(m+1)} = y^{(m)} + \delta^{(m+1)} \tag{10.2}$$

where m is the Newton iteration index, and the Newton update  $\delta^{(m+1)}$  is the solution of the linear system

$$A(y^{(m)})\delta^{(m+1)} = -F(y^{(m)}), (10.3)$$

in which A is the Jacobian matrix

$$A \equiv \partial F/\partial y. \tag{10.4}$$

Depending on the linear solver used, the Sunnonlinsol\_newton module will employ either a Modified Newton method, or an Inexact Newton method [5, 9, 18, 20, 33]. When used with a direct linear solver, the Jacobian matrix A is held constant during the Newton iteration, resulting in a Modified Newton method. With a matrix-free iterative linear solver, the iteration is an Inexact Newton method.

In both cases, calls to the integrator-supplied SUNNonlinSollSetupFn function are made infrequently to amortize the increased cost of matrix operations (updating A and its factorization within direct linear solvers, or updating the preconditioner within iterative linear solvers). Specifically, SUNNONLINSOL\_NEWTON will call the SUNNonlinSollSetupFn function in two instances:

- (a) when requested by the integrator (the input callLSetSetup is SUNTRUE) before attempting the Newton iteration, or
- (b) when reattempting the nonlinear solve after a recoverable failure occurs in the Newton iteration with stale Jacobian information (jcur is SUNFALSE). In this case, SUNNONLINSOL\_NEWTON will set jbad to SUNTRUE before calling the SUNNonlinSollSetupFn function.

Whether the Jacobian matrix A is fully or partially updated depends on logic unique to each integrator-supplied SUNNonlinSolSetupFn routine. We refer to the discussion of nonlinear solver strategies provided in Chapter 2 for details on this decision.

The default maximum number of iterations and the stopping criteria for the Newton iteration are supplied by the SUNDIALS integrator when SUNNONLINSOL\_NEWTON is attached to it. Both the maximum number of iterations and the convergence test function may be modified by the user by calling the SUNNonlinSolSetMaxIters and/or SUNNonlinSolSetConvTestFn functions after attaching the SUNNONLINSOL\_NEWTON object to the integrator.

#### 10.2.2 SUNNonlinearSolver\_Newton functions

The SUNNONLINSOL\_NEWTON module provides the following constructors for creating a SUNNonlinearSolver object.

#### SUNNonlinSol\_Newton

Description The function SUNNonlinSol\_Newton creates a SUNNonlinearSolver object for use with

SUNDIALS integrators to solve nonlinear systems of the form F(y) = 0 using Newton's

method.

Arguments y (N\_Vector) a template for cloning vectors needed within the solver.

Return value The return value NLS (of type SUNNonlinearSolver) will be a SUNNONLINSOL object if

the constructor exits successfully, otherwise NLS will be NULL.

F2003 Name This function is callable as FSUNNonlinSol\_Newton when using the Fortran 2003 inter-

face module.

# SUNNonlinSol\_NewtonSens

Call NLS = SUNNonlinSol\_NewtonSens(count, y);

Description The function SUNNonlinSol\_NewtonSens creates a SUNNonlinearSolver object for use

with SUNDIALS sensitivity enabled integrators (CVODES and IDAS) to solve nonlinear

systems of the form F(y) = 0 using Newton's method.

Arguments count (int) the number of vectors in the nonlinear solve. When integrating a system containing Ns sensitivities the value of count is:

• Ns+1 if using a *simultaneous* corrector approach.

• Ns if using a *staggered* corrector approach.

y (N\_Vector) a template for cloning vectors needed within the solver.

Return value The return value NLS (of type SUNNonlinearSolver) will be a SUNNONLINSOL object if the constructor exits successfully, otherwise NLS will be NULL.

F2003 Name This function is callable as FSUNNonlinSol\_NewtonSens when using the Fortran 2003 interface module.

The SUNNONLINSOL\_NEWTON module implements all of the functions defined in sections 10.1.1 - 10.1.3 except for the SUNNonlinSolSetup function. The SUNNONLINSOL\_NEWTON functions have the same names as those defined by the generic SUNNONLINSOL API with \_Newton appended to the function name. Unless using the SUNNONLINSOL\_NEWTON module as a standalone nonlinear solver the generic functions defined in sections 10.1.1 - 10.1.3 should be called in favor of the SUNNONLINSOL\_NEWTON-specific implementations.

The SUNNONLINSOL\_NEWTON module also defines the following additional user-callable function.

# ${\tt SUNNonlinSolGetSysFn\_Newton}$

Call retval = SUNNonlinSolGetSysFn\_Newton(NLS, SysFn);

Description The function SUNNonlinSolGetSysFn\_Newton returns the residual function that defines

the nonlinear system.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object

SysFn (SUNNonlinSolSysFn\*) the function defining the nonlinear system.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

Notes This function is intended for users that wish to evaluate the nonlinear residual in a

custom convergence test function for the SUNNONLINSOL\_NEWTON module. We note that SUNNONLINSOL\_NEWTON will not leverage the results from any user calls to SysFn.

F2003 Name This function is callable as FSUNNonlinSolGetSysFn\_Newton when using the Fortran

2003 interface module.

#### 10.2.3 SUNNonlinearSolver\_Newton Fortran interfaces

The SUNNONLINSOL\_NEWTON module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

#### FORTRAN 2003 interface module

The fsunnonlinsol\_newton\_mod FORTRAN module defines interfaces to all SUNNONLINSOL\_NEWTON C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNNonlinSol\_Newton is interfaced as FSUNNonlinSol\_Newton.

The FORTRAN 2003 SUNNONLINSOL\_NEWTON interface module can be accessed with the use statement, i.e. use fsunnonlinsol\_newton\_mod, and linking to the library

libsundials\_fsunnonlinsolnewton\_mod.lib in addition to the C library. For details on where the library and module file fsunnonlinsol\_newton\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunnonlinsolnewton\_mod library.

#### FORTRAN 77 interface functions

For SUNDIALS integrators that include a FORTRAN 77 interface, the SUNNONLINSOL\_NEWTON module also includes a Fortran-callable function for creating a SUNNonlinearSolver object.

### FSUNNEWTONINIT

Call FSUNNEWTONINIT(code, ier);

Description The function FSUNNEWTONINIT can be called for Fortran programs to create a

SUNNonlinearSolver object for use with SUNDIALS integrators to solve nonlinear sys-

tems of the form F(y) = 0 with Newton's method.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, and 4 for ARKODE).

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See printed message for details in case of failure.

## 10.2.4 SUNNonlinearSolver\_Newton content

The SUNNONLINSOL\_NEWTON module defines the *content* field of a SUNNonlinearSolver as the following structure:

```
struct _SUNNonlinearSolverContent_Newton {
```

```
SUNNonlinSolSysFn
                          Sys;
  SUNNonlinSolLSetupFn
                          LSetup;
  SUNNonlinSolLSolveFn
                          LSolve;
  SUNNonlinSolConvTestFn CTest;
  N_Vector
               delta:
  booleantype jcur;
  int
               curiter;
  int
              maxiters:
  long int
              niters;
  long int
              nconvfails;
};
```

These entries of the *content* field contain the following information:

Sys - the function for evaluating the nonlinear system,

LSetup - the package-supplied function for setting up the linear solver,
 LSolve - the package-supplied function for performing a linear solve,
 CTest - the function for checking convergence of the Newton iteration,

delta - the Newton iteration update vector,

jcur - the Jacobian status (SUNTRUE = current, SUNFALSE = stale),

curiter - the current number of iterations in the solve attempt,

maxiters - the maximum number of Newton iterations allowed in a solve, and

niters - the total number of nonlinear iterations across all solves.

nconvfails - the total number of nonlinear convergence failures across all solves.

## 10.3 The SUNNonlinearSolver\_FixedPoint implementation

This section describes the SUNNONLINSOL implementation of a fixed point (functional) iteration with optional Anderson acceleration. To access the SUNNONLINSOL\_FIXEDPOINT module, include the header file sunnonlinsol/sunnonlinsol\_fixedpoint.h. We note that the SUNNONLINSOL\_FIXEDPOINT module is accessible from SUNDIALS integrators without separately linking to the libsundials\_sunnonlinsolfixedpoint module library.

## 10.3.1 SUNNonlinearSolver\_FixedPoint description

To find the solution to

$$G(y) = y \tag{10.5}$$

given an initial guess  $y^{(0)}$ , the fixed point iteration computes a series of approximate solutions

$$y^{(n+1)} = G(y^{(n)}) (10.6)$$

where n is the iteration index. The convergence of this iteration may be accelerated using Anderson's method [3, 45, 21, 36]. With Anderson acceleration using subspace size m, the series of approximate solutions can be formulated as the linear combination

$$y^{(n+1)} = \sum_{i=0}^{m_n} \alpha_i^{(n)} G(y^{(n-m_n+i)})$$
(10.7)

where  $m_n = \min\{m, n\}$  and the factors

$$\alpha^{(n)} = (\alpha_0^{(n)}, \dots, \alpha_{m_n}^{(n)}) \tag{10.8}$$

solve the minimization problem  $\min_{\alpha} \|F_n \alpha^T\|_2$  under the constraint that  $\sum_{i=0}^{m_n} \alpha_i = 1$  where

$$F_n = (f_{n-m_n}, \dots, f_n) \tag{10.9}$$

with  $f_i = G(y^{(i)}) - y^{(i)}$ . Due to this constraint, in the limit of m = 0 the accelerated fixed point iteration formula (10.7) simplifies to the standard fixed point iteration (10.6).

Following the recommendations made in [45], the SUNNONLINSOL\_FIXEDPOINT implementation computes the series of approximate solutions as

$$y^{(n+1)} = G(y^{(n)}) - \sum_{i=0}^{m_n - 1} \gamma_i^{(n)} \Delta g_{n-m_n+i}$$
(10.10)

with  $\Delta g_i = G(y^{(i+1)}) - G(y^{(i)})$  and where the factors

$$\gamma^{(n)} = (\gamma_0^{(n)}, \dots, \gamma_{m-1}^{(n)}) \tag{10.11}$$

solve the unconstrained minimization problem  $\min_{\gamma} \|f_n - \Delta F_n \gamma^T\|_2$  where

$$\Delta F_n = (\Delta f_{n-m_n}, \dots, \Delta f_{n-1}), \tag{10.12}$$

with  $\Delta f_i = f_{i+1} - f_i$ . The least-squares problem is solved by applying a QR factorization to  $\Delta F_n = Q_n R_n$  and solving  $R_n \gamma = Q_n^T f_n$ .

The acceleration subspace size m is required when constructing the SUNNONLINSOL\_FIXEDPOINT object. The default maximum number of iterations and the stopping criteria for the fixed point iteration are supplied by the SUNDIALS integrator when SUNNONLINSOL\_FIXEDPOINT is attached to it. Both the maximum number of iterations and the convergence test function may be modified by the user by calling SUNNonlinSolSetMaxIters and SUNNonlinSolSetConvTestFn functions after attaching the SUNNONLINSOL\_FIXEDPOINT object to the integrator.

#### 10.3.2 SUNNonlinearSolver\_FixedPoint functions

The SUNNONLINSOL\_FIXEDPOINT module provides the following constructors for creating a SUNNonlinearSolver object.

## SUNNonlinSol\_FixedPoint

Description The function SUNNonlinSol\_FixedPoint creates a SUNNonlinearSolver object for use with SUNDIALS integrators to solve nonlinear systems of the form G(y) = y.

 $Arguments \quad \hbox{$\tt y$ (N\_Vector)$ a template for cloning vectors needed within the solver}$ 

m (int) the number of acceleration vectors to use

Return value The return value NLS (of type SUNNonlinearSolver) will be a SUNNONLINSOL object if the constructor exits successfully, otherwise NLS will be NULL.

F2003 Name This function is callable as FSUNNonlinSol\_FixedPoint when using the Fortran 2003 interface module.

## SUNNonlinSol\_FixedPointSens

Call NLS = SUNNonlinSol\_FixedPointSens(count, y, m);

Description The function SUNNonlinSol\_FixedPointSens creates a SUNNonlinearSolver object for

use with SUNDIALS sensitivity enabled integrators (CVODES and IDAS) to solve nonlinear

systems of the form G(y) = y.

Arguments count (int) the number of vectors in the nonlinear solve. When integrating a system

containing Ns sensitivities the value of count is:

• Ns+1 if using a *simultaneous* corrector approach.

• Ns if using a *staggered* corrector approach.

y (N\_Vector) a template for cloning vectors needed within the solver.

m (int) the number of acceleration vectors to use.

Return value The return value NLS (of type SUNNonlinearSolver) will be a SUNNONLINSOL object if the constructor exits successfully, otherwise NLS will be NULL.

F2003 Name This function is callable as FSUNNonlinSol\_FixedPointSens when using the Fortran 2003 interface module.

Since the accelerated fixed point iteration (10.6) does not require the setup or solution of any linear systems, the SUNNONLINSOL\_FIXEDPOINT module implements all of the functions defined in sections 10.1.1 – 10.1.3 except for the SUNNonlinSolSetup, SUNNonlinSolSetLSetupFn, and

SUNNonlinSolSetLSolveFn functions, that are set to NULL. The SUNNONLINSOL\_FIXEDPOINT functions have the same names as those defined by the generic SUNNONLINSOL API with <code>FixedPoint</code> appended to the function name. Unless using the SUNNONLINSOL\_FIXEDPOINT module as a standalone nonlinear solver the generic functions defined in sections 10.1.1 - 10.1.3 should be called in favor of the SUNNONLINSOL\_FIXEDPOINT-specific implementations.

The  ${\tt SUNNONLINSOL\_FIXEDPOINT}$  module also defines the following additional user-callable function.

## SUNNonlinSolGetSysFn\_FixedPoint

Call retval = SUNNonlinSolGetSysFn\_FixedPoint(NLS, SysFn);

Description The function SUNNonlinSolGetSysFn\_FixedPoint returns the fixed-point function that

defines the nonlinear system.

Arguments NLS (SUNNonlinearSolver) a SUNNONLINSOL object

SysFn (SUNNonlinSolSysFn\*) the function defining the nonlinear system.

Return value The return value retval (of type int) should be zero for a successful call, and a negative

value for a failure.

Notes This function is intended for users that wish to evaluate the fixed-point function in a

custom convergence test function for the  ${\tt SUNNONLINSOL\_FIXEDPOINT}$  module. We note that  ${\tt SUNNONLINSOL\_FIXEDPOINT}$  will not leverage the results from any user calls to

SysFn.

F2003 Name This function is callable as FSUNNonlinSolGetSysFn\_FixedPoint when using the For-

tran 2003 interface module.

## 10.3.3 SUNNonlinearSolver\_FixedPoint Fortran interfaces

The SUNNONLINSOL\_FIXEDPOINT module provides a FORTRAN 2003 module as well as FORTRAN 77 style interface functions for use from FORTRAN applications.

### FORTRAN 2003 interface module

The fsunnonlinsol\_fixedpoint\_mod FORTRAN module defines interfaces to all SUNNONLINSOL\_FIXEDPOINT C functions using the intrinsic iso\_c\_binding module which provides a standardized mechanism for interoperating with C. As noted in the C function descriptions above, the interface functions are named after the corresponding C function, but with a leading 'F'. For example, the function SUNNonlinSol\_FixedPoint is interfaced as FSUNNonlinSol\_FixedPoint.

The FORTRAN 2003 SUNNONLINSOL\_FIXEDPOINT interface module can be accessed with the use statement, i.e. use fsunnonlinsol\_fixedpoint\_mod, and linking to the library libsundials\_fsunnonlinsolfixedpoint\_mod.lib in addition to the C library. For details on where the library and module file fsunnonlinsol\_fixedpoint\_mod.mod are installed see Appendix A. We note that the module is accessible from the FORTRAN 2003 SUNDIALS integrators without separately linking to the libsundials\_fsunnonlinsolfixedpoint\_mod library.

#### FORTRAN 77 interface functions

For SUNDIALS integrators that include a FORTRAN 77 interface, the SUNNONLINSOL\_FIXEDPOINT module also includes a Fortran-callable function for creating a SUNNonlinearSolver object.

### FSUNFIXEDPOINTINIT

```
Call FSUNFIXEDPOINTINIT(code, m, ier);
```

Description The function FSUNFIXEDPOINTINIT can be called for Fortran programs to create a

SUNNonlinearSolver object for use with SUNDIALS integrators to solve nonlinear sys-

tems of the form G(y) = y.

Arguments code (int\*) is an integer input specifying the solver id (1 for CVODE, 2 for IDA, and 4

for ARKODE).

m (int\*) is an integer input specifying the number of acceleration vectors.

Return value ier is a return completion flag equal to 0 for a success return and -1 otherwise. See printed message for details in case of failure.

## 10.3.4 SUNNonlinearSolver\_FixedPoint content

The SUNNONLINSOL\_FIXEDPOINT module defines the *content* field of a SUNNonlinearSolver as the following structure:

```
struct _SUNNonlinearSolverContent_FixedPoint {
```

```
SUNNonlinSolSysFn
                       Sys;
SUNNonlinSolConvTestFn CTest;
int
          m;
int
         *imap;
realtype *R;
realtype *gamma;
realtype *cvals;
N_Vector *df;
N_Vector *dg;
N_Vector *q;
N_Vector *Xvecs;
N_Vector yprev;
N_Vector gy;
N_Vector fold;
N_Vector gold;
N_Vector delta;
```

- N\_Vector used in acceleration algorithm

gold

```
int
              curiter;
  int
              maxiters;
              niters;
  long int
  long int
              nconvfails;
};
The following entries of the content field are always allocated:
             - function for evaluating the nonlinear system,
             - function for checking convergence of the fixed point iteration,
CTest
             - N_Vector used to store previous fixed-point iterate,
yprev
             - N_Vector used to store G(y) in fixed-point algorithm,
gу
delta
             - N_Vector used to store difference between successive fixed-point iterates,
curiter
             - the current number of iterations in the solve attempt,
maxiters
             - the maximum number of fixed-point iterations allowed in a solve, and
             - the total number of nonlinear iterations across all solves.
niters
nconvfails - the total number of nonlinear convergence failures across all solves.
             - number of acceleration vectors.
If Anderson acceleration is requested (i.e., m > 0 in the call to SUNNonlinSol_FixedPoint), then the
following items are also allocated within the content field:
          - index array used in acceleration algorithm (length m)
R
          - small matrix used in acceleration algorithm (length m*m)
          - small vector used in acceleration algorithm (length m)
gamma
          - small vector used in acceleration algorithm (length m+1)
cvals
df
          - array of N_Vectors used in acceleration algorithm (length m)
          - array of N_Vectors used in acceleration algorithm (length m)
dg
          - array of N_Vectors used in acceleration algorithm (length m)
q
          - N_Vector pointer array used in acceleration algorithm (length m+1)
Xvecs
          - N_Vector used in acceleration algorithm
fold
```

## Appendix A

## SUNDIALS Package Installation Procedure

The installation of any SUNDIALS package is accomplished by installing the SUNDIALS suite as a whole, according to the instructions that follow. The same procedure applies whether or not the downloaded file contains one or all solvers in SUNDIALS.

The SUNDIALS suite (or individual solvers) are distributed as compressed archives (.tar.gz). The name of the distribution archive is of the form *solver-x.y.z.tar.gz*, where *solver* is one of: sundials, cvode, cvodes, arkode, ida, idas, or kinsol, and x.y.z represents the version number (of the SUNDIALS suite or of the individual solver). To begin the installation, first uncompress and expand the sources, by issuing

% tar xzf solver-x.y.z.tar.gz

This will extract source files under a directory *solver*-x.y.z.

Starting with version 2.6.0 of SUNDIALS, CMake is the only supported method of installation. The explanations of the installation procedure begins with a few common observations:

• The remainder of this chapter will follow these conventions:

solverdir is the directory solver-x.y.z created above; i.e., the directory containing the SUNDI-ALS sources.

builddir is the (temporary) directory under which SUNDIALS is built.

instdir is the directory under which the SUNDIALS exported header files and libraries will be installed. Typically, header files are exported under a directory instdir/include while libraries are installed under instdir/CMAKE\_INSTALL\_LIBDIR, with instdir and CMAKE\_INSTALL\_LIBDIR specified at configuration time.

- For SUNDIALS CMake-based installation, in-source builds are prohibited; in other words, the build directory builddir can **not** be the same as solverdir and such an attempt will lead to an error. This prevents "polluting" the source tree and allows efficient builds for different configurations and/or options.
- The installation directory *instdir* can **not** be the same as the source directory *solverdir*.
- By default, only the libraries and header files are exported to the installation directory *instdir*. If enabled by the user (with the appropriate toggle for CMake), the examples distributed with SUNDIALS will be built together with the solver libraries but the installation step will result in exporting (by default in a subdirectory of the installation directory) the example sources and sample outputs together with automatically generated configuration files that reference the *installed* SUNDIALS headers and libraries. As such, these configuration files for the SUNDIALS examples can be used as "templates" for your own problems. CMake installs CMakeLists.txt files



and also (as an option available only under Unix/Linux) Makefile files. Note this installation approach also allows the option of building the SUNDIALS examples without having to install them. (This can be used as a sanity check for the freshly built libraries.)

• Even if generation of shared libraries is enabled, only static libraries are created for the FCMIX modules. (Because of the use of fixed names for the Fortran user-provided subroutines, FCMIX shared libraries would result in "undefined symbol" errors at link time.)

## A.1 CMake-based installation

CMake-based installation provides a platform-independent build system. CMake can generate Unix and Linux Makefiles, as well as KDevelop, Visual Studio, and (Apple) XCode project files from the same configuration file. In addition, CMake also provides a GUI front end and which allows an interactive build and installation process.

The SUNDIALS build process requires CMake version 3.1.3 or higher and a working C compiler. On Unix-like operating systems, it also requires Make (and curses, including its development libraries, for the GUI front end to CMake, ccmake), while on Windows it requires Visual Studio. CMake is continually adding new features, and the latest version can be downloaded from http://www.cmake.org. Build instructions for CMake (only necessary for Unix-like systems) can be found on the CMake website. Once CMake is installed, Linux/Unix users will be able to use ccmake, while Windows users will be able to use CMakeSetup.

As previously noted, when using CMake to configure, build and install SUNDIALS, it is always required to use a separate build directory. While in-source builds are possible, they are explicitly prohibited by the SUNDIALS CMake scripts (one of the reasons being that, unlike autotools, CMake does not provide a make distclean procedure and it is therefore difficult to clean-up the source tree after an in-source build). By ensuring a separate build directory, it is an easy task for the user to clean-up all traces of the build by simply removing the build directory. CMake does generate a make clean which will remove files generated by the compiler and linker.

## A.1.1 Configuring, building, and installing on Unix-like systems

The default CMake configuration will build all included solvers and associated examples and will build static and shared libraries. The *instdir* defaults to /usr/local and can be changed by setting the CMAKE\_INSTALL\_PREFIX variable. Support for FORTRAN and all other options are disabled.

CMake can be used from the command line with the cmake command, or from a curses-based GUI by using the ccmake command. Examples for using both methods will be presented. For the examples shown it is assumed that there is a top level SUNDIALS directory with appropriate source, build and install directories:

```
% mkdir (...)sundials/instdir
% mkdir (...)sundials/builddir
% cd (...)sundials/builddir
```

### Building with the GUI

Using CMake with the GUI follows this general process:

- Select and modify values, run configure (c key)
- New values are denoted with an asterisk
- To set a variable, move the cursor to the variable and press enter
  - If it is a boolean (ON/OFF) it will toggle the value
  - If it is string or file, it will allow editing of the string

- For file and directories, the <tab> key can be used to complete
- Repeat until all values are set as desired and the generate option is available (g key)
- Some variables (advanced variables) are not visible right away
- To see advanced variables, toggle to advanced mode (t key)
- To search for a variable press / key, and to repeat the search, press the n key

To build the default configuration using the GUI, from the *builddir* enter the ccmake command and point to the *solverdir*:

#### % ccmake ../solverdir

The default configuration screen is shown in Figure A.1.

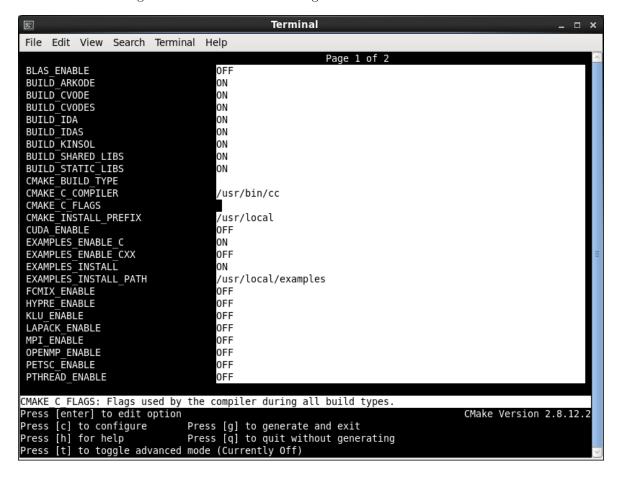


Figure A.1: Default configuration screen. Note: Initial screen is empty. To get this default configuration, press 'c' repeatedly (accepting default values denoted with asterisk) until the 'g' option is available.

The default *instdir* for both SUNDIALS and corresponding examples can be changed by setting the CMAKE\_INSTALL\_PREFIX and the EXAMPLES\_INSTALL\_PATH as shown in figure A.2.

Pressing the (g key) will generate makefiles including all dependencies and all rules to build SUNDIALS on this system. Back at the command prompt, you can now run:

#### % make

To install SUNDIALS in the installation directory specified in the configuration, simply run:

## % make install

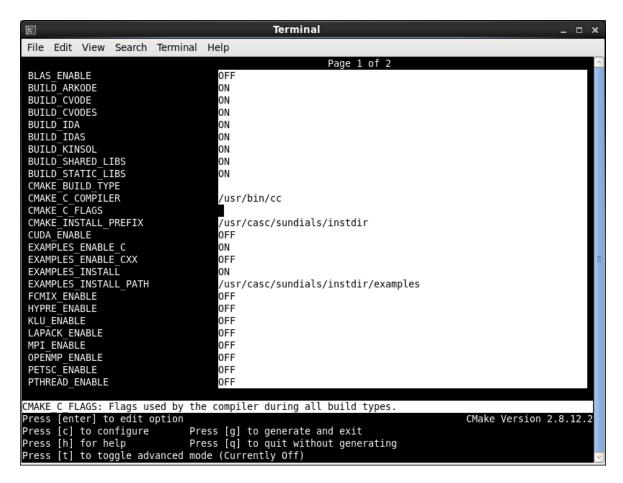


Figure A.2: Changing the *instdir* for SUNDIALS and corresponding examples

## Building from the command line

Using CMake from the command line is simply a matter of specifying CMake variable settings with the cmake command. The following will build the default configuration:

```
% cmake -DCMAKE_INSTALL_PREFIX=/home/myname/sundials/instdir \
> -DEXAMPLES_INSTALL_PATH=/home/myname/sundials/instdir/examples \
> ../solverdir
% make
% make install
```

## A.1.2 Configuration options (Unix/Linux)

A complete list of all available options for a CMake-based SUNDIALS configuration is provide below. Note that the default values shown are for a typical configuration on a Linux system and are provided as illustration only.

```
BLAS_ENABLE - Enable BLAS support
Default: OFF
```

Note: Setting this option to ON will trigger additional CMake options. See additional information on building with BLAS enabled in A.1.4.

```
BLAS_LIBRARIES - BLAS library
Default: /usr/lib/libblas.so
```

Note: CMake will search for libraries in your LD\_LIBRARY\_PATH prior to searching default system paths.

BUILD\_ARKODE - Build the ARKODE library

Default: ON

BUILD\_CVODE - Build the CVODE library

Default: ON

BUILD\_CVODES - Build the CVODES library

Default: ON

BUILD\_IDA - Build the IDA library

Default: ON

BUILD\_IDAS - Build the IDAS library

Default: ON

BUILD\_KINSOL - Build the KINSOL library

Default: ON

BUILD\_SHARED\_LIBS - Build shared libraries

Default: ON

BUILD\_STATIC\_LIBS - Build static libraries

Default: ON

CMAKE\_BUILD\_TYPE - Choose the type of build, options are: None (CMAKE\_C\_FLAGS used), Debug, Release, RelWithDebInfo, and MinSizeRel

Default:

Note: Specifying a build type will trigger the corresponding build type specific compiler flag options below which will be appended to the flags set by CMAKE\_<language>\_FLAGS.

 ${\tt CMAKE\_C\_COMPILER\ -\ C\ compiler}$ 

Default: /usr/bin/cc

CMAKE\_C\_FLAGS - Flags for C compiler

Default:

CMAKE\_C\_FLAGS\_DEBUG - Flags used by the C compiler during debug builds

Default: -g

CMAKE\_C\_FLAGS\_MINSIZEREL - Flags used by the C compiler during release minsize builds

Default: -Os -DNDEBUG

CMAKE\_C\_FLAGS\_RELEASE - Flags used by the C compiler during release builds

Default: -O3 -DNDEBUG

CMAKE\_CXX\_COMPILER - C++ compiler

Default: /usr/bin/c++

Note: A C++ compiler (and all related options) are only triggered if C++ examples are enabled (EXAMPLES\_ENABLE\_CXX is ON). All SUNDIALS solvers can be used from C++ applications by default without setting any additional configuration options.

CMAKE\_CXX\_FLAGS - Flags for C++ compiler

Default:

 ${\tt CMAKE\_CXX\_FLAGS\_DEBUG~- Flags~used~by~the~C++~compiler~during~debug~builds}$ 

Default: -g

CMAKE\_CXX\_FLAGS\_MINSIZEREL - Flags used by the C++ compiler during release minsize builds Default: -Os -DNDEBUG

### CMAKE\_Fortran\_COMPILER - Fortran compiler

Default: /usr/bin/gfortran

Note: Fortran support (and all related options) are triggered only if either Fortran-C support is enabled (FCMIX\_ENABLE is ON) or BLAS/LAPACK support is enabled (BLAS\_ENABLE or LAPACK\_ENABLE is ON).

## CMAKE\_Fortran\_FLAGS - Flags for Fortran compiler

Default:

## 

CMAKE\_Fortran\_FLAGS\_MINSIZEREL - Flags used by the Fortran compiler during release minsize builds Default: -Os

## $\begin{tabular}{ll} $\tt CMAKE\_Fortran\_FLAGS\_RELEASE - Flags used by the Fortran compiler during release builds \\ Default: -O3 \\ \end{tabular}$

## CMAKE\_INSTALL\_PREFIX - Install path prefix, prepended onto install directories

Default: /usr/local

Note: The user must have write access to the location specified through this option. Exported SUNDIALS header files and libraries will be installed under subdirectories include and CMAKE\_INSTALL\_LIBDIR of CMAKE\_INSTALL\_PREFIX, respectively.

### CMAKE\_INSTALL\_LIBDIR - Library installation directory

Default:

Note: This is the directory within CMAKE\_INSTALL\_PREFIX that the SUNDIALS libraries will be installed under. The default is automatically set based on the operating system using the GNUInstallDirs CMake module.

## Fortran\_INSTALL\_MODDIR - Fortran module installation directory

Default: fortran

## CUDA\_ENABLE - Build the SUNDIALS CUDA vector module.

Default: OFF

## EXAMPLES\_ENABLE\_C - Build the SUNDIALS C examples

Default: ON

## ${\tt EXAMPLES\_ENABLE\_CUDA}$ - ${\tt Build}$ the ${\tt SUNDIALS}$ ${\tt CUDA}$ examples

Default: OFF

Note: You need to enable CUDA support to build these examples.

## EXAMPLES\_ENABLE\_CXX - Build the SUNDIALS C++ examples

Default: OFF

## EXAMPLES\_ENABLE\_RAJA - Build the SUNDIALS RAJA examples

Default: OFF

Note: You need to enable CUDA and RAJA support to build these examples.

## EXAMPLES\_ENABLE\_F77 - Build the SUNDIALS Fortran77 examples

Default: ON (if F77\_INTERFACE\_ENABLE is ON)

EXAMPLES\_ENABLE\_F90 - Build the SUNDIALS Fortran90/Fortran2003 examples

Default: ON (if F77\_INTERFACE\_ENABLE or F2003\_INTERFACE\_ENABLE is ON)

EXAMPLES\_INSTALL - Install example files

Default: ON

Note: This option is triggered when any of the SUNDIALS example programs are enabled (EXAMPLES\_ENABLE\_<language> is ON). If the user requires installation of example programs then the sources and sample output files for all SUNDIALS modules that are currently enabled will be exported to the directory specified by EXAMPLES\_INSTALL\_PATH. A CMake configuration script will also be automatically generated and exported to the same directory. Additionally, if the configuration is done under a Unix-like system, makefiles for the compilation of the example programs (using the installed SUNDIALS libraries) will be automatically generated and exported to the directory specified by EXAMPLES\_INSTALL\_PATH.

EXAMPLES\_INSTALL\_PATH - Output directory for installing example files

Default: /usr/local/examples

Note: The actual default value for this option will be an examples subdirectory created under CMAKE\_INSTALL\_PREFIX.

F77\_INTERFACE\_ENABLE - Enable Fortran-C support via the Fortran 77 interfaces

Default: OFF

F2003\_INTERFACE\_ENABLE - Enable Fortran-C support via the Fortran 2003 interfaces

Default: OFF

HYPRE\_ENABLE - Enable hypre support

Default: OFF

Note: See additional information on building with hypre enabled in A.1.4.

HYPRE\_INCLUDE\_DIR - Path to hypre header files

HYPRE\_LIBRARY\_DIR - Path to hypre installed library files

KLU\_ENABLE - Enable KLU support

Default: OFF

Note: See additional information on building with KLU enabled in A.1.4.

KLU\_INCLUDE\_DIR - Path to SuiteSparse header files

KLU\_LIBRARY\_DIR - Path to SuiteSparse installed library files

LAPACK\_ENABLE - Enable LAPACK support

Default: OFF

Note: Setting this option to ON will trigger additional CMake options. See additional information on building with LAPACK enabled in A.1.4.

LAPACK\_LIBRARIES - LAPACK (and BLAS) libraries

Default: /usr/lib/liblapack.so;/usr/lib/libblas.so

Note: CMake will search for libraries in your LD\_LIBRARY\_PATH prior to searching default system paths.

MPI\_ENABLE - Enable MPI support (build the parallel nvector).

Default: OFF

Note: Setting this option to ON will trigger several additional options related to MPI.

MPI\_C\_COMPILER - mpicc program

Default:

MPI\_CXX\_COMPILER - mpicxx program

Default:

Note: This option is triggered only if MPI is enabled (MPI\_ENABLE is ON) and C++ examples are enabled (EXAMPLES\_ENABLE\_CXX is ON). All SUNDIALS solvers can be used from C++ MPI applications by default without setting any additional configuration options other than MPI\_ENABLE.

MPI\_Fortran\_COMPILER - mpif77 or mpif90 program

Default:

Note: This option is triggered only if MPI is enabled (MPI\_ENABLE is ON) and Fortran-C support is enabled (F77\_INTERFACE\_ENABLE or F2003\_INTERFACE\_ENABLE is ON).

MPIEXEC\_EXECUTABLE - Specify the executable for running MPI programs

Default: mpirun

Note: This option is triggered only if MPI is enabled (MPI\_ENABLE is ON).

OPENMP\_ENABLE - Enable OpenMP support (build the OpenMP nvector).

Default: OFF

OPENMP\_DEVICE\_ENABLE - Enable OpenMP device offloading (build the OpenMPDEV nvector) if supported by the provided compiler.

Default: OFF

SKIP\_OPENMP\_DEVICE\_CHECK - advanced option - Skip the check done to see if the OpenMP provided by the compiler supports OpenMP device offloading.

Default: OFF

PETSC\_ENABLE - Enable PETSc support

Default: OFF

Note: See additional information on building with PETSc enabled in A.1.4.

PETSC\_INCLUDE\_DIR - Path to PETSc header files

PETSC\_LIBRARY\_DIR - Path to PETSc installed library files

PTHREAD\_ENABLE - Enable Pthreads support (build the Pthreads nvector).

Default: OFF

RAJA\_ENABLE - Enable RAJA support (build the RAJA nvector).

Default: OFF

Note: You need to enable CUDA in order to build the RAJA vector module.

SUNDIALS\_F77\_FUNC\_CASE - advanced option - Specify the case to use in the Fortran name-mangling scheme, options are: lower or upper

Default:

Note: The build system will attempt to infer the Fortran name-mangling scheme using the Fortran compiler. This option should only be used if a Fortran compiler is not available or to override the inferred or default (lower) scheme if one can not be determined. If used, SUNDIALS\_F77\_FUNC\_UNDERSCORES must also be set.

SUNDIALS\_F77\_FUNC\_UNDERSCORES - advanced option - Specify the number of underscores to append in the Fortran name-mangling scheme, options are: none, one, or two

Note: The build system will attempt to infer the Fortran name-mangling scheme using the Fortran compiler. This option should only be used if a Fortran compiler is not available or to override the inferred or default (one) scheme if one can not be determined. If used, SUNDIALS\_F77\_FUNC\_CASE must also be set.

SUNDIALS\_INDEX\_TYPE - advanced option - Integer type used for SUNDIALS indices. The size must match the size provided for the

SUNDIALS\_INDEX\_SIZE option.

Default:

Note: In past SUNDIALS versions, a user could set this option to INT64\_T to use 64-bit integers, or INT32\_T to use 32-bit integers. Starting in SUNDIALS 3.2.0, these special values are deprecated. For SUNDIALS 3.2.0 and up, a user will only need to use the SUNDIALS\_INDEX\_SIZE option in most cases.

SUNDIALS\_INDEX\_SIZE - Integer size (in bits) used for indices in SUNDIALS, options are: 32 or 64

Default: 64

Note: The build system tries to find an integer type of appropriate size. Candidate 64-bit integer types are (in order of preference): int64\_t, \_\_int64, long long, and long. Candidate 32-bit integers are (in order of preference): int32\_t, int, and long. The advanced option, SUNDIALS\_INDEX\_TYPE can be used to provide a type not listed here.

SUNDIALS\_PRECISION - Precision used in SUNDIALS, options are: double, single, or extended Default: double

SUPERLUMT\_ENABLE - Enable SuperLU\_MT support

Default: OFF

Note: See additional information on building with SuperLU\_MT enabled in A.1.4.

SUPERLUMT\_INCLUDE\_DIR - Path to SuperLU\_MT header files (typically SRC directory)

SUPERLUMT\_LIBRARY\_DIR - Path to SuperLU\_MT installed library files

SUPERLUMT\_THREAD\_TYPE - Must be set to Pthread or OpenMP

Default: Pthread

USE\_GENERIC\_MATH - Use generic (stdc) math libraries

Default: ON

## **xSDK** Configuration Options

SUNDIALS supports CMake configuration options defined by the Extreme-scale Scientific Software Development Kit (xSDK) community policies (see https://xsdk.info for more information). xSDK CMake options are unused by default but may be activated by setting USE\_XSDK\_DEFAULTS to ON.

When xSDK options are active, they will overwrite the corresponding SUNDIALS option and may have different default values (see details below). As such the equivalent SUNDIALS options should not be used when configuring with xSDK options. In the GUI front end to CMake (ccmake), setting USE\_XSDK\_DEFAULTS to ON will hide the corresponding SUNDIALS options as advanced CMake variables. During configuration, messages are output detailing which xSDK flags are active and the equivalent SUNDIALS options that are replaced. Below is a complete list xSDK options and the corresponding SUNDIALS options if applicable.

## TPL\_BLAS\_LIBRARIES - BLAS library

Default: /usr/lib/libblas.so

SUNDIALS equivalent: BLAS\_LIBRARIES

Note: CMake will search for libraries in your LD\_LIBRARY\_PATH prior to searching default system paths.

TPL\_ENABLE\_BLAS - Enable BLAS support

Default: OFF

SUNDIALS equivalent: BLAS\_ENABLE

TPL\_ENABLE\_HYPRE - Enable hypre support

Default: OFF

SUNDIALS equivalent: HYPRE\_ENABLE



TPL\_ENABLE\_KLU - Enable KLU support

Default: OFF

SUNDIALS equivalent: KLU\_ENABLE

TPL\_ENABLE\_PETSC - Enable PETSc support

Default: OFF

SUNDIALS equivalent: PETSC\_ENABLE

TPL\_ENABLE\_LAPACK - Enable LAPACK support

Default: OFF

SUNDIALS equivalent: LAPACK\_ENABLE

TPL\_ENABLE\_SUPERLUMT - Enable SuperLU\_MT support

Default: OFF

SUNDIALS equivalent: SUPERLUMT\_ENABLE

 ${\tt TPL\_HYPRE\_INCLUDE\_DIRS}$  - Path to hypre header files

SUNDIALS equivalent: HYPRE\_INCLUDE\_DIR

TPL\_HYPRE\_LIBRARIES - hypre library

SUNDIALS equivalent: N/A

 ${\tt TPL\_KLU\_INCLUDE\_DIRS}$  - Path to KLU header files

SUNDIALS equivalent: KLU\_INCLUDE\_DIR

TPL\_KLU\_LIBRARIES - KLU library

SUNDIALS equivalent: N/A

TPL\_LAPACK\_LIBRARIES - LAPACK (and BLAS) libraries

Default: /usr/lib/liblapack.so;/usr/lib/libblas.so

SUNDIALS equivalent: LAPACK\_LIBRARIES

Note: CMake will search for libraries in your LD\_LIBRARY\_PATH prior to searching default system

paths.

TPL\_PETSC\_INCLUDE\_DIRS - Path to PETSc header files

SUNDIALS equivalent: PETSC\_INCLUDE\_DIR

TPL\_PETSC\_LIBRARIES - PETSc library

SUNDIALS equivalent: N/A

TPL\_SUPERLUMT\_INCLUDE\_DIRS - Path to SuperLU\_MT header files

SUNDIALS equivalent: SUPERLUMT\_INCLUDE\_DIR

TPL\_SUPERLUMT\_LIBRARIES - SuperLU\_MT library

SUNDIALS equivalent: N/A

 ${\tt TPL\_SUPERLUMT\_THREAD\_TYPE} \ - \ {\tt SuperLU\_MT} \ library \ thread \ type$ 

 ${\tt SUNDIALS\ equivalent:\ SUPERLUMT\_THREAD\_TYPE}$ 

USE\_XSDK\_DEFAULTS - Enable xSDK default configuration settings

Default: OFF

SUNDIALS equivalent: N/A

Note: Enabling xSDK defaults also sets CMAKE\_BUILD\_TYPE to Debug

XSDK\_ENABLE\_FORTRAN - Enable SUNDIALS Fortran interfaces

Default: OFF

SUNDIALS equivalent: F77\_INTERFACE\_ENABLE/F2003\_INTERFACE\_ENABLE

```
    XSDK_INDEX_SIZE - Integer size (bits) used for indices in SUNDIALS, options are: 32 or 64
        Default: 32
        SUNDIALS equivalent: SUNDIALS_INDEX_SIZE
    XSDK_PRECISION - Precision used in SUNDIALS, options are: double, single, or quad
        Default: double
        SUNDIALS equivalent: SUNDIALS_PRECISION
```

## A.1.3 Configuration examples

The following examples will help demonstrate usage of the CMake configure options.

To configure SUNDIALS using the default C and Fortran compilers, and default mpic and mpif77 parallel compilers, enable compilation of examples, and install libraries, headers, and example sources under subdirectories of /home/myname/sundials/, use:

```
% cmake \
> -DCMAKE_INSTALL_PREFIX=/home/myname/sundials/instdir \
> -DEXAMPLES_INSTALL_PATH=/home/myname/sundials/instdir/examples \
> -DMPI_ENABLE=ON \
> -DFCMIX_ENABLE=ON \
> /home/myname/sundials/solverdir
%
% make install
%
```

To disable installation of the examples, use:

```
% cmake \
> -DCMAKE_INSTALL_PREFIX=/home/myname/sundials/instdir \
> -DEXAMPLES_INSTALL_PATH=/home/myname/sundials/instdir/examples \
> -DMPI_ENABLE=ON \
> -DFCMIX_ENABLE=ON \
> -DEXAMPLES_INSTALL=OFF \
> /home/myname/sundials/solverdir
%
% make install
%
```

## A.1.4 Working with external Libraries

The SUNDIALS suite contains many options to enable implementation flexibility when developing solutions. The following are some notes addressing specific configurations when using the supported third party libraries. When building SUNDIALS as a shared library external libraries any used with SUNDIALS must also be build as a shared library or as a static library compiled with the -fPIC flag.



#### Building with BLAS

SUNDIALS does not utilize BLAS directly but it may be needed by other external libraries that SUNDIALS can be built with (e.g. LAPACK, PETSc, SuperLU\_MT, etc.). To enable BLAS, set the BLAS\_ENABLE option to ON. If the directory containing the BLAS library is in the LD\_LIBRARY\_PATH environment variable, CMake will set the BLAS\_LIBRARIES variable accordingly, otherwise CMake will attempt to find the BLAS library in standard system locations. To explicitly tell CMake what libraries to use, the BLAS\_LIBRARIES variable can be set to the desired library. Example:

```
% cmake \
> -DCMAKE_INSTALL_PREFIX=/home/myname/sundials/instdir \
```

```
> -DEXAMPLES_INSTALL_PATH=/home/myname/sundials/instdir/examples \
> -DBLAS_ENABLE=ON \
> -DBLAS_LIBRARIES=/myblaspath/lib/libblas.so \
> -DSUPERLUMT_ENABLE=ON \
> -DSUPERLUMT_INCLUDE_DIR=/mysuperlumtpath/SRC
> -DSUPERLUMT_LIBRARY_DIR=/mysuperlumtpath/lib
> /home/myname/sundials/solverdir
%
% make install
%
```



When allowing CMake to automatically locate the LAPACK library, CMake may also locate the corresponding BLAS library.

If a working Fortran compiler is not available to infer the Fortran name-mangling scheme, the options SUNDIALS\_F77\_FUNC\_CASE and SUNDIALS\_F77\_FUNC\_UNDERSCORES must be set in order to bypass the check for a Fortran compiler and define the name-mangling scheme. The defaults for these options in earlier versions of SUNDIALS were lower and one respectively.

## Building with LAPACK

To enable LAPACK, set the LAPACK\_ENABLE option to ON. If the directory containing the LAPACK library is in the LD\_LIBRARY\_PATH environment variable, CMake will set the LAPACK\_LIBRARIES variable accordingly, otherwise CMake will attempt to find the LAPACK library in standard system locations. To explicitly tell CMake what library to use, the LAPACK\_LIBRARIES variable can be set to the desired libraries. When setting the LAPACK location explicitly the location of the corresponding BLAS library will also need to be set. Example:

```
% cmake \
> -DCMAKE_INSTALL_PREFIX=/home/myname/sundials/instdir \
> -DEXAMPLES_INSTALL_PATH=/home/myname/sundials/instdir/examples \
> -DBLAS_ENABLE=ON \
> -DBLAS_LIBRARIES=/mylapackpath/lib/libblas.so \
> -DLAPACK_ENABLE=ON \
> -DLAPACK_LIBRARIES=/mylapackpath/lib/liblapack.so \
> /home/myname/sundials/solverdir
%
make install
```



%

When allowing CMake to automatically locate the LAPACK library, CMake may also locate the corresponding BLAS library.

If a working Fortran compiler is not available to infer the Fortran name-mangling scheme, the options SUNDIALS\_F77\_FUNC\_CASE and SUNDIALS\_F77\_FUNC\_UNDERSCORES must be set in order to bypass the check for a Fortran compiler and define the name-mangling scheme. The defaults for these options in earlier versions of SUNDIALS were lower and one respectively.

## Building with KLU

The KLU libraries are part of SuiteSparse, a suite of sparse matrix software, available from the Texas A&M University website: http://faculty.cse.tamu.edu/davis/suitesparse.html. SUNDIALS has been tested with SuiteSparse version 4.5.3. To enable KLU, set KLU\_ENABLE to ON, set KLU\_INCLUDE\_DIR to the include path of the KLU installation and set KLU\_LIBRARY\_DIR to the lib path of the KLU installation. The CMake configure will result in populating the following variables: AMD\_LIBRARY, AMD\_LIBRARY\_DIR, BTF\_LIBRARY\_DIR, COLAMD\_LIBRARY, COLAMD\_LIBRARY\_DIR, and KLU\_LIBRARY.

## Building with SuperLU\_MT

The SuperLU\_MT libraries are available for download from the Lawrence Berkeley National Laboratory website: http://crd-legacy.lbl.gov/~xiaoye/SuperLU/#superlu\_mt. SUNDIALS has been tested with SuperLU\_MT version 3.1. To enable SuperLU\_MT, set SUPERLUMT\_ENABLE to ON, set SUPERLUMT\_INCLUDE\_DIR to the SRC path of the SuperLU\_MT installation, and set the variable SUPERLUMT\_LIBRARY\_DIR to the lib path of the SuperLU\_MT installation. At the same time, the variable SUPERLUMT\_THREAD\_TYPE must be set to either Pthread or OpenMP.

Do not mix thread types when building SUNDIALS solvers. If threading is enabled for SUNDIALS by having either OPENMP\_ENABLE or PTHREAD\_ENABLE set to ON then SuperLU\_MT should be set to use the same threading type.



#### Building with PETSc

The PETSc libraries are available for download from the Argonne National Laboratory website: http://www.mcs.anl.gov/petsc. SUNDIALS has been tested with PETSc version 3.7.2. To enable PETSc, set PETSC\_ENABLE to ON, set PETSC\_INCLUDE\_DIR to the include path of the PETSc installation, and set the variable PETSC\_LIBRARY\_DIR to the lib path of the PETSc installation.

### Building with hypre

The hypre libraries are available for download from the Lawrence Livermore National Laboratory website: http://computation.llnl.gov/projects/hypre. SUNDIALS has been tested with hypre version 2.11.1. To enable hypre, set HYPRE\_ENABLE to ON, set HYPRE\_INCLUDE\_DIR to the include path of the hypre installation, and set the variable HYPRE\_LIBRARY\_DIR to the lib path of the hypre installation.

## Building with CUDA

SUNDIALS CUDA modules and examples have been tested with version 8.0 of the CUDA toolkit. To build them, you need to install the Toolkit and compatible NVIDIA drivers. Both are available for download from the NVIDIA website: https://developer.nvidia.com/cuda-downloads. To enable CUDA, set CUDA\_ENABLE to ON. If CUDA is installed in a nonstandard location, you may be prompted to set the variable CUDA\_TOOLKIT\_ROOT\_DIR with your CUDA Toolkit installation path. To enable CUDA examples, set EXAMPLES\_ENABLE\_CUDA to ON.

### Building with RAJA

RAJA is a performance portability layer developed by Lawrence Livermore National Laboratory and can be obtained from https://github.com/LLNL/RAJA. SUNDIALS RAJA modules and examples have been tested with RAJA version 0.3. Building SUNDIALS RAJA modules requires a CUDA-enabled RAJA installation. To enable RAJA, set CUDA\_ENABLE and RAJA\_ENABLE to ON. If RAJA is installed in a nonstandard location you will be prompted to set the variable RAJA\_DIR with the path to the RAJA CMake configuration file. To enable building the RAJA examples set EXAMPLES\_ENABLE\_RAJA to ON.

## A.1.5 Testing the build and installation

If SUNDIALS was configured with EXAMPLES\_ENABLE\_<language> options to ON, then a set of regression tests can be run after building with the make command by running:

% make test

Additionally, if EXAMPLES\_INSTALL was also set to ON, then a set of smoke tests can be run after installing with the make install command by running:

% make test\_install

## A.2 Building and Running Examples

Each of the SUNDIALS solvers is distributed with a set of examples demonstrating basic usage. To build and install the examples, set at least of the EXAMPLES\_ENABLE\_<language> options to ON, and set EXAMPLES\_INSTALL to ON. Specify the installation path for the examples with the variable EXAMPLES\_INSTALL\_PATH. CMake will generate CMakeLists.txt configuration files (and Makefile files if on Linux/Unix) that reference the *installed* SUNDIALS headers and libraries.

Either the CMakeLists.txt file or the traditional Makefile may be used to build the examples as well as serve as a template for creating user developed solutions. To use the supplied Makefile simply run make to compile and generate the executables. To use CMake from within the installed example directory, run cmake (or ccmake to use the GUI) followed by make to compile the example code. Note that if CMake is used, it will overwrite the traditional Makefile with a new CMake-generated Makefile. The resulting output from running the examples can be compared with example output bundled in the SUNDIALS distribution.



NOTE: There will potentially be differences in the output due to machine architecture, compiler versions, use of third party libraries etc.

## A.3 Configuring, building, and installing on Windows

CMake can also be used to build SUNDIALS on Windows. To build SUNDIALS for use with Visual Studio the following steps should be performed:

- 1. Unzip the downloaded tar file(s) into a directory. This will be the solverdir
- 2. Create a separate builddir
- 3. Open a Visual Studio Command Prompt and cd to builddir
- 4. Run cmake-gui ../solverdir
  - (a) Hit Configure
  - (b) Check/Uncheck solvers to be built
  - (c) Change CMAKE\_INSTALL\_PREFIX to instdir
  - (d) Set other options as desired
  - (e) Hit Generate
- 5. Back in the VS Command Window:
  - (a) Run msbuild ALL\_BUILD.vcxproj
  - (b) Run msbuild INSTALL.vcxproj

The resulting libraries will be in the *instdir*. The SUNDIALS project can also now be opened in Visual Studio. Double click on the ALL\_BUILD.vcxproj file to open the project. Build the whole *solution* to create the SUNDIALS libraries. To use the SUNDIALS libraries in your own projects, you must set the include directories for your project, add the SUNDIALS libraries to your project solution, and set the SUNDIALS libraries as dependencies for your project.

## A.4 Installed libraries and exported header files

Using the CMake SUNDIALS build system, the command

will install the libraries under *libdir* and the public header files under *includedir*. The values for these directories are *instdir*/CMAKE\_INSTALL\_LIBDIR and *instdir*/include, respectively. The location can be changed by setting the CMake variable CMAKE\_INSTALL\_PREFIX. Although all installed libraries reside under *libdir*/CMAKE\_INSTALL\_LIBDIR, the public header files are further organized into subdirectories under *includedir*/include.

The installed libraries and exported header files are listed for reference in Table A.1. The file extension .lib is typically .so for shared libraries and .a for static libraries. Note that, in the Tables, names are relative to libraries and to includedir for header files.

A typical user program need not explicitly include any of the shared SUNDIALS header files from under the <code>includedir/include/sundials</code> directory since they are explicitly included by the appropriate solver header files (e.g., <code>cvode\_dense.h</code> includes <code>sundials\_dense.h</code>). However, it is both legal and safe to do so, and would be useful, for example, if the functions declared in <code>sundials\_dense.h</code> are to be used in building a preconditioner.

Table A.1: SUNDIALS libraries and header files

	Libraries	ALS libraries and header files			
SHARED	Header files	n/a			
	Header files	sundials/sundials_config.h			
		sundials/sundials_fconfig.h			
		sundials/sundials_types.h			
		sundials/sundials_math.h			
		sundials/sundials_nvector.h			
		sundials/sundials_fnvector.h			
		sundials/sundials_matrix.h			
		sundials/sundials_linearsolver.h			
		sundials/sundials_iterative.h			
		sundials/sundials_direct.h			
		sundials/sundials_dense.h			
		sundials/sundials_band.h			
		sundials/sundials_nonlinearsolver.h			
		sundials/sundials_version.h			
		sundials/sundials_mpi_types.h			
NVECTOR_SERIAL	Libraries	libsundials_nvecserial.lib			
		$libsundials\_fnvecserial\_mod.lib$			
		libsundials_fnvecserial.a			
	Header files	nvector/nvector_serial.h			
	Module	fnvector_serial_mod.mod			
	files				
NVECTOR_PARALLEL	Libraries	$lib sundials\_nvec parallel. \\ lib sundials\_fnvec parallel. \\ a$			
	Header files	nvector/nvector_parallel.h			
NVECTOR_OPENMP	Libraries	$libsundials\_nvecopenmp. lib$			
		${\it libsundials\_fnvecopenmp\_mod.} lib$			
		libsundials_fnvecopenmp.a			
	Header files	nvector/nvector_openmp.h			
	Module	fnvector_openmp_mod.mod			
	files				
NVECTOR_OPENMPDEV	Libraries	$libsundials\_nvecopenmpdev. \it lib$			
	Header files	nvector/nvector_openmpdev.h			
NVECTOR_PTHREADS	Libraries	$libsundials\_nvecpthreads. \\ lib$			
		$libsundials\_fnvecpthreads\_mod. \it lib$			
		libsundials_fnvecpthreads.a			
	Header files	nvector/nvector_pthreads.h			
	Module	fnvector_pthreads_mod.mod			
	files				
NVECTOR_PARHYP	Libraries	libsundials_nvecparhyp.lib			
	Header files	nvector/nvector_parhyp.h			
		continued on next page			

NVECTOR_PETSC	Libraries	libsundials_nvecpetsc.lib
Bereiter Bibe	Header files	nvector/nvector_petsc.h
NVECTOR_CUDA	Libraries	libsundials_nveccuda.lib
IV ECTORECODA	Libraries	libsundials_nvecmpicuda.lib
	Header files	nvector/nvector_cuda.h
	Trouder mes	nvector/nvector_mpicuda.h
		nvector/cuda/ThreadPartitioning.hpp
		nvector/cuda/Vector.hpp
		nvector/cuda/VectorKernels.cuh
NVECTOR_RAJA	Libraries	libsundials_nveccudaraja.lib
NVECTOR_RAJA	Libraries	libsundials_nveccudampiraja.lib
	Header files	nvector/nvector_raja.h
	Header mes	nvector/nvector_mpiraja.h
		nvector/raja/Vector.hpp
SUNMATRIX_BAND	Libraries	libsundials_sunmatrixband.lib
SUNMATRIX_BAND	Libraries	libsundials_fsunmatrixband_mod.lib
		libsundials_fsunmatrixband.a
	Header files	
	Module  Module	sunmatrix/sunmatrix_band.h fsunmatrix_band_mod
		isunmatrix_band_mod.mod
	files	111 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
SUNMATRIX_DENSE	Libraries	libsundials_sunmatrixdense.lib
		libsundials_fsunmatrixdense_mod.lib
	TT 1 01	libsundials_fsunmatrixdense.a
	Header files	sunmatrix/sunmatrix_dense.h
	Module	fsunmatrix_dense_mod.mod
	files	
SUNMATRIX_SPARSE	Libraries	$libsundials\_sunmatrixsparse.lib$
		$libsundials\_fsunmatrixsparse\_mod. lib$
		libsundials_fsunmatrixsparse.a
	Header files	sunmatrix_sparse.h
	Module	fsunmatrix_sparse_mod.mod
	files	
SUNLINSOL_BAND	Libraries	libsundials_sunlinsolband.lib
		$libsundials\_fsunlinsolband\_mod.lib$
		libsundials_fsunlinsolband.a
	Header files	sunlinsol/sunlinsol_band.h
	Module	fsunlinsol_band_mod.mod
	files	
SUNLINSOL_DENSE	Libraries	libsundials_sunlinsoldense.lib
		libsundials_fsunlinsoldense_mod.lib
		libsundials_fsunlinsoldense.a
	Header files	sunlinsol/sunlinsol_dense.h

continued from last page		
	Module	fsunlinsol_dense_mod.mod
	files	
SUNLINSOL_KLU	Libraries	libsundials_sunlinsolklu.lib
		libsundials_fsunlinsolklu_mod.lib
		libsundials_fsunlinsolklu.a
	Header files	sunlinsol/sunlinsol_klu.h
	Module	fsunlinsol_klu_mod.mod
	files	
SUNLINSOL_LAPACKBAND	Libraries	$lib sundials\_sunlin sollapack band. \\ lib$
		libsundials_fsunlinsollapackband.a
	Header files	sunlinsol/sunlinsol_lapackband.h
SUNLINSOL_LAPACKDENSE	Libraries	libsundials_sunlinsollapackdense.lib
		libsundials_fsunlinsollapackdense.a
	Header files	sunlinsol/sunlinsol_lapackdense.h
SUNLINSOL_PCG	Libraries	libsundials_sunlinsolpcg.lib
		libsundials_fsunlinsolpcg_mod.lib
		libsundials_fsunlinsolpcg.a
	Header files	sunlinsol/sunlinsol_pcg.h
	Module	fsunlinsol_pcg_mod.mod
	files	
SUNLINSOL_SPBCGS	Libraries	$libsundials\_sunlinsolspbcgs. lib$
		libsundials_fsunlinsolspbcgs_mod.lib
		libsundials_fsunlinsolspbcgs.a
	Header files	sunlinsol/sunlinsol_spbcgs.h
	Module	fsunlinsol_spbcgs_mod.mod
	files	
SUNLINSOL_SPFGMR	Libraries	libsundials_sunlinsolspfgmr.lib
		libsundials_fsunlinsolspfgmr_mod.lib
	TT 1 01	libsundials_fsunlinsolspfgmr.a
	Header files	sunlinsol/sunlinsol_spfgmr.h
	Module	fsunlinsol_spfgmr_mod.mod
	files	111 11 11 11 11
SUNLINSOL_SPGMR	Libraries	libsundials_sunlinsolspgmr.lib
		libsundials_fsunlinsolspgmr_mod.lib
	TT 1 C1	libsundials_fsunlinsolspgmr.a
	Header files	sunlinsol/sunlinsol_spgmr.h
	Module	fsunlinsol_spgmr_mod.mod
GINI INGOL GERESOVE	files	library dialo gunling alantformy 1/1
SUNLINSOL_SPTFQMR	Libraries	libsundials_sunlinsolsptfqmr.lib
		libsundials_fsunlinsolsptfqmr_mod.lib
	Header files	libsundials_fsunlinsolsptfqmr.a
	neader mes	sunlinsol/sunlinsol_sptfqmr.h
		continued on next page

continued from last page				
<del></del>	Module	fsunlinsol_sptfqmr_mod.mo	d	
	files			
SUNLINSOL_SUPERLUMT	Libraries	libsundials_sunlinsolsuperlumt.lib		
		libsundials_fsunlinsolsuperlumt.a		
	Header files	sunlinsol/sunlinsol_superlumt.h		
SUNNONLINSOL_NEWTON	Libraries	libsundials_sunnonlinsolnewton.lib		
		libsundials_fsunnonlinsolne	${ m wton\_mod.} lib$	
		libsundials_fsunnonlinsolne	wton.a	
	Header files	sunnonlinsol/sunnonlinsol_newton.h		
	Module	fsunnonlinsol_newton_mod.mod		
	files			
SUNNONLINSOL_FIXEDPOINT	Libraries	libsundials_sunnonlinsolfixe	edpoint. lib	
		libsundials_fsunnonlinsolfix		
		libsundials_fsunnonlinsolfix	$\operatorname{edpoint} \operatorname{\mod} . lib$	
	Header files	sunnonlinsol/sunnonlinsol_	fixedpoint.h	
	Module	fsunnonlinsol_fixedpoint_me	od.mod	
	files			
CVODE	Libraries	$libsundials\_cvode.lib$	libsundials_fcvode.a	
	Header files	cvode/cvode.h	cvode/cvode_impl.h	
		cvode/cvode_direct.h	$cvode/cvode\_ls.h$	
		cvode/cvode_spils.h	$cvode/cvode\_bandpre.h$	
		$cvode/cvode\_bbdpre.h$		
	Module	fcvode_mod.mod		
	files			
CVODES	Libraries	libsundials_cvodes.lib		
	Header files	cvodes/cvodes.h	$cvodes/cvodes\_impl.h$	
		cvodes/cvodes_direct.h	$cvodes/cvodes\_ls.h$	
		cvodes_spils.h	$cvodes/cvodes\_bandpre.h$	
		cvodes_bbdpre.h		
ARKODE	Libraries	$libsundials\_arkode.lib$	$libsundials\_farkode.a$	
	Header files	arkode/arkode.h	$arkode/arkode\_impl.h$	
		arkode/arkode_ls.h	$arkode/arkode\_bandpre.h$	
		arkode/arkode_bbdpre.h		
IDA	Libraries	libsundials_ida.lib	libsundials_fida.a	
	Header files	ida/ida.h	ida/ida_impl.h	
		ida/ida_direct.h	ida/ida_ls.h	
		ida/ida_spils.h	ida/ida_bbdpre.h	
IDAS	Libraries	libsundials_idas.lib		
	Header files	idas/idas.h	idas/idas_impl.h	
		idas/idas_direct.h	idas/idas_ls.h	
		idas/idas_spils.h	$idas/idas\_bbdpre.h$	
KINSOL	Libraries	libsundials_kinsol.lib	libsundials_fkinsol.a	
			continued on next page	

continued from last page			
	Header files	kinsol/kinsol.h	kinsol/kinsol_impl.h
		kinsol/kinsol_direct.h	kinsol/kinsol_ls.h
		kinsol/kinsol_spils.h	$kinsol/kinsol\_bbdpre.h$

## Appendix B

## **IDAS** Constants

Below we list all input and output constants used by the main solver and linear solver modules, together with their numerical values and a short description of their meaning.

## B.1 IDAS input constants

		IDAS main solver module
IDA_NORMAL	1	Solver returns at specified output time.
IDA_ONE_STEP	2	Solver returns after each successful step.
IDA_SIMULTANEOUS	1	Simultaneous corrector forward sensitivity method.
IDA_STAGGERED	2	Staggered corrector forward sensitivity method.
IDA_CENTERED	1	Central difference quotient approximation $(2^{nd} \text{ order})$ of the sensitivity RHS.
IDA_FORWARD	2	Forward difference quotient approximation $(1^{st} \text{ order})$ of the sensitivity RHS.
IDA_YA_YDP_INIT	1	Compute $y_a$ and $\dot{y}_d$ , given $y_d$ .
IDA_Y_INIT	2	Compute $y$ , given $\dot{y}$ .
	II	DAS adjoint solver module
IDA_HERMITE	1	Use Hermite interpolation.
IDA_POLYNOMIAL	2	Use variable-degree polynomial interpolation.
	Ite	rative linear solver module
PREC_NONE	0	No preconditioning
PREC_LEFT	1	Preconditioning on the left.
MODIFIED_GS	1	Use modified Gram-Schmidt procedure.
CLASSICAL_GS	2	Use classical Gram-Schmidt procedure.

## B.2 IDAS output constants

IDAS main solver module

340 IDAS Constants

IDA_SUCCESS	0	Successful function return.
IDA_TSTOP_RETURN	1	IDASolve succeeded by reaching the specified stopping point.
IDA_ROOT_RETURN	2	IDASolve succeeded by reaching the specified stopping point.  IDASolve succeeded and found one or more roots.
	99	IDASolve succeeded but an unusual situation occurred.
IDA_WARNING IDA_TOO_MUCH_WORK	-1	
IDA_IUU_MUCH_WURK	-1	The solver took mxstep internal steps but could not reach tout.
IDA_TOO_MUCH_ACC	-2	
IDA_IUU_MUCH_ACC	-2	The solver could not satisfy the accuracy demanded by the
TDA EDD EATI	-3	user for some internal step.
IDA_ERR_FAIL	-9	Error test failures occurred too many times during one inter-
IDA_CONV_FAIL	-4	nal time step or minimum step size was reached.
IDA_CONV_FAIL	-4	Convergence test failures occurred too many times during one internal time step or minimum step size was reached.
TDA I TNITE DATI	-5	The linear solver's initialization function failed.
IDA_LINIT_FAIL	-6	
IDA_LSETUP_FAIL	-0	The linear solver's setup function failed in an unrecoverable
TDA I GOLVE EATI	-7	manner.  The linear solver's solve function failed in an unrecoverable
IDA_LSOLVE_FAIL	-1	
TDA DEG EATI	0	manner.
IDA_RES_FAIL	-8	The user-provided residual function failed in an unrecoverable
TDA DED DEC EATI	-9	manner. The user provided residuel function repeatedly returned a re-
IDA_REP_RES_FAIL	-9	The user-provided residual function repeatedly returned a re-
TDA DECIMA CATI	10	coverable error flag, but the solver was unable to recover.
IDA_RTFUNC_FAIL	-10	The rootfinding function failed in an unrecoverable manner.
IDA_CONSTR_FAIL	-11	The inequality constraints were violated and the solver was
TDA BIDOM DEG BATI	10	unable to recover.
IDA_FIRST_RES_FAIL	-12	The user-provided residual function failed recoverably on the
TDA I TNEGRADGII RATI	19	first call.
IDA_LINESEARCH_FAIL	-13	The line search failed.
IDA_NO_RECOVERY	-14	The residual function, linear solver setup function, or linear
		solver solve function had a recoverable failure, but IDACalcIC could not recover.
IDA_NLS_INIT_FAIL	-15	The nonlinear solver's init routine failed.
IDA_NLS_SETUP_FAIL	-16	The nonlinear solver's setup routine failed.
IDA_MEM_NULL	-10 -20	The ida_mem argument was NULL.
IDA_MEM_FAIL	-21	A memory allocation failed.
IDA_ILL_INPUT	-22	One of the function inputs is illegal.
IDA_NO_MALLOC	-23	The IDAS memory was not allocated by a call to IDAInit.
IDA_BAD_EWT	-24	Zero value of some error weight component.
IDA_BAD_K	-25	The k-th derivative is not available.
IDA_BAD_T	-26	The time $t$ is outside the last step taken.
IDA_BAD_DKY	-27	The vector argument where derivative should be stored is
	21	NULL.
IDA_NO_QUAD	-30	Quadratures were not initialized.
IDA_QRHS_FAIL	-31	The user-provided right-hand side function for quadratures
-		failed in an unrecoverable manner.
IDA_FIRST_QRHS_ERR	-32	The user-provided right-hand side function for quadratures
		failed in an unrecoverable manner on the first call.

IDA_REP_QRHS_ERR	-33	The user-provided right-hand side repeatedly returned a recoverable error flag, but the solver was unable to recover.
IDA_NO_SENS	-40	Sensitivities were not initialized.
IDA_SRES_FAIL	-41	The user-provided sensitivity residual function failed in an unrecoverable manner.
IDA_REP_SRES_ERR	-42	The user-provided sensitivity residual function repeatedly returned a recoverable error flag, but the solver was unable to recover.
IDA_BAD_IS	-43	The sensitivity identifier is not valid.
IDA_NO_QUADSENS	-50	Sensitivity-dependent quadratures were not initialized.
IDA_QSRHS_FAIL	-51	The user-provided sensitivity-dependent quadrature right- hand side function failed in an unrecoverable manner.
IDA_FIRST_QSRHS_ERR	-52	The user-provided sensitivity-dependent quadrature right- hand side function failed in an unrecoverable manner on the first call.
IDA_REP_QSRHS_ERR	-53	The user-provided sensitivity-dependent quadrature right- hand side repeatedly returned a recoverable error flag, but the solver was unable to recover.
	ID	AS adjoint solver module
IDA_NO_ADJ	-101	The combined forward-backward problem has not been initialized.
IDA_NO_FWD	-102	IDASolveF has not been previously called.
IDA_NO_BCK	-103	No backward problem was specified.
IDA_BAD_TBO	-104	The desired output for backward problem is outside the interval over which the forward problem was solved.
IDA_REIFWD_FAIL	-105	No checkpoint is available for this hot start.
IDA_FWD_FAIL	-106	IDASolveB failed because IDASolve was unable to store data between two consecutive checkpoints.
IDA_GETY_BADT	-107	Wrong time in interpolation function.
	IDA	ALS linear solver interface
IDALS_SUCCESS	0	Successful function return.
IDALS_MEM_NULL	-1	The ida_mem argument was NULL.
IDALS_LMEM_NULL	-2	The IDALS linear solver has not been initialized.
IDALS_ILL_INPUT	-3	The IDALS solver is not compatible with the current NVECTOR module, or an input value was illegal.
IDALS_MEM_FAIL	-4	A memory allocation request failed.
IDALS_PMEM_NULL	-5	The preconditioner module has not been initialized.
IDALS_JACFUNC_UNRECVR	-6	The Jacobian function failed in an unrecoverable manner.
IDALS_JACFUNC_RECVR	-7	The Jacobian function had a recoverable error.
IDALS_SUNMAT_FAIL	-8	An error occurred with the current Sunmatrix module.
IDALS_SUNLS_FAIL	-9	An error occurred with the current sunlinsol module.

342 IDAS Constants

IDALS_NO_ADJ	-101	The combined forward-backward problem has not been ini-
		tialized.
IDALS_LMEMB_NULL	-102	The linear solver was not initialized for the backward phase.

## **Bibliography**

- [1] KLU Sparse Matrix Factorization Library. http://faculty.cse.tamu.edu/davis/suitesparse.html.
- [2] SuperLU\_MT Threaded Sparse Matrix Factorization Library. http://crd-legacy.lbl.gov/xiaoye/-SuperLU/.
- [3] D. G. Anderson. Iterative procedures for nonlinear integral equations. J. Assoc. Comput. Machinery, 12:547–560, 1965.
- [4] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. SIAM, Philadelphia, Pa, 1996.
- [5] P. N. Brown. A local convergence theory for combined inexact-Newton/finite difference projection methods. SIAM J. Numer. Anal., 24(2):407–434, 1987.
- [6] P. N. Brown and A. C. Hindmarsh. Reduced Storage Matrix Methods in Stiff ODE Systems. J. Appl. Math. & Comp., 31:49–91, 1989.
- [7] P. N. Brown, A. C. Hindmarsh, and L. R. Petzold. Using Krylov Methods in the Solution of Large-Scale Differential-Algebraic Systems. SIAM J. Sci. Comput., 15:1467–1488, 1994.
- [8] P. N. Brown, A. C. Hindmarsh, and L. R. Petzold. Consistent Initial Condition Calculation for Differential-Algebraic Systems. SIAM J. Sci. Comput., 19:1495–1512, 1998.
- [9] P. N. Brown and Y. Saad. Hybrid Krylov Methods for Nonlinear Systems of Equations. SIAM J. Sci. Stat. Comput., 11:450–481, 1990.
- [10] G. D. Byrne. Pragmatic Experiments with Krylov Methods in the Stiff ODE Setting. In J.R. Cash and I. Gladwell, editors, Computational Ordinary Differential Equations, pages 323–356, Oxford, 1992. Oxford University Press.
- [11] G. D. Byrne and A. C. Hindmarsh. User Documentation for PVODE, An ODE Solver for Parallel Computers. Technical Report UCRL-ID-130884, LLNL, May 1998.
- [12] G. D. Byrne and A. C. Hindmarsh. PVODE, An ODE Solver for Parallel Computers. *Intl. J. High Perf. Comput. Apps.*, 13(4):254–365, 1999.
- [13] Y. Cao, S. Li, L. R. Petzold, and R. Serban. Adjoint Sensitivity Analysis for Differential-Algebraic Equations: The Adjoint DAE System and its Numerical Solution. *SIAM J. Sci. Comput.*, 24(3):1076–1089, 2003.
- [14] M. Caracotsios and W. E. Stewart. Sensitivity Analysis of Initial Value Problems with Mixed ODEs and Algebraic Equations. Computers and Chemical Engineering, 9:359–365, 1985.
- [15] S. D. Cohen and A. C. Hindmarsh. CVODE, a Stiff/Nonstiff ODE Solver in C. Computers in Physics, 10(2):138–143, 1996.
- [16] A. M. Collier, A. C. Hindmarsh, R. Serban, and C.S. Woodward. User Documentation for KINSOL v4.0.0. Technical Report UCRL-SM-208116, LLNL, 2018.

344 BIBLIOGRAPHY

[17] T. A. Davis and P. N. Ekanathan. Algorithm 907: KLU, a direct sparse solver for circuit simulation problems. *ACM Trans. Math. Softw.*, 37(3), 2010.

- [18] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact Newton Methods. SIAM J. Numer. Anal., 19:400–408, 1982.
- [19] J. W. Demmel, J. R. Gilbert, and X. S. Li. An asynchronous parallel supernodal algorithm for sparse gaussian elimination. SIAM J. Matrix Analysis and Applications, 20(4):915–952, 1999.
- [20] J. E. Dennis and R. B. Schnabel. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. SIAM, Philadelphia, 1996.
- [21] H. Fang and Y. Saad. Two classes of secant methods for nonlinear acceleration. *Numer. Linear Algebra Appl.*, 16:197–221, 2009.
- [22] W. F. Feehery, J. E. Tolsma, and P. I. Barton. Efficient Sensitivity Analysis of Large-Scale Differential-Algebraic Systems. *Applied Numer. Math.*, 25(1):41–54, 1997.
- [23] R. W. Freund. A Transpose-Free Quasi-Minimal Residual Algorithm for Non-Hermitian Linear Systems. SIAM J. Sci. Comp., 14:470–482, 1993.
- [24] M. R. Hestenes and E. Stiefel. Methods of Conjugate Gradients for Solving Linear Systems. J. Research of the National Bureau of Standards, 49(6):409–436, 1952.
- [25] K. L. Hiebert and L. F. Shampine. Implicitly Defined Output Points for Solutions of ODEs. Technical Report SAND80-0180, Sandia National Laboratories, February 1980.
- [26] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward. SUNDIALS, suite of nonlinear and differential/algebraic equation solvers. ACM Trans. Math. Softw., (31):363–396, 2005.
- [27] A. C. Hindmarsh and R. Serban. User Documentation for CVODE v4.0.0. Technical Report UCRL-SM-208108, LLNL, 2018.
- [28] A. C. Hindmarsh and R. Serban. User Documentation for CVODES v4.0.0. Technical report, LLNL, 2018. UCRL-SM-208111.
- [29] A. C. Hindmarsh, R. Serban, and A. Collier. Example Programs for IDA v4.0.0. Technical Report UCRL-SM-208113, LLNL, 2018.
- [30] A. C. Hindmarsh, R. Serban, and A. Collier. User Documentation for IDA v4.0.0. Technical Report UCRL-SM-208112, LLNL, 2018.
- [31] A. C. Hindmarsh, R. Serban, and D. R. Reynolds. Example Programs for CVODE v4.0.0. Technical report, LLNL, 2018. UCRL-SM-208110.
- [32] A. C. Hindmarsh and A. G. Taylor. PVODE and KINSOL: Parallel Software for Differential and Nonlinear Systems. Technical Report UCRL-ID-129739, LLNL, February 1998.
- [33] C. T. Kelley. Iterative Methods for Solving Linear and Nonlinear Equations. SIAM, Philadelphia, 1995.
- [34] S. Li, L. R. Petzold, and W. Zhu. Sensitivity Analysis of Differential-Algebraic Equations: A Comparison of Methods on a Special Problem. Applied Num. Math., 32:161–174, 2000.
- [35] X. S. Li. An overview of SuperLU: Algorithms, implementation, and user interface. *ACM Trans. Math. Softw.*, 31(3):302–325, September 2005.
- [36] P. A. Lott, H. F. Walker, C. S. Woodward, and U. M. Yang. An accelerated Picard method for nonlinear systems related to variably saturated flow. Adv. Wat. Resour., 38:92–101, 2012.

BIBLIOGRAPHY 345

[37] T. Maly and L. R. Petzold. Numerical Methods and Software for Sensitivity Analysis of Differential-Algebraic Systems. *Applied Numerical Mathematics*, 20:57–79, 1997.

- [38] D.B. Ozyurt and P.I. Barton. Cheap second order directional derivatives of stiff ODE embedded functionals. SIAM J. of Sci. Comp., 26(5):1725–1743, 2005.
- [39] Daniel R. Reynolds. Example Programs for ARKODE v3.0.0. Technical report, Southern Methodist University, 2018.
- [40] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. SIAM J. Sci. Comput., 14(2):461–469, 1993.
- [41] Y. Saad and M. H. Schultz. GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. SIAM J. Sci. Stat. Comp., 7:856–869, 1986.
- [42] R. Serban and A. C. Hindmarsh. CVODES, the sensitivity-enabled ODE solver in SUNDIALS. In Proceedings of the 5th International Conference on Multibody Systems, Nonlinear Dynamics and Control, Long Beach, CA, 2005. ASME.
- [43] R. Serban and A. C. Hindmarsh. Example Programs for IDAS v3.0.0. Technical Report LLNL-TR-437091, LLNL, 2018.
- [44] H. A. Van Der Vorst. Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems. SIAM J. Sci. Stat. Comp., 13:631–644, 1992.
- [45] H. F. Walker and P. Ni. Anderson acceleration for fixed-point iterations. SIAM Jour. Num. Anal., 49(4):1715–1735, 2011.

# Index

adjoint sensitivity analysis	FSUNMASSDENSELINSOLINIT, 251
checkpointing, 22	FSUNMASSKLUINIT, 263
implementation in IDAS, 23, 26	FSUNMASSKLUREINIT, 263
mathematical background, 20–23	FSUNMASSKLUSETORDERING, 264
quadrature evaluation, 144	FSUNMASSLAPACKBANDINIT, 259
residual evaluation, 142, 143	FSUNMASSLAPACKDENSEINIT, 256
sensitivity-dependent quadrature evaluation,	FSUNMASSPCGINIT, 296
145	FSUNMASSPCGSETMAXL, 297
	FSUNMASSPCGSETPRECTYPE, 297
BIG_REAL, 32, 165	FSUNMASSSPBCGSINIT, 285
booleantype, 32	FSUNMASSSPBCGSSETMAXL, 286
	FSUNMASSSPBCGSSETPRECTYPE, 286
eh_data, 73	FSUNMASSSPFGMRINIT, 279
error control	FSUNMASSSPFGMRSETGSTYPE, 279
sensitivity variables, 19	FSUNMASSSPFGMRSETMAXRS, 280
error messages, 45	FSUNMASSSPFGMRSETPRECTYPE, 280
redirecting, 45	FSUNMASSSPGMRINIT, 272
user-defined handler, 47, 73	FSUNMASSSPGMRSETGSTYPE, 273
fnvector_serial_mod, 174	FSUNMASSSPGMRSETMAXRS, 274
forward sensitivity analysis	FSUNMASSSPGMRSETPRECTYPE, 273
absolute tolerance selection, 19	FSUNMASSSPTFQMRINIT, 290
correction strategies, 18, 26, 96, 97	FSUNMASSSPTFQMRSETMAXL, 291
mathematical background, 17–20	FSUNMASSSPTFQMRSETPRECTYPE, 291
residual evaluation, 108	FSUNMASSSUPERLUMTINIT, 267
right hand side evaluation, 20	FSUNMASSUPERLUMTSETORDERING, 268
right-hand side evaluation, 19	${\tt fsunmatrix\_band\_mod}, 227$
FSUNBANDLINSOLINIT, 253	$fsunmatrix\_dense\_mod, 222$
FSUNDENSELINSOLINIT, 251	${\tt fsunmatrix\_sparse\_mod}, 235$
FSUNFIXEDPOINTINIT, 316	FSUNNEWTONINIT, 312
FSUNKLUINIT, 262	fsunnonlinsol_fixedpoint_mod, 316
FSUNKLUREINIT, 263	${\tt fsunnonlinsol\_newton\_mod}, {\tt 312}$
FSUNKLUSETORDERING, 264	FSUNPCGINIT, 296
FSUNLAPACKBANDINIT, 258	FSUNPCGSETMAXL, 297
FSUNLAPACKDENSEINIT, 256	FSUNPCGSETPRECTYPE, 297
fsunlinsol_band_mod, 253	FSUNSPBCGSINIT, 285
fsunlinsol_dense_mod, 251	FSUNSPBCGSSETMAXL, 286
fsunlinsol_klu_mod, 262	FSUNSPBCGSSETPRECTYPE, 285
fsunlinsol_pcg_mod, 296	FSUNSPFGMRINIT, 278
fsunlinsol_spbcgs_mod, 284	FSUNSPFGMRSETGSTYPE, 279
fsunlinsol_spfgmr_mod, 278	FSUNSPFGMRSETMAXRS, 280
fsunlinsol_spgmr_mod, 272	FSUNSPFGMRSETPRECTYPE, 280
fsunlinsol_sptfqmr_mod, 290	FSUNSPGMRINIT, 272
FSUNMASSBANDLINSOLINIT, 254	FSUNSPGMRSETGSTYPE, 273

FSUNSPGMRSETMAXRS, 274	IDA_NO_RECOVERY, 43
FSUNSPGMRSETPRECTYPE, 273	IDA_NO_SENS, 98, 99, 101-103, 105-108, 111, 112,
FSUNSPTFQMRINIT, 290	114
FSUNSPTFQMRSETMAXL, 291	IDA_NORMAL, 44, 122, 126, 133
FSUNSPTFQMRSETPRECTYPE, 291	IDA_ONE_STEP, 44, 122, 126, 133
FSUNSUPERLUMTINIT, 267	IDA_POLYNOMIAL, 125
FSUNSUPERLUMTSETORDERING, 267	IDA_QRHS_FAIL, 82, 86, 118
T BONDOT ENDOTTED FOR DELICING, 201	IDA_QRHSFUNC_FAIL, 144, 145
half-bandwidths, 89	IDA_QSRHS_FAIL, 112
header files, 33, 88	IDA_REIFWD_FAIL, 133
ileader illes, 50, 00	IDA_REP_QRHS_ERR, 82
ida/ida_ls.h, 33	IDA_REP_QSRHS_ERR, 112
IDA_BAD_DKY, 58, 83, 101–103, 113, 114	
	IDA_REP_RES_ERR, 45
IDA_BAD_EWT, 43	IDA_REP_SRES_ERR, 101
IDA_BAD_IS, 102, 103, 114	IDA_RES_FAIL, 43, 45
IDA_BAD_ITASK, 133	IDA_RESFUNC_FAIL, 143, 144
IDA_BAD_K, 58, 83, 102, 103, 113, 114	IDA_ROOT_RETURN, 44, 127
IDA_BAD_T, 58, 83, 102, 103, 113, 114	IDA_RTFUNC_FAIL, 45, 74
IDA_BAD_TB0, 128, 129	IDA_SIMULTANEOUS, 26, 97
IDA_BAD_TBOUT, 133	IDA_SOLVE_FAIL, 133
IDA_BCKMEM_NULL, 133	IDA_SRES_FAIL, 101, 109
IDA_CENTERED, 104	IDA_STAGGERED, 26, 97
IDA_CONSTR_FAIL, 43, 45	IDA_SUCCESS, 38, 39, 42-44, 47-51, 55-58, 67,
$IDA\_CONV\_FAIL, 43, 45$	70, 72, 81-85, 97-108, 111-117, 125-130,
IDA_CONV_FAILURE, 127, 133	133, 134, 139-142
IDA_ERR_FAIL, 45	IDA_TOO_MUCH_ACC, 45, 127, 133
IDA_ERR_FAILURE, 127, 133	IDA_TOO_MUCH_WORK, 45, 127, 133
IDA_FIRST_QRHS_ERR, 82, 86	IDA_TSTOP_RETURN, 44, 127
IDA_FIRST_QSRHS_ERR, 112, 118	IDA_WARNING, 73
IDA_FIRST_RES_FAIL, 43, 109	IDA_Y_INIT, 43
IDA_FORWARD, 104	IDA_YA_YDP_INIT, 42
IDA_FWD_FAIL, 133	IDAAdjFree, 125
IDA_GETY_BADT, 139	IDAAdjInit, 122, 125
IDA_HERMITE, 125	IDAAdjReInit, 125
IDA_ILL_INPUT, 38, 39, 42-44, 48-51, 55-57, 66,	IDAAdjSetNoSensi, 126
72, 84, 97–100, 103, 104, 108, 111, 112,	IDABBDPRE preconditioner
115, 125, 127–130, 133, 134, 140–142	description, 86–87
IDA_LINESEARCH_FAIL, 43	optional output, 91
IDA_LINIT_FAIL, 43, 45	usage, 88–89
IDA_LMEM_NULL, 70	usage with adjoint module, 154–157
IDA_LSETUP_FAIL, 43, 45, 127, 133, 146, 147, 156,	user-callable functions, 89–91, 154–155
157	user-supplied functions, 87–88, 156–157
IDA_LSOLVE_FAIL, 43, 45, 127	IDABBDPrecGetNumGfnEvals, 91
IDA_MEM_FAIL, 38, 49, 65, 66, 81, 97, 98, 105, 107,	IDABBDPrecGetWorkSpace, 91
108, 111, 125, 127, 128, 140, 141	IDABBDPrecInit, 89
IDA_MEM_NULL, 38, 39, 42–44, 47–51, 55–58, 60–	IDABBDPrecInitB, 154
67, 70, 72, 81–85, 97–108, 111–117, 126,	IDABBDPrecReInit, 90
128-130, 133, 134, 139-142	IDABBDPrecReInitB, 155
IDA_NO_ADJ, 125-134, 140-142	IDACalcIC, 42
IDA_NO_BCK, 133	IDACalcICB, 131
IDA_NO_FWD, 133	IDACalcICBS, 131, 132
IDA_NO_MALLOC, 39, 43, 72, 127–130	IDACreate, 38
IDA_NO_QUAD, 82-85, 115, 141	$\mathtt{IDACreateB},\ 122,\ 127$
TDA NO QUADSENS 112-117	IDAD1sGetLastFlag 71

IDADlsGetNumJacEvals, 68	IDAGetQuadDky, 83
IDAD1sGetNumRhsEvals, 68	IDAGetQuadErrWeights, 85
IDAD1sGetReturnFlagName, 71	IDAGetQuadNumErrTestFails, 85
IDAD1sGetWorkspace, 68	IDAGetQuadNumRhsEvals, 84
IDAD1sJacFn, 76	IDAGetQuadSens, 113
IDAD1sJacFnB, 147	IDAGetQuadSens1, 113
IDAD1sJacFnBS, 148	IDAGetQuadSensDky, 113
IDAD1sSetJacFn, 52	IDAGetQuadSensDky1, 114
IDAD1sSetJacFnB, 135	IDAGetQuadSensErrWeights, 116
IDAD1sSetJacFnBS, 135	IDAGetQuadSensNumErrTestFails, 116
IDAD1sSetLinearSolver, 41	IDAGetQuadSensNumRhsEvals, 116
IDAD1sSetLinearSolverB, 131	IDAGetQuadSensStats, 117
IDAErrHandlerFn, 73	IDAGetQuadStats, 85
IDAEwtFn, 73	IDAGetReturnFlagName, 66
IDAFree, 37, 38	IDAGetRootInfo, 67
IDAGetActualInitStep, 63	IDAGetSens, 96, 101
IDAGetAdjCheckPointsInfo, 139	IDAGetSens1, 96, 102
IDAGetAdjIDABmem, 138	IDAGetSensConsistentIC, 108
IDAGetAdjY, 139	IDAGetSensDky, 96, 101, 102
IDAGetB, 133	IDAGetSensDky1, 96, 102
IDAGetConsistentIC, 66	IDAGetSensErrWeights, 107
IDAGetConsistentICB, 140	IDAGetSensNonlinSolvStats, 107
IDAGetCurrentOrder, 62	IDAGetSensNumErrTestFails, 106
IDAGetCurrentStep, 63	IDAGetSensNumLinSolvSetups, 106
IDAGetCurrentTime, 63	IDAGetSensNumNonlinSolvConvFails, 107
IDAGetDky, 58	IDAGetSensNumNonlinSolvIters, 107
IDAGetErrWeights, 64	IDAGetSensNumResEvals, 105
IDAGetEstLocalErrors, 64	IDAGetSensStats, 106
IDAGetIntegratorStats, 64	IDAGetTolScaleFactor, 64
IDAGetLastLinFlag, 70	IDAGetWorkSpace, 60
IDAGetLastOrder, 62	IDAInit, 38, 71
IDAGetLastStep, 63	IDAInitB, 123, 128
IDAGetLinReturnFlagName, 71	IDAInitBS, 123, 128
IDAGetLinWorkSpace, 67	IDALS linear solver interface
IDAGetNonlinSolvStats, 65	convergence test, 54
IDAGetNumBacktrackOps, 66	Jacobian approximation used by, 51, 52
IDAGetNumErrTestFails, 62	memory requirements, 67
IDAGetNumGEvals, 67	optional input, 51–55, 134–138
IDAGetNumJacEvals, 68	optional output, 67–71
IDAGetNumJtimesEvals, 70	preconditioner setup function, 54, 78
IDAGetNumJTSetupEvals, 70	preconditioner setup function (backward), 153
IDAGetNumLinConvFails, 69	preconditioner solve function, 54, 78
IDAGetNumLinIters, 68	preconditioner solve function (backward), 151
IDAGetNumLinResEvals, 68	IDALS_ILL_INPUT, 41, 53, 54, 90, 131, 135-138,
IDAGetNumLinSolvSetups, 62	155
IDAGetNumNonlinSolvConvFails, 65	IDALS_JACFUNC_RECVR, 146, 147
IDAGetNumNonlinSolvIters, 65	IDALS_JACFUNC_UNRECVR, 146-148
IDAGetNumPrecEvals, 69	IDALS_LMEM_NULL, 52-54, 67-70, 90, 91, 135-138,
IDAGetNumPrecSolves, 69	155
IDAGetNumResEvals, 61	IDALS_MEM_FAIL, 41, 90, 131, 155
IDAGetNumResEvalsSEns, 105	IDALS_MEM_NULL, 41, 52-54, 67-70, 90, 91, 131,
IDAGetNumSteps, 61	135–138, 155
IDAGetQuad, 82, 141	IDALS_NO_ADJ, 131, 135-138
TDAGetQuadB 124 141	TDALS PMEM NULL 91 155

${\tt IDALS\_SUCCESS},\ 41,\ 52-54,\ 67-70,\ 90,\ 91,\ 131,$	IDASensFree, 98
134-138, 155	${\tt IDASensInit},95–97$
${\tt IDALS\_SUNLS\_FAIL},41,53,54$	${\tt IDASensReInit},97$
IDALsJacFn, 74	IDASensResFn, 97, 108
IDALsJacFnB, 146	IDASensSStolerances, 99
IDALsJacFnBS, 146	IDASensSVtolerances, 99
IDALsJacTimesSetupFn, 77	IDASensToggleOff, 98
IDALsJacTimesSetupFnB, 150	${ t IDASetConstraints}, 51$
IDALsJacTimesSetupFnBS, 150	IDASetEpsLin, 54
IDALsJacTimesVecFn, 76	IDASetEpsLinB, 138
IDALsJacTimesVecFnB, 148	IDASetErrFile, 47
IDALsJacTimesVecFnBS, 148	IDASetErrHandlerFn, 47
IDALsPrecSetupFn, 78	IDASetId, 51
IDALsPrecSolveFn, 78	${\tt IDASetIncrementFactor},53$
IDAQuadFree, 82	IDASetIncrementFactorB, 136
IDAQuadInit, 81	IDASetInitStep, 48
IDAQuadInitB, 140	IDASetJacFn, 52
IDAQuadInitBS, 141	IDASetJacFnB, 134
IDAQuadReInit, 81	IDASetJacFnBS, 135
IDAQuadReInitB, 141	IDASetJacTimes, 52
IDAQuadRhsFn, 81, 85	IDASetJacTimesB, 135
IDAQuadRhsFnB, 140, 144	IDASetJacTimesBS, 136
IDAQuadRhsFnBS, 141, 145	IDASetLinearSolver, 36, 41, 74, 218
IDAQuadSensEEtolerances, 115, 116	IDASetLinearSolverB, 123, 130, 146, 154
IDAQuadSensFree, 112	IDASetLineSearchOffIC, 56
IDAQuadSensInit, 111	IDASetMaxBacksIC, 56
IDAQuadSensReInit, 112	IDASetMaxConvFails, 50
IDAQuadSensRhsFn, 111, 117	IDASetMaxErrTestFails, 49
IDAQuadSensSStolerances, 115	IDASetMaxNonlinIters, 49
IDAQuadSensSVtolerances, 115	IDASetMaxNumItersIC, 56
IDAQuadSStolerances, 84	IDASetMaxNumJacsIC, 55
IDAQuadSVtolerances, 84	IDASetMaxNumSteps, 48
IDAReInit, 71, 72	IDASetMaxNumStepsIC, 55
IDAReInitB, 129	IDASetMaxOrd, 48
IDAResFn, 38, 72	IDASetMaxStep, 48
IDAResFnB, 128, 142	IDASetNoInactiveRootWarn, 57
IDAResFnBS, 128, 143	IDASetNonlinConvCoef, 50
IDARootFn, 74	IDASetNonlinConvCoefIC, 55
IDARootInit, 43	IDASetNonLinearSolver, 42
IDAS	IDASetNonlinearSolver, 36, 42, 100, 101
motivation for writing in C, 2	IDASetNonlinearSolverB, 124
package structure, 25	IDASetNonLinearSolverSensSim, 100
relationship to IDA, 1–2	IDASetNonlinearSolverSensSim, 100
IDAS linear solver interface	IDASetNonLinearSolverSensStg, 100
IDALS, $41, 130$	IDASetNonlinearSolverSensStg, 100
IDAS linear solvers	IDASetPreconditioner, 53, 54
header files, 33	IDASetPrecSolveFnB, 137
implementation details, 29	IDASetPrecSolveFnBS, 137
NVECTOR compatibility, 31	IDASetQuadErrCon, 83
selecting one, 41	IDASetQuadSensErrCon, 114
usage with adjoint module, 130	IDASetRootDirection, 57
IDA linear solver interfaces, 26	IDASetSensDQMethod, 104
idas/idas.h, 33	IDASetSensErrCon, 104
IDASensEttolerances 99	IDASetSensMaxNonlinIters 104

IDASetSensParams, 103	Jacobian-vector product
IDASetStepToleranceIC, 57	user-supplied (backward), 135, 148
IDASetStopTime, 49	Jacobian-vector setup
IDASetSuppressAlg, 50	user-supplied, 77–78
IDASetUserData, 47	user-supplied (backward), 150
IDASolve, 36, 44, 116	user-supplied, 52, 74–76
IDASolveB, 124, 132, 133	user-supplied (backward), 134, 146
IDASolveF, 122, 126	( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( (
IDASpilsGetLastFlag, 71	maxord, 71
IDASpilsGetNumConvFails, 69	memory requirements
IDASpilsGetNumJtimesEvals, 70	IDABBDPRE preconditioner, 91
IDASpilsGetNumJTSetupEvals, 70	IDALS linear solver interface, 67
IDASpilsGetNumLinIters, 69	IDAS solver, 81, 97, 111
IDASpilsGetNumPrecEvals, 69	IDAS solver, 61
IDASpilsGetNumPrecSolves, 70	12/16/2017(1)
IDASpilsGetNumRhsEvals, 68	N_VCloneVectorArray, 160
IDASpilsGetReturnFlagName, 71	N_VCloneVectorArray_OpenMP, 180
IDASpilsGetWorkspace, 68	N_VCloneVectorArray_OpenMPDEV, 207
IDASpilsJacTimesSetupFn, 78	N_VCloneVectorArray_Parallel, 176
	N_VCloneVectorArray_ParHyp, 190
IDASpils JacTimes Setup FnB, 150	
IDASpilsJacTimesSetupFnBS, 151	N_VCloneVectorArray_Petsc, 193
IDASpilsJacTimesVecFn, 77	N_VCloneVectorArray_Pthreads, 185
IDASpilsJacTimesVecFnB, 149	N_VCloneVectorArray_Serial, 171
IDASpilsJacTimesVecFnBS, 149	N_VCloneVectorArrayEmpty, 160
IDASpilsPrecSetupFn, 79	N_VCloneVectorArrayEmpty_OpenMP, 181
IDASpilsPrecSetupFnB, 153	N_VCloneVectorArrayEmpty_OpenMPDEV, 207
IDASpilsPrecSetupFnBS, 154	N_VCloneVectorArrayEmpty_Parallel, 176
IDASpilsPrecSolveFn, 78	N_VCloneVectorArrayEmpty_ParHyp, 190
IDASpilsPrecSolveFnB, 152	N_VCloneVectorArrayEmpty_Petsc, 193
IDASpilsPrecSolveFnBS, 153	N_VCloneVectorArrayEmpty_Pthreads, 185
IDASpilsSetEpsLin, 55	N_VCloneVectorArrayEmpty_Serial, 171
IDASpilsSetEpsLinB, 138	<code>N_VCopyFromDevice_Cuda</code> , $199$
IDASpilsSetIncrementFactor, 53	<code>N_VCopyFromDevice_OpenMPDEV</code> , $208$
IDASpilsSetIncrementFactorB, 137	<code>N_VCopyFromDevice_Raja</code> , $204$
${\tt IDASpilsSetJacTimes}, 53$	$ t N_VCopyToDevice\_Cuda,199$
IDASpilsSetJacTimesB, 136	${ t N_VCopyToDevice\_OpenMPDEV, } 208$
IDASpilsSetJacTimesBS, 136	<code>N_VCopyToDevice_Raja</code> , $204$
IDASpilsSetLinearSolver, 41	$ exttt{N\_VDestroyVectorArray}, 160$
IDASpilsSetLinearSolverB, 131	N_VDestroyVectorArray_OpenMP, 181
IDASpilsSetPreconditioner, 54	${ t N\_VDestroyVectorArray\_OpenMPDEV,\ 207}$
IDASpilsSetPreconditionerB, 137	N_VDestroyVectorArray_Parallel, 176
IDASpilsSetPreconditionerBS, 138	N_VDestroyVectorArray_ParHyp, 190
IDASStolerances, 39	N_VDestroyVectorArray_Petsc, 193
IDASStolerancesB, 130	N_VDestroyVectorArray_Pthreads, 186
IDASVtolerances, 39	N_VDestroyVectorArray_Serial, 171
IDASVtolerancesB, 130	N_Vector, 33, 159
IDAWFtolerances, 39	N_VEnableConstVectorArray_Cuda, 201
itask, 44, 126	N_VEnableConstVectorArray_OpenMP, 182
,,	N_VEnableConstVectorArray_OpenMPDEV, 209
Jacobian approximation function	N_VEnableConstVectorArray_Parallel, 177
difference quotient, 51	N_VEnableConstVectorArray_ParHyp, 191
Jacobian times vector	N_VEnableConstVectorArray_Petsc, 195
difference quotient, 52	N_VEnableConstVectorArray_Pthreads, 187
user-supplied, 52, 76–77	N_VEnableConstVectorArray_Raja, 205

N_VEnableConstVectorArray_Serial, 172	N_VEnableScaleAddMulti_Cuda, 200
N_VEnableDotProdMulti_Cuda, 200	N_VEnableScaleAddMulti_OpenMP, 182
N_VEnableDotProdMulti_OpenMP, 182	N_VEnableScaleAddMulti_OpenMPDEV, 209
N_VEnableDotProdMulti_OpenMPDEV, 209	N_VEnableScaleAddMulti_Parallel, 177
N_VEnableDotProdMulti_Parallel, 177	N_VEnableScaleAddMulti_ParHyp, 191
N_VEnableDotProdMulti_ParHyp, 191	N_VEnableScaleAddMulti_Petsc, 194
N_VEnableDotProdMulti_Petsc, 194	N_VEnableScaleAddMulti_Pthreads, 187
N_VEnableDotProdMulti_Pthreads, 187	N_VEnableScaleAddMulti_Raja, 205
N_VEnableDotProdMulti_Serial, 172	N_VEnableScaleAddMulti_Serial, 172
N_VEnableFusedOps_Cuda, 200	N_VEnableScaleAddMultiVectorArray_Cuda, 201
N_VEnableFusedOps_OpenMP, 181	N_VEnableScaleAddMultiVectorArray_OpenMP, 183
N_VEnableFusedOps_OpenMPDEV, 208	N_VEnableScaleAddMultiVectorArray_OpenMPDEV,
N_VEnableFusedOps_Parallel, 177	210
N_VEnableFusedOps_ParHyp, 190	N_VEnableScaleAddMultiVectorArray_Parallel,
N_VEnableFusedOps_Petsc, 194	178
N_VEnableFusedOps_Pthreads, 186	N_VEnableScaleAddMultiVectorArray_ParHyp, 192
N_VEnableFusedOps_Raja, 204	N_VEnableScaleAddMultiVectorArray_Petsc, 195
N_VEnableFusedOps_Serial, 172	N_VEnableScaleAddMultiVectorArray_Pthreads,
N_VEnableLinearCombination_Cuda, 200	188
N_VEnableLinearCombination_OpenMP, 181	N_VEnableScaleAddMultiVectorArray_Raja, 205
N_VEnableLinearCombination_OpenMPDEV, 208	N_VEnableScaleAddMultiVectorArray_Serial, 173
N_VEnableLinearCombination_Parallel, 177	N_VEnableScaleVectorArray_Cuda, 200
•	N_VEnableScaleVectorArray_OpenMP, 182
N_VEnableLinearCombination_ParHyp, 190	N_VEnableScaleVectorArray_OpenMPDEV, 209
N_VEnableLinearCombination_Petsc, 194	N_VEnableScaleVectorArray_Parallel, 177
N_VEnableLinearCombination_Pthreads, 186	· · · · · · · · · · · · · · · · · · ·
N_VEnableLinearCombination_Raja, 204	N_VEnableScaleVectorArray_ParHyp, 191
N_VEnableLinearCombination_Serial, 172	N_VEnableScaleVectorArray_Petsc, 194
${\tt N\_VEnableLinearCombinationVectorArray\_Cuda},$	
201	N_VEnableScaleVectorArray_Raja, 205
N_VEnableLinearCombinationVectorArray_OpenM	
183	N_VEnableWrmsNormMaskVectorArray_Cuda, 201
	PDNVEnableWrmsNormMaskVectorArray_OpenMP, 182
210	N_VEnableWrmsNormMaskVectorArray_OpenMPDEV,
N_VEnableLinearCombinationVectorArray_Paral	
178	N_VEnableWrmsNormMaskVectorArray_Parallel, 178
	pN_VEnableWrmsNormMaskVectorArray_ParHyp, 191
192	N_VEnableWrmsNormMaskVectorArray_Petsc, 195
	,N_VEnableWrmsNormMaskVectorArray_Pthreads, 187
195	N_VEnableWrmsNormMaskVectorArray_Serial, 173
${\tt N\_VEnableLinearCombinationVectorArray\_Pthree}$	· · · · · · · · · · · · · · · · · · ·
188	N_VEnableWrmsNormVectorArray_OpenMP, 182
${\tt N\_VEnableLinearCombinationVectorArray\_Raja},$	
205	N_VEnableWrmsNormVectorArray_Parallel, 178
${\tt N\_VEnableLinearCombinationVectorArray\_Seria}$	· · · · · · · · · · · · · · · · · · ·
173	${\tt N\_VEnableWrmsNormVectorArray\_Petsc}, 195$
${\tt N\_VEnableLinearSumVectorArray\_Cuda}, {\tt 200}$	N_VEnableWrmsNormVectorArray_Pthreads, 187
${\tt N\_VEnableLinearSumVectorArray\_OpenMP},182$	N_VEnableWrmsNormVectorArray_Serial, 173
${\tt N\_VEnableLinearSumVectorArray\_OpenMPDEV}, {\tt 209}$	N_VGetDeviceArrayPointer_Cuda, 197
${\tt N\_VEnableLinearSumVectorArray\_Parallel}, \frac{177}{}$	N_VGetDeviceArrayPointer_OpenMPDEV, 208
${\tt N\_VEnableLinearSumVectorArray\_ParHyp,}~191$	N_VGetDeviceArrayPointer_Raja, 203
${\tt N\_VEnableLinearSumVectorArray\_Petsc}, 194$	${\tt N\_VGetHostArrayPointer\_Cuda},197$
${\tt N\_VEnableLinearSumVectorArray\_Pthreads}, 187$	${\tt N\_VGetHostArrayPointer\_OpenMPDEV},\ 208$
N_VEnableLinearSumVectorArray_Raja, 205	W WG . W . A . D D
::- v = :: a y = :: a	N_VGetHostArrayPointer_Raja, 202

N_VGetLength_OpenMP, 181	$ exttt{N_VPrintFile_OpenMPDEV}, 208$
N_VGetLength_OpenMPDEV, 207	$ exttt{N_VPrintFile\_Parallel}, 176$
N_VGetLength_Parallel, 176	$ exttt{N_VPrintFile\_ParHyp}, 190$
N_VGetLength_Pthreads, 186	${\tt N\_VPrintFile\_Petsc}, 194$
N_VGetLength_Raja, 202	$ exttt{N_VPrintFile_Pthreads},186$
N_VGetLength_Serial, 171	$ exttt{N\_VPrintFile\_Raja}, 204$
N_VGetLocalLength_Cuda, 197	N_VPrintFile_Serial, 171
N_VGetLocalLength_Parallel, 176	N_VSetCudaStream_Cuda, 199
N_VGetLocalLength_Raja, 202	NV_COMM_P, 175
N_VGetMPIComm_Cuda, 197	NV_CONTENT_OMP, 179
N_VGetMPIComm_Raja, 203	NV_CONTENT_OMPDEV, 206
N_VGetVector_ParHyp, 189	NV_CONTENT_P, 174
N_VGetVector_Petsc, 193	NV_CONTENT_PT, 184
N_VIsManagedMemory_Cuda, 197	NV_CONTENT_S, 169
N_VMake_Cuda, 198	NV_DATA_DEV_OMPDEV, 206
N_VMake_OpenMP, 180	NV_DATA_HOST_OMPDEV, 206
N_VMake_OpenMPDEV, 207	NV_DATA_OMP, 179
N_VMake_Parallel, 176	NV_DATA_P, 174
N_VMake_ParHyp, 189	NV_DATA_PT, 184
N_VMake_Petsc, 193	NV_DATA_S, 170
N_VMake_Pthreads, 185	NV_GLOBLENGTH_P, 174
N_VMake_Raja, 204	NV_Ith_OMP, 180
N_VMake_Serial, 171	NV_Ith_P, 175
N_VMakeManaged_Cuda, 199	NV_Ith_PT, 185
N_VNew_Cuda, 197	NV_Ith_S, 170
N_VNew_OpenMP, 180	NV_LENGTH_OMP, 179
N_VNew_OpenMPDEV, 207	NV_LENGTH_OMPDEV, 206
N_VNew_Parallel, 175	NV_LENGTH_PT, 184
N_VNew_Pthreads, 185	NV_LENGTH_S, 170
N_VNew_Raja, 203	NV_LOCLENGTH_P, 174
N_VNew_SensWrapper, 309	NV_NUM_THREADS_OMP, 179
N_VNew_Serial, 170	NV_NUM_THREADS_PT, 184
N_VNewEmpty_Cuda, 198	NV_OWN_DATA_OMP, 179
N_VNewEmpty_OpenMP, 180	NV_OWN_DATA_OMPDEV, 206
N_VNewEmpty_OpenMPDEV, 207	NV_OWN_DATA_P, 174
N_VNewEmpty_Parallel, 175	NV_OWN_DATA_PT, 184
N_VNewEmpty_ParHyp, 189	NV_OWN_DATA_S, 170
N_VNewEmpty_Petsc, 193	NVECTOR module, 159
N_VNewEmpty_Pthreads, 185	nvector_openmp_mod, 183
N_VNewEmpty_Raja, 203	nvector_pthreads_mod, 188
N_VNewEmpty_SensWrapper, 309	nvooror_ponrodds_mod, 100
N_VNewEmpty_Serial, 170	optional input
N_VNewManaged_Cuda, 198	backward solver, 134
N_VPrint_Cuda, 199	forward sensitivity, 103–105
N_VPrint_OpenMP, 181	generic linear solver interface, 51–55, 134–138
N_VPrint_OpenMPDEV, 208	initial condition calculation, 55–57
N_VPrint_Parallel, 176	iterative linear solver, 55, 137–138
N_VPrint_ParHyp, 190	iterative-free linear solver, 53
N_VPrint_Petsc, 193	matrix-based linear solver, 51–52, 134–135
N_VPrint_Pthreads, 186	matrix-free linear solver, 52–53, 135–137
N_VPrint_Raja, 204	quadrature integration, 83–84, 142
N_VPrint_Serial, 171	rootfinding, 57
N_VPrintFile_Cuda, 200	sensitivity-dependent quadrature integration,
N_VPrintFile_OpenMP, 181	114–116
· · · · · · · · · · · · · · · · · · ·	The state of the s

solver, $45-51$	SM_COLUMNS_D, 219
optional output	SM_COLUMNS_S, 231
backward initial condition calculation, 139–	SM_CONTENT_B, 224
140	SM_CONTENT_D, 219
backward solver, 138–139	SM_CONTENT_S, 231
band-block-diagonal preconditioner, 91	SM_DATA_B, 224
forward sensitivity, 105–108	SM_DATA_D, 219
generic linear solver interface, 67–71	SM_DATA_S, 231
initial condition calculation, 66, 108	SM_ELEMENT_B, 75, 224
interpolated quadratures, 83	SM_ELEMENT_D, 75, 220
interpolated sensitivities, 101	SM_INDEXPTRS_S, 231
interpolated sensitivity-dep. quadratures, 113	SM_INDEXVALS_S, 231
interpolated solution, 58	SM_LBAND_B, 224
quadrature integration, 84–85, 142	SM_LDATA_B, 224
sensitivity-dependent quadrature integration,	SM_LDATA_D, 219
116–117	SM_LDIM_B, 224
solver, 60–66	SM_NNZ_S, 76, 231
version, 58–60	SM_NP_S, 231
output mode, 126, 133	SM_ROWS_B, 224
output mode, 120, 100	SM_ROWS_D, 219
partial error control	
explanation of IDAS behavior, 118	SM_ROWS_S, 231
portability, 32	SM_SPARSETYPE_S, 231
preconditioning	SM_SUBAND_B, 224
advice on, 15–16, 26	SM_UBAND_B, 224
band-block diagonal, 86	SMALL_REAL, 32
setup and solve phases, 26	step size bounds, 48–49
user-supplied, 53–54, 78, 137–138, 151, 153	SUNBandMatrix, 35, 225
aser supplied, 55 51, 76, 157 156, 151, 155	SUNBandMatrix_Cols, 227
quadrature integration, 17	SUNBandMatrix_Column, 227
forward sensitivity analysis, 20	SUNBandMatrix_Columns, 226
v v	SUNBandMatrix_Data, 227
RCONST, 32	SUNBandMatrix_LDim, 226
realtype, $32$	SUNBandMatrix_LowerBandwidth, 226
reinitialization, 71, 129	SUNBandMatrix_Print, 226
residual function, 72	SUNBandMatrix_Rows, 226
backward problem, 142, 143	SUNBandMatrix_StoredUpperBandwidth, 226
forward sensitivity, 108	SUNBandMatrix_UpperBandwidth, 226
quadrature backward problem, 144	${\tt SUNBandMatrixStorage}, 225$
sensitivity-dep. quadrature backward prob-	${\tt SUNDenseMatrix},35,220$
lem, 145	SUNDenseMatrix_Cols, 221
right-hand side function	SUNDenseMatrix_Column, 221
quadrature equations, 85	SUNDenseMatrix_Columns, 220
sensitivity-dependent quadrature equations,	SUNDenseMatrix_Data, 221
117	SUNDenseMatrix_LData, 221
Rootfinding, 16, 36, 43	SUNDenseMatrix_Print, 220
3, 2, 22, 2	SUNDenseMatrix_Rows, 220
second-order sensitivity analysis, 23	sundials/sundials_linearsolver.h, 237
support in IDAS, 24	sundials_nonlinearsolver.h, 33
SM_COLS_B, 224	sundials_nvector.h, 33
SM_COLS_D, 219	sundials_types.h, 32, 33
SM_COLUMN_B, 75, 224	SUNDIALSGetVersion, 60
SM_COLUMN_D, 75, 220	SUNDIALSGetVersionNumber, 60
SM_COLUMN_ELEMENT_B, 75, 224	sunindextype, 32
SM_COLUMNS_B, 224	SUNLinearSolver, 237, 244
	, ,

SUNLinearSolver module, 237	SUNNonlinearSolver, 33, 301
SUNLINEARSOLVER_DIRECT, 239, 246	SUNNonlinearSolver module, 301
SUNLINEARSOLVER_ITERATIVE, 239, 247	SUNNONLINEARSOLVER_FIXEDPOINT, 302
SUNLINEARSOLVER_MATRIX_ITERATIVE, 239, 247	SUNNONLINEARSOLVER_ROOTFIND, 302
sunlinsol/sunlinsol_band.h, 33	SUNNonlinSol_FixedPoint, 314, 317
sunlinsol/sunlinsol_dense.h, 33	SUNNonlinSol_FixedPointSens, 315
sunlinsol/sunlinsol_klu.h, 33	SUNNonlinSol_Newton, 311
sunlinsol/sunlinsol_lapackband.h, 33	SUNNonlinSol_NewtonSens, 311
sunlinsol/sunlinsol_lapackdense.h, 33	SUNNonlinSolFree, 37, 303
sunlinsol/sunlinsol_pcg.h, 34	SUNNonlinSolGetCurIter, 305
sunlinsol/sunlinsol_spbcgs.h, 33	SUNNonlinSolGetNumConvFails, 305
sunlinsol/sunlinsol_spfgmr.h, 33	SUNNonlinSolGetNumIters, 305
sunlinsol/sunlinsol_spgmr.h, 33	SUNNonlinSolGetSysFn_FixedPoint, 315
sunlinsol/sunlinsol_sptfqmr.h, 34	SUNNonlinSolGetSysFn_Newton, 312
sunlinsol/sunlinsol_superlumt.h, 33	${\tt SUNNonlinSolGetType},302$
SUNLinSol_Band, 41, 252	${\tt SUNNonlinSolInitialize},302$
SUNLinSol_Dense, 41, 250	SUNNonlinSolLSetupFn, 303
SUNLinSol_KLU, 41, 260	SUNNonlinSolSetConvTestFn, 304
SUNLinSol_KLUReInit, 261	SUNNonlinSolSetLSolveFn, 304
SUNLinSol_KLUSetOrdering, 263	SUNNonlinSolSetMaxIters, 304
SUNLinSol_LapackBand, 41, 257	SUNNonlinSolSetSysFn, 303
SUNLinSol_LapackDense, 41, 255	SUNNonlinSolSetup, 302
SUNLinSol_PCG, 41, 294, 296	SUNNonlinSolSolve, 302
SUNLinSol_PCGSetMax1, 295	SUNSparseFromBandMatrix, 232
SUNLinSol_PCGSetPrecType, 295	SUNSparseFromDenseMatrix, 232
SUNLinSol_SPBCGS, 41, 283, 285	SUNSparseMatrix, 35, 232
SUNLinSol_SPBCGSSetMaxl, 284	SUNSparseMatrix_Columns, 233
•	SUNSparseMatrix_Data, 234
SUNLinSol_SPECGSSetPrecType, 284	SUNSparseMatrix_IndexPointers, 234
SUNLinSol_SPFGMR, 41, 276, 279	SUNSparseMatrix_IndexValues, 234
SUNLinSol_SPFGMRSetMaxRestarts, 278	SUNSparseMatrix_NNZ, 76, 233
SUNLinSol_SPFGMRSetPrecType, 277	SUNSparseMatrix_NP, 234
SUNLinSol_SPGMR, 41, 269, 272	SUNSparseMatrix_Print, 233
SUNLinSol_SPGMRSetMaxRestarts, 271	SUNSparseMatrix_Realloc, 233
SUNLinSol_SPGMRSetPrecType, 271	SUNSparseMatrix_Reallocate, 233
SUNLinSol_SPTFQMR, 41, 288, 290	- · · · · · · · · · · · · · · · · · · ·
SUNLinSol_SPTFQMRSetMax1, 289	SUNSparseMatrix_Rows, 233
SUNLinSol_SPTFQMRSetPrecType, 289	SUNSparseMatrix_SparseType, 234
SUNLinSol_SuperLUMT, 41, 265	tolerances, 13, 39, 40, 73, 84, 115
SUNLinSol_SuperLUMTSetOrdering, 267	tolerances, 19, 99, 40, 79, 04, 119
SUNLinSolFree, 37, 238, 240	UNIT_ROUNDOFF, 32
SUNLinSolGetType, 238	User main program
SUNLinSolInitialize, 238, 239	Adjoint sensitivity analysis, 121
SUNLinSolLastFlag, 242	forward sensitivity analysis, 93
SUNLinSolNumIters, 241	IDABBDPRE usage, 88
SUNLinSolResNorm, 241	IDAS usage, 34
SUNLinSolSetATimes, 239, 240, 247	integration of quadratures, 79
SUNLinSolSetPreconditioner, 241	
SUNLinSolSetScalingVectors, 241	integration of sensitivitiy-dependent quadra
SUNLinSolSetup, 238, 239, 247	tures, 109
SUNLinSolSolve, 238, 239	user_data, 47, 73, 74, 86, 88, 117
SUNLinSolSpace, 242	user_dataB, 156, 157
SUNMatDestroy, 37	weighted root-mean-square norm, 12–13
SUNMatrix, 215	weighted 100t-mean-square norm, 12–15
SUNMatrix module, 215	