# STRUMPACK Users' Guide

Pieter Ghysels\* Xiaoye S. Li\* Christopher Gorman<sup>†</sup> Gustavo Chávez\* François-Henry Rouet<sup>‡</sup>

Version 2.3.0, ?? 2018

<sup>&</sup>lt;sup>1</sup>Lawrence Berkeley National Laboratory, Computational Research Division, MS 50F-1650, One Cyclotron Road, Berkeley  ${CA94720.\ \{pghysels,xsli,gichavez\}@lbl.gov}\\ {^2UC\ Santa\ Barbara.\ gorman@math.ucsb.edu}$ 

 $<sup>^3</sup>$ Livermore Software Technology Corporation. fhrouet@lstc.com

# Contents

1	STRUMPACK Overview	3
2	Installation and Requirements	3
3	Algorithm	5
4	Using STRUMPACK Sparse  4.1 StrumpackSparseSolver Example  4.2 StrumpackSparseSolverMPI Example  4.3 StrumpackSparseSolverMPIDist Example  4.4 Initializing the Solver Object  4.5 Sparse Matrix Format  4.6 Setting and Parsing Options  4.7 Reordering  4.7.1 Reordering for numerical stability  4.7.2 Nested dissection reordering  4.8 Factorization  4.9 Solve  4.10 All Options for the Sparse Solver  4.10.1 SPOptions <scalar> Interface  4.10.2 Command Line Options</scalar>	6 6 7 8 9 10 11 11 12 12 12 13 13 15
5 6	HSS Preconditioning 5.1 HSSOptions <scalar> Interface</scalar>	16 17 18 18
7	Examples	18
8	C Interface	19
9	Advanced Usage Tips	19
10	FAQ	19
11	Acknowledgements	20
<b>12</b>	Copyright notice	<b>20</b>
13	License agreement	20

## 1 STRUMPACK Overview

STRUMPACK – STRUctured Matrix PACKage – is a C++ library for computations with dense and sparse matrices. It uses so-called *structured matrices*, i.e., matrices that exhibit some kind of low-rank property, in this case with Hierarchically Semi-Separable matrices (HSS), to speed up linear algebra operations. This version of STRUMPACK unifies two main components that were separate in previous versions: a package for dense matrix computations (STRUMPACK-dense) and a package (STRUMPACK-sparse) for sparse linear systems. The algorithms for solving dense linear systems are described in [7] while the algorithms for solving sparse linear systems are described in [5, 4]. STRUMPACK can be used as a general algebraic sparse direct solver (based on the multifrontal factorization method), or as an efficient preconditioner for sparse matrices obtained by discretization of partial differential equations. Included in STRUMPACK are also the GMRES and BiCGStab iterative Krylov solvers, that use the approximate, HSS-accelerated, sparse solver as a preconditioner for the efficient solution of sparse linear systems.

The STRUMPACK project started at the Lawrence Berkeley National Laboratory in 2014 and is supported by the FASTMath SciDAC Institute funded by the Department of Energy and by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

Check the STRUMPACK website for more information and for the latest code:

http://portal.nersc.gov/project/sparse/strumpack/

## 2 Installation and Requirements

The STRUMPACK package uses the CMake build system (CMake version >= 2.8). The recommended way of building the STRUMPACK library is as follows:

```
> tar -xvzf strumpack-x.y.z.tar.gz
> cd strumpack-x.y.x
> mkdir build
> cd build
> cmake ... / -DCMAKE BUILD TYPE=Release \
             -DCMAKE INSTALL PREFIX=/path/to/install \
             -DCMAKE CXX COMPILER=<C++ compiler> \
                                                                                                                                                                                                                         # this and below are optional,
             -DCMAKE_C_COMPILER=<C compiler> \
                                                                                                                                                                                                                         # CMake will try to autodetect
             -DCMAKE Fortran COMPILER=<Fortran77 compiler> \
             -DSCALAPACK_LIBRARIES="/path/to/scalapack/libscalapack.a;/path/to/blacs/libblacs.a" \
             -DMETIS INCLUDES=/path/to/metis/incluce \
             -DMETIS LIBRARIES=/path/to/metis/libmetis.a \
             -DSTRUMPACK USE PARMETIS=ON \
                                                                                                                                                                                                                         # optional
             -DPARMETIS_INCLUDES=/path/to/parmetis/include \
             -DPARMETIS_LIBRARIES=/path/to/parmetis/libparmetis.a \
             -DSTRUMPACK USE SCOTCH=ON \
                                                                                                                                                                                                                         # optional
             -DSCOTCH INCLUDES=/path/to/scotch/include \
             -DSCOTCH LIBRARIES="/path/to/ptscotch/libscotch.a;...libscotcherr.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...libptscotch.a;...lib
> make test
                                           # optional, takes a while
> make install
```

The above will only work if you have the following dependencies, and CMake can find them:

• C++11, C and FORTRAN77 compilers. CMake looks for these compilers in the standard locations, if they are installed elsewhere, you can specify them as follows:

```
> cmake ../ -DCMAKE_BUILD_TYPE=Release \
    -DCMAKE CXX COMPILER=g++ \
```

```
-DCMAKE_C_COMPILER=gcc \
-DCMAKE Fortran COMPILER=gfortran
```

- MPI (Message Passing Interface) library. You should not need to manually specify the MPI compiler wrappers. CMake will look for MPI options and libraries and set the appropriate compiler and linker flags.
- OpenMP v4.5 support is required in the C++ compiler to use the shared-memory parallelism in the code. OpenMP support can be disabled by adding the CMake option

```
-DSTRUMPACK USE OPENMP=OFF
```

OpenMP v3.1 introduces task parallelism, which is used extensively throughout the code. OpenMP v4.5 added the taskloop construct. CMake will check whether your compiler supports OpenMP and sets the appropriate compiler and linker flags.

• BLAS, LAPACK and Scalapack libraries. For performance it is crucial to use optimized BLAS/LAPACK libraries like for instance Intel<sup>®</sup> MKL, AMD<sup>®</sup> ACML, Cray<sup>®</sup> LibSci or OpenBLAS. The default versions of the Intel<sup>®</sup> MKL and Cray<sup>®</sup> LibSci BLAS libraries will use multithreaded kernels, unless when they are called from within an OpenMP parallel region, in which case they run sequentially. This is the behavior STRUMPACK relies upon to achieve good performance when running in MPI+OpenMP hybrid mode. Scalapack depends on the Blacs communication library and on PBLAS (parallel Blas), both of which are typically included with the Scalapack installation (from Scalapack 2.0.2, the blacs library is included in the Scalapack library file). If CMake cannot locate these libraries, you can specify their path by setting the environment variable \$SCALAPACKDIR or by specifying the libraries manually:

```
> cmake ../ -DCMAKE_BUILD_TYPE=Release \
 -DBLAS_LIBRARIES="..." \
 -DLAPACK_LIBRARIES="..." \
 -DSCALAPACK_LIBRARIES="..."
```

Or one can also directly modify the linker flags to add for instance the ScaLAPACK and BLACS libraries:

```
> cmake ../ -DCMAKE_BUILD_TYPE=Release \
   -DCMAKE EXE LINKER FLAGS="-L/usr/lib64/mpich/lib/ -lscalapack -lmpiblacs"
```

When using Intel MKL we recommend using the the link advisor: https://software.intel.com/en-us/articles/intel-mkl-link-line-advisor

• **METIS** (≥ 5.1.0 **required**) for the nested dissection matrix reordering. Metis can be obtained from: http://glaros.dtc.umn.edu/gkhome/metis/metis/download.

CMake looks for the Metis inlude files the library in the default locations as well as in \$metisdir/include and \$metisdir/lib. Using the Bash shell, the METISDIR environment variable can be set as export METISDIR=/usr/local/metis/. Alternatively, you can specify the location of the header and library as follows:

```
> cmake ../ -DCMAKE_BUILD_TYPE=Release \
 -DMETIS_INCLUDES=/usr/local/metis/include \
 -DMETIS_LIBRARIES=/usr/local/metis/lib/libmetis.a
```

• **PARMETIS** (optional) for parallel nested dissection. ParMetis can be download from http://glaros.dtc.umn.edu/gkhome/metis/parmetis/download

The steps to make sure CMake can find ParMetis are similar as for Metis. Enable with -dstrumpack\_use\_parmetis. The CMake variables are \$parmetisdir or parmetis\_includes and parmetis\_libraries.

• SCOTCH and PT-SCOTCH (≥ 5.1.12) (optional) for matrix reordering. Scotch can be downloaded from:

http://www.labri.fr/perso/pelegrin/scotch/

Configuring CMake to find (PT-)Scotch is similar to Metis. Enable with -DSTRUMPACK\_USE\_SCOTCH For (PT-)Scotch the CMake variables are \$SCOTCHDIR or SCOTCH\_INCLUDES and SCOTCH\_LIBRARIES. Make sure to specify all libraries: libscotch, libscotcherr, libptscotch and libptscotcherr.

- getopt long: This is a GNU extension to the POSIX getopt() C library function.
- Combinatorial BLAS [1] (optional) can be used for parallel reordering for stability, i.e., to get nonzeros on the diagonal of the matrix, as an alternative to the sequential MC64 (included). Get it from:

https://people.eecs.berkeley.edu/~aydin/CombBLAS/html/

Enable by specifying -dstrumpack\_use\_combblas and set combblas\_includes and combblas\_libraries.

• TCMalloc, TBB Malloc or jemalloc (optional): these can lead to dramatic performance improvements for multithreaded code that performs frequent memory allocations. Link with the one of these libraries, e.g.:

```
-DCMAKE_EXE_LINKER_FLAGS="-ltcmalloc"
```

to replace the default memory allocator (C++ new) with a more scalable implementation. See also Section 9.

The code was tested on GNU/Linux with the GNU and Intel<sup>®</sup> compilers and the OpenBLAS, Intel<sup>®</sup> MKL<sup>®</sup> and Cray<sup>®</sup> LibSci<sup>®</sup> numerical libraries. If you encounter issues on other platforms or with other BLAS/LAPACK implementations, please let us know. Successful compilation will create a libstrumpack.a file.

## 3 Algorithm

The algorithm used in STRUMPACK is described in detail in [5], and is based on the work by Jianlin Xia [9]. Here we summarize the main algorithm features. Section 5 has more information on the low-rank compression strategy and how to tune this to get a good preconditioner for your specific problem. There are three main steps in the algorithm: matrix reordering, factorization and solve.

Matrix reordering: There are three distinct matrix reordering steps: one for stability, one to limit fill-in and one to reduce HSS-ranks. First, the matrix is reordered and possibly scaled for numerical stability by the MC64 code [3] or by a code from Combinatorial BLAS [1]. This step is called the matching phase. For many matrices, this reordering can safely be disabled. By default, MC64 is used to maximize the product of the diagonal values of the matrix, and to scale the rows and columns of the matrix. Alternatively, MC64 can be used to maximize the smallest diagonal value or to maximize the sum of the diagonals. Next, a nested dissection reordering is applied to limit fill-in. Both (Par)Metis and (PT-)Scotch are supported. We expose one user tunable parameter which controls the size of the smallest separators. Finally, when HSS compression is used, there is an extra reordering step to reduce the HSS-ranks. This reordering uses Metis and does not require user tuning.

Factorization: Before the actual numerical factorization, there is a symbolic factorization step to construct the elimination tree. After that, the multifrontal factorization procedure traverses this elimination tree from bottom (smallest separators) to top (root separator). With each node of the elimination tree a dense matrix is associated, referred to as a frontal matrix, or simply front. These fronts can possibly be compressed as Hierarchically Semi-Separable (HSS) matrices. This compression will only pay off for

fronts that are large enough, which are typically the frontal matrices at the nodes in the elimination tree close to the root. Without any HSS compression, the solver acts as a standard multifrontal direct solver. HSS approximations are constructed using a randomized sampling algorithm.

Solve: Once the matrix is factorized, the factors can be used to efficiently solve a linear system of equations by doing a forward and a backward solve sweeps. When no HSS compression is used, this is a direct solver. The multifrontal solve procedure is then used within an iterative refinement loop, with typically only 1 or very few iterations. However, when the factors are compressed using HSS, a single multifrontal solve is only approximate and the solve is by default used as a preconditioner for GMRES(30). The required number of GMRES iterations will depend strongly on the quality of the HSS approximation.

## 4 Using STRUMPACK Sparse

This section gives an overview on the basic usage of the sparse solvers in STRUMPACK. Many STRUMPACK options can be set from the command line. Running with --help or -h, will give you a list of supported runtime options.

An example Makefile is available in the examples/ directory. This Makefile is generated by the cmake command, see Section 2.

The STRUMPACK package is written in C++, and offers a simple C++ interface. See Section 8 if you prefer a C interface. STRUMPACK-sparse has three different solver classes, all interaction happens through objects of these classes:

#### • StrumpackSparseSolver<scalar,integer=int>

This class represents the sparse solver for a single computational node, optionally using OpenMP parallelism. Use this if you are running the code sequentially, on a (multicore) laptop or desktop or on a single node of a larger cluster. This class is defined in StrumpackSparseSolver.hpp, so include this header if you intend to use it.

#### StrumpackSparseSolverMPI<scalar,integer=int>

This solver has (mostly) the same interface as StrumpackSparseSolver<scalar,integer> but the numerical factorization and multifrontal solve phases run in parallel using MPI and ScaLAPACK. However, the inputs (sparse matrix, right-hand side vector) need to be available completely on every MPI process. The reordering phase uses Metis or Scotch (not ParMetis or PTScotch) and the symbolic factorization is threaded, but not distributed. The (multifrontal) solve is done in parallel, but the right-hand side vectors need to be available completely on every processor. Make sure to call MPI\_Init[\_thread] before instantiating an object of this class and include the header file StrumpackSparseSolverMPI.hpp. We do not recommend this solver, instead, use StrumpackSparseSolverMPIDist whenever possible.

#### StrumpackSparseSolverMPIDist<scalar,integer=int>

This solver is fully distributed. The numerical factorization and solve as well as the symbolic factorization are distributed. The input is now a block-row distributed sparse matrix and a correspondingly distributed right-hand side. For matrix reordering, ParMetis or PT-Scotch are used. Include the header file StrumpackSparseSolverMPIDist.hpp and call MPI\_Init[\_thread].

The three solver classes StrumpackSparseSolver, StrumpackSparseSolverMPI and StrumpackSparseSolverMPIDist depend on two template parameters <scalar,integer>: the type of a scalar and an integer type. The scalar type can be float, double, std::complex<float> or std::complex<double>. It is recommended to first try to simply use the default integer=int type, unless you run into 32 bit integer overflow problems. In that case one can switch to for instance int64 t (a signed integer type).

### 4.1 StrumpackSparseSolver Example

The following shows the typical way to use a (sequential or multithreaded) STRUMPACK sparse solver:

$$A = \begin{bmatrix} 8.2 & 0.1 & & & & & & \\ 0 & & -4.8 & & & & \\ 6.2 & 1.1 & & 2.6 & & & \\ & & -1.0 & & & & \\ & & & 99.9 & 4.0 \end{bmatrix} \qquad \begin{array}{c} \mathsf{row\_ptr} = [0, \, 3, \, 5, \, 8, \, 9, \, 11] \\ \mathsf{col\_ind} = [0, \, 1, \, 4 \, | \, 0, \, 2 \, | \, 0, \, 1, \, 3 \, | \, 2 \, | \, 3, \, 4] \\ \mathsf{values} = [8.2, \, 0.1, \, 3.1 \, | \, 0, \, -4.8 \, | \, 6.2, \, 1.1, \, 2.6 \, | \, -1.0 \, | \, 99.9, \, 4.0] \\ \end{array}$$

Figure 1: Illustration of a small  $5 \times 5$  sparse matrix with 11 nonzeros and its Compressed Sparse Row (CSR) or Yale format representation. We always use 0-based indexing! Let N=5 denote the number of rows. The row\_ptr array has N+1 elements, with element i denoting the start of row i in the col\_ind and values arrays. Element row\_ptr[N] = nnz, i.e., the total number of nonzero elements in the matrix. The values array holds the actual matrix values, ordered by row. The corresponding elements in col\_ind give the column indices for each nonzero. There can be explicit zero elements in the matrix. The nonzero values and corresponding column indices need not be sorted by column (within a row).

```
#include "StrumpackSparseSolver.hpp"
using namespace strumpack; // all strumpack code is in the strumpack namespace,
int main(int argc, char* argv[]) {
 int N = ...;
                          // construct an NxN CSR matrix with nnz nonzeros
 int* row ptr = ...;
                           // N+1 integers
 int* col ind = ...;
                           // nnz integers
 double* val = ...; // nnz scalars
 double* x = new double[N]; // will hold the solution vector
 double* b = ...;
                            // set a right-hand side b
 StrumpackSparseSolver<double> sp;
                                                // create solver object
  sp.options().set_rel_tol(1e-10);
                                                // set options
  sp.options().set gmres restart(10);
                                                // ...
  sp.options().enable HSS();
                                                // enable HSS compression, see section 5
 sp.options().set from command line(argc, argv); // parse command line options
 sp.set csr matrix(N, row ptr, col ind, val);
                                                // set the matrix (copy)
 sp.reorder();
                                                 // reorder matrix
 sp.factor();
                                                 // numerical factorization
                                                 // solve Ax=b
 sp.solve(b. x):
  ... // check residual/error and cleanup
```

The main steps are: create solver object, set options (parse options from the command line), set matrix, reorder, factor and finally solve. The matrix should be in the Compressed Sparse Row (CSR) format, also called Yale format, with 0 based indices. Figure 1 illustrates the CSR format. In the basic scenario, it is not necessary to explicitly call reorder and factor, since trying to solve with a StrumpackSparseSolver object that is not factored yet, will internally call the factor routine, which will call reorder if necessary.

The above code should be linked with -lstrumpack and with the Metis, ParMetis, Scotch, PT-Scotch, BLAS, LAPACK, and ScaLAPACK libraries.

### 4.2 StrumpackSparseSolverMPI Example

Usage of the StrumpackSparseSolverMPI<scalar,integer=int> solver is very similar:

```
#include "StrumpackSparseSolverMPI.hpp"
using namespace strumpack;
int main(int argc, char* argv[]) {
 int thread level, rank;
 // StrumpackSparseSolverMPI uses OpenMP so we should ask for MPI_THREAD_FUNNELED at least
 MPI Init thread(&argc, &argv, MPI THREAD FUNNELED, &thread level);
 MPI Comm rank(MPI COMM WORLD, &rank);
 if (thread level != MPI THREAD FUNNELED && rank == 0)
    std::cout << "MPI_implementation_does_not_support_MPI_THREAD_FUNNELED" << std::endl;</pre>
   // define the same CSR matrix as for StrumpackSparseSolver
                     // construct an NxN CSR matrix with nnz nonzeros
   int N = ...;
   int* row_ptr = ...;  // N+1 integers
int* col_ind = ...;  // nnz integers
double* val = ...
   double* val = ...;
                               // nnz scalars
   // allocate entire solution and right-hand side vectors on each MPI process
   double* x = new double[N]; // will hold the solution vector
   double* b = ...;
                                // set a right-hand side b
   // construct solver and specify the MPI communicator
   StrumpackSparseSolverMPI<double> sp(MPI COMM WORLD);
   sp.options().set matching(MatchingJob::NONE);
   sp.options().set from command line(argc, argv);
   sp.set csr matrix(N, row ptr, col ind, val);
   sp.solve(b, x);
    ... // check residual/error, cleanup
 Cblacs_exit(1);
 MPI_Finalize();
```

The only difference here is the use of StrumpackSparseSolverMPI instead of StrumpackSparseSolver and the calls to MPI Init thread, Cblacs exit and MPI Finalize.

## 4.3 StrumpackSparseSolverMPIDist Example

Finally, we illustrate the usage of StrumpackSparseSolverMPIDist<scalar,integer=int> solver. This interface takes a block-row distributed compressed sparse row matrix as input. This matrix format is illustrated in Figure 2.

```
#include "StrumpackSparseSolverMPI.hpp"
using namespace strumpack;

int main(int argc, char* argv[]) {
   int thread_level, rank, P;
   MPI_Init_thread(&argc, &argv, MPI_THREAD_FUNNELED, &thread_level);
   MPI_Comm_rank(MPI_COMM_WORLD, &rank);
   MPI_Comm_size(MPI_COMM_WORLD, &P);
   {
      // define a block-row distributed CSR matrix
      int* dist = new int[P];
      // set dist such that processor p owns rows [dist[p], dist[p+1]) of the sparse matrix
      for (int p=0; p<P; p++) dist[p] = ...;
      // local_n is the number of rows of the input matrix assigned to me</pre>
```

$$A = \begin{bmatrix} 8.2 & 0.1 & & & 3.1 \\ \hline 0 & & -4.8 & & \\ 6.2 & 1.1 & & 2.6 & & \\ \hline & & -1.0 & & \\ & & 99.9 & 4.0 \end{bmatrix} P_0$$
 
$$\text{dist} = \begin{bmatrix} 0, 1, 3, 5 \end{bmatrix}$$
 
$$P_0$$
 
$$P_1$$
 
$$\text{row\_ptr} = \begin{bmatrix} 0, 3 \end{bmatrix}$$
 
$$\text{row\_ptr} = \begin{bmatrix} 0, 2, 5 \end{bmatrix}$$
 
$$\text{row\_ptr} = \begin{bmatrix} 0, 1, 3 \end{bmatrix}$$
 
$$\text{col\_ind} = \begin{bmatrix} 0, 1, 4 \end{bmatrix}$$
 
$$\text{col\_ind} = \begin{bmatrix} 0, 2 \mid 0, 1, 3 \end{bmatrix}$$
 
$$\text{col\_ind} = \begin{bmatrix} 2 \mid 3, 4 \end{bmatrix}$$
 
$$\text{values} = \begin{bmatrix} 8.2, 0.1, 3.1 \end{bmatrix}$$
 
$$\text{values} = \begin{bmatrix} 0, -4.8 \mid 6.2, 1.1, 2.6 \end{bmatrix}$$
 
$$\text{values} = \begin{bmatrix} -1.0 \mid 99.9, 4.0 \end{bmatrix}$$

Figure 2: Illustration of a small  $5 \times 5$  sparse matrix with 11 nonzeros and its block-row distributed compressed sparse row representation. We always use 0-based indexing! Process  $P_0$  owns row 0, process  $P_1$  has rows 1 and 2 and process  $P_2$  has rows 3 and 4. This distribution of rows over the processes is represented by the dist array. Process p owns rows [dist[p],dist[p+1]). If N=5 is the number of rows in the entire matrix and P is the total number of processes, then dist[P]=N. The (same) dist array is stored on every process. Each process holds a CSR representation of only its local rows of the matrix, see Figure 1.

```
int local n = dist[rank+1] - dist[rank];
 int* row ptr = new int[local n+1];
  .. // set the sparse matrix row pointers in row ptr
 int local_nnz = row_ptr[local_n+1] - row_ptr[0];
 int* col ind = new int[local nnz];
  .. // set the sparse matrix column indices in col ind
 double* val = new double[local nnz];
  ... // set the matrix nonzero value in val
 double* x
               = new double[local n];
                                             // local part of solution
               = new double[local_n];
 double* b
                                             // local part of rhs
 for (int i=0; i<local_n; i++) b[i] = ..;</pre>
                                             // set the rhs
 StrumpackSparseSolverMPIDist<double> sp(MPI_COMM_WORLD);
  sp.options().set reordering method(ReorderingStrategy::PARMETIS);
 sp.options().set from command line(argc, argv);
 sp.set_distributed_csr_matrix(local_n, row_ptr, col_ind, val, dist);
 sp.solve(b, x);
  ... // check residual/error, cleanup
Cblacs exit(1);
MPI_Finalize();
```

### 4.4 Initializing the Solver Object

Let

```
typedef strumpack::StrumpackSparseSolver<scalar,integer> Sp;
typedef strumpack::StrumpackSparseSolverMPI<scalar,integer> SpMPI;
typedef strumpack::StrumpackSparseSolverMPIDist<scalar,integer> SpMPIDist;
```

Each of the solver classes has two constructors:

```
Sp::StrumpackSparseSolver(bool verbose=true, bool root=true);
Sp::StrumpackSparseSolver(int argc, char* argv[], bool verbose=true, bool root=true);
```

```
SpMPI::StrumpackSparseSolverMPIDist(MPI_Comm comm, bool verbose=true);
SpMPI::StrumpackSparseSolverMPIDist(MPI_Comm comm, int argc, char* argv[], bool verbose=true);
```

```
SpMPIDist::StrumpackSparseSolverMPIDist(MPI_Comm comm, bool verbose=true);
SpMPIDist::StrumpackSparseSolverMPIDist(MPI_Comm comm, int argc, char* argv[], bool verbose=true);
```

where argc and argv contain the command line options and the verbose option can be set to false to suppress output of the solver. Note that since SpMPIDist is a subclass of SpMPI, which is a subclass of Sp, all public members of Sp are also members of SpMPI and SpMPIDist. The public interface to the SpMPI class is exactly the same as that for the Sp class.

### 4.5 Sparse Matrix Format

The sparse matrix should be specified in compressed sparse row format [8]:

Internally, the matrix is copied, so it will not be modified. Previous versions of STRUMPACK also supported the CSC format, but this is now deprecated. If the sparsity pattern of the matrix is symmetric (the values do not have to be symmetric), then you can set symmetric\_pattern=true. This saves some work in the setup phase of the solver.

For the SpMPIDist solver the input is a block-row distributed compressed sparse row matrix (as illustrated in the example above):

```
void SpMPIDist::set_distributed_csr_matrix
  (integer local_rows, integer* row_ptr, integer* col_ind,
    scalar* values, integer* dist, bool symmetric_pattern=false);
```

Alternatively, you can also specify a sequential CSR matrix to the SpMPIDist solver:

```
void SpMPIDist::set_csr_matrix
    (integer N, integer* row_ptr, integer* col_ind,
    scalar* values, bool symmetric_pattern=false);
```

For this routine, the matrix only needs to be specified completely on the root process. Other processes can pass NULL for the arrays.

## 4.6 Setting and Parsing Options

The solver class has an object of type SPOptions<scalar>, which can be accessed through:

```
SPOptions<scalar>& Sp::options();
```

The SPOptions<scalar> class is defined in SPOptions.hpp. The complete public interface for the SPOptions<scalar> class is given in Section 4.10.1. The following subsections describe some of the options available from SPOptions<scalar> in more detail.

## 4.7 Reordering

There are three types of matrix reordering: for numerical stability, to reduce fill-in and to reduce the HSS-ranks. These reorderings are all performed when calling

```
ReturnCode Sp::reorder();
```

The return value is of type ReturnCode (defined in strumpack parameters.hpp) and can be

```
enum class ReturnCode {
   SUCCESS,     /*!< Operation completed successfully. */
   MATRIX_NOT_SET,     /*!< The input matrix was not set.     */
   REORDERING_ERROR     /*!< The matrix reordering failed.     */
};</pre>
```

#### 4.7.1 Reordering for numerical stability

The reordering for numerical stability is performed using MC64 or Combinatorial BLAS. For many matrices, this reordering is not necessary and can safely be disabled! MC64 supports 5 different modes and there is one option to select the Combinatorial BLAS code:

```
enum class MatchingJob {
  NONE,
                                  /*!< Don't do anything
                                                                                                           */
  MAX CARDINALITY.
                                  /*!< Maximum cardinality</pre>
                                                                                                           */
  MAX SMALLEST DIAGONAL,
                                  /*!< Maximum smallest diagonal value</pre>
  MAX SMALLEST DIAGONAL 2,
                                  /*!< Same as MAX SMALLEST DIAGONAL, different algorithm
  MAX DIAGONAL SUM,
                                  /*!< Maximum sum of diagonal values</pre>
  MAX DIAGONAL PRODUCT SCALING, /*!< Maximum product of diagonal values and row and column scaling
  COMBBLAS
                                  /*!< Use AWPM from Combinatorial BLAS</pre>
};
```

which can be selected via

```
void SPOptions::set_matching(MatchingJob job);
MatchingJob SPOptions::matching() const;
```

where matching() queries the currently selected strategy (the default is MAX\_DIAGONAL\_PRODUCT\_SCALING maximum product of diagonal values plus row and column scaling). The command line option

```
--sp_matching [0-6]
```

can also be used, where the integers are defined as:

- 0: no reordering for stability, this disables MC64/matching
- 1: MC64(1): currently not supported
- 2: MC64(2): maximize the smallest diagonal value
- 3: MC64(3): maximize the smallest diagonal value, different strategy
- 4: MC64(4): maximize sum of diagonal values
- 5: MC64(5): maximize product of diagonal values and apply row and column scaling
- 6: Combinatorial BLAS: approximate weight perfect matching

The MC64 code is sequential, so when using this option in parallel, the graph is first gathered to the root process. The Combinatorial BLAS code can currently only be used in parallel, and only with a square number of processes.

#### 4.7.2 Nested dissection reordering

The STRUMPACK sparse solver supports both (Par)Metis and (PT-)Scotch for the matrix reordering. The following functions can set the preferred method or check the currently selected method:

```
void SPOptions::set_reordering_method(ReorderingStrategy m);
ReorderingStrategy SPOptions::reordering_method() const;
```

The options for MatrixReorderingStrategy are

```
enum class ReorderingStrategy {
  NATURAL.
              /*!< Do not reorder the system
  METIS,
              /*!< Use Metis nested-dissection reordering</pre>
  PARMETIS,
              /*!< Use ParMetis nested-dissection reordering
  SCOTCH,
              /*!< Use Scotch nested-dissection reordering
  PTSCOTCH,
              /*!< Use PT-Scotch nested-dissection reordering
              /*!< Use RCM reordering</pre>
  RCM.
  GEOMETRIC
              /*!< A simple geometric nested dissection code that
                only works for regular meshes. (see Sp::reorder)
};
```

When the solver is an object of Sp, PARMETIS or PTSCOTCH are not supported. When the solver is parallel, either an SpMPI or SpMPIDist object, and METIS, SCOTCH or RCM are chosen, then the graph of the complete matrix will be gathered onto the root process and the root process will call the (sequential) Metis, Scotch or RCM reordering routine. For large graphs this might fail due to insufficient memory.

The GEOMETRIC option is only allowed for regular grids. In this case, the dimensions of the grid should be specified in the function

```
ReturnCode Sp::reorder(int nx=1, int ny=1, int nz=1);
```

For instance for a regular 2d  $2000 \times 4000$  grid, you can call this as sp.reorder(2000, 4000). In the general algebraic case, the grid dimensions don't have to be provided. The reordering method can also be specified via the command line option

--sp\_reordering\_method [metis|parmetis|scotch|ptscotch|geometric|rcm]

#### 4.8 Factorization

Compute the factorization by calling

```
ReturnCode Sp::factor();
```

where the possible return values are the same as for Sp::reorder(). If Sp::reorder() was not called already, it is called automatically. When HSS compression is not enabled, this will compute an exact LU factorization of the (permuted) sparse input matrix. If HSS compression is enabled (with SPOptions::enable\_HSS() or --sp\_enable\_HSS, see Section 5), the factorization is only approximate.

#### 4.9 Solve

Solve the linear system Ax = b by calling

```
ReturnCode Sp::solve(scalar* b, scalar* x, bool use_initial_guess=false);
```

By default (bool use\_initial\_guess=false) the input in x is ignored. If bool use\_initial\_guess=true, x is used as initial guess for the iterative solver (if an iterative solver is used, for instance iterative refinement or GMRES). If the Sp::factor() was not called, it is called automatically. The return values are the same as for Sp::reorder().

The iterative solver can be chosen through:

```
void SPOptions::set_Krylov_solver(KrylovSolver s);
```

where KrylovSolver can take the following values:

```
enum class KrylovSolver {
  AUTO,
                 /*!< Use iterative refinement if no HSS compression is used, otherwise PGMRES.
  DIRECT,
                 /*!< No outer iterative solver, just a single application of the multifrontal solver.
  REFINE.
                 /*!< Iterative refinement.</pre>
  PREC GMRES,
                 /*!< Preconditioned GMRES. The preconditioner is the (approx) multifrontal solver.
                                                                                                            */
  GMRES.
                 /*!< UN-preconditioned GMRES. (for testing mainly)</pre>
  PREC_BICGSTAB, /*!< Preconditioned BiCGStab. The preconditioner is the (approx) multifrontal solver.
  BTCGSTAB
                 /*!< UN-preconditioned BiCGStab. (for testing mainly)</pre>
};
```

with KrylovSolver::AUTO being the default value. The KrylovSolver::AUTO setting will use iterative refinement when HSS compression is not enabled, and preconditioned GMRES when HSS compression is enabled, see Section 5. To use the solver as a preconditioner, or a single (approximate) solve, set the solver to KrylovSolver::DIRECT. When calling SpMPIDist::solve, the right-hand side and solution vectors should only point to the local parts!

### 4.10 All Options for the Sparse Solver

The HSS specific options are stored in an object of type HSSOptions<scalar>, inside the SPOptions object. These options are described in Section 5.

#### 4.10.1 SPOptions<scalar> Interface

The complete public interface to the options class is as follows, wher the real type is the real part of a scalar, i.e., decltype(std::real(scalar(0))).

```
template<typename scalar> class SPOptions {
public:
 SPOptions();
 SPOptions(int argc, char* argv[]);
 /* print statistics?
 void set_verbose(bool verbose);
                                                        bool verbose() const;
  /* maximum iterations in iterative solver
 void set_maxit(int maxit);
                                                        int maxit() const;
  /* relative residual stopping criterion for iterative solver
                                                        real rel_tol() const;
 void set_rel_tol(real rel_tol);
  /* absolute residual stopping criterion for iterative solver
                                                                   */
 void set_abs_tol(real abs_tol);
                                                        real abs_tol() const;
  /* type of iterative solver to use, see section 4.9
                                                                   */
 void set_Krylov_solver(KrylovSolver s);
                                                        KrylovSolver Krylov_solver() const;
  /* GMRES restart
 void set gmres restart(int m);
                                                        int gmres_restart() const;
 /* type of Gram-Schmidt used in GMRES
                                                        GramSchmidtType GramSchmidt_type() const;
 void set_GramSchmidt_type(GramSchmidtType t);
 /* nested-dissection code, see section 4.7.2
 void set_reordering_method(ReorderingStrategy m);
                                                        ReorderingStrategy reordering_method() const;
  /* stop nested-dissection when domains are smaller than nd_param */
 void set nd param(int nd param);
                                                        int nd param() const;
  /* use the internal (undocumented) metis routine METIS NodeNDP
 void enable METIS NodeNDP();
                                                        bool use_METIS_NodeNDP() const;
  /* do not use METIS_NodeNDP, use METIS_NodeND instead
```

```
void disable METIS NodeNDP();
/* use METIS NodeND instead of METIS NodeDNP
                                                      bool use METIS NodeND() const;
void enable METIS NodeND();
/* do not use METIS_NodeND, use METIS_NodeNDP instead
                                                                 */
void disable METIS NodeND();
/* build the supernodal tree using the MUMPS SYMQAMD code
                                                                 */
void enable MUMPS SYMQAMD();
                                                      bool use MUMPS SYMQAMD() const;
/* do not use MUMPS SYMQAMD, use fundamental supernodes
                                                                 */
void disable MUMPS SYMQAMD();
                                                                 */
/* when using MUMPS_SYMQAMS, enable aggressive amalgamation
void enable agg amalg();
                                                      bool use agg amalg() const;
/* when using MUMPS SYMQAMS, disable aggressive amalgamation
void disable agg amalg();
/* set the job to be used for static pivoting, see section 4.7.1 */
void set matching(MatchingJob job);
                                                      MatchingJob matching() const;
/* not used at the moment
void enable assembly tree log();
                                                      bool log_assembly_tree() const;
void disable assembly tree log();
/* enable HSS compression, see section 5
                                                                  */
void enable HSS();
                                                      bool use HSS() const;
/* disable HSS compression
                                                                  */
void disable HSS();
                                                                  */
/* set the minimum size of a separator for HSS compression
void set_HSS_min_sep_size(int s);
                                                     int HSS_min_sep_size() const;
/* set level to 1 to enable length 2 connections in the separator
 * before computing separator reordering to reduce HSS ranks.
 * Set to to disable length 2 connections.
void set_separator_ordering_level(int l);
                                                    int separator_ordering_level() const;
/* best not to touch this
void enable_indirect_sampling();
void disable indirect sampling();
                                                     bool indirect sampling() const;
                                                      bool replace_tiny_pivots() const;
void enable_replace_tiny_pivots();
void disable_replace_tiny_pivots(;
                                                                  */
/* get the HSS specific options, see section 5.
const HSS::HSSOptions<scalar>& HSS_options() const;
HSS::HSSOptions<scalar>& HSS_options();
/* parse the options in argc/argv set in the constructor
void set_from_command_line();
/* parse the options in argc/argv
void set_from_command_line(int argc, char* argv[]);
/* print out message listing all command line options
void describe_options() const;
```

This uses the following (scoped) enumeration for the Gram-Schmidt type used in GMRES:

```
enum class GramSchmidtType {
   CLASSICAL, /*!< Classical Gram-Schmidt is faster, more scalable. */
   MODIFIED /*!< Modified Gram-Schmidt is slower, but stable. */
};</pre>
```



Figure 3: Illustration of a Hierarchically Semi-Separable (HSS) matrix. Gray blocks are dense matrices. Off-diagonal blocks, on different levels of the HSS hierarchy, are low-rank. The low-rank factors of off-diagonal blocks of different levels are related.

#### 4.10.2 Command Line Options

To get a list of all available options, make sure to pass "int argc, char\* argv[]" when initializing the StrumpackSparseSolver or when calling SPOptions::set\_from\_command\_line and run the application with --help or -h. Some default values listed here are for double precision and might be different when running in single precision.

```
STRUMPACK options:
```

```
--sp_maxit int (default 5000)
--sp_rel_tol real (default 1e-06)
--sp_abs_tol real (default 1e-10)
--sp Krylov solver auto|direct|refinement|pgmres|gmres|pbicgstab|bicgstab
--sp gmres restart int (default 30)
--sp GramSchmidt type modified|classical
--sp_reordering_method natural|metis|scotch|parmetis|ptscotch|rcm|geometric
--sp_nd_param int (default 8)
--sp_enable_METIS_NodeNDP (default true)
--sp disable METIS NodeNDP (default false)
--sp enable METIS NodeND (default false)
--sp disable METIS NodeND (default true)
--sp enable MUMPS SYMQAMD (default false)
--sp_disable_MUMPS_SYMQAMD (default true)
--sp_enable_agg_amalg (default false)
--sp_disable_agg_amalg (default true)
--sp matching 0-6 (default 5)
--sp enable hss (default false)
--sp disable hss (default true)
--sp hss min sep size int (default 256)
--sp_separator_ordering_level (default 1)
--sp enable indirect sampling
--sp disable indirect sampling
--sp enable replace tiny pivots
--sp disable_replace_tiny_pivots
--sp_verbose or -v (default true)
--sp_quiet or -q (default false)
--help or -h
```

#### 5 HSS Preconditioning

The sparse multifrontal solver can optionally use Hierarchically Semi-Separable, rank-structured matrices to compress the fill-in. In the multifrontal method, computations are performed on dense matrices called frontal matrices. A frontal matrix can be approximated as an HSS matrix, but this will only be beneficial (compared to storing the frontal as a standard dense matrix and operating on it with BLAS/LAPACK routines) if the frontal matrix is large enough.

Figure 3 illustrates the HSS matrix format. The matrix is partitioned as a  $2 \times 2$  block matrix, with the partitioning recursively applied on the diagonal blocks, until diagonal blocks are smaller than a specified leaf size. The off-diagonal block on each level of the hierarchy are approximated by a low-rank product. This low-rank storage format asymptotically reduces memory usage and floating point operations, while introducing approximation errors. HSS compression is not used by default in the STRUMPACK sparse solver (the default is to perform exact LU factorization), but can be turned on/off via the command line:

```
--sp_enable_hss
                  (no argument)
--sp_disable_hss
                   (no argument)
```

or via the C++ API as follows

```
void SPOptions<scalar>::enable HSS();
void SPOptions<scalar>::disable HSS();
bool SPOptions<scalar>::use HSS();
                                       // check whether HSS compression is enabled
```

When HSS compression is enabled, the default STRUMPACK behavior is to use the HSS enabled approximate LU factorization as a preconditioner within GMRES. This behavior can also be changed, see Section 4.9.

However, HSS compression has a considerable overhead and only pays off for sufficiently large matrices. Therefore STRUMPACK has a tuning parameter to specify the minimum size a dense matrix needs to be to be considered a candidate for HSS compression. The minimum dense matrix size for HSS compression is set via the command line via

```
--sp hss min sep size int (default 256)
or via the C++ API as follows
  void SPOptions<scalar>::set HSS min sep size(int s);
```

int SPOptions<scalar>::HSS\_min\_sep\_size() const;

```
The routine set_HSS_min_sep_size(int s) refers to the size of the top-left block of the front only. This top-left
block is the part that corresponds to a separator, as given by the nested dissection reordering algorithm.
```

// get the current value

This top-left block is also referred to as the block containing the fully-summed variable. Factorization (LU in the dense case, ULV in the HSS case) is only applied to this top-left block. Tuning the value for the minimum separator size can have a big impact on performance and memory usage!

The above options affect the use of HSS within the multifrontal solver. There are more, HSS specific, options which are stored in an object of type HSS::HSSOptions<scalar>. An object of this type is stored in the SPOptions<scalar> object stored in the StrumpackSparseSolver. It can be accessed via the HSS options() routine as follows:

```
StrumpackSparseSolver<double> sp;
                                                // create solver object
sp.options().enable HSS();
                                                // enable HSS compression in the multifrontal solver
sp.options().HSS options().set leaf size(256); // set the HSS leaf size
```

In STRUMPACK, HSS matrices are constructed using a randomized sampling algorithm [6]. To construct an HSS approximation for a matrix A, sampling of the rows and columns of A is computed by multiplication with a tall and skinny random matrix R as follows:  $S^r = AR$  and  $S^c = A^*R$ . Ideally, the number of columns in the matrix R is d = r + p, with r the maximum off-diagonal block rank in the HSS matrix and p a small oversampling parameter. Unfortunately, the HSS rank is not known a-priori, so it needs to determined adaptively. The adaptive sampling scheme used in STRUMPACK starts with an initial number of random vector  $d_0$ , and increases this in steps of  $\Delta d$ , until the compression quality reaches the desired user specified tolerance, or until the maximum rank is reached. The compression tolerances can greatly impact performance. They can be set using:

```
--hss_rel_tol real (default 0.01)
--hss_abs_tol real (default 1e-08)
or via the C++ API
```

```
void HSSOptions<scalar>::set_rel_tol(real rel_tol);
void HSSOptions<scalar>::set_abs_tol(real abs_tol);
real HSSOptions<scalar>::rel_tol() const; // get the current value
real HSSOptions<scalar>::abs_tol() const;
```

## 5.1 HSSOptions<scalar> Interface

Other options are available to tune for instance the initial number of random vectors  $d_0$ , the increment  $\Delta d$ , the random number generator or the random number distribution. The complete public interface for the HSSOptions<scalar> class is:

```
template<typename scalar> class HSSOptions {
public:
 /* relative compression tolerance
 void set rel tol(real rel tol);
                                            real rel_tol() const;
                                                          */
 /* absolute compression tolerance
 void set_abs_tol(real abs_tol);
                                            real abs_tol() const;
 /* size of the smallest blocks in the HSS hierarchy
                                                        */
 void set_leaf_size(int leaf_size);
                                            int leaf_size() const;
 /* initial number of random vectors used in the
     adaptive randomized compression algorithm
 void set d0(int d0);
                                            int d0() const;
 /* number of random vectors added in each step of the
     adaptive randomized HSS compression algorithm
 void set_dd(int dd);
                                            int dd() const;
 /* currently not used
                                            int q() const;
 void set q(int q);
 /* maximum rank in the HSS representation
 void set_max_rank(int max_rank);
                                            int max_rank() const;
 /* random engine/generator to use, see below
 void set_random_engine(random::RandomEngine random_engine);
 random::RandomEngine random engine() const;
 /* the random number distribution, see below
 void set random distribution
  (random::RandomDistribution random distribution);
 random::RandomDistribution random_distribution() const;
 /st the compression algorithm to use
 void set_compression_algorithm(CompressionAlgorithm a);
 CompressionAlgorithm compression_algorithm() const;
 /* for expert users
 void set_user_defined_random(bool user_defined_random);
 bool user defined random() const;
 /* for expert users
 void set_synchronized_compression(bool sync);
 bool synchronized compression() const;
 /* currently not used
 void set log ranks(bool log ranks);
                                            bool log ranks() const;
```

## 5.2 HSS Command Line Options

The HSS specific command line options are:

```
HSS Options:
  --hss rel tol real (default 0.01)
  --hss abs tol real (default 1e-08)
  --hss leaf size int (default 128)
  --hss d0 int (default 128)
  --hss dd int (default 32)
  --hss_q int (default 0)
  --hss max rank int (default 5000)
  --hss random distribution normal|uniform (default normal(0,1))
  --hss random engine linear|mersenne (default minstd rand)
  --hss compression algorithm original|stable (default stable)
  --hss user defined random (default false)
  --hss enable sync (default true)
  --hss_disable_sync (default false)
  --hss log ranks (default false)
  --hss verbose or -v (default false)
  --hss quiet or -q (default true)
  --help or -h
```

## 6 HSS Approximation of Dense Matrices

The HSS code can be found in the src/HSS/ subdirectory. All HSS code is in a namespace called HSS. The class for a sequential/multithreaded HSS matrix is HSSMatrix<scalar>, while the distributed memory HSS class is HSSMatrix<scalar>. For examples of the usage of these classes, see the test code in test/test\_HSS\_seq.cpp and test/test\_HSS\_mpi.cpp respectively. There is also one sequential example in examples/MLkernel.cpp, which uses HSS compression for kernel matrices as used in certain machine learning applications, see for instance [2]. This part of the code will be better documented in future STRUMPACK releases.

## 7 Examples

A number of examples are available in the examples/ folder. This folder, with the sources, is copied to the build directory when running cmake. However, the examples are not part of the CMake build system, so they will not be compiled and linked when running make. A simple Makefile is generated in the examples/ folder in the build directory, with the compiler setting based on what CMake has discovered. However, this Makefile is rather simplistic and in certain cases it might need some manual edits. However, it can serve as an example for how to compile code using STRUMPACK. Check the README file in the examples/ directory for more details.

## 8 C Interface

The C interface is defined in the header file StrumpackSparseSolver.h and is very similar to the C++ interface. For example usage see the programs sexample.c, dexample.c, cexample.c and zexample.c in the examples/directory, for simple single and double precision real and complex example programs. Note that since the STRUMPACK code is written in C++, even when using the C interface you should link with a C++ aware linker or link with the standard C++ library. For instance when using the GNU toolchain, link with g++ instead of gcc or link with gcc and include -lstdc++.

## 9 Advanced Usage Tips

- To keep track of the number of floating point operations performed in the STRUMPACK Sparse Solver, you can run CMake with -DSTRUMPACK\_COUNT\_FLOPS=ON. Then, when running, do not set the quiet flag in the StrumpackSparseSolver constructor or on the command line and the solver will print some statistics. This will also enable a counter for data movement in the solve phase, from which the (approximately) attained bandwidth usage is derived. This is done because the solve phase is typically bandwidth limited, while the factorization is flop limited.
- There is also some support for PAPI. Run CMake with -DSTRUMPACK\_USE\_PAPI=ON and specify the PAPI include folders and libraries via -DPAPI INCLUDES=... and -DPAPI LIBRARIES=...
- We have added timers all throughout the code. These can be enabled with -DSTRUMPACK\_TASK\_TIMERS=ON. Running the code will generate a file time.log. A script to visualize these timings is provided.
- If you compile with MKL or OpenBLAS, you can take advantage of some extra optimized routines by specifying -D\_HAVE\_MKL or -D\_HAVE\_OPENBLAS respectively.
- The code is not completely thread safe at the moment: do not call solve on the same StrumpackSparseSolve object from different threads simultaneously.
- For comments, feature requests or bug reports: {pghysels,xsli,gichavez}@lbl.gov

## 10 FAQ

• Help, I get this compilation error:

catastrophic error: cannot open source file "chrono"
#include <chrono>

You need a C++11 capable compiler, and also a C++11 enabled standard library. For instance suppose you are using the Intel 15.0 C++ compiler with GCC 4.4 headers. The Intel 15.0 C++ compiler supports the C++11 standard, but the GCC 4.4 headers do not implement the C++11 standard library. You should install/load a newer GCC version (or just the headers). On cray machines, this can be done with module unload gcc; module load gcc/4.9.3 for instance.

• When running make test, many of the tests fail!

The parallel execution in ctest is invoked by the MPIEXEC command as discovered by CMake. On many HPC clusters, this does not run unless it is executed from within a batch script. In this case all parallel tests will fail. At the moment, a small number of tests still fail. This is normal behavior.

## 11 Acknowledgements

The new, more robust version of the adaptive randomized HSS compression algorithm was developed in collaboration with Theo Mary from the Université de Toulouse.

The code for the STRUMPACK-sparse is based on the sequential code StruMF, originally developed by Artem Napov. We wish to thank people who sent us test problems and helped testing the code: Alex Druinsky, Yvan Notay and Shen Wang.

Partial support for this work was provided through Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research (and Basic Energy Sciences/Biological and Environmental Research/High Energy Physics/Fusion Energy Sciences/Nuclear Physics). This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

## 12 Copyright notice

STRUMPACK – STRUctured Matrices PACKage, Copyright (c) 2014, The Regents of the University of California, through Lawrence Berkeley National Laboratory (subject to receipt of any required approvals from the U.S. Dept. of Energy). All rights reserved.

If you have questions about your rights to use or distribute this software, please contact Berkeley Lab's Technology Transfer Department at TTD@lbl.gov.

NOTICE. This software is owned by the U.S. Department of Energy. As such, the U.S. Government has been granted for itself and others acting on its behalf a paid-up, nonexclusive, irrevocable, worldwide license in the Software to reproduce, prepare derivative works, and perform publicly and display publicly. Beginning five (5) years after the date permission to assert copyright is obtained from the U.S. Department of Energy, and subject to any subsequent five (5) year renewals, the U.S. Government is granted for itself and others acting on its behalf a paid-up, nonexclusive, irrevocable, worldwide license in the Software to reproduce, prepare derivative works, distribute copies to the public, perform publicly and display publicly, and to permit others to do so.

## 13 License agreement

"STRUMPACK – STRUctured Matrices PACKage, Copyright (c) 2014, The Regents of the University of California, through Lawrence Berkeley National Laboratory (subject to receipt of any required approvals from the U.S. Dept. of Energy). All rights reserved."

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 following disclaimer.
- 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- 3. Neither the name of the University of California, Lawrence Berkeley National Laboratory, U.S. Dept. of Energy 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 THE COPYRIGHT OWNER 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.

You are under no obligation whatsoever to provide any bug fixes, patches, or upgrades to the features, functionality or performance of the source code ("Enhancements") to anyone; however, if you choose to make your Enhancements available either publicly, or directly to Lawrence Berkeley National Laboratory, without imposing a separate written license agreement for such Enhancements, then you hereby grant the following license: a non-exclusive, royalty-free perpetual license to install, use, modify, prepare derivative works, incorporate into other computer software, distribute, and sublicense such enhancements or derivative works thereof, in binary and source code form.

## References

- [1] A. Buluç and J. R. Gilbert, *The Combinatorial BLAS: Design, implementation, and applications*, The International Journal of High Performance Computing Applications, 25 (2011), pp. 496–509.
- [2] G. B. D YU CHENHAN, WILLIAM B MARCH, An N log N Parallel Fast Direct Solver for Kernel Matrices, in 2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS), IEEE, May 2017, pp. 886–896.
- [3] I. S. Duff and J. Koster, The design and use of algorithms for permuting large entries to the diagonal of sparse matrices, SIAM Journal on Matrix Analysis and Applications, 20 (1999), pp. 889–901.
- [4] P. GHYSELS, X. S. LI, C. GORMAN, AND F. H. ROUET, A robust parallel preconditioner for indefinite systems using hierarchical matrices and randomized sampling, in 2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS), IEEE, May 2017, pp. 897–906.
- [5] P. GHYSELS, X. S. LI, F.-H. ROUET, S. WILLIAMS, AND A. NAPOV, An efficient multi-core implementation of a novel HSS-structured multifrontal solver using randomized sampling, Submitted to SIAM SISC, (2015).
- [6] P.-G. Martinsson, A fast randomized algorithm for computing a hierarchically semiseparable representation of a matrix, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 1251–1274.
- [7] F.-H. ROUET, X. S. LI, AND P. GHYSELS, A distributed-memory package for dense hierarchically semi-separable matrix computations using randomization, ACM Transactions on Mathematical Software, (2016). to appear.
- [8] Y. Saad, Iterative methods for sparse linear systems, Society for Industrial Mathematics, 2003.
- [9] J. XIA, Randomized sparse direct solvers, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 197–227.