

Large-Scale direct numerical simulations of turbulence using GPUs and modern Fortran

Martin Karp¹ , Daniele Massaro² , Niclas Jansson³ ,
Alistair Hart⁴, Jacob Wahlgren¹, Philipp Schlatter² , and
Stefano Markidis¹

The International Journal of High
Performance Computing Applications
2023, Vol. 0(0) 1–16
© The Author(s) 2023



Article reuse guidelines:

sagepub.com/journals-permissions
DOI: 10.1177/10943420231158616
journals.sagepub.com/home/hpc



Abstract

We present our approach to making direct numerical simulations of turbulence with applications in sustainable shipping. We use modern Fortran and the spectral element method to leverage and scale on supercomputers powered by the Nvidia A100 and the recent AMD Instinct MI250X GPUs, while still providing support for user software developed in Fortran. We demonstrate the efficiency of our approach by performing the world's first direct numerical simulation of the flow around a Flettner rotor at $Re = 30,000$ and its interaction with a turbulent boundary layer. We present a performance comparison between the AMD Instinct MI250X and Nvidia A100 GPUs for scalable computational fluid dynamics. Our results show that one MI250X offers performance on par with two A100 GPUs and has a similar power efficiency based on readings from on-chip energy sensors.

Keywords

GPU, Fortran, spectral element method, computational fluid dynamics, direct numerical simulation, high performance computing

1. Introduction

High-performance computing resources drive discoveries in fields of crucial importance for society. Environmental sustainability and energy consumption reduction are societal issues of increasing importance. Our dependence on non-renewable energy has become an increasingly pressing issue due to climate change and increasing global temperatures (Pachauri and Reisinger, 2007).

Global shipping is responsible for more than 2% of all carbon emissions and the largest cargo ships in the world make up a significant portion of these emissions (Smith et al., 2015). As the number of tonnes of goods transported by waterway is only expected to increase in the coming years, reducing fuel consumption will have a significant impact on both costs, reliability, and emissions (Smith et al., 2015). Active research is concerned with several unconventional designs being considered to make ships more energy efficient. One of these approaches is the Flettner rotor, a rotating cylinder that can reduce the fuel consumption in ships by more than 20% by using the Magnus force, effectively working as a sail (Magnus 1853; Seddiek and Ammar 2021). Rotors are now being tested in practice on ships as illustrated in Figure 1. However, as far as numerical simulations go we

have so far been limited to less detailed and accurate simulations based on the Reynolds-Averaged Navier-Stokes (RANS) approach (De Marco et al., 2016). Direct numerical simulations (DNS) that resolve all scales of the flow in both time and space have not been used because of their high computational cost, both concerning time and power usage.

Modern computing clusters are now approaching a size and power efficiency where DNS is possible within a reasonable amount of time and number of nodes. Especially,

¹Division of Computational Science and Technology, EECS, KTH Royal Institute of Technology, Stockholm, Sweden

²SimEx/FLOW, Engineering Mechanics, KTH Royal Institute of Technology, Stockholm, Sweden

³PDC Centre for High Performance Computing, EECS, KTH Royal Institute of Technology, Stockholm, Sweden

⁴Hewlett Packard Enterprise (HPE), UK

Corresponding author:

Martin Karp, Division of Computational Science and Technology, EECS, KTH Royal Institute of Technology, Lindstedtsvägen 5, Stockholm 100 44, Sweden.

Email: makarp@kth.se

Correction (April 2023): The Funding statement has been revised in this article after its original publication.

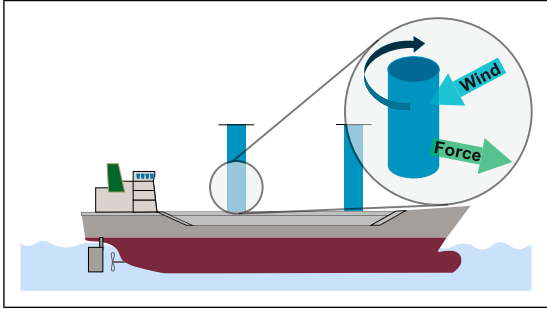


Figure 1. Illustration of a rotor ship which uses two Flettner rotors (in blue) to reduce fuel consumption and carbon emissions. We also highlight how the rotor and the wind coming from the side of the ship generates a force directed forward.

as Graphics Processing Units (GPUs) have become commonplace and outmatch older computing solutions with regards to power efficiency and performance (9 out of the top 10 Green 500 computers are powered by GPUs in November 2021 (Green500, 2021)). However, while our computer systems are more performant and energy-efficient than ever, computer systems have also grown more complex and heterogeneous.

The increased complexity of modern computer systems has limited the usage of trusted legacy software that scales and performs exemplary on multi-core processors. Therefore, several new frameworks that automate parts of the porting process to heterogeneous hardware, or non-intrusive pragma based approaches, have been introduced (Medina et al., 2014; Keryell et al., 2015; Wienke et al., 2012). However, these packages either require large code changes if the legacy code is written in Fortran 77, or offers relatively low performance. We take another approach, where we use modern Fortran to accommodate various backends, but not be dependent on any single overarching software package. Instead we aim to provide the flexibility to use the most suitable software for a given device. With the Fortran core, we can also efficiently utilize older, verified, CPU code without major changes. We base our solver on the numerical methods and schemes from the spectral element method and Nek5000 (Fischer et al., 2008) and leverage them in Neko – a portable spectral element framework targeting several computer architectures (Jansson et al., 2021).

Neko can both exploit the computing power of modern accelerators and use older tools for Nek5000. Using the spectral element method, Neko provides high-order convergence on unstructured meshes while having a low communication cost and high performance. In this work, we present our development and usage of Neko to leverage two state-of-the-art GPUs, the Nvidia A100 and the very recent AMD Instinct MI250X. In addition, we interface with older toolboxes written for Fortran 77. We present the efficiency of our solver by making the first DNS of a Flettner rotor in a turbulent

boundary layer. We illustrate how our developments lead to energy and time savings, both by strong-scaling and reducing the run time of the simulation, but also how a solver in modern Fortran can make the porting process to new hardware easier.

Our contribution is the following:

- We illustrate the software and algorithmic considerations in developing Neko, a highly scalable solver that caters both to increasingly heterogeneous supercomputers and legacy user software.
- We present the first direct numerical simulation of a Flettner rotor and its interaction with a turbulent boundary layer at Reynolds number of 30,000.
- We show the scaling, performance, and power measurements of our simulations on both the new AMD Instinct MI250X and Nvidia A100 GPUs, showing a matching power efficiency and that one MI250X performs similarly to two A100s.

The paper is organized as follows, in Section 2 we cover the related work, and in Section 3 we provide some background for the Flettner rotor. In Section 4, we describe Neko, its numerical methods and software design. In Section 5 we present our simulation setup and computer platforms used for our experiments. We then go on to present the simulation results and compare the performance and power efficiency of our computer platforms. In the final section we make our concluding remarks.

2. Related work

Our work targets high-fidelity incompressible flow simulations and is based on a spectral element method. The spectral element method was first introduced for CFD by Patera (1984); Maday and Patera (1989). The method has then been developed and derived into several different frameworks such as Nektar++ (Cantwell et al., 2015) and Nek5000 (Fischer et al., 2008). For DNS of turbulence in moderately complex geometries, Nek5000 has been well used because of its extreme scalability and accuracy but uses static memory and Fortran 77, making it most suitable for CPUs (Tufo and Fischer 1999), although acceleration using OpenACC directives has been considered (Otero et al., 2019). Other methods that provide similar functionality, in that they provide low numerical dissipation and accommodate complex geometries fall into the family of high-order finite element methods, where the spectral element method is one of them. Software in this category are, for example, deal.II (Arndt et al., 2020) and MFEM (Anderson et al., 2021). Neko shares its roots with Nek5000, but unlike Nek5000, Neko makes use of object-oriented modern Fortran, dynamic memory, and uses a new communication backend. Neko also provides a device-abstraction layer, where the underlying

implementation for the accelerator is hidden to the user. Beneath this layer, Neko uses optimized kernels written in CUDA, HIP, OpenCL for the highest performance, but can be easily extended to other programming models such as Kokkos, SYCL or RAJA (Edwards et al., 2014; Reyes and Lomüller 2016; Hornung and Keasler 2014) without changing the entire codebase, as illustrated with MAMBA (Dykes et al., 2021). Another approach based on Nek5000, NekRS (Fischer et al., 2021), also provides GPU support, but is rewritten in C++ and uses OCCA extensively to generate code for different backends (Medina et al., 2014). As our simulation makes use of the openly available KTH Toolboxes originally developed for Nek5000 (Peplinski 2022), we could with Neko’s device layer make very limited changes to port the necessary parts to Neko and GPUs.

3. Flettner rotor

In this work, we consider the flow around a Flettner rotor through numerical simulation with Neko. The Flettner rotor is named after the German inventor Anton Flettner and uses the Magnus effect to generate lift, a force perpendicular to the incoming flow. The physical phenomenon has been known in the literature since the 19th century and was introduced first by Magnus (1853). The Magnus effect consists of the generation of sidewise force by a rotating object, for example, a cylinder or sphere, submerged in a flow. The force results from the velocity changes induced by the spinning motion. The rotation leads to a pressure difference at the cylinder’s opposite sides, generating a force pointing from the high to the low-pressure side. In a rotor ship, the cylinder stands vertically and is mechanically driven to develop lift in the direction of the ship, as illustrated in Figure 1 and in the video by Massaro et al. (2022). Over the last century, various designs using this effect have been proposed for both ships and aeroplanes, but with limited success (Seifert 2012). Recently, Flettner rotors have gained renewed interest as a potential alternative to sails to reduce the fuel consumption of ships on our way towards net-zero carbon emissions. This is already being tested in practice for a select number of shipping routes. Because of Flettner rotors’ promise in reducing carbon emissions, the DNS of a Flettner rotor and its interaction with a turbulent boundary layer are of fundamental interest.

Flettner rotors have recently been studied by conducting massive experimental campaigns in a large wind tunnel (Bordogna et al., 2019, 2020). Numerical simulations allow to set up of a virtual wind tunnel, where the instantaneous solution of the entire flow field is available, unlike wind tunnel experiments where only some local probes or in-plane particle images velocimetry (PIV) measurements are extracted. Despite the growing interest in this rotor configuration, the literature is limited to RANS simulations (see, e.g., De Marco et al. (2016)). In these numerical

studies, only the mean flow dynamics is solved and a turbulence model needs to be introduced to close the problem. The presence of a model itself compromises the level of fidelity and accuracy. We present here the first DNS for the flow around a spinning cylinder immersed in a turbulent boundary layer, and no periodic boundary conditions are imposed. We do not rely on any model and all the scales are spatially and temporally resolved. While the lift generation mechanism has been well known since the last century (Magnus 1853), several physical aspects remain unclear, for example, the boundary layer interaction with the rotating cylinder, and are only possible to study through high-fidelity simulations.

4. Neko

Neko is a spectral element framework targeting incompressible flows. With its roots and main inspiration in Nek5000, a CFD solver widely acclaimed for its numerical methods and scalability on multicore processors (Tufo and Fischer 1999), Neko is a new code making use of object-oriented Fortran to control memory allocation and allow multi-tier abstractions of the solver stack. Neko has been used to accommodate various hardware backends ranging from general-purpose processors, accelerators, vector processors (Jansson 2021), as well as (limited) FPGA support (Karp et al., 2021, 2022b). In our work, we extend and optimize Neko to perform a large-scale DNS of practical industrial relevance on GPUs. The spectral element method used in Neko is a high-order, matrix-free, finite element method which provides high-order convergence on unstructured meshes. Here we provide a brief overview of Neko, the spectral element discretization, time-stepping schemes, numerical solvers, and an overview of the overarching software structure.

4.1. Discretization

We integrate the incompressible, non-dimensional, Navier-Stokes equation in time

$$\nabla \cdot \mathbf{v} = 0, \quad (1)$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \frac{1}{Re} \nabla^2 \mathbf{v} + \mathbf{F} \quad (2)$$

where $Re = UL/\nu$ is a Reynolds number defined by a characteristic velocity U , length scale L , and kinematic viscosity ν , \mathbf{v} is the instantaneous velocity field, p the non-dimensional pressure and \mathbf{F} an external forcing/source term. We solve the problem on a domain Ω with suitable boundary conditions on $\partial\Omega$. As this system of equations is coupled, we use the so-called $P_N - P_N$ splitting scheme to separate the pressure and velocity solves at each time step.

This splitting scheme was originally proposed by Karniadakis et al. (1991) and has shown to have the same order of convergence as the time stepping scheme that is used, with the error being primarily isolated to the boundaries of the domain. The $P_N - P_N$ scheme discretizes the momentum equation in time in the following fashion

$$\frac{\gamma_0 \mathbf{v}^{n+1}}{\Delta t} = \frac{\hat{\mathbf{v}}}{\Delta t} - \nabla p^{n+1} + \frac{1}{\text{Re}} \nabla^2 \mathbf{v}^{n+1} \quad (3)$$

with

$$\hat{\mathbf{v}} = \Delta t \sum_{q=0}^{J_e-1} \beta_q \mathbf{v}^{n-q} \cdot \nabla \mathbf{v}^{n-q} + \sum_{q=0}^{J_i-1} \alpha_q \mathbf{v}^{n-q} + \mathbf{F} \quad (4)$$

We use an explicit scheme of order J_e for the nonlinear terms and an implicit scheme of order J_i for the linear terms. In our case, we use a standard backward differentiation scheme of order three for the implicit terms and an extrapolation scheme of order three for the explicit terms. To decouple the velocity and pressure solves we take the divergence of (3) and use incompressibility, our explicit treatment of \mathbf{v} and the identity

$$\nabla^2 \mathbf{v}^{n+1} = \nabla (\nabla \cdot \mathbf{v}^{n+1}) - \nabla \times (\nabla \times \mathbf{v}^{n+1}) \quad (5)$$

to arrive at a pressure-Poisson equation for the pressure

$$-\nabla^2 p^{n+1} = -\frac{\nabla \cdot \hat{\mathbf{v}}}{\Delta t} + \frac{1}{\text{Re}} \nabla \cdot (\nabla \times \omega) \quad (6)$$

with ω being the vorticity computed explicitly

$$\omega = \sum_{q=0}^{J_e-1} \beta_q (\nabla \times \mathbf{v}^{n-q}) \quad (7)$$

With this approach, we have decoupled the pressure and velocity solve and each time step can thus be computed by first solving the system for the pressure according to (6) followed by using the resulting pressure field p^{n+1} in equation (3) and obtaining the velocity at step $n+1$. Crucial to this splitting approach is choosing suitable boundary conditions for the pressure. It is found that the temporal error of the splitting scheme is highest along the boundary and directly depends on properly imposed pressure boundary conditions. Using Neumann boundary conditions it is possible to limit the temporal splitting error for the pressure and divergence to $\mathcal{O}(\Delta t^{J_e})$ and thus limit the splitting error for the velocity to $\mathcal{O}(\Delta t^{J_e+1})$ as shown in Orszag et al. (1986); Karniadakis et al. (1991). In addition, it can be shown that the error in the divergence of the velocity at step $n+1$ decays exponentially from the boundary proportionally to $e^{-s/l}$, where s is a coordinate normal to the boundary and $l = \sqrt{\gamma_0 \nu \Delta t}$ is the so-called numerical boundary layer thickness. As $l \rightarrow 0$ for $\text{Re} \rightarrow \infty$ the error decays rapidly from the boundary for highly turbulent flows

(Karniadakis et al., 1991). Unique for this splitting scheme, compared to the more conventional $P_N - P_{N-2}$ splitting for the spectral element method presented by Maday and Patera (1989) is that the velocity and pressure solve can be done with the same function space, and thus use the same polynomial degree N . With our temporal discretization in place, we apply our spectral element discretization in space.

The spectral element method is used because of its high-order convergence and high accuracy with relatively few grid points in space compared to, for example, finite volume methods, while still accommodating unstructured meshes (Deville et al., 2002). The linear system is evaluated in a matrix-free fashion, using optimized tensor-operations, which yields a high operational intensity, performance, and low communication costs. These aspects make the spectral element method well-suited for DNS in complex geometries.

For the spectral element discretization we consider the weak form of the Navier-Stokes equations and decompose the domain into E non-overlapping hexahedral elements $\Omega = \cup \Omega^e$. On these elements we represent the solution with polynomial basis functions based on the Legendre polynomials of degree N interpolated on the Gauss-Lobatto-Legendre (GLL) points ζ_i , $i = 0, N$. In 3D the number of GLL points per element is $(N+1)^3$. A function on the reference element can be represented according to

$$u^e(\zeta, \eta, \gamma) = \sum_{i,j,k=0}^N u_{i,j,k}^e l_i(\zeta) l_j(\eta) l_k(\gamma) \quad (8)$$

where l_i are the polynomial basis functions and ζ, η, γ are the location within the reference element. As we use non-overlapping hexahedral elements, differentiation can be evaluated efficiently through tensor-products Orszag (1979). For instance, the weak form of the Laplace operator can be written as

$$(\nabla \mathbf{v}, \nabla \mathbf{u}) = \sum_e (\mathbf{v}^e)^T \mathbf{D}^T \mathbf{G}^e \mathbf{D} \mathbf{u}^e = \sum_e (\mathbf{v}^e)^T \mathbf{A}^e \mathbf{u}^e \quad (9)$$

where we introduce the differentiation matrices \mathbf{D} and the geometric factors \mathbf{G}^e . These two matrices are defined as

$$\mathbf{D}_{ii} = \frac{dl_i(\zeta)}{d\zeta} \Big|_{\zeta_i}, (\mathbf{G}_{st}^e)_{ijk} = \rho_i \rho_j \rho_k \mathcal{J}_{ijk} \sum_{l=1}^3 \frac{\partial r_s}{\partial x_l} \frac{\partial r_t}{\partial x_l}$$

where $\mathcal{J} = \partial \mathbf{x} / \partial \mathbf{r}$ is the Jacobian for the mapping to the reference element. We have that $\mathbf{G}_{st}^e = \mathbf{G}_{ts}^e$ and thus we have that \mathbf{G}^e is symmetric and contains six non-trivial values per Gauss-Lobatto-Legendre (GLL) point. In Neko, the global system \mathbf{A}^e is never formed as this would be prohibitively expensive, but \mathbf{A}^e is instead evaluated in a matrix-free fashion (Deville et al., 2002). For this, we introduce the local representation $u_L = \mathcal{Q} \mathbf{u}$ where \mathcal{Q} is known as the

scatter operator. The local representation u_L contains duplicate values along element boundaries, but enables for efficient tensor-operations. By ensuring that

$$u_{ijk}^e = u_{i'j'k'}^{e'} \quad \text{if} \quad \mathbf{x}_{ijk}^e = \mathbf{x}_{i'j'k'}^{e'} \quad (10)$$

and using the local representation we can therefore evaluate the Laplace operator as

$$\sum_e^E (v^e)^T A^e u^e = (Qv)^T A_L Qu \quad (11)$$

with A_L being the block diagonal matrix of A^e . The final discrete system that we solve for can thus be written as

$$QQ^T A_L u_L = A u_L \quad (12)$$

A similar approach is applied to the Helmholtz equation for the velocity arising from the implicit treatment of the viscous term. Using this discretization and time splitting we have all the components to set up our flow case. To solve the systems for the velocity and pressure we use heavily optimized numerical solvers.

4.2. Numerical solvers

The pressure solve is the main source of stiffness in incompressible flow and because of this, we use a restarted generalized minimal residual method (GMRES) (Saad and Schultz 1986) combined with a preconditioner based on an overlapping additive Schwarz method and a coarse grid solve (Lottes and Fischer 2005). In Neko, the coarse grid solve is done with 10 iterations of the preconditioned conjugate gradient (P-CG) method. This combination of GMRES + multigrid preconditioner with Schwarz smoothing together with a coarse grid solver converges in a few number of iterations at the expense of increased work per iteration. The system matrix for the pressure solve is in general ill-conditioned and non-symmetric.

For the velocity solver, we utilize P-CG with a block-Jacobi preconditioner. For both the velocity and pressure solve we utilize a projection scheme where we project the solution of the current time step on previous solution vectors to decrease the start residual significantly. This has been observed to reduce the number of iterations for the Krylov solver significantly (Fischer 1998). All of our solvers are designed with locality across the memory hierarchy in mind, to be able to as efficiently as possible use modern GPUs with a significant machine imbalance. Parts of this optimization process and the theoretical background is described in Karp et al. (2022a).

We illustrate how each time step is computed on a high level in Figure 2. Here we also illustrate the software structure of Neko, which we will cover more in-depth in the next section.

4.3. Software design

The primary consideration in Neko is how to efficiently utilize different computer backends, without re-implementing the whole framework for each backend and while maintaining the core of the solver in modern Fortran. We solve this problem by considering the weak form of the equations used in the spectral element method. The weak formulation allows us to formulate equations as abstract problems which enable us to keep the abstractions at the top level of the software stack and reduce the amount of platform-dependent kernels to a minimum. The multi-tier abstractions in Neko are realized using abstract Fortran types, with deferred implementations of required procedures. For example, to allow for different formulations of a simulation's governing equations, Neko provides an abstract type `ax_t`, defining the abstract problem's matrix-vector product. The abstract type is shown in Figure 2 under Abstract types and requires any derived, concrete type to provide an implementation of the deferred procedure `compute`. This procedure would return the action of multiplying the stiffness matrix of a given equation with a vector. For the Poisson equation, the `compute` subroutine corresponds to computing expression (9).

In a typical object-oriented fashion, whenever a routine needs a matrix-vector product, it is always expressed as a call to `compute` on the abstract base type and never on the actual concrete implementation. Abstract types are all defined at the top level in the solver stack during initialization and represent large, compute-intensive kernels, thus reducing overhead costs associated with the abstraction layer.

Furthermore, this abstraction also accommodates the possibility of providing tuned matrix-vector products (`ax_t`) for specific hardware, only providing a particular implementation of `compute` without having to modify the entire solver stack. The ease of supporting various hardware backends is the key feature behind the performance portability of Neko.

However, a portable matrix-vector multiplication backend is not enough to ensure performance portability. Therefore, abstract types are used to describe a flow solver's common parts. These abstract types describe all the parts necessary to compute a time step and provide deferred procedures for essential subroutines. These building blocks then make up the entire solver used to calculate one-time step. We illustrate this structure in Figure 2 where a solver such as `fluid_pnpn_t` implements the deferred subroutine step. The solver is then ready to solve a particular case. Each case (`case_t`) is defined based on a mesh (`mesh_t`), parameters (`param_t`), together with the abstract solver (`solver_t`), later defined as an actual extended derived type, such as `fluid_pnpn_t`, based on the simulation parameters. The abstract solvers contain a set of defined derived types necessary for a

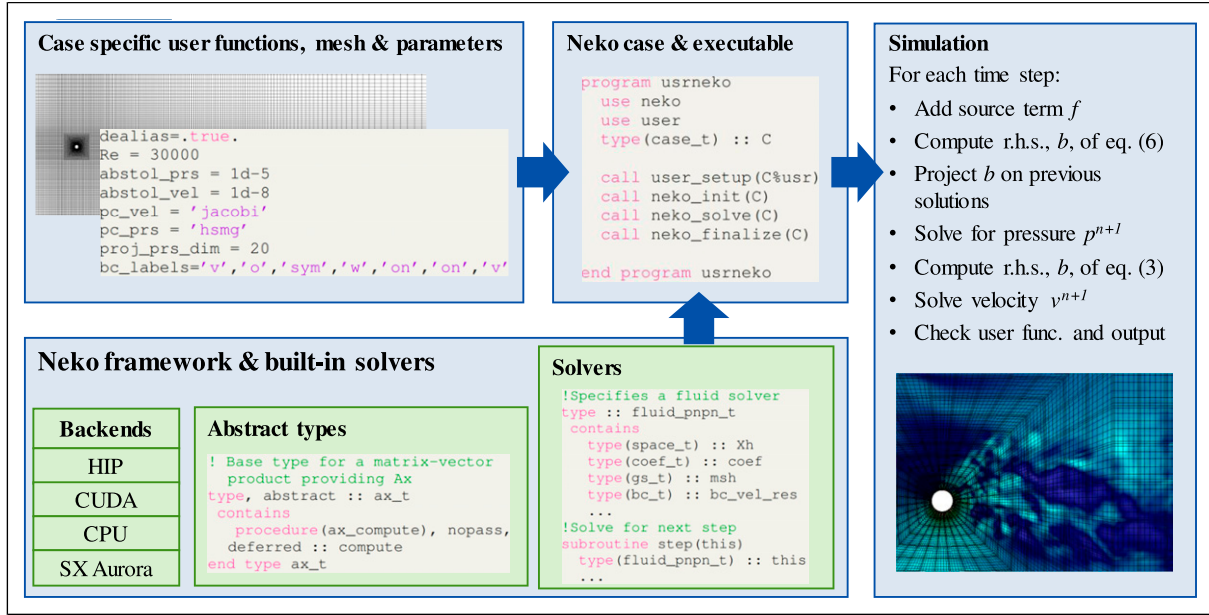


Figure 2. A chart illustrating the structure of the Neko framework. The Neko framework, with various abstract types and backends make up different solvers which are used to solve a case. The case is defined by user provided functions, as well as the input parameters and mesh. In our case, the simulation is performed using the $P_N - P_N$ method to perform a DNS of a Flettner rotor.

spectral element simulation, such as a function space (`space_t`), coefficients (`coef_t`), and various fields (`field_t`). Additionally, they contain further abstract types for defining matrix-vector products (`ax_t`) and gather-scatter QQ^T kernels (`gs_t`). Each of these abstract types is associated with an actual implementation in an extended derived type, allowing for hardware or programming model specific implementations, all interchangeable at runtime. This way, Neko can accommodate different backends with both native and offloading type execution models without unnecessary code duplication in the solver stack.

4.4. GPU implementation considerations

As Neko provides abstractions (abstract derived types) on several levels it is possible to utilize different libraries to leverage accelerators when implementing a hardware-specific derived type in the solver stack (e.g. `ax_t`). In our work, we do not rely on vendor-specific solutions (e.g. CUDA Fortran) or directives-based approaches; instead, Neko provides a device abstraction layer to manage device memory, data transfer, and kernel launches from Fortran via its C interoperability layer. Behind this interface, Neko calls the native accelerator implementation written in, for example, CUDA, HIP or OpenCL. By using this layer, the majority of the code-base can be written in a vendor-

neutral way, and in Fortran, with only limited sections (the extended derived types mentioned in the preceding section) implemented in, for example, CUDA or HIP, or some other accelerator programming model, together with a small Fortran module, extending the particular derived type and providing the call to launch the kernel. In our work we extend and optimize the CUDA and HIP backends in Neko.

As we extend several features of the device layer in Neko in this work, a few aspects of the code development are important to point out regarding the GPU backend. In particular, we observed a considerable difference in how HIP and CUDA kernels map to their respective platform. As several important kernels in Neko follow a similar tensor-product structure to the kernels considered by Świrzydowicz et al. (2019), and perform close to the memory-bound roofline on Nvidia GPUs, the same kernels ported to HIP needed to be changed to better exploit AMD GPUs. We found that a 1D thread structure and manual calculation of indices in a thread block enabled higher performance on the AMD Instinct MI250X, rather than launching a 2D/3D thread block per element as we did previously on Nvidia. We illustrate this difference between the two approaches in Listing 1 where we show how the indices i, j, k in element e are computed depending on if the kernel is launched with a 2D or 1D thread structure.

Listing 1: Pseudocode for how the indices i, j, k in element e are computed for a 1D and 2D kernel respectively. N is the polynomial order.

```
#2D structure
#More readable, but caused performance degradation on AMD
void 2D_kernel(...) {
    int LX = N + 1;
    int e = blockIdx.x;
    int j = threadIdx.y;
    int i = threadIdx.x;
    int ijk = i + j * LX;

    for (int k = 0; k < LX; k++){
        int ijk = ijk + k * LX * LX;
        ...
    }
}

#Preferred structure on AMD/HIP
void 1D_kernel(...) {
    int LX = N + 1;
    int e = blockIdx.x;
    int iii = threadIdx.x;
    int nchunks = (LX * LX * LX - 1) / blockDim.x + 1;

    for (int n = 0; n < nchunks; n++) {
        int ijk = iii + n * blockDim.x;
        int jk = ijk / LX;
        int i = ijk - jk * LX;
        int k = jk / LX;
        int j = jk - k * LX;
        ...
    }
}
```

It would also appear that there is a difference in the order the threads execute their operations on AMD, meaning that we at one point needed to synchronize the thread-block in HIP, while the same code performed deterministically on Nvidia GPUs. These differences between CUDA/Nvidia and HIP/AMD warrant further and more systematic investigation as it is not immediately evident why we observe such differences.

One contribution of the GPU backend in Neko is that we limit data movement to and from the device as far as possible. In practice, this means that only data needed for communication over the network or to do I/O is sent back to the host. As the bandwidth between host and device is significantly lower than that of the high-bandwidth memory (HBM) on the GPUs, limiting data movement to the host is essential to obtain high performance. However, this also means that the HBM memory of the GPUs must fit all the data necessary to carry out the simulation. While it might be beneficial to utilize both CPU and GPUs in the future, the GPU nodes evaluated in this work have significantly more powerful GPUs than the host CPU with regards to both flop/s and memory bandwidth (more than 10 \times). We found that the most effective way to leverage this computing power was to use the host only as a coordinator and offload the entire calculation to the GPUs.

In addition to limiting the data movement to the host, we also avoid data movement from the GPU global memory as far as possible. We do this by merging kernels and by utilizing shared memory and registers in compute heavy kernels as detailed in, for example, [Wahib and Maruyama \(2014\)](#). For modern GPUs, the spectral element method is in the memory-bound domain as discussed by [Kolev et al. \(2021\)](#) and optimizing the code for temporal and spatial locality is our main priority when designing kernels for the GPU backend in Neko, this was recently considered in

depth for the CG method used in Neko in [Karp et al. \(2022a\)](#).

Using the device abstraction layer, from the user perspective, it is possible to write user-specified functions without having to implement handwritten accelerator code, for example, CUDA or HIP. This enables tools from Nek5000 to be used on the device with relatively minor changes as we introduce device kernels with the same functionality as the original subroutines in Nek5000. Neko provides support for different backends for the gather-scatter operation and Ax as previously mentioned, but also for mathematical operators such as the gradient and curl. With the device layer, we can with only minor changes incorporate verified toolboxes and functions necessary for our flow case into Neko and run them directly on the GPU without sending data back to the host.

4.5. Modern Fortran

As is clear from the description of Neko, modern Fortran (e.g. Fortran 03, 08, 18) offers similar object-oriented features to other programming languages such as C++ ([Reid 2008](#)). Most importantly, modern Fortran supports dynamic memory allocation. Other features of Fortran are the native support for the quad data-type, FP128, that all arrays are by default not aliased, pure functions, and that modern Fortran both supports older Fortran code and can interface with C-code. As Fortran is tailored to high-performance numerical computations it caters to the functionalities necessary to perform high-fidelity flow simulations.

5. Experimental setup

We describe here our precise problem description to accurately capture the physics of a Flettner rotor in a turbulent boundary layer. In addition, we detail the computational setup for our scaling runs on the GPUs. We also describe our baseline CPU platform and provide some further information regarding our final production run that was used to collect the simulation results.

5.1. Flow configuration

In marine applications, the Flettner rotor is usually placed on the deck of the ship and hit by an incoming flow, that is, a turbulent atmospheric boundary layer with a certain thickness. In the current study, we simulate such a configuration as an open channel flow where the Flettner rotor is located at the origin of our domain. The reference system is oriented with the y -axis along with the cylinder (vertical) and the x and z being the streamwise and spanwise directions respectively. The flow is fully characterized by three non-dimensional parameters:

- the Reynolds number based on the center-line velocity u_{cl} and the height h of the open channel: $Re_{cl} = u_{cl}h/\nu$, where ν is the kinematic fluid viscosity
- the ratio between the height h and the cylinder diameter $\gamma = h/D$
- the ratio between the centre-line and the spinning cylinder velocity $\alpha = u_{sp}/u_{cl}$.

We set those parameters as $Re_{cl} = 30,000$, $\gamma = 10$, and $\alpha = 3$, aiming to reproduce a realistic configuration, within the limits of DNS where all the scales are simulated and no turbulence model is introduced.

The mesh is Cartesian with a block-structured configuration, taking into account the resolution requirements in the different regions and the flow physics. It counts 930,070 spectral elements, which turns into $n \approx 0.48$ billion unique grid points since the polynomial order is $N = 7$ and we have 8^3 GLL points per element. We show the spectral element mesh with domain dimensions, without GLL points, in Figure 3. Close to the cylinder, high resolution is necessary to properly resolve the smallest scales in the boundary layer and we show a snapshot of the large number of GLL points in Figure 4. For the simulation, a time step of $\Delta t^* = 2 \cdot 10^{-5}$ is used, corresponding to a Courant—Friedrichs—Lewy (CFL) number around 0.41. With regards to our iterative solvers, we use a residual tolerance of 10^{-5} for the pressure and 10^{-8} for the velocity. The choice of tolerances is case-dependent. We consider different orders of magnitude as the pressure solver is the most time-consuming. However, the relaxed tolerance for the pressure is still small enough compared to other error sources. The grid is designed to avoid any blockage effects on

the rotor, that is, the domain sizes are adequately, and conservatively, large: $(-50D, 120D)$ in x , $(-25D, 25D)$ in z and $(0, 10D)$ in y , where x, z, y are the streamwise, spanwise and vertical directions respectively.

At the inflow ($x = -50D$) the Dirichlet boundary condition (BC) prescribes a turbulent channel velocity profile with the power-law $u_x/u_{cl} = (y/h)^{1/7}$. Unlike the parabolic laminar profile, it is flatter in the central part of the channel and drops rapidly at the walls. On the rotor, a Dirichlet-like BC is used as well, prescribing the wall impermeability and setting a rotational velocity $u_\theta = u_{sp} = \alpha u_{cl}$ with $\alpha = 3$ and $u_\rho = 0$. Here, the BC is expressed in cylindrical coordinates, being u_θ and u_ρ the velocities in the azimuthal and radial direction, respectively. To avoid incompatibility conditions at the base of the cylinder the spinning velocity is smoothed as a function of y

$$\begin{aligned} s(y=0) &= 0, \\ s(0 < y < \delta) &= u_{sp} / (1 + e^{1/(q-1)+1/q}) \\ s(y \geq \delta) &= u_{sp} \end{aligned} \quad (13)$$

where $q = y/\delta$ and $\delta = 0.02h$. The outflow at $x = 120D$ consists of natural boundary condition $(-p\mathbf{I} + \nu\nabla\mathbf{u}) \cdot \mathbf{n} = 0$. We consider an infinitely long cylinder in the y direction, as a consequence symmetric Robin (or mixed) boundary conditions are prescribed at the top surface ($\mathbf{u} \cdot \mathbf{n} = 0$ with $(\nabla\mathbf{u} \cdot \mathbf{t}) \cdot \mathbf{n} = 0$). For the spanwise boundaries mixed conditions allowing transportation are used, similar to the open boundary, but prescribing no velocity changes in non-normal directions.

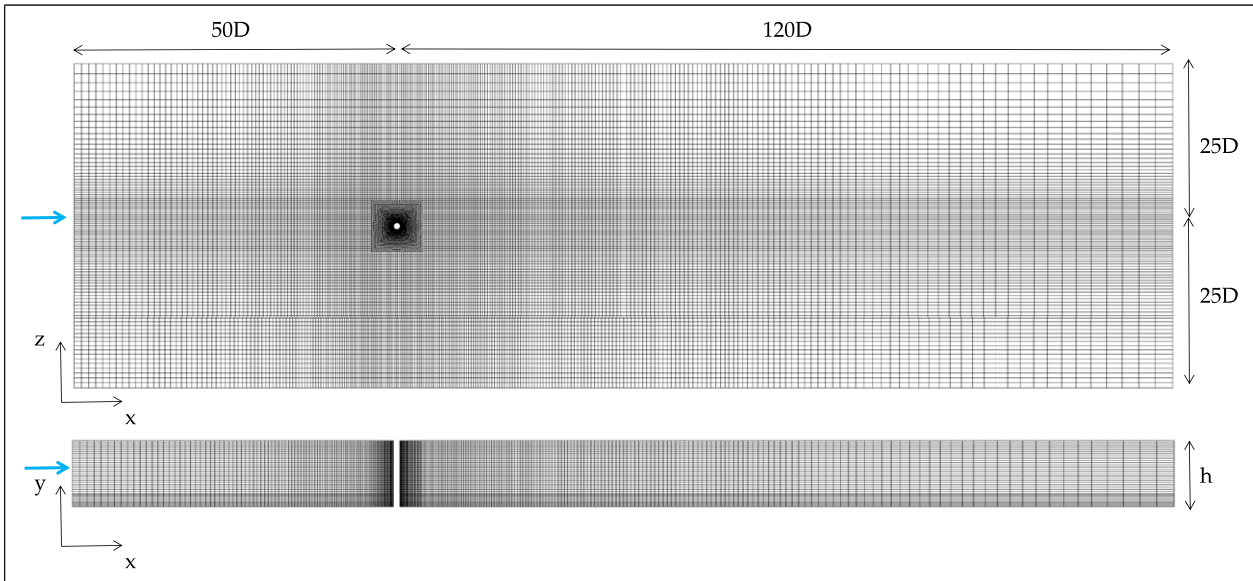


Figure 3. The spectral elements grid is shown. The domain sizes are reported in terms of cylinder diameter D and open channel height h ($\gamma = h/D = 10$). The blue arrows indicate the flow direction.

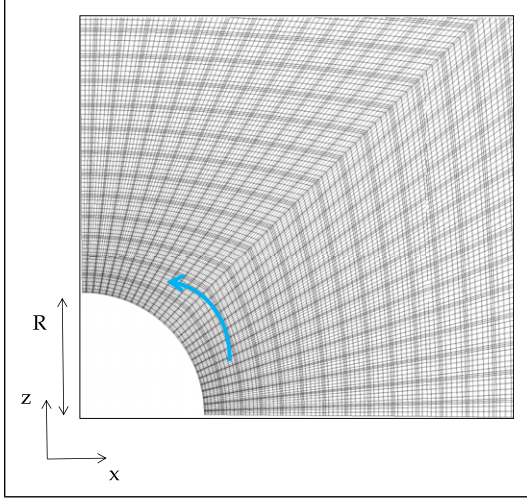


Figure 4. The figure shows the Gauss-Lobatto-Legendre points around the cylinder. The cylinder rotates along the vertical y axis, and the blue arrow indicates the direction of rotation.

To keep the front streamwise extent of the computational domain as short as possible, reduce the computational cost, and sustain the incoming turbulence, the laminar-turbulent transition is initiated via the boundary layer tripping introduced by Schlatter and Örlü (2012). It consists of a stochastic forcing term which is added in a small elliptical region close to the wall. The region is centred along a user-defined line at $x_0 = -4.5$ and it extends from $z = -1$ to $z = 1$. The implementation follows Schlatter and Örlü (2012), using a weak random volume force that acts in the y direction. The term which enters the Navier–Stokes equation is

$$F_2 = g(z, t) \cdot \exp\left(\frac{(x - x_0)^2}{l_x^2} - \frac{y^2}{l_y^2}\right) \quad (14)$$

where l_x and l_y are the spatial Gaussian attenuation of the forcing region. The function $g(z, t)$ consists of two terms, corresponding to steady and unsteady perturbations, with amplitudes T_s and T_u respectively. The form is the following

$$g(z, t) = T_s g(z) + T_u [(1 - b(t))h^i(z) + b(t)h^{i+1}(z)] \quad (15)$$

with $i = \text{int}(t/t_s)$, $p = t/t_s - i$, and $b(t) = 3p^2 - 2p^3$. Third-order Lagrangian interpolants with time scale t_s are used for the temporal fluctuations and the forcing has a continuous first-order derivative in time. The functions $g(z)$ and $h(z)$ are Fourier series of unit amplitude with N_m random coefficients. The noise is generated with a uniform distribution over all frequencies lower than the cutoff frequency, corresponding to $2\pi/t_s$. The values of the parameters have to be properly chosen; here we use $T_s = 0$, $T_u = 0.3$, $t_s = 0.14$ and $N_m = 40$.

This tripping procedure has been used extensively in Nek5000 (see Tanarro et al. (2019); Atzori et al. (2021)) and keeps the same original implementation of the pseudo-spectral code SIMSON (Chevalier et al., 2007). We implement the interpolation on the GPU using the Neko device layer to avoid unnecessary data movement to and from the device. This can be done in Neko without making too many changes to the original implementation and we can rely on a tripping procedure that is well-used and validated in our simulation.

5.2. Computational setup

We perform our experiments on two different GPU clusters, Alvis at C3SE at Chalmers University of Technology in Gothenburg (Sweden) and on an internal HPE Cray EX system. We present the technical details of both computer setups in Table 1. The largest difference between the two setups is the Peak double precision (FP64) performance of their respective GPUs, but only a factor of two difference for the available memory bandwidth. As each MI250X consists of two graphics compute dies (GCDs), each GCD corresponding to one logical GPU, we will throughout the discussion compare logical GPUs, comparing one GCD to one A100. We note that the HPE Cray EX system is an internal engineering test system at HPE. Although our results are in line with expectations, it is possible that the performance or scalability are affected by temporary hardware settings and configuration. In addition to the GPU runs we also performed CPU measurements on Dardel, also a Cray HPE EX system at PDC at KTH Royal Institute of Technology.

5.3. Experimental methodology

For our scaling runs we scale between 8 and 32 Alvis nodes each with four A100 GPUs. We measure the performance over 500-time steps and use the last 100 to collect the average wall time per time step as well as the standard deviation. The same procedure is used on the HPE cluster where we scale from 4 – 16 nodes, each equipped with four MI250X GPUs. Since they are dual-chip GPUs, we use eight MPI ranks per node. On Dardel we scale from 16 – 64 nodes, each with two CPUs, using one MPI rank per core. As the flow case is large, we could not measure the performance at a fewer number of nodes.

We also compare the power consumption of the two different GPU clusters. For these results we rerun the scaling runs, but also measure the power on the GPUs during the runs. On the Alvis cluster, a report with the average power usage for each GPU is automatically generated after each run through the NVIDIA Datacenter GPU Management interface. In addition to this automatic system-provided report, we poll nvidia-smi during the

Table 1. Hardware and software details for our experimental platforms. Hardware listed is per node while bandwidth, performance, and TDP is per physical GPU card. FP64 performance from tensor/matrix cores omitted.

| Setup | Alvis | HPE Cray EX | Dardel |
|---------------|-------------------------|-------------------------------|-------------------------------|
| GPU | 4 Nvidia A100 SXM4 | 4 AMD Instinct MI250X | - |
| CPU | 2 Intel Xeon Gold 6338 | AMD EPYC 7A53 | AMD EPYC 7742 |
| DRAM | 512 GB DDR4 | 512 GB DDR4 | 256 GB DDR4 |
| Cores | 2 × 32 (2 threads/core) | 64 (2 threads/core) | 2 × 64 (2 threads/core) |
| Sockets | 2 | 1 | 2 |
| MPI | OpenMPI 4.0.5 | Cray MPICH 8.1.14 | Cray MPICH 8.1.11 |
| Interconnect | Mellanox ConnectX-6 | HPE Slingshot Interconnect 11 | HPE Slingshot Interconnect 10 |
| — | (2x200 Gb/s) | 200 GbE NICs (4x200 Gb/s) | 100 GbE NICs (1x100 Gb/s) |
| Compiler | GCC 10.2 | CCE 13.0 | CCE 11.0.4 |
| GPU Driver | 510.39.01 | 5.11.32 | - |
| CUDA/ROCM | CUDA 11.1.1 | ROCM 4.5.2 | - |
| Bandwidth/GPU | 1550 GB/s | 3277 GB/s | - |
| GPU Peak FP64 | 9.746 TFLOPS | 47.87 TFLOPS | - |
| GPU TDP | 400W | 560 W | - |

run to continually measure the power of the GPUs. With these two measurements, we check that the average power usage during the run is consistent between the two measurements. As initialization does not properly represent the power draw during a simulation we use the measurements from nvidia-smi to compute the average power usage for the GPUs during the run after initialization.

On the internal HPE Cray EX system, the power draw of the node is measured using a similar method to that on the Cray XC systems described by Hart et al. (2014). On the HPE Cray EX system, separate user-readable counters are provided on each node for the CPU, memory, and each of the four accelerator sockets. By measuring the energy counters and the timestamp at the start and end of the job script, a mean power draw for each component can be computed and averaged across the nodes used. We also poll the GPUs with rocm-smi during the run to get an overview of the power draw during the run. With these measurements, we verify that the average power draw between the two measurements during the course of the run is consistent for the GPUs. We then use the measurements from rocm-smi to compute the average power usage during the computation, excluding initialization. We note that the differences in methodology mean that some care should be taken when comparing time-averaged power draws between the two GPU systems. For the CPU measurements on Dardel, we used CrayPAT to make power measurements during the run after init.

For our pilot production simulation, we use Alvis at C3SE to generate over a thousand samples of the field and produce over 10 TB of flow field data over the course of only 3 days using between 64 and 128 Nvidia A100 GPUs.

6. Simulation results

After running the simulation for around $15t^*$ flow time units, where $t^* = tu_{cl}/h$ is the non-dimensional time, and using around 10 000 GPU hours, we show the flow field in Figure 5. We can observe a significant interaction between the rotor and the turbulent boundary layer, with the wake deflected by the spinning motion. The Flettner rotor generates a spanwise force which could be used to significantly reduce the amount of required thrust of a ship.

A first estimation has been performed by time-averaging the surface integral of the pressure and viscous stress tensor. The resulting aerodynamic force, \mathbf{F}_a , is computed as

$$\mathbf{F}_a = \int_S -p\mathbf{n} + \boldsymbol{\tau}\mathbf{n}dA \quad (16)$$

where S is the cylinder surface, p the pressure, \mathbf{n} the vector normal to the cylinder surface and $\boldsymbol{\tau}$ the viscous stress $\tau_{ij} = \mu\partial u_i/\partial u_j$. According to the standard convention, the force is normalized with $0.5\rho u_{cl}hD$. As a preliminary result, we present the most relevant integral quantities, that is, the drag and lift coefficients. They are the streamwise and spanwise normalized force components, respectively. The spanwise lift coefficient $C_l = 7.464$ shows an agreement with the experimental data that was measured to be between 7 and 8 for the same $\alpha = 3$ (Bordogna et al., 2019). The Reynolds number in our case is smaller, but its influence on C_l seems to be limited. Quite the opposite, the drag coefficient $C_d = 1.092$ is dependent on the Reynolds number. Both standard deviations are small (order of magnitude $\approx 10^{-3}$), showing no large temporal fluctuations for these integral estimations. This first pilot investigation shows interesting flow features, as the wake of the cylinder interacts with the boundary layer, encouraging us to

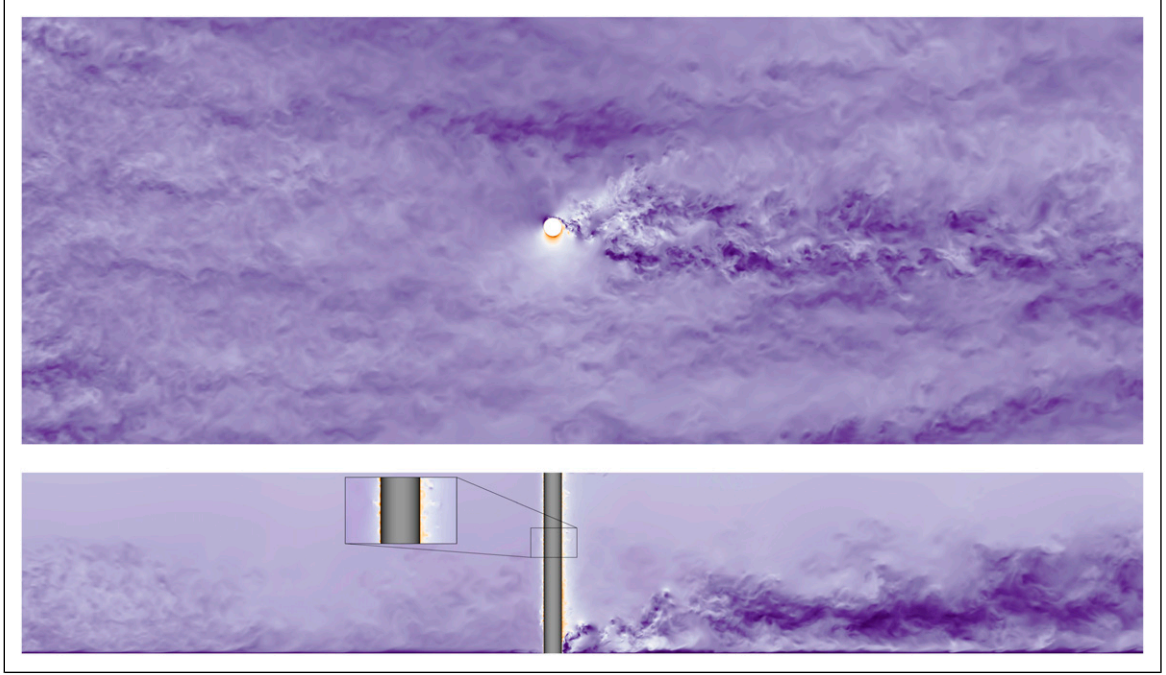


Figure 5. Two zoomed-in visualizations of the flow around our simulated Flettner rotor. We show the velocity magnitude where a lower velocity is darker violet and a higher velocity becomes white then orange. At the top we show a cut at $z = 0$ and at the bottom we show a plane at $y = 0.1$. The coordinate axes reflect the mesh as shown in [Figure 3](#), but are zoomed in to better visualize the turbulence close to the cylinder.

continue the study by observing the rotation speed influence or testing different incoming velocity orientations. We aim to highlight new flow characteristics by means of statistical flow analysis and the introduction of modal decomposition, for example, the proper orthogonal decomposition (POD) to extract the most energetic coherent structures of the flow. More in-depth knowledge of the Flettner rotor dynamics turns out to be extremely relevant in many engineering applications. While our results in this work are preliminary, we have shown how DNS is a valuable tool in understanding how we can leverage Flettner rotors for more sustainable transportation and how DNS can be used to understand flow features not visible in physical experiments.

7. Performance results

For the performance results, we compare the two GPU platforms to a CPU baseline by strong scaling with our Flettner rotor case. We also compare the power efficiency of the Nvidia A100 and the AMD Instinct 250X GPUs. We also provide the measured energy consumption for the nodes used in the CPU baseline.

The shaded area shows the standard deviation between GPUs used in the run for their average power usage.

7.1. Scaling performance

In [Figures 6](#) and [7](#) we present the achieved strong scaling performance and speedup for both the AMD Instinct MI250X, Nvidia A100, and the AMD EPYC 7742 CPU baseline. First, comparing the two GPUs, it is clear from the time per time step that two logical GPUs on the MI250X correspond to two A100s with regard to performance. We observe that the performance of one A100 and one GCD of the MI250X matches very well in all measurements. The average time per time step between the two architectures differs by less than 5%, comparing two A100 to one MI250X. This is in contrast to a recent report by [Kolev et al. \(2022\)](#) where they observed that one GCD of the MI250X performs closer to 79% of one A100 for a single rod simulation with NekRS, a similar code based on Nek5000 ([Fischer et al., 2021](#)). They also share our impression during code development that kernels that work well with CUDA/Nvidia do not necessarily map well to HIP/AMD, but need to be modified to properly exploit AMD GPUs. Comparing the two architectures in [Table 1](#), and considering that we expect to be memory-bound as has been discussed in [Karp et al. \(2021\)](#); [Kolev et al. \(2021\)](#), we see that the fraction between the bandwidth of two A100 versus one MI250X GPU is 1.05, indicating that they should perform similarly. As many solvers in the wider HPC community are memory-bound, we anticipate that several other codes will obtain similar

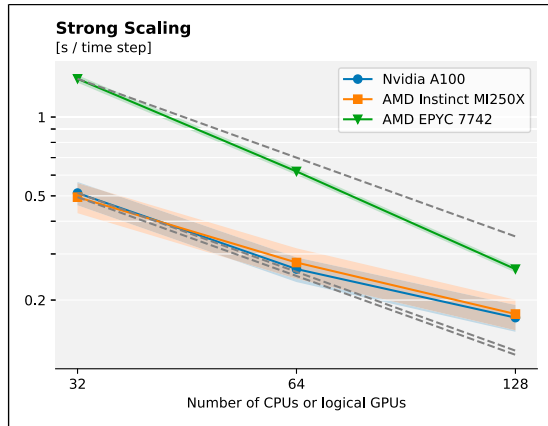


Figure 6. Performance in time per time step as well as shaded areas for the standard deviation. The orange and blue lines represent the GPU systems, while the green line with triangle markers represents the baseline CPU system. We illustrate linear scaling with dotted lines for each platform. The x-axis corresponds to the number of CPUs used for the CPU-only run or the number of GCDs/A100 GPUs used in the GPU runs.

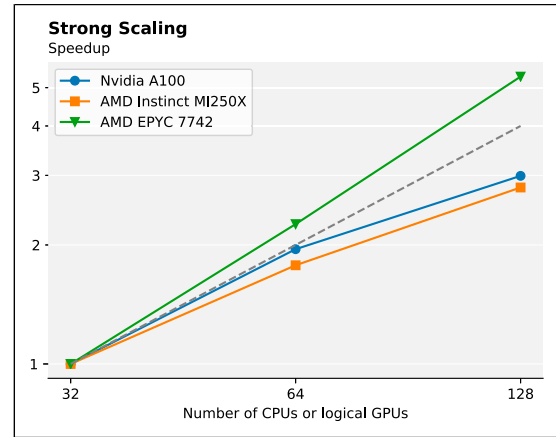


Figure 7. Speedup for the different systems. The orange and blue lines represent the GPU systems, while the green line with triangle markers represents the baseline CPU system. We illustrate linear scaling with a dotted line. The x-axis corresponds to the number of CPUs used for the CPU-only run or the number of GCDs/A100 GPUs used in the GPU runs.

performance numbers to ours on upcoming GPU supercomputers powered by the AMD Instinct MI250X relative to systems using the Nvidia A100. The relation between the floating-point performance is significantly larger than the observed difference, considering we do not use the tensor cores. The difference in peak FP64 performance suggests that compute-bound applications in FP64 might see larger performance improvements with the MI250X.

We observe a slightly higher standard deviation for the GPUs, compared to the CPU, and this can primarily be attributed to the projection scheme that is used in Neko. This scheme resets every 20-time steps on the GPU, effectively leading to a time variation with a period of 20 for the wall time per time step. On the CPU the projection space is not reset but a more complex scheme where the space is re-orthogonalized is used instead, and we see a lower standard deviation. Other than this difference, the numerical methods used are identical on CPU and GPU.

Comparing the two GPUs to the CPU baseline we see that for a large problem size such as this, the GPUs offer significantly higher performance than modern CPUs at lower node counts. The CPU requires more nodes to achieve the same performance as only a few GPU nodes. In particular, 128 CPUs, totalling 8192 cores performs similarly to only 64 Nvidia A100 GPUs or 32 AMD Instinct MI250X for this specific problem. While the strong scaling of the CPU is higher than the GPUs and running on a large number of cores can offer even higher performance, the possibility of using fewer nodes to get a similar performance is an alluring aspect of the GPUs. For even larger problems with even more elements, these results indicate that GPUs might

be the only way to obtain satisfactory performance for smaller computing clusters with a limited number of nodes.

We compute the overall parallel efficiency for the AMD Instinct MI250X and Nvidia A100 to 70% and 75% for 128 GCDs and 128 GPUs respectively, with around 3.75 million grid points (7000 elements) per A100 or GCD. This parallel efficiency is comparable and within one standard deviation. Also, we observe a slightly better average performance for the MI250X with fewer nodes than the A100, adding to the slightly lower parallel efficiency, which is also visible in the speedup plot. We anticipate that we will have a satisfactory parallel efficiency until 3500 elements per logical GPU. For the CPU on the other hand we see a superlinear parallel efficiency as the problem size per CPU is very large, between 200 000 grid points (450 elements) per core when using 32 CPUs, leading to an almost 5× speedup when using 128 CPUs. There is a significant cache effect as the number of points per core is decreased, leading to its superlinear scaling. This superlinear behavior for the strong scaling of spectral element codes has been observed previously and is also described by Fischer et al. (2021) and Offermans et al. (2016). Considering the large L3 cache on the AMD EPYC, this effect seems to be even more pronounced than in previous experiments. Overall, our parallel efficiency for all platforms is very high and shows how high-order, matrix-free methods can be used for large-scale DNS, both with conventional CPUs and when GPU acceleration is considered.

It is clear from these scaling results that our problem size is large for the considered number of CPU cores, and that the case requires at least 16k cores to run efficiently on the CPUs, but could use even more, the scaling limit on CPUs has previously been measured to around 50–100 elements

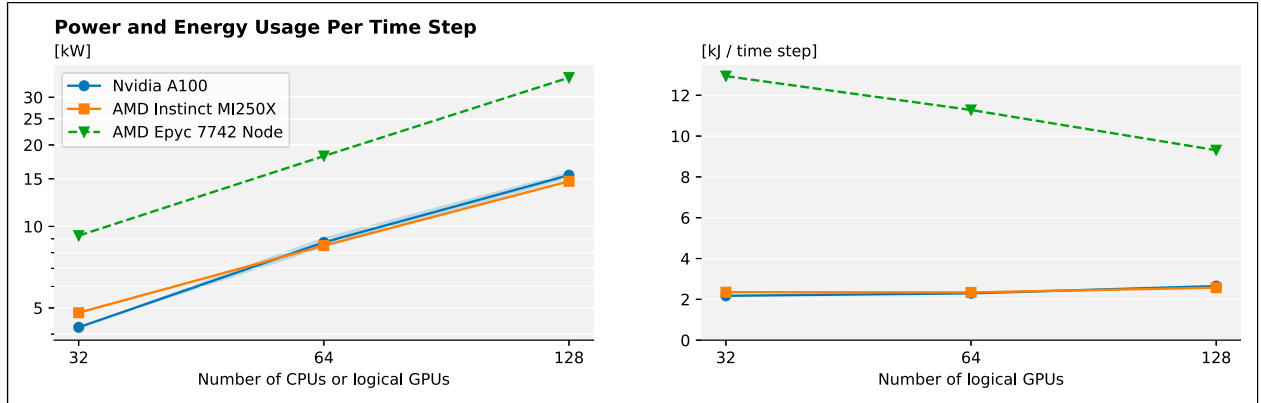


Figure 8. Power usage in watts and the energy consumption in joules during a time step for the GPUs, during a simulation. We also provide the measured power for the entire node for the CPU baseline, but stress that the CPU power measurements are for the whole node while the GPU measurements are only for the GPUs, and the measurements are therefore not equivalent.

per core (Offermans et al., 2016). Potentially, the CPUs could overtake the GPUs for a large number of cores with regard to performance. However, it is clear that GPU acceleration is necessary to run a case such as this on a fewer number of nodes, yielding a higher performance when the problem size per node is large. For more realistic flow cases, such as simulating the entire rotor ship illustrated in Figure 1, we would have a Re of several million, and thus at least a factor 10^5 increase in the number of gridpoints, which for the foreseeable future is out of reach. With upcoming exascale machines, we anticipate that GPU acceleration enables simulations at $Re = 300,000$, requiring thousands of GPUs. To properly simulate the entire ship configuration, accurate wall models will be essential to decrease the computational cost as was recently discussed by Bae and Koumoutsakos (2022).

7.2. Energy efficiency

We compare the MI250X and A100 GPUs' power consumption in Figure 8. Here, we see that the MI250X and the A100 perform on par with regard to power efficiency for Neko. The average energy consumed per time step is consistently within 10% between the two GPUs, with a slight edge for the A100 for a smaller number of nodes and the MI250X for a larger number of nodes. The standard deviation for the average power between GPUs is small (less than 5%) and the average power measured was highly consistent across GPUs on different nodes during the runs. As the AMD MI250X is expected to power several of the largest upcoming supercomputers in the world, the impact of its power efficiency should not be understated, both concerning operating costs and sustainability impact. As the A100 was previously unmatched with regards to power efficiency for this type of application (Karp et al., 2021), our results indicate that

the AMD MI250X is among the most energy-efficient options for large-scale computational fluid dynamics (CFD) currently available. Further supporting the power efficiency of the A100 and MI250X is that nine of the 10 most power-efficient supercomputers in the world in November 2021 utilize Nvidia A100 GPUs (Green500, 2021). We should note that we omit the energy usage of the network, CPUs, and other peripherals in our power measurements and isolate our focus on only the accelerators. The power measurement methodology also differs and this adds uncertainty to our measurements. However, as a contrast we also include the power measurements for the Dardel CPU nodes, where we see that one node with four MI250X will draw more than three times the power of one Dardel node. We note that the overarching computer system will have a significant impact on the energy efficiency of the computations, but for accelerator-heavy systems, the GPUs will draw a significant amount of the overall power.

In Figure 8 we also show the energy consumed by the GPUs during one time step. It is computed by taking the power usage and multiplying it by the time per time step. We note that the energy usage per time step that we observe is almost constant as we increase the number of GPUs, unlike the parallel efficiency. For the two GPUs, the energy consumed stays within 20% of the energy used with 32 GPUs/GCDs. The CPUs behave differently in that they draw the same amount of power, regardless of the problem size per core and the energy consumption thus having the same super-linear scaling as the run time. The CPU measurements are per node, and therefore not equivalent to the GPU measurements. One thing we can observe though is that while GPUs do not offer the same strong scaling characteristics as CPUs, the absolute run time is competitive or better and the energy needed for the GPUs during the simulation is roughly constant as we scale up. It

might be more important to consider energy usage per simulation, rather than run-time per simulation when comparing different architectures and algorithms. This more correctly corresponds with the operating costs of the computation and sustainability impact. While we have used the power counters available to us through `rocm/nvidia-smi` for the GPUs and `CrayPAT` for the CPUs, we note that obtaining accurate power usage is a challenge. Currently, direct power meter measurements are not widespread and used in production systems and it is an open question how we should compare different components of heterogeneous computer systems at scale with regards to power.

Another aspect of our power measurements is that our simulations are made in double precision. Using lower precision arithmetic to reduce power consumption and increase performance is gaining traction as more and more computing units support lower precision arithmetic. Incorporating other precision formats into Neko is an active area of research, both to increase performance, but also to decrease the energy consumption of large-scale DNS even further. Going forward, Neko will be used to assess the impact of numerical precision on direct numerical simulations. In addition, including adaptive mesh refinement and data-compression techniques to reduce the load on the I/O subsystem will let us use Neko for a wider range of cases and more complex geometries.

8. Conclusion

We have extended and optimized Neko, a Fortran-based solver with support for modern accelerators, ready for AMD and Nvidia GPUs. We leveraged modern Fortran to use previous tools developed for Fortran 77 and performed the world's first direct numerical simulation of a Flettner rotor in a turbulent boundary layer. We have presented some initial integral quantities and shown that DNS can further our understanding of turbulent coherent structures and how Flettner rotors can be used to increase sustainability in shipping. We have also presented a performance and energy comparison between the new AMD Instinct MI250X and Nvidia A100. Overall, we see that modern GPUs enable us to perform relevant, large-scale DNS on relatively few nodes within a time frame and energy budget that was previously not possible.

Acknowledgements

We are grateful to HPE for providing access to their system for this research. We thank C3SE for the opportunity to scale and perform the majority of our simulation using their Alvis system at Chalmers University of Technology in Gothenburg.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Vetenskapsrådet (2019-04723); the Swedish National Infrastructure for Computing (SNIC) at C3SE and PDC Centre for High Performance Computing, partially funded by Vetenskapsrådet (2018-05973); and the Swedish e-Science Research Centre (Exascale Simulation Software Initiative (SESSI)).

ORCID iDs

Martin Karp  <https://orcid.org/0000-0003-3374-8093>
 Daniele Massaro  <https://orcid.org/0000-0002-6712-8944>
 Niclas Jansson  <https://orcid.org/0000-0002-5020-1631>
 Philipp Schlatter  <https://orcid.org/0000-0001-9627-5903>

References

- Anderson R, Andrej J, Barker A, et al. (2021) MFEM: A modular finite element methods library. *Computers & Mathematics with Applications* 81: 42–74.
- Arndt D, Bangerth W, Blais B, et al. (2020) The deal.II library, version 9.2. *Journal of Numerical Mathematics* 28(3): 131–146.
- Atzori M, Vinuesa R, Stroh A, et al. (2021) Uniform blowing and suction applied to nonuniform adverse-pressure-gradient wing boundary layers. *Physical Review Fluids* 6: 113904.
- Bae HJ and Koumoutsakos P (2022) Scientific multi-agent reinforcement learning for wall-models of turbulent flows. *Nature Communications* 13(1): 1443–1449.
- Bordogna G, Muggiasca S, Giappino S, et al. (2019) Experiments on a Flettner rotor at critical and supercritical Reynolds numbers. *Journal of Wind Engineering and Industrial Aerodynamics* 188: 19–29.
- Bordogna G, Muggiasca S, Giappino S, et al. (2020) The effects of the aerodynamic interaction on the performance of two flettner rotors. *Journal of Wind Engineering and Industrial Aerodynamics* 196: 104024.
- Cantwell C, Moxey D, Comerford A, et al. (2015) Nektar++: An open-source spectral/hp element framework. *Computer Physics Communications* 192: 205–219.
- Chevalier M, Lundbladh A and Henningson D (2007) SIMSON—a pseudo-spectral solver for incompressible boundary layer flow. Technical report.
- De Marco A, Mancini S, Pensa C, et al. (2016) Flettner rotor concept for marine applications: A systematic study. *International Journal of Rotating Machinery* 2016: 1–12.
- Deville MO, Fischer PF and Mund E (2002) *High-order methods for incompressible fluid flow*. Cambridge University Press.

- Dykes T, Foyer C, Richardson H, et al. (2021) Mamba: Portable array-based abstractions for heterogeneous high-performance systems. In: 2021 International Workshop on Performance, Portability and Productivity in HPC, pp. 10–21. DOI: [10.1109/P3HPC54578.2021.00005](https://doi.org/10.1109/P3HPC54578.2021.00005)
- Carter Edwards H, Trott CR and Sunderland D (2014) Kokkos: Enabling manycore performance portability through polymorphic memory access patterns. *Journal of Parallel and Distributed Computing* 74(12): 3202–3216.
- Fischer P, Kerkemeier S, Min M, et al. (2021) *NekRS, a GPU-Accelerated Spectral Element Navier-Stokes Solver*. arXiv preprint arXiv:2104.05829.
- Fischer PF (1998) Projection techniques for iterative solution of with successive right-hand sides. *Computer Methods in Applied Mechanics and Engineering* 163(1–4): 193–204.
- Fischer PF, Lottes JW and Kerkemeier SG (2008) *nek5000 Web Page*. <http://nek5000.mcs.anl.gov>
- Green500 (2021) *Green500 Web Page*. <https://www.top500.org/lists/green500/>
- Hart A, Richardson H, Doleschal J, et al. (2014) User-level power monitoring and application performance on cray xc30 supercomputers. *Proceedings of the Cray User Group (CUG)* 1.
- Hornung RD and Keasler JA (2014) The RAJA portability layer: overview and status.
- Jansson N (2021) Spectral Element Simulations on the NEC SX-Aurora TSUBASA. In: The International Conference on High Performance Computing in Asia-Pacific Region. New York, NY, USA: HPC Asia 2021ACM, pp. 32–39.
- Jansson N, Karp M, Podobas A, et al. (2021) *Neko: A Modern, Portable, and Scalable Framework for High-Fidelity Computational Fluid Dynamics*. arXiv preprint arXiv:2107.01243.
- Kamiadakis GE, Israeli M and Orszag SA (1991) High-order splitting methods for the incompressible Navier-Stokes equations. *Journal of Computational Physics* 97(2): 414–443.
- Karp M, Jansson N, Podobas A, et al. (2022a) Reducing communication in the conjugate gradient method: A case study on high-order finite elements. In: Platform for Advanced Scientific Computing Conference. PASC '22. DOI: [10.1145/3539781.3539785](https://doi.org/10.1145/3539781.3539785)
- Karp M, Podobas A, Jansson N, et al. (2021) High-performance spectral element methods on field-programmable gate arrays : Implementation, evaluation, and future projection. In: 2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 1077–1086. DOI: [10.1109/IPDPS49936.2021.00116](https://doi.org/10.1109/IPDPS49936.2021.00116)
- Karp M, Podobas A, Kenter T, et al. (2022b) A high-fidelity flow solver for unstructured meshes on field-programmable gate arrays: Design, evaluation, and future challenges. In: International Conference on High Performance Computing in Asia-Pacific Region. New York, NY, USA: HPCAsia2022Association for Computing Machinery, pp. 125–136. DOI: [10.1145/3492805.3492808](https://doi.org/10.1145/3492805.3492808)
- Keryell R, Reyes R and Howes L (2015) Khronos SYCL for OpenCL: a tutorial. In: Proceedings of the 3rd International Workshop on OpenCL, pp. 1–1.
- Kolev T, Fischer P, Abdelfattah A, et al. (2022) *ECP milestone report high-order algorithmic developments and optimizations for more robust exascale applications wbs 2.2, milestone ceed-ms38*.
- Kolev T, Fischer P, Min M, et al. (2021) Efficient exascale discretizations: High-order finite element methods. *The International Journal of High Performance Computing Applications* 35(6): 527–552.
- Lottes JW and Fischer PF (2005) Hybrid multigrid/Schwarz algorithms for the spectral element method. *Journal of Scientific Computing* 24(1): 45–78.
- Maday Y and Patera AT (1989) *Spectral Element Methods for the Incompressible Navier-Stokes Equations. State-of-the-Art Surveys on Computational Mechanics (A90-47176 21-64)*. New York, pp. 71–143.
- Magnus G (1853) Über die Abweichung der Geschosse, und: Über eine abfallende Erscheinung bei rotierenden Körpern. *Annalen der Physik und Chemie* 164: 1–29.
- Massaro D, Jiahui L, Martin K, et al. (2022) The flettner rotor: leveraging the magnus effect for sustainable transport. In: *Gallery of fluid motion, APS Division of Fluid Dynamics*. GFM.V0113. DOI: [10.1103/APS.DFD.2022](https://doi.org/10.1103/APS.DFD.2022)
- Medina DS, St-Cyr A and Warburton T (2014) *OCCA: A Unified Approach to Multi-Threading Languages*. arXiv preprint arXiv:1403.0968.
- Offermans N, Marin O, Schanen M, et al. (2016) On the strong scaling of the spectral element solver nek5000 on petascale systems. In: Proceedings of the Exascale Applications and Software Conference, pp. 1–10.
- Orszag SA (1979) Spectral methods for problems in complex geometries. In: *Numerical methods for partial differential equations*. Elsevier, pp. 273–305.
- Orszag SA, Israeli M and Deville MO (1986) Boundary conditions for incompressible flows. *Journal of Scientific Computing* 1(1): 75–111.
- Otero E, Gong J, Min M, et al. (2019) OpenACC acceleration for the PN–PN-2 algorithm in Nek5000. *Journal of Parallel and Distributed Computing* 132: 69–78.
- Pachauri RK, Reisinger A, et al. (2007) *IPCC Fourth Assessment Report*. Geneva: IPCC.
- Patera AT (1984) A spectral element method for fluid dynamics: laminar flow in a channel expansion. *Journal of Computational Physics* 54(3): 468–488.
- Peplinski A (2022) KTH Framework for Nek5000. <https://github.com/KTH-Nek5000/KTH-textunderscoreFramework>
- Reid J (2008) The new features of Fortran 2008. New York, NY, USA: ACM, pp. 8–21.
- Reyes R and Lomüller V (2016) SYCL: Single-source c++ accelerator programming. In: *Parallel Computing: On the Road to Exascale*. IOS Press, pp. 673–682.

- Saad Y and Schultz MH (1986) GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 7(3): 856–869.
- Schlatter P and Örlü R (2012) Turbulent boundary layers at moderate reynolds numbers: Inflow length and tripping effects. *Journal of Fluid Mechanics* 710: 5–34.
- Seddiek IS and Ammar NR (2021) Harnessing wind energy on merchant ships: case study Flettner rotors onboard bulk carriers. *Environmental Science and Pollution Research International* 28(25): 32695–32707.
- Seifert J (2012) A review of the Magnus effect in aeronautics. *Progress in Aerospace Sciences* 55: 17–45.
- Smith T, Traut M, Bows-Larkin A, et al. (2015) *CO2 Targets, Trajectories and Trends for International Shipping*.
- Świrzydowicz K, Chalmers N, Karakus A, et al. (2019) Acceleration of tensor-product operations for high-order finite element methods. *The International Journal of High Performance Computing Applications* 33(4): 735–757.
- Tanarro A, Vinuesa R and Schlatter P (2019) Effect of adverse pressure gradients on turbulent wing boundary layers. *Journal of Fluid Mechanics* 883: A8.
- Tufo HM and Fischer PF (1999) Terascale spectral element algorithms and implementations. In: *Proceedings of the 1999 ACM/IEEE Conference on Supercomputing*, p. 68.
- Wahib M and Maruyama N (2014) Scalable kernel fusion for memory-bound gpu applications. In: *SC'14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE, pp. 191–202.
- Wienke S, Springer P, Terboven C, et al. (2012) Openacc—first experiences with real-world applications. In: *European Conference on Parallel Processing*. Springer, pp. 859–870.

Author biographies

Martin Karp is a doctoral student in Computer Science at the division of Computational Science and Technology at KTH Royal Institute of Technology. Previously, he received a MSc in Engineering Physics from Lund University. His research interests revolve around large-scale simulations of turbulence and he is closely involved in the development of the spectral element code Neko.

Daniele Massaro is a PhD student studying Fluid Mechanics at KTH Royal Institute of Technology in Stockholm, Sweden, within the SimEx/FLOW group in the Engineering Mechanics department. Previously, He

received a Bachelor and a Master of Science in Aerospace Engineering from Politecnico di Milano.

Niclas Jansson is a researcher at PDC Center for High-Performance Computing, KTH Royal Institute of Technology. He received his MSc in computer science and PhD in numerical analysis from KTH Royal Institute of Technology. His research interests include scalable numerical methods and adaptive finite and spectral element methods. He has extensive experience in extreme-scale computing as a developer of RIKEN's multiphysics framework CUBE, the HPC branch of FEniCS, and the next-generation spectral element framework Neko.

Alistair Hart is a High-Performance Computing specialist based in the UK, working for Hewlett Packard Enterprise as a member of the EMEA Performance Team. Before this, he has an academic background researching theoretical high-energy physics.

Jacob Wahlgren is a PhD student in computer science at KTH Royal Institute of Technology. He is part of the ScaLab research group, with a focus on heterogeneous computing and memory for high-performance and cloud environments. He previously received a BSc in computer engineering and an MSc in computer science from KTH.

Philipp Schlatter is Professor in Fluid Mechanics at the Royal Institute of Technology (KTH) in Stockholm, Sweden. He obtained his PhD degree from ETH Zurich (Switzerland) in 2005, before he moved to Sweden. He has published over 150 papers in the fields of fluid mechanics, turbulence and simulations. Recent interests include modern computational methods applied to fluid problems, including aspects of high-performance computing and large-scale turbulence simulations, uncertainty quantification, modal decompositions, and turbulence modeling.

Stefano Markidis is an Associate Professor of Computer Science with a specialization in High-Performance Computing at the KTH Royal Institute of Technology. He holds a PhD in Nuclear Engineering from the University of Illinois at Urbana-Champaign and an MSc from the Politecnico di Torino. He is one of the recipients of the 2005 and 2017 R&D100 awards. His research interests are diverse and include large-scale simulations of space and astrophysical systems, parallel programming models, integration of machine-learning techniques into traditional numerical approaches, and quantum computing.