Augmented Lagrangian Preconditioner for Large-Scale Hydrodynamic Stability Analysis
Johann Moulin, Pierre Jolivet, Olivier Marquet

To cite this version:

HAL Id: hal-02088738
https://hal.archives-ouvertes.fr/hal-02088738
Submitted on 24 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Augmented Lagrangian Preconditioner for Large-Scale Hydrodynamic Stability Analysis

Johann Moulin* Pierre Jolivet† Olivier Marquet*

Abstract

Hydrodynamic linear stability analysis of large-scale three-dimensional configurations is usually performed with a “time-stepping” approach, based on the adaptation of existing solvers for the unsteady incompressible Navier–Stokes equations. We propose instead to solve the nonlinear steady equations with the Newton method and to determine the largest growth-rate eigenmodes of the linearized equations using a shift-and-invert spectral transformation and a Krylov–Schur algorithm. The solution of the shifted linearized Navier–Stokes problem, which is the bottleneck of this approach, is computed via an iterative Krylov subspace solver preconditioned by the modified augmented Lagrangian (mAL) preconditioner [12]. The well-known efficiency of this preconditioned iterative strategy for solving the real linearized steady-state equations is assessed here for the complex shifted linearized equations. The effect of various numerical and physical parameters is investigated numerically on a two-dimensional flow configuration, confirming the reduced number of iterations over state-of-the-art steady-state and time-stepping-based preconditioners. A parallel implementation of the steady Navier–Stokes and eigenvalue solvers, developed in the FreeFem++ language, suitably interfaced with the PETSc/SLEPc libraries, is described and made openly available to tackle three-dimensional flow configurations. Its application on a small-scale three-dimensional problem shows the good performance of this iterative approach over a direct LU factorization strategy, in regards of memory and computational time. On a large-scale three-dimensional problem with 75 million unknowns, a 80% parallel efficiency on 256 up to 2,048 processes is obtained.

Keywords Navier–Stokes equations, Newton method, Krylov–Schur method, linear stability analysis, recycled Krylov methods, distributed computing

1 Introduction

Over the past century, hydrodynamic linear stability theory was developed to understand the early stage of laminar-turbulence transition in parallel flows, such as boundary layers and shear flows [28]. In the local stability theory, the growth or decay of perturbations developing on parallel flows, described with mono-dimensional velocity profiles, is investigated assuming a normal form decomposition. The resulting eigenproblem is of small size and
does not require large computational resources to be solved. Although the local stability theory (mono-dimensional) was then extended to the description of spatially developing flows [37], the linear stability analysis of truly two- and three-dimensional flows has gained in popularity since the beginning of the century [68, 21, 66, 69, 26, 47] thanks to the development of computational resources and numerical tools allowing (a) to compute steady solutions of the governing equations and (b) to determine the most unstable eigenmodes of the linearized equations around this steady solution. An efficient and highly-parallel numerical tool is proposed in the present paper to achieve these two steps in the case of the incompressible Navier–Stokes equations.

Two main numerical approaches exist to carry out a linear stability analysis. The first one is the “time-stepping” [47] or “matrix-free” [4] approach based on the use of existing unsteady nonlinear solvers, developed in Computational Fluid Dynamics (CFD). The “matrix-free” denomination indicates that the action of matrices onto vectors is obtained without assembling them. The unsteady solvers are adapted to compute steady solutions and to extract the eigenmodes of largest growth rate, relevant in linear hydrodynamic stability analysis. For computing steady (stable or unstable) solutions, stabilization procedures, such as the recursive projection method [65], the selective frequency damping method [2], or more recently the BoostConv algorithm [24], are applied together with the unsteady nonlinear solver. The computation of leading eigenvalues is then achieved by noticing that the operations performed at each iteration of the linearized time-stepping solver correspond to an exponential-based transformation of the Jacobian operator [4, 47]. Classical Krylov subspace-based methods like Arnoldi or Krylov–Schur are then commonly used to compute the eigenvalues of the exponential operator with largest magnitude, which are also the leading (rightmost) eigenvalues of the Jacobian operator. One of the advantages of this approach is the computational-time efficiency of applying one time-step of the unsteady solvers. Indeed, these solvers are often highly optimized, not only thanks to very scalable parallel implementation, but also because efficient numerical algorithms have been developed for solving the time-discretized problems (e.g., splitting [41] or fractional step [44] methods). The drawback of this approach is the slow convergence of the Arnoldi method induced by the use of time-steppers. Indeed, small time-steps are required for an application of the linearized time-stepper to approximate accurately the exponential transformation [72]. This leads to a large number of so-called “outer” iterations (in the $10^3–10^4$ range) to converge only a few eigenvalues. The efficiency of the “time-stepping” approach is thus mainly based on fast outer iterations at the expense of a large number of such iterations in the Arnoldi process. Note that other strategies for computing matrix exponential allow to relax the small time-step constraint and thus provide better convergence properties [19, 60].

The second existing numerical approach to perform linear stability analysis in hydrodynamics [66] is referred here to as the “matrix-based” approach. It relies on the assembly of sparse matrices resulting from the spatial discretization of the underlying problem and the solution of corresponding linear systems using existing parallel libraries that implement direct sparse $LU$ factorization of those matrices (MUMPS [3], SuperLU [45]). The steady-state solutions are then computed by solving the steady nonlinear equations with a (quasi-)Newton method. An invert-based spectral transformation of the Jacobian operator, like the shift-and-invert [22, 29, 63, 66] or Cayley [25, 52, 49] transformations, is then applied with a Krylov subspace-based method (typically, the Arnoldi method [61]) to
determine the leading eigenvalues. The Newton method and the shift-and-invert strategy allow, respectively, to achieve fast convergence towards the steady solution and the leading eigenvalues. Usually, the number of Newton iterations is around 10 or so to compute a steady solution, while it may require a few hundred outer iterations in the Arnoldi algorithm to compute a few eigenvalues. This reduced number of applications comes at a price: it requires the ability to invert the linearized steady (and generally shifted) Navier–Stokes equations. Consequently, it has mainly been used for linear stability analysis of two-dimensional flow configurations, for which the number of unknowns remains limited (not much greater than $10^5$ unknowns) and the Jacobian matrices remain sparse, so that the LU factorization is affordable. The “matrix-based” strategy is particularly efficient for the eigenvalue computation, since the time-consuming LU factorization of the sparse Jacobian matrix is done once for all, while only the forward eliminations and back substitutions are repeated at each outer iteration. The main drawback of this approach is the large amount of memory needed to perform factorization, especially for three-dimensional flow configurations [51]. For large-scale hydrodynamics problems, the high cost of forming the Jacobian matrix explicitly, and the prohibitive memory requirements of direct solvers drove many authors [5, 48, 23] towards the “matrix-free” strategy.

At the early beginning of nineties, Tuckerman [70, 50, 72] proposed to improve the slow convergence of the “matrix-free” approach by using a Newton method (resp. a shift-and-invert strategy) for the steady-state (resp. eigenvalue) computation, while still using an existing unsteady solver. This method is based on the observation that one can adapt a (linearized) unsteady solver in order to apply, to some given vector, the steady Navier–Stokes Jacobian operator, left-preconditioned by the (unsteady) Stokes operator. Thus, this technique provides a cheap “matrix-free” way of preconditioning the Navier–Stokes Jacobian operator by the (unsteady) Stokes operator, for use inside Krylov subspace linear solvers typically. The method is nowadays known as the “Stokes” preconditioning technique and has been largely applied during the last decades for the computation of steady-state and leading eigenvalues [7, 15, 73, 53, 71]. Recently, it has been adapted and applied to the determination of resolvent modes in large-scale three-dimensional configurations [17]. In the Stokes preconditioning technique, the time-step of the linearized unsteady solver becomes a parameter of the preconditioner. Large time-steps usually provide better preconditioning, but make the application of one linearized time iteration harder. More details and improvements of the method can be found in [8]. In any case, the performance of this method remains limited by the efficiency of the (unsteady) Stokes operator to precondition the linearized steady Navier–Stokes operator.

In the present paper, we propose to develop a “matrix-based” specific solver for performing linear stability analysis, which relies on state-of-the-art preconditioners for the linearized steady incompressible Navier–Stokes equations, thus avoiding the use of direct solvers on the full problem. Over the last decades, various promising approaches have been developed aiming at overcoming the two main difficulties of this problem: the saddle-point structure of the equations deriving from the incompressibility constraint and the absence of a (small) time-step parameter that greatly enhances the convergence of iterative algorithms, due to the resulting diagonal dominance of the matrix. Among those steady preconditioners are the well-known SIMPLE preconditioner [58], the more recent Pressure Convection–Diffusion (PCD) preconditioner proposed by [13], as well as the original augmented Lagrangian (AL) [10, 11, 35] and modified augmented Lagrangian (mAL) [12, 13, 11, 14] pre-
conditioners. Several authors showed the superiority of the modified augmented Lagrangian approach over other state-of-the-art alternatives for solving the Oseen and linearized incompressible Navier–Stokes equations \[64, 33\]. Moreover, a very recent work \[31\] proposed an efficient and highly scalable steady Navier–Stokes solver based on the original augmented Lagrangian preconditioner. If the augmented Lagrangian strategy has been regularly used for steady-state computations, it was never tested on practical case of eigenvalue computations. A work in that direction was however proposed by Olshanskii and Benzi \[56\], who adapted the original augmented Lagrangian preconditioner to solve the shifted linearized Navier–Stokes equations. They showed theoretically and numerically that AL was robust to a real-valued shift on a variety of 2D flow configurations. Complex-valued shifts, as needed in practice to efficiently explore the complex plane with a shift-and-invert strategy, were not considered.

The first objective of the present paper is to assess the efficiency of the modified augmented Lagrangian preconditioner for the computation of steady-state solutions with a Newton method and leading eigenvalues with a complex shift-and-invert strategy. The second objective is to describe, and test on a three-dimensional flow configuration, an open-source parallel implementation of the modified augmented Lagrangian preconditioner for linear stability analysis purposes, using the FreeFem++ finite element library \[34\] interfaced with PETSc \[6\] and SLEPc \[36\]. The full code is made available at https://github.com/prj-/moulin2019al.

The paper is organized as follows. The governing equations required to carry out the linear stability analysis of incompressible flows are introduced in section 2. The Newton method used to solve the steady nonlinear equations and the eigensolver based on the shift-and-invert strategy are also described. The preconditioning technique and the modified augmented Lagrangian preconditioner are introduced in section 3. The parallel implementation is detailed in section 4. Numerical results are given in section 5. First we examine, on a two-dimensional problem, the effect of various numerical and physical parameters on the performance of the mAL preconditioner for solving the complex shifted linearized Navier–Stokes problem. Then we compare the performance of mAL with other state-of-the-art preconditioners. Finally, we evaluate the performance of the proposed parallel implementation by first comparing it to a sparse direct solver on a small-scale three-dimensional test case and then, by testing its scalability on a large-scale configuration that cannot be afforded with a direct sparse solver.

2 Methods for linear stability analysis in hydrodynamics

2.1 Governing equations

Let us consider an incompressible flow, described by the two-dimensional (resp. three-dimensional) velocity field \(u = [u, v]^T\) (resp. \(u = [u, v, w]^T\)) and the pressure field \(p\), that satisfy the incompressible Navier–Stokes equations:

\[
\frac{\partial u}{\partial t} + (u \cdot \nabla)u + \nabla p - \frac{1}{Re} \nabla^2 u = 0 , \quad -\nabla \cdot u = 0
\]
The Reynolds number is defined as $Re = U_\infty L/\nu$, where $U_\infty$ and $L$ are characteristic velocity and length used to make non-dimensional the velocity and pressure fields, and $\nu$ is the kinematic viscosity. For conciseness, the Navier–Stokes equations are rewritten in a state-space form as follows

$$M \frac{\partial \mathbf{q}}{\partial t} + \mathcal{R}(\mathbf{q}) = 0,$$

where $\mathbf{q} = (u, p)^T$ is the state-space vector. Base flows, denoted hereinafter $\mathbf{q}_b(x)$, are time-independent (steady) solutions of the Navier–Stokes equations (1) and thus satisfy the nonlinear steady Navier–Stokes equations

$$\mathcal{R}(\mathbf{q}_b) = 0.$$ (2)

Linear stability of base flows is investigated by superimposing infinitesimal perturbations $\mathbf{q}'(x, t)$ to the base flow solution, i.e. $\mathbf{q}(x, t) = \mathbf{q}_b(x) + \epsilon \mathbf{q}'(x, t)$, where $\epsilon$ is an infinitesimal parameter. After inserting this decomposition into eq. (1), using the definition eq. (2) of the base flow and neglecting high-order terms in $\epsilon$, one obtains the linearized Navier–Stokes equations governing the temporal evolution of the infinitesimal perturbation,

$$M \frac{\partial \mathbf{q}'}{\partial t} + \mathcal{J}(\mathbf{q}_b) \mathbf{q}' = 0,$$

where $\mathcal{J}(\mathbf{q}_b) = \left( \begin{array}{cc} (u_b \cdot \nabla)(\bullet) + (\bullet \cdot \nabla)u_b - Re^{-1}\nabla^2 \bullet - \nabla \cdot \bullet & 0 \\ -\nabla \cdot \bullet & 0 \end{array} \right)$

is the Jacobian operator defined around the base flow $\mathbf{q}_b$. The long-term evolution of any infinitesimal perturbation is conveniently described by assuming a spectral decomposition of perturbations as $\mathbf{q}' = \hat{\mathbf{q}}(x) e^{\sigma t} + c.c.$, where $\hat{\mathbf{q}}(x)$ is a complex spatial field whose temporal evolution is exponential and given by the complex number $\sigma = \lambda + i \omega$. $\lambda$ is the growth rate and $\omega$ is the angular frequency. Inserting this modal decomposition into the above linearized equations shows that $\sigma$ and $\hat{\mathbf{q}}$ are respectively eigenvalues and eigenmodes of the generalized eigenproblem:

$$\sigma M \hat{\mathbf{q}} + \mathcal{J}(\mathbf{q}_b) \hat{\mathbf{q}} = 0.$$ (3)

The stability of the base flow is then determined by considering the leading eigenmode $\hat{\mathbf{q}}_0$ associated to the eigenvalue $\sigma_0 = \lambda_0 + i \omega_0$ with the largest real part $\lambda_0$. When the growth rate of the leading eigenmode is negative ($\lambda_0 < 0$), all the eigenvalues have negative real parts, and the base flow is linearly stable since any perturbations superimposed to the base flow is damped at sufficiently large time. On the other hand, when the growth rate of the leading eigenmode is positive ($\lambda_0 > 0$), the perturbation will grow in time and the base flow is linearly unstable [66].

A linear stability analysis thus consists first in computing a base flow, which is a solution of the steady Navier–Stokes eq. (2), and then in determining the leading eigenvalues/eigenmodes of the eigenproblem (3) with the largest growth rate.

### 2.2 Spatial discretization

In the present paper, a finite element method is used for the spatial discretization of the nonlinear steady equations (2) and of the linear eigenproblem (3) on a $d$-dimensional ($d = 2, 3$) domain $\Omega$. A grad-div stabilized weak formulation [54] of eq. (2) is used, which
consists in finding \( \mathbf{u}_b \) in \( \mathcal{V}_\Gamma = \{ \mathbf{u} \in (H^1(\Omega))^d, \text{ s.t. } \mathbf{u} = \mathbf{u}_\Gamma \text{ on } \Gamma \} \) and \( p_b \) in \( Q = L^2(\Omega) \) such that:

\[
\begin{align*}
\mathcal{R}_u(\mathbf{q}_b; \mathbf{u}) &= \langle \mathbf{u}_b \cdot \nabla \mathbf{u}_b, \mathbf{u} \rangle + \langle Re^{-1} \nabla \mathbf{u}_b, \nabla \mathbf{u} \rangle - \langle p_b, \nabla \cdot \mathbf{u} \rangle + \gamma \langle \nabla \cdot \mathbf{u}_b, \nabla \cdot \mathbf{u} \rangle = 0 \quad (4a) \\
\mathcal{R}_p(\mathbf{q}_b; \bar{p}) &= -\langle \nabla \cdot \mathbf{u}_b, \bar{p} \rangle = 0 \quad (4b)
\end{align*}
\]

for all \( (\mathbf{u}, \bar{p}) \) in \( \mathcal{V}_0 \times Q \), where \( (\cdot, \cdot) \) denotes the \( L^2 \) inner-product and \( \mathcal{V}_0 = \{ \mathbf{u} \in (H^1(\Omega))^d, \text{ s.t. } \mathbf{u} = 0 \text{ on } \Gamma \} \) is the velocity space with vanishing velocity on the boundary \( \Gamma \). The weak residuals of the momentum and mass conservation equations are \( \mathcal{R}_u \) and \( \mathcal{R}_p \), respectively. The last term in the momentum residual \( \mathcal{R}_u \) is the grad–div stabilization (also called augmentation) term that corresponds to the weak form of \(-\gamma \nabla (\cdot \mathbf{u}_b)\), with \( \gamma \geq 0 \) a numerical parameter. In the above continuous weak formulation, the stabilization term strictly vanishes on the solution: \( \langle \nabla \cdot \mathbf{u}_b, \nabla \cdot \mathbf{u} \rangle = 0 \). Indeed, the divergence-free condition \( \langle \nabla \cdot \mathbf{u}_b, \bar{p} \rangle = 0 \) is satisfied for all \( \bar{p} \in Q \), and in particular for \( \nabla \cdot \mathbf{u} \in Q \).

A Delaunay triangulation of the domain \( T_h = \{ K \} \), consisting in triangular \((d = 2)\) or tetrahedral \((d = 3)\) elements \( K \), is used. In order to satisfy the inf–sup Ladyženskaja–Babuška–Brezzi (LBB) condition (see [16]), the Taylor–Hood finite element pair is chosen, so that the discrete velocity \( \mathbf{u}^h \) and pressure \( p^h \) are sought respectively in \( \mathcal{V}^h = \{ \mathbf{u}^h \in C^0(\Omega), \text{ s.t. } \mathbf{u}^h|_K \in P_2(K), \forall K \in T_h \, , \mathbf{u}^h = \mathbf{u}_\Gamma \text{ on } \Gamma \} \) and \( Q^h = \{ p^h \in C^0(\Omega), \text{ s.t. } p^h|_K \in P_1(K), \forall K \in T_h \} \). Note that, with the Taylor–Hood finite element pair, the discrete divergence of the velocity test functions does not belong to the discrete pressure space, i.e. \( \nabla \cdot \mathbf{u}^h \not\in Q^h \). Therefore, contrary to the continuous case, the stabilization term does not vanish from the discrete momentum equation \((\nabla \cdot \mathbf{u}^h, \nabla \cdot \mathbf{v}^h) \neq 0\), and the discrete solution depends on the value of the stabilization parameter \( \gamma \). Here, the grad–div stabilization is mainly introduced to improve, thanks to an efficient preconditioner (see section [3]), the iterative solution of linear systems involved when solving the nonlinear discrete equations [5]. The question of whether the grad–div stabilized discrete solution is closer or further from the continuous weak solution is out of the scope of this paper. However, several studies (e.g. [55, 57, 54, 46, 35]) showed that the grad–div stabilization often improves the mass conservation property and the velocity error of the discrete solution, for adequate values of \( \gamma \). Numerical experiments are performed in section [5.2.1] to assess the accuracy of the stabilized discrete solution and to determine adequate values of \( \gamma \).

Such a discretization yields the following discrete version of the nonlinear base flow eq. [2]:

\[
\mathbf{R}(\mathbf{q}_b^h) = 0
\]

where \( \mathbf{q}_b^h \) denotes now the vector of coefficients of \( \mathbf{u}_b^h \) and \( p_b^h \) in the finite elements basis.

The generalized eigenproblem [3] is discretized similarly, yielding

\[
\sigma \mathbf{M} \mathbf{q}^h + \mathbf{J}(\mathbf{q}_b^h) \mathbf{q}^h = 0
\]

where \( \mathbf{M} \) and \( \mathbf{J}(\mathbf{q}_b^h) \), the finite element matrices obtained after discretization of the mass \( \mathbf{M} \) and Jacobian operator \( \mathbf{J}(\mathbf{q}_b^h) \), are respectively defined as

\[
\mathbf{M} = \begin{pmatrix}
\mathbf{M}_u & 0 \\
0 & 0
\end{pmatrix}
\quad \text{and} \quad
\mathbf{J}(\mathbf{q}_b^h) = \begin{pmatrix}
\mathbf{A}_\gamma & \mathbf{B}^T \\
\mathbf{B} & 0
\end{pmatrix}.
\]

The rectangular matrix \( \mathbf{B} \) is the discretization of the divergence operator and its transpose \( \mathbf{B}^T \) represents the discrete gradient. The mass matrix on the velocity space \( \mathbf{M}_u \) can be
written as a 3-by-3 block diagonal matrix corresponding to the three velocity components. The 3-by-3 block matrix $A_\gamma = A + \gamma \Gamma$ is the sum of $A$, which represents the linearized diffusion and convection terms in the momentum conservation equation, and $\Gamma$, obtained after discretization of the grad–div stabilization term. They write:

$$
A_\gamma = \begin{pmatrix}
A_{uu} + \gamma \Gamma_{uu} & A_{uw} + \gamma \Gamma_{uw} & A_{ww} + \gamma \Gamma_{ww} \\
A_{vu} + \gamma \Gamma_{vu} & A_{vv} + \gamma \Gamma_{vv} & A_{ww} + \gamma \Gamma_{ww} \\
A_{wu} + \gamma \Gamma_{wu} & A_{wv} + \gamma \Gamma_{wv} & A_{ww} + \gamma \Gamma_{ww}
\end{pmatrix}, \quad M = \begin{pmatrix}
M_u & 0 & 0 \\
0 & M_v & 0 \\
0 & 0 & M_w
\end{pmatrix}.
$$

In the following, we will mostly refer to the discrete solutions. Therefore, the superscript $^h$ is dropped unless confusion is possible.

### 2.3 Nonlinear steady-state solver

The nonlinear solution $q_b$ of the discrete problem eq. (5) is obtained by the classical Newton method. The approximated solution at the $k$th iteration is obtained as

$$
q_b^k = q_b^{k-1} + \delta q_b^k,
$$

where $\delta q_b^k$ denotes the solution increment, obtained by solving the linear problem

$$
J(q_b^{k-1}) \delta q_b^k = -R(q_b^{k-1})
$$

where $J(q_b^{k-1})$ is the Jacobian matrix defined in eq. (7) with the known approximation of the steady solution $q_b^{k-1}$. The solution of this linear system is repeated for each iteration of the Newton algorithm, that is considered to be converged when the $l^2$ norm of the residual $\|R(q_b)\|_2$ is below some numerical tolerance.

### 2.4 Linear eigensolver

The Krylov–Schur algorithm \[67\] is used in the present study to solve the generalized eigenproblem (6). In order to compute the leading eigenvalues, which lie in the complex plane close to the zero growth-rate axis ($\lambda = 0$) for any frequency $\omega$, a shift-and-invert spectral transformation is first applied, yielding the transformed eigenproblem

$$
\mu \hat{q} + T \hat{q} = 0, \quad T = (J(q_b) + sM)^{-1} M
$$

where $s = s_r + is_i$ is the so-called complex shift. The eigenvalues $\mu$ of the transformed problem are related to the eigenvalues $\sigma$ of eq. (6) through $\mu = (\sigma - s)^{-1}$ while the eigenvectors are left unchanged by the spectral transformation. Like the classical power method, the Krylov–Schur algorithm allows to compute the eigenvalues of largest magnitude. When applied to the transformed problem, it gives the eigenvalues $\mu$ of largest magnitude, which correspond to the eigenvalues $\sigma$ closest to the complex shift $s$. To determine the leading eigenvalue of eq. (6), the eigenproblem (11) is solved for several values of the complex $s$ close to the real axis, spanning appropriately the imaginary axis. For each eigenvalue computation, the Krylov–Schur algorithm requires multiple “matrix–vector” applications of the matrix $T$. In other words, repeated solutions of the linear system $(J(q_b) + sM)q_i = q_i$ are required, where the right-hand side vectors $q_i$ are given by the Krylov–Schur algorithm.
In the present work, linear stability analysis is thus performed using a nonlinear steady-state solver and a linear eigensolver, that both rely on multiple solutions of linear systems involving the complex shifted Jacobian matrix \((J + sM)\). For the steady-state solver, this matrix reduces to the real Jacobian matrix \(J\) as the complex shift vanishes \(s = 0\). The next section introduces a preconditioned iterative method used to solve efficiently such systems.

3 An augmented Lagrangian approach for the shifted Jacobian matrix

As explained in the previous section, the main challenge of an hydrodynamic stability analysis is to solve efficiently the following linear equation:

\[
\begin{pmatrix}
A_{\gamma,uu} + sM_u & A_{\gamma,uv} & A_{\gamma,uw} & B_u^T \\
A_{\gamma,vu} & A_{\gamma,vv} + sM_v & A_{\gamma,vw} & B_v^T \\
A_{\gamma,wu} & A_{\gamma,wv} & A_{\gamma,ww} + sM_w & B_w^T \\
B_u & B_v & B_w & 0
\end{pmatrix}
\begin{pmatrix}
u_o \\
v_o \\
w_o \\
p_o
\end{pmatrix} =
\begin{pmatrix}
u_i \\
v_i \\
w_i \\
p_i
\end{pmatrix}
\]

(12)

where \(A_{\gamma,\alpha\beta} = A_{\alpha\beta} + \gamma \Gamma_{\alpha\beta} \) (\(\alpha, \beta = u, v, w\)), \(q_o = (u_o, v_o, w_o, p_o)^T\) is the solution vector and \(q_i = (u_i, v_i, w_i, p_i)^T\) is a right-hand side vector. In the perspective of large-scale computations, we must avoid the use of direct solvers applied directly to eq. (12), due to their huge memory cost [51]. Instead, we use the flexible Generalized Minimal Residual algorithm (GMRES) [62] for solving iteratively eq. (12). The shifted-Jacobian matrix being indefinite and ill-conditioned, the use of an iterative method without preconditioning is inefficient as it requires a very large number of iterations [71]. To improve the numerical efficiency of the iterative solution, the above linear system is replaced by the right-preconditioned linear system:

\[
\begin{pmatrix}
A_{\gamma,uu} + sM_u & A_{\gamma,uv} & A_{\gamma,uw} & B_u^T \\
A_{\gamma,vu} & A_{\gamma,vv} + sM_v & A_{\gamma,vw} & B_v^T \\
A_{\gamma,wu} & A_{\gamma,wv} & A_{\gamma,ww} + sM_w & B_w^T \\
B_u & B_v & B_w & 0
\end{pmatrix} P^{-1}
\begin{pmatrix}
u_o \\
v_o \\
w_o \\
p_o
\end{pmatrix} =
\begin{pmatrix}
u_i \\
v_i \\
w_i \\
p_i
\end{pmatrix}
\]

(13)

where the matrix \(P\) is the so-called preconditioner and \((\tilde{u}_o, \tilde{v}_o, \tilde{w}_o, \tilde{p}_o)^T\) is an intermediate solution. The final solution is found by solving the following linear system

\[
P
\begin{pmatrix}
u_o \\
v_o \\
w_o \\
p_o
\end{pmatrix} =
\begin{pmatrix}
\tilde{u}_o \\
\tilde{v}_o \\
\tilde{w}_o \\
\tilde{p}_o
\end{pmatrix}
\]

(14)

The GMRES algorithm is applied to the right-preconditioned eq. (13) which, in addition to matrix–vector products with the shifted-Jacobian matrix, requires the repeated application of \(P^{-1}\), i.e., the solution of eq. (14). A good preconditioner achieves a compromise between a fast application of the preconditioner and a small number of iterations to solve
the preconditioned system.

The augmented Lagrangian preconditioner allows to solve iteratively the Oseen \[10, 12, 11, 35\] and linearized Navier–Stokes equations \[13, 56\] in a very limited number of iterations, regardless of the mesh refinement and the Reynolds number value. Nevertheless, these interesting properties are counterbalanced by the difficulty of solving iteratively the coupled (two or three-dimensional) velocity subproblem arising in the application of the original preconditioner, as it requires highly specific multigrid solvers \[10, 31\]. In order to circumvent this particular issue, the so-called modified augmented Lagrangian (mAL) preconditioner was introduced \[12\]. It is derived from the original augmented Lagrangian preconditioner by neglecting either the lower block matrices \[12\] or the upper block matrices \[33\], as follows

\[
P_{\text{mAL}} = \begin{pmatrix}
A_{\gamma,uu} + sM_u & 0 & 0 & 0 \\
A_{\gamma,uv} & A_{\gamma,uv} + sM_v & 0 & 0 \\
A_{\gamma,wu} & A_{\gamma,wv} & A_{\gamma,ww} + sM_w & 0 \\
B_u & B_v & B_w & S_p
\end{pmatrix},
\]

where \(S_p\) is an approximation of the pressure Schur complement \(-B(A_{\gamma} + sM_u)^{-1}B^T\). Rather than being explicitly specified, this matrix is defined by the action of its inverse as

\[
S_p^{-1} = -(\gamma + Re^{-1})M_p^{-1} - sL_p^{-1},
\]

where \(M_p\) is the mass matrix and \(L_p\) the Laplacian matrix, both defined on the discrete pressure space. Note that, for base flow computations \(s = 0\), only the first term remains in the definition of the approximated Schur complement eq. \[16\]. The lower block-triangular version of the preconditioner is chosen for practical reasons explained in section \[4\]. In the original preconditioner proposed by \[10, 12\], the augmentation term of the Jacobian matrix \(\Gamma\) was defined algebraically as \(B^T M_p^{-1} B\), thus requiring two sparse matrix products to be constructed explicitly. As proposed in \[35\], the construction cost can significantly be reduced by building the matrix \(\Gamma\) from the finite element discretization of the grad–div stabilization term. They showed that \(\Gamma\) is then the sum of the algebraic augmentation \(B^T M_p^{-1} B\) and of a stabilization matrix. The efficiency of the mAL preconditioner is thus conserved while significantly reducing the construction costs. Finally, the grad–div augmentation matrix \(\Gamma\) is much more sparse than its algebraic counterpart, thus motivating our choice for this implementation of the mAL preconditioner, in the perspective of large-scale three-dimensional computations.

4 Parallel implementation with FreeFem++ and its interface to PETSc/SLEPc

For realistic three-dimensional geometries, the approach derived in the previous sections requires the solution of nonlinear systems and generalized eigenproblems of large dimensions. Thus, high-performance computing becomes necessary. The goal of this section is to show how this is done using a finite element domain specific language, FreeFem++ \[34, 39\],
interfaced with distributed linear algebra backends, PETSc [6] and SLEPc [36]. A thorough introduction of these libraries may be found in their respective manuals [1, 2, 3]. Our implementation is openly available at https://github.com/prj-/moulin2019al and the rest of this section follows the available source code.

4.1 Outer solvers

In this section, we describe how the outer solvers (i.e. the nonlinear steady-state solver and the eigensolver) are implemented.

The Newton method described in section 2.3 is implemented using FreeFem++. Only the inversion of the Jacobian matrix $J$ of eq. (10) is performed by PETSc. Given a FreeFem++ distributed version of the Jacobian matrix $dJ$, PETSc options defining the linear solver for eq. (10) are set using the following FreeFem++ syntax:

```cpp
set(dJ, sparams = params, fields = vX[], names = names,
    schurPreconditioner = S, schurList = listX[]);
```

The keyword `sparams` is a string defined by the user gathering the PETSc runtime options for the Krylov subspace solver (KSP) and preconditioner (PC). The interested reader should refer to PETSc manual for details on the use of runtime options. The keywords `fields`, `names`, `schurPreconditioner`, and `schurList` allow to implement specific block preconditioners, like mAL, and their use is detailed in the next sections.

For the eigenvalue computation presented in section 2.4, only the finite element matrices are built by FreeFem++. Then, the Krylov–Schur algorithm is performed entirely by SLEPc through the use of the eigenvalue problem solver framework (EPS), which is called from within FreeFem++.

```cpp
int k = zeigensolver(dJ, dM, vectors = vec, values = val, sparams = params,
    fields = vX[], names = names, schurPreconditioner = S, schurList = listX[]);

// solves the eigenvalue problem $dJq = \sigma dMq$
```

Contrary to the case of the linear solver interface, two matrices $dJ$ and $dM$ that define the generalized eigenproblem (6) must now be passed to SLEPc. In addition, `sparams` must also contain the SLEPc runtime options defining the eigensolver.

4.2 Inner mAL-preconditioned linear solvers

The inner linear solves of system eq. (12) with a mAL-preconditioned GMRES require the implementation of the block structure of the preconditioner (eq. (15)). This is done in PETSc by using the so-called fieldsplit structure that gives to the users a high-level of abstraction to define operators by blocks. The following PETSc runtime options define such a preconditioner:

```cpp
string params = paramsXYZ + " " + paramsP + " " + paramsKrylov + " -pc_type fieldsplit -pc_fieldsplit_type multiplicative";
```

The desired lower block-triangular structure of the preconditioner is obtained by the use of PETSc keyword *multiplicative*. The strings `paramsXYZ` and `paramsP` respectively contain the innermost velocity and pressure block solvers options that will be detailed later on. The string `paramsKrylov` contains the definition of the Krylov subspace linear solver. For example, one should simply write `paramsKrylov = "-ksp_type fgmres"` to use the flexible GMRES. In order to implement the modified augmented Lagrangian preconditioner $P_{mAL}$ through the fieldsplit structure, in 3D, the four fields $u, v, w,$ and $p$ must be defined. Assuming the problem is formulated in the full vectorial finite element space $WH$, containing the velocities and pressure unknowns, one must be able to differentiate the degrees of freedom belonging to each field. To that aim a finite element function taking a different integer value for each one of the four fields is defined in FreeFem++ and passed to PETSc/SLEPc through the keyword *fields*. Then, for simplicity, each field is attributed a name that will be used to identify it when defining the different innermost solvers associated to the diagonal blocks of $P_{mAL}$, c.f. section 4.3. Those names are contained in an array of strings, that is provided to the solver through the keyword *names*.

```c
Wh[vX, vY, vZ, p] = [1, 2, 3, 4]; // numbering of each field
string[int] names(4); // prefix of each field
names[0] = "vX"; // x-velocity
names[1] = "vY"; // y-velocity
names[2] = "vZ"; // z-velocity
names[3] = "p"; // pressure
```

**Approximate Schur complements** The default setting in PETSc, when using a multiplicative fieldsplit preconditioner, is to define the preconditioner as the lower block triangular part of the system matrix in eq. (12). Thus, on the block diagonal of such a preconditioner, one would have $A_{\gamma,uu} + sM_u$, $A_{\gamma,vv} + sM_v$, $A_{\gamma,ww} + sM_w$ and the null matrix. In order to implement $P_{mAL}$, one must replace, in the preconditioner only, the default operator for the pressure field (the null matrix) by the ones necessary to implement the desired Schur complement approximation eq. (16). This is done in PETSc using `PCFieldSplitGetSubKSP` to retrieve the operators linked to each field of the fieldsplit structure and then `KSPSetOperators` to set the new operators that define eq. (16).

When the shift $s$ is null, for base flow computations, the approximate Schur complement only requires the assembly of one operator: $(\gamma + Re^{-1})^{-1}M_p$. This is done in FreeFem++ as shown below, and then passed to PETSc with the keyword *schurPreconditioner*.

```c
matrix[int] S(1); // array with a single matrix
varf vSchur(p, q) = int3d(th, qforder = 3)
  (-1.0/(gamma + 1.0/Re) * p * q); // eq. (16) with s = 0
S[0] = vSchur(Qh, Qh); // matrix assembly
```

For eigenvalue computations, two auxiliary operators are now needed: $(\gamma + Re^{-1})^{-1}M_p$ and $s^{-1}L_p$. The construction in FreeFem++ is performed using the following lines, and then again passed to SLEPc with the keyword *schurPreconditioner*.
matrix<complex>[int] S(2); // array with two matrices

complex scale;

varf vMp(p, q) = int3d(th, qforder = 3)(scale * p * q); // eq. [7]
scale = 1.0/(gamma + 1.0/Re);
S[0] = vMp(Qh, Qh); // first matrix assembly

macro grad(p) [dx(p), dy(p), dz(p)] // macro for computing \nabla p

varf vLp(p, q) = on(3, p = 1) // inlet boundary condition
+ int3d(th, qforder = 2)(scale * (grad(p)' * grad(q)));

// shift value s
complex s = getARGV("-shift_real", 1.0e-6) + getARGV("-shift_imag", 0.6) * 1i;

class scale = 1.0/s;
S[1] = vLp(Qh, Qh); // second matrix assembly

Finally, we note that the operators needed for the Schur complement approximation are built on the pressure space Qh. However, in FreeFem++, it is not possible to know a priori the correspondence between the numbering of \( \text{Wh} \), where the full solution is defined, and Qh. To circumvent this issue, we compute this correspondence in FreeFem++ from an interpolation between Qh and \( \text{Wh} \), and then pass it to PETSc/SLEPc with the keyword schurList:

Qh pIdx; // function from the pressure space
pIdx[] = 1:pIdx[].n; // numbering of the unknowns of Qh
// renumbering into the complete space by doing an interpolation on Wh
Wh [listX, listY, listZ, listP] = [0, 0, 0, pIdx];

4.3 Innermost velocity and pressure linear solvers

The approximate inverse of the diagonal blocks in eq. (15) are defined using off-the-shelf iterative methods from PETSc. For each velocity field, the GMRES is used, right-preconditioned by an additive Schwarz method (ASM) with one-level of algebraic overlap, as well as exact \( LU \) factorizations for each subdomain solver. These factorizations are carried out by MUMPS [3]. A maximum Krylov dimension of 50 is prescribed and the GMRES is stopped when the relative unpreconditioned residual norm is lower than \( 10^{-1} \).

In our implementation, the PETSc runtime options defining the approximate inverse of the diagonal velocity blocks are contained in the string \( \text{paramsXYZ} \) detailed below:

real tolV = getARGV("-velocity_tol", 1.0e-1); // default to \( 10^{-1} \)
// monodimensional velocity solver

string paramsV = 
"-ksp_type gmres 
-ksp_converged_reason -ksp_pc_side right " + 
"-ksp_rtol " + tolV + " -ksp_gmres_restart 50 -pc_type asm " + 
"-pc_asm_overlap 1 -sub_pc_type lu -sub_pc_factor_mat_solver_type mumps";
// each velocity component gets the same monodimensional solver
// defined by paramsV

string paramsXYZ = 
"-prefix_push fieldsplit_vX_ " + paramsV + " -prefix_pop" + 
" -prefix_push fieldsplit_vY_ " + paramsV + " -prefix_pop" + 
" -prefix_push fieldsplit_vZ_ " + paramsV + " -prefix_pop";

For the pressure Schur complement approximate inverse eq. (16), the PETSc runtime options defining the solver are contained in the string \( \text{paramsP} \). We must distinguish the cases of the base flow and eigensolvers. For the former \( (s = 0) \), only the action of \( M_p^{-1} \) has to
be evaluated. For that purpose, we use at most five iterations of the Jacobi-preconditioned conjugate gradient algorithm:

```
string paramsP = "-prefix_push fieldsplit_p_ " +
                "-ksp_type cg -ksp_max_it 5 -pc_type jacobi -prefix_pop";
```

For the eigensolver ($s \neq 0$), the action of the inverse of the Schur complement is approximated by the sum of the action of $(\gamma + Re^{-1})M_p^{-1}$ and $sL_p^{-1}$. This is done through PETSc composite preconditioner:

```
string paramsP = "-prefix_push st_fieldsplit_p_ " +
                "-ksp_type preonly -pc_type composite " +
                "-prefix_push sub_0_ " + // action of $(\gamma + Re^{-1})M_p^{-1}
                "-pc_type bjacobi -sub_pc_type icc -prefix_pop " +
                "-prefix_push sub_1_ " + // action of $sL_p^{-1}
                "-pc_type gamg -pc_gamg_square_graph 10 -prefix_pop " +
                "-prefix_pop";
```

Here only one application of the block Jacobi preconditioner with ICC(0) subsolvers [20] is used for approximating the mass matrix inverse while one V-cycle of GAMG [1] is used for the Laplacian term.

5 Numerical results

The efficiency of the modified augmented Lagrangian (mAL) preconditioner is investigated in this section by performing the linear stability analysis of two- and three-dimensional flow configurations described in section 5.1. The two-dimensional computations are always performed on one process as they are of limited size. For the three-dimensional case, the fully parallel implementation presented in section 4 is used. The influence of various numerical and physical parameters, such as the augmentation parameter, the mesh size and the Reynolds number, is first assessed in section 5.2 for the two-dimensional configuration, before comparing the performance of mAL preconditioner with other block preconditioners (PCD, SIMPLE, Stokes) in section 5.3. The efficiency of the parallel implementation is finally investigated in section 5.4 for the three-dimensional configuration.

5.1 Two- and three-dimensional test cases

The two-dimensional flow configuration is sketched in fig. 1a. A thin plate of height $h$ and thickness $t = h/6$ is immersed in an incoming flow of uniform velocity $U_\infty$. The size of the computational box indicated in the figure and the flow variables are made non-dimensional using $h$ as characteristic length and $U_\infty$ as characteristic velocity, so that the Reynolds number is defined as $Re = U_\infty h/\nu$, where $\nu$ is the kinematic viscosity.

Triangulations of the computational domain are obtained with the internal mesh generator of FreeFem++. The no-slip boundary condition $u = v = 0$ is applied on the plate, symmetry boundary conditions ($\partial_y u = 0$ and $v = 0$) are applied at the top and bottom boundaries of the computational domain, and a stress-free boundary condition is applied at the outlet boundary.

A typical steady solution of the incompressible Navier–Stokes equation is displayed in fig. 2a.
Figure 1: Two-dimensional and three-dimensional flow configurations. Sketch of the computational domains used for (a) the two-dimensional plate of height $h = 1$ and thickness $t = 1/6$ and (b) the three-dimensional plate of span $L = 2.5$ immersed in an upstream uniform streamwise flow $U_\infty$.

The flow recirculates in two symmetric regions in the wake of the plate, as indicated by the streamlines. The linear stability analysis of this base flow yields the eigenvalue spectrum shown in fig. 2c with circles. A pair of complex conjugate unstable eigenvalues is found ($\lambda > 0$) characterized by an angular frequency $\omega = 0.70$. For a lower value of the Reynolds number $Re = 30$, this eigenvalue is stable as shown by the square symbols. The real part of the eigenmode associated to this leading eigenvalue is depicted in fig. 2b with isocontours of the streamwise velocity. The spatial structure of this eigenmode breaks the symmetry of the steady solution and is responsible for the onset of the well-known Von Kármán vortex-street that becomes visible behind bluff bodies once the exponential growth of the linear instability saturates due to nonlinearities.
The three-dimensional flow configuration is a plate of height and thickness identical to the two-dimensional plate, but of finite length $L$ in the spanwise direction $z$, as sketched in fig. 1b. The computational domain is discretized using Gmsh [32] by a Delaunay mesh composed of 17 million tetrahedra, which are then partitioned between processes with ParMETIS [42]. Using Taylor–Hood finite element pair, cf. section 2.2, this leads to a total of 75 million unknowns. The boundary conditions are similar to those detailed above for the two-dimensional configuration.

The linear stability analysis of this flow configuration has been performed by [51] who determined the neutral curves of various unstable eigenmodes in the range of Reynolds number $40 \leq \text{Re} \leq 200$ and length $1 \leq L \leq 6$. Here, we specifically investigate the plate of length $L = 2.5$ for the Reynolds number $\text{Re} = 100$. The steady solution, depicted in fig. 3a, exhibits a large three-dimensional recirculation region in the wake of the plate. The stability analysis performed in [51] revealed that two pairs of complex eigenvalues get unstable above $\text{Re} \approx 101$ for this parameter choice, with respective angular frequencies of $\omega \approx 0.3$ and $\omega \approx 0.57$. Hereinafter, we focus on the high-frequency eigenmode, depicted in fig. 3b. As shown in the figure, the three-dimensional eigenmode breaks the top/bottom symmetry of the steady-state solution, as for the two-dimensional plate.
Figure 3: Linear stability analysis for the three-dimensional flow around of plate of spanwise length $L = 2.5$ and $Re = 100$. Streamwise velocity contours of (a) the steady solution and (b) the high-frequency unstable eigenmode ($\omega = 0.57$) are presented.

5.2 Influence of numerical and physical parameters

We investigate in this section the influence of various numerical and physical parameters on the performance of the mAL preconditioner. Tests are performed on the two-dimensional flow configuration previously introduced. The effect of the augmentation parameter on the preconditioner efficiency and solution accuracy is reported in section 5.2.1. The performance of the preconditioner is tested in section 5.2.2 for many values of the complex shift parameter used in the shift-and-invert strategy to compute the leading eigenvalues. The behavior of the preconditioner in regards to the mesh refinement and the Reynolds number is finally tested in section 5.2.3.

In all the numerical tests performed in this section, the full GMRES without restart is used in order to fairly assess the performance of the preconditioner. The diagonal blocks defined by the mAL preconditioner eq. (15) are here inverted using the sparse direct solver MUMPS.

5.2.1 Effect of the augmentation parameter

The effect of the augmentation parameter $\gamma$ on the performance of the mAL preconditioner is first assessed by considering the number of GMRES iterations. Using the Newton method described in section 2.3, steady solutions are computed for the Reynolds number $Re = 40$ and several values of the augmentation parameters reported in the first columns of table 1a and table 1b, that correspond to results obtained with a coarse mesh (14,674 triangles) and a finer mesh (1.29 · 10⁵ triangles), respectively. The GMRES relative tolerance being fixed to $10^{-6}$, the average numbers of inner (GMRES) iterations per outer (Newton) iteration are reported in the second columns of those tables. For both meshes, there exists an optimal value of the augmentation parameter, $\gamma \approx 1$, for which a minimum number of iterations is reached. Similar observations are reported in other studies [12, 13, 35, 14, 33] for different flow configurations such as the lid-driven cavity flow, the backward facing step or the flow over a flat plate. Note also that the number of iterations is quite similar for the coarse and fine meshes, regardless of the augmentation parameter value. For the optimal $\gamma$, the
average number of inner iterations is around 50.

As briefly discussed in section 2.2, the introduction of the grad–div stabilization term in the weak formulation does not modify the conservation of momentum at the continuous level, since the continuous solution is divergence-free. However, with the spatial discretization chosen in the present study (Taylor–Hood finite element), the divergence of the velocity is only weakly satisfied and the grad–div stabilization term modifies the discrete momentum equation. The augmentation parameter has therefore an influence on the accuracy of the discrete solution. To assess this effect, a reference steady solution, denoted \( (\mathbf{u}_b^*, p_b^*) \), is computed without stabilization parameter \( \gamma = 0 \) on a very-fine mesh made of \( 5.13\cdot10^9 \) triangles. The corresponding leading eigenvalue denoted \( \sigma^r \) is also computed. The two last columns of table 1a and table 1b report the relative errors of the steady velocity and the leading eigenvalue computed with the coarse and fine meshes, respectively, for several values of \( \gamma \). Examining first the results obtained with the coarse mesh (see table 1a), a minimal error is obtained for \( \gamma \simeq 1 \), not only for the steady solution but also for the leading eigenvalue. When the mesh is refined (see table 1b), a minimal error is still obtained for \( \gamma \simeq 1 \), although less pronounced. Compared to results obtained with the coarse mesh, the relative error is decreased whatever the value of the augmentation parameter. As expected, the augmentation parameter less affects the accuracy of the discrete solution when the mesh is refined, since the discrete solution tends towards the continuous solution. It is worth noticing that the use of the stabilization term can significantly improve the accuracy of the solution. For instance, the accuracy of the eigenvalue obtained for the coarse mesh with \( \gamma = 1 \) is identical to the one obtained for the fine mesh without stabilization \( \gamma = 0 \). In other words, the same accuracy is obtained but with ten times fewer mesh elements.

The present results clearly indicate that \( \gamma \simeq 1 \) is an optimal value from both the solution accuracy point of view and the preconditioning efficiency point of view when considering not only steady solutions, as reported before [35], but also leading eigenvalues. As a consequence, in the following, we consider that \( \gamma \) can be chosen on preconditioning efficiency criteria only without compromising accuracy.

| \( \gamma \) | \# of GMRES iterations | \( \frac{\|\mathbf{u}_b^h - \mathbf{u}_b^*\|_2}{\|\mathbf{u}_b^*\|_2} \) | \( \frac{|\sigma^h - \sigma^r|}{|\sigma^r|} \) |
|---|---|---|---|
| 0 | 860 | 2.8 \cdot 10^{-4} | 9.8 \cdot 10^{-4} |
| 10^{-1} | 191 | 2.1 \cdot 10^{-4} | 7.9 \cdot 10^{-4} |
| 10^0 | 52 | 8.4 \cdot 10^{-5} | 1.1 \cdot 10^{-4} |
| 10^1 | 337 | 4.4 \cdot 10^{-4} | 1.5 \cdot 10^{-3} |
| 10^2 | >1,000 | 1.3 \cdot 10^{-3} | 4.6 \cdot 10^{-3} |

(a) Coarse mesh (14,674 triangles)

| \# of GMRES iterations | \( \frac{\|\mathbf{u}_b^h - \mathbf{u}_b^*\|_2}{\|\mathbf{u}_b^*\|_2} \) | \( \frac{|\sigma^h - \sigma^r|}{|\sigma^r|} \) |
|---|---|---|
| 873 | 3.2 \cdot 10^{-5} | 1.1 \cdot 10^{-4} |
| 194 | 2.7 \cdot 10^{-5} | 1.1 \cdot 10^{-4} |
| 53 | 1.4 \cdot 10^{-5} | 5.4 \cdot 10^{-5} |
| 363 | 2.1 \cdot 10^{-5} | 8.5 \cdot 10^{-5} |
| >1,000 | 6 \cdot 10^{-5} | 2.1 \cdot 10^{-4} |

(b) Fine mesh (1.29 \cdot 10^5 triangles)

Table 1: Effect of the grad–div augmentation parameter \( \gamma \) on the mAL preconditioning efficiency and the solution accuracy. For both tables, the first column indicates values of \( \gamma \). The second column represents the average number of mAL preconditioned GMRES iterations per Newton iteration. The last two columns give the relative errors of the steady solution and the leading eigenvalue compared with a reference solution \( (\mathbf{u}_b^*, \sigma^r) \) computed on a very fine mesh without stabilization (\( \gamma = 0 \)).
5.2.2 Effect of the shift parameter

The shift-and-invert strategy, adopted in the present study to compute the eigenvalues with largest real part, requires to specify the complex value $s = s_r + i s_i$ that appears in the spectral transformation section 2.4. When investigating the transition of a steady solution from a stable to an unstable state, a common practice is to choose the shift parameter as a pure imaginary number, i.e. $s = i s_i$, and to vary the imaginary part in order to compute complex eigenvalues with growth rates close to $\lambda = 0$. Depending on the flow configuration investigated, the steady solution may get unstable for eigenmodes characterized by very different frequencies. Ideally, the number of preconditioned GMRES iterations should be insensitive to the value of the complex shift, for the Krylov–Schur algorithm to converge rapidly whatever the eigenvalue of interest. To the best of our knowledge, only the case of a real-valued shift has been considered so far, either positive when solving the unsteady Oseen problem [12, 35] or negative when solving the linearized Navier–Stokes equation [56].

Here, we vary $s$ in the whole complex plane and assess its effect on the efficiency of the mAL preconditioner by performing the following numerical experiment. The linear system eq. (12) is solved with right-hand side vectors whose coefficients are randomly generated in $[0, 1] + [0, 1]i$, as done for instance in [9]. The Jacobian matrix $J$ of the linear system is computed with the steady solution at $Re = 40$ and the augmentation parameter $\gamma = 0.7$. In other words, only the inner solver is studied, no outer iteration (Newton or Krylov–Schur) is performed. The isocontours shown in fig. 4 in the complex plane $(s_r, s_i)$ correspond to the number of inner (GMRES) iterations required to decrease the relative residual to $10^{-6}$. The red circles are eigenvalues of the Jacobian matrix. First, the number of iterations increases when the shift gets closer to any eigenvalue. In that case, the matrix $J + s M$ involved in the spectral transformation (2.4) becomes singular, leading to a very ill-conditioned linear system and thus high iteration counts. Second, the number of iterations is reduced when increasing $s_r$. Solving the linear system for $s_r < 0$ is generally more expensive than for $s_r > 0$. According to [56], this is due to the indefiniteness of the velocity block $A_\gamma + s M_u$ in eq. (12) for sufficiently large negative values of $s_r$. On the contrary, the velocity block is definite when $s_r > 0$. The contribution of the shift can be interpreted as a (positive definite) time-step term, which reinforces the diagonal dominance of the problem and thus improve its spectral properties. For more details, the reader can refer to [12, section 2.6], where the mAL preconditioner is used to solve the unsteady Oseen problem. Finally, no particular trend is observed in the number of inner iterations when fixing the real part $s_r$ and varying the imaginary part $s_i$, except when getting closer to an eigenvalue. For $s_r = 0$, the number of iterations is roughly constant for $s_i < 0.5$ and increases around $s_i = 0.6$ due to the proximity of the unstable eigenvalue marked by the red circle. By further increasing $s_i$, the number of iterations then decreases.

5.2.3 Effect of the mesh refinement and Reynolds number

The modified augmented Lagrangian preconditioner allows to compute steady solution in a number of GMRES iterations independent of the mesh refinement, as previously observed in table 1 and mildly dependent of the Reynolds number [12, 11]. The influence of the mesh refinement and Reynolds number on the number of iterations needed to solve the complex-shifted linear system eq. (12) has not been investigated so far. The numerical
Figure 4: Influence of the complex shift. Isocontours represent the number of GMRES iterations needed to solve the linear system (13) with the mAL preconditioner, depicted in the complex plane \((s_r, s_i)\), and computed for \(Re = 40\) and \(\gamma = 0.7\). The red circles are the eigenvalues of the Jacobian matrix \([6]\).

**Experiment:**

The experiment consists in solving the linear system to the relative tolerance of \(10^{-6}\) for right-hand side vectors with randomly generated coefficients as explained before. First, the Reynolds number is fixed \((Re = 40)\) as well as the augmentation parameter \((\gamma = 0.7)\) while the mesh refinement changes. The number of inner GMRES iterations is reported in fig. 5a as a function of the number of triangles. The curves correspond to different values of the (purely imaginary) shift. Clearly, the iteration count is independent of the mesh refinement, regardless of the shift. Second, a fixed mesh refinement is chosen \((14,828\) triangles) and the linear system is solved for different values of the Reynolds number in the range \([10; 500]\) for different shift values. As reported in \([13]\) when computing steady solutions, the optimal value of the augmentation parameter that minimizes the number of iterations depends on the Reynolds number. The optimal value of \(\gamma\) has been first determined for each value of \(Re\). For \(Re = 10\), the optimal value is \(\gamma \approx 1.2\) and it decreases to \(\gamma \approx 0.4\) for \(Re = 500\). These optimal values of \(\gamma\) have been determined for \(s = 0\) but are used in the following regardless of the values of \(s\). The number of iterations is depicted in fig. 5b as a function of the Reynolds number. The mAL preconditioner shows a mild degradation of its performance as \(Re\) increases, independently of the values of the shift. The increase of the number of inner iterations is proportional to \(Re^{0.5}\) in this numerical experiment.
As a conclusion, the modified augmented Lagrangian preconditioner exhibits interesting properties for performing efficiently a linear stability analysis using a shift-and-invert strategy: robustness with respect to a complex-valued shift, mesh independence, and a mild deterioration as $Re$ increases.

5.3 Comparison with other block preconditioners

The mAL preconditioner is one of many other preconditioners developed to solve the steady incompressible Navier–Stokes equations. Among them, we select the Pressure Convection–Diffusion (PCD) [43] and the SIMPLE [58] preconditioners, widely used and easily implemented, and compare their performance with those of the mAL preconditioner. In addition, we also test the unsteady Stokes preconditioner [70] which has gained in popularity in the hydrodynamic stability community [74], as it can be easily implemented using existing time-steppers so as to compute base flows and leading eigenvalues [72]. In our implementation however, the Stokes preconditioner is itself applied using a nested Krylov subspace method instead of an existing time-stepper. More details on those preconditioners are given in section B but it is worth recalling here that they are designed for classical Galerkin discretization of the Navier–Stokes equations. Therefore, their application to the iterative solution of eq. (12) is meant for $\gamma = 0$, i.e. without grad–div stabilization terms.

The numerical test case consists in solving iteratively eq. (12) with a random right-hand side vector, as detailed before, for a large relative tolerance equal to $10^{-3}$ to limit the number of iterations. The shift is fixed to $s = 0$ since it was shown in section 5.2.3 that the mAL preconditioner depends very weakly on the shift parameter when the mesh is refined or the Reynolds number is increased. All innermost block solutions are performed using exact LU factorizations. The PCD and SIMPLE preconditioners are parameter-free, unlike the mAL and Stokes preconditioners. For the latter, the optimal values of the parameter ($\gamma$ for mAL and a time-step like parameter for Stokes) are determined for each values of the Reynolds number.
The effect of the mesh refinement and Reynolds number on the number of inner iterations, studied in the previous paragraph for the mAL preconditioner, is assessed here for all the other preconditioners. Results are compared in fig. 6. This number of iterations is a good measure to compare the efficiency of the different preconditioners, in a first approximation, because for each inner iteration, the application of all preconditioners requires the solution of subproblems with similar complexities. Therefore, the computational time of one inner iteration is roughly similar for all preconditioners.

All the preconditioners are independent of the mesh refinement, as shown in fig. 6a, except for the SIMPLE preconditioner for which the number of inner iterations slightly increases with the number of triangles. Interestingly, the number of iterations is significantly less for mAL and PCD (around 50) than for Stokes and SIMPLE (around 1,000). Note that for the two preconditioners depending on a parameter (mAL and Stokes), their optimal value was found to be independent of the mesh refinement.

The effect of the Reynolds number is reported in fig. 6b. For all tested preconditioners, the number of iterations increases with the Reynolds number, but with different slopes. The mAL preconditioner exhibits the best performance for all Reynolds numbers, except for low Reynolds number (\( Re < 20 \)) where PCD is more efficient. However, the number of iterations obtained with the PCD preconditioner increases strongly for larger values of the Reynolds number (\( Re > 80 \)). The mAL and SIMPLE preconditioners exhibit a similar trend: the number of iterations scales with the Reynolds number as \( Re^{0.5} \). However, it is significantly larger with SIMPLE than with mAL, regardless of the Reynolds number. At low \( Re \), the Stokes preconditioner behaves similarly to the SIMPLE preconditioner, but for larger \( Re \), it degrades significantly and exhibits the same trend as the PCD preconditioner. Finally, when considering the number of iterations, the mAL preconditioner is undoubtedly the best preconditioner. We note that, contrary to the mesh dependence study, the optimal parameters of mAL and Stokes showed some variations with respect to \( Re \).

\[^4\text{two scalar velocity solves and one pressure Schur complement solve for mAL and Stokes; one vectorial velocity solve and one pressure Schur complement solve for PCD and SIMPLE}\]
Let us now compare the computational time for applying the four preconditioners. To
that aim, the direct $LU$ factorizations used until now for the innermost velocity blocks
solvers are replaced by GMRES right-preconditioned with an $ILU(2)$ method (as imple-
mented in PETSc). The choice of an innermost iterative solution allows for a more com-
prehensive interpretation of the computational time, since it accounts for the various com-
plexities in solving iteratively the velocity blocks involved in the different preconditioners.
Moreover, such an innermost iterative solution is necessary when considering large-scale
three-dimensional problems, as shown in the next section. The relative tolerance of the
inner (resp. innermost) GMRES is fixed to $10^{-3}$ (resp. $10^{-2}$). The computational times ob-
tained with the four preconditioners are reported in fig. 7 for $Re = 40$ (left) and $Re = 100$
(right). The total time is split into the time spent in computing matrix–vector products,
in applying the global preconditioner, and in constructing the global Krylov subspace. The
inner iteration counts are given between parenthesis. Note that it may be slightly higher
than what is presented in fig. 6 since the velocity blocks are now solved only approximately.
All computations are run on a standard laptop computer.

For the low Reynolds number $Re = 40$, the mAL and PCD preconditioners are about
ten times faster than the SIMPLE and Stokes preconditioners. For this Reynolds number,
the PCD preconditioner is comparable with the mAL preconditioner, as it is only 40% slower.
However, when the Reynolds number is increased to $Re = 100$, the performance
of PCD degrades significantly with respect to mAL, as it is now about five times slower.
The computational times are not given for the SIMPLE and Stokes preconditioners because
they largely exceed 2,100 seconds. The deterioration in the computational time of the PCD

\[ \alpha Re \]

\[ \alpha Re^{0.5} \]

Figure 6: Influence of (a) the number of mesh elements and (b) the Reynolds number $Re$
on the number of inner iterations required to solve eq. (12) with the mAL, PCD,
SIMPLE, and Stokes preconditioners. The relative tolerance is set to
$10^{-3}$ and the shift to zero.

\footnote{Note that the time for building the inner Krylov subspace is very small for the Stokes preconditioner at $Re = 40$, despite a large number of iterations. This is due to the fact that, in the implementation detailed in section B, the Stokes preconditioner is itself applied with a nested iterative method. Therefore, no Krylov subspace of dimension 947 is actually built.}
preconditioner when the Reynolds number is increased, is in agreement with the growth in the number of iterations observed before. For even higher Reynolds numbers $Re > 100$, the mAL preconditioner is expected to be increasingly more interesting than its competitors.

As a conclusion, this benchmark shows that, compared to other widely used preconditioners, mAL provides a more efficient approach for solving eq. (12) on a configuration typical of two-dimensional external flows around bluff bodies. In particular, among the alternatives tested here, it is the only preconditioner combining a mesh-independent iteration count and a mild degradation with $Re^{0.5}$, making it the most efficient preconditioner for $Re \geq 20$.

Figure 7: Comparison of the performance of various preconditioners on the two-dimensional test case (14,828 triangles). The total time is split between matrix–vector products, applying the preconditioner, and building the Krylov subspace. The number of global GMRES iterations is given between parenthesis. System eq. (12) is solved to a relative tolerance of $10^{-3}$. For $Re = 100$, the hatched bars correspond to preconditioners for which the total time largely exceeded 2,100 seconds and is not reported in details. The velocity blocks in the preconditioners are solved iteratively to a relative tolerance of $10^{-2}$ using an innermost GMRES, right-preconditioned with $ILU(2)$. The pressure blocks are solved exactly with MUMPS. These times will depend on the particular preconditioners used for solving the diagonal blocks. Therefore, those results should be considered qualitatively.

5.4 Performance of the parallel implementation

In this section, the performance of the parallel implementation detailed in section 5 is tested on the three-dimensional configuration presented in section 5.1. First, a coarse mesh is used, in order to be able to compare our approach with the direct parallel solver MUMPS. Then, the full size 3D configuration presented before is considered to test the parallel performance of our approach on a problem that a direct solver could not handle at a reasonable memory cost.
5.4.1 Comparison with a direct solver on a small-scale 3D configuration

Despite its large memory requirements, some authors have used the “matrix-based” approach, combined with direct solvers for the arising linear systems, to perform the stability analysis of three-dimensional flows \[51, 38\]. In this section, we aim at comparing the performance of this approach to ours. To that end, the three-dimensional test case is considered using a coarse mesh of 1.1 million tetrahedra (4.8 million unknowns), in order to keep the memory consumption of the direct solver reasonable. The computations are performed on Sator, an ONERA cluster composed of 620 nodes with two fourteen-core Intel Broadwell clocked at 2.4 GHz. The direct solver we compare ourselves to is MUMPS.

5.4.1.1 Nonlinear solver

In this section, the Newton nonlinear tolerance and the GMRES relative tolerance are both set to $10^{-6}$. In fig. 8a we report the average wall-clock time per Newton iteration for the mAL and MUMPS approaches, as a function of the number of processes. On the top $x$-axis, we report the amount of available memory corresponding to each number of processes. First, as expected, the memory requirements of our approach are lower than with MUMPS: we observe that MUMPS cannot be run on less than 224 processes, which corresponds to an available memory of 1,024 GB, whereas the mAL approach can be run on 28 processes (128 GB). Note that the memory requirements of the mAL approach could be even lower by using iterative methods for the subdomain solvers of the innermost ASM-preconditioned GMRES iterations. Moreover, thanks to good scalability properties and the absence of a full $LU$ factorization at each Newton iteration, the mAL approach is clearly faster than MUMPS (about ten times with 448 processes).

5.4.1.2 Eigensolver

For the eigen solver, the Krylov–Schur tolerance is $10^{-6}$ whereas the inner relative tolerance is $10^{-3}$. Note that we use a larger tolerance for the inner linear solve than for the outer Krylov–Schur solver. Indeed, contrary to what is often recommended in the literature (e.g., \[59 \S 3.4.1\]), we observed that it was not necessary to use a smaller tolerance for the inner solution of eq. (12) in order to keep a satisfying accuracy on the computed eigenvalues. More details on that aspect may be found in section C.

We show in fig. 8b the total wall-clock time for computing 5 eigenvalues closest to the shift $s = 0.6i$, using mAL and MUMPS as inner solvers, as a function of the number of processes. The available memory is again reported on the top $x$-axis. Similar conclusions as for the nonlinear solver can be made for memory consumption with a multiplication factor of two, due to the use complex instead of real algebra. From a wall-clock time point of view, we observe, as for the nonlinear solver, that the mAL approach possesses much better scalability than MUMPS, which leads to a faster computation. We note however that MUMPS is harder to beat with an iterative approach when used in the eigensolver than in the nonlinear solver. The reason is that, in the Krylov–Schur method, the very high cost of forming the full $LU$ factorization is greatly amortized by the many inner solves realized with it, whereas in the Newton method, each inner solve requires to build the factorization again.
As a conclusion, the mAL approach presents the double advantage of being much less memory-intensive than the direct solver and also faster, even for the unfavorable case of eigenvalue computations.

![Available memory (GB)](image_url)

(a) Nonlinear solver (average time per Newton (b) Eigensolver (overall time for 5 eigenvalues, iteration)

Figure 8: Comparison of mAL (——) and MUMPS (—) as inner solvers on (a) the nonlinear solver and (b) the eigensolver. The computation are performed on the three-dimensional test case using a coarse mesh (4.8 million unknowns) and $\gamma = 0.3$. The grey area indicates when the $LU$ factorization is not feasible due to too large memory requirements. The dashed line represents ideal scalability.

5.4.2 Parallel performance on a large-scale 3D configuration

In this section, the 3D plate configuration is used, with a fine mesh, resulting in 75 million unknowns. The parallel performance of our implementation is investigated for the nonlinear base flow solver and eigenvalue solver.

Results were obtained on Curie, a system composed of 5,040 nodes with two eight-core Intel Sandy Bridge clocked at 2.7 GHz. The interconnect is an InfiniBand QDR full fat tree and the MPI implementation exploited was bullxMPI version 1.2.9.2. All binaries and shared libraries were compiled with Intel compilers and Math Kernel Library support (for dense linear algebra computations) version 18.0.1.163. Recent releases of FreeFem++ and PETSc/SLEPc were used (version 3.61 and 3.9.3 respectively). In all following plots and tables, the time spent in the finite element kernel is never accounted for because we are mostly interested in the performance of the preconditioner. Only the time spent in PETSc or SLEPc is reported.

5.4.2.1 Nonlinear solver

In this paragraph, we investigate the parallel performance of the nonlinear steady-state solver. The inner Krylov solver is the flexible GMRES algorithm [62], which is stopped when the relative unpreconditioned residual norm is lower than $10^{-1}$. The Newton method
is stopped when the $l^2$ norm of the residual is lower than $10^{-6}$. As an initial guess for the computation at $Re = 100$, a solution at a lower Reynolds number $Re = 50$ is first computed using a higher nonlinear outer tolerance of $10^{-4}$. Also note that in this preprocessing step, all the domain decomposition information obtained from ParMETIS partitioning is dumped and will be used in successive runs for the nonlinear and generalized eigenvalue solvers. In table 2 the numerical performance of the nonlinear solver are reported. One may notice that even if a high relative tolerance is used to stop the flexible GMRES, very few Newton iterations (second column) are needed for the solver to converge, independently of the number of subdomains (first column). The number of mAL-preconditioned inner iterations needed to reach convergence, averaged over all Newton iterations, is reported in the third column. It is seen not to depend on the number of processes. Eventually, in the last three columns, we show the average number of ASM-preconditioned innermost iterations needed for each velocity block of $P_{mAL}$ to reach the desired convergence tolerance of $10^{-1}$ (see section 4.3). A slight increase is observed with the number of processes. This is an expected feature of simple one-level domain decomposition methods, like the additive Schwarz method, that are known to not scale numerically [27].

<table>
<thead>
<tr>
<th>$P$</th>
<th>Setup (s)</th>
<th>Solve (s)</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>256</td>
<td>364.6</td>
<td>5,739.9</td>
<td>–</td>
</tr>
<tr>
<td>512</td>
<td>88.9</td>
<td>2,994.9</td>
<td>2.0</td>
</tr>
<tr>
<td>1,024</td>
<td>56.6</td>
<td>1,523.6</td>
<td>3.9</td>
</tr>
<tr>
<td>2,048</td>
<td>13.9</td>
<td>895.1</td>
<td>6.7</td>
</tr>
</tbody>
</table>

Table 2: Numerical performance of the 3D nonlinear solver with respect to the number of processes ($Re = 100$). The second column represents the number of Newton outer iterations, the third is the number of mAL-preconditioned inner GMRES iterations per Newton step. The last three columns correspond to the average number of ASM-preconditioned innermost GMRES iterations for each velocity block per inner iteration.

In fig. 9 the scalability of our implementation is shown, using the run with 256 processes.
as the reference, and going up to 2,048 processes. The parallel efficiency of this approach remains above 83%. The fact that one additional Newton iteration is needed with 256 processes has to be highlighted, since it does improve the efficiency. Other than that, because exact LU factorizations are used as subdomain solvers in the additive Schwarz method used for each velocity field, the setup phase scales superlinearly (see the second column of the table in fig. 9). Moreover, because the number of iterations needed for the corresponding solvers only grows slightly, as shown in the three last columns from table 2, the solution phase also scales appropriately.

5.4.2.2 Eigensolver

We now evaluate the parallel performance of the eigensolver. The tolerance on the Krylov–Schur algorithm is $10^{-6}$ whereas the relative tolerance for the inner linear solver is set to $10^{-4}$. The main difference with the Newton method is that the multiple inner linear solves involve the very same shifted operator $(J + sM)$ and preconditioner $P_{mAL}$. To improve the performance of the eigensolver, let us show first the effect of using a recycled Krylov method for solving these systems, instead of the standard GMRES. This can be done by switching from the KSP objects of PETSc to the iterative methods of the HPDDM library [40] which handle subspace recycling. In particular, the flexible GCRO-DR method is used, with a recycled subspace between each linear solves of dimension five. The following lines allow to switch between PETSc and HPDDM Krylov methods from within FreeFem++:

```cpp
int recycle = getARGV("-recycle", 0); // use GMRES by default
int restart = getARGV("-restart", 200); // default to 200
real tolInner = getARGV("-inner_tol", 1.0e-4); // default to 10^{-4}
string paramsKrylov = (recycle == 0 ? "-st_ksp_type fgmres " +
    "-st_ksp_monitor -st_ksp_rtol " + tolInner +
    " -st_ksp_gmres_restart " + restart + " -st_ksp_max_it 1000"
    :
    "-st_ksp_type hpddm -hpddm_st_krylov_method gcrodr " +
    "-hpddm_st_recycle " + recycle + " -hpddm_st_max_it 1000" +
    " -hpddm_st_verbosity 4 -hpddm_st_gmres_restart " + restart +
    " -hpddm_st_tol " + tolInner + " -hpddm_st_variant flexible");
```

In fig. 10, the number of mAL-preconditioned inner linear iterations needed for each iterative method (FGMRES or FGCRO-DR) to solve the sequence of linear systems of the first iteration of the Krylov–Schur algorithm is reported. For this particular Krylov–Schur iteration, fifteen systems have to be solved. When using FGMRES, it corresponds to a total of 3,751 inner linear iterations. When using FGCRO-DR, it corresponds to only 2,209 inner linear iterations. Even though the solutions of all fifteen systems are not rigorously equal when switching from FGMRES to FGCRO-DR, after the first iteration of the eigensolver, convergence is reached for the two eigenpairs closest to the shift: $-1.03 \cdot 10^{-2} + 0.57i$ and $-7.81 \cdot 10^{-2} + 0.57i$.

In all the following runs, FGCRO-DR is used in order to reduce the number of inner iterations. The number of Krylov–Schur iterations needed to retrieve the requested eigenpairs is reported in the second column of table 3. In the third column is the average number of

---

6for the same reasons explained in 5.4.1.2 we use a larger tolerance for the inner linear solves than for the outer Krylov–Schur algorithm (see also section C)
solved linear systems per eigensolver iteration. The average number of mAL-preconditioned FGCRO-DR iterations (inner iterations) per linear solve (outer iteration) is presented in the fourth column whereas the last three columns contain the average number of innermost iterations for each velocity field, per application of $P_{mAL}$. It was not possible to have the code run on 256 processes due to memory requirements significantly higher than for the nonlinear solver. Indeed, we switch from a real-valued to a complex-valued problem and the additional operators $M$ and $L_p$ are assembled explicitly. In table 4, the scalability of our implementation is shown, using the run with 512 processes as the reference, and going up to 2,048 processes. The parallel efficiency of this approach is approximately the same as for the nonlinear solver, though on a narrower range of process counts, remaining above 82%.

<table>
<thead>
<tr>
<th>$P$</th>
<th># of eigensolver iterations</th>
<th># of linear solves</th>
<th># of iterations per linear solve</th>
<th># of iterations per field ($x, y, z$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>7</td>
<td>7</td>
<td>120</td>
<td>7 10 11</td>
</tr>
<tr>
<td>1,024</td>
<td>7</td>
<td>8</td>
<td>127</td>
<td>7 10 13</td>
</tr>
<tr>
<td>2,048</td>
<td>7</td>
<td>8</td>
<td>119</td>
<td>10 13 17</td>
</tr>
</tbody>
</table>

Table 3: Numerical performance of the 3D eigensolver with respect to the number of processes. The second and third columns represent respectively the number of Krylov–Schur iterations and the number of linear solves (outer iterations) per Krylov–Schur iteration. The fourth column is the number of mAL-preconditioned FGCRO-DR inner iterations per linear solve. The last three columns correspond to the average number of innermost ASM-preconditioned GMRES iterations for each velocity sub-block per inner iteration.
Table 4: Scalability of the 3D eigensolver with respect to the number of processes.

<table>
<thead>
<tr>
<th></th>
<th>Setup (s)</th>
<th>Solve (s)</th>
<th>Speedup</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>55.3</td>
<td>39,160.7</td>
<td>—</td>
</tr>
<tr>
<td>1,024</td>
<td>25.7</td>
<td>24,508.3</td>
<td>1.6</td>
</tr>
<tr>
<td>2,048</td>
<td>27.1</td>
<td>11,849.9</td>
<td>3.3</td>
</tr>
</tbody>
</table>

6 Conclusion

The stationary base flow as well as the eigenvalue computations involved in hydrodynamic linear stability analysis require multiple solutions of linear systems based on the (shifted) Jacobian operator of the incompressible steady Navier–Stokes equations. To solve such systems on large-scale configurations involving hundreds of millions of unknowns, we proposed to use a Krylov subspace linear solver like the flexible GMRES algorithm, preconditioned by the modified augmented Lagrangian (mAL) preconditioner [12]. On a two-dimensional bluff-body flow, we studied numerically the performance of the mAL preconditioner for linear stability analysis purposes. We showed in particular that this approach handles efficiently complex-valued shifts and thus is well-suited for the computations of eigenvalues with possibly large frequencies, using the shift-and-invert spectral transformation. Then, the mAL preconditioner was tested against some other widely used steady-state (PCD, SIMPLE) and time-stepping-based (Stokes) preconditioners, all of them used in a sequential version. The mAL preconditioner was shown to require lower numbers of GMRES iterations and to be faster than all its competitors.

In order to perform large-scale three-dimensional stability analysis computations, a parallel implementation of the mAL preconditioner was presented and is made available online: [https://github.com/prj-/moulin2019al](https://github.com/prj-/moulin2019al). The FreeFem++ finite element language was used as a discretization kernel whereas PETSc and SLEPc were used as distributed linear algebra backends. First, a comparison with the parallel direct solver MUMPS was presented on a three-dimensional bluff-body flow configuration, using a coarse enough mesh to make the LU factorization possible. The mAL approach required about one tenth as much memory and had better strong scaling properties. Despite the attractiveness of a direct linear solver — when it can be afforded — for the eigenvalue computations (the factorization is done once and re-used multiple times), the mAL approach turned out to be a faster alternative than the direct approach, thanks to its much better parallel performance. Finally, the implementation was used on a fine mesh resulting in 75 million unknowns and showed satisfying strong scaling properties up to 2,048 processes, for both the base flow and eigenvalue computations. The role of subspace recycling between the multiple consecutive linear solves, inside the Krylov–Schur eigensolver, was tested and allowed significant performance gains.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement number
Moreover, this work was granted access to the HPC resources of TGCC@CEA under the allocation A0030607519 made by GENCI.

Appendix A  Reproducibility

In addition to the few extracts of the code used in the paper, the interested reader can find the complete FreeFem++ code in the following repository: https://github.com/prj-/moulin2019al.

Appendix B  Definition of other block preconditioners

Here are defined the classical block preconditioners for incompressible Navier–Stokes that are compared to the modified augmented Lagrangian approach in section 5.3. Contrary to mAL, those preconditioners do not require an augmentation. Thus, they are used without grad–div stabilization ($\gamma = 0$). Versions that incorporate a complex shift $s$ are proposed here.

**Pressure convection–diffusion preconditioner**  The pressure convection–diffusion (PCD) preconditioner was proposed by [13]:

$$
P_{\text{PCD}} = \begin{pmatrix} A + sM_u & B^T \\ 0 & S_p \end{pmatrix},
$$

(17)

with $S_p^{-1} = -M_p^{-1}(F_p + sM_p)L_p^{-1}$, where $F_p$ is a convection–diffusion operator built on the pressure space. Compared to the classical PCD preconditioner for steady-state Navier–Stokes equations, the shift contribution $sM_p$ is added to the pressure Schur complement approximation.

**SIMPLE preconditioner**  The SIMPLE preconditioner was proposed as a solver by [58]. We use its preconditioner version [30]:

$$
P_{\text{SIMPLE}} = \begin{pmatrix} A + sM_u & 0 \\ B & S_p \end{pmatrix} \left( \begin{array}{c} I \\ 0 \end{array} \right) \left( \begin{array}{c} \text{diag}(A + sM_u)^{-1}B^T \\ I \end{array} \right),
$$

(18)

and $S_p^{-1} = -[B\text{diag}(A + sM_u)^{-1}B^T]^{-1}$.

**Stokes preconditioner**  The Stokes preconditioning approach was popularized by [73] in the hydrodynamic stability community. Tuckerman’s idea is two-fold:

1. preconditioning the linearized Navier–Stokes problem eq. (12) by the Stokes problem, i.e.,

$$
P_{\text{Stokes}} = \begin{pmatrix} A_{\text{Stokes}} + sM_u & B^T \\ B & 0 \end{pmatrix},
$$
with \( A_{\text{Stokes}} = D + \Delta t^{-1}M_u \) and \( D \) contains only the diffusion terms. Note that the time-step contribution \( \Delta t \) has no physical meaning here: it only represents some numerical parameter of the preconditioner. A case-dependent optimal value may exist, as reported in [8]. We determined this optimal value numerically.

2. applying the preconditioner by adapting a pre-existing time-stepping code, which significantly reduces the development costs. In this work however, we prefer to apply \( P_{\text{Stokes}}^{-1} \) by using a few inner iterations of GMRES, preconditioned by:

\[
P_{\text{Stokes, inner}} = \begin{pmatrix} A_{\text{Stokes}} + sM_u & B^T \\ 0 & S_p^{-1} \end{pmatrix},
\]

with \( S_p^{-1} = -Re^{-1}M_p^{-1} - (\Delta t + s)L_p^{-1} \) [18]. A large relative tolerance of \( 10^{-2} \) is set for the inner iterations, as we observed that further convergence of the inner iterations did not improve the convergence of the outer GMRES iterations. Obviously, \( A_{\text{Stokes}} \) being block diagonal, applying \( P_{\text{Stokes, inner}}^{-1} \) naturally requires two scalar velocity solves (in 2D) and one pressure solve.

Note that, in section 5.3, the GMRES iteration count reported for the Stokes preconditioner corresponds to the total number of applications of \( P_{\text{Stokes, inner}}^{-1} \) necessary to converge eq. (12) to the desired tolerance.

Appendix C Linear solver tolerance and eigenvalue convergence criterion

In cases where an iterative linear solver is used, the action of \((J + sM)^{-1}\), required when applying the spectrally transformed operator \( T \) in the Krylov–Schur algorithm (see section 2.4), is approximated using some user-defined tolerance. As a consequence, matrix \( T \) in eq. (11) is replaced by some approximation \( \tilde{T} \). It is usually recommended to set the tolerance for the linear solver lower than the one prescribed to the eigensolver, so that the imprecision of the linear solver does not pollute the eigensolver accuracy (see e.g. [59] § 3.4.1). Here, we re-evaluate this statement numerically on the two-dimensional test case presented in section 5.1.

The following numerical experiment is performed. We solve the eigenproblem using a mAL-preconditioned GMRES algorithm to apply \((J + sM)^{-1}\). The relative tolerance of the GMRES algorithm \( \varepsilon_{\text{lin}} \) is varied between \( 10^{-8} \) and \( 10^{-1} \) while the tolerance of the Krylov–Schur algorithm \( \varepsilon_{\text{eig}} \) is kept constant to \( 10^{-6} \). Only one eigenvalue, closest to the shift \( s = 0.7i \), is demanded. In the second column of table 5 the number of GMRES iterations required to apply \((J + sM)^{-1}\) is shown. The value is averaged over all applications of \((J + sM)^{-1}\) to compute the demanded eigenvalue. In the third column, the total number of applications of \( \tilde{T} \) is shown. In the last three columns, we monitor the eigenvalue and the discrete \( l^2 \) norm of the eigenproblem residual. It is observed that one can in practice increase \( \varepsilon_{\text{lin}} \) well above \( \varepsilon_{\text{eig}} = 10^{-6} \), without compromising significantly the accuracy of the computed eigenvalue. At least up to \( \varepsilon_{\text{lin}} = 10^{-3} \), the computed eigenvalue is converged to satisfying accuracy, for a cost divided by two with respect to the “safe choice” \( \varepsilon_{\text{lin}} = 10^{-6} \). Increasing \( \varepsilon_{\text{lin}} \) may thus allow some performance improvement.
Finally, the interpretation of the last column of table 5 deserves some further explanation. Indeed, one can observe that, for $\varepsilon_{\text{lin}} = 10^{-1}$ and $\varepsilon_{\text{lin}} = 10^{-2}$, despite the fact that we kept $\varepsilon_{\text{eig}} = 10^{-6}$, the Krylov–Schur algorithm considered it had converged to an appropriate eigenvalue, while the residual was still above $\varepsilon_{\text{eig}}$. The reason to that observation is that most Arnoldi-based eigensolver packages, such as SLEPc, use convergence criteria based on the residual of the transformed problem eq. (11), not the original one eq. (6). As a consequence, the effect of using a large linear tolerance $\varepsilon_{\text{lin}}$ is to apply an approximate operator $\tilde{T}$ far from the exact one. But the eigenvalues of $\tilde{T}$ can be computed to any prescribed tolerance by the Krylov–Schur algorithm. Note that, in SLEPc, a workaround is to use the option -eps_true_residual which forces the computation of the residual on the original problem eq. (6) and thus is free from any approximation linked to the underlying linear solver. In this case, the effect of using a large linear tolerance $\varepsilon_{\text{lin}}$ would be to make the convergence of the Krylov–Schur algorithm increasingly slow (or even impossible). This option being more costly, it should however be avoided for large-scale computations.

<table>
<thead>
<tr>
<th>$\varepsilon_{\text{lin}}$</th>
<th># of GMRES iterations per linear solve</th>
<th># of linear solves</th>
<th>$\lambda$</th>
<th>$\omega$</th>
<th>$|\sigma\mathbf{Mq} + J(\mathbf{q}_b)\mathbf{q}|_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-1}$</td>
<td>15</td>
<td>12</td>
<td>$4.48 \cdot 10^{-2}$</td>
<td>$6.90 \cdot 10^{-1}$</td>
<td>$1.5 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>28</td>
<td>10</td>
<td>$3.77 \cdot 10^{-2}$</td>
<td>$7.01 \cdot 10^{-1}$</td>
<td>$2.6 \cdot 10^{-6}$</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>36</td>
<td>8</td>
<td>$3.75 \cdot 10^{-2}$</td>
<td>$7.01 \cdot 10^{-1}$</td>
<td>$9.5 \cdot 10^{-7}$</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>45</td>
<td>8</td>
<td>$3.75 \cdot 10^{-2}$</td>
<td>$7.01 \cdot 10^{-1}$</td>
<td>$9.2 \cdot 10^{-7}$</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>53</td>
<td>8</td>
<td>$3.75 \cdot 10^{-2}$</td>
<td>$7.01 \cdot 10^{-1}$</td>
<td>$9.2 \cdot 10^{-7}$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>62</td>
<td>8</td>
<td>$3.75 \cdot 10^{-2}$</td>
<td>$7.01 \cdot 10^{-1}$</td>
<td>$9.2 \cdot 10^{-7}$</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>71</td>
<td>8</td>
<td>$3.75 \cdot 10^{-2}$</td>
<td>$7.01 \cdot 10^{-1}$</td>
<td>$9.2 \cdot 10^{-7}$</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>80</td>
<td>8</td>
<td>$3.75 \cdot 10^{-2}$</td>
<td>$7.01 \cdot 10^{-1}$</td>
<td>$9.2 \cdot 10^{-7}$</td>
</tr>
</tbody>
</table>

Table 5: Effect of linear solver tolerance on eigenvalue computation ($Re = 40$, $\gamma = 0.7$, $\varepsilon_{\text{eig}} = 10^{-6}$, $s = 0.7i$)
References


