Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2016-0070

Vol. **22**, No. 5, pp. 1508-1532
November 2017

Computing Optimal Forcing Using Laplace Preconditioning

M. Brynjell-Rahkola[1,*], L. S. Tuckerman[2], P. Schlatter[1] and D. S. Henningson[1]

[1] *Linné FLOW Centre and Swedish e-Science Research Centre (SeRC), KTH Mechanics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden.*
[2] *PMMH (UMR 7636 CNRS – ESPCI – UPMC Paris 6 – UPD Paris 7), 10 rue Vauquelin, 75005 Paris, France.*

Received 13 May 2016; Accepted (in revised version) 10 April 2017

Abstract. For problems governed by a non-normal operator, the leading eigenvalue of the operator is of limited interest and a more relevant measure of the stability is obtained by considering the harmonic forcing causing the largest system response. Various methods for determining this so-called optimal forcing exist, but they all suffer from great computational expense and are hence not practical for large-scale problems. In the present paper a new method is presented, which is applicable to problems of arbitrary size. The method does not rely on timestepping, but on the solution of linear systems, in which the inverse Laplacian acts as a preconditioner. By formulating the search for the optimal forcing as an eigenvalue problem based on the resolvent operator, repeated system solves amount to power iterations, in which the dominant eigenvalue is seen to correspond to the energy amplification in a system for a given frequency, and the eigenfunction to the corresponding forcing function. Implementation of the method requires only minor modifications of an existing timestepping code, and is applicable to any partial differential equation containing the Laplacian, such as the Navier-Stokes equations. We discuss the method, first, in the context of the linear Ginzburg-Landau equation and then, the two-dimensional lid-driven cavity flow governed by the Navier-Stokes equations. Most importantly, we demonstrate that for the lid-driven cavity, the optimal forcing can be computed using a factor of up to 500 times fewer operator evaluations than the standard method based on exponential timestepping.

AMS subject classifications: 76M25, 65F08, 76E15, 76E09, 65F15

Key words: Hydrodynamic stability, optimal forcing, Laplace preconditioner, iterative methods, eigenvalue problem, Ginzburg-Landau equation, lid-driven cavity flow.

*Corresponding author. *Email addresses:* mattiasbr@mech.kth.se (M. Brynjell-Rahkola), laurette.tuckerman@espci.fr (L. S. Tuckerman), pschlatt@mech.kth.se (P. Schlatter), henning@mech.kth.se (D. S. Henningson)

 ©2017 Global-Science Press

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

1 Introduction

In hydrodynamic stability, a classical analysis generally consists of two parts — the determination of a basic state about which the governing equations may be linearized, and the calculation of eigenvalues of the Jacobian $\mathcal{A}$. For non-normal operators, other considerations may be more relevant. For example, solutions may experience transient growth even when all of the eigenvalues are located in the left half of the complex plane, and in a non-linear framework trigger subcritical transition [10,30,31]. Another type of analysis concerns the amplification due to a harmonic driving force $f(x)e^{i\omega t}$, where we seek to determine the temporal frequency ω and spatial profile f that cause the largest energy amplification in the system.

The purpose of this paper is to introduce a novel iterative matrix-free method for computing the optimal forcing of a driven system. This method is best explained by placing it in the context of those used to carry out linear stability analysis, so we begin by surveying these techniques. Denoting by $\mathcal{A}$ the governing operator linearized about a basic state, perturbations $q(x,t)$ obey

$$\frac{\partial q}{\partial t} = \mathcal{A}q. \tag{1.1}$$

The governing operator $\mathcal{A}$ is considered to be spatially dependent, either via the geometrical specifications of the problem, or through a spatially-dependent basic state about which the evolution equations have been linearized, or both. Perturbations q may depend on one, two, or three spatial dimensions. If there is only one spatial dimension, the governing operator can be formulated and treated explicitly. For higher-dimensional systems, if one or two of the spatial directions are homogeneous, then the eigenfunctions are trigonometric or exponential in those directions and the linearized operator is banded or block-diagonal [12]. In such cases, it may still be possible to determine the eigenvalues and eigenfunctions (denoted by eigenpairs) of $\mathcal{A}$ through direct methods.

With increased geometrical complexity, an explicit representation and a full diagonalization of the operator are usually too costly in terms of storage and computational power and it becomes necessary to use matrix-free methods to find the desired eigenpairs. A timestepping algorithm for solving (1.1), which carries out the action of an approximation to the exponential operator $\exp(\mathcal{A}\Delta t)$, is a natural means for doing so. Integrating the linearized equations (1.1) in time is equivalent to carrying out the power method on $\exp(\mathcal{A}\Delta t)$, and will converge to the leading eigenfunction.

Turning to the topic of this paper, when a system is linearly stable, it may nevertheless undergo amplification due to a harmonic driving force, as described by

$$\frac{\partial q}{\partial t} = \mathcal{A}q + fe^{i\omega t}. \tag{1.2}$$

If all of the eigenvalues of $\mathcal{A}$ have negative real part, then $q(x,t) \to -(\mathcal{A}-i\omega\mathcal{I})^{-1}f(x)e^{i\omega t}$

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

as $t\to\infty$. Defining the amplification of (1.2) as

$$G(\omega)\equiv\max_{\|f\|\neq 0}\frac{\|(\mathcal{A}-i\omega\mathcal{I})^{-1}f\|}{\|f\|}, \tag{1.3}$$

we seek to determine the frequency ω and f that lead to the largest amplification in the system. The solution to the optimization problem (1.3) will be referred to as the *optimal forcing* of the system.

In order to find the optimal forcing, (1.2) can be integrated in time, followed by integration of the corresponding adjoint problem, forced by q

$$-\frac{\partial g}{\partial t}=\mathcal{A}^{\dagger}g+2q. \tag{1.4}$$

Algorithms of this type have been employed by Monokrousos et al. [26] to determine the linear optimal initial condition and optimal forcing for the Blasius boundary layer. Techniques based on the approximate exponential are straightforward, robust, and general, but slow and computationally expensive. The reason for this is that the numerical timestepping operator only approximates the exponential operator $\exp(\mathcal{A}\Delta t)$ in the limit of small Δt. When Δt is small, $\exp(\mathcal{A}\Delta t)$ is near the identity, and so each action of the operator has only an incremental effect. Moreover, (1.2) and (1.4) may need to be integrated for a very long time horizon until initial transients have decayed and the solution has converged to an approximate asymptotic time-periodic state.

Another approach, which is equivalent to timestepping with (1.2) and (1.4), is to seek the singular value decomposition (SVD) of $-(\mathcal{A}-i\omega\mathcal{I})^{-1}$, where the right singular vectors correspond to the spatial forcing profiles and the left singular vectors correspond to the spatial structure of the flow responses. This is suitable for problems involving operators small or sparse enough to be formulated explicitly [33,34] or whose dynamics can be captured by a reduced order model involving a modest number of eigenfunctions [1,2]. In practice, the latter approach involving a reduced order model may require a basis of as many as $\sim 10^3$ eigenfunctions even for simple two-dimensional geometries, which renders also this approach intractable for large problems.

In what follows, we describe a matrix-free iterative method for computing optimal forcing that circumvents these issues and is applicable to problems of arbitrary size and complexity. The method can conveniently be implemented by minor modifications of a pre-existing time-integration code, and hence does not require development of any new software. Variants of it have already been used to calculate steady states and eigenpairs in fluid dynamics [6–9, 25, 38, 39] as well as in condensed matter physics [22, 23]. The main purpose of the current article is to demonstrate for the first time that iterative inversion with Laplacian preconditioning can also be used to calculate the frequency and spatial function that yield the optimal forcing. We will show that our method only requires a small fraction of the number of operator evaluations compared to the method of timestepping, as has been shown to be the case for the calculation of steady states and eigenpairs.

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

Previous research on Laplacian preconditioning has been almost entirely applications-oriented, reporting mainly the physical results of the computations; very little analysis of the method itself has been published [8, 38]. A secondary purpose of this article is to carry out a detailed study of its performance and convergence properties, in particular the effect of the value of Δt on the spectra and pseudospectra of the preconditioned operator.

The cases we investigate are the linearized one-dimensional Ginzburg-Landau equation, which is a commonly used model for the Navier-Stokes equations [4, 14, 15, 17], and the two-dimensional lid-driven cavity flow [3]. In Section 2 the concept and properties of the Laplace preconditioner are reviewed and extended, and in Section 3 the optimal forcing algorithm is outlined and applied to the chosen test cases. The article concludes in Section 4 with a discussion.

Eigenvalue spectra are denoted by $\Lambda(\cdot)$ and individual eigenvalues by λ. Inner products $\langle\cdot,\cdot\rangle$, and their corresponding norms $\|\cdot\|=\sqrt{\langle\cdot,\cdot\rangle}$, will refer to those of the L^2-space. The norm $\|\cdot\|_2$ refers to the usual Euclidean norm. A superscript asterisk denotes complex conjugate, and in the case of finite dimensional matrices conjugate transpose. Adjoint operators (possibly infinite dimensional) derived using the L^2-inner product are denoted with a superscript dagger.

2 Laplace preconditioner

2.1 Operator definitions

In what follows we study a linear partial differential equation whose evolution operator $\mathcal{A}$ is the sum of two parts, $\mathcal{L}$ and $\mathcal{N}$, i.e.

$$\frac{\partial q}{\partial t}=\mathcal{A}q=\mathcal{N}q+\mathcal{L}q. \tag{2.1}$$

For the Navier-Stokes equations, $\mathcal{L}$ is usually taken to be the diffusive terms, with the incorporation of the pressure, and $\mathcal{N}$ is taken to be the linearization of the advective terms about the basic state. For the Ginzburg-Landau equation (see Appendix A.1), we take $\mathcal{L}$ to contain the linear reaction coefficient $\mu(x)$ as well as the diffusive term, giving $\mathcal{L}\equiv\gamma\partial^2/\partial x^2+\mu(x)$ and $\mathcal{N}$ to be $\mathcal{N}\equiv-\nu\partial/\partial x$. When the Reynolds number is low or moderate, equations such as (2.1) are often discretized in time via an explicit scheme for $\mathcal{N}$ and an implicit scheme for $\mathcal{L}$. The reason for this is that $\mathcal{L}$ poses a much more stringent stability requirement than $\mathcal{N}$ on Δt. In these cases, our method takes advantage of the implicit timestepping scheme for $\mathcal{L}$ by interpreting it as a preconditioner for the combined operator $\mathcal{N}+\mathcal{L}$. In other cases, such as for high Reynolds numbers, a mixed explicit-implicit scheme is not advantageous. Our method does not apply to such flows.

Choosing the first order forward/backward Euler time discretization leads to

$$\frac{q(x,t+\Delta t)-q(x,t)}{\Delta t}=\mathcal{N}q(x,t)+\mathcal{L}q(x,t+\Delta t), \tag{2.2}$$

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

from which the timestepping operator $\mathcal{B}(\Delta t)$ can be expressed as

$$q(x,t+\Delta t)=(\mathcal{I}-\Delta t\mathcal{L})^{-1}(\mathcal{I}+\Delta t\mathcal{N})q(x,t)\equiv\mathcal{B}(\Delta t)q(x,t). \tag{2.3}$$

Given an existing timestepping code, we can calculate the difference between two consecutive solution fields separated by a single timestep Δt:

$$\begin{aligned} q(x,t+\Delta t)-q(x,t)&=(\mathcal{B}(\Delta t)-\mathcal{I})q(x,t)\\ &=\left[(\mathcal{I}-\Delta t\mathcal{L})^{-1}(\mathcal{I}+\Delta t\mathcal{N})-\mathcal{I}\right]q(x,t)\\ &=(\mathcal{I}-\Delta t\mathcal{L})^{-1}\Delta t(\mathcal{N}+\mathcal{L})q(x,t). \end{aligned} \tag{2.4}$$

Defining

$$\mathcal{P}(\Delta t)\equiv(\mathcal{I}-\Delta t\mathcal{L})^{-1}\Delta t, \tag{2.5}$$

the calculation of (2.4) shows that the difference between two consecutive solution fields provides the action of the evolution operator $\mathcal{A}$ left-multiplied by $\mathcal{P}(\Delta t)$, i.e.

$$\mathcal{B}(\Delta t)-\mathcal{I}=\mathcal{P}(\Delta t)\mathcal{A}. \tag{2.6}$$

The main point of this approach is that the operator $\mathcal{PA}$ is far better conditioned than $\mathcal{A}$ and thus that iterative methods solving linear systems involving $\mathcal{PA}$ converge far more quickly than those involving $\mathcal{A}$.

We can interpret the cause of this to be the fact that $\mathcal{P}$ is a good preconditioner for $\mathcal{A}$. As will be discussed in Section 2.2, the operator $\mathcal{L}$ is responsible for the ill-conditioning of $\mathcal{A}$. The preconditioning property of $\mathcal{P}$ can be understood by considering the limit of large timesteps Δt for which $\mathcal{P}(\Delta t)=(\mathcal{I}-\Delta t\mathcal{L})^{-1}\Delta t\approx-\mathcal{L}^{-1}$, thus counteracting the poor conditioning of $\mathcal{L}$. For small timesteps, $\mathcal{P}(\Delta t)\approx\Delta t\mathcal{I}$, and hence provides no preconditioning. Thus $\mathcal{P}$ can be viewed as interpolating between $\Delta t\mathcal{I}$ and $-\mathcal{L}^{-1}$ as Δt is increased. We will refer to $\mathcal{P}$ and $\mathcal{P}^\dagger$ as the direct and adjoint Laplace preconditioner. Considering the L^2-inner product, it is straightforward to show that $\mathcal{P}$ is self-adjoint if $\mathcal{L}$ is. This will for instance be the case for the Navier-Stokes flow case considered later, where $\mathcal{L}$ is defined by (2.12b), but not the case for the Ginzburg-Landau equation with $\mathcal{L}=\gamma\partial^2/\partial x^2+\mu(x)$.

We stress that we do not carry out the actions of $\mathcal{A}$ and of $\mathcal{P}$ separately. Given the existence of a timestepping code which effectively carries out $\mathcal{B}$, the action of $\mathcal{B}-\mathcal{I}$ is more accessible than that of $\mathcal{A}$. The action of $\mathcal{P}$ by itself may be performed by taking a single implicit timestep with the linear operator $\mathcal{L}$. Our assumption is that none of the operators $\mathcal{A}$, $\mathcal{B}$, $\mathcal{L}$, $\mathcal{N}$, $\mathcal{P}$ are stored as matrices, and that only the actions of $\mathcal{B}$, $\mathcal{P}$ and their adjoint counterparts $\mathcal{B}^\dagger$, $\mathcal{P}^\dagger$ are available to us via the timestepping code. We emphasize that the Laplace preconditioner is an intrinsic part of our operator and thus requires no additional computational cost, unlike many other preconditioners.

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

2.2 Preconditioning properties

To empirically examine the properties of $\mathcal{P}$, we consider the one-dimensional Ginzburg-Landau equation and solve a linear system involving operator $\mathcal{PA}$ with a random right-hand side using various Δt. Since the matrix $\mathcal{A}$ in general is non-Hermitian (see Section A.1), a method designed for non-Hermitian systems such as bi-conjugate gradients stabilized (Bi-CGSTAB) [40] or generalized minimal residual (GMRES) [32] is required. For the present analysis we will use GMRES and a relative tolerance of 10^{-13}. The results of these calculations are shown in Fig. 1(a), where it can be seen that the iteration count goes down steadily with increasing Δt. To gain better insight in the properties of the preconditioner, we turn to the convergence theory of GMRES [11, 18, 19, 32].

The residual r_k is bounded by

$$\frac{\|r_k\|_2}{\|r_0\|_2} \leq \min_{p_k \in \mathbb{P}_k^0} \|p_k(\mathcal{PA})\|_2, \tag{2.7}$$

where $\mathbb{P}_k^0 = \{\text{polynomials } p \text{ of order} \leq k \text{ satisfying } p(0)=1\}$ is the space of polynomials over which the minimization is carried out. GMRES finds the optimal polynomial $p_k \in \mathbb{P}_k^0$ which realizes the minimum in (2.7).

The roots of the approximating polynomial p_k are the harmonic Ritz values [37] and are given by the eigenvalues of the matrix $\mathbf{H}_k + h_{k+1,k}^2 (\mathbf{H}_k^*)^{-1} e_k e_k^*$, where e_k is the kth dimensional unit vector and $\mathbf{H}_k$ is an upper Hessenberg matrix satisfying the Arnoldi relation:

$$(\mathcal{PA})\mathbf{W}_k = \mathbf{W}_k \mathbf{H}_k + h_{k+1,k} w_{k+1} e_k^*. \tag{2.8}$$

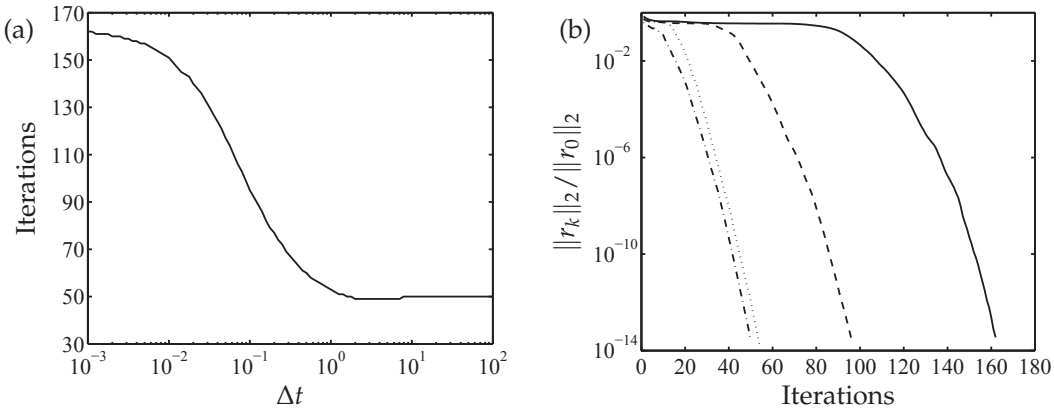


Figure 1: Preconditioning of the Ginzburg-Landau equation with Laplace preconditioner. Frame (a) shows the number of GMRES-iterations required to solve the system $\mathcal{PA}$ with a random right-hand side for different Δt. Frame (b) plots residual as a function of iteration count for $\Delta t = 10^{-3}$ (solid line), $\Delta t = 0.1$ (dashed line), $\Delta t = 1$ (dotted line), $\Delta t = 10$ (dot-dashed line).

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

The right-hand side of (2.7) can be bounded in terms of the eigenvalues of $\mathcal{PA}$, i.e.

$$\min_{p_k\in\mathbb{P}_k^0}\|p_k(\mathcal{PA})\|_2\leq\kappa(\mathbf{X})\min_{p_k\in\mathbb{P}_k^0}\max_{\lambda\in\Lambda(\mathcal{PA})}|p_k(\lambda)|, \tag{2.9}$$

where $\kappa(\mathbf{X})$ denotes the 2-norm condition number of the eigenfunction matrix $\mathbf{X}$. If a matrix is normal (i.e. if it has a complete set of orthogonal eigenvectors [37]), then $\kappa(\mathbf{X})=1$. Thus a requirement for fast convergence of GMRES for a reasonably normal matrix $\mathcal{PA}$ is that the polynomial p_k can be made small on the spectrum $\Lambda(\mathcal{PA})$ [19, 32]. If the operator $\mathcal{PA}$ is strongly non-normal, as is the case for the Ginzburg-Landau operator with the parameter values specified in Section A.1, $\kappa(\mathbf{X})$ will be very large, and the bound given by (2.9) may be uninformative [21]. In such cases a more suitable convergence bound is obtained by looking at the ε-pseudospectra of $\mathcal{PA}$ [37], defined as

$$\Lambda_\varepsilon(\mathcal{PA})\equiv\{\sigma\in\mathbb{C}\ \mid\ \|(\sigma\mathcal{I}-\mathcal{PA})^{-1}\|_2\geq\varepsilon^{-1}\}. \tag{2.10}$$

A convergence bound based on the ε-pseudospectra rather than the eigenvalues can be derived to be

$$\min_{p_k\in\mathbb{P}_k^0}\|p_k(\mathcal{PA})\|_2\leq\frac{L(\Gamma_\varepsilon)}{2\pi\varepsilon}\min_{p_k\in\mathbb{P}_k^0}\max_{\sigma\in\Lambda_\varepsilon(\mathcal{PA})}|p_k(\sigma)|, \tag{2.11}$$

in which Γ_ε is the bounding contour of $\Lambda_\varepsilon(\mathcal{PA})$ for a given ε and $L(\Gamma_\varepsilon)$ signifies the length of this contour. As shown in (2.11), the pseudospectral convergence bounds depend on the particular value of ε, and different values may characterize the convergence at different stages of the iteration [18, 19].

In Fig. 2, the effect of different Δt on the eigenvalue spectra and ε-pseudospectra of the operator $\mathcal{PA}$ is shown. We focus here on the direct operator as similar behavior is observed for the adjoint operator. To monitor the convergence, along with the spectra and pseudospectra, we also plot the harmonic Ritz values [37].

Recall that, according to (2.5), for small timesteps, application of $\mathcal{P}$ amounts to a mere scaling by Δt of $\mathcal{A}$. Comparison of Fig. 2(a) and Fig. 5(a) shows that the former spectrum essentially resembles a scaled down copy of the latter. In the limit of small Δt the spectrum and pseudospectrum of $\mathcal{PA}$ will be very close to the origin, where the approximating polynomials are normalized to have value one. Therefore, in order for the polynomial to attain a small value on the spectra and pseudospectra, a large number of iterations are required before it has sufficiently many degrees of freedom to satisfy both requirements. Smaller values of Δt will scale down the spectra even further, while increasing Δt moves the spectrum away from the origin and changes its shape, hence facilitating polynomial fitting (Figs. 2(b)-2(c)). Eventually, several eigenvalues are expelled from the core of the spectrum (Fig. 2(d)). However, given that they are located far away from the origin, each of them will require only one zero in the approximating polynomial and may be annihilated in a single iteration. As discussed in [18], outliers do not affect the asymptotic convergence rate. This is also seen in Fig. 1(b), where the asymptotic convergence rate for $\Delta t=1$ and $\Delta t=10$ are approximately the same, although the spectrum for $\Delta t=10$

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

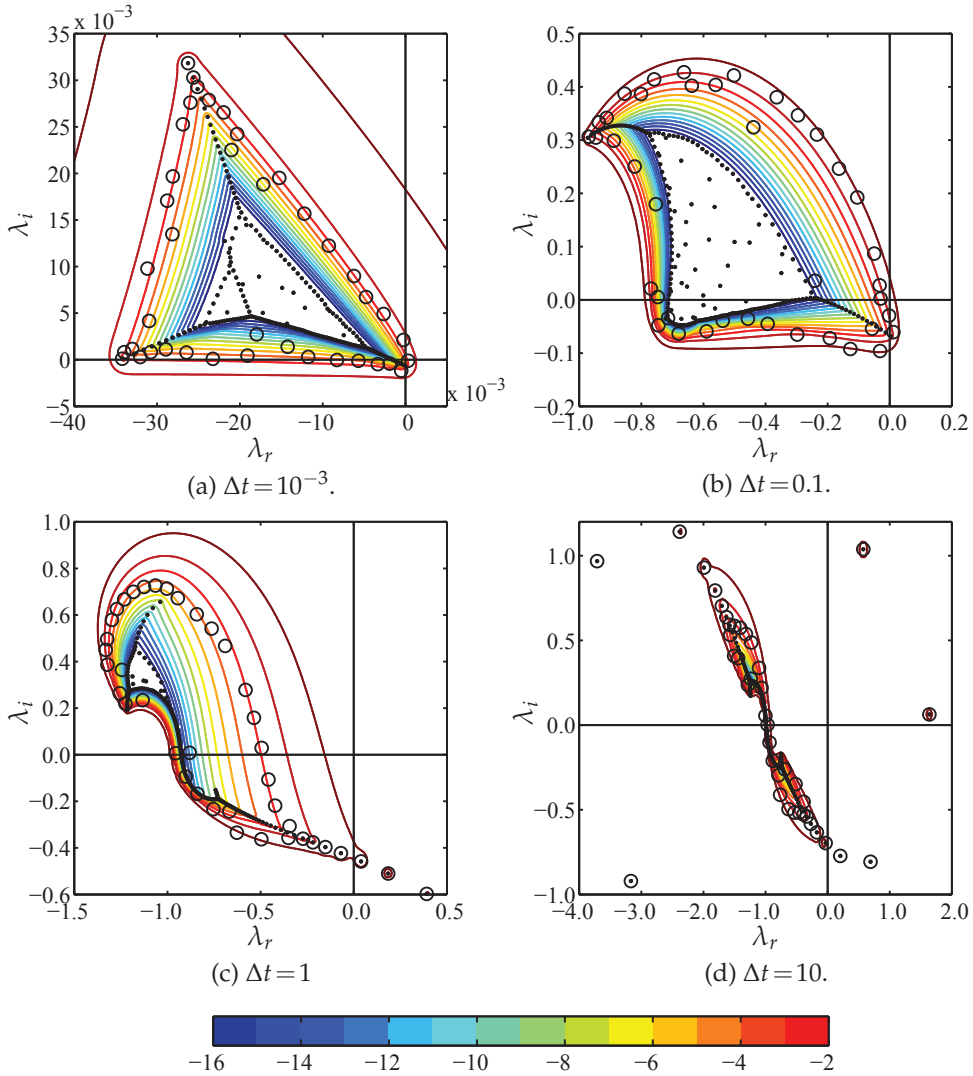


Figure 2: Effect of Δt on the eigenvalue spectra and ε-pseudospectra of $\mathcal{PA}$. Plotted are the eigenvalue spectrum (dots), harmonic Ritz values of GMRES-iteration k=40 (circles), and the contours of the ε-pseudospectra ranging from 10^{-2} to 10^{-15}. For $\Delta t \geq 1$ the core of the spectrum remains centered around $\lambda_r = -1$, $\lambda_i = 0$ and does not move relative to the origin. Notice how the harmonic Ritz values tend to be roughly aligned with the contours Γ_ε following (2.11), and for large Δt encircle an increasing number of eigenvalues.

contains many more outliers. Fig. 1(b) also displays an initial region of stagnation in convergence for small values of Δt. This may partly be explained by the non-normality of the matrix $\mathcal{PA}$ [19], but given the rapid shortening of this region with increasing Δt, this more likely is due to near-origin clustering. As a final remark, we notice no significant change in the shape and location of the main part of the spectra for $\Delta t \gtrsim 5$, which is in line with what can be expected given the asymptotic behavior shown in Fig. 1(a).

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

2.3 Extension to Navier-Stokes equations

In the present section, the concepts introduced in the previous sections for the one-dimensional Ginzburg-Landau equation will be verified in the multi-dimensional setting governed by the Navier-Stokes equations. The direct and adjoint fields will be denoted $\mathbf{q}(\mathbf{x},t)$ and $\mathbf{g}(\mathbf{x},t)$, where $\mathbf{x}\in\mathbb{R}^d$ (d is the number of dimensions). As discussed in Appendix A.2, the incompressible Navier-Stokes equations can be formulated in the same form as (2.1), namely $\partial\mathbf{q}/\partial t=\mathcal{N}\mathbf{q}+\mathcal{L}\mathbf{q}$, where

$$\mathcal{N}\mathbf{q}=\mathscr{P}_\nabla(-(\bar{\mathbf{q}}\cdot\nabla)\mathbf{q}-(\mathbf{q}\cdot\nabla)\bar{\mathbf{q}}), \tag{2.12a}$$

$$\mathcal{L}\mathbf{q}=\mathscr{P}_\nabla(Re^{-1}\nabla^2\mathbf{q}), \tag{2.12b}$$

and the operator $\mathscr{P}_\nabla(\cdot)$ is used to project the velocities onto a divergence-free field [16].

Although the method presented in Section 3 and all concepts discussed in this paper apply to cases of arbitrary size and complexity, we will illustrate it on the well-known two-dimensional lid-driven cavity flow (see Section A.2). To verify that the desired behavior of the preconditioner is also obtained with the Navier-Stokes equations, a figure equivalent to Fig. 1 for the system $\mathcal{PA}$ with $\mathcal{L}$ and $\mathcal{N}$ defined in (2.12) is generated. In Fig. 3(a), the system is solved to a relative tolerance of 10^{-10} for different values of Δt with a random right-hand side. Indeed, the same trend as observed in Fig. 1(a) is obtained. Variation of the Reynolds number from $Re=100$ to $Re=500$ increases the iteration count required for convergence and Δt at which the preconditioner saturates. Fig. 3(b) shows how the residual of the calculation varies with iteration for different Δt. Again, the same behavior with respect to Δt as seen in Fig. 1(b) is observed (here only the results for $Re=100$ are shown for clarity).

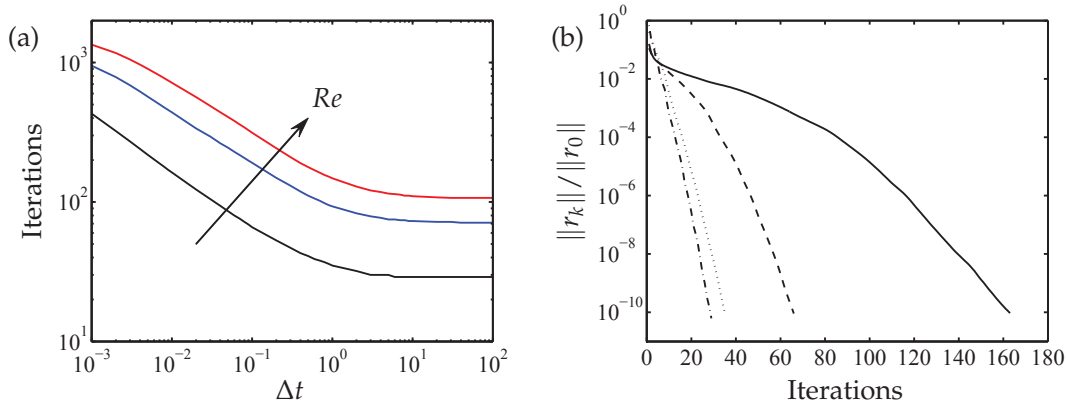


Figure 3: Convergence of GMRES for the lid-driven cavity. Frame (a) shows the iteration count versus timestep Δt for $Re=100$ (black), $Re=300$ (blue) and $Re=500$ (red), and frame (b) illustrates the decrease in residual with iteration count for $Re=100$ and $\Delta t=10^{-2}$ (solid line), $\Delta t=0.1$ (dashed line), $\Delta t=1$ (dotted line), $\Delta t=10$ (dot-dashed line).

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

3 Optimal forcing

With the increased understanding of the properties of Laplace preconditioner from Section 2, we now explain how to apply these techniques to the problem of optimal forcing. As in the previous section, we start by deriving the method for the one-dimensional Ginzburg-Landau equation, and then apply it to the two-dimensional Navier-Stokes equations.

3.1 Application of Laplace preconditioner

We recall from Section 1 that we are interested in the harmonically driven system

$$\frac{\partial q}{\partial t} = \mathcal{A}q + fe^{i\omega t},$$

where $q = q(x,t)$ and x represents the spatial coordinate. The linear operator A depends on x and usually contains spatial derivatives. This problem has as its general solution

$$q(x,t) = e^{\mathcal{A}t}c(x) - (\mathcal{A} - i\omega\mathcal{I})^{-1}f(x)e^{i\omega t}, \tag{3.1}$$

where $c(x) = q(x,0) + (\mathcal{A} - i\omega\mathcal{I})^{-1}f(x)$. Assuming that all of the eigenvalues of $\mathcal{A}$ have negative real part, then as $t \to \infty$, $q(x,t) \to s(x)e^{i\omega t}$, where

$$s(x) \equiv -(\mathcal{A} - i\omega\mathcal{I})^{-1}f(x) \equiv -\mathcal{R}(i\omega)f(x). \tag{3.2}$$

Thus, an input field f is mapped into an output field s by the resolvent operator $\mathcal{R}$. Defining the amplification (gain) of (1.2) as

$$G(\omega) \equiv \max_{\|f\| \neq 0} \frac{\|s(x)\|}{\|f(x)\|} = \max_{\|f\| \neq 0} \sqrt{\frac{\langle -\mathcal{R}(i\omega)f(x), -\mathcal{R}(i\omega)f(x)\rangle}{\langle f(x), f(x)\rangle}}, \tag{3.3}$$

we seek to determine the forcing frequency ω and forcing profile f that yield the largest amplification. The problem of determining the amplification $G(\omega)$ and f, given a frequency ω, is thus that of finding the dominant eigenpair of the operator

$$\mathcal{R}^\dagger(i\omega)\mathcal{R}(i\omega) = \Big((\mathcal{A} - i\omega I)(\mathcal{A}^\dagger + i\omega I)\Big)^{-1}. \tag{3.4}$$

The operator $\mathcal{R}^\dagger\mathcal{R}$ is self-adjoint, which implies that all of its eigenvalues are real and can be ordered in descending order. Physically, $\langle f, \mathcal{R}^\dagger\mathcal{R}f\rangle$ is a measure of the kinetic energy in the domain, and will thus always be larger than or equal to zero.

As previously stated, it is generally not possible to construct and treat the operators $\mathcal{A}$ or $\mathcal{A}^\dagger$ explicitly. However, as shown in (2.4)-(2.6), the matrix-vector products are realized

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

through the difference between solution fields separated by one direct or adjoint first-order Euler timestep. Given an estimate $f^{(k)}$, we produce $f^{(k+1)} \equiv \mathcal{R}^\dagger \mathcal{R} f^{(k)}$ as follows:

$$f^{(k+1)} = \Big((\mathcal{A}-i\omega\mathcal{I})(\mathcal{A}^\dagger+i\omega\mathcal{I})\Big)^{-1} f^{(k)} \qquad \Longleftrightarrow \tag{3.5a}$$

$$(\mathcal{A}-i\omega\mathcal{I})(\mathcal{A}^\dagger+i\omega\mathcal{I}) f^{(k+1)} = f^{(k)} \qquad \Longleftrightarrow \tag{3.5b}$$

$$\mathcal{P}(\mathcal{A}-i\omega\mathcal{I})\mathcal{P}^{\dagger-1}\mathcal{P}^\dagger(\mathcal{A}^\dagger+i\omega\mathcal{I}) f^{(k+1)} = \mathcal{P} f^{(k)}. \tag{3.5c}$$

Formation of each power iterate $f^{(k)}$ thus requires the solution of two linear systems, which will be well-conditioned given a suitably chosen Δt, and one application of the direct and adjoint Laplace preconditioner. We emphasize again that we neither form nor factorize any of the operators in system (3.5). Our method is entirely matrix-free, meaning that all linear systems are solved using GMRES, which requires only the action of the operator on a field.

The complete algorithm for solving (3.5c) and carrying out power iterations is given in Algorithm 1. Steps 4–7 essentially implement (3.5c), whereas step 8 evaluates the Rayleigh quotient. As discussed in Section 2.1, the action of $\mathcal{PA}$ (and $\mathcal{P}^\dagger\mathcal{A}^\dagger$) are obtained by (2.4). The action of $\mathcal{P}$ (and $\mathcal{P}^\dagger$) can be obtained by simply taking an implicit timestep with $\mathcal{L}$. If $\mathcal{PA}$, $\mathcal{P}$ and their corresponding adjoint operators are real, and the complex arithmetic needs to be handled explicitly, the sizes of the systems are doubled, so that on each iteration the system

$$\begin{bmatrix} \mathcal{PA} & \omega\mathcal{P} \\ -\omega\mathcal{P} & \mathcal{PA} \end{bmatrix} \begin{pmatrix} v_r \\ v_i \end{pmatrix} = \begin{pmatrix} w_r \\ w_i \end{pmatrix} \tag{3.6}$$

is solved, followed by a similar equation for the adjoint problem.

Due to the self-adjointness of $\mathcal{R}^\dagger\mathcal{R}$, eigenvectors corresponding to different eigenvalues will be pairwise orthogonal. If sub-optimal forcing profiles are of interest, for instance if different forcing profiles have similar amplification rates, or, if a basis on which an arbitrary forcing can be represented is sought, the solution of (3.5c) (steps 4–7 of Algorithm 1) can be coupled with the Lanczos algorithm [28] in a straightforward manner.

In order to obtain the optimal forcing frequency and forcing function, the optimization problem (3.3) must be solved for a range of frequencies ω. Hence, computational savings can be achieved with a good choice of initial vector $f^{(0)}$. Techniques exist for recycling Krylov subspaces when solving a sequence of slightly varying linear systems [27], but here a sequence of slightly varying eigenvalue problems needs to be considered. Assuming that $\Delta\omega$ is small, then the operators $\mathcal{R}^\dagger(i\omega+i\Delta\omega)\mathcal{R}(i\omega+i\Delta\omega)$ and $\mathcal{R}^\dagger(i\omega)\mathcal{R}(i\omega)$ will be close to one another (a first order Taylor expansion shows that $\mathcal{R}^\dagger(i\omega+i\Delta\omega)\mathcal{R}(i\omega+i\Delta\omega) = \mathcal{R}^\dagger(i\omega)\mathcal{R}(i\omega) - \Delta\omega[(2\omega\mathcal{I}+i(\mathcal{A}-\mathcal{A}^\dagger))(\mathcal{R}^\dagger(i\omega)\mathcal{R}(i\omega))^2]$), and as a result, it is reasonable to expect the eigenpairs of the two operators also to be close. Hence, a good candidate for the initial vector of Algorithm 1 applied to $\omega+\Delta\omega$ is the optimal forcing profile obtained for the preceding frequency ω. Alternatively, if the Lanczos algorithm

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

Algorithm 1: Optimal forcing through application of the Laplace preconditioner and the inverse power method.

Input : Forcing frequency ω, initial guess $f^{(0)}$
Output: Amplification G, force profile f

$\tilde{f}^{(0)} \leftarrow f^{(0)}/\|f^{(0)}\|$;
$\alpha^{(0)} \leftarrow 0$;
for $k=1,\dots$ **do**
  $w \leftarrow \mathcal{P}\tilde{f}^{(k-1)}$;
  solve $(\mathcal{P}\mathcal{A}-i\omega\mathcal{P})v=w$;
  $w \leftarrow \mathcal{P}^\dagger v$;
  solve $(\mathcal{P}^\dagger\mathcal{A}^\dagger+i\omega^*\mathcal{P}^\dagger)v=w$;
  $\alpha^{(k)} \leftarrow \langle \tilde{f}^{(k-1)}, v\rangle$;
  $\tilde{f}^{(k)} \leftarrow v/\|v\|$;
  if *satisfied* **then** break;
end
$G \leftarrow \sqrt{\alpha^{(k)}}$;
$f \leftarrow \tilde{f}^{(k)}$;

is carried out, the initial vector can be chosen as a linear combination of the dominant eigenvectors for the preceding frequency.

3.2 Validation case: Ginzburg-Landau equation

As an initial validation case, Algorithm 1 is implemented and applied to the Ginzburg-Landau equation. Given the preconditioning capabilities of the Laplace preconditioner shown in Fig. 1(a), a timestep of $\Delta t = 10.0$ is chosen. In order to validate the results, we also consider an SVD of the resolvent operator, which is a convenient way of computing the optimal forcing and its corresponding response for a small problem like the present one. Given (3.3), and the Cholesky factorization of the weight matrix associated with the L^2-inner product, $\mathbf{M}$, the gain can be written as

$$G(\omega) = \|\mathbf{M}(\mathcal{A}-i\omega I)^{-1}\mathbf{M}^{-1}\|_2. \tag{3.7}$$

An SVD of the operator appearing on the right-hand side of (3.7) yields

$$\mathbf{M}(\mathcal{A}-i\omega I)^{-1}\mathbf{M}^{-1} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \Longleftrightarrow (\mathcal{A}-i\omega I)^{-1}(\mathbf{M}^{-1}\mathbf{V}) = (\mathbf{M}^{-1}\mathbf{U})\mathbf{\Sigma}, \tag{3.8}$$

where the weighted right singular vectors $(\mathbf{M}^{-1}\mathbf{V})$ are to be interpreted as forcing profiles, the weighted left singular vectors $(\mathbf{M}^{-1}\mathbf{U})$ as the spatial distributions of the corresponding flow responses, and the singular values $\mathbf{\Sigma}$ as their amplifications.

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

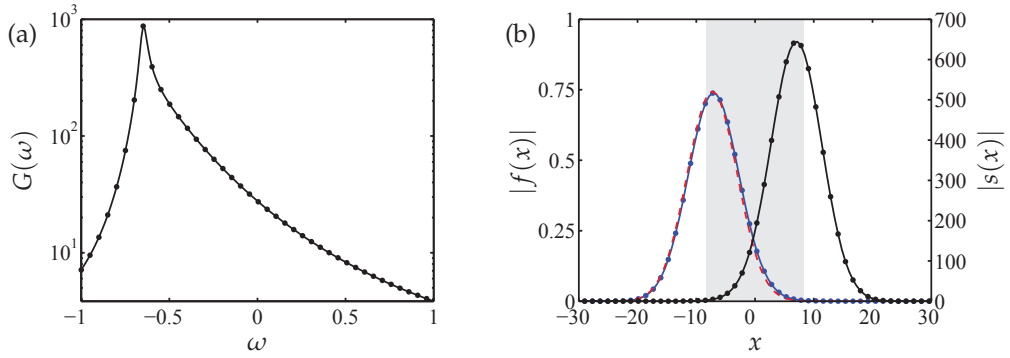


Figure 4: Optimal forcing of the Ginzburg-Landau operator. (a) Frequency response with a peak energy amplification for $\omega=-0.648$. (b) The modulus of the optimal forcing function (solid blue line), the modulus of the adjoint eigenfunction corresponding to eigenvalue $\lambda=-0.018+i0.648$ (dashed red line) scaled according to the left vertical axis, and the optimal response (solid black line) to this forcing scaled according to the right vertical axis. The spatial region of exponential instability is shown in gray. The solid dots are results from an SVD, i.e. σ_1 in (a), and $\mathbf{M}^{-1}u_1$ (blue) and $\sigma_1\mathbf{M}^{-1}v_1$ (black) for $\omega=-0.648$ in (b) (for a description, see the text).

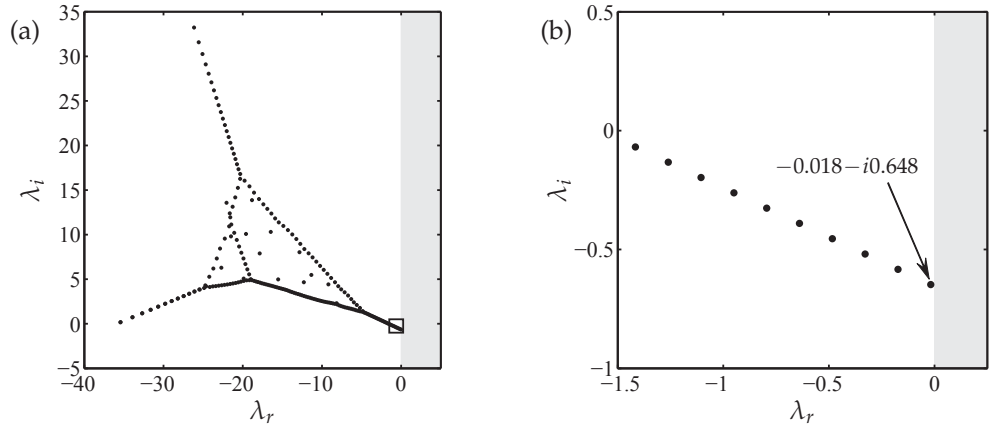


Figure 5: Eigenvalues of the discretized direct Ginzburg-Landau operator. A close-up of the leading ones marked with a rectangle in frame (a), is plotted in frame (b). The region of instability is shown in gray.

The frequency response is shown in Fig. 4(a) and has an energy peak at $\omega=-0.648$ due to resonance with the leading eigenvalue (see Fig. 5). An amplification peak for the same frequency was also obtained by Bagheri et al. [4] when applying a Gaussian force centered around the upstream instability branch (branch I, $x=-8.246$) for the same configuration.

The corresponding optimal forcing profile is plotted in Fig. 4(b) with a blue solid line. The profile is seen to have a peak at $x=-7.202$ around the upstream instability branch and to closely resemble the adjoint eigenfunction corresponding to eigenvalue $\lambda=-0.018+i0.648$, shown with a red dashed line. The optimal response to this forcing

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

is recovered by solving (3.2), which in practice amounts to solving

$$-(\mathcal{P}\mathcal{A}-i\omega\mathcal{P})s=\mathcal{P}f. \tag{3.9}$$

The optimal response is plotted with a black solid line in Fig. 4(b). We see that the optimal response is oriented towards the downstream instability branch (branch II, $x=8.246$) and has a norm that matches the peak in Fig. 4(a).

Comparison of these results with the largest singular value σ_1 (which is equivalent to the resolvent norm), and the corresponding left and right singular vectors $\mathbf{M}^{-1}u_1$ and $\sigma_1\mathbf{M}^{-1}v_1$ shows a perfect agreement.

3.3 Validation case: Navier-Stokes equations (lid-driven cavity flow)

Next, Algorithm 1 is applied to the lid-driven cavity flow. In addition to the operators defined in (2.12), we now require their adjoint counterparts, i.e.

$$\mathcal{N}^{\dagger}\mathbf{g}=\mathscr{P}_{\nabla}((\bar{\mathbf{q}}\cdot\nabla)\mathbf{g}-(\nabla\bar{\mathbf{q}})^{T}\mathbf{g}), \tag{3.10a}$$

$$\mathcal{L}^{\dagger}\mathbf{g}=\mathscr{P}_{\nabla}(Re^{-1}\nabla^{2}\mathbf{g}) \tag{3.10b}$$

(see Section A.2 for details and references).

The outcome of the computation for $Re=100$ is shown in Fig. 6(b). As seen, the largest amplification, $G(\omega)=1.9879$, is achieved for a steady forcing, which resonates with the eigenmode corresponding to the leading eigenvalue $\lambda=-0.5425+i0.0$ (see Fig. 6(a)). To recover the optimal flow response, one can either solve (3.9), or substitute the optimal forcing into the governing equation and integrate the solution in time. Here we choose

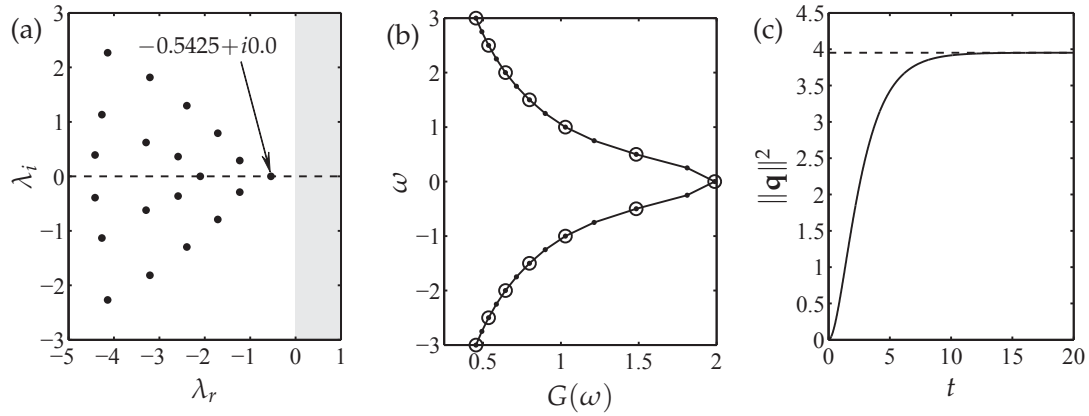


Figure 6: Frequency response of the lid-driven cavity at $Re=100$. The eigenvalue spectra showing the 20 leading eigenvalues of the lid-driven cavity is plotted in frame (a) (the region of instability is colored in gray). The energy amplification for different frequencies is shown in frame (b), where results obtained with Algorithm 1 and 2 are plotted with dots and circles, respectively. Frame (c) shows the energy evolution in the system when driven by a steady force corresponding to the amplification peak in frame (b).

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

Algorithm 2: Optimal forcing through time integration.

Input : Forcing frequency ω, initial guess $f^{(0)}$, optimization horizon T
Output: Amplification G, force profile f

$\tilde{f}^{(0)} \leftarrow f^{(0)}/\|f^{(0)}\|$;
$\alpha^{(0)} \leftarrow 0$;
for $k=1,\dots$ **do**
  $q|_{t=0} \leftarrow 0$;
  integrate $q|_{t=T} = \int_0^T \left(\mathcal{A}q + \tilde{f}^{(k-1)} e^{i\omega t}\right) \mathrm{d}t$,
  obtain the response $s(x)$ during the last period of ω in $t \in [0,T]$;
  $g|_{t=T} \leftarrow 0$;
  integrate $g|_{t=0} = -\int_T^0 \left(\mathcal{A}^\dagger g + 2s e^{i\omega t}\right) \mathrm{d}t$,
  obtain the response $b(x)$ during the last period of ω in $t \in [0,T]$;
  $\alpha^{(k)} \leftarrow \|b\|/2$;
  $\tilde{f}^{(k)} \leftarrow b/(2\alpha^{(k)})$;
  if *satisfied* **then** break;
end
$G \leftarrow \sqrt{\alpha^{(k)}}$;
$f \leftarrow \tilde{f}^{(k)}$;

the latter approach, which gives the solution $\mathbf{q}(\mathbf{x},t) = (e^{\mathcal{A}t} - \mathcal{I})\mathcal{A}^{-1}\mathbf{f}(\mathbf{x})$ (since $\omega = 0$). As expected, $\mathbf{q} \to -\mathcal{R}(i\omega)\mathbf{f}$ and $\|\mathbf{q}\|^2 \to G^2$ after a brief transient phase (see Fig. 6(c)).

As a validation, we compare our results with those obtained with the method of [26]. This method is derived by seeking a stationary point of the Lagrangian functional (see the derivation in Appendix B). On each iteration, the direct equation is integrated forward in time and during the last period of the forcing frequency, the flow response is computed, e.g. through a Fourier transformation. This flow response is then used as a forcing of the adjoint equation, which is integrated backward in time. The details of the different steps are outlined in Algorithm 2 and the result of this algorithm for a sample of forcing frequencies are shown in Fig. 6(b). As seen, the agreement between the results obtained with the two algorithms is good.

To show the efficiency of the novel method, the computational cost associated with calculating the largest eigenpair of the operator $\mathcal{R}^\dagger(i\omega)\mathcal{R}(i\omega)$ using Algorithm 1 and Algorithm 2 is investigated. An important note regarding the two methods is the role of Δt, which in Algorithm 1 is an algebraic parameter that should be chosen large enough to ensure convergence of the linear systems. (Absolute and relative tolerances of 10^{-10} and 10^{-13}, respectively, are used as a convergence criterion for the iterative solvers.) In contrast, when carrying out the time integration in Algorithm 2, we march the solution using the timestepping operator, i.e. $\mathbf{q}(\mathbf{x},t+\Delta t) = \mathcal{B}(\Delta t)\mathbf{q}(\mathbf{x},t) + \mathcal{P}(\Delta t)\Re(\mathbf{f}e^{i\omega t})$. The timestep

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

Δt must be chosen small so as to achieve sufficient precision in the timestepping scheme and to achieve stability via the CFL-condition. Moreover, the time horizon T in Algorithm 2 should be chosen long enough such that the homogeneous solution in (3.1) has decayed and the output is dominated by the particular solution. Convergence is measured by $|\alpha^{(k)}-\alpha^{(k-1)}|$, and the computation is halted when this value is below 10^{-6}. We furthermore choose T to be the smallest integer number of forcing periods that yields a change in the gain that is below this tolerance. Correspondingly, for the case of $\omega=0.0$, the integration time was gradually increased in steps of 500 timesteps until the further increases led to a change in gain that was below the tolerance. For consistency of the comparison, time integration is carried out using the first order forward/backward Euler time discretization defined in (2.3). The initial condition for the forcing profile is random noise in both algorithms.

Since $\mathcal{N}$ can be applied explicitly, the cost of each timestep is mainly due to solving $(\mathcal{I}-\Delta t\mathcal{L})^{-1}$. Therefore, each application of $\mathcal{B}$, $\mathcal{PA}$ and $\mathcal{P}$ and their adjoint counterparts is measured as 1 cost unit, meaning that the price of each timestep in Algorithm 2 is 1 unit (the force term can be grouped together with the advective term), and the price of each iteration required to solve the systems in Algorithm 1 is 4 units (or 1 unit if $\omega=0$), assuming that the complex arithmetic must be handled explicitly by doubling the system size according to (3.6).

The resulting number of operator evaluations are presented in Table 1. Both methods are shown to give the same amplification to the third or fourth digit, but the number of operator evaluations associated with Algorithm 1 is between 30 and 490 times lower than that associated with Algorithm 2 depending on the forcing frequency. For generality, we do not compare the wall-clock time associated with executing the different algorithms since this is strongly dependent on the spatial discretization. However, one could expect these figures to be reflected in the execution time for a numerical scheme in which the cost of evaluating the above operators is independent of the timestep. This is the case, for example, if the linear problems are solved directly, as they are in a spectral method with a tensor-product basis. Although the above comparison indicates that the inverse power method has the potential for being far less costly than time integration for the same accuracy, the exact figures are strongly dependent on the flow case, tolerance and mesh.

Table 1: Comparison of the results and the number of operator evaluations associated with Algorithm 1 and Algorithm 2 for different ω.

	Time integration			Inverse power method	
ω	T	cost	$G(\omega)$	cost	$G(\omega)$
0.0	29.00	261,000	1.987883	533	1.987877
1.0	31.42	471,240	1.029109	3,764	1.029304
3.0	14.66	425,140	0.454935	7,208	0.454855
5.0	7.54	452,400	0.275763	15,020	0.275634

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

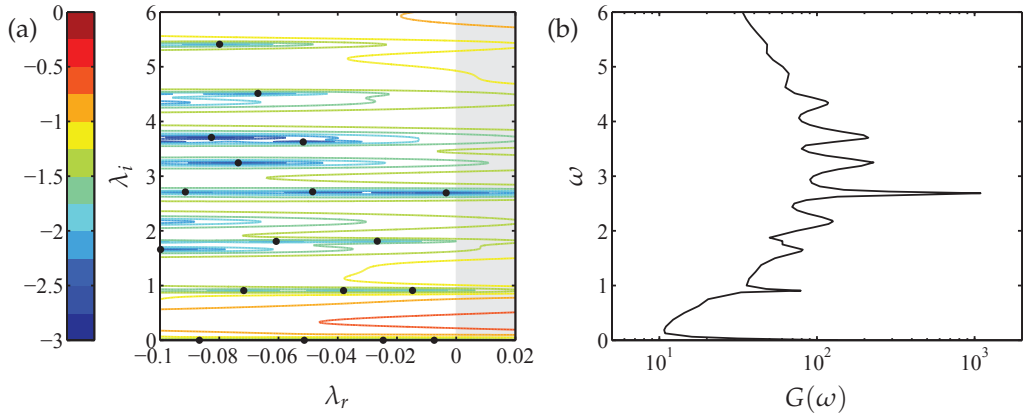


Figure 7: Frequency response of the lid-driven cavity at $Re=8015$. The leading eigenvalues of the lid-driven cavity and the ε-pseudospectrum (logarithmically spaced colored contours) are plotted in frame (a) (the region of instability is colored in gray). The energy amplification for different frequencies is shown in frame (b).

As a final example, we increase the Reynolds number to $Re=8015$, which is very close to the critical Reynolds number $Re_c \in [8017.6, 8018.8)$ at which the flow undergoes a Hopf bifurcation [3]. Since the flow is asymptotically stable, but close to the first bifurcation point, we expect a large energy amplification in the presence of harmonic forcing, which makes this configuration a suitable prototype problem for studying optimal forcing. The streamfunction of the baseflow, visualized with black contours in Figs. 8 and 9, shows a large primary vortex in the center of the cavity and a weaker secondary vortex in each corner. (The secondary vortex in the upper right corner stems from the treatment of the corner singularities, see Appendix A.2.) There is also a tertiary vortex in each lower corner of the cavity, as well as the sign of a very weak quaternary vortex (not visible in the figures) in the lower right corner. The eigenvalue spectra for this case, shown in Fig. 7(a), reveals that the least stable eigenvalue is $\lambda = -3.3681 \times 10^{-3} \pm i\,2.6935$, which corresponds to a temporal frequency of 0.4287. This is comparable with the frequency 0.4496 of the limit cycle at $Re=8018.8$, reported by [3].

A convergence study similar to that presented in Figs. 1(a) and 3(a) suggests that good convergence for this case is obtained with $\Delta t = 10^3$. The computed energy amplification is plotted in Fig. 7(b). Since the flow is so close to the first bifurcation point, all computed eigenvalues are weakly damped, causing a strong energy amplification ranging between 10 and 1000 for every forcing frequency. From the gain curve, several local maxima corresponding to forcing frequencies $\omega \approx \{0.0, 0.91, 1.66, 2.17, 2.69, 3.25, 3.70, 4.34, 4.88\}$ can be identified. For most of these peaks, the bandwidth of the resonance frequency is very narrow, which is common for weakly damped systems. The sharp transition between a peak and a valley is also reflected in the ε-pseudospectrum, plotted in Fig. 7(a). The present pseudospectrum is not evaluated for the full operator, but for the low-dimensional Hessenberg matrix arising from the Arnoldi-factorization used to compute the eigenvalue

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

spectra in Fig. 7(a) [36]. As seen, the ε-pseudospectrum consists of several regions of alternating sensitivity, which roughly coincide with the peaks and valleys of the gain curve in Fig. 7(b).

The largest response is obtained for $\omega=2.6875$, which is very close to the frequency of the least stable eigenvalue. This is the optimal forcing of the system and the amplification at this point is $G(\omega)\approx 1100$. The streamfunction of the corresponding optimal forcing profile is visualized in Figs. 8(a) and 8(b). As can be seen, the profile is localized on the shear layers of the right side of the cavity immediately above the lower secondary vortex and around the primary vortex located in the center of the cavity. Since the velocity of the lid is in the positive x-direction, the primary vortex assumes a clockwise rotation. Hence, the flow response due to this forcing (evaluated by solving (3.9)) is emphasized towards the left side of the cavity, and the shear layers separating the primary vortex from the secondary ones (see Figs. 8(c) and 8(d)). The reason for this strong amplification can be understood by comparing the forcing profile in Fig. 8 to the forcing profile corresponding to $\omega=2.9375$ shown in Fig. 9. This frequency represents a valley in the gain curve in Fig. 7(b), and as can be seen, the profile is entirely localized to the shear layers on the right side of the cavity (the color scale of Figs. 8(a), 8(b), and 9 is the same). Inspection of the forcing profiles corresponding to the other peaks and valleys, shows that most of these (except those corresponding to low frequencies, $\omega<1.0$) resemble the structure visualized in Fig. 9, and attain their maximum on the right side of the cavity immediately above the lower secondary vortex. One may thus conclude that the strong system response around $\omega=2.6875$ is mainly caused by excitation of the primary vortex, which appears to be very receptive to this frequency and insensitive to other driving frequencies.

Furthermore, as was the case for the Ginzburg-Landau equation, Figs. 8(a) and 8(b) have a structure similar to that of the adjoint eigenfunction corresponding to the least stable eigenmode (not shown). It can be shown that in order to optimally excite an eigenmode with eigenvalue λ, the forcing frequency should be $\omega\approx\Im(\lambda)$ and the shape of the forcing function be close to the adjoint eigenfunction corresponding to this mode (see [35]).

The accuracy of these results is estimated a posteriori by substituting the computed eigenpairs into the eigenvalue relation $\|((\mathcal{A}-i\omega\mathcal{I})(\mathcal{A}^\dagger+i\omega\mathcal{I}))^{-1}f^{(k)}-G(\omega)^2f^{(k)}\|$, which can be evaluated using the operators $\mathcal{PA}$, $\mathcal{P}$ and their adjoints. The magnitude of this residual is $\sim 10^{-3}-10^{-5}$ for all the frequencies.

4 Summary and conclusions

We have presented a method for calculating the optimal input for a harmonically forced linear problem, and its resulting flow response. The core of the method is the classic inverse power method applied on the resolvent, which is in turn preconditioned by the inverse Laplacian. The method can readily be implemented by adapting a pre-existing time integration code. It can therefore be used in the same circumstances as time inte-

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

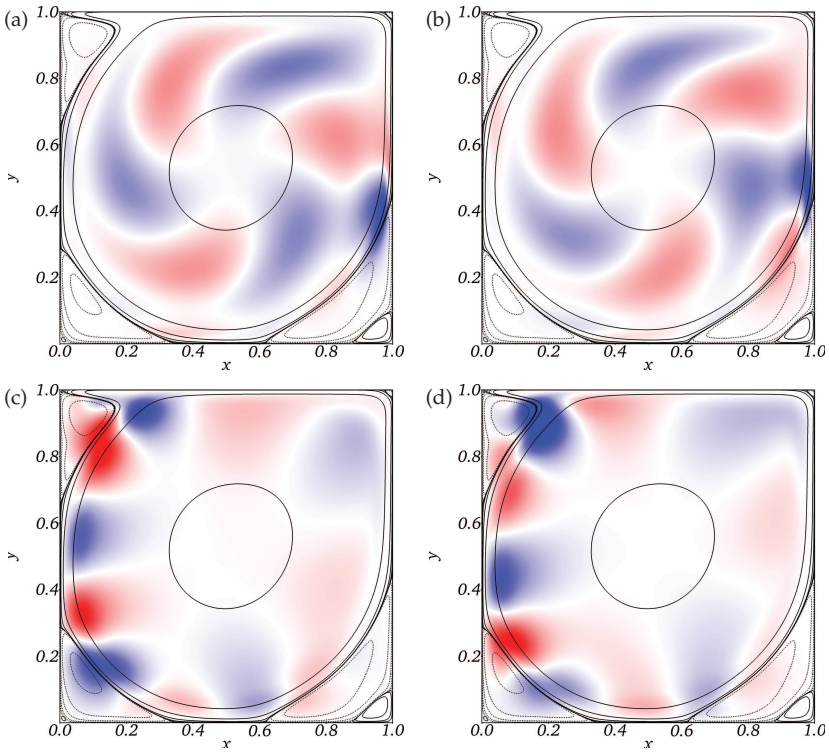


Figure 8: Optimal forcing profile at $\omega = 2.6875$ for the two-dimensional lid-driven cavity, $Re = 8015$. Real and imaginary parts of the optimal forcing profile are shown in (a) and (b), respectively, together with real and imaginary part of the resulting flow response shown in (c) and (d). The color shows the streamfunction of the optimal forcing and response, with red and blue indicating positive and negative values, respectively. The solid and dashed lines represent negative and positive values, respectively, of logarithmically distributed contours of the baseflow streamfunction.

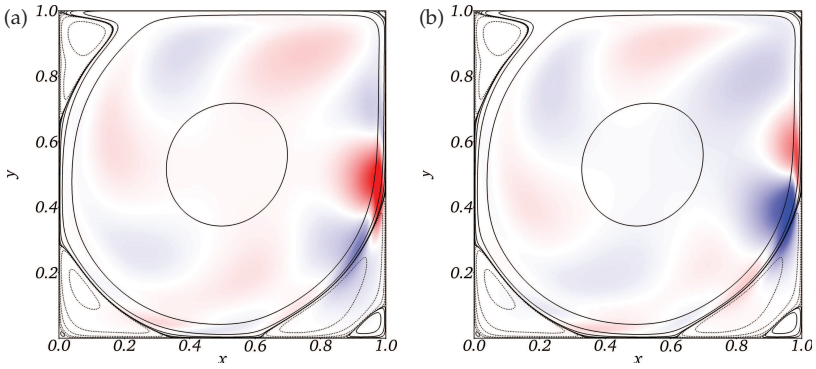


Figure 9: Streamfunction of the forcing profile corresponding to $\omega = 2.9375$ for the two-dimensional lid-driven cavity, $Re = 8015$. Real and imaginary parts of the forcing profile are shown in (a) and (b). The color shows the optimal forcing with red and blue indicating positive and negative values, respectively. The solid and dashed lines represent negative and positive values, respectively, of logarithmically distributed contours of the baseflow streamfunction.

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

gration, and can take advantage of all of the computational advances developed for time integration for various spatial discretizations, including finite differences, finite elements, spectral and spectral elements.

Versions of this method have already been used to calculate steady states and eigenvalues [6, 7, 9, 22, 23, 25, 38, 39]. In this exploratory paper, we have investigated the one-dimensional linear Ginzburg-Landau equation, which is a common model for the Navier-Stokes equations; and the two-dimensional lid-driven cavity flow, which is a simple Navier-Stokes case that is suitable for testing novel computational methods.

In the Ginzburg-Landau case, all operators can be stored, inverted, and diagonalized explicitly. This has enabled us to study the properties and conditioning of the operators in great detail, as well as to validate the results obtained with the proposed method against SVD. With the final goal of studying the optimal forcing of large three-dimensional complex flow problems that are governed by the Navier-Stokes equations, we have implemented the proposed method in the spectral element code `Nek5000` [20]. Since matrices cannot be stored and decomposed for a large flow case, the alternative to SVD that has been used is to integrate the direct and adjoint Navier-Stokes equations in time, which upon convergence yields a stationary point of the corresponding Lagrangian functional. As a proof of concept, the two-dimensional lid-driven cavity flow at $Re=\{100,300,500\}$ is studied. This simple test case shows convergence behavior that is consistent with that of the Ginzburg-Landau equation. A comparison of the optimal forcing computed with the novel method to that obtained through time-integration for $Re=100$ shows that the proposed method gives similar results but requires one to two orders of magnitude fewer operator evaluations. In order to further demonstrate its applicability to higher Reynolds number flows, the forcing of a marginally stable lid-driven cavity at $Re=8015$ is investigated. The results show that the flow is susceptible to a wide range of forcing frequencies, but that the largest energy amplification ($G(\omega)\approx 1100$) is obtained for a harmonic forcing with frequency $\omega=\pm 2.6875$, corresponding to the least stable eigenvalues $\lambda=-3.3681\cdot 10^{-3}\pm i2.6935$. It is shown that this strong energy gain is due to excitation of the primary vortex, which is very receptive to a driving frequency close to the frequency of these eigenvalues.

In order to keep the discussion general, we have deliberately chosen to omit technical details associated with the solution and spatial discretization of the selected model problems. Such issues, and application of the method to larger problems involving complex geometries, will be covered in future publications.

Acknowledgments

We wish to thank Dr. Adam Peplinski, Prof. Paul F. Fischer and Prof. Martin Berggren for valuable discussions and help during the course of this project. We are also grateful for the constructive comments on the manuscript provided by the anonymous referees. Simulations have been performed at the National Supercomputer Centre (NSC) in Swe-

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

den with computer time granted by the Swedish National Infrastructure for Computing (SNIC).

A Test problems

A.1 Ginzburg-Landau equation

The one-dimensional linear Ginzburg-Landau equation reads

$$\frac{\partial q}{\partial t}=\left(-\nu\frac{\partial}{\partial x}+\gamma\frac{\partial^2}{\partial x^2}+\mu(x)\right)q,\quad x\in[-\infty,\infty], \tag{A.1}$$

with boundary conditions $q(x,t)\to 0$ as $x\to\pm\infty$. The corresponding adjoint equation derived with the L^2-inner product is

$$-\frac{\partial g}{\partial t}=\left(\nu^*\frac{\partial}{\partial x}+\gamma^*\frac{\partial^2}{\partial x^2}+\mu^*(x)\right)g,\quad x\in[-\infty,\infty], \tag{A.2}$$

and is subject to boundary conditions $g(x,t)\to 0$ as $x\to\pm\infty$. Eq. (A.1) is of advection-diffusion-reaction type, where the function $\mu(x)$ can be considered to be a spatially-dependent reaction rate. Depending on the form of $\mu(x)$, (A.1) may exhibit stability, or instability [17], which makes it a suitable model for the Navier-Stokes equations.

The reaction term is chosen to be $\mu(x)=(\mu_0-c_u^2)+\mu_2(x^2/2)$ [15]. The coefficient μ_0 is a bifurcation parameter analogous to the Reynolds number and μ_2 determines the degree of non-parallelism. The quadratic form of $\mu(x)$ results in an unstable region bounded by $\pm((-2/\mu_2)(\mu_0-c_u^2))^{1/2}$ (see the shaded regions in Fig. 4(b)). The advection coefficient is $\nu=U+2ic_u$ and the most unstable wave number is c_u. The diffusion coefficient is $\gamma=1+ic_d$, where c_d is the dispersion coefficient. The significance of c_u and c_d can be understood through the dispersion relation $D(k,\omega,\mu_0)=0$, by neglecting the dependence on x and expressing q as a superposition of normal modes (see Bagheri et al. [4] for details). If $\mathrm{d}\mu(x)/\mathrm{d}x\neq 0$, i.e. $\mu_2\neq 0$, and if either $U\neq 0$ or $c_d\neq 0$, the Ginzburg-Landau operator can be shown to be non-normal and to have non-orthogonal eigenfunctions.

In order to compute optimal forcing, a stable flow case is necessary. We have therefore chosen the set of numerical parameters $U=2.0$, $c_u=0.2$, $c_d=-1.0$, $\mu_0=0.38$ and $\mu_2=-0.01$, which corresponds to the sub-critical case of [4]. Eqs. (A.1) and (A.2) are discretized using Hermite polynomials and the differentiation suite provided by Weideman and Reddy [41].

A.2 Navier-Stokes equations (lid-driven cavity flow)

The non-dimensionalized linear incompressible Navier-Stokes equations are given by

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

$$\frac{\partial \mathbf{q}}{\partial t} = -\nabla p - (\bar{\mathbf{q}}\cdot\nabla)\mathbf{q} - (\mathbf{q}\cdot\nabla)\bar{\mathbf{q}} + Re^{-1}\nabla^2\mathbf{q}, \tag{A.3a}$$

$$\nabla\cdot\mathbf{q} = 0, \tag{A.3b}$$

where the perturbation field is denoted $\mathbf{q}=(u,v)^T$, and the steady base flow around which the stability analysis is performed is denoted $\bar{\mathbf{q}}=(\bar{u},\bar{v})^T$. By following the Helmholtz-Hodge decomposition theorem [16] and introducing an orthogonal projector $\mathscr{P}_\nabla(\cdot)$ onto a divergence-free field, (A.3a) may be rewritten schematically as

$$\frac{\partial \mathbf{q}}{\partial t} = \mathscr{P}_\nabla\left(-(\bar{\mathbf{q}}\cdot\nabla)\mathbf{q} - (\mathbf{q}\cdot\nabla)\bar{\mathbf{q}}\right) + \mathscr{P}_\nabla\left(Re^{-1}\nabla^2\mathbf{q}\right), \tag{A.4}$$

which brings (A.3) into the same form as (2.1). In addition to the operators defined by (A.4), we also consider the adjoint counterparts defined in a similar fashion as

$$-\frac{\partial \mathbf{g}}{\partial t} = \mathscr{P}_\nabla\left((\bar{\mathbf{q}}\cdot\nabla)\mathbf{g} - (\nabla\bar{\mathbf{q}})^T\mathbf{g}\right) + \mathscr{P}_\nabla\left(Re^{-1}\nabla^2\mathbf{g}\right) \tag{A.5}$$

(see [5] for details on the derivation of the adjoint Navier-Stokes equations).

As a test case, the two-dimensional lid-driven square cavity is considered. The method is incorporated into the Navier-Stokes solver `Nek5000` [20], which is based on the spectral element method [29] and uses the $\mathbb{P}_N-\mathbb{P}_{N-2}$ discretization for velocity and pressure [24]. The base flow obeys homogeneous Dirichlet conditions on all the boundaries except at the top boundary, where the y-component is homogeneous and the x-component is given by

$$\bar{u}(\tilde{x}) = \begin{cases} 1, & \text{if } |\tilde{x}| \le (1/2 - 1/x_{rise}), \\ \left(1+\exp\left(\frac{1}{x_{rise}(1/2-|\tilde{x}|)-1} + \frac{1}{x_{rise}(1/2-|\tilde{x}|)}\right)\right)^{-1}, & \text{if } (1/2-1/x_{rise}) < |\tilde{x}| < (1/2-\epsilon), \\ 0, & \text{otherwise,} \end{cases} \tag{A.6}$$

in which $\tilde{x}=x-1/2$ ($x\in[0,1]$). Eq. (A.6) represents a two-sided symmetric step function [13] used to treat the singularities that arise in the top corners due to the discontinuous boundary conditions (see e.g. [3] for another treatment of this issue). In order to have a smooth velocity distribution along the boundaries, the tuning parameters of (A.6) are chosen to be $x_{rise}=6.0$, and $\epsilon=0.001$. For the perturbation, homogeneous Dirichlet conditions are imposed on all the boundaries for $\mathbf{q}$ and $\mathbf{g}$. With these boundary conditions, $\mathcal{P}(\Delta t)$ as defined in (2.5) will be self-adjoint.

B Derivation of the optimal forcing time integration method

Consider a dynamical system such as (1.2), $\partial q/\partial t = \mathcal{A}q + fe^{i\omega t}$, and introduce the ansatz function $q(x,t)=s(x)e^{i\omega t}$. The Lagrangian governing the problem of determining the

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

optimal forcing reads [26]

$$\mathscr{L}(s,b,\alpha,f)=\langle s,s\rangle-\Re\{\langle b,(i\omega\mathcal{I}-\mathcal{A})s-f\rangle\}-\alpha(\langle f,f\rangle-1), \tag{B.1}$$

where $\langle s,s\rangle$ is the energy of the response (the objective to be maximized), and the other two terms represent linear equality constraints that must be satisfied by a feasible solution, i.e. the solution must satisfy the governing equation, and the force should have unit norm. The variables $b(x)$ and α are Lagrange multipliers of these constraints, and we assume that $\omega,\alpha\in\mathbb{R}$ and $s(x),b(x),f(x)\in\mathbb{C}$. An optimal solution necessarily has to be a stationary point of (B.1) in order to fulfill the first order optimality conditions. This implies that the first variations of (B.1) with respect to its arguments must vanish simultaneously,

$$\langle\delta\mathscr{L},\delta s\rangle=\Re\left\{\langle 2s+(i\omega\mathcal{I}+\mathcal{A}^\dagger)b,\delta s\rangle\right\}=0 \qquad \forall\delta s, \tag{B.2a}$$
$$\langle\delta\mathscr{L},\delta b\rangle=\Re\{\langle(i\omega\mathcal{I}-\mathcal{A})s-f,\delta b\rangle\}=0 \qquad \forall\delta b, \tag{B.2b}$$
$$\langle\delta\mathscr{L},\delta\alpha\rangle=\delta\alpha(\langle f,f\rangle-1)=0 \qquad \forall\delta\alpha, \tag{B.2c}$$
$$\langle\delta\mathscr{L},\delta f\rangle=\Re\{\langle b-2\alpha f,\delta f\rangle\}=0 \qquad \forall\delta f, \tag{B.2d}$$

which gives

$$2s=-(\mathcal{A}^\dagger+i\omega\mathcal{I})b, \tag{B.3a}$$
$$f=(-\mathcal{A}+i\omega\mathcal{I})s, \tag{B.3b}$$
$$\langle f,f\rangle=1, \tag{B.3c}$$
$$b=2\alpha f. \tag{B.3d}$$

These equations represent in turn the adjoint and the direct equation, the normalization and the optimality condition. From (B.3) the different steps of Algorithm 2 follows. Substituting (B.3a) and (B.3d) into (B.3b) gives

$$\Big((\mathcal{A}-i\omega I)(\mathcal{A}^\dagger+i\omega I)\Big)^{-1}f=\alpha f, \tag{B.4}$$

where the operator on the left-hand side is equal to (3.4), and $\alpha=G(\omega)^2$.

References

[1] E. Åkervik, U. Ehrenstein, F. Gallaire, and D. S. Henningson. Global two-dimensional stability measures of the flat plate boundary-layer flow. *Eur. J. Mech. B-Fluids*, 27(5):501–513, 2008.

[2] F. Alizard, S. Cherubini, and J.-C. Robinet. Sensitivity and optimal forcing response in separated boundary layer flows. *Phys. Fluids*, 21(6), 2009.

[3] F. Auteri, N. Parolini, and L. Quartapelle. Numerical investigation on the stability of singular driven cavity flow. *J. Comput. Phys.*, 183(1):1–25, 2002.

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

[4] S. Bagheri, D. S. Henningson, J. Hœpffner, and P. J. Schmid. Input-output analysis and control design applied to a linear model of spatially developing flows. *Appl. Mech. Rev.*, 62:020803, 2009.
[5] D. Barkley, H. M. Blackburn, and S. J. Sherwin. Direct optimal growth analysis for timesteppers. *Int. J. Numer. Meth. Fl.*, 57(9):1435–1458, 2008.
[6] D. Barkley and L. S. Tuckerman. Stokes preconditioning for the inverse power method. In P. Kutler, J. Flores, and J.-J. Chattot, editors, *Lecture Notes in Physics: Proc. of the Fifteenth Int'l. Conf. on Numerical Methods in Fluid Dynamics*, pages 75–76. Springer, New York, 1997.
[7] O. Batiste, E. Knobloch, A. Alonso, and I. Mercader. Spatially localized binary-fluid convection. *J. Fluid Mech.*, 560:149–158, 2006.
[8] C. Beaume. Adaptive Stokes preconditioning for steady incompressible flows. *Commun. Comput. Phys.*, 22(2):494–516, 2017.
[9] A. Bergeon, D. Henry, H. BenHadid, and L. S. Tuckerman. Marangoni convection in binary mixtures with Soret effect. *J. Fluid Mech.*, 375:143–177, 1998.
[10] K. M. Butler and B. F. Farrell. Three-dimensional optimal perturbations in viscous shear flow. *Phys. Fluids A*, 4(8):1637–1650, 1992.
[11] S. L. Campbell, I. C. F. Ipsen, C. T. Kelley, and C. D. Meyer. GMRES and the minimal polynomial. *BIT Numer. Math.*, 36(4):664–675, 1996.
[12] C. Canuto, Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods in Fluid Dynamics*. Springer Series in Computational Physics. Springer Berlin Heidelberg, 1988.
[13] M. Chevalier, P. Schlatter, A. Lundbladh, and D. S. Henningson. SIMSON – A Pseudo-Spectral Solver for Incompressible Boundary Layer Flows. Technical Report TRITA-MEK 2007:07, KTH Mechanics, 2007.
[14] J. M. Chomaz. Global instabilities in spatially developing flows: non-normality and nonlinearity. *Annu. Rev. Fluid Mech.*, 37:357–392, 2005.
[15] J. M. Chomaz, P. Huerre, and L. G. Redekopp. Models of hydrodynamic resonances in separated shear flows. In *Proceedings of the 6th Symp. Turb. Shear Flows*, September 1987.
[16] A. J. Chorin and J. E. Marsden. *A Mathematical Introduction to Fluid Mechanics*. Texts in Applied Mathematics. Springer New York, 2000.
[17] C. Cossu and J. M. Chomaz. Global measures of local convective instabilities. *Phys. Rev. Lett.*, 78:4387–4390, Jun 1997.
[18] T. A. Driscoll, K. C. Toh, and L. N. Trefethen. From potential theory to matrix iterations in six steps. *SIAM Rev.*, 40(3):547–578, 1998.
[19] M. Embree. How descriptive are GMRES convergence bounds? Technical Report 99/08, Oxford University Computing Laboratory, June 1999.
[20] P. F. Fischer, J. W. Lottes, and S. G. Kerkemeier. nek5000 Web page, 2008. http://nek5000.mcs.anl.gov.
[21] A. Greenbaum, V Pták, and Z. Strakoš. Any nonincreasing convergence curve is possible for GMRES. *SIAM J. Matrix Anal. Appl.*, 17(3):465–469, 1996.
[22] C. Huepe, S. Métens, P. Borckmans, and M. E. Brachet. Decay rates in attractive Bose-Einstein condensates. *Phys. Rev. Lett.*, 82(8):1616–1619, 1999.
[23] C. Huepe, L.S. Tuckerman, S. Métens, and M. E. Brachet. Stability and decay rates of non-isotropic attractive Bose-Einstein condensates. *Phys. Rev. A*, 68:023609, 2003.
[24] Y. Maday, A. T. Patera, and E. M. Rønquist. The $\mathbb{P}_N \times \mathbb{P}_{N-2}$ method for the approximation of the Stokes problem. Technical report, Department of Mechanical Engineering, MIT, Cambridge, MA.
[25] C. K. Mamun and L. S. Tuckerman. Asymmetry and Hopf bifurcation in spherical Couette

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press

flow. *Phys. Fluids*, 7(1):80–91, 1995.
[26] A. Monokrousos, E. Åkervik, L. Brandt, and D. S. Henningson. Global three-dimensional optimal disturbances in the Blasius boundary-layer flow using time-steppers. *J. Fluid Mech.*, 650:181–214, 5 2010.
[27] M. L. Parks, E. de Sturler, G. Mackey, D. D. Johnson, and S. Maiti. Recycling Krylov subspaces for sequences of linear systems. *SIAM J. Sci. Comput.*, 28(5):1651–1674, 2006.
[28] B. Parlett. *The Symmetric Eigenvalue Problem*. Society for Industrial and Applied Mathematics, 1998.
[29] A. T. Patera. A spectral element method for fluid dynamics: Laminar flow in a channel expansion. *J. Comput. Phys.*, 54(3):468 – 488, 1984.
[30] S. C. Reddy, P. J. Schmid, J. S. Baggett, and D. S. Henningson. On stability of streamwise streaks and transition thresholds in plane channel flows. *J. Fluid Mech.*, 365:269–303, 1998.
[31] S. C. Reddy, P. J. Schmid, and D. S. Henningson. Pseudospectra of the Orr–Sommerfeld operator. *SIAM J. Appl. Math.*, 53(1):15–47, 1993.
[32] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7(3):856–869, 1986.
[33] D. Sipp. Open-loop control of cavity oscillations with harmonic forcings. *J. Fluid Mech.*, 708:439–468, 2012.
[34] D. Sipp and O. Marquet. Characterization of noise amplifiers with global singular modes: the case of the leading-edge flat-plate boundary layer. *Theor. Comp. Fluid Dyn.*, 27(5):617–635, 2013.
[35] D. Sipp, O. Marquet, P. Meliga, and A. Barbagallo. Dynamics and control of global instabilities in open-flows: a linearized approach. *Appl. Mech. Rev.*, 63(3):030801, 2010.
[36] K. C. Toh and L. N. Trefethen. Calculation of pseudospectra by the Arnoldi iteration. *SIAM J. Sci. Comp.*, 17(1):1–15, 1996.
[37] L. N. Trefethen and M. Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, 2005.
[38] L. S. Tuckerman. Laplacian preconditioning for the inverse Arnoldi method. *Commun. Comput. Phys.*, 18:1336–1351, 2015.
[39] L. S. Tuckerman and D. Barkley. Bifurcation analysis for timesteppers. In E. Doedel and L. S. Tuckerman, editors, *Numerical Methods for Bifurcation Problems and Large-Scale Dynamical Systems*. Springer, New York, 2000.
[40] H. van der Vorst. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 13(2):631–644, 1992.
[41] J. A. Weideman and S. C. Reddy. A MATLAB differentiation matrix suite. *ACM Trans. Math. Software*, 26(4):465–519, 2000.

https://doi.org/10.4208/cicp.OA-2016-0070 Published online by Cambridge University Press