Science and Technology Facilities Council

Review of adjoint method in plasma physics for NEPTUNE

Technical Report 2068625-TN-08
Deliverable 5.1

Josh Williams* Ubaid Ali Qadri† Sue Thorne‡

July 2023 (Revised September 2023)

1 Introduction

Tokamak reactors aim to use confined plasma to create energy by nuclear fusion [57]. Due to the complex combination of interactions at multiple length-scales, flow instabilities can arise that cause reduced performance and damage to the reactor's walls [41, 48]. The interplay between multiple interacting design and operating parameters creates a difficult optimisation problem to maximise power output of future tokamak reactors. These parameters are also likely to have some degree of uncertainty, which will need to be quantified by future users of software in the NEPTUNE codebase. Additionally, the complex plasma flow is prone to instabilities that limit the confinement achieved by tokamaks. Identifying the onset of instabilities and developing passive/active control strategies is therefore crucial for the success of nuclear fusion reactors. The adjoint method has been used extensively in stability analysis, uncertainty quantification, optimisation and control problems [37, 43, 64]. Sensitivity information obtained from the adjoint method is also useful for adaptive mesh refinement [4]. This report aims to review the adjoint method to determine how it may be used in optimising plasma flows within the NEPTUNE project.
This report is structured as follows. In Section 2, we present the background and theory of the adjoint method. We then compare different methods of computing the adjoint variables, which can be done using a continuous or discrete adjoint method (Section 2.1). We discuss some recent advancements in algorithms used to solve the adjoint equations in Section 2.2. In Section 3, we review a variety of applications of the adjoint method in the literature that are relevant to NEPTUNE. This includes (i) sensitivity and stability analysis, (ii) parameter optimisation and sensitivity, (iii) uncertainty quantification, (iv) adaptive mesh refinement and (v) shape optimisation. Finally, in Section 4, we draw our conclusions.

2 Adjoint method

The adjoint method computes the sensitivity of a cost function (based on a quantity of interest) relative to a change in any relevant parameters (operating conditions, design parameters, modelling parameters) [42]. This information can be used to optimise parameters at a significantly lower cost than iteratively performing simulations in a non-intrusive way.
As explained by Giles and Pierce [29], the theory behind the adjoint method is easiest understood in the context of a linear problem. In the series of problems discussed in this report, we are aiming to find the change in some cost function (or quantity of interest), $\mathcal{J}$, relative to a small perturbation in the operating

*Josh Williams is with the Hartree Centre, STFC Royal Observatory Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK.
†Ubaid Ali Qadri is with the Hartree Centre, STFC Daresbury Laboratory, Daresbury Science and Innovation Campus, Warrington, Cheshire WA4 4AD. Email contact: ubaid.qadri@stfc.ac.uk
‡Sue Thorne is with the Hartree Centre, STFC Rutherford Appleton Laboratory, Harwell Campus, Didcot, OX11 0QX, UK. Email contact: sue.thorne@stfc.ac.uk

parameters $\boldsymbol{p}$ ($d\mathbf{J}/d\boldsymbol{p}$) which will change the flow field. The state vector containing different flow field variables is $\boldsymbol{q}$, the perturbed state vector is $\boldsymbol{q}^{'}$, and the perturbed cost function is

$$\boldsymbol{g}^T\boldsymbol{q}^{'} = \frac{\partial \mathcal{J}}{\partial \boldsymbol{q}}\boldsymbol{q}^{'}, \tag{1}$$

where $\boldsymbol{g}$ is a vector describing the cost function gradient.
Therefore, we are aiming to find $\boldsymbol{g}^T\boldsymbol{q}^{'}$, where the perturbed state vector is described by the linear problem

$$\boldsymbol{L}\boldsymbol{q}^{'} = \boldsymbol{f} = \frac{\partial \boldsymbol{q}^{'}}{\partial t}, \tag{2}$$

where $\boldsymbol{L}$ is a linear operator, and $\boldsymbol{f}$ is a vector acting as a source term. The dual of this problem being

$$\boldsymbol{L}^T\boldsymbol{q}^{+} = \boldsymbol{g}, \tag{3}$$

where $\boldsymbol{q}^{+}$ is the state vector containing the adjoint solutions.
The equivalence between (2) and (3) is

$$\left(\boldsymbol{q}^{+}\right)^T\boldsymbol{f} = \left(\boldsymbol{q}^{+}\right)^T\boldsymbol{L}\boldsymbol{q}^{'} = \left(\boldsymbol{L}^T\boldsymbol{q}^{+}\right)^T\boldsymbol{q}^{'} = \boldsymbol{g}^T\boldsymbol{q}^{'}, \tag{4}$$

which is simply

$$\left(\boldsymbol{q}^{+}\right)^T\boldsymbol{f} = \boldsymbol{g}^T\boldsymbol{q}^{'}. \tag{5}$$

The sensitivity of $\mathcal{J}$ to a set of parameters can be found by the chain rule [43]

$$\frac{d\mathcal{J}}{d\boldsymbol{p}} = \frac{\partial \mathcal{J}}{\partial \boldsymbol{p}} + \frac{\partial \mathcal{J}}{\partial \boldsymbol{q}}\frac{d\boldsymbol{q}}{d\boldsymbol{p}}. \tag{6}$$

The issue is that computing $d\boldsymbol{q}/d\boldsymbol{p}$ is very expensive when the dimension of $\boldsymbol{p}$ is large. Hence, the challenge is then how to compute this without directly solving $d\boldsymbol{q}/d\boldsymbol{p}$ for all values of $\boldsymbol{p}$.
We know that the state vector must satisfy some governing equations that vary in space based on the flow parameters ($\boldsymbol{\mathcal{R}}(\boldsymbol{q},\boldsymbol{p}) = 0$), where $\boldsymbol{\mathcal{R}}$ is the residual of the governing equations. The term $d\boldsymbol{q}/d\boldsymbol{p}$ should also satisfy

$$\frac{d\boldsymbol{\mathcal{R}}}{d\boldsymbol{p}} = \frac{\partial \boldsymbol{\mathcal{R}}}{\partial \boldsymbol{p}} + \frac{\partial \boldsymbol{\mathcal{R}}}{\partial \boldsymbol{q}}\frac{d\boldsymbol{q}}{d\boldsymbol{p}} = 0. \tag{7}$$

In the explanation of Giles and Pierce [29], (7) is related to the linear variables in (4) by

$$\boldsymbol{q}^{'} = \frac{d\boldsymbol{q}}{d\boldsymbol{p}}, \tag{8}$$

$$\boldsymbol{L} = \frac{\partial \boldsymbol{\mathcal{R}}}{\partial \boldsymbol{q}}, \tag{9}$$

$$\boldsymbol{g}^T = \frac{\partial \mathcal{J}}{\partial \boldsymbol{q}}, \tag{10}$$

$$\boldsymbol{f} = -\frac{\partial \boldsymbol{\mathcal{R}}}{\partial \boldsymbol{p}}, \tag{11}$$

which can be combined with (1) and (4) to give

$$\frac{d\mathcal{J}}{d\boldsymbol{p}} = \frac{\partial \mathcal{J}}{\partial \boldsymbol{p}} + \boldsymbol{g}^T\boldsymbol{q}^{'} \tag{12}$$

$$= \frac{\partial \mathcal{J}}{\partial \boldsymbol{p}} + \left(\boldsymbol{q}^{+}\right)^T\boldsymbol{f}. \tag{13}$$

For multiple design/operating parameters ($\boldsymbol{p}$), we note that each design variable has a different $\boldsymbol{f}$ but the same $\boldsymbol{g}$**.** Therefore, instead of solving multiple equations of the form $\boldsymbol{L}\boldsymbol{q}^{'} = \boldsymbol{f}$ and using (12), the computational cost can be significantly reduced by solving the single equation $\boldsymbol{L}^T\boldsymbol{q}^{+} = \boldsymbol{g}$ to compute $\boldsymbol{q}^{+}$ and then use (13), which is known as the adjoint method.

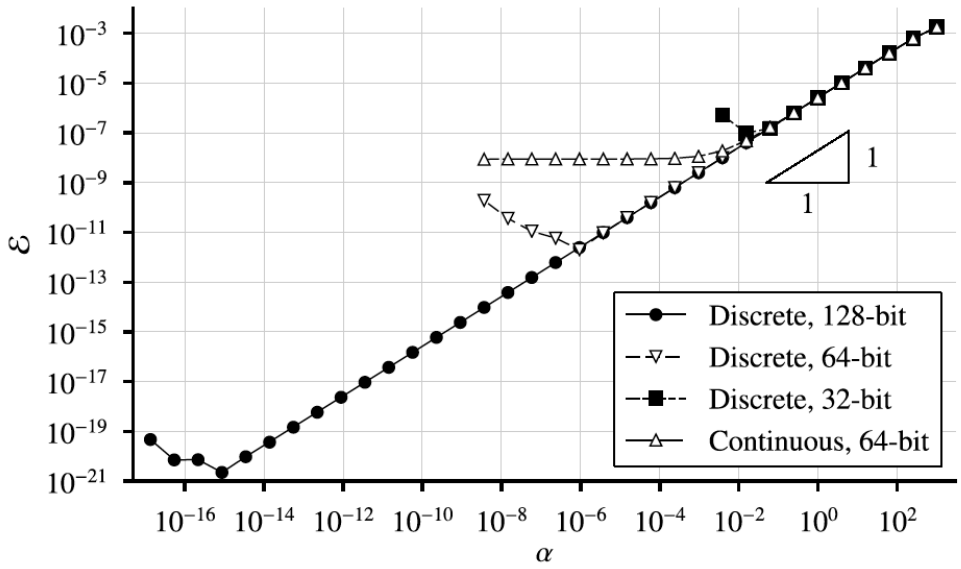

Figure 1: Gradient error, ε, compared to step-size, α, for discrete adjoint with varying floating point precision, and continuous adjoint (64-bit precision) [77].

2.1 Continuous or discrete adjoint method

The gradients can be obtained by (i) the continuous method, where the user derives the derivatives of the governing equations analytically prior to discretisation or (ii) the discrete method, where the derivatives are calculated after the governing equations are discretised [54], which is simpler to implement [43]. In the latter, the exact gradients can be computed to machine precision [43]. Since the discrete adjoint method computes derivatives from the discretised governing equations, we can ensure no inconsistencies are introduced to the simulations during discretisation. This property is extremely attractive for chaotic turbulent flows [9] (such as might be found in tokamak reactors), since minor inaccuracies can propagate/amplify to large errors [38, 77]. In aerodynamic optimisation, it has been shown that the continuous adjoint method approaches the exact gradients when the mesh is sufficiently resolved (as discretisation error decreases) [52, 53]. However, for aeroacoustic control of a compressible flow mixing layer, it was shown by Vishnampet *et al.* that the continuous adjoint gradient error saturated well above the error of numerical precision (continuous adjoint error was $\mathcal{O}(10^{-8})$ compared to $\mathcal{O}(10^{-12})$ for discrete adjoint, both with 64 bit precision, Figure 1) [77]. This may be attributed to the complexity and chaotic nature of the compressible flow which amplifies errors, or the low amplitude of the pressure fluctuations responsible for aeroacoustic noise creating gradients that are highly susceptible to discretisation errors [77]. Importantly, the authors mention that even errors introduced by the numerical imprecision in the discrete adjoint will eventually accumulate to corrupt the solution for longer simulations due to the chaotic nature of turbulence [77]. In stellarator parameter optimisation, the error from the continuous adjoint was found to be around one order of magnitude larger than the discrete adjoint due to discretisation error [60]. However, the authors state that the continuous adjoint error is still very small ($\approx 0.1\%$) and acceptable for their application. This suggests that discretisation errors in the continuous adjoint are broadly negligible except in extremely chaotic flows.
Comparing the discrete and continuous adjoint methods, Chandler [11] showed qualitatively the same global modes in sensitivity analysis of non-isothermal low density jets, see Figure 2. However, Chandler's results showed some numerical artefacts for the discrete adjoint method that were not present in the continuous adjoint formulation. Chandler also reported that the continuous adjoint formulation converged faster than the discrete adjoint but with increased debugging issues.

2.2 Adjoint algorithm developments

Although the adjoint equation is significantly less expensive than computing sensitivities with finite differences, in-practice it can be computationally expensive to solve. Here we review recent algorithmic

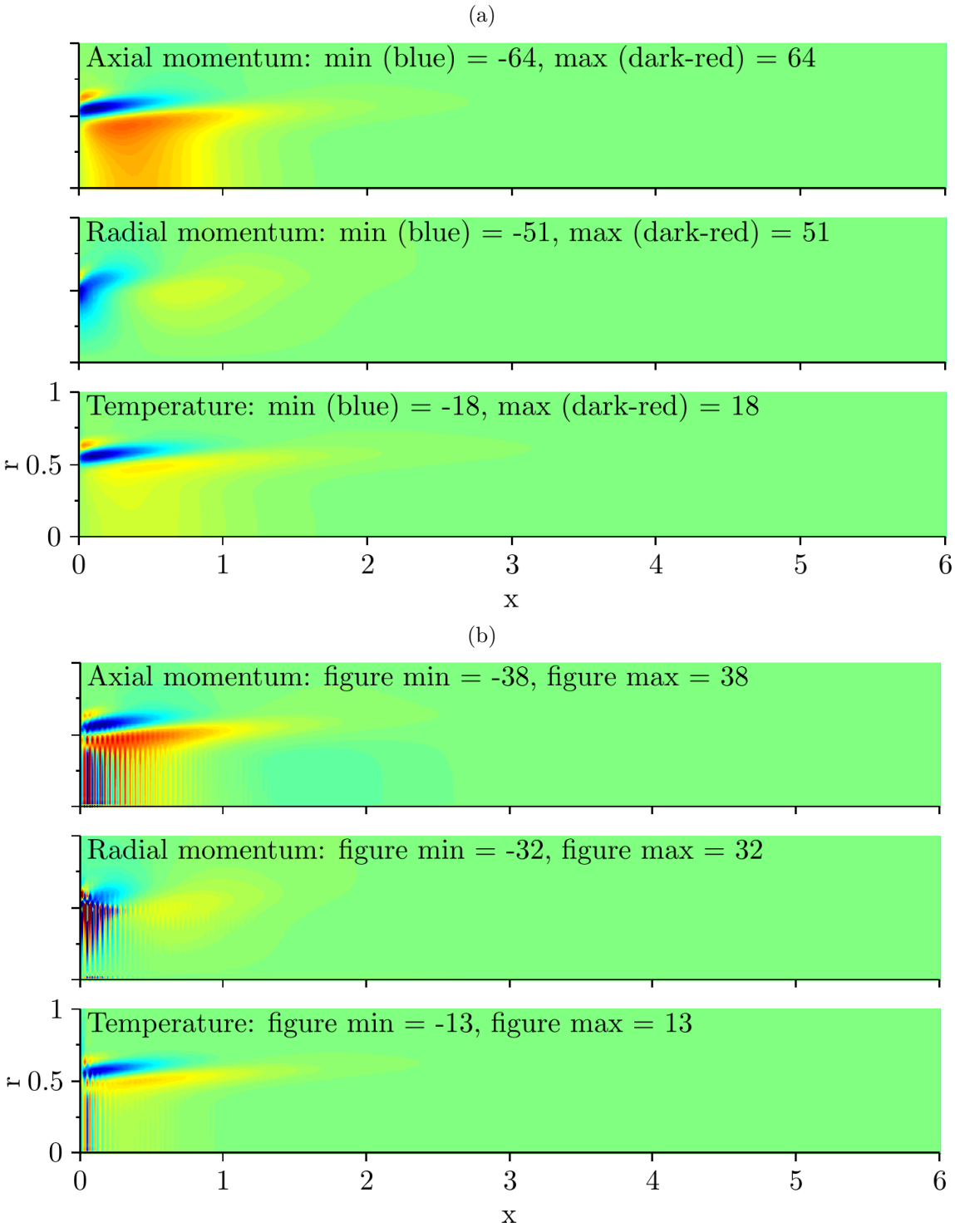

Figure 2: Real part of most unstable global mode from a (a) continuous adjoint and (b) discrete adjoint, in simulations performed by Chandler [11].

developments improving the ease of implementing and solving adjoint problems.

2.2.1 Parallel-in-time

The iterative nature of forward-backward computations in the adjoint method can create a computationally expensive process. Recently, a parallel-in-time algorithm for linear [23] and non-linear [24] initial-value problems was extended to improve performance of adjoint optimisation [14, 71]. The original algorithm decomposes a linear initial-value problem into a homogeneous and inhomogeneous part. The homogeneous equation can be integrated with significantly greater speed and efficiency. The inhomogeneous part (usually dependent on a source term) is decomposed in time and distributed to each processor [23]. The parallel-in-time adjoint method was tested on a forced advection-diffusion equation (linear) and harmonically-forced Burgers equation (non-linear) [71]. This method was recently extended to solve the adjoint equation in problems governed by the incompressible Navier-Stokes equations [14]. However, it is not yet clear if this approach is suitable for multiscale, coupled problems such as plasma transport in tokamaks. Furthermore, the parallel-in-time approach would require significant modification of the codes developed throughout the NEPTUNE project thus far. Another ExCALIBUR-funded project is currently working on advancing parallel-in-time algorithms and determining optimal problem-specific parallel-in-time approaches for different applications such as weather and fusion modelling [22], which may be useful in the future.

2.2.2 Checkpointing unsteady adjoint problems

One major challenge in using adjoint optimisation of time-dependent problems is that the adjoint equation requires knowledge of a state variable computed from the primal solution, u, and its adjoint, u^*, at each time step [56]. For long time series, this is not feasible as the full time-series of the primal solution must either be (1) stored in memory and accessed at each step of the adjoint calculation, which may not be possible for simulations with a long time series or many degrees of freedom [30], or (2) written to the disk and read at each step, which is prohibitively slow. To counteract this, checkpointing algorithms can be used such as `revolve` [30], which saves `checkpoints` at a pre-defined time interval (based on available memory and number of timesteps), and the intermediate timesteps are recomputed. For simulations where the number of timesteps is not known *a priori*, an adaptive checkpointing scheme must be used (referred to as 'online' checkpointing) [33, 74]. A checkpointing algorithm has been implemented in the spectral element code Nek5000 [66]. The authors also improved the standard `revolve` algorithm to include disk I/O checkpoint (two-level checkpointing), which significantly reduced the number of recomputed forward steps in simulations with a large number of timesteps [66]. The lower number of recomputation is only beneficial if the I/O time is not prohibitively large [66]. The I/O time can be reduced by using efficient file formats such as HDF5 [79]. A two-level checkpointing algorithm based on `revolve` has been implemented in PETSc: the implementation was described in [45, 80].
Recently, an alternative method to checkpointing was proposed that uses the incremental singular value decomposition (iSVD) to store a reduced-order model of the primal solution [39, 76]. In this method, snapshots of the flow are added to the SVD at each time step. Since the rank of the SVD is reduced, only storing the main modes of the flow, the memory consumption is significantly reduced and this gives improved performance (CPU time) compared to checkpointing with the same memory consumption. In contrast to the checkpointing method where time-instants of an exact primal field are stored in memory and the remaining time-series is re-computed (sacrificing CPU time), using iSVD does not require expensive re-computations since an approximation of the primal's full time-series can be stored. This method obviously introduces some error due to the reduced amount of information contained in the primal solution but it has been shown that this has very little influence on the objective function [76]. For example, when performing a shape optimisation study, Vezyris *et al.* found the iSVD approach to be within 0.01% of the objective function obtained with checkpointing but with a 29% lower CPU time and 0.5% of the memory overhead [76]. Reduced-order modelling can be leveraged using the recently developed libROM library from Lawrence Livermore National Lab [12] which has an iSVD implementation. However, there is little work published using libROM and, therefore, its performance level is unknown.

3 Applications of adjoint method relevant to plasma modelling

The adjoint method has been applied to many problems in fluid mechanics. Many of the studies discussed crossover multiple applications because topics such as parameter optimisation often make use of the adjoint method's ability to interpret sensitivity and stability properties for understanding optimisation results. The adjoint sensitivity field is also a highly useful error indicator for adaptive mesh refinement. In this section, we will discuss various applications of the adjoint method in the literature.

3.1 Sensitivity and stability analysis

A suitable toolkit to extract information on the response of a base flow to a perturbation is stability, receptivity and sensitivity analysis [68]. The stability can be computed by obtaining the eigenvalues of the direct problem, which is done by first expressing the perturbed state-vector using a Fourier decomposition

$$\boldsymbol{q}^{'} = e^{\lambda t}\hat{\boldsymbol{q}} \tag{14}$$

where the notation $\hat{\cdot}$ represents a complex vector field that describes the spatial structure of the perturbation (eigenvectors of $\boldsymbol{L}$), and $\lambda = \sigma + i\omega$ is a complex scalar describing the temporal evolution of the perturbation (eigenvalues of $\boldsymbol{L}$), with a growth rate, ω, and a frequency, ω. By substituting (14) into (2), we arrive at the eigenvalue problem

$$\boldsymbol{L}\hat{\boldsymbol{q}} = \lambda\hat{\boldsymbol{q}}, \tag{15}$$

which represents the 'direct global modes'. If $\sigma > 0$ (or $\mathrm{Re}(\lambda) > 0$) for any single mode, the flow is asymptotically unstable [70]. These instabilities can be absolute (growing from a fixed point in space) or convective (growing in magnitude but propagating in space) [34]. The growth of convective instabilities was the subject of several works by Blackburn *et al.* [5–7]. In stability problems, $\mathcal{J}$ is based on the amplification of energy (energy at time T compared to energy in the base flow) [67].
From the adjoint problem, we obtain the receptivity, which represents the flow's response to (external) open-loop forcing [43]. Similar to the direct method (15), the adjoint global modes can be expressed as [47]

$$\boldsymbol{L}^{+}\hat{\boldsymbol{q}}^{+} = \lambda^{+}\hat{\boldsymbol{q}}^{+}, \tag{16}$$

where $\boldsymbol{L}^{+}$ is the adjoint of $\boldsymbol{L}$. When both the direct and adjoint global modes are stable but short-term instabilities are present, this suggests that the direct and adjoint operators are non-normal [70].
The ('structural') sensitivity is then computed as the product of the stability and receptivity, and can be used to obtain information on instabilities such as the location of a 'wavemaker' [28]. The wavemaker identifies the flow region where a change in problem's structure would produce the largest eigenvalue drift. This structural sensitivity is computed simply as $\boldsymbol{S} = \hat{\boldsymbol{q}}\hat{\boldsymbol{q}}^{+}$ [62]. Qadri used the Frobenius norm of the tensor to identify the wavemaker region but stated that each component of the tensor should be visualised to better understand the flow's feedback mechanism to perturbations [62].
Giannetti and Luchini [28] studied the wavemaker region in a low Reynolds number flow over a cylinder near the critical Reynolds number ($Re_{crit} \approx 47$), where the steady structure of the flow is broken. They compared the structural sensitivity field with previously reported experimental data [73] that showed the placement of a small control cylinder a few diameters downstream of the wake can completely suppress the onset of unsteadiness. The exact recommended location of the control cylinder was identified as being within the separation bubble from the structural sensitivity, providing a theoretical explanation for the instability suppression observed experimentally that was not possible from analysing the adjoint or direct modes alone [28]. This approach has been used in many studies since the original development [28], including an extension to 3D flow over a cylinder with a periodic base flow [27] and other geometries such as flow over an open cavity [13]. Citro *et al.* [13] considered stability analysis of flow over an open cavity. They evaluated adjoint eigenmodes to show the region where an instability is most receptive to momentum forcing (Figure 3 right). They then computed the product of the direct and adjoint fields to find the structural sensitivity and the wavemaker that followed a streamline around the cavity (Figure 3 bottom). Clearly, the sensitivity field is strongest where the adjoint and direct modes overlap. Murali [49] studied hydrodynamic and magnetohydrodynamic (MHD) flow in cooling blanket ducts with repeated wedge-shaped protrusions, relevant for fusion reactors. They computed the structural sensitivity and growth rate of perturbations in the flow [51], this was then used to later study the generation of quasi-2D vortices to enhance heat transfer within the duct at various Reynolds numbers [49, 50]. The use of the

adjoint method to evaluate receptivity and (when combined with the direct modes) structural sensitivity is relevant to the study of fusion reactors as it can be used to evaluate areas for passive or active controls to prevent instabilities that would hinder the confinement [72].
Several works have been done recently using the adjoint method for sensitivity analysis of stellarator-based problems [2, 60]. Paul *et al.* used the adjoint method to evaluate the sensitivity of ion particle flux and bootstrap current to changes in the Fourier amplitudes of the magnetic field strength [60]. The same group also studied shape gradients of MHD equilibria within a stellarator [2], and the sensitivity of magnetic island size to magnetic field variations [26]. Similar methods have recently been applied, where Nies *et al.* [55] derived a shape-gradient-like sensitivity function to quantify quasi-symmetry on a flux surface, with the aim of minimising quasi-symmetry-breaking modes in the flow. This is similar to the study of Murali on quasi-2D vortices breaking into full 3D flows in cooling blanket ducts of fusion reactors [49, 50].

3.2 Parameter optimisation and sensitivity

When experimental data is available, the adjoint method has been used to tune empirically-determined parameters in plasma edge-physics models [3]. The authors used the adjoint method to attempt to optimise coefficients of a simple transport model with neutral density, ion density and temperature profiles obtained from high fidelity simulations (performed with B2-EIRENE code [69]). This did not appear to work well at matching profiles of plasma in the edge region of a JET configuration. Particularly, the temperature profiles obtained from the simplified model with optimised parameters significantly over-estimated the high-fidelity simulations. This was likely due to the simplified form of the transport model used [3].
Adjoint methods were used to control the electron profile in ITER with electron cyclotron current drive (ECCD) [32]. The authors varied the driving current beam injection angles to evaluate the current density sensitivity and determine the optimal beam position. This approach can be used to stabilise MHD wave modes.
In tokamak and stellarator reactors, MHD instabilities can arise that hinder fusion performance by reducing the plasma confinement [72]. Paul *et al.* [61] developed an adjoint method to minimise MHD instabilities in stellarator equilibria, demonstrating it on three optimisation problems: (i) a target rotational transform profile that would enable low shear flow, (ii) creation of a magnetic well, (iii) and quasi-symmetry on a magnetic axis. This was then extended by Gaur *et al.* [25], who focused on the 'balloon mode' MHD instability in a fixed geometry. Specifically, they optimised the field line geometry to find MHD equilibria that were stable against this mode by minimising the maximum eigenvalue. Their final results gave a maximum eigenvalue with a stable growth rate.
Blommaert *et al.* [8] performed sensitivity analysis of uncertain parameters in a plasma edge model in a tokamak configuration. Their model had 34 parameters such as radial transport coefficients, boundary conditions, rate coefficients (such as charge exchange rate), and parameters controlling the magnetic equilibrium profile. They considered two cost functions: one related to the heat load and the other related to the temperature on the outer strike point of the divertor. Their sensitivity analysis showed strong dependence of both objectives on the boundary conditions and equilibrium profile parameters. Using the adjoint method, this only required 2 simulations instead of 34 (one for each cost function). Their sensitivities were also validated against finite difference calculations.
In reactive flows, the adjoint method was used for optimisation of spark location for laminar flame ignition by Qadri *et al.* [64]. As is common in gradient-based methods, they showed that the optimal solution is dependent on the initial condition [64]. Kord and Capecelatro used the adjoint method to determine a momentum source term that could control a mixing layer, limiting the Kelvin-Helmholtz instability that was observed in the baseline solution to a small spatial region [36]. In a previous study, the same authors used the adjoint method to control a Rayleigh-Taylor instability [37]. Also in reactive flows, Qadri *et al.* [63] studied low density jets and their global modes. They used an adjoint analysis to identify regions receptive to external forcing that may control the growth rate/frequency of the unstable mode.
Mao *et al.* [44] studied control strategies for base flow perturbations using the adjoint method in a stenotic flow and Batchelor flow. They used a boundary condition that varied in time and space, given by $\boldsymbol{u}_{\Gamma_c} = \boldsymbol{c}(\boldsymbol{x})f(t,\omega)$, where Γ_c is the control boundary, $\boldsymbol{c}$ is the spatial change in the velocity amplitude of the control boundary (to be optimised), f is the temporal change in velocity with a circular frequency ω. The cost function (6) was based on the kinetic energy of the flowfield, which was minimised with

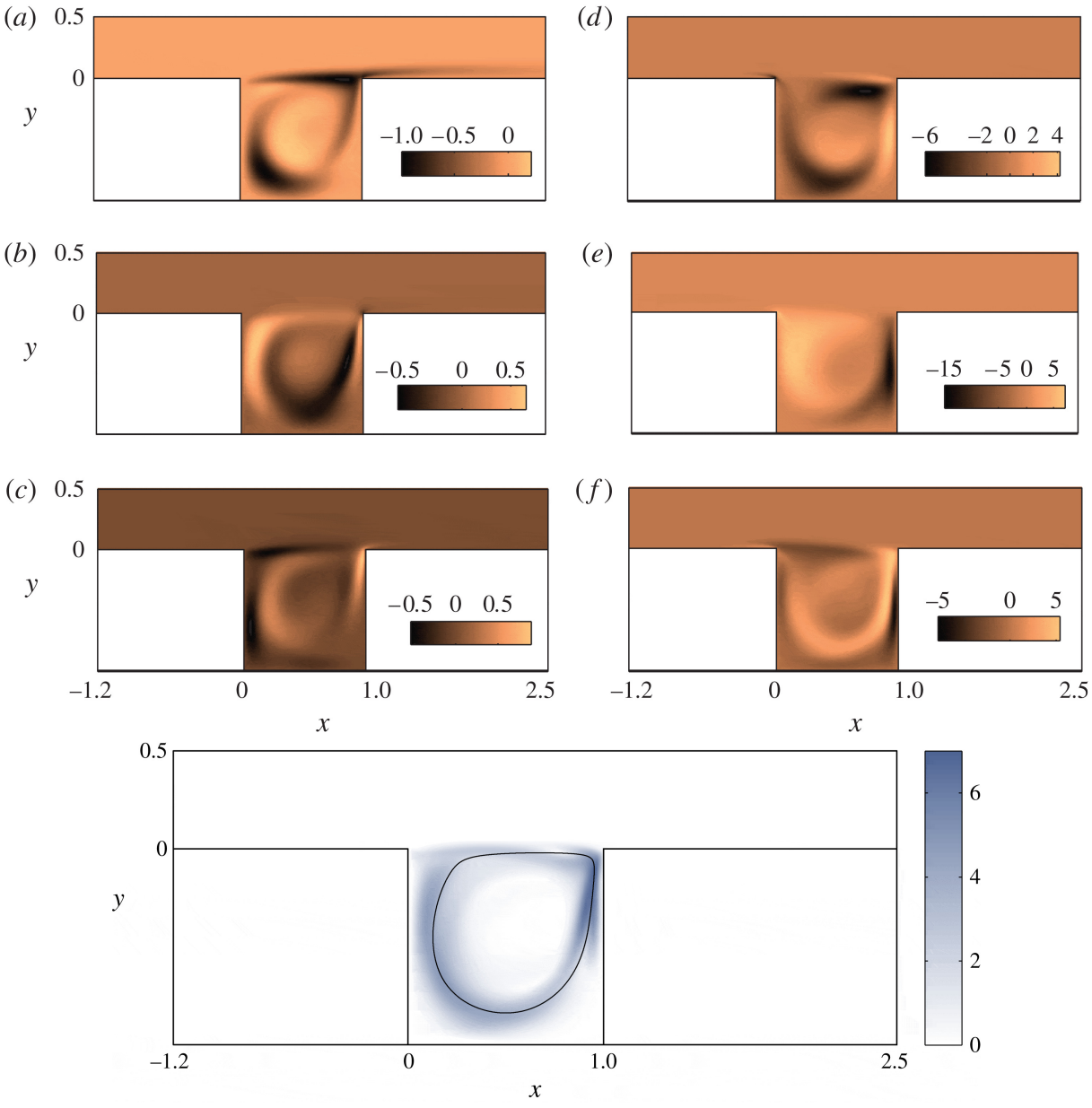

Figure 3: Contours of stability (direct mode, left), receptivity (adjoint mode, right) and structural sensitivity (bottom row) in flow over an open cavity at a Reynolds number near the transition from 2D to 3D flow. Panels show (a,d) streamwise direction, (b,e) wall-normal direction, (c,f) spanwise direction. Figures taken from Citro *et al.* [13].

respect to the amplitude of $\boldsymbol{c}$. In tokamaks, the flow may depend on wall temperature, which could be optimised/controlled in a similar way.
Related to parameter optimisation, the adjoint method is also widely used in numerical weather prediction to perform four-dimensional data assimilation [40]. This was discussed in a previous NEPTUNE report on model order reduction [21]. In the case of numerical weather prediction, flow data of varying type and location can be combined with a parameterised mathematical model to determine unknown parameters which describe the flow. The unknown parameters are determined through a gradient-based optimisation process where the cost function is based on the mean-square error of the model and data [31]. The assimilated parameters can then be used to forecast and predict future weather (or more generally, future flow states). Given the large number of diagnostic devices in tokamaks which provide measurements, the use of adjoint methods for data assimilation is highly relevant to NEPTUNE.

3.3 Uncertainty quantification

Uncertainties can enter complex simulations such as plasma models through a variety of sources including initial conditions, boundary conditions, geometrical parameters, modelling errors [20], numerical errors, etc [78]. The latter point on numerical errors is also relevant for adaptive mesh refinement (Section 3.4). Wang [78] studied flow past a cylinder, where the cylinder is subject to small but finite, random and unknown oscillations in its angular velocity (around a mean of 0), given as

$$\omega(t) = \sum_{i=1}^{10} \xi_{2i-1} \cos(2\pi i f_V t) + \xi_{2i} \sin(2\pi i f_V t), \tag{17}$$

where ω is the angular velocity, $\xi_0 ... \xi_{20}$ are random variables from a distribution $\mathcal{N}(0, 0.01)$, f_V is the vortex shedding frequency. The cost function gradient w.r.t the random variables is $\nabla_{\boldsymbol{\xi}} \mathcal{J}$. Wang found the sensitivity of the drag coefficient, C_D, using the adjoint method, and used this in combination with Monte Carlo sampling to determine the probability of the drag coefficient passing a certain critical threshold, $C_{D,crit}$, due to the random perturbations in (17) [78].
Wang's study aimed to use the sensitivity information to improve the efficiency and accuracy of uncertainty quantification done with a brute force Monte Carlo estimator. Instead of simulating random instances of $\boldsymbol{\xi}$ and the resultant cost function, $\mathcal{J}$, Wang simply evaluated the cost function once for one set of $\boldsymbol{\xi}$ and used the following linear approximation to estimate the probability of $P(C_D > C_{D,crit})$:

$$\mathcal{J} \approx \widetilde{\mathcal{J}} = \mathcal{J}_0 + \nabla_{\boldsymbol{\xi}} \mathcal{J} \cdot \boldsymbol{\xi}, \tag{18}$$

where the approximation $\widetilde{\mathcal{J}}$ can be computed with minimal cost (particularly compared to evaluating $\mathcal{J}$ by brute force methods). This is extremely fast since the forward and adjoint problems should only be solved once, but the confidence interval of $P(C_D > C_{D,crit})$ was similar to a brute force Monte Carlo. The confidence interval on $P(C_D > C_{D,crit})$ was reduced one order of magnitude using an importance sampling method, which used $\widetilde{\mathcal{J}}$ to lower the variance of the Monte Carlo solution in the heavy-tailed region [78].
For plasma problems, Carli *et al.* [10] used a Bayesian approach to quantify uncertainty in turbulence modelling parameters in a plasma-edge code (SOLPS-ITER), where the Bayesian posterior distribution was found through gradient-descent with adjoint mode algorithmic differentiation.

3.4 Adaptive mesh refinement

Several studies have researched using adjoint sensitivity as an error estimator for adaptive mesh refinement [39, 58, 59, 75]. Classical adaptive mesh refinement has used heuristic error estimators such as velocity gradients (for wall-bounded turbulence) [39]. This has been shown to have poor performance since all flow features are treated equally and the user must iteratively determine controls for the adaptation process [39]. In contrast, the adjoint method can be used as a 'goal-orientated' error estimator that can identify regions where a cost function is sensitive to perturbations and can refine there to avoid introducing numerical artefacts into a sensitive region [4, 65]. This 'dual-weighted residual' method uses the residuals of the primal solution, weighted by the adjoint sensitivities to estimate local error and identify regions that would benefit from mesh refinement [65]. Initial tests of steady flow past a cylinder showed that

one could obtain one order of magnitude improvement of drag coefficient error (with the same number number of cells) by using an error estimator based on adjoint sensitivities [65]. Adjoint methods were used to estimate local error in adaptive mesh-refinement in spectral element simulations of a turbulent flow over a periodic hill with the incompressible Navier-Stokes equations [58]. Their adjoint problem followed the cost function

$$\mathcal{J} = \int_\Omega \boldsymbol{u} \cdot \boldsymbol{j}_{\Omega,\boldsymbol{u}} dV + \int_\Omega p \cdot \boldsymbol{j}_{\Omega,p} dV + \int_{\Gamma_D} \left(\frac{1}{Re} \nabla \boldsymbol{u} \cdot \boldsymbol{n} - p\boldsymbol{n} \right) \cdot \boldsymbol{j}_{\Gamma_D} ds + \int_{\Gamma_o} \boldsymbol{u} \cdot \boldsymbol{j}_{\Gamma_o} ds, \tag{19}$$

where $\boldsymbol{u}$ is the fluid velocity, p is the fluid pressure, Re is the Reynolds number, $\boldsymbol{n}$ is a unit vector normal to the boundary, Ω is the internal domain, Γ_o is the outlet boundary and Γ_D is a Dirichlet boundary (e.g. walls with no-slip condition). $\boldsymbol{j}$ is a problem-specific set of coefficients. For example, in a lid-driven cavity problem, Offermans *et al.* used $\boldsymbol{u}_{lid} = -\boldsymbol{j}_{\Gamma_D}$, and all other $\boldsymbol{j} = (0, 0)$ (for a 2D case) [59].
Offermans *et al.* then computed the error on the spectral element solution of the cost function (19) when a polynomial of order N is used [59] with the following approximation, considering the domain is split into E subdomains, Ω_e, with boundary Γ_e:

$$\delta\mathcal{J} = \mathcal{J} - \mathcal{J}_N \approx \sum_{e=1}^{E} \left[\left\langle \mathcal{R}_1(\boldsymbol{q}), \boldsymbol{u}^+ - \boldsymbol{u}_N^+ \right\rangle_{\Omega_e} + \left\langle \mathcal{R}_2(\boldsymbol{q}), \boldsymbol{u}^+ - \boldsymbol{u}_N^+ \right\rangle_{\Gamma_e} + \left\langle \mathcal{R}_3(\boldsymbol{q}), p^+ - p_N^+ \right\rangle_{\Omega_e} \right] \tag{20}$$

$$\delta\mathcal{J} \approx \sum_{e=1}^{E} \left(||\mathcal{R}_1|| \, ||\boldsymbol{u}^+ - \boldsymbol{u}_N^+|| \right)_{\Omega_e} + \left(||\mathcal{R}_2|| \, ||\boldsymbol{u}^+ - \boldsymbol{u}_N^+|| \right)_{\Gamma_e} + \left(||\mathcal{R}_3|| \, ||p^+ - p_N^+|| \right)_{\Omega_e}. \tag{21}$$

where $\langle \cdot \rangle$ is the inner product, $|| \cdot ||$ is the L^2-norm, and each $\mathcal{R}$ is a residual from either the cell solution ($\mathcal{R}_1$ is from the momentum equation and $\mathcal{R}_3$ is from the continuity equation) or the face residual between two elements ($\mathcal{R}_2$). The terms $\boldsymbol{u}^+ - \boldsymbol{u}_N^+$ and $p^+ - p_N^+$ are based on the interpolation error and are approximated with a relationship based on the cell size (volume and edge length) and polynomial order [59]. As we can see from (21), the error is based on the residual of the primal solutions, which are weighted by the adjoint (also known as 'dual') solution, giving the name 'dual-weighted residual' [4].
Their results showed that the adjoint method focused the mesh refinement to areas where the objective function sensitivity was high and left the mesh unchanged in low-sensitivity regions of the domain. This was beneficial compared to the alternative 'spectral error indicator' error measure that refined the mesh based on the truncation error and quadrature error, which are approximated based on the local decay rate of the spectral coefficients [46]. This computes the influence of the polynomial coefficients, where a large influence of high-order terms corresponds to a high polynomial truncation error and low grid resolution. The spectral error indicator was shown to be much simpler to compute but provided a more uniform mesh [58]. The computation of the adjoint error estimator was complex but provided a more focused and localized refinement. Using the adjoint error estimator for grid refinement was shown to converge to the correct solution with fewer refinement iterations [59] but took twice the time for one refinement round due to increased computational complexity for a steady simulation. For an unsteady simulation, the adjoint method was shown to be only 10% more time-consuming than the spectral error indicator due to the low number of timesteps used for mesh refinement relative to the total simulation duration (3 iterations of mesh refinement took 21 flow-through times compared to 140 flow-through times on the final mesh to get turbulent statistics) [58]. The adjoint error estimator was shown to improve the accuracy of a quantity of interest after fewer rounds of mesh refinement, but the increased computational cost of refinement iterations suggests that adjoint-based mesh refinement is not yet superior to other error estimators (with regards to cost-accuracy trade off).
Similar to as discussed in Section 2.2.2, computing the adjoint for a long time series on a fine mesh invokes computational challenges surrounding how to manage storing the full primal solution. For their adaptive mesh refinement study on unsteady turbulent flows, Offermans *et al.* used checkpointing with the `revolve` library [58]. An alternative approach, also discussed in Section 2.2.2, is storing a reduced-order model of the primal solution and using this when computing the adjoint equations, as was done by Li *et al.* [39]. This removes the bottleneck of time- and memory-consuming checkpointing schemes (at the cost of lower accuracy). However, this reduced-order model approach makes the most sense in an adaptive mesh refinement problem, since the point of adaptive mesh refinement is to minimise the computational cost required to obtain the desired solution. Offermans *et al.* [58] also mentioned issues in stability when solving the adjoint equations due to highly turbulent flow, which they speculated may

be resolved by using a reduced-order model to store the primal, since the small and chaotic motion is generally not included in reduced-order models. Therefore, the benefit of using a reduced-order model for storing the primal flow field in highly turbulent flows may be two-fold: (1) reduced storage, memory and re-computation requirements, (2) improved stability when back-integrating the adjoint problem.

3.5 Shape optimisation

Early adoption of the adjoint methods in fluid flows were mainly focused on shape optimisation of airfoils [35]. This application was later extended to fusion by optimising tokamak divertors by Dekeyser [15]. Their cost function aimed to generate a uniform heat flux profile by minimising the least squares difference between the computed flux and a flux below the critical value of the material [17, 18]. Different approaches for treating the moving shape include: (1) a density-based method, where the entire shape design space is meshed and cells beyond a boundary are given a high density/low permeability, (2) a body fitted mesh where the mesh must be morphed to new target shapes, or (3) a level-set method which requires special numerical treatment at the wall [1]. Dekeyser et al. [17, 18] use a body-fitted mesh that is orthogonal to the field surfaces which is common in plasmas. The divertor region was then parameterised by a function ϕ and the new geometry in the divertor region was generated by shrinking/stretching cells here. They also use a high-order (nine-point) discretisation stencil to calculate the convective fluxes in the plasma edge transport model [16]. The nine-point stencil includes contributions from diagonally adjacent cells, compared to only contributions from cells which are adjacent in the horizontal and vertical directions [15]. This five-point stencil gave some inaccuracies in shape sensitivities compared to finite differences, whereas the nine-point stencil agreed perfectly [16]. Later, a 1D version of the finite volume code used in Dekeyser's previous work for edge simulations was coupled to a Monte Carlo code for a kinetic simulation of neutrals [19]. Even in a 1D analysis, there was a severe issue computing the cost function gradients due to Monte Carlo statistical error. The continuous adjoint required around a factor of 100 more particles to obtain similar cost function accuracy as the discrete adjoint, although their discrete adjoint code was 10 times more expensive (when using the same number of particles). This shows the use of adjoint methods for shape optimisation and other applications mentioned in Section 3 becomes problematic when coupled to a kinetic simulation. This may lead to the requirement of an unfeasibly large number of particles for accurate adjoint analyses in NEPTUNE.

4 Conclusions

We have reviewed the adjoint method, its developments and various applications within the context of plasma flows. The adjoint method is an effective tool for optimisation and sensitivity analysis in flow configurations with a large number of operating parameters or modelling parameters in tokamak simulations. A key issue when using the adjoint method to study unsteady flows is that the primal solution must be known at each timestep, which presents some computational issues that have been addressed through checkpointing, reduced-order modelling and parallel-in-time algorithms. Of these, checkpointing appears the least intrusive and is already implemented in libraries such as PETSc, meaning it is a viable fit for NEPTUNE developments in the Nektar++ framework. Applications of the adjoint method mainly consist of analysing flow sensitivity to perturbations. Such information about sensitivity is often used to optimise design parameters, shape or control strategies. Sensitivity information can also be used to understand and quantify error in the presence of unknown parameters, or as an error estimator in adaptive mesh refinement algorithms.

References

1. Alexandersen, J. & Andreasen, C. S. A review of topology optimisation for fluid-based problems. *Fluids* **5,** 29 (2020).
2. Antonsen, T., Paul, E. J. & Landreman, M. Adjoint approach to calculating shape gradients for three-dimensional magnetic confinement equilibria. *Journal of Plasma Physics* **85,** 905850207 (2019).

3. Baelmans, M., Blommaert, M., De Schutter, J., Dekeyser, W. & Reiter, D. Efficient parameter estimation in 2D transport models based on an adjoint formalism. *Plasma Physics and Controlled Fusion* **56,** 114009 (2014).
4. Bangerth, W. & Rannacher, R. *Adaptive finite element methods for differential equations* (Springer Science & Business Media, 2003).
5. Blackburn, H. M., Barkley, D. & Sherwin, S. J. Convective instability and transient growth in flow over a backward-facing step. *Journal of Fluid Mechanics* **603,** 271–304 (2008).
6. Blackburn, H. M. & Sherwin, S. J. Instability modes and transition of pulsatile stenotic flow: pulse-period dependence. *Journal of Fluid Mechanics* **573,** 57–88 (2007).
7. Blackburn, H. M., Sherwin, S. J. & Barkley, D. Convective instability and transient growth in steady and pulsatile stenotic flows. *Journal of Fluid Mechanics* **607,** 267–277 (2008).
8. Blommaert, M., Reiter, D. & Baelmans, M. An efficient methodology to analyze plasma edge model parameter sensitivities. *Nuclear Materials and Energy* **12,** 1049–1054 (2017).
9. Capecelatro, J., Bodony, D. J. & Freund, J. B. Adjoint-based sensitivity and ignition threshold mapping in a turbulent mixing layer. *Combustion Theory and Modelling* **23,** 147–179 (2019).
10. Carli, S. *et al.* Bayesian maximum a posteriori-estimation of κ turbulence model parameters using algorithmic differentiation in SOLPS-ITER. *Contributions to Plasma Physics* **62,** e202100184 (2022).
11. Chandler, G. J. *Sensitivity analysis of low-density jets and flames* PhD thesis (University of Cambridge, 2011).
12. Choi, Y., Arrighi, W. J., Copeland, D. M., Anderson, R. W. & Oxberry, G. M. *libROM* version 1.0.0. Oct. 2019. `https://github.com/LLNL/libROM`.
13. Citro, V., Giannetti, F., Brandt, L. & Luchini, P. Linear three-dimensional global and asymptotic stability analysis of incompressible open cavity flow. *Journal of Fluid Mechanics* **768,** 113–140 (2015).
14. Costanzo, S., Sayadi, T., de Pando, M. F., Schmid, P. & Frey, P. Parallel-in-time adjoint-based optimization–application to unsteady incompressible flows. *Journal of Computational Physics* **471,** 111664 (2022).
15. Dekeyser, W. Optimal Plasma Edge Configurations for Next-Step Fusion Reactors (2014).
16. Dekeyser, W., Reiter, D. & Baelmans, M. A one shot method for divertor target shape optimization. *PAMM* **14,** 1017–1022 (2014).
17. Dekeyser, W., Reiter, D. & Baelmans, M. Automated divertor target design by adjoint shape sensitivity analysis and a one-shot method. *Journal of Computational Physics* **278,** 117–132 (2014).
18. Dekeyser, W., Reiter, D. & Baelmans, M. Divertor target shape optimization in realistic edge plasma geometry. *Nuclear fusion* **54,** 073022 (2014).
19. Dekeyser, W. *et al.* Divertor design through adjoint approaches and efficient code simulation strategies. *Contributions to Plasma Physics* **58,** 643–651 (2018).
20. Dow, E. & Wang, Q. *Quantification of Structural Uncertainties in the k-w Turbulence Model* in *52nd AIAA/ASME/ASCE/AHS/ASC Structures, structural dynamics and materials conference 19th AIAA/ASME/AHS adaptive structures conference 13t* (2011), 1762.
21. Ed Threlfall, W. A. *Select techniques for MOR (Model Order Reduction)* tech. rep. CD/EXCALIBUR-FMS/0031-M2.5.1 (UK Atomic Energy Authority, Mar. 15, 2021). `https://github.com/ExCALIBUR-NEPTUNE/Documents/blob/main/reports/ukaea_reports/CD-EXCALIBUR-FMS0031-M2.5.1.pdf`.
22. *Exposing parallelism: Parallel in Time* Excalibur projects. `https://excalibur.ac.uk/projects/exposing-parallelism-parallel-in-time`.
23. Gander, M. J. & Güttel, S. PARAEXP: A parallel integrator for linear initial-value problems. *SIAM Journal on Scientific Computing* **35,** C123–C142 (2013).
24. Gander, M. J., Güttel, S. & Petcu, M. *A nonlinear ParaExp algorithm* in *Domain Decomposition Methods in Science and Engineering XXIV 24* (2018), 261–270.

25. Gaur, R. *et al.* An adjoint-based method for optimizing MHD equilibria against the infinite-n, ideal ballooning mode. *arXiv preprint arXiv:2302.07673* (2023).

26. Geraldini, A., Landreman, M. & Paul, E. An adjoint method for determining the sensitivity of island size to magnetic field variations. *Journal of Plasma Physics* **87,** 905870302 (2021).

27. Giannetti, F., Camarri, S. & Luchini, P. Structural sensitivity of the secondary instability in the wake of a circular cylinder. *Journal of Fluid Mechanics* **651,** 319–337 (2010).

28. Giannetti, F. & Luchini, P. Structural sensitivity of the first instability of the cylinder wake. *Journal of Fluid Mechanics* **581,** 167–197 (2007).

29. Giles, M. B. & Pierce, N. A. An introduction to the adjoint approach to design. *Flow, turbulence and combustion* **65,** 393–415 (2000).

30. Griewank, A. & Walther, A. Algorithm 799: revolve: an implementation of checkpointing for the reverse or adjoint mode of computational differentiation. *ACM Transactions on Mathematical Software (TOMS)* **26,** 19–45 (2000).

31. Griffith, A. K. & Nichols, N. K. Adjoint methods in data assimilation for estimating model error. *Flow, turbulence and combustion* **65,** 469–488 (2000).

32. Hamamatsu, K. & Fukuyama, A. Controllability of driven current profile in ECCD on ITER. *Fusion engineering and design* **53,** 53–58 (2001).

33. Heuveline, V. & Walther, A. *Online checkpointing for parallel adjoint computation in PDEs: Application to goal-oriented adaptivity and flow control* in *Euro-par* (2006), 689–699.

34. Huerre, P. & Monkewitz, P. A. Absolute and convective instabilities in free shear layers. *Journal of Fluid Mechanics* **159,** 151–168 (1985).

35. Jameson, A. Aerodynamic design via control theory. *Journal of scientific computing* **3,** 233–260 (1988).

36. Kord, A. & Capecelatro, J. A discrete-adjoint framework for optimizing high-fidelity simulations of turbulent reacting flows. *Proceedings of the Combustion Institute* (2022).

37. Kord, A. & Capecelatro, J. Optimal perturbations for controlling the growth of a Rayleigh–Taylor instability. *Journal of Fluid Mechanics* **876,** 150–185 (2019).

38. Lea, D. J., Allen, M. R. & Haine, T. W. Sensitivity analysis of the climate of a chaotic system. *Tellus A: Dynamic Meteorology and Oceanography* **52,** 523–532 (2000).

39. Li, X., Hulshoff, S. & Hickel, S. Towards adjoint-based mesh refinement for Large Eddy Simulation using reduced-order primal solutions: Preliminary 1D Burgers study. *Computer Methods in Applied Mechanics and Engineering* **379,** 113733 (2021).

40. Lorenc, A. C. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society* **112,** 1177–1194 (1986).

41. Luce, T. Realizing steady-state tokamak operation for fusion energy. *Physics of Plasmas* **18,** 030501 (2011).

42. Luchini, P. & Bottaro, A. Adjoint equations in stability analysis. *Annual Review of fluid mechanics* **46,** 493–517 (2014).

43. Magri, L. Adjoint methods as design tools in thermoacoustics. *Applied mechanics reviews* **71** (2019).

44. Mao, X., Blackburn, H. & Sherwin, S. Optimal suppression of flow perturbations using boundary control. *Computers & Fluids* **121,** 133–144 (2015).

45. Marin, O., Constantinescu, E. & Smith, B. A scalable matrix-free spectral element approach for unsteady PDE constrained optimization using PETSc/TAO. *Journal of Computational Science* **47,** 101207 (2020).

46. Mavriplis, C. Adaptive mesh strategies for the spectral element method. *Computer methods in applied mechanics and engineering* **116,** 77–86 (1994).

47. Meliga, P., Chomaz, J.-M. & Sipp, D. Global mode interaction and pattern selection in the wake of a disk: a weakly nonlinear expansion. *Journal of Fluid Mechanics* **633,** 159–189 (2009).

48. Mordijck, S. Taking Control of Fusion Reactor Instabilities. *Physics* **15,** 154 (2022).

49. Murali, S. *Stability and Heat Transfer Characteristics of Hydrodynamic Channel Flows and Quasi-two-dimensional Magnetohydrodynamic Duct Flows with Repeated Wedge-shaped Protrusions* PhD thesis (Monash University, 2022).

50. Murali, S., Hussam, W. K. & Sheard, G. J. Heat transfer enhancement in quasi-two-dimensional magnetohydrodynamic duct flows using repeated flow-facing wedge-shaped protrusions. *International Journal of Heat and Mass Transfer* **171,** 121066 (2021).

51. Murali, S., Ng, Z. Y. & Sheard, G. J. Stability of flow in a channel with repeated flow-facing wedge-shaped protrusions. *Journal of Fluid Mechanics* **941,** A59 (2022).

52. Nadarajah, S. & Jameson, A. *A comparison of the continuous and discrete adjoint approach to automatic aerodynamic optimization* in *38th Aerospace sciences meeting and exhibit* (2000), 667.

53. Nadarajah, S. & Jameson, A. *Studies of the continuous and discrete adjoint approaches to viscous automatic aerodynamic shape optimization* in *15th AIAA computational fluid dynamics conference* (2001), 2530.

54. Newman III, J. C., Taylor III, A. C., Barnwell, R. W., Newman, P. A. & Hou, G. J.-W. Overview of sensitivity analysis and shape optimization for complex aerodynamic configurations. *Journal of Aircraft* **36,** 87–96 (1999).

55. Nies, R., Paul, E. J., Hudson, S. R. & Bhattacharjee, A. Adjoint methods for quasi-symmetry of vacuum fields on a surface. *Journal of Plasma Physics* **88,** 905880106 (2022).

56. Nobis, H., Schlatter, P., Wadbro, E., Berggren, M. & Henningson, D. S. Topology optimization of unsteady flows using the spectral element method. *Computers & Fluids* **239,** 105387 (2022).

57. Nordlund, K. H. *et al.* European research roadmap to the realisation of fusion energy (2018).

58. Offermans, N., Massaro, D., Peplinski, A. & Schlatter, P. Error-driven adaptive mesh refinement for unsteady turbulent flows in spectral-element simulations. *Computers & Fluids* **251,** 105736 (2023).

59. Offermans, N., Peplinski, A., Marin, O. & Schlatter, P. Adaptive mesh refinement for steady flows in Nek5000. *Computers & Fluids* **197,** 104352 (2020).

60. Paul, E. J., Abel, I. G., Landreman, M. & Dorland, W. An adjoint method for neoclassical stellarator optimization. *Journal of Plasma Physics* **85,** 795850501 (2019).

61. Paul, E. J., Landreman, M. & Antonsen, T. Gradient-based optimization of 3D MHD equilibria. *Journal of Plasma Physics* **87,** 905870214 (2021).

62. Qadri, U. A. *Global stability and control of swirling jets and flames* PhD thesis (University of Cambridge, 2014).

63. Qadri, U. A., Chandler, G. J. & Juniper, M. P. Passive control of global instability in low-density jets. *European Journal of Mechanics-B/Fluids* **72,** 311–319 (2018).

64. Qadri, U. A., Magri, L., Ihme, M. & Schmid, P. J. Using adjoint-based optimization to enhance ignition in non-premixed jets. *Proceedings of the Royal Society A* **477,** 20200472 (2021).

65. Rannacher, R. Adaptive finite element discretization of flow problems for goal-oriented model reduction. *Computational Fluid Dynamics 2008,* 31–45 (2008).

66. Schanen, M., Marin, O., Zhang, H. & Anitescu, M. Asynchronous two-level checkpointing scheme for large-scale adjoints in the spectral-element solver Nek5000. *Procedia Computer Science* **80,** 1147–1158 (2016).

67. Schmid, P. J. Nonmodal stability theory. *Annu. Rev. Fluid Mech.* **39,** 129–162 (2007).

68. Schmid, P. J. & Brandt, L. Analysis of Fluid Systems: Stability, Receptivity, Sensitivity. *Applied Mechanics Review* **66,** 024803 (2014).

69. Schneider, R. *et al.* Plasma edge physics with B2-eirene. *Contributions to Plasma Physics* **46,** 3–191 (2006).

70. Sipp, D., Marquet, O., Meliga, P. & Barbagallo, A. Dynamics and control of global instabilities in open-flows: a linearized approach. *Applied Mechanics Reviews* **63** (2010).

71. Skene, C. S., Eggl, M. F. & Schmid, P. J. A parallel-in-time approach for accelerating direct-adjoint studies. *Journal of Computational Physics* **429,** 110033 (2021).

72. Strait, E. J. Magnetic control of magnetohydrodynamic instabilities in tokamaks. *Physics of Plasmas* **22,** 021803 (2015).

73. Strykowski, P. J. & Sreenivasan, K. R. On the formation and suppression of vortex 'shedding'at low Reynolds numbers. *Journal of Fluid Mechanics* **218,** 71–107 (1990).

74. Stumm, P. & Walther, A. New algorithms for optimal online checkpointing. *SIAM Journal on Scientific Computing* **32,** 836–854 (2010).

75. Venditti, D. A. & Darmofal, D. L. Adjoint error estimation and grid adaptation for functional outputs: Application to quasi-one-dimensional flow. *Journal of Computational Physics* **164,** 204–227 (2000).

76. Vezyris, C., Papoutsis-Kiachagias, E. & Giannakoglou, K. On the incremental singular value decomposition method to support unsteady adjoint-based optimization. *International Journal for Numerical Methods in Fluids* **91,** 315–331 (2019).

77. Vishnampet, R., Bodony, D. J. & Freund, J. B. A practical discrete-adjoint method for high-fidelity compressible turbulence simulations. *Journal of Computational Physics* **285,** 173–192 (2015).

78. Wang, Q. *Uncertainty quantification for unsteady fluid flow using adjoint-based approaches* PhD thesis (Stanford University, 2009).

79. Zhang, H. & Constantinescu, E. *Revolve-based adjoint checkpointing for multistage time integration* in *International Conference on Computational Science* (2021), 451–464.

80. Zhang, H., Constantinescu, E. M. & Smith, B. F. PETSc TSAdjoint: a discrete adjoint ODE solver for first-order and second-order sensitivity analysis. *SIAM Journal on Scientific Computing* **44,** C1–C24 (2022).