Robust control of resistive wall modes using pseudospectra

M Sempf, P Merkel, E Strumberger , C Tichmann and S Günter

Max-Planck-Institut für Plasmaphysik, EURATOM Association, Boltzmannstraße 2, D-85748 Garching, Germany

E-mail: msempf@ipp.mpg.de

Abstract. A novel methodology to design robust feedback controllers for the stabilisation of resistive wall modes (RWMs) in tokamaks is presented. A linear state space model is used which describes the system composed of the plasma, the resistive wall, the active coils, and the magnetic field sensors. The full 3D geometry of the wall and the coils is taken into account. The control system is represented by a parametrised matrix whose robust stability properties are optimised under variations of the parameters. To make the optimisation process feasible, a reduced state space model is constructed by means of a novel technique involving an orthogonal projection. The orthogonality is essential to the robust stability concept used. The latter is based upon the idea of matrix pseudospectra and therefore accounts for the sensitivity of eigenvalues. Furthermore, the transient growth of perturbations is investigated. Namely, in some theoretically stable linear systems, initial perturbations can grow by large factors before they eventually decay, rendering them practically unstable.

A detailed analysis of a simple, ITER-like test case provides the following general, conceptual insights into the RWM stabilisation problem: (i) it is important to consider not only the eigenvalues themselves, but also their sensitivity, (ii) transient amplification might be an issue requiring consideration, and (iii) the transient peak can be substantially reduced by a careful choice of the optimisation objective and the sensor configuration.

PACS numbers: 52.35.Py, 07.05.Dz

AMS classification scheme numbers: 76W05, 37N35, 49N35, 93B11, 93B35, 93B51, 93B52, 93B60, 93C05, 93D09, 93D15, 93D20, 93D21

Submitted to: *New J. Phys.*

Contents

1. Introduction

An important figure of merit for plasma confinement in nuclear fusion devices is the parameter $\beta = 2\mu_0\langle p\rangle/\langle B^2\rangle$, where $\langle\cdot\rangle$ denotes the volume average. Since the fusion power output increases with β, a future reactor must reach a sufficiently high β in order to be efficient and economically attractive. However, ideal magnetohydrodynamic (MHD) instabilities impose hard limits on the achievable β. In tokamaks, if internal instabilities (i. e., those which occur within the plasma core) are avoided by a suitable choice of the plasma current profile, the β limit comes about by the onset of *external kink modes*. These are long-wavelength instabilities which are driven by the radial gradients of the toroidal current and the pressure. They cause a deformation of the plasma boundary, grow on the Alfvénic time scale of order 10^{-6} s and can terminate the plasma discharge abruptly. A superconducting wall sufficiently close to the plasma boundary would stabilise these modes and therefore substantially increase the β limit. But in the realistic case that such a wall has non-vanishing resistivity, the plasma configuration becomes unstable again, the stability limit being virtually the same as in the case without wall. However, the modes grow much more slowly, namely, on the resistive time scale of the wall which is typically of order 10^{-2} s. These decelerated external kink modes are denoted *resistive wall modes* (RWMs). Therefore, in the presence of a resistive wall the active feedback stabilisation of RWMs by means of magnetic field sensors and a system of additional correction coils becomes technologically feasible, and successful experiments, e. g., in the DIII-D tokamak, have already been conducted [1, 2, 3].

Given a set of sensors and control coils, the coils have to be connected to the sensors by means of some feedback logics, or controller. Finding an appropriate controller is always a difficult task, but at least it is strongly facilitated by numerical modelling of RWMs

and their feedback stabilisation. In this paper, a novel approach to the computational design of robust RWM controllers is presented.

The numerical treatment of RWM control has already been the subject of numerous studies, and several codes have been developed and extensively used to compute the linear stability of RWMs in the presence of a feedback system (in particular for the upcoming fusion experiment ITER): VALEN [4, 5], DCON coupled to VACUUM [6], MARS-F [7, 8, 9], CarMa [10], and STARWALL [11, 12]. STARWALL is used also in the present work. All these codes differ from each other in the complexity of the representation of the plasma, the conducting wall, and the feedback coils. In the VALEN code, the plasma state is represented by a single unstable eigenmode only, whereas in the other codes an expansion in terms of a finite, but large set of basis functions is used. Only MARS-F and CarMa take toroidal rotation of the plasma into account. The conducting structures have a full 3D representation in VALEN, CarMa, and STARWALL, whereas the other codes are restricted to axisymmetric configurations. In VALEN and STARWALL, however, a thin-wall approximation is used. Finally, STARWALL is the only code being able to account for more than one toroidal Fourier index n simultaneously in the expansion of the plasma state, thus resolving the coupling between different n's in non-axisymmetric setups.

All the above mentioned studies, except those related to STARWALL, tackle the control problem in the same way as it is quite common in control engineering: by the use of *transfer functions*. The underlying methodology shall be briefly reviewed here although a different approach is taken is this study. The transfer function concept is based upon regarding the system to be controlled as a "signal processor", namely, a black box into which an input signal $\mathbf{v}$ (scalar or vector-valued) can be fed and which produces an output signal $\mathbf{y}$ (also scalar or vector-valued). Inside the black box there is a linear

dynamical system described by a matrix $\mathbf{A}_0$, whose internal dynamics is subjected to an additional external forcing if a non-zero input signal is applied. This means, the control system is described by

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}_0\mathbf{x} + \mathbf{B}\mathbf{v}, \\ \mathbf{y} &= \mathbf{C}\mathbf{x}, \end{aligned} \tag{1.1}$$

so that the input matrix $\mathbf{B}$ describes the influence of the input $\mathbf{v}$ on the system state vector $\mathbf{x}$, and the output signal $\mathbf{y}$ is a measurement taken from $\mathbf{x}$ by means of the output matrix $\mathbf{C}$. Without any feedback control, the input-output characteristics of the system is described by the *open-loop transfer function* $\mathbf{G}_0(z)$, which is defined by $\{\mathcal{L}\mathbf{y}\}(z) = \mathbf{G}_0(z)\{\mathcal{L}\mathbf{v}\}(z)$, where the operator $\mathcal{L}$ denotes the Laplace transform (not the z-transform!), and $z \in \mathbb{C}$. If the input and output are both scalars, the open-loop transfer function is a scalar, as well, otherwise it is a vector or a matrix. It is given by

$$\mathbf{G}_0(z) = \mathbf{C}(z\mathbb{1} - \mathbf{A}_0)^{-1}\mathbf{B}. \tag{1.2}$$

In many control applications, a feedback controller with its own transfer function $\mathbf{G}_\mathrm{K}(z)$ is applied which processes the output signal $\mathbf{y}$ and produces an appropriate input signal $\mathbf{v}_\mathrm{K}$ with $\{\mathcal{L}\mathbf{v}_\mathrm{K}\}(z) = \mathbf{G}_\mathrm{K}(z)\{\mathcal{L}\mathbf{y}\}(z)$ to be added to the external input $\mathbf{v}$ in order to stabilise the system. The resulting, feedback-controlled system has a *closed-loop transfer function* given by

$$\mathbf{G}(z) = [\mathbb{1} - \mathbf{G}_0(z)\mathbf{G}_\mathrm{K}(z)]^{-1}\mathbf{G}_0(z), \tag{1.3}$$

satisfying $\{\mathcal{L}\mathbf{y}\}(z) = \mathbf{G}(z)\{\mathcal{L}\mathbf{v}\}(z)$. The poles of the closed-loop transfer function are related to the system's stability properties. The system is stable if and only if all poles lie in the open left complex half-plane. When restricting to the imaginary axis, $z = \mathrm{i}\omega$,

the transfer function describes the frequency response of the system.

An important aspect of control theory is the *robustness* of stability. A stabilising controller is not very useful if it is not stabilising when applied to a slightly different system. A system is called robustly stable if it remains stable under any "moderately large" perturbation of the system properties (not of the system state!). Clearly, robustness is important for several reasons. First of all, the "background parameters" of the physical system under consideration (e. g., the plasma equilibrium current profile or pressure profile in the case of a tokamak) are uncertain to some extent. Furthermore, "real" sensor and actuator signals are always noisy. Finally, the computer model introduces uncertainties due to numerical approximations as well as due to physical simplifications. At first sight, it seems hard to find some general concept of robustness which accounts for all these entirely different kinds of uncertainties at a time. What is needed is a measure of the "size" of systems and system perturbations, that is, a norm in the space of all possible systems, in order to decide whether a perturbation is "large" or not. In the context of using transfer functions, the so-called $\mathcal{H}_\infty$-*norm* has become a widely accepted concept to quantify robustness, and a comprehensive theory of $\mathcal{H}_\infty$-optimal control has been developed [13]. The $\mathcal{H}_\infty$-norm of a stable system is defined in terms of its transfer function $\mathbf{G}(z)$ via

$$||\mathbf{G}||_\infty = \sup_\omega \; s_{\max}[\mathbf{G}(\mathrm{i}\omega)], \tag{1.4}$$

where $s_{\max}$ denotes the largest singular value. In the case of scalar input and output, $||\mathbf{G}||_\infty$ is equal to the frequency gain maximised over all input frequencies ω, including infinity. The details of $\mathcal{H}_\infty$ control theory are complicated. Its fundamental idea consists in considering a stable system robust if it remains stable under all perturbations whose $\mathcal{H}_\infty$-norm does not exceed some value, which is, however, large enough. The $\mathcal{H}_\infty$ theory

has already been applied to the RWM stabilisation problem using the MARS-F code [14, 15].

Concerning system norms and robustness of stability, a different philosophy is followed in this study. Here, the focus of interest is what happens inside the black box. This means, the internal dynamics of the system is primarily investigated rather than its input-output behaviour. The methodology to be derived relies on the following basic assumption:

> *Among the various possible vector norms in state space, there is one particular norm $||\cdot||$ which is of special interest to the investigator. It provides an accepted measure of the "size" of system states and, hence, of differences between states.*

Typically, such a norm can be established if the state space variables of the dynamical system are quantities which are physically measurable (in principle), or functions of such quantities. Then, any norm of interest will be based on physical grounds, since it is a function of physically measurable quantities. In the case of the numerical treatment of RWM stabilisation as presented in this paper, the above assumption is assumed to hold because a physically based state space norm will be derived in section 3.

A norm which should be used to measure the size of systems and system perturbations can be derived immediately from the above assumption: because the linear system is represented by a matrix, the system norm should be the matrix norm which is vector-bound to the vector norm $||\cdot||$ provided by the assumption. In the following, the system matrix $\mathbf{A}$ will be assumed to include the effects of the feedback controller (in the case without feedback, $\mathbf{A} = \mathbf{A}_0$ except for a similarity transformation to be explained in section 3), so that the system dynamics is governed by $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. The system norm will

then be defined by

$$||\mathbf{A}|| = \max_{||x||=1} ||\mathbf{A}\mathbf{x}||. \tag{1.5}$$

Consequently, a stable system $\mathbf{A}$ (meaning that all eigenvalues of $\mathbf{A}$ lie in the open left complex half-plane) is considered robustly stable if $\mathbf{A}+\mathbf{E}$ is stable for all perturbations $\mathbf{E}$ satisfying $||\mathbf{E}|| < \epsilon$ with ϵ being "sufficiently" large. For any stable $\mathbf{A}$, the *stability radius* $\rho(\mathbf{A})$ is defined to be equal to the supremum of all ϵ values for which the above condition holds, and it serves as a measure of robustness. These definitions of system norms and robustness are meaningful because $||\mathbf{E}||$ is equal to the maximum possible norm of the $\mathbf{E}$-induced change in the tendency $\dot{\mathbf{x}}$ provided that $||\mathbf{x}|| = 1$. Finally, following the methodology introduced in [16], optimally robust control is achieved by maximising $\rho(\mathbf{A})$ over the set of all feasible feedback logics, i. e., over all the corresponding modifications of $\mathbf{A}$. As explained in [16], this approach is closely related to the concept of matrix *pseudospectra* [17], which is also the fundament of the present study. The ϵ-pseudospectrum of $\mathbf{A}$ is the union of all the eigenvalue spectra occurring for all perturbed matrices $\mathbf{A}+\mathbf{E}$ with $||\mathbf{E}|| < \epsilon$.

It is important to note that the usage of transfer functions and the application of the $\mathcal{H}_\infty$ theory are incompatible with the concept of robustness proposed here. The input-output characteristics have to be independent of the internal state space coordinate system used to represent $\mathbf{x}$ and $\mathbf{A}_0$. Indeed, using (1.1), the transfer functions (1.2) and (1.3) can be shown to be invariant under any similarity transformation $\mathbf{x}' \leftarrow \mathbf{T}^{-1}\mathbf{x}$, $\mathbf{A}_0' \leftarrow \mathbf{T}^{-1}\mathbf{A}_0\mathbf{T}$. Hence, the system's $\mathcal{H}_\infty$-norm (1.4) is invariant under oblique coordinate transformations, but it's matrix norm (1.5) is not. ‡ The latter is invariant

‡ Typically, in robust control applications using the $\mathcal{H}_\infty$-norm, more generalised versions of the problem (1.1), including more than one (scalar or vector-valued) input and output, are solved. The resulting transfer functions differ from (1.2) and (1.3), but the issue remains that they are invariant under oblique

only under orthogonal (or unitary) transformations, because only these transformations preserve the given state space norm and the induced metrics. It follows that, for systems where a certain state space norm is of particular interest, the information contained in the $\mathcal{H}_\infty$-norm is independent of any properties of that state space norm and is therefore insufficient to assess robustness in the sense explained above. The transfer function is an incomplete characterisation of such kind of a system.

The paper is structured as follows. In section 2, the fundamentals of the used numerical model of the linearised RWM dynamics (the STARWALL code) are briefly explained, and it is shown in detail how the implementation of the feedback logics casts the RWM control problem into the structure of a parametrised eigenvalue problem. These parameters need to be optimised in order to achieve robust stability. The optimisation task is feasible only if the state space dimension of the model is substantially reduced. To this end, a novel model reduction procedure has been developed which is described in section 3. The reduction is based upon projection onto an appropriate subspace. This projection has to be orthogonal in order to preserve the state space norm of interest which is introduced in the same section. In section 4, several concepts and objectives of controller optimisation are formulated, with emphasis on pseudospectra and related ideas. Their numerical implementation in the newly developed OPTIM code package is briefly discussed. Because the proposed objective functions have some peculiar properties, their optimisation as performed by OPTIM requires special explanation, which is given in section 5. The methodology is applied to a computational example related to ITER, but slightly simplified, as described in section 6. Finally, the work is summarised and discussed in section 7.

coordinate transformations, and hence the $\mathcal{H}_\infty$-norm.

2. Structure of the feedback control problem

In the following three subsections, the mathematical structure of the RWM feedback control problem is explained. The result is an eigenvalue problem of a matrix depending on several parameters, which require optimisation in order to robustly stabilise the system.

Subsection 2.1 contains a brief description of the physics underlying the STARWALL code which is used to compute the stability of RWMs in the presence of a conducting wall and a system of active feedback coils. The structure of the corresponding eigenvalue equation will be presented. The feedback logics relating the actions of the coils to the signals measured by a set of magnetic field sensors is described in subsection 2.2. This feedback logics will be encoded by a small number of free design parameters. In subsection 2.3, inequality constraints for the parameters will be derived which can be used to account for technological limits of the feedback system. Finally, in subsection 2.4 it is demonstrated how the eigenvalue equation has to be modified in order to account for the effect of a time delay between the sensor measurements and the actions taken by the coils. The feedback control problem is presented in its final form before the model reduction step.

2.1. RWM stability calculations including a feedback system

STARWALL is a 3D ideal MHD stability code specialised to resistive wall modes [11, 12, 18]. It forms a self-consistent MHD-electrodynamical model of the system composed of the plasma, the conducting wall, and the active coils. The plasma dynamics is linearised about a prescribed ideal MHD equilibrium. The basic equations of the code

stem from the Extended Energy Principle of ideal MHD [19, 20], from the boundary conditions at the plasma-vacuum interface and the wall, and from circuit equations for the coils. The Extended Energy Principle is modified by neglecting the kinetic energy of the plasma. This is admissible for plasma motions on typical RWM timescales. Under this approximation, the plasma state can be eliminated from the equations, and the system state is uniquely determined by the current density distribution on the conducting wall and by the coil currents.

The coils are modelled as thin, ribbon-like conductors with only one winding. The wall is also assumed to be thin, and the currents flowing in it are divergence-free. The surface current density $\mathbf{j}$ is determined by a current potential ϕ via

$$\mathbf{j} = \mathbf{n} \times \nabla\phi, \tag{2.1}$$

where $\mathbf{n}$ is a surface normal pointing outward. The wall is discretised into a triangular mesh. The current density is assumed to be constant on each triangle and can be derived from the ϕ values at the corresponding three vertices. Thus, together with the coil currents, the current potentials at the nodes of the wall mesh are the state variables of the discretised system. The wall mesh can have holes and may consist of more than one connected component. The current potential takes the same value on all nodes attached to a hole, and this value forms a single state variable for each hole, respectively. On one particular node per connected component, the ϕ value is set to zero and removed from the state vector. This corresponds to fixing a gauge constant for ϕ on that connected component and prevents singularity of some of the matrices introduced below.

In addition the 3D geometry of the wall and the coils, the positions and orientations of

the sensors, which measure deviations from the equilibrium magnetic field, are required as input to STARWALL. Furthermore, the shape of the plasma-vacuum interface is needed as well as the MHD force operator matrix. These data are provided as output of the 3D ideal MHD stability code CAS3D [21]. In CAS3D, the plasma displacement vector is expanded in terms of Fourier modes (characterised by poloidal and toroidal Fourier indices m and n) on each flux surface accounted for in the radial discretisation, giving rise to the matrix representation of the force operator. Recently, the possibility to refine the radial discretisation near flux surfaces where the safety factor becomes rational has been added to CAS3D. This grid refinement avoids unphysical spikes in the displacement vector field near the rational surfaces and improves accuracy of the computations. In the case that non-axisymmetric geometries are considered in the STARWALL computation, it is advisable to include in the CAS3D Fourier expansion at least all those n's simultaneously for which the equilibrium is unstable with respect to external kinks in absence of the wall. The coupling between different n's induced by the breaking of the axisymmetry can be significant [12].

The linear dynamics of the plasma-wall-coils system is governed by

$$\mathbf{L}\dot{\mathbf{x}} = \mathbf{R}\mathbf{x}, \tag{2.2}$$

where $\mathbf{x} \in \mathbb{R}^N$ is the state vector of the system. The first N_c components of $\mathbf{x}$ contain the coil currents (N_c is the number of coils), and the remaining components represent the current potential values at the nodes of the wall mesh. The symmetric matrix $\mathbf{L}$ is the sum of an inductance matrix and a matrix describing the effects due to the presence of the plasma, and

$$\mathbf{R} = -\mathbf{R}_0 + \mathbf{U} \tag{2.3}$$

consists of the symmetric, positive definite resistance matrix $\mathbf{R}_0$ and the matrix $\mathbf{U}$ mapping the system state $\mathbf{x}$ onto external voltages applied to the feedback coils. $\mathbf{U}$ has non-zero entries in the first N_c rows only and will be specified in the next subsection. Substituting a time dependence $\mathbf{x} \sim \mathrm{e}^{\gamma t}$ and inverting $\mathbf{L}$, one obtains the eigenvalue problem

$$\mathbf{A}_F \mathbf{x} = \gamma \mathbf{x}, \tag{2.4}$$

where

$$\mathbf{A}_F = \mathbf{A}_0 + \mathbf{F} \tag{2.5}$$

is the system matrix composed of the open-loop part

$$\mathbf{A}_0 = -\mathbf{L}^{-1}\mathbf{R}_0 \tag{2.6}$$

and the feedback part

$$\mathbf{F} = \mathbf{L}^{-1}\mathbf{U}. \tag{2.7}$$

The system is stable if and only if all eigenvalues of $\mathbf{A}_F$ have negative real parts. Without feedback ($\mathbf{U} = 0$), all eigenvalues are real, and positive ones belong to unstable RWMs. The feedback control problem consists in choosing $\mathbf{U}$ properly so that $\mathbf{A}_F$ becomes robustly stable.

2.2. The feedback controller model

As already adumbrated, a voltage control feedback model is adopted in this study. That means, depending on the system state, or, more precisely, on sensor measurements,

voltages are applied to the feedback coils. This dependence is described by a matrix of gain factors which maps the vector of sensor measurements onto the vector of coil voltages. The elements of this gain matrix will be optimised within a physically reasonable matrix subspace. However, for such kind of a proportional gain voltage controller, the promptness of the feedback system response is limited by the L/R time constants of the coils. To circumvent this problem, the coil currents are taken as additional measured quantities which are fed back onto the coil voltages by additional gain factors. These act exactly like additional artificial resistances serially connected with the coils and therefore effectively reduce the coil time constants. Larger total dc resistances ("natural" plus artificial resistances) have to be compensated by larger voltage gains to some extent. The artificial resistances are not fixed a priori, but are used as additional degrees of freedom for optimisation. Namely, it is not clear a priori if the fastest available response is also the most favourable one, or if the integrating character of a slower response can also have desirable effects.

To construct the gain matrix, all coils and sensors, respectively, are subdivided into one or more toroidal arrays, depending on the configuration. All coils within an array are assumed to be identical in construction and to have the same poloidal and radial position, and the same orientation. The same assumption should hold for all sensors in an array. Each coil array k ($k = 1, \ldots, K$) is linked to each sensor array l ($l = 1, \ldots, L$) via a gain sub-matrix $\mathbf{G}^{kl}$. In addition, the current flowing in each coil in the k-th array is fed back onto that coil's voltage by an additional gain factor $-\tilde{R}_k$ which is equal for each member of the array. It can be shown that this procedure, in principle, is equivalent to current control by means of cascade control.§ Summarising, for $k = 1, \ldots, K$, the

§ Application of a controller voltage U to a serial connection of a coil with resistance R and an additional resistance $\tilde{R}$ gives a current $I(t)$ and a voltage drop $U(t) = (R + \tilde{R})I - \tilde{R}I(t)$ at the coil, where $I = U/(R + \tilde{R})$ is the saturation current. This voltage controller is equivalent to a current controller subdivided into an "exterior" and an "interior" controller, where $\tilde{R}$ is merely a numerical

voltage vector

$$\mathbf{u}^k = \sum_{l=1}^{L} \mathbf{G}^{kl}\mathbf{s}^l - \tilde{R}_k \mathbf{i}^k, \tag{2.8}$$

is applied to the coils in the k-th toroidal array, where $\mathbf{i}^k$ is the vector of currents already flowing in these coils, and $\mathbf{s}^l$ is the signal vector measured by the l-th sensor array. If the system is in state $\mathbf{x}$, the sensor matrix $\mathbf{S}^l$ delivers $\mathbf{s}^l$:

$$\mathbf{s}^l = \mathbf{S}^l\mathbf{x}, \quad l = 1, \ldots, L. \tag{2.9}$$

To reduce the number of design parameters in a physically reasonable manner, the gain sub-matrix $\mathbf{G}^{kl}$ is constructed for each $k = 1, \ldots, K$, $l = 1, \ldots, L$ in such a way that coil array k produces a field with toroidal Fourier index n in response to magnetic field perturbations with the same n, measured by sensor array l. This response is a linear combination of a response having the same toroidal phase as the perturbation, and another one which is toroidally phase shifted by $90/n$ degrees. The in-phase response is described by a basic gain matrix $\mathbf{G}^{kl}_{n,\alpha}$, the phase-shifted response by $\mathbf{G}^{kl}_{n,\beta}$. $\mathbf{G}^{kl}$ is a linear combination of these basic matrices:

$$\mathbf{G}^{kl} = \sum_n \alpha_n^{kl}\mathbf{G}^{kl}_{n,\alpha} + \beta_n^{kl}\mathbf{G}^{kl}_{n,\beta}, \tag{2.10}$$

where the sum runs over all n's to be controlled. The elements of $\mathbf{G}^{kl}_{n,\alpha}$ and $\mathbf{G}^{kl}_{n,\beta}$ are given by

$$(\mathbf{G}^{kl}_{n,\alpha})_{ij} = \cos(n\varphi^{kl}_{ij}) \tag{2.11}$$

constant determining these controllers' properties. In response to a sensor signal, the exterior controller "demands" a current I, where the current gain equals the voltage gain divided by $R + \tilde{R}$. The interior controller measures the actual coil current $I(t)$, computes the auxiliary quantity $I'(t) = (\tilde{R}/(R+\tilde{R}))I(t)$ and applies the voltage $U(t) = -(R + \tilde{R})(I'(t) - I)$, being equal to the voltage drop mentioned above, to the coil, so that $I(t)$ approaches I.

and

$$(\mathbf{G}^{kl}_{n,\beta})_{ij} = \sin(n\varphi^{kl}_{ij}), \tag{2.12}$$

where φ^{kl}_{ij} is the toroidal angle between coil i of coil array k and sensor j of sensor array l. The free parameters α^{kl}_n (gains for in-phase responses) and β^{kl}_n (gains for phase-shifted responses) are left for optimisation, together with $\tilde{R}_k$.

As will be shown in the following, this feedback controller model can be compactly written down by deriving an expression for the voltage matrix $\mathbf{U}$ introduced in equation (2.3). With proper numbering of the coil current components in the state vector $\mathbf{x}$ (the first N_c components) and considering (2.8), (2.9), and (2.10), one finds

$$\begin{aligned}\mathbf{U} &= \begin{pmatrix} \mathbf{U}^1 \\ \mathbf{U}^2 \\ \vdots \\ \mathbf{U}^K \\ 0 \in \mathbb{R}^{(N-N_\mathrm{c})\times N} \end{pmatrix} - \tilde{R}_k \mathbf{I}^k \\ &= \sum_{k=1}^{K} \left(\sum_n \sum_{l=1}^{L} \alpha^{kl}_n \mathbf{U}^{kl}_{n,\alpha} + \beta^{kl}_n \mathbf{U}^{kl}_{n,\beta} \right) - \tilde{R}_k \mathbf{I}^k, \end{aligned} \tag{2.13}$$

that means, $\mathbf{U}^k = \sum_n \sum_{l=1}^{L} (\alpha^{kl}_n \mathbf{G}^{kl}_{n,\alpha} + \beta^{kl}_n \mathbf{G}^{kl}_{n,\beta}) \mathbf{S}^l$ for $k = 1, \ldots, K$, so that $\mathbf{U}$ is a linear

combination of the basic voltage matrices

$$\mathbf{U}_{n,\alpha}^{kl} = \begin{pmatrix} 0 \\ \vdots \\ \mathbf{G}_{n,\alpha}^{kl}\mathbf{S}^{l} \\ \vdots \\ 0 \\ 0 \in \mathbb{R}^{(N-N_\mathrm{c})\times N} \end{pmatrix}, \quad \mathbf{U}_{n,\beta}^{kl} = \begin{pmatrix} 0 \\ \vdots \\ \mathbf{G}_{n,\beta}^{kl}\mathbf{S}^{l} \\ \vdots \\ 0 \\ 0 \in \mathbb{R}^{(N-N_\mathrm{c})\times N} \end{pmatrix}, \tag{2.14}$$

and the basic current matrices $\mathbf{I}^k$ with the elements

$$(\mathbf{I}^k)_{ij} = \begin{cases} \delta_{ij} & , \quad i \leq N_\mathrm{c} \text{ and coil } i \text{ belongs to array } k \\ 0 & , \quad \text{otherwise.} \end{cases} \tag{2.15}$$

Finally, it follows from (2.7) and (2.13) that the feedback matrix $\mathbf{F}$ contributing to the system matrix $\mathbf{A}_\mathrm{F}$, cf. (2.5), depends on the parameters as follows:

$$\mathbf{F} = \sum_{k=1}^{K} \left(\sum_{n} \sum_{l=1}^{L} \alpha_n^{kl}\mathbf{F}_{n,\alpha}^{kl} + \beta_n^{kl}\mathbf{F}_{n,\beta}^{kl} \right) - \tilde{R}_k\mathbf{F}_\mathrm{I}^k \tag{2.16}$$

with the basic feedback matrices $\mathbf{F}_{n,\alpha}^{kl} = \mathbf{L}^{-1}\mathbf{U}_{n,\alpha}^{kl}$, $\mathbf{F}_{n,\beta}^{kl} = \mathbf{L}^{-1}\mathbf{U}_{n,\beta}^{kl}$, and $\mathbf{F}_\mathrm{I}^k = \mathbf{L}^{-1}\mathbf{I}^k$. Thus, the eigenvalue problem (2.4) has become a parametrised one.

2.3. Constraints

In most cases, the parameter optimisation should be carried out subject to constraints. Such constraints are useful in ensuring technological feasibility of the parameters. Furthermore, termination of the optimisation procedure can only be guaranteed if the allowable parameter values form a bounded set in parameter space. In principle, it is possible to limit the coils' time constant reciprocals, voltages, and currents.

It makes sense to prevent the time constants from becoming arbitrarily small by imposing a lower bound $T^k_{\min}$ for each coil array k:

$$\frac{L^k}{R^k + \tilde{R}^k} \geq T^k_{\min}, \tag{2.17}$$

where L^k is the self-inductance of each coil in array k, and R_k is its "natural" dc resistance.

Constraining the voltage and current is more difficult. Because the control problem is a linear one and the voltages applied in response to a measured sensor signal are proportional to that signal, some estimate concerning the maximum signal magnitude $S_{\max}$ has to be given in order to formulate the voltage constraint. Here, the following simplifying assumptions are made, similar to those described in [9]: $S_{\max}$ is equal to an RWM detection limit, or noise level. As long as the sensor signal is below $S_{\max}$, the feedback system does not react at all. As soon as the threshold $S_{\max}$ is exceeded, the feedback is switched on and counteracts the origin of the sensor signal. Soon, the latter drops below $S_{\max}$ again, so that $S_{\max}$ is never significantly exceeded.

Once a value for $S_{\max}$ is fixed, the voltage constraint can be written for each coil array $k = 1, \ldots, K$ as

$$\left\{ \frac{1}{2} \sum_n \sum_{l=1}^{L} \left(\frac{N^l_{\mathrm{s}}}{2} \right)^2 \left[(\alpha^{kl}_n)^2 + (\beta^{kl}_n)^2 \right] \right\}^{\frac{1}{2}} S_{\max} \leq U^k_{\max}, \tag{2.18}$$

where N^l_{s} is the number of sensors in array l, and $U^k_{\max}$ is some, for example technologically caused voltage limit for the coils in array k. The expression on the left hand side is approximately equal to the root mean square voltage applied to coil array k (mean taken over all coils in the array) in response to a measured magnetic field perturbation with amplitude $S_{\max}$; it has been derived from (2.8), (2.10), (2.11), and

(2.12). It makes sense to consider the root mean square voltage and current because the voltage and current can be assumed to be alternating. Stabilised RWMs, and therefore also the distribution of voltages and currents among coils, typically rotate in the toroidal direction, with frequencies comparable to the RWM growth rates without feedback, but typically much lower than the coil time constant reciprocals. From the latter fact it follows that the instantaneous current is always approximately equal to the saturation current which corresponds to the instantaneous voltage divided by the total dc resistance. The current distribution among coils has virtually the same toroidal phase as the voltage distribution. Consequently, a current constraint can be derived from (2.18) and reads

$$\frac{\left\{\frac{1}{2}\sum_n \sum_{l=1}^{L} \left(\frac{N_{\mathrm{S}}^l}{2}\right)^2 \left[(\alpha_n^{kl})^2 + (\beta_n^{kl})^2\right]\right\}^{\frac{1}{2}} S_{\max}}{R_k + \tilde{R}_k} \leq I_{\max}^k, \tag{2.19}$$

with $I_{\max}^k$ being the current limit to be set for the coils in array k.

The above assumptions concerning the evolution of the sensor signal magnitude after it exceeds $S_{\max}$ have to be taken with an pinch of salt, since its actual evolution strongly depends on the initial condition. If the initial state (the state at the moment when the sensors start to detect a signal, as it exceeds $S_{\max}$) is largely an RWM, the assumptions should be reasonably justified, since the active coils indeed counteract the origin of the detected signal, namely, the RWM. This is no longer true for different initial states, where significant transient amplification of the sensor signal might actually occur after exceeding the detection limit. First of all, an initial sensor signal might be due to small-scale eddy currents, not related to an RWM, in the wall parts close to the sensors. In such a case, the action of the feedback system will erroneously and rapidly excite an RWM, which, however, will be controlled later on if the system is stable (an example will

be discussed in section 6). This process will probably be accompanied with a transient increase of the sensor signal magnitude. Second, one has to take into account the fact that the sensors "see" only a low dimensional subspace of the state space, that means, there is a high-dimensional subspace which is invisible to the sensors. To any initial state vector, an arbitrarily large portion from the invisible subspace can be added without changing the instantaneous sensor measurement. Since the invisible subspace cannot be expected to be invariant under the system dynamics, the large, initially invisible portion will increasingly project onto the visible subspace as it dynamically evolves. By this means, initial states with arbitrarily large transient amplification of the sensor signal can be constructed theoretically. However, such initial state vectors have to be large themselves.

Fortunately, initial states which produce large sensor signal amplifications lie in the stable subspace of the open-loop system, because they correspond to current distributions in the wall which are unrelated to RWMs. Their spontaneous excitation can be caused only by physical processes not included in the model used here (e. g., *edge localised modes (ELMs)*). Estimating the amplitude and relevance of such perturbations would require nonlinear simulations of more complete, forced-dissipative MHD models, but is not possible here. Therefore, the possibly overoptimistic constraints (2.18) and (2.19) are relied upon, noting that, if some amount of transient sensor signal amplification is expected, it could be incorporated as a factor into the estimate of $S_{\max}$.

2.4. Effect of time delay between sensors and actuators

In the following, a technically caused time delay between an "event" detected by the sensors and the coil voltage application in response to that event is accounted for. This

delay can be characterised by an impulse response $r(t)$ satisfying $r(t) = 0$ for $t < 0$ and

$$\int_0^\infty r(t)\mathrm{d}t = 1. \tag{2.20}$$

For simplicity, the impulse response is here assumed to be the same for all magnetic field sensors and all ammeters at the coils. But the following explanations can readily be generalised. The delay causes that, at time t, the sensors do not "see" the present system state, but a combination of states from the past. It follows from (2.5) that the system dynamics is then governed by

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0\mathbf{x}(t) + \int_0^\infty r(t')\mathbf{F}\mathbf{x}(t-t')\mathrm{d}t'. \tag{2.21}$$

Now, a new matrix $\hat{\mathbf{A}}_\mathrm{F}$ is sought which describes the system dynamics so that $\dot{\mathbf{x}}(t) = \hat{\mathbf{A}}_\mathrm{F}\mathbf{x}(t)$ and therefore $\mathbf{x}(t) = \exp(t\hat{\mathbf{A}}_\mathrm{F})\mathbf{x}(0)$. It follows that $\mathbf{x}(t-t') = \exp(-t'\hat{\mathbf{A}}_\mathrm{F})\mathbf{x}(t)$ and

$$\hat{\mathbf{A}}_\mathrm{F} = \mathbf{A}_0 + \int_0^\infty r(t')\mathbf{F}\exp(-t'\hat{\mathbf{A}}_\mathrm{F})\mathrm{d}t'. \tag{2.22}$$

Depending on the particular form of $r(t)$, an exact solution for $\hat{\mathbf{A}}_\mathrm{F}$ might be difficult. But, under the assumption that $r(t)$ is essentially zero for all $t > q$ and some q which is small compared to all time scales occurring in the system dynamics, the exponential in (2.22) can be approximated to first order in t'. With (2.20) and the "average delay time"

$$\tau = \int_0^\infty r(t)t\mathrm{d}t, \tag{2.23}$$

one finally arrives at the solution

$$\hat{\mathbf{A}}_\mathrm{F} = (\mathbb{1} + \tau\mathbf{F})^{-1}(\mathbf{A}_0 + \mathbf{F}). \tag{2.24}$$

Under the approximation used, $r(t)$ has exactly the same effect as a pure time delay $r_{\mathrm{d}}(t) = \delta(t - \tau)$. The final form of the unreduced, parametrised eigenvalue problem reads

$$\hat{\mathbf{A}}_{\mathrm{F}}\mathbf{x} = \gamma\mathbf{x}, \tag{2.25}$$

with $\hat{\mathbf{A}}_{\mathrm{F}}$ given by (2.24) and the parameter dependence residing in $\mathbf{F}$ as given by (2.16).

3. Model reduction

For the full-sized matrix $\hat{\mathbf{A}}_{\mathrm{F}}$, the parameter optimisation is computationally not feasible since the state space dimension N is typically of order 10^4. A model reduction procedure, which dramatically reduces the system dimension, but retains as much information of the original system as possible, is absolutely necessary. More precisely, such kind of a reduction should be arranged in such a way that the neglected part of the system dynamics is virtually unaffected by the presence of a feedback system, whereas the retained part responds essentially in the same way to the feedback as the original system does. A typical approach, which is also adopted here, consists of two steps: First, the unstable subspace of the system without feedback is split off. Second, the stable subspace is projected onto a further, "dominant" subspace by means of a similarity transform and truncation (deletion of "non-dominant" rows and columns). Finally, the unstable subspace is given back to the reduced system. The various truncation-based model reduction methods which can be found in the literature [22, 23, and references therein] differ from each other in the choice of the similarity transform, or, in other words, in the measure of dominance. For reasons already explained in the introduction, a novel reduction method will be introduced here which is based on an orthogonal similarity

transform and thus preserves the metrics in the state space, in contrast to the methods available in the literature. This procedure shall be denoted *isometric truncation*.

As the first step, the separation of the unstable subspace is accomplished by computing all eigenvalues and eigenvectors of the open-loop system. The unstable subspace is spanned by the eigenvectors belonging to positive eigenvalues. But the solution of the eigenvalue problem also serves another purpose. Namely, a physically based state space metrics is established, after the system is subjected to a preliminary similarity transform defined by the eigenvector matrix, which is orthogonal not in the strict, but in a more general sense, as it will be explained in subsection 3.1. The second step consists of the isometric truncation and will be described in subsection 3.2.

3.1. Open-loop eigenvalue problem and the state space norm

The open-loop eigenvalue problem $\mathbf{A}_0\mathbf{x}_i = \gamma_i\mathbf{x}_i$, $i = 1, \ldots, N$ is most efficiently solved by first computing the solution of the generalised eigenvalue problem

$$\mathbf{L}\mathbf{x}_i = \mu_i\mathbf{R}_0\mathbf{x}_i, \tag{3.1}$$

where $\mathbf{L}$ and $\mathbf{R}_0$ are symmetric and $\mathbf{R}_0$ is positive definite, and then substituting $\gamma_i = -\mu_i^{-1}$, cf. (2.2), (2.3), and (2.4). The ordering $\gamma_i \geq \gamma_{i+1}$ is assumed in the following. The eigenvectors $\mathbf{x}_i$ are normalised so that they satisfy the generalised orthonormality condition

$$\mathbf{x}_i^{\mathrm{T}}\mathbf{R}_0\mathbf{x}_j = \delta_{ij}. \tag{3.2}$$

For any system state $\mathbf{x}$, $\mathbf{x}^{\mathrm{T}}\mathbf{R}_0\mathbf{x}$ corresponds to the ohmic loss produced by the current distribution represented by $\mathbf{x}$, since $\mathbf{R}_0$ is the resistance matrix. All eigenvectors are

normalised to generate unit ohmic loss. The norm

$$||\mathbf{x}||_{\mathbf{R}_0} := \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{R}_0\mathbf{x}} \tag{3.3}$$

is defined in terms of a physically measurable quantity, and it is considered the norm of interest. This motivates a transformation of the system states to eigenvector coefficients: $\tilde{\mathbf{x}} = \mathbf{X}^{-1}\mathbf{x}$, where $\mathbf{X}^{-1} = \mathbf{X}^{\mathrm{T}}\mathbf{R}_0$ and $\mathbf{X} = (\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_N)$. In the new coordinate system, the square root of the ohmic loss can be calculated in terms of the canonical scalar product: $||\tilde{\mathbf{x}}|| = \sqrt{\tilde{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{x}}} = ||\mathbf{x}||_{\mathbf{R}_0}$. All the system matrices are transformed as $\tilde{\mathbf{M}} = \mathbf{X}^{-1}\mathbf{M}\mathbf{X}$, where $\mathbf{M}$ has to be replaced by $\mathbf{A}_0$, $\mathbf{F}^{kl}_{n,\alpha}$, $\mathbf{F}^{kl}_{n,\beta}$, and $\mathbf{F}^{k}_{\mathrm{I}}$, respectively, cf. (2.16). Clearly, $\tilde{\mathbf{A}}_0 = \mathrm{diag}(\gamma_1, \gamma_2, \ldots, \gamma_N)$.

3.2. Isometric truncation

The stable part of the open-loop system in the new eigenvector coordinates is described by the matrix $\bar{\mathbf{A}}_0 \in \mathbb{R}^{(N-u)\times(N-u)}$, where u is the number of nonnegative, i. e., unstable eigenvalues. $\bar{\mathbf{A}}_0$ is generated from $\tilde{\mathbf{A}}_0$ by deleting the first u rows and columns: $\bar{\mathbf{A}}_0 = \mathrm{diag}(\gamma_{u+1}, \gamma_{u+2}, \ldots, \gamma_N)$. The stable part of the transformed state vector $\tilde{\mathbf{x}}$ is denoted $\bar{\mathbf{x}}$ and contains the last $N - u$ components of $\tilde{\mathbf{x}}$. To derive the isometric truncation procedure, the stable part of the control system is described by the standard "black box" formalism of control theory already described in the introduction. The system is assumed to have an actuator input vector $\mathbf{v} \in \mathbb{R}^{N_{\mathrm{i}}}$ and a sensor output vector $\mathbf{y} \in \mathbb{R}^{N_{\mathrm{o}}}$, resulting in the description

$$\begin{aligned} \dot{\bar{\mathbf{x}}} &= \bar{\mathbf{A}}_0\bar{\mathbf{x}} + \bar{\mathbf{B}}\mathbf{v}, \\ \mathbf{y} &= \bar{\mathbf{C}}\bar{\mathbf{x}} \end{aligned} \tag{3.4}$$

with input matrix $\bar{\mathbf{B}}$ output matrix $\bar{\mathbf{C}}$. In the following it will be shown how these matrices can be derived from their counterparts $\mathbf{B}$ and $\mathbf{C}$ of the full system (including the unstable part) in the original coordinate system, where the system description is given by (1.1). The matrix $\mathbf{B} \in \mathbb{R}^{N \times N_\mathrm{i}}$ ($N_\mathrm{i} = 2KN_\mathrm{t}$, where N_t is the number of different toroidal Fourier indices n to be controlled) describes the excitation of cosine and sine current patterns in the different coil arrays and with different toroidal "wave numbers" n. It has the following block structure:

$$\mathbf{B} = \mathbf{L}^{-1} \begin{pmatrix} \mathbf{B}^1 & 0 & \cdots & 0 \\ 0 & \mathbf{B}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{B}^K \\ 0 \in \mathbb{R}^{(N-N_\mathrm{c}) \times N_\mathrm{i}} & & & \end{pmatrix}. \tag{3.5}$$

For each $k = 1, \ldots, K$, the block $\mathbf{B}^k \in \mathbb{R}^{N_\mathrm{c}^k \times 2N_\mathrm{T}}$ describes the action of the k-th coil array, where N_c^k is the number of coils in array k. The block is given by

$$\mathbf{B}^k = \begin{pmatrix} \cos n_1\theta_1^k & \sin n_1\theta_1^k & \cdots & \cos n_{N_\mathrm{t}}\theta_1^k & \sin n_{N_\mathrm{t}}\theta_1^k \\ \cos n_1\theta_2^k & \sin n_1\theta_2^k & \cdots & \cos n_{N_\mathrm{t}}\theta_2^k & \sin n_{N_\mathrm{t}}\theta_2^k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \cos n_1\theta_{N_\mathrm{c}^k}^k & \sin n_1\theta_{N_\mathrm{c}^k}^k & \cdots & \cos n_{N_\mathrm{t}}\theta_{N_\mathrm{c}^k}^k & \sin n_{N_\mathrm{t}}\theta_{N_\mathrm{c}^k}^k \end{pmatrix} \tag{3.6}$$

with n_1, n_2, …, n_{N_t} being the different n's to be controlled and θ_i^k being the toroidal position (angle) of the barycentre of the i-th coil in array k. Similarly, the output matrix $\mathbf{C} \in \mathbb{R}^{N_\mathrm{o} \times N}$ ($N_\mathrm{o} = 2LN_\mathrm{t}$) projects the output vectors of the sensor arrays onto cosine

and sine patterns. It takes the form

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}^1 \\ \mathbf{C}^2 \\ \vdots \\ \mathbf{C}^L \end{pmatrix}. \tag{3.7}$$

For each $l = 1, \ldots, L$, $\mathbf{C}^l$ provides the outputs of the l-th sensor array:

$$\mathbf{C}^l = \mathbf{P}^l \mathbf{S}^l, \tag{3.8}$$

where $\mathbf{S}^l$ is the l-th sensor matrix, cf. (2.9). $\mathbf{P}^l$ is a projector onto the cosine and sine components under consideration, which reads

$$\mathbf{P}^l = \begin{pmatrix} \cos n_1\vartheta_1^l & \cos n_1\vartheta_2^l & \cdots & \cos n_1\vartheta_{N_\mathrm{s}^l} \\ \sin n_1\vartheta_1^l & \sin n_1\vartheta_2^l & \cdots & \sin n_1\vartheta_{N_\mathrm{s}^l} \\ \cos n_2\vartheta_1^l & \cos n_2\vartheta_2^l & \cdots & \cos n_2\vartheta_{N_\mathrm{s}^l} \\ \sin n_2\vartheta_1^l & \sin n_2\vartheta_2^l & \cdots & \sin n_2\vartheta_{N_\mathrm{s}^l} \\ \vdots & \vdots & \ddots & \vdots \\ \cos n_{N_\mathrm{t}}\vartheta_1^l & \cos n_{N_\mathrm{t}}\vartheta_2^l & \cdots & \cos n_{N_\mathrm{t}}\vartheta_{N_\mathrm{s}^l} \\ \sin n_{N_\mathrm{t}}\vartheta_1^l & \sin n_{N_\mathrm{t}}\vartheta_2^l & \cdots & \sin n_{N_\mathrm{t}}\vartheta_{N_\mathrm{s}^l} \end{pmatrix}. \tag{3.9}$$

Here, N_s^l is the number of sensors in array l, and ϑ_j^l is the toroidal position of the j-th sensor in that array. The corresponding matrices $\bar{\mathbf{B}}$ and $\bar{\mathbf{C}}$ in the stable subspace are generated by first computing the transformed matrices $\tilde{\mathbf{B}} = \mathbf{X}^{-1}\mathbf{B}$ and $\tilde{\mathbf{C}} = \mathbf{C}\mathbf{X}$ and then deleting the first u rows in $\tilde{\mathbf{B}}$ to obtain $\bar{\mathbf{B}}$ and the first u columns in $\tilde{\mathbf{C}}$ to get $\bar{\mathbf{C}}$.

Before returning to the model reduction problem, it should be noted that the above definitions of the matrices $\mathbf{B}$ and $\mathbf{C}$ are consistent with the derivation of the controller

model in subsection 2.2. This can be seen by concluding that, in the case that the artificial coil resistances $\tilde{R}_k$ are zero for all $k = 1, \ldots, K$ (this case is assumed throughout the model reduction procedure, yielding reduced models which work equally well for nonzero $\tilde{R}_k$ unless the natural dc resistances R_k are extremely small, as for superconducting coils), the controller model is equivalent to a particular *static output feedback* applied to (1.1). That means, the feedback matrix $\mathbf{F}$ derived in subsection 2.2 can be reconstructed by connecting the input $\mathbf{v}$ with the output $\mathbf{y}$ via a suitably chosen controller matrix $\mathbf{K}$, so that $\mathbf{v} = \mathbf{K}\mathbf{y} = \mathbf{K}\mathbf{C}\mathbf{x}$ and therefore $\mathbf{F} = \mathbf{B}\mathbf{K}\mathbf{C}$, cf. (1.1) and (2.5). Using the definition $\varphi_{ij}^{kl} = \theta_i^k - \vartheta_j^l$, see (2.11) and (2.12), it can be shown that $\mathbf{B}\mathbf{K}\mathbf{C}$ is equal to (2.16) for $\tilde{R}_k = 0$, $k = 1, \ldots, K$ if the controller is chosen as

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}^{11} & \mathbf{K}^{12} & \cdots & \mathbf{K}^{1L} \\ \mathbf{K}^{21} & \mathbf{K}^{22} & \cdots & \mathbf{K}^{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{K}^{K1} & \mathbf{K}^{K2} & \cdots & \mathbf{K}^{KL} \end{pmatrix}, \tag{3.10}$$

where $\mathbf{K}^{kl} = \mathrm{diag}(\mathbf{k}_{n_1}^{kl}, \mathbf{k}_{n_2}^{kl}, \ldots, \mathbf{k}_{n_{N_\mathrm{t}}}^{kl})$ with the (2×2)-blocks

$$\mathbf{k}_{n_i}^{kl} = \begin{pmatrix} \alpha_{n_i}^{kl} & -\beta_{n_i}^{kl} \\ \beta_{n_i}^{kl} & \alpha_{n_i}^{kl} \end{pmatrix}, \quad i = 1, \ldots, N_\mathrm{t}. \tag{3.11}$$

Almost all truncation-based model reduction methods currently available in the literature rely on oblique transformations. Among the most popular methods are *balanced truncation* and related procedures [24, 25, 26, 27, 28, 29, 30, 23, 22]. The isometric truncation presented here is based upon an orthogonal transformation, but the basic idea used to define this transformation is exactly the same as in Moore's original work on balanced truncation [24], namely, the analysis of the response to two different kinds of test signals injected into the control system (3.4) with the initial condition $\bar{\mathbf{x}} = 0$

for all $t < 0$. The first kind of test signals is defined by $\mathbf{v}(t) = \mathbf{e}_i \delta(t)$, $i = 1, \ldots, N_\mathrm{i}$ with the i-th canonical unit vector $\mathbf{e}_i \in \mathbb{R}^{N_\mathrm{i}}$. The corresponding state vector response time series $\bar{\mathbf{x}}_i(t)$ are collected into a matrix $\bar{\mathbf{X}}(t) = (\bar{\mathbf{x}}_1(t)\ \bar{\mathbf{x}}_2(t)\ \ldots\ \bar{\mathbf{x}}_{N_\mathrm{i}}(t)) = \exp(t\bar{\mathbf{A}}_0)\bar{\mathbf{B}}$. It can be expected that, for all $t > 0$, the columns of $\bar{\mathbf{X}}(t)$ project most strongly onto those state space directions which can most easily be controlled by the actuators, and the aim is to extract these directions. Similarly, the second kind of test signals is defined by adding a term $\mathbf{e}_i \delta(t)$ for $i = 1, \ldots, N$, with $\mathbf{e}_i \in \mathbb{R}^N$ now being the i-th canonical unit vector in state space, to the right hand side of the equation for $\dot{\bar{\mathbf{x}}}$ in (3.4). The corresponding output vector responses $\mathbf{y}_i(t)$ are collected into $\mathbf{Y}(t) = (\mathbf{y}_1(t)\ \mathbf{y}_2(t)\ \ldots\ \mathbf{y}_N(t)) = \bar{\mathbf{C}} \exp(t\bar{\mathbf{A}}_0)$. For all $t > 0$, the columns of $\mathbf{Y}^\mathrm{T}(t)$ should project most strongly onto those state space directions which can most easily be observed by the sensors, and, again, one aims at finding these directions. In the model reduction procedure, the directions which are both almost (or exactly) uncontrollable and unobservable are discarded.

The solution of the problem to find low dimensional subspaces which the columns of $\bar{\mathbf{X}}(t)$ and $\mathbf{Y}^\mathrm{T}(t)$ primarily project onto is given by *principal component analysis* (PCA) [31], a statistical data analysis method which is widely used in various scientific disciplines ranging from meteorology to psychology. Given a matrix $\mathbf{Z}(t) \in \mathbb{R}^{P \times Q}$ ($t \in [t_1\ t_2]$), whose columns contain Q vector time series, PCA yields a set of basis vectors $\mathbf{w}_i \in \mathbb{R}^P$ and time-dependent coefficient vectors $\mathbf{c}_i(t) \in \mathbb{R}^Q$, $i = 1, \ldots, P$, so that the columns of $\mathbf{Z}(t)$ are expanded in terms of the $\mathbf{w}_i$'s:

$$\mathbf{Z}(t) = \sum_{i=1}^{P} \mathbf{w}_i \mathbf{c}_i^\mathrm{T}(t). \tag{3.12}$$

The $\mathbf{w}_i$'s are chosen in such a way that, for all $k < N$, the error introduced by a projection of the columns of $\mathbf{Z}(t)$ onto the k-dimensional subspace spanned by

$\mathbf{w}_1, \ldots, \mathbf{w}_k$ becomes minimal in a least squares sense:

$$\int_{t_1}^{t_2} \sum_{i=k+1}^{P} ||\mathbf{w}_i \mathbf{c}_i^{\mathrm{T}}(t)||_{\mathrm{F}}^2 \mathrm{d}t = \min, \tag{3.13}$$

where $||\cdot||_{\mathrm{F}}$ is the Frobenius norm. The solution of this optimisation problem is to choose the $\mathbf{w}_i$'s as eigenvectors of the *Gramian*, or *covariance matrix*

$$\mathbf{W} = \int_{t_1}^{t_2} \mathbf{Z}(t)\mathbf{Z}^{\mathrm{T}}(t)\mathrm{d}t. \tag{3.14}$$

Since $\mathbf{W}$ is symmetric and positive semidefinite, its eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_P$ are orthogonal, and its eigenvalues $w_1 \geq w_2 \geq \ldots \geq w_P$ are nonnegative. For all $i = 1, \ldots, P$, the ratio $w_i / \sum_{i=1}^{P} w_i$ describes the fraction at which the variance of the matrix time series $\mathbf{w}_i \mathbf{c}_i^{\mathrm{T}}(t)$ contributes to the total variance of $\mathbf{Z}(t)$. Therefore, w_i is an "importance" measure for $\mathbf{w}_i$.

It follows that, when looking for the system states which are most easily controllable, one should carry out a PCA of $\bar{\mathbf{X}}(t) = \exp(t\bar{\mathbf{A}}_0)\bar{\mathbf{B}}$ and consider the leading eigenvectors of the *controllability Gramian*

$$\mathbf{W}_{\mathrm{c}} = \int_0^{\infty} \exp(t\bar{\mathbf{A}}_0)\bar{\mathbf{B}}\bar{\mathbf{B}}^{\mathrm{T}} \exp(t\bar{\mathbf{A}}_0^{\mathrm{T}})\mathrm{d}t. \tag{3.15}$$

Analogously, a PCA of $\bar{\mathbf{Y}}^{\mathrm{T}}(t) = \exp(t\bar{\mathbf{A}}_0^{\mathrm{T}})\bar{\mathbf{C}}^{\mathrm{T}}$ can be performed to find the most easily observable states as the leading eigenvectors of the *observability Gramian*

$$\mathbf{W}_{\mathrm{o}} = \int_0^{\infty} \exp(t\bar{\mathbf{A}}_0^{\mathrm{T}})\bar{\mathbf{C}}^{\mathrm{T}}\bar{\mathbf{C}} \exp(t\bar{\mathbf{A}}_0)\mathrm{d}t. \tag{3.16}$$

In many PCA applications, the time integration in (3.14) has to be carried out numerically by sampling‖ $\mathbf{Z}(t)$, but in the case of (3.15) and (3.16) there exists a

‖ Actually, in most cases the data to be analysed are present as samples from the very beginning, and sometimes the discretised "time" merely plays the role of sample numbering.

convenient alternative. Namely, $\mathbf{W}_\mathrm{c}$ and $\mathbf{W}_\mathrm{o}$ are solutions of the Lyapunov equations

$$\begin{aligned}\bar{\mathbf{A}}_0\mathbf{W}_\mathrm{c} + \mathbf{W}_\mathrm{c}\bar{\mathbf{A}}_0^\mathrm{T} &= -\bar{\mathbf{B}}\bar{\mathbf{B}}^\mathrm{T},\\ \bar{\mathbf{A}}_0^\mathrm{T}\mathbf{W}_\mathrm{o} + \mathbf{W}_\mathrm{o}\bar{\mathbf{A}}_0 &= -\bar{\mathbf{C}}^\mathrm{T}\bar{\mathbf{C}}\end{aligned} \tag{3.17}$$

[22]. Normally, solving a Lyapunov equation involves the computation of the Schur form of $\bar{\mathbf{A}}_0$ [22], but in the case at hand, the necessary work has been done before, since $\bar{\mathbf{A}}_0 = \mathrm{diag}(\gamma_{u+1}, \ldots, \gamma_N)$ is already diagonal, and (3.17) can be solved at virtually no extra cost:

$$\begin{aligned}(\mathbf{W}_\mathrm{c})_{ij} &= -(\gamma_{u+i} + \gamma_{u+j})^{-1}(\bar{\mathbf{B}}\bar{\mathbf{B}}^T)_{ij},\\ (\mathbf{W}_\mathrm{o})_{ij} &= -(\gamma_{u+i} + \gamma_{u+j})^{-1}(\bar{\mathbf{C}}^T\bar{\mathbf{C}})_{ij}.\end{aligned} \tag{3.18}$$

To proceed with the model reduction, all the system states which are essentially both uncontrollable and unobservable have to be excluded. Hence, $\mathbf{W}_\mathrm{c}$ and $\mathbf{W}_\mathrm{o}$ need to be combined somehow in order to account for controllability and observability simultaneously. In the case of balanced truncation and related methods, oblique transformations are defined which are associated with the product $\mathbf{W}_\mathrm{c}\mathbf{W}_\mathrm{o}$. Here, a different, somewhat simpler approach is taken in order to obtain an orthogonal transformation. Instead of analysing $\bar{\mathbf{X}}(t)$ and $\mathbf{Y}^\mathrm{T}(t)$ separately, both matrix time series are analysed at a time. To this end, the *joint controllability and observability Gramian*

$$\mathbf{W}_\mathrm{co} = \frac{1}{2}[\mathrm{tr}(\mathbf{W}_\mathrm{c})^{-1}\mathbf{W}_\mathrm{c} + \mathrm{tr}(\mathbf{W}_\mathrm{o})^{-1}\mathbf{W}_\mathrm{o}] \tag{3.19}$$

is defined here, corresponding to a PCA of $\frac{1}{2}[\mathrm{tr}(\mathbf{W}_\mathrm{c})^{-1}\bar{\mathbf{X}}(t) + \mathrm{tr}(\mathbf{W}_\mathrm{o})^{-1}\mathbf{Y}^\mathrm{T}(t)]$. In this expression, the matrix addition includes that the matrix which has less columns than the other is augmented with zeros to fit the size of the larger matrix. The trace prefactors before $\bar{\mathbf{X}}(t)$ and $\mathbf{Y}^\mathrm{T}(t)$ ensure that controllability and observability are given equal

weight in the sense that both sets of time series are normalised to have the same total signal energy of 0.5.

To perform the isometric truncation, the eigenvalue problem

$$\mathbf{W}_{\mathrm{co}}\boldsymbol{\theta}_i = \lambda_i \boldsymbol{\theta}_i, \quad i = 1, \ldots, N-u \tag{3.20}$$

is solved with the eigenvalues λ_i sorted in descending order. For all i, λ_i is equal to the variance fraction explained by the projection of all the time series onto $\boldsymbol{\theta}_i$, since $\sum_{i=1}^{N-u} \lambda_i = 1$. The orthogonal eigenvectors $\boldsymbol{\theta}_i$ are normalised to unity and define the transformation matrix $\bar{\mathbf{T}}$ acting in the stable subspace:

$$\bar{\mathbf{T}} = (\boldsymbol{\theta}_1 \; \boldsymbol{\theta}_2 \; \ldots \; \boldsymbol{\theta}_{N-u}). \tag{3.21}$$

In the full space represented by the open-loop eigenvector coordinates, the orthogonal transformation matrix is given by

$$\mathbf{T} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \bar{\mathbf{T}} \end{pmatrix} \tag{3.22}$$

where the upper left block contains the identity matrix $\mathbb{1} \in \mathbb{R}^{u \times u}$ acting in the unstable subspace. The open-loop matrix and the basic feedback matrices, which already have been pre-transformed into the open-loop eigenvector coordinate system (cf. the end of subsection 3.1), are subjected to a further, orthogonal transform as $\check{\mathbf{M}} = \mathbf{T}^{\mathrm{T}}\tilde{\mathbf{M}}\mathbf{T}$, where $\tilde{\mathbf{M}}$ has to be replaced by $\tilde{\mathbf{A}}_0$, $\tilde{\mathbf{F}}^{kl}_{n,\alpha}$, $\tilde{\mathbf{F}}^{kl}_{n,\beta}$, and $\tilde{\mathbf{F}}^{k}_{\mathrm{I}}$. Together with the free parameters α^{kl}_n, β^{kl}_n, and $\tilde{R}_k$ these transformed matrices define the transformed system matrix

$$\mathbf{A} = (\mathbb{1} + \tau\check{\mathbf{F}})^{-1}(\check{\mathbf{A}}_0 + \check{\mathbf{F}}), \tag{3.23}$$

cf. (2.24), where $\check{\mathbf{F}}$ is constructed from $\check{\mathbf{F}}^{kl}_{n,\alpha}$, $\check{\mathbf{F}}^{kl}_{n,\beta}$, and $\check{\mathbf{F}}^{k}_{\mathrm{I}}$ analogous to (2.16). Finally, in the truncation step the lowermost rows and rightmost columns in $\check{\mathbf{A}}_0$ and $\check{\mathbf{F}}$ are deleted, thus retaining only an upper left $N_{\mathrm{red}} \times N_{\mathrm{red}}$-block, respectively. The truncated version of $\mathbf{A}$ is constructed according to (3.23), but substituting the truncated versions of $\check{\mathbf{A}}_0$ and $\check{\mathbf{F}}$. If $\sum_{i=N_{\mathrm{red}}+1-u}^{N-u} \lambda_i \ll 1$, it can be expected that the truncated system mimics the full system quite accurately in responding to the feedback. Using the truncated model, an optimally stabilising parameter set is searched for by means of the methodology described in sections 4 and 5. Afterwards, the parameters found are used to construct the full-sized matrix $\mathbf{A}$. For quality control, the properties (eigenvalues, pseudospectra, $||\mathrm{e}^{t\mathbf{A}}||$ curves, cf. section 4) of the full model and the reduced one have to be compared. If there are unacceptable deviations, the optimisation step has to be repeated using a less severely truncated model, i. e., a higher value of N_{red}.

Since the system states correspond to current distributions on the conducting wall and in the coils, the state space directions, which the system is projected onto during the isometric truncation procedure, can be interpreted as current patterns. Represented by state vectors in the original, untransformed coordinate system, these current patterns are given by the columns of the "pattern matrix"

$$\mathbf{\Pi} = \mathbf{X}\mathbf{T}, \tag{3.24}$$

which comprises the preliminary transform to open-loop eigenvector coordinates and the subsequent orthogonal transform. The column vectors of $\mathbf{\Pi} = (\boldsymbol{\pi}_1\ \boldsymbol{\pi}_2\ \ldots\ \boldsymbol{\pi}_N)$ can be used for visualisation of the patterns and are orthonormal with respect to the "ohmic loss" scalar product: $\boldsymbol{\pi}_i^{\mathrm{T}}\mathbf{R}_0\boldsymbol{\pi}_j = \delta_{ij}$. The leading patterns can be interpreted as being the most important physical processes involved in the control system.

In the following, $\mathbf{A}$ will denote both the full-sized, transformed system matrix and its truncated counterpart. The distinction between the two will be explicitly noted wherever necessary.

4. Concepts for controller optimisation

The question remains, what is "optimal stability"? When attempting to stabilise a parametrised matrix, two basic objectives are obvious: On the one hand, it is desirable to obtain good asymptotic stability. That means, all eigenvalues lie in the left complex half-plane and are comfortably bounded away from the imaginary axis, so that initial perturbations decay reasonably fast as $t \to \infty$. On the other hand, the stability should be robust with respect to uncertainties in the parameters, or, more generally, to imperfections in the mathematical model. Assuming that such uncertainties are reflected by uncertainties in the matrix elements, robust stability means that, if the matrix $\mathbf{A}$ is stable, then $(\mathbf{A} + \mathbf{E})$ should be stable for any matrix perturbation $\mathbf{E}$ of "moderate strength". Unfortunately, achievement of the first objective does not necessarily imply achievement of the second. A large distance of the eigenvalues from the imaginary axis does not ensure robustness if the eigenvalues are very sensitive to perturbations of $\mathbf{A}$.

Eigenvalues of *normal* matrices, i. e., matrices with an orthogonal eigensystem, are most insensitive. If the matrix is *nonnormal*, which is the case for the RWM problem with feedback, the sensitivity can increase dramatically. Furthermore, there is a second unfavourable feature: if $\mathbf{A}$ is stable, but nonnormal, then $||\mathrm{e}^{t\mathbf{A}}||$ may be much greater than unity for some $t > 0$ (this holds for any matrix norm, but the matrix 2-norm is used throughout this study). This means that the initial state at $t = 0$ can be chosen in such a way that it exhibits strong transient growth before it eventually decays [17] (of course,

not all possible initial states suffer from transient growth, in general). Systems which are theoretically stable might be practically unstable if the transient amplification is too strong. This fact gives rise to a third objective in matrix stabilisation: the transient peak, i. e., the supremum of $||e^{t\mathbf{A}}||$ for $t > 0$, should be kept at a moderate level.

In the following, these three objectives are put into quantitative terms. In subsection 4.1, measures to assess the asymptotic stability of a matrix are introduced. Following [17], measures of robust stability are derived in subsection 4.2, which are based on the central concept of matrix pseudospectra. In subsection 4.3 it is explained how to tackle the problem of transient peaks, which turn out to be related to pseudospectra, as well. Finally, the numerical realization of these concepts in the OPTIM code is briefly explained in subsection 4.4.

4.1. Measures of asymptotic stability

The traditional measure of asymptotic stability is the *spectral abscissa* $\sigma(\mathbf{A})$, that is, the largest real part of all eigenvalues of $\mathbf{A}$:

$$\sigma(\mathbf{A}) = \max_{i=1,\ldots,N} \mathrm{Re}\, \gamma_i. \tag{4.1}$$

Minimising $\sigma(\mathbf{A})$ under variations of the parameters upon which $\mathbf{A}$ depends gives the best available asymptotic decay rate of initial disturbances. However, this function "sees" only the leading eigenvalue(s) (those with largest real parts), and in many cases several eigenvalues are simultaneously on the leading position when the optimal solution has been found, thus giving the "least stable subspace" a rather high dimension. In such cases, it may be possible to push some of these eigenvalues much further to the left in the complex plane under additional parameter variations without affecting other eigenvalues

very much. Therefore, a second function $\eta(\mathbf{A})$ which accounts for all the eigenvalues, but gives more weight to the leading ones, has been implemented into OPTIM in addition to $\sigma(\mathbf{A})$:

$$\eta(\mathbf{A}) = \sum_{i=1}^{N} \exp(\text{Re } \gamma_i). \tag{4.2}$$

Here, this function is given the name *exponential spectral function*. Minimising $\eta(\mathbf{A})$ establishes a compromise in attempting to shift all eigenvalues as far as possible to the left. Neither $\sigma(\mathbf{A})$ nor $\eta(\mathbf{A})$ guarantee a high level of robustness of the stability after optimisation, because eigenvalue sensitivity is not taken into account. However, under the assumption that the sensitivity of the optimal solution is similar in both cases, optimisation of $\eta(\mathbf{A})$ appears to be likely to produce somewhat more robust results, because generally most of the eigenvalues should have a greater distance to the imaginary axis than they would have after $\sigma(\mathbf{A})$ optimisation. But that assumption might be not true.

4.2. Eigenvalue sensitivity and robust stability

To assess the robustness of stability it is necessary to consider the sensitivity of eigenvalues. An elegant way to capture the sensitivity of the entire spectrum is given by the definition of the ϵ-*pseudospectrum* $\gamma_\epsilon(\mathbf{A})$:

$$\begin{aligned}\gamma_\epsilon(\mathbf{A}) = \{z \in \mathbb{C} \ : \ & z \text{ is an eigenvalue of } \mathbf{A} + \mathbf{E} \\ & \text{for some } \mathbf{E} \in \mathbb{C}^{N \times N} \text{ with } ||\mathbf{E}|| < \epsilon\}.\end{aligned} \tag{4.3}$$

While the eigenvalue spectrum of $\mathbf{A}$ is a discrete set of points in $\mathbb{C}$, $\gamma_\epsilon(\mathbf{A})$ is an open set in $\mathbb{C}$ containing the spectrum. Since (4.3) involves a matrix norm (in this study, the 2-norm), the ϵ-pseudospectra depend on the state space coordinate system. This

is exactly the reason why the similarity transform underlying the model reduction was demanded to be orthogonal.

One might immediately ask why the perturbations $\mathbf{E}$ are allowed to be complex. In physical systems, perturbations are generally real, not complex. Indeed, the *real structured pseudospectrum*, which is obtained when replacing $\mathbf{E} \in \mathbb{C}^{N\times N}$ by $\mathbf{E} \in \mathbb{R}^{N\times N}$ in (4.3), attracts more and more interest in the control theory community. But considering complex pseudospectra instead of real ones has certain advantages. The former are easier to deal with than the latter, both theoretically and practically. More important, however, is the fact that complex pseudospectra provide insight in the transient behaviour of the system, as it will be discussed in subsection 4.3, while real pseudospectra do not. Finally, a system proven to be robust under complex perturbations will certainly be robust under real ones.

Based upon the concept of complex pseudospectra, the robustness of a stable system $\mathbf{A}$ can be measured by the *complex stability radius* $\rho(\mathbf{A})$:

$$\rho(\mathbf{A}) = \sup\{\epsilon : \mathbf{A} + \mathbf{E} \text{ is stable } \forall\ \mathbf{E} \in \mathbb{C}^{N\times N} \text{ with } ||\mathbf{E}|| < \epsilon\}. \tag{4.4}$$

In other words, $\rho(\mathbf{A})$ is equal to that particular ϵ for which the closure of $\gamma_\epsilon(\mathbf{A})$ touches the imaginary axis.

Another measure of robustness is the spectral abscissa of the system $\mathbf{A}$ if subjected to a worst case perturbation $\mathbf{E}$ smaller than a given ϵ. This measure is denoted the *ϵ-pseudospectral abscissa* $\sigma_\epsilon(\mathbf{A})$ and corresponds to the largest real part of all points on the closure of $\gamma_\epsilon(\mathbf{A})$:

$$\sigma_\epsilon(\mathbf{A}) = \sup\{\text{Re } z : z \in \gamma_\epsilon(\mathbf{A})\}. \tag{4.5}$$

Both functions, $\rho(\mathbf{A})$ and $\sigma_\epsilon(\mathbf{A})$, are implemented in OPTIM and can be chosen as objective for optimisation. The consequences of the choice of ϵ when optimising $\sigma_\epsilon(\mathbf{A})$ will be briefly discussed in the following subsection.

4.3. Transient growth

There are various theorems relating the transient behaviour to pseudospectra. An important one is the following lower bound for the transient peak [17]:

$$\sup_{t\geq 0} ||\mathrm{e}^{t\mathbf{A}}|| \geq \frac{\sigma_\epsilon(\mathbf{A})}{\epsilon} \quad \forall \epsilon > 0. \tag{4.6}$$

That means, if the ϵ-pseudospectrum extends significantly into the right complex half-plane for some ϵ so that $\sigma_\epsilon(\mathbf{A}) > \epsilon$, then there is transient growth. Furthermore, it is shown in [17] that, for large values of $\sigma_\epsilon(\mathbf{A})/\epsilon$, $||\mathrm{e}^{t\mathbf{A}}||$ grows exponentially at least with a rate close to $\sigma_\epsilon(\mathbf{A})$ for small t. The larger the value of $\sigma_\epsilon(\mathbf{A})/\epsilon$ is, the longer this behaviour continues.

It follows from these results that it is necessary, albeit not sufficient, to keep $\sigma_\epsilon(\mathbf{A})/\epsilon$ at a minimum for all ϵ in order to diminish transient growth. This can be attempted by minimising $\sigma_\epsilon(\mathbf{A})$ for ϵ significantly larger than $\rho(\mathbf{A})$, at the expense of robustness. The choice of ϵ requires some experimentation. If ϵ is chosen too large, the optimal solution may even be unstable. Generally, optimising $\sigma_\epsilon(\mathbf{A})$ for $\epsilon > \rho(\mathbf{A})$ corresponds to a tradeoff between robustness and favourable transient behaviour, whereas the optimisation for $0 < \epsilon < \rho(\mathbf{A})$ is a compromise between robustness and asymptotic stability.

4.4. Functionality of OPTIM

The OPTIM code package offers four different features to be applied to parametrised matrices $\mathbf{A}$: i) Optimisation of stability under parameter variations by means of the algorithm described in section 5, ii) computation of pseudospectra for a given parameter set, iii) computation of $||\mathrm{e}^{t\mathbf{A}}||$ on a given time interval $0 \leq t \leq T$ and for a given parameter set, and iv) simple evaluation of a chosen objective function for a given parameter set. In principle, all these functionalities can be applied to the reduced as well as to the full-sized model. The code is fully parallelised based upon the PBLAS and ScaLAPACK libraries for use on a distributed-memory architecture. It has to be noted, however, that the order of the reduced model used in this study is so small that parallelisation is not efficient, so that the optimisations are carried out on a single processor only. But the computations made for the full system (pseudospectra, $||\mathrm{e}^{t\mathbf{A}}||$ calculations, evaluations of objective functions, see section 6) take full advantage of the parallel execution on 32 processors. Due to its modular structure, the code is not restricted to the RWM stabilisation problem presented here, but can readily be applied to any user-implemented parametrised matrix and therefore offers a wide area of application in control theory and related fields.

Let $\mathbf{p} = (p_1\ p_2\ \ldots\ p_P)$ be the vector of parameters which $\mathbf{A}$ depends on, $\mathbf{A} = \mathbf{A}(\mathbf{p})$. When optimising the stability of $\mathbf{A}$, one of the following four objective functions can be chosen for minimisation, cf. (4.1), (4.2), (4.4), and (4.5):

$$\mathcal{F}_1(\mathbf{p}) = \sigma[\mathbf{A}(\mathbf{p})],$$

$$\mathcal{F}_2(\mathbf{p}) = \eta[\mathbf{A}(\mathbf{p})],$$

$$\mathcal{F}_3(\mathbf{p}) = \begin{cases} \sigma[\mathbf{A}(\mathbf{p})] & , \quad \mathbf{A}(\mathbf{p}) \text{ is unstable} \\ -\rho[\mathbf{A}(\mathbf{p})] & , \quad \mathbf{A}(\mathbf{p}) \text{ is stable,} \end{cases}$$

$$\mathcal{F}_4(\mathbf{p}) = \sigma_\epsilon[\mathbf{A}(\mathbf{p})]. \tag{4.7}$$

By the case differentiation, $\mathcal{F}_3$ becomes a function which is continuous everywhere in the parameter space. To evaluate $\mathcal{F}_1$ and $\mathcal{F}_2$, the eigenvalue problem of $\mathbf{A}$ has to be solved. The algorithm to calculate $\mathcal{F}_3$ for stable $\mathbf{A}$ is closely adapted from the two-step $\mathcal{H}_\infty$-norm computation method described in [32]. Namely, it can be shown, by means of the alternative definition (4.8) of $\gamma_\epsilon(\mathbf{A})$ introduced below, that $\rho(\mathbf{A})$ is equal to the reciprocal of the $\mathcal{H}_\infty$-norm (1.4) of the "transfer function" $\mathbf{G}(z) = (\mathbf{A} - z\mathbb{1})^{-1}$. $\mathcal{F}_4$ is computed in terms of a criss-cross algorithm [33], which is quite similar to the two-step method.

In the computation of pseudospectra, an alternative definition of $\gamma_\epsilon(\mathbf{A})$ is used, which can be shown to be equivalent to the definition (4.3) as long as the matrix 2-norm is used there [17]:

$$\gamma_\epsilon(\mathbf{A}) = \{z \in \mathbb{C} \ : \ s_{\min}(\mathbf{A} - z\mathbb{1}) < \epsilon\} \tag{4.8}$$

with $s_{\min}(\mathbf{A} - z\mathbb{1})$ being the smallest singular value of $(\mathbf{A} - z\mathbb{1})$. By calculating $s_{\min}(\mathbf{A} - z\mathbb{1})$ for various z values on a grid in $\mathbb{C}$ and visualising them in a contour plot, one obtains different ϵ-pseudospectra boundaries given by the corresponding ϵ-contours. An efficient method for this computation is used, which has been adapted from [17, 34], but with the inverse iteration step replaced by an explicitly restarted Lanczos procedure as described in [35] to improve convergence in the case that the smallest singular values are densely spaced.

In the computation of $||\mathrm{e}^{t\mathbf{A}}||$, the matrix exponential is evaluated by means of the scaling and squaring method as described in [36], which involves a Padé approximation of the exponential.

5. Optimisation method

In general, all the objective functions in (4.7) have the unfavourable property of being not smooth everywhere. Even though these functions are continuous, their gradients can be discontinuous or may even diverge on sub-manifolds in parameter space. Conventional nonlinear optimisation methods are likely to get stuck near such manifolds, and a special algorithm is required to successfully minimise functions like (4.7). The *gradient bundle method* [37, 38] implemented in OPTIM fits this purpose very well and will be briefly described below. Like other optimisation methods, it requires the gradient of the respective objective function, and it is therefore highly beneficial to derive analytic expressions for the gradients of (4.7) which are valid everywhere except on sub-manifolds such as mentioned above.

The functions $\mathcal{F}_1(\mathbf{p})$ and $\mathcal{F}_2(\mathbf{p})$ are related to the eigenvalues of $\mathbf{A}(\mathbf{p})$. It can be inferred from (4.1) and (4.2) that the components of their gradients are given by

$$\frac{\partial \mathcal{F}_1}{\partial p_j} = \mathrm{Re}\, \frac{\partial \gamma_m}{\partial p_j}, \tag{5.1}$$

where γ_m is an eigenvalue of $\mathbf{A}$ with maximum real part, and

$$\frac{\partial \mathcal{F}_2}{\partial p_j} = \sum_{i=1}^{N} \mathrm{Re}\, \frac{\partial \gamma_i}{\partial p_j} \exp(\mathrm{Re}\, \gamma_i), \tag{5.2}$$

$j = 1, \ldots, P$. For the derivatives of the eigenvalues with respect to the parameters, first-order perturbation theory yields

$$\frac{\partial \gamma_i}{\partial p_j} = \frac{\mathbf{z}_i^{\mathrm{H}} (\partial \mathbf{A} / \partial p_j) \mathbf{x}_i}{\mathbf{z}_i^{\mathrm{H}} \mathbf{x}_i}, \tag{5.3}$$

where $\mathbf{x}_i$ is the right eigenvector and $\mathbf{z}_i$ the left eigenvector of $\mathbf{A}$ associated with the eigenvalue γ_i which has to be nondegenerate [39]. Finally, the matrix derivative

$(\partial \mathbf{A}/\partial p_j)$ needs to be determined. The parameter dependence of the transformed feedback matrix $\check{\mathbf{F}}$, which contributes to $\mathbf{A}$ according to (3.23), can be compactly written as $\check{\mathbf{F}} = \sum_{i=1}^{P} p_i \check{\mathbf{F}}_i$, where the $\check{\mathbf{F}}_i$'s are the transformed basic feedback matrices. A short calculation gives

$$\frac{\partial \mathbf{A}}{\partial p_j} = (\mathbb{1} + \tau \check{\mathbf{F}})^{-1} \check{\mathbf{F}}_j (\mathbb{1} - \tau \mathbf{A}). \tag{5.4}$$

The components of the gradient of $\mathcal{F}_3$ are given by

$$\frac{\partial \mathcal{F}_3}{\partial p_j} = \begin{cases} \partial \mathcal{F}_1 / \partial p_j & , \quad \mathbf{A}(\mathbf{p}) \text{ is unstable} \\ -(\partial \mathbf{A}/\partial p_j) \cdot \mathrm{Re}(\boldsymbol{\upsilon}\boldsymbol{\omega}^{\mathrm{H}}) & , \quad \mathbf{A}(\mathbf{p}) \text{ is stable,} \end{cases} \tag{5.5}$$

where $\boldsymbol{\upsilon}$ and $\boldsymbol{\omega}$ are the left and right singular vectors belonging to the smallest singular value of $(\mathbf{A} - z\mathbb{1})$ and $z \in \mathbb{C}$ is a point where the pseudospectrum $\gamma_{\rho(\mathbf{A})}$ touches the imaginary axis [16, 40]; z is computed automatically as a by-product of the algorithm which evaluates $\rho(\mathbf{A})$. The product $(\partial \mathbf{A}/\partial p_j) \cdot \mathrm{Re}(\boldsymbol{\upsilon}\boldsymbol{\omega}^{\mathrm{H}})$ is a "matrix scalar product" defined as an ordinary scalar product while interpreting the matrices as vectors with N^2 components. Similarly, the gradient components of $\mathcal{F}_4$ are computed as

$$\frac{\partial \mathcal{F}_4}{\partial p_j} = (\partial \mathbf{A}/\partial p_j) \cdot \mathrm{Re}\left(\frac{1}{\boldsymbol{\upsilon}^{\mathrm{H}}\boldsymbol{\omega}} \boldsymbol{\upsilon}\boldsymbol{\omega}^{\mathrm{H}} \right). \tag{5.6}$$

Again, $\boldsymbol{\upsilon}$ and $\boldsymbol{\omega}$ are the left and right singular vectors corresponding to the smallest singular value of $(\mathbf{A} - z\mathbb{1})$, but z is now a point on the closure of $\gamma_\epsilon(\mathbf{A})$ with maximum real part [16, 40], automatically determined during $\sigma_\epsilon(\mathbf{A})$ computation.

In principle, the gradient bundle method is similar to the simple steepest descent algorithm. In both methods, each iteration consists of determining a search direction in parameter space, followed by a line search along this direction in order to find a significantly lower objective function value. In the case of steepest descent, the search

direction is simply the negative gradient evaluated at the current point in parameter space. The gradient bundle method additionally accounts for gradients evaluated at some other points in the neighbourhood. The search direction is defined to be the negative of the shortest vector on the convex hull of all gradient vectors. If the objective function is smooth in the neighbourhood of the current parameter point, all gradient vectors will be much the same, and the search direction will resemble the steepest descent. But if the current parameter point is close to a manifold where the gradient is discontinuous, gradients will be sampled from both sides of that manifold. The resulting search direction vector will then be aligned mostly parallel to the manifold, but still directed towards decreasing function values. This prevents the procedure from being forced to make tiny steps or getting stuck at the manifold. A well developed theory guarantees convergence against a local minimum of the objective function [37, 38]. In OPTIM, the required convex hull computations are arranged by calling the external program LRS [41]. The shortest vector on the respective convex hull is determined by a suitable quadratic programming subroutine [42].

The algorithm version implemented in OPTIM differs from the original one [37, 38] in two small, but noteworthy details. First of all, in the original procedure the additional gradients are evaluated at $2P$ points randomly chosen within some sampling diameter around the current point in parameter space. In OPTIM, the gradients are computed at the vertices of a simplex with given edge length, positioned in such a way that the current parameter point coincides with the barycentre of the simplex. This involves only $P+1$ instead of $2P$ additional gradient computations and therefore gives a considerable speedup. The convex hull problem is simplified by this means, as well. Despite the smaller number of points, the probability that information from both sides of a critical manifold is accounted for is still very high. To further stabilise the procedure, the

simplex is always rotated in a different, but deterministic manner in each iteration. Secondly, the selection of the "neighbourhood size" (simplex size or sampling diameter) and the termination criteria in OPTIM are somewhat different from the original. Like in the original, the simplex size is decreased by a prescribed factor if the convex hull happens to be very close to the origin or contains it, or if the line search step becomes too small. But what is new is that the simplex is also slightly enlarged from iteration to iteration unless a prescribed maximum size is already reached. Thus, the simplex size is dynamically adapted to the local scales on which the objective function varies significantly. The program terminates when the simplex size is already at a prescribed minimum, and a further shrinking is requested.

Finally, the implementation of constraints, such as those discussed in subsection 2.3, should be mentioned. In OPTIM, constraints are realized in terms of user-implemented penalty functions to be added to the objective function, and their respective gradients. These penalty functions should be identically zero in the allowed parameter domain and become increasingly positive the more the constraints are violated. The gradients must exist almost everywhere, but may be discontinuous on sub-manifolds.

6. ITER-like example

In the following, the derived methodology is applied to an ITER-related example. Because the principal aim of this paper is to demonstrate the capabilities of the method and of the code package and to gain general, conceptual insights rather than to produce precise results directly relevant to the ongoing ITER design, some simplifications are made in order to obtain a test case allowing for good visualisation and simple analysis, and requiring only moderate numerical effort.

The chosen plasma equilibrium is the steady state scenario 4 from the ITER design, with β set to 2.29%. Without a conducting wall, this equilibrium is unstable to external kinks with $n = 1$ only. Toroidal Fourier indices higher than $n = 1$ will not be included in the expansion of the plasma state, and the feedback control will also act on $n = 1$ only. In the strict sense, the 3D conducting structures break the axisymmetry and the different n's become dynamically coupled, so that it might be possible that the feedback system drives intrinsically stable RWMs with $n > 1$ unstable. But this possibility is not investigated in this study as the entire analysis is restricted to $n = 1$.

While a nested double wall is scheduled for ITER, the conducting wall model used here accounts only for the interior wall and also neglects some other geometric details. The effects of the exterior wall on the RWMs are minor, and a single wall greatly facilitates the visualisation of current patterns. Figure 1 displays the poloidal cross section of the configuration including the plasma, wall, coils, and sensors. The 3D triangular mesh of the conducting wall and the feedback coils are shown in figure 2. The holes in the wall correspond to the various ports which penetrate the ITER wall. The original wall will be made of steel with conductivity $\sigma = 1.212 \cdot 10^6\ \Omega^{-1}\mathrm{m}^{-1}$ and thickness $d = 0.06$ m. Since the thin-wall approximation is used in the STARWALL code, it is the surface conductivity σd which is accounted for. The coils are assumed to be made of copper ($\sigma = 6 \cdot 10^7\ \Omega^{-1}\mathrm{m}^{-1}$) with quadratic cross section having an edge length of 0.1 m. In the computations, the coil resistances are set accordingly. The coils are located inside some of the equatorial ports, equidistantly spaced by 40 degrees in the toroidal direction, and grouped into a single toroidal array, i. e., $K = 1$. Two neighbouring coils are missing in the array due to collision with neutral beam injection devices, so that only seven coils are remaining. This feedback coil set is one of several sets which have been under discussion in the ITER design process. The recent final decision gave preference to a different coil

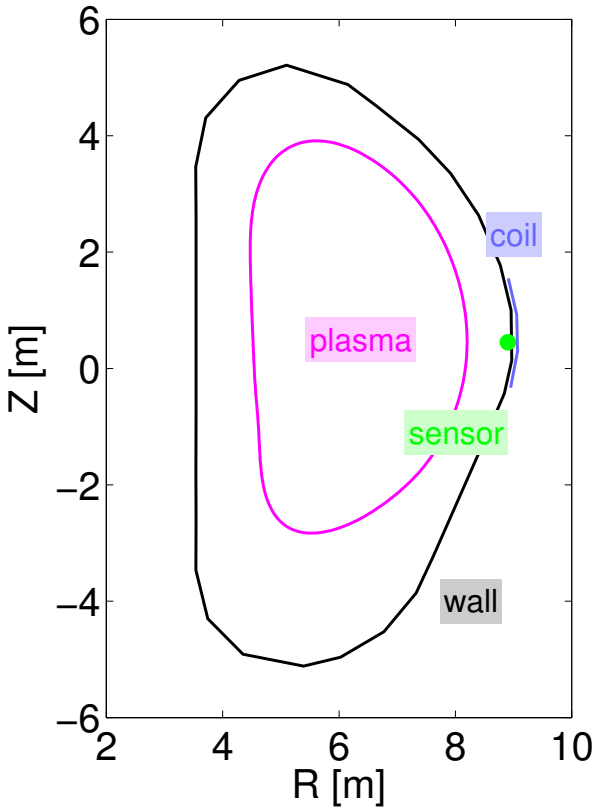

Figure 1. Poloidal cross section showing the geometry or position of the plasma boundary, the conducting wall, the feedback coils, and the sensors.

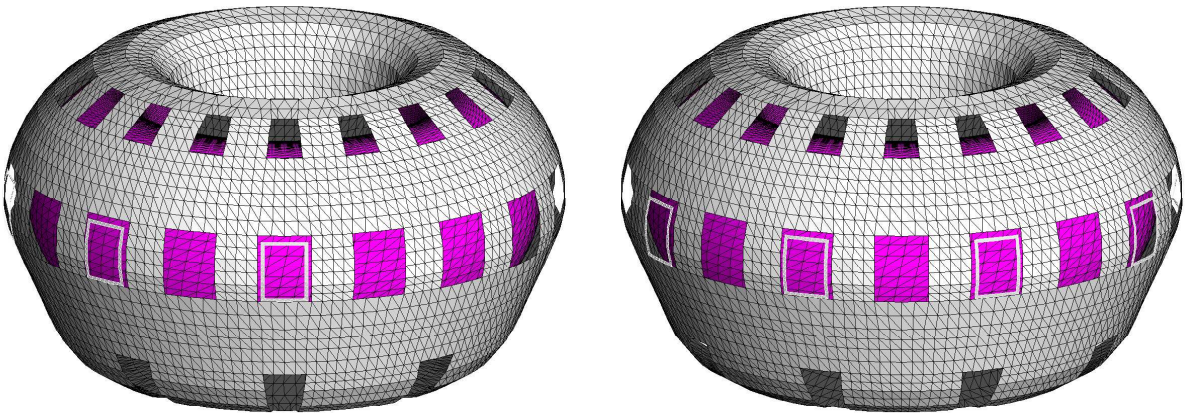

Figure 2. Triangular mesh of the conducting wall, and the feedback coil set. The plasma is also shown. The image on the right hand side differs from the one on the left by toroidal rotation of the device by 180 degrees.

set which will be mounted inside the wall, but the set used here is most suitable for the visualisation of coil currents. The 18 sensors also form a single array ($L = 1$) and are positioned at equidistant toroidal angles, centred between the equatorial ports and mounted inside, but very close to the wall (see figure 1). The sensors are oriented in the vertical direction and therefore measure the poloidal component of the magnetic field perturbation. Poloidal sensors were shown to be superior to radial sensors in several studies [4, 7, 43, 15]. The indices k, l, n are omitted in the following because, in this simple test case, they take the value 1 only.

The solution of the open-loop eigenvalue problem $\mathbf{A}_0\mathbf{x}_i = \gamma_i\mathbf{x}_i$, $i = 1, \ldots, N$, where $N = 5190$, results in two unstable RWMs $\mathbf{x}_1$, $\mathbf{x}_2$ with growth rates $\gamma_1 = 21.9\ \mathrm{s}^{-1}$ and $\gamma_2 = 21.7\ \mathrm{s}^{-1}$, whereas all other modes $\mathbf{x}_3, \ldots, \mathbf{x}_N$ are stable. The wall current patterns corresponding to $\mathbf{x}_1$ and $\mathbf{x}_2$ are shown in figure 3. The two patterns have a helical $n = 1$ structure and are very similar, but toroidally phase shifted against each other by 90 degrees. The eigenvalues γ_1 and γ_2 are not perfectly degenerate because the axisymmetry is broken due to the presence of the holes and, what is more important, due to the two missing coils (the coils act as short-circuited passive conductors in the open-loop computation).

The model reduction procedure yields a collection of current patterns $\boldsymbol{\pi}_3, \ldots, \boldsymbol{\pi}_N$ spanning the stable subspace of $\mathbf{A}_0$, which constitute all except the first two columns of the pattern matrix $\boldsymbol{\Pi}$ as defined by (3.24). Note that the first two columns of $\boldsymbol{\Pi}$ correspond to the unstable RWMs, i. e., $\boldsymbol{\pi}_1 = \mathbf{x}_1$ and $\boldsymbol{\pi}_2 = \mathbf{x}_2$. Each pattern $\boldsymbol{\pi}_i$, $i = 3, \ldots, N$ is associated with a variance fraction λ_{i-2}, or "importance measure", as computed during the solution of the eigenvalue problem (3.20). The twelve leading stable patterns $\boldsymbol{\pi}_3, \ldots, \boldsymbol{\pi}_{14}$ and their corresponding variance fractions $\lambda_1, \ldots, \lambda_{12}$ are displayed in figures 4 and 5. As explained in the following, several different physical processes can

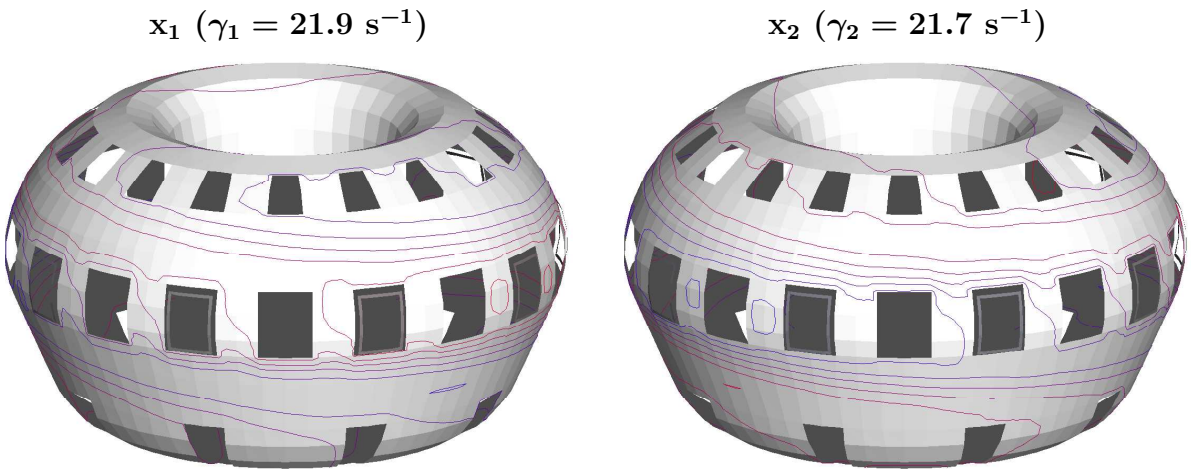

Figure 3. Current patterns corresponding to the two unstable RWMs $\mathbf{x}_1$ and $\mathbf{x}_2$, represented by lines of equal current potential ϕ, cf. (2.1). The wall currents flow parallel to the lines, and the current density is proportional to the line density. Blue lines circumvent areas of low ϕ, and red lines enclose areas of high ϕ, where the rotational direction of the currents is opposite to the areas of low ϕ. In both cases, toroidal rotation by 180 degrees (not shown) results in virtually the same pattern structure, but with reversed sign, reflecting the $n = 1$ structure of the modes. The angle of view is the same as in the right panel of figure 2.

be readily identified in these patterns, and it appears plausible that these processes are the most important ones involved in the RWM feedback control problem. Concerning their variance fractions, the first four patterns $\boldsymbol{\pi}_3, \ldots, \boldsymbol{\pi}_6$ are well separated from the others. Altogether, they comprise four processes, namely, (i) an $n = 1$ pattern of coil currents, (ii) the corresponding counterpart toroidally rotated by roughly 90 degrees, (iii) a pattern of small scale, dipole-like eddy currents localised in the neighbourhood of the sensors and modulated sinusoidally with "wave number" $n = 1$ in the toroidal direction, and (iv) the corresponding pattern toroidally rotated by 90 degrees. While $\boldsymbol{\pi}_3$ is an almost pure coil current pattern, $\boldsymbol{\pi}_4$, $\boldsymbol{\pi}_5$ and $\boldsymbol{\pi}_6$ are linear combinations of the remaining coil and eddy current processes. The importance of the coil current patterns is clear since these are just the actuator patterns generated by the gain matrices (2.11) and (2.12). But the small eddy currents are also very important because they produce a strong $n = 1$ signal detected by the sensors. It can be expected that such current patterns provoke an unwanted, fierce response of the feedback system, and it will be

shown later in this section that this is indeed the case. The patterns $\boldsymbol{\pi}_7$ and $\boldsymbol{\pi}_8$ represent the small-scale part of the image currents associated with the excitation of the coil currents. Another small-scale eddy structure near the sensors can be found in $\boldsymbol{\pi}_9$ and $\boldsymbol{\pi}_{10}$, whereas $\boldsymbol{\pi}_{11}$ and $\boldsymbol{\pi}_{12}$ mainly represent the large-scale part of the image currents excited by the coils. Finally, $\boldsymbol{\pi}_{13}$ and $\boldsymbol{\pi}_{14}$ are almost purely large-scale $n = 1$ patterns, probably strongly interacting with the plasma.

As already implied by the values given in figures 4 and 5, the variance fractions λ_i fall off rapidly as i increases. This indicates the model reduction being very efficient. The truncation limit N_{red} (reduced model dimension) is set to the value for which $\lambda_{N_{\mathrm{red}}-2} > 10^{-12} > \lambda_{N_{\mathrm{red}}-1}$ holds. This results in $N_{\mathrm{red}} = 56$, allowing for both very fast optimisation runs and excellent reproduction of the control system properties, as shown below.

In the following feedback computations, the time delay (2.23) is set to $\tau = 0.1$ ms. The feedback optimisation is conducted subject to the constraint that the effective current in each coil must not exceed the prescribed value $I_{\max} = 74$ kA to be substituted into the inequality (2.19). By this choice and by the estimate $S_{\max} = 1.5$ mT [9] as an RWM detection limit, the saturation current gain magnitude

$$\mathcal{I} = 9\frac{\sqrt{\frac{1}{2}(\alpha^2 + \beta^2)}}{R + \tilde{R}}, \tag{6.1}$$

corresponding to the left hand side of (2.19), is limited to twice the minimum gain needed to obtain stabilisation. The chosen value of $I_{\max}$ is not inconsistent with the fact that the coils in the set were designated to carry a total current of about 300 kA, where, however, a large portion of this current was intended to be used for ELM suppression. As an additional constraint, the time constant $T = L/(R + \tilde{R})$ of the coils is demanded

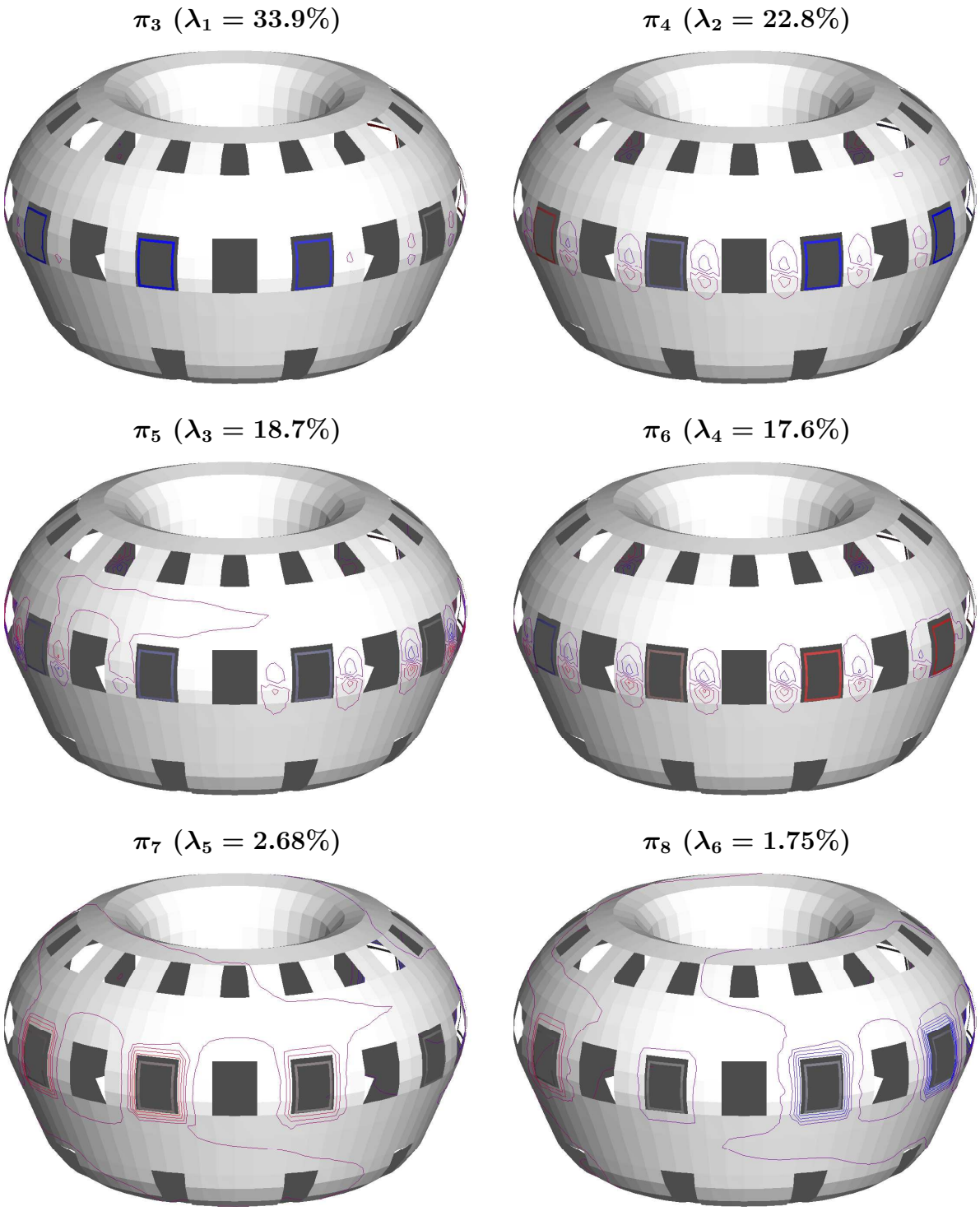

Figure 4. Stable current patterns $\boldsymbol{\pi}_3, \ldots, \boldsymbol{\pi}_8$ represented by current potential isolines in the same way as the unstable RWMs in figure 3. These patterns explain variance fractions $\lambda_1, \ldots, \lambda_6$ as indicated. Coil currents are visualised by colouring of the coils. The intensity of the colour corresponds to the strength of the current, and the choice of blue or red colour indicates the direction of the coil current in the same way as the rotational direction of wall currents is represented by the colouring of the equipotential lines. Like in the case of RWMs, toroidal rotation of these patterns by 180 degrees results mainly in a sign reversal (not shown).

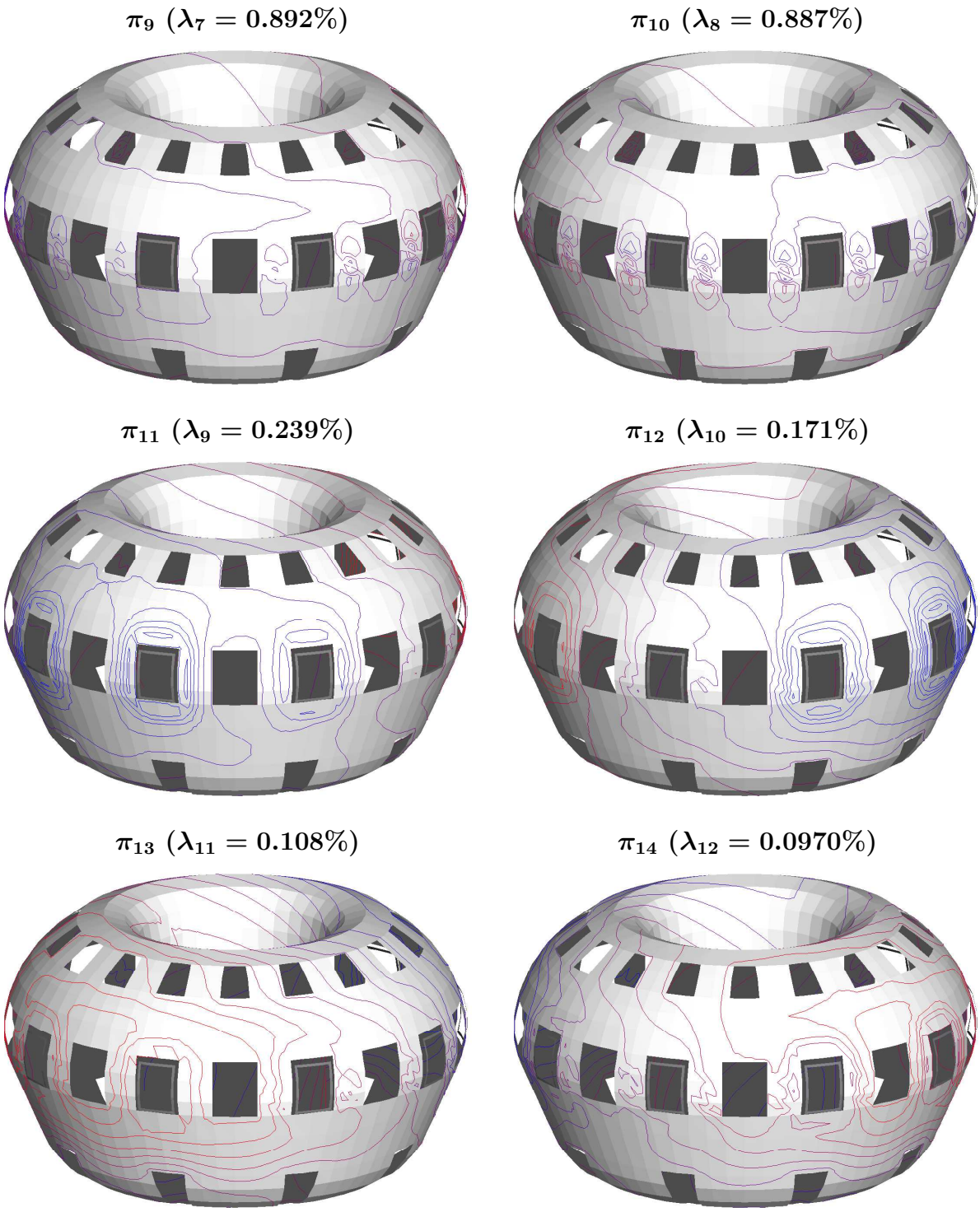

Figure 5. As figure 4, but showing the stable current patterns $\boldsymbol{\pi}_9, \ldots, \boldsymbol{\pi}_{14}$ and their variance fractions $\lambda_7, \ldots, \lambda_{12}$.

to be not smaller than $T_{\min} = 1$ ms, cf. (2.17). This choice of a minimum value ensures that the approximation used in deriving the effect of the time delay τ (section 2.4) is justified. Furthermore, the combination of the two constraints renders the set of feasible parameter values bounded and ensures convergence of the optimisation algorithm. A voltage constraint like (2.18) is not used here.

Each of the four objective functions (4.7) is minimised in subsequent OPTIM runs. Optimisation of $\mathcal{F}_4 = \sigma_\epsilon(\mathbf{A})$ is performed setting $\epsilon = 2\rho_{\text{opt}}$ after having obtained an optimal complex stability radius ρ_{opt} from the minimisation of $\mathcal{F}_3$.

Since the minima found in these nonlinear optimisation problems generally cannot be guaranteed to be global, the computations are repeated for each objective using several different starting parameter sets. Optimisations are started with the initial value of $\tilde{R}$ chosen such that $T = 100$ ms, 30 ms, 10 ms, 3 ms, and 1 ms, and the voltage gains α, β set to zero. The result with the lowest objective value is considered the best available one and is retained for further analysis, respectively, and all other runs are discarded. Each of the four optimal solutions is further examined in terms of the values of all four objective functions, and concerning its eigenvalues, pseudospectra, and $||\mathrm{e}^{t\mathbf{A}}||$ curves for t between 0 and 0.1 s. All these quantities are computed for both the full and the reduced model in order to show sufficient agreement, so that the chosen reduced model size is ensured to be large enough. Based on all these analyses, the four objective functions will be judged with respect to their capabilities to produce controllers achieving good asymptotic stability, or robust stability, or moderate transient behaviour.

The sets of parameter values α, β, $\tilde{R}$ resulting from the optimisations are listed in table 1, together with derived quantities which are somewhat easier to interpret, namely, the coil time constant T, the saturation current gain $\mathcal{I}$ as defined by (6.1), and the toroidal phase

optimal	α [V/mT]	β [V/mT]	$\tilde{R}$ [mΩ]	T [ms]	$\mathcal{I}$ [kA/mT]	φ [°]
$\mathcal{F}_1$	0.372	0.638	0.117	36.2	36.6	59.6
$\mathcal{F}_2$	19.7	-18.9	4.64	1.00	37.3	-43.8
$\mathcal{F}_3$	34.1	-11.5	4.64	1.00	49.3	-18.7
$\mathcal{F}_4$	29.4	20.8	4.64	1.00	49.3	35.3

Table 1. Optimal parameter values α, β, $\tilde{R}$ obtained after optimisation of the objective functions $\mathcal{F}_1$ through $\mathcal{F}_4$. Derived from these values, the following quantities are also listed: the coil time constants $T = L/(R + \tilde{R})$ corresponding to $\tilde{R}$, where $L = 4.65$ μH, $R = 10.9$ $\mu\Omega$, the saturation current gain magnitude $\mathcal{I}$ as defined by (6.1), and the toroidal response phase $\varphi = \arctan(\beta/\alpha)$. Note the constraints $T \geq 1$ ms and $\mathcal{I} \leq 49.3$ kA/mT.

angle φ between the measured $n = 1$ perturbation and the coils' $n = 1$ response pattern. The four different solutions differ strongly from each other concerning the response phase, and also considerably concerning the current gain magnitude. Furthermore, the $\mathcal{F}_1$-optimal solution is characterised by a slow response (large T), whereas the response is fast in the other three cases.

The values of all functions $\mathcal{F}_1$ through $\mathcal{F}_4$ evaluated using the full model for all four parameter sets can be found in table 2. For the reduced model, the values coincide with those computed with the full model up to a relative error of about 10^{-3} in the worst case (the $\mathcal{F}_1$ value after optimisation of $\mathcal{F}_1$), but to 10^{-4} or less in all the other cases (not shown). The $\mathcal{F}_2$ values are the only exception. They are considerably higher in case of the full model because much more eigenvalues are contributing to the sum (4.2) which are, however, largely invariant under parameter variations. Of course, the $\mathcal{F}_1$-optimal solution exhibits the best asymptotic stability, but the latter is not much worse in the other cases. Robustness, as expressed by the $\mathcal{F}_3$ value, varies strongly among the different cases. Expectedly, the $\mathcal{F}_3$- and $\mathcal{F}_4$-optimal solutions are considerably more robust than the other two. Robustness is by far the weakest after $\mathcal{F}_2$ optimisation. Not surprisingly, optimising $\mathcal{F}_4$ leads to a slightly less robust result than optimising $\mathcal{F}_3$, but

optimal	$\mathcal{F}_1$ [s^{-1}]	$\mathcal{F}_2$	$\mathcal{F}_3$ [s^{-1}]	$\mathcal{F}_4$ [s^{-1}]
$\mathcal{F}_1$	-7.41	22.5	-0.97	15.9
$\mathcal{F}_2$	-6.26	21.2	-0.29	357.7
$\mathcal{F}_3$	-6.26	21.4	-2.70	166.5
$\mathcal{F}_4$	-6.28	21.8	-2.37	12.0

Table 2. Values of the functions $\mathcal{F}_1$ through $\mathcal{F}_4$ as indicated by the column label, evaluated for the respective parameter set optimising the function corresponding to the row label, substituted into the full model. When evaluating $\mathcal{F}_2$, the eigenvalues have been scaled to internal, dimensionless units.

this is compensated to some extent by improved transient behaviour, as it will be shown below.

Figures 6, 7, 8, and 9 show spectra and pseudospectra for all four parameter sets substituted into the reduced as well as the full model. These plots reveal the sensitivity of individual eigenvalues as well as the sensitivity of the system's stability as a whole. For example, the $\mathcal{F}_3$ values given in table 1 can also be crudely estimated by inspection of figure 6. Furthermore, as already discussed in subsection 4.3 and also examined below, the pseudospectra parts in the right complex half-plane are related to the transient behaviour.

Comparing for each parameter set, respectively, the pseudospectra for the two models, one finds virtually perfect agreement in the right half-plane and in the neighbourhood of eigenvalues with large imaginary parts (corresponding to rotating, damped modes), which turn out to be the most sensitive ones. Noticeable deviations are found only in regions close to the negative real axis, far away from the imaginary axis, where lots of eigenvalues of the full model are missing in the reduced model. But these deviations are unimportant, since these eigenvalues have small imaginary parts, indicating that the corresponding eigenmodes are likely to be largely unaffected by the feedback system. Furthermore, they decay rapidly. The exclusion of these eigenmodes from the reduced

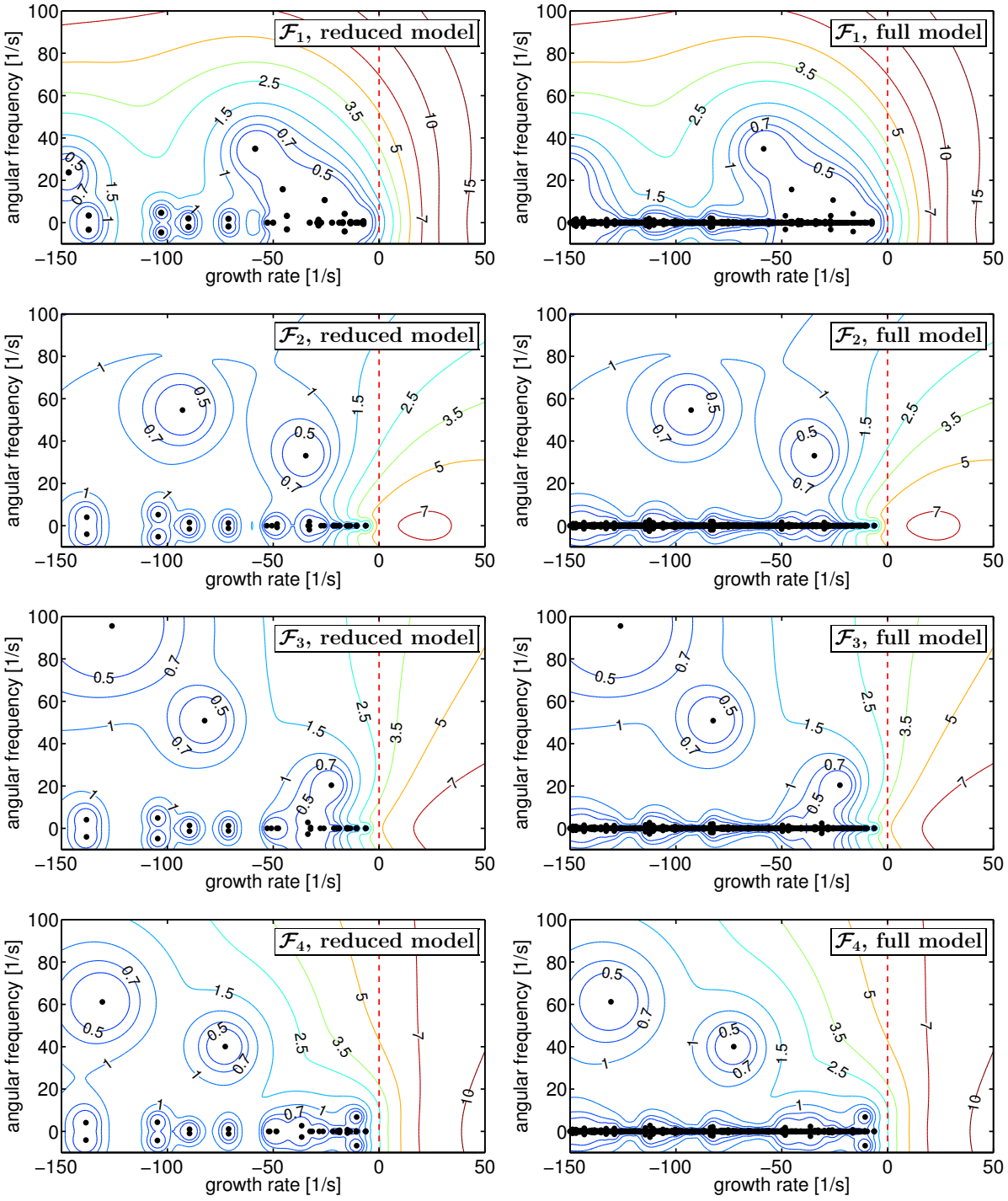

Figure 6. Spectra and boundaries of ϵ-pseudospectra $\gamma_\epsilon(\mathbf{A})$ after optimisation of the objective functions $\mathcal{F}_1$ through $\mathcal{F}_4$ and substitution of the optimal parameter values (table 1) into the reduced and the full model, as indicated; ϵ-values as given by the contour labels in units of s^{-1}. Black dots correspond to eigenvalues. Note that all the spectra and pseudospectra are symmetric with respect to the real axis because the system matrix $\mathbf{A}$ is real.

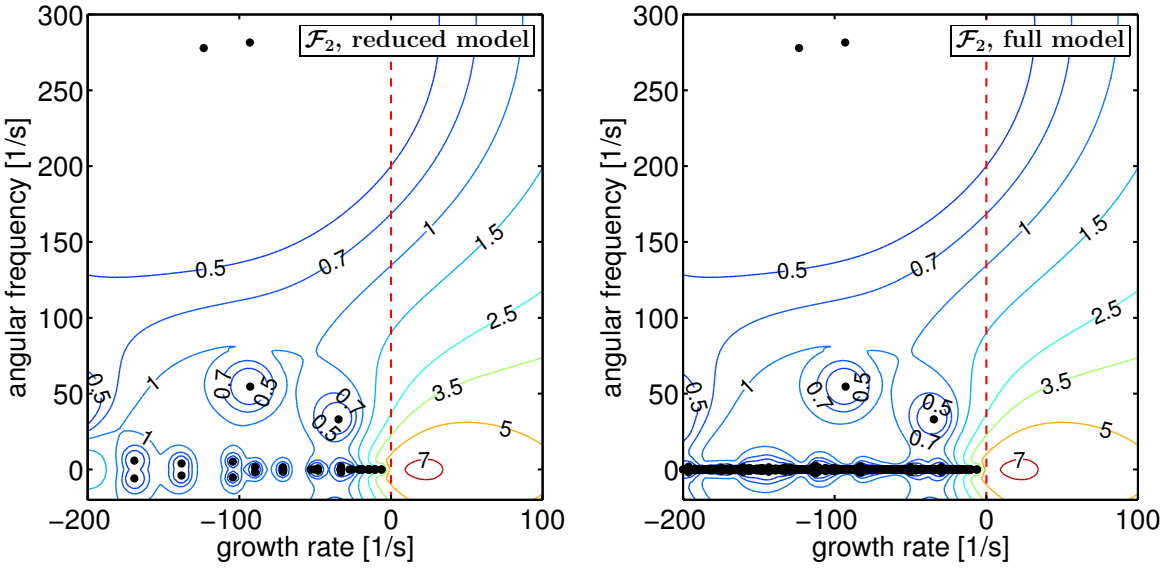

Figure 7. As figure 6, but for a larger domain in the complex plane and for the function $\mathcal{F}_2$ only.

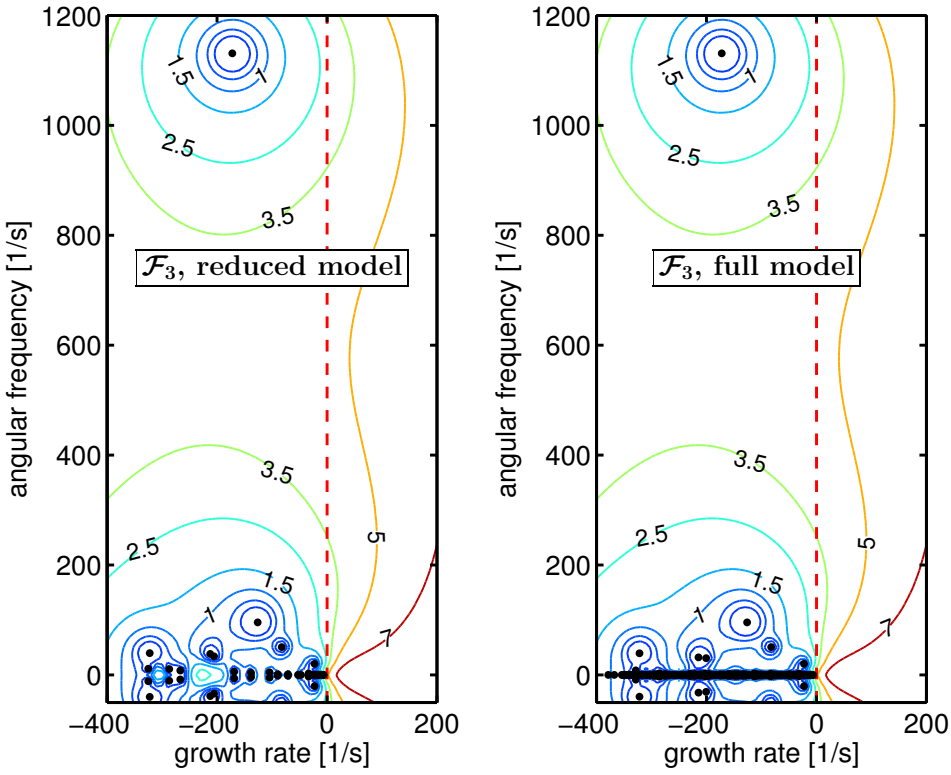

Figure 8. As figure 6, but for a larger domain in the complex plane and for the function $\mathcal{F}_3$ only.

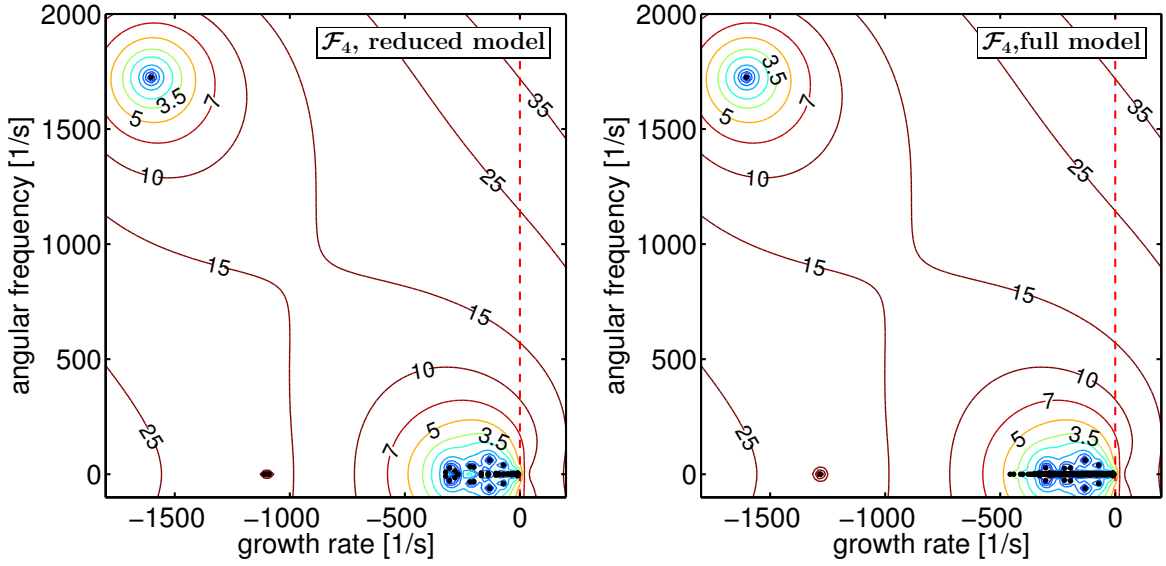

Figure 9. As figure 6, but for a larger domain in the complex plane and for the function $\mathcal{F}_4$ only.

model shows that the reduction method works as intended. The reduction is well suited for the pseudospectral approach, since the reduced model reproduces accurately the locations and sensitivities of the "most dangerous" eigenvalues. Since the optimization runs are finished within minutes in the simplified test case presented here, is is to be expected that the method can readily be applied to realistic ITER geometry with high level of detail. It is noteworthy that the truncation criterion used here (include all patterns whose variance fraction is greater than 10^{-12}) is extremely conservative and can be considerably relaxed in any case when the reduced models are found to be too large.

In the case of the $\mathcal{F}_1$-optimal parameter set, the ϵ-pseudospectra cover a strikingly small domain even for large ϵ values, compared to the other parameter sets. The eigenvalues are not extremely sensitive. Estimating the ratio on the right hand side of (4.6) by inspection of the contours for some of the larger ϵ-values given in figure 6, one finds values exceeding unity, but not by a large amount. It follows that the transient peak

does not necessarily need to be immoderate. However, the eigenvalues close to the imaginary axis are not insensitive enough to obtain very robust stability.

The solution after optimising $\mathcal{F}_2$ is characterised by two extremely sensitive complex eigenvalue pairs with large magnitude of imaginary parts (see figure 7), which deteriorate the robustness and also the transient behaviour. E. g., for $\epsilon = 1$ s^{-1}, $\gamma_\epsilon(\mathbf{A})$ overlaps a large region in the right half-plane due to the strong influence of the dangerous eigenvalues. By estimating $\sigma_\epsilon(\mathbf{A})$ for $\epsilon = 1$ s^{-1} from figure 7 and considering (4.6), one finds that some initial states must exist which are transiently amplified at least by a factor of about 80.

Optimisation of $\mathcal{F}_3$ and $\mathcal{F}_4$, respectively, results in a complex, sensitive eigenvalue pair with large imaginary part, as well (figures 8 and 9). These eigenvalues, however, have been pushed far enough to the left so that robustness is not affected by them. But in the case of the $\mathcal{F}_3$-optimal parameter set, the sensitive eigenvalue pair causes a significant protrusion of $\gamma_\epsilon(\mathbf{A})$ into the right half-plane for $\epsilon > \rho(\mathbf{A})$ and thus determines $\sigma_\epsilon(\mathbf{A})$ for these ϵ values. A large transient peak can be expected. For the $\mathcal{F}_4$-optimal solution, the sensitive pair is located so far on the left so that no influence on $\sigma_\epsilon(\mathbf{A})$ for any ϵ value considered in figure 9, and the transient amplification could be more moderate.

The $||\mathrm{e}^{t\mathbf{A}}||$ curves obtained for each optimal parameter set substituted into the full model are displayed in figure 10. The corresponding curves for the reduced model (not shown) are indistinguishable from those shown in the figure. Together with the excellent agreement concerning the objective function values and pseudospectra, this shows that the reduced model size has definitely been chosen sufficiently large. The $||\mathrm{e}^{t\mathbf{A}}||$ curves confirm the predictions made above. The transient amplification is quite modest for $\mathcal{F}_1$. On the other hand, $||\mathrm{e}^{t\mathbf{A}}||$ decays much slower than in all the other cases. For $\mathcal{F}_2$, the

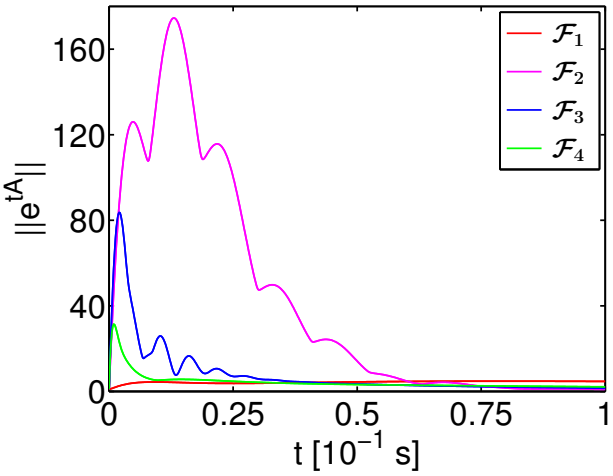

Figure 10. Transient amplification ($||e^{t\mathbf{A}}||$ curves) after optimisation of $\mathcal{F}_1$ through $\mathcal{F}_4$ and substituting each respective optimal parameter set into the full model.

transient behaviour is disastrous. There are amplifications by a factor of more than 170 and, in addition, the $||e^{t\mathbf{A}}||$ curve is comparatively broad. In the $\mathcal{F}_3$-optimal case, the peak is sharper, but the amplification factor still exceeds 80. Finally, the peak value is alleviated to about 30 for $\mathcal{F}_4$, which, however, is still somewhat unsatisfactory.

In order to understand which physical processes are responsible for transient growth, the initial state which suffers the strongest transient amplification in the $\mathcal{F}_3$-optimal case is presented in figure 11 as a current pattern. This state vector is given by the leading eigenvector of $(e^{t\mathbf{A}})^{\mathrm{T}}e^{t\mathbf{A}}$ with t being equal to the time where $||e^{t\mathbf{A}}||$ attains its maximum. The vector has been normalised to unity. The main feature of the current pattern is a structure composed of small-scale eddy currents, toroidally modulated with $n = 1$, which closely resembles the corresponding structures contributing to the stable current patterns $\boldsymbol{\pi}_4$ through $\boldsymbol{\pi}_6$ (figure 4). For the other parameter sets, the corresponding equivalents to the state shown in figure 11 are very similar (not shown), except that a significant RWM contribution is admixed in the $\mathcal{F}_1$ case. The mechanism of transient amplification is therefore much the same in all four cases: The small-scale eddies generate a strong $n = 1$ sensor signal resulting in a vigorous response of the coils and an erroneous,

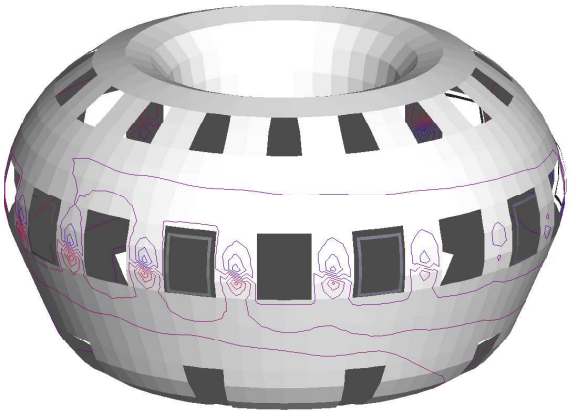

Figure 11. Normalised initial state which suffers the strongest transient amplification in the $\mathcal{F}_3$-optimal case.

rapid excitation of a high-amplitude RWM, which has to be controlled afterwards. This behaviour can be observed in a video sequence (`evolution_smalleddy.mpg`, 2.8 MB) showing the time evolution starting from the initial state displayed in figure 11 and using the $\mathcal{F}_3$-optimal feedback controller. Note that the normalisation of the coil colouring in the video is the same as in figure 4 in order to visualise the coil current distribution also in the later stage of the evolution. In the early stage, the colour range is "overdriven", i. e., the coil currents are much stronger than indicated by the colours.

For comparison, another video (`evolution_rwm.mpg`, 2.8 MB) shows the evolution for the same controller, but with a pure RWM (the normalised state $\mathbf{x}_1$, see left panel of figure 3) as the initial condition. Significant transient amplification (increase of the overall current intensity) is also present here, but it is much less severe than in the case of the small-eddy initial state. Interestingly, in both time evolutions the damped, slowly rotating RWM is superposed by a rapidly rotating, strongly damped mode in the initial phase of the evolution. This mode has a frequency of about 180 Hz and corresponds to the eigenvalue pair with large imaginary part magnitudes (the eigenvalue with large positive imaginary part is visible in figure 8).

Since the physical mechanism of transient amplification has now been revealed, new strategies can be developed to avoid transient amplification while retaining much of the robustness, or even improve the latter. These strategies can be used as an alternative, or in addition, to optimising $\mathcal{F}_4$ for large ϵ. Four different strategies, denoted A, B, C, and D in the following, shall be tested here: (A) use the constraint $T \geq 30$ ms, (B) shift the sensors from the radial position $R = 8.9$ m to 8.6 m, (C) keep the sensors at their original position $(R, Z) = (8.9\text{ m}, 0.45\text{ m})$ and add an additional array of 18 sensors at the same toroidal and radial positions, but at $Z = 0.0$ m, (D) shift both sensor arrays to $R = 8.6$ m. The motivation of keeping the feedback response slow, as done in case A, is to prevent the response growing large before small eddies as shown in figure 11 decay. Shifting the sensors away from the conducting wall decreases the strength of the eddy-generated magnetic field at the sensors. Finally, the main intention of adding the second sensor array is to improve robustness rather than to decrease transient amplification. Observability is enhanced by the fact that not only the poloidal field is measured, but also the poloidal derivative. Therefore, two sensor arrays should do better in distinguishing RWMs from intrinsically stable plasma modes than a single array. But it is also hoped that a double array is somewhat harder to be misled by small-scale currents in the wall than a single array, so that the transient behaviour should also benefit from the introduction of the second array.

The different approaches are tested by optimising $\mathcal{F}_3$ for all cases A through D subject to the same current constraint $I_{\max} = 74$ kA as before, and subject to $T \geq 1$ ms for the cases B through D. For C and D, the additional sensor array increases the number of free parameters from 3 to 5. The optimal $\mathcal{F}_3$ values are -2.27 s^{-1} (A), -4.28 s^{-1} (B), -3.99 s^{-1} (C), and -5.31 s^{-1} (D). The corresponding $||\mathrm{e}^{t\mathbf{A}}||$ plots are displayed in figure 12. Using the slow-response constraint (case A) dramatically reduces

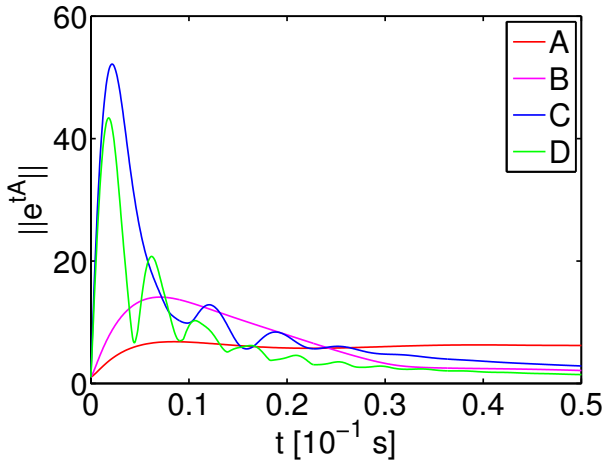

Figure 12. Transient amplification for strategies A through D, see text.

the transient peak, but at the cost of robustness and the decay rate of the $||e^{t\mathbf{A}}||$ curve. Compared to the $\mathcal{F}_3$-optimal standard case shown in figure 10, the peak value is reduced substantially in case B, but only slightly in case C. In case D, it is considerably larger than in case B, but D is more robust. Compared to the standard case, both strategies B and D lead to a considerable improvement of both robustness and transient behaviour, whereas strategy C is less efficient. The transient behaviour might still be considered not entirely satisfactory even for case B, but it can be further improved at the expense of robustness, by combining, e. g., strategy B with slowing down of the feedback response (like in strategy A) or optimising $\mathcal{F}_4$ for ϵ larger than the optimal complex stability radius.

7. Summary and discussion

A novel approach to the numerical treatment of stabilising RWMs in tokamaks by means of magnetic field sensors and active feedback coils has been presented. Instead of characterising the control system in terms of its input-output behaviour, as usual in control theory, the analysis has been focused on the internal dynamics of the

system. This approach is led by the insight that the state space representation of the system, which emerges from physical modelling and therefore exists for its own sake, contains much more information than needed to generate a particular input-output behaviour (many different systems can have the same input-output characteristics). This information surplus is exploited to analyse quantitatively, by means of a physically based measure, the control system's sensitivity to uncertainties of the model, or, in other words, the robustness of the system's stability with respect to changes of the system.

In this approach, the choice of the coordinate system used to define the system's state space is of particular significance. While the coordinate system is unimportant to the input-output characteristics, it defines a metrics in the state space, a measure of difference between system states. The basic assumption of this study is the special interest in a particular state vector norm which is typically based on physically measurable quantities. This norm is considered distinguished compared to all the other possible choices of norms. The induced metrics enables the investigator to decide whether the difference between system states is small or large. Therefore it can also be decided whether a change of the system (represented by changes of the system's matrix elements) can produce a large change in the tendency of the system state or not. The corresponding measure is the matrix norm which is vector-bound to the state vector norm of interest.

The control system model is formulated in terms of a parametrised matrix. The parameter values describe the rule how the active coils react to the sensor signals. These parameters can be varied in order to stabilise the system and optimise the system's stability with respect to certain criteria. To make such a parameter optimisation computationally feasible it is inevitable to reduce the dimension of the control system model while retaining the characteristics of its response to the feedback. Powerful

standard methods which serve this purpose are available in the literature, but they are based on oblique transformations. They do not preserve the state space metrics and are therefore not suitable for use in combination with the methodology of this study. For this reason, a novel, highly efficient orthogonal model reduction procedure denoted "isometric truncation" has been designed and utilised in this work.

For optimisation of the feedback parameters, the OPTIM code has been developed. Four different objective functions given by (4.7) are available which can be chosen to be optimised. Two of them, $\mathcal{F}_1$ and $\mathcal{F}_2$, can be considered as measures for asymptotic stability and take only the eigenvalues of the parametrised matrix into account. The other two functions, $\mathcal{F}_3$ and $\mathcal{F}_4$, are related not only to the eigenvalues but also to their sensitivity. Therefore, they act as measures for the robustness of the stability. They involve the matrix norm mentioned above and are in close relationship to matrix pseudospectra. The ϵ-pseudospectrum of a matrix is the union of eigenvalue spectra occurring for all perturbed matrices where the perturbation is smaller than ϵ. Optimising $\mathcal{F}_4$ for a suitably chosen ϵ value may help to attenuate possible transient amplification of initial states.

The objective functions are not everywhere smooth, in general. Non-standard techniques are required for their optimisation. The method implemented in the OPTIM code is based on sampling gradients from the neighbourhood of the current point in parameter space.

For a simple, ITER-like test case, the feedback parameters have been optimised with respect to $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$, and $\mathcal{F}_4$, respectively. The four resulting, feedback controlled systems have been compared with respect to asymptotic stability, eigenvalue sensitivity, robust stability, and transient behaviour by inspecting the values of the

different objective functions, the eigenvalues, pseudospectra, and $||e^{t\mathbf{A}}||$ plots. The solutions which have been optimised with respect to $\mathcal{F}_1$ and $\mathcal{F}_2$ are characterised by inferior robustness. The $\mathcal{F}_2$-optimal solution has extremely sensitive eigenvalues and a catastrophic transient peak. Optimising $\mathcal{F}_3$ produces a result which is by far more robust. Finally, the $\mathcal{F}_4$-optimal system is slightly less robust but exhibits less severe transient amplification than the $\mathcal{F}_3$-optimal case. The asymptotic stability of the latter two cases is not substantially worse than in the former two cases.

As a physical mechanism of the transient amplification, small-scale eddy currents in the conducting wall parts in the neighbourhood of the sensors have been identified. Based on this finding, additional strategies to mitigate the transient amplification have been developed and tested, namely, slowing down the response of the feedback system, moving the sensors away from the wall, and using an additional sensor array. All these strategies are successful to some extent and can be combined with each other. Taking these actions can either degrade or improve robustness. Testing such strategies in the case of a realistic ITER geometry is left for a future study.

It has been convincingly shown that it can be extremely dangerous to be satisfied with merely having stabilised the RWM control system. Inspecting the system's eigenvalues alone and finding that all of them are located well in the left complex half-plane tells nothing about the sensitivity of eigenvalues, robustness of the stability, and possible transient amplification. Pseudospectra plots are a useful tool to get an impression of the eigenvalue sensitivity. Computation of the complex stability radius (i. e., $\mathcal{F}_3$) provides a quantitative measure of robustness. Plotting $||e^{t\mathbf{A}}||$ reveals possibly dangerous transient amplification. It appears recommendable to use these tools routinely when designing RWM feedback controllers, and to use objective functions like $\mathcal{F}_3$ or $\mathcal{F}_4$ for optimisation.

To the authors' knowledge, initial conditions which are not RWMs have not yet been considered in any other RWM stabilisation study, and the possibility of severe transient amplification of these states has not yet been recognised by the fusion research community. On the other hand, it has to be noted that it cannot be inferred from this study to which extent the transient growth as predicted by $||e^{t\mathbf{A}}||$ is indeed relevant. The question remains if the states which are significantly transiently amplified, like the one shown in figure 11, are excited at any time in a "real world" situation. Since such states project mainly onto the stable subspace of the open-loop system, they can be expected to be driven only by physical processes not included in the model (see also the discussion concerning transient amplification of the sensor signal in subsection 2.3). Again, nonlinear, forced-dissipative MHD simulations would be required to estimate the significance of such "external" processes. Anyway, if the transient amplification will be found to be an issue, strategies like those developed in this study could be useful to mitigate this problem to some extent.

Acknowledgments

Many thanks to J. Bialek and J. Menard for providing the ITER wall and coil data, and to R. Meyer-Spasche and W. Suttrop for useful and interesting discussions. This work, supported by the European Communities under the contract of Association between EURATOM/IPP, was carried out within the framework of the European Fusion Development Agreement (under EFDA Contract No. 07-1705/1602). The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] A. M. Garofalo, A. D. Turnbull, E. J. Strait, M. E. Austin, J. Bialek, M.S. Chu, E. Frederickson, R. J. La Haye, G. A. Navratil, L. L. Lao, E. A. Lazarus, M. Okabayashi, B. W. Rice, S. A. Sabbagh, J. T. Scoville, T. S. Taylor, M. L. Walker, and the DIII-D Team. Stabilization of the external kink and control of the resistive wall mode in tokamaks. *Phys. Plasmas*, 6:1893–1898, 1999.

[2] A. M. Garofalo, M. S. Chu, E. D. Frederickson, M. Gryaznevich, T.H. Jensen, L.C. Johnson, R. J. La Haye, G. A. Navratil, M. Okabayashi, J. T. Scoville, E. J. Strait, A. D. Turnbull, and the DIII-D Team. Resistive wall mode dynamics and active feedback control in DIII-D. *Nucl. Fusion*, 41:1171–1176, 2001.

[3] M. Okabayashi, J. Bialek, M. S. Chance, M. S. Chu, E. D. Frederickson, A. M. Garofalo, M. Gryaznevich, R. E. Hatcher, T. H. Jensen, L. C. Johnson, R. J. La Haye, E. A. Lazarus, M. A. Makovski, J. Manickam, G. A. Navratil, J. T. Scoville, E. J. Strait, A. D. Turnbull, M. L. Walker, and the DIII-D Team. Active feedback stabilization of the resistive wall mode on the DIII-D device. *Phys. Plasmas*, 8:2071–2082, 2001.

[4] J. Bialek, A. H. Boozer, M. E. Mauel, and G. A. Navratil. Modeling of active control of external magnetohydrodynamic instabilities. *Phys. Plasmas*, 8:2170–2180, 2001.

[5] O. Katsuro-Hopkins, J. Bialek, D. A. Maurer, and G. A. Navratil. Enhanced ITER resistive wall mode feedback performance using optimal control techniques. *Nucl. Fusion*, 47:1157–1165, 2007.

[6] M. S. Chu, M. S. Chance, A. H. Glaser, and M. Okabayashi. Normal mode approach to modelling of feedback stabilization of the resistive wall mode. *Nucl. Fusion*, 43:441–454, 2003.

[7] A. Bondeson, Y. Q. Liu, C. M. Fransson, B. Lennartson, C. Breitholz, and T. S. Taylor. Active feedback stabilization of high beta modes in advanced tokamaks. *Nucl. Fusion*, 41:455–463, 2001.

[8] A. Bondeson, Y. Q. Liu, D. Gregoratto, C. M. Franson, and Y. Gribov. Physics and stabilization of resistive wall modes in tokamaks. *Plasma Phys. Control. Fusion*, 45:A253–A267, 2003.

[9] Y. Liu, A. Bondeson, Y. Gribov, and R. Polevoi. Stabilization of resistive wall modes in ITER by active feedback and toroidal rotation. *Nucl. Fusion*, 44:77–86, 2004.

[10] F. Villone, G. Rubinacci, Y. Q. Liu, and Y. Gribov. Analysis of RWM with a 3D model of conducting structures. *32nd EPS conference on Plasma Physics. Tarragona*, 29C:5049–5052,

2005.

[11] P. Merkel and M. Sempf. Feedback stabilization of resistive wall modes in the presence of multiply-connected wall structures. *21st IAEA Fusion Energy Conference 2006, Chengdu, China, paper TH/P3-8*, 2006.

[12] E. Strumberger, P. Merkel, M. Sempf, and S. Günter. On fully three-dimensional resistive wall mode and feedback stabilization studies. *Phys. Plasmas*, 15:056110, 2008. DOI:10.1063/1.2884579.

[13] M. Green and D. J. N. Limebeer. *Linear Robust Control*. Prentice Hall, Englewood Cliffs, New Jersey, 1995.

[14] Y. Q. Liu and A. Bondeson. Robust control of resistive wall modes in tokamaks. *Plasma Phys. Control. Fusion*, 44:L21–L28, 2002.

[15] Y. Q. Liu, A. Bondeson, D. Gregoratto, C. M. Fransson, Y. Gribov, and R. Paccagnella. Feedback control of resistive wall modes in toroidal devices. *Nucl. Fusion*, 44:77–86, 2004.

[16] J. V. Burke, A. S. Lewis, and M. L. Overton. A nonsmooth, nonconvex optimization approach to robust stabilization by static output feedback and low-order controllers. *Proc. IFAC Symp. Robust Control Design, Milan, Italy*, 2003.

[17] L. N. Trefethen and M. Embree. *Spectra and Pseudospectra - The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, Princeton and Oxford, 2005. 606 pp.

[18] P. Merkel. 2009. In preparation.

[19] I. B. Bernstein, E. A. Frieman, M. D. Kruskal, and B. M. Kulsrud. An energy principle for hydromagnetic stability problems. *Proc. R. Soc. London, Ser. A*, 244:17–40, 1958.

[20] J. P. Freidberg. *Ideal Magnetohydrodynamics*. Plenum Press, New York and London, 1987. 489 pp.

[21] C. Nührenberg. Global ideal magnetohydrodynamic stability analysis for the configurational space of Wendelstein 7-X. *Phys. Plasmas*, 3:2401–2410, 1996.

[22] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Rice University, Houston, Texas, 2005.

[23] G. Obinata and B. D. O. Anderson. *Model reduction for control system design*. Springer, London, 2001.

[24] B. C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE T. Automat. Contr.*, AC-26:17–32, 1981.

[25] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *Int. J. Control*, 39:1115–1193, 1984.

[26] A. J. Laub, M. T. Heath, C. C. Paige, and R. C. Ward. Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms. *IEEE T. Automat. Contr.*, AC-32:115–122, 1987.

[27] M. S. Tombs and I. Postlethwaite. Truncated balanced realization of a stable non-minimal state-space system. *Int. J. Control*, 46:1319–1330, 1987.

[28] M. G. Safonov and R. Y. Chiang. A Schur method for balanced-truncation model reduction. *IEEE T. Automat. Contr.*, 34:729–733, 1989.

[29] A. Varga. Balancing-free square-root algorithm for computing singular perturbation approximations. *Proc. 30th IEEE Conf. Dec. Control, Brighton, UK*, pages 1062–1065, 1991.

[30] A. Varga. Efficient minimal realization procedure based on balancing. *Prepr. of the IMACS Symp. on Modelling and Control of Technological Systems*, 2:42–47, 1991.

[31] H. Hotelling. Analysis of a complex of statistical variables into principal components. *J. Educ. Psych.*, 24:417–441, 1933.

[32] N. A. Bruinsma and M. Steinbuch. A fast algorithm to compute the H_∞-norm of a transfer function matrix. *Syst. Contr. Lett.*, 14:287–293, 1990.

[33] J. V. Burke, A. S. Lewis, and M. L. Overton. Robust stability and a criss-cross algorithm for pseudospectra. *IMA J. Numer. Anal.*, 23:359–375, 2003.

[34] R. Meyer-Spasche and S. Huber. Spektren und Pseudospektren - Theorie, Numerik und Anwendungen (Spectra and pseudospectra: Theory, Numerics and Applications). Technical Report IPP 5/120, Max-Planck-Institut für Plasmaphysik, Garching, Germany, 2008. available online at `http://www.ipp.mpg.de/ippcms/de/kontakt/bibliothek/ipp_reports/index.html`. In German.

[35] M. Szularz, J. Weston, and M. Clint. *Euro-Par'97 Parallel Processing, Lecture Notes in Computer Science*, chapter Robust parallel Lanczos methods for clustered eigenvalues. Springer, Berlin, Heidelberg, 1997.

[36] N. J. Higham. The scaling and squaring method for the matrix exponential revisited. *SIAM J. Matrix Anal. Appl.*, 26:1179–1193, 2005.

[37] J. V. Burke, A. S. Lewis, and M. L. Overton. Two numerical methods for optimizing matrix stability. *Linear Algebra Appl.*, 351-352:117–145, 2002.

[38] J. V. Burke, A. S. Lewis, and M. L. Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM J. Optimiz.*, 15:751–779, 2005.

[39] C. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[40] J. V. Burke, A. S. Lewis, and M. L. Overton. Optimization and pseudospectra, with applications to robust stability. *SIAM J. Matrix Anal. Appl.*, 25:80–104, 2003.

[41] D. Avis. lrs (software for vertex enumeration/convex hull problems), 2005. available at `http://cgm.cs.mcgill.ca/~avis/C/lrs.html`.

[42] NAG Fortran Library, Mark 21, The Numerical Algorithms Group Ltd., 2005.

[43] M. S. Chu, V. S. Chan, M. S. Chance, D. H. Edgell, A. M. Garofalo, A. H. Glasser, S. C. Guo, D. A. Humphreys, T. H. Jensen, J. S. Kim, R. J. La Haye, L. Lao, G. A. Navratil, M. Okabayashi, F. W. Perkins, H. Reimerdes, H. E. St. John, E. Soon, E. J. Strait, A. D. Turnbull, M. L. Walker, and S. K. Wong. Modelling of feedback and rotation stabilization of the resistive wall mode in tokamaks. *Nucl. Fusion*, 43:196–201, 2003.