A NONSMOOTH, NONCONVEX OPTIMIZATION APPROACH TO ROBUST STABILIZATION BY STATIC OUTPUT FEEDBACK AND LOW-ORDER CONTROLLERS

James V. Burke [*,1] **Adrian S. Lewis** [**,2]
Michael L. Overton [***,3]

[*] *University of Washington, Seattle, WA*
[**] *Simon Fraser University, Burnaby, BC*
[***] *New York University, New York, NY*

Abstract: Stabilization by static output feedback (SOF) is a long-standing open problem in control: given an n by n matrix A and rectangular matrices B and C, find a p by q matrix K such that $A + BKC$ is stable. Low-order controller design is a practically important problem that can be cast in the same framework, with $(p+k)(q+k)$ design parameters instead of pq, where k is the order of the controller, and $k << n$. Robust stabilization further demands stability in the presence of perturbation and satisfactory transient as well as asymptotic system response. We formulate two related nonsmooth, nonconvex optimization problems over K, respectively with the following objectives: minimization of the ϵ-pseudospectral abscissa of $A + BKC$, for a fixed $\epsilon \geq 0$, and maximization of the complex stability radius of $A + BKC$. Finding global optimizers of these functions is hard, so we use a recently developed gradient sampling method that approximates local optimizers. For modest-sized systems, local optimization can be carried out from a large number of starting points with no difficulty. The best local optimizers may then be investigated as candidate solutions to the static output feedback or low-order controller design problem. We show results for two problems published in the control literature. The first is a turbo-generator example that allows us to show how different choices of the optimization objective lead to stabilization with qualitatively different properties, conveniently visualized by pseudospectral plots. The second is a well known model of a Boeing 767 aircraft at a flutter condition. For this problem, we are not aware of any SOF stabilizing K published in the literature. Our method was not only able to find an SOF stabilizing K, but also to locally optimize the complex stability radius of $A + BKC$. We also found locally optimizing order–1 and order–2 controllers for this problem. All optimizers are visualized using pseudospectral plots.

Keywords: Static output feedback, low-order controller, nonsmooth optimization, H-infinity norm, pseudospectra, spectral abscissa, stability radius

1. STATIC OUTPUT FEEDBACK AND LOW-ORDER CONTROLLERS

Given an $n \times n$ matrix A, an $n \times p$ matrix B and a $p \times q$ matrix C, the *stabilization by static output feedback* (SOF) problem is to find a $p \times q$ matrix K such that $A + BKC$ is stable (has all its eigenvalues in the left half-plane). How to efficiently find such a K (or show that it does not exist) is a long-standing open problem in control (Blondel *et al.*, 1995).

Given the same data, together with an integer $k < n$, the *low-order controller* design problem is to find K_1, K_2, K_3 and K_4, respectively with dimensions $p \times q$, $p \times k$, $k \times q$ and $k \times k$, such that

$$\begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} K_1 & K_2 \\ K_3 & K_4 \end{bmatrix} \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} \qquad (1)$$

is stable. The problem is mainly of interest when k, the order of the controller, is much less than n. When $k = n$, the task of finding a stabilizing solution or showing that it does not exist is efficiently solvable by well known techniques based on linear matrix inequalities and convex optimization (Boyd *et al.*, 1994). The low-order controller design problem is an SOF problem with $(p+k)(q+k)$ design variables instead of pq, with the case $k = 0$ being the original SOF problem.

2. ROBUST STABILIZATION

As is well known, stabilization is a necessary but far from sufficient property for a controlled system to be well behaved. The notion of "robust stability" has emerged over the past two decades as an important concept in control and dynamical systems. In the robust control community, the *complex stability radius* (Hinrichsen and Pritchard, 1986) has become a very well known measure of robust stability. This notion may be interpreted in terms of *pseudospectra* (Trefethen, 1990), an approach to measuring system behavior that has become well known in the numerical analysis and PDE communities, and for which powerful graphical tools have recently become available.

For a real $\epsilon \geq 0$, the ϵ-*pseudospectrum* of a matrix $X \in \mathbf{R}^{n \times n}$ is the subset of the complex plane defined by

$$\Lambda_\epsilon(X) = \{z \in \mathbf{C} : \sigma_{\min}(X - zI) \leq \epsilon\},$$

where $\sigma_{\min}$ denotes smallest singular value. Equivalently (Trefethen, 1990),

$$\Lambda_\epsilon(X) = \{z \in \mathbf{C} : \sigma_{\min}(Y - zI) = 0, \text{ for some } Y \in \mathbf{C}^{n \times n} \text{ with } \|Y - X\| \leq \epsilon\}.$$

Here $\|\cdot\|$ denotes the 2-norm (largest singular value). Note that Y may be complex. Thus the pseudospectrum of a matrix is the union of the spectra of nearby matrices. When $\epsilon = 0$, the pseudospectrum reduces to the spectrum.

In this paper we focus on two measures of robust stability of a matrix X. One is the ϵ-*pseudospectral abscissa* $\alpha_\epsilon(X)$, defined as the largest real part of all elements of the pseudospectrum $\Lambda_\epsilon(X)$, for a fixed ϵ. The case $\epsilon = 0$ is the *spectral abscissa.* The other measure is the *complex stability radius* $\beta(X)$, also known as the *distance to the unstable matrices.* It is defined to be 0 if X is not stable and otherwise as the largest ϵ for which the pseudospectrum $\Lambda_\epsilon(X)$ lies in the left half-plane. More directly, without explicit reference to pseudospectra, we have

$$\alpha_\epsilon(X) = \max_{z \in \mathbf{C}} \{\mathrm{Re}\ z : \sigma_{\min}(X - zI) \leq \epsilon\}, \quad (2)$$

and

$$\beta(X) = \min_{z \in \mathbf{C}} \{\sigma_{\min}(X - zI) : \mathrm{Re}\ z \geq 0\}. \quad (3)$$

It is not hard to see that the connection between these two concepts is summarized by the facts that

$$\alpha_\epsilon(X) \geq 0 \quad \Leftrightarrow \quad \beta(X) \leq \epsilon.$$

For a matrix X that is not stable, we have $\beta(X) = 0$, while for stable X, $\alpha_{\beta(X)}(X) = 0$, and the minimization in (3) may as well be done over z on the imaginary axis. As is well known, $\beta(X)^{-1}$ is the $\mathbf{H}_\infty$ norm of the associated transfer function $(sI - X)^{-1}$.

The relationships between the quantities $\alpha_\epsilon(X)$ and $\beta(X)$ and the *robustness* of the stability of X with respect to perturbation are clear from the definitions. Less obvious is that these functions also measure the *transient response* of the associated dynamical system $\dot{v} = Xv$ (Trefethen, 1997). By choosing $\epsilon < \beta(X)$ in α_ϵ, we place more emphasis on asymptotic behavior than when we choose $\epsilon = \beta(X)$. Stated differently, we choose ϵ according to the size of perturbation we are prepared to tolerate, and measure what kind of asymptotic response this allows us to guarantee, instead of measuring the largest perturbation that can be tolerated while still guaranteeing stability. In robust control, another important consideration is that perturbations are typically *structured*, but we do not address this in the present work. In particular, since perturbations are typically real, not complex, another interesting measure of robust stability is the *real* stability radius; however, this does *not* provide a measure of transient response.

[1] Partially supported by NSF Grant DMS-0203175
[2] Partially supported by NSERC
[3] Partially supported by NSF Grant CCR-0098145

Efficient algorithms to compute the pseudospectral abscissa $\alpha_\epsilon(X)$ and the complex stability radius $\beta(X)$ are available. For the latter, which is a special case of more general $\mathbf{H}_\infty$ norm computations, fast algorithms based on computing eigenvalues of Hamiltonian matrices are well known (Boyd and Balakrishnan, 1990; Bruinsma and Steinbuch, 1990), and software is available in the MATLAB control toolboxes and the SLICOT library. For α_ϵ, a related algorithm has been developed recently (Burke *et al.*, 2003*c*), and software is also available. [4]

3. NONSMOOTH, NONCONVEX OPTIMIZATION

With these measures of robust stability at hand, we are now in a position to approach the static output feedback stabilization problem via nonsmooth, nonconvex optimization. Consider the optimization problems

$$\min_{K\in\mathbf{R}^{p\times q}} \alpha_\epsilon(A+BKC) \tag{4}$$

and

$$\min_{K\in\mathbf{R}^{p\times q}} -\beta(A+BKC). \tag{5}$$

Both α_ϵ and $-\beta$ are nonsmooth, nonconvex functions. However, they are locally Lipschitz, except α_0 (the spectral abscissa). At the opposite extreme, the limit of $\alpha_\epsilon(X)-\epsilon$ as $\epsilon\to\infty$ is the convex function $\frac{1}{2}\lambda_{\max}(X+X^T)$ (the largest eigenvalue of the symmetric part of X). In fact, as ϵ is increased from zero to an arbitrarily large quantity, the optimization problem (4) evolves from minimization of $\alpha_0(A+BKC)$ (enhancing the *asymptotic* decay rate of the associated dynamical system) to the minimization of $\lambda_{\max}(A+BKC+A^T+C^TK^TB^T)$ (minimizing the *initial* growth rate of the associated system). For ϵ equal to the optimal value of (5), the optimization problem (4) has the same solution set as (5). These issues are explored at greater length in (Burke *et al.*, 2003*a*).

We approach robust stabilization by searching for minimizers of (4) and (5). However, because these problems are not convex, finding global minimizers can be expected to be difficult. Indeed, it is known that if we add interval bounds on the entries of K, the problem (4) is NP-hard in the case $\epsilon = 0$ (Blondel and Tsitsiklis, 1997). Consequently, we focus on approximating local minimizers, which may then be investigated as candidate solutions for robust stabilization by static output feedback.

[4] http://www.cs.nyu.edu/faculty/overton/software/

Not only are the optimization problems (4) and (5) nonconvex, they are also nonsmooth, so that standard local optimization methods such as steepest descent and BFGS are inapplicable (when tried, they typically "jam" at a point of discontinuity of the gradient of the objective, forcing the line search to repeatedly take tiny steps). The state of the art for both smooth, nonconvex optimization and for nonsmooth, convex optimization is quite advanced, and reliable software is in wide use, but nonsmooth, nonconvex optimization problems are much more difficult. However, we have devised a method based on *gradient sampling* that is very effective in practice and for which we have developed a local convergence theory (Burke *et al.*, 2002; Burke *et al.*, 2003*b*). This method is intended for minimizing functions f that are continuous and for which the gradient exists and is readily computable *almost everywhere* on the design parameter space, even though the gradient may be (and often is) discontinuous at an optimizing solution. Briefly, the method generates a sequence of points $\{x^\nu\}$ in the design space, say $\mathbf{R}^m$, as follows. Given x^ν, the gradient ∇f is computed at x^ν and at randomly generated points near x^ν within a sampling diameter η, and the convex combination of these gradients with smallest 2-norm, say d, is computed by solving a quadratic program. One should view $-d$ as a kind of stabilized steepest descent direction. A line search is then used to obtain $x^{\nu+1} = x^\nu - td/\|d\|$, with $f(x^{\nu+1}) < f(x^\nu)$, for some $t \leq 1$. If $\|d\|$ is below a prescribed tolerance, or a prescribed iteration limit is exceeded, the sampling diameter η is reduced by a prescribed factor, and the process is repeated. For the numerical examples in the next section, we used sampling diameters 10^{-j}, $j = 1,\ldots,6$, with a maximum of 100 iterates per sampling diameter and a tolerance 10^{-6} for $\|d\|$, and we set the number of randomly generated sample points to $2m$ (twice the number of design variables) per iterate. Besides its simplicity and wide applicability, a particularly appealing feature of the gradient sampling algorithm is that it provides approximate "optimality certificates": if $\|d\|$ is small for a small sampling diameter η, one can be reasonably sure that a local minimizer has been approximated.

In order to apply the gradient sampling algorithm to the optimization problems (4) and (5), we need to know how to compute the relevant gradients. References for methods to evaluate the objective functions α_ϵ and β were given at the end of the previous section; implementations respectively return a maximizing z in (2) and minimizing z in (3) as well as the value of the function. These allow us to compute the gradients of α_ϵ and β as functions on $\mathbf{R}^{n\times n}$, as follows. Except on a set of measure zero in $\mathbf{R}^{n\times n}$, we can assume that the rightmost points of the pseudospectrum $\Lambda_\epsilon(X)$ (the set of maximizers in (2)) consist either of a single real point z or a complex conjugate pair $(z, \bar{z})$. It follows (Burke *et al.*, 2003*a*) that

$$\nabla\alpha_\epsilon(X) = \mathrm{Re}\left(\frac{1}{u^*v}uv^*\right),$$

where u and v are respectively left and right singular vectors corresponding to $\sigma_{\min}(X - zI)$. Typically, there is never a tie for the definition of the maximizing z for iterates generated by the gradient sampling optimization algorithm, although at an exact minimizer of (4), there very often *is* a tie — hence the nonsmoothness of the function. A second source of nonsmoothness could arise if the singular value $\sigma_{\min}(X - zI)$ is multiple, but this does not usually occur, even at an exact minimizer of (4).

When $\epsilon = 0$, the function $\alpha_\epsilon(X)$ is simply the maximum of the real parts of the eigenvalues of X. In this case, the singular vectors u and v are left and right eigenvectors of X for the eigenvalue z. For α_0, there are actually two different sources for the nonsmoothness of the function: the possibility of ties for the rightmost eigenvalue, and the possibility of multiple eigenvalues, which in fact often occur at minimizers of spectral abscissa optimization problems (Burke *et al.*, 2001).

Likewise, for stable X, except on a set of measure zero, we can assume that the set of points where the pseudospectrum $\Lambda_{\beta(X)}(X)$ is tangent to the imaginary axis (the set of minimizers in (3)) consists either of a single real point z or a complex conjugate pair $(z, \bar{z})$. It follows that

$$\nabla\left(-\beta(X)\right) = \mathrm{Re}\left(uv^*\right),$$

where u and v are left and right singular vectors satisfying $u^*(X - zI)v = \sigma_{\min}(X - zI)$. Again, typically there is never a tie for the definition of the minimizing z at points generated by the algorithm, although at an exact minimizer of (5) there very often is a tie.

Finally, we need the gradients in the design parameter space. Let $\mathcal{A}$ denote the affine function that maps $K \in \mathbf{R}^{p\times q}$ to $A + BKC \in \mathbf{R}^{n\times n}$. From the ordinary chain rule, the gradient of the composition of any function ϕ on $\mathbf{R}^{n\times n}$ with $\mathcal{A}$ is

$$\nabla\left(\phi \circ \mathcal{A}\right)(K) = B^T \nabla\phi\left(A + BKC\right)C^T,$$

a matrix in $\mathbf{R}^{p\times q}$ that can also be interpreted as a vector in the design parameter space $\mathbf{R}^{pq}$. Applying this with ϕ equal to either α_ϵ or β provides the gradients needed to implement the gradient sampling algorithm in $\mathbf{R}^m$, with $m = pq$.

4. NUMERICAL EXAMPLES

We have applied gradient sampling optimization to many static output feedback (SOF) stabilization problems published in the control literature, all kindly provided to us by F. Leibfritz. We now describe results for two problems that we think are especially instructive. For the first, we consider only the pure SOF problem, while for the second we consider SOF and low-order controllers. [5]

The first set of data matrices A, B, C comes from a turbo-generator model (Hung and MacFarlane, 1982, Appendix E). For this problem, $n = 10$ and $p = q = 2$, so there are four design parameters in the static output feedback matrix K. Each of Figures 1 through 4 shows a pseudospectral plot in the complex plane, showing, for a particular K, the boundary of $\Lambda_\epsilon(A + BKC)$ for four different values of ϵ. The legend at the right of each figure shows the logarithms (base 10) of the four values of ϵ used in the plot. A particular pseudospectrum $\Lambda_\epsilon(A + BKC)$ may or may not be connected, but each connected component must contain one or more eigenvalues, shown as solid dots. The figures do not show all 10 eigenvalues; in particular, they do not show a conjugate pair of eigenvalues with large imaginary part, whose corresponding pseudospectral components are well to the left of the ones appearing in the figures. In Figures 3 and 4, the smallest real eigenvalue is also outside the region shown. The pseudospectral contours were plotted by T. Wright's EigTool, an extremely useful graphical interface for interpreting spectral and pseudospectral properties of nonsymmetric matrices. [6]

Figure 1 shows the pseudospectra of the original matrix A, that is, with $K = 0$. Although A is stable, since its eigenvalues are to the left of the imaginary axis, it is not robustly stable, since three connected components of the 10^{-2}-pseudospectrum cross the imaginary axis.

Figure 2 shows the pseudospectra of $A + BKC$ when K solves (4) with $\epsilon = 0$, or, in other words, when the rightmost eigenvalue of $A + BKC$ is pushed as far as possible to the left. Notice that six eigenvalues of $A + BKC$ are now arranged on a line parallel to the imaginary axis, and that two of the conjugate pairs are quite close to each other, indicating the possibility that at an exact minimizer there is a double conjugate pair of eigenvalues as well as a simple conjugate pair with the same real part. The 10^{-2}-pseudospectrum is now contained in the left half-plane, but the $10^{-1.5}$-pseudospectrum is not. The optimizing K is approximately

$$\begin{bmatrix} -2.2227 & -5.6120 \\ -0.0087 & -0.0230 \end{bmatrix}.$$

[5] The problem data matrices A, B and C are available at http://www.cs.nyu.edu/faculty/overton/papers/SOFdata/

[6] Freely available from the Pseudospectra Gateway, http://web.comlab.ox.ac.uk/projects/pseudospectra/

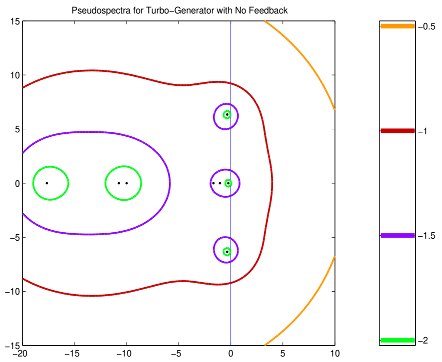

Fig. 1. Pseudospectra for Turbo-Generator with No Feedback

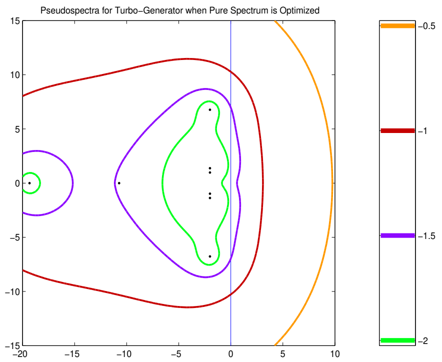

Fig. 2. Pseudospectra for Turbo-Generator when Pure Spectrum is Optimized

Figure 3 shows the pseudospectra of $A + BKC$ when K solves (4) with $\epsilon = 10^{-1.5}$, or, in other words, when the rightmost part of the $10^{-1.5}$-pseudospectrum of $A + BKC$ is pushed as far as possible to the left. Now the $10^{-1.5}$-pseudospectrum is to the left of the imaginary axis, but the eigenvalues have moved back towards the right, compared to Figure 2. There is apparently a three-way tie for the maximizing z in (2) at the local minimizer — one real value in its own small pseudospectral component, and two conjugate pairs in a much larger pseudospectral component. The optimizing K is approximately

$$\begin{bmatrix} -0.0641 & -0.3684 \\ -0.0341 & -0.0398 \end{bmatrix}.$$

Figure 4 shows the pseudospectra of $A + BKC$ when K solves (5) (maximizes the complex stability radius β), or, in other words, when K is chosen to maximize the ϵ for which the ϵ-pseudospectrum of $A + BKC$ is contained in the left half-plane. For this optimal value, $\epsilon = 10^{-1.105}$, the ϵ-pseudospectrum is tangent to the imaginary axis at five points, a real point and two conjugate pairs, indicating (as previously) a three-way tie for the minimizing z in (5). Each minimizing z has its own pseudospectral component. This ϵ-pseudospectrum crosses the imaginary axis in the previous three figures. On the other hand, the $10^{-1.5}$-pseudospectrum is now further to the right than it was in Figure 3. The optimizing K is approximately

$$\begin{bmatrix} -0.7763 & -0.7193 \\ -0.0935 & -0.1515 \end{bmatrix}.$$

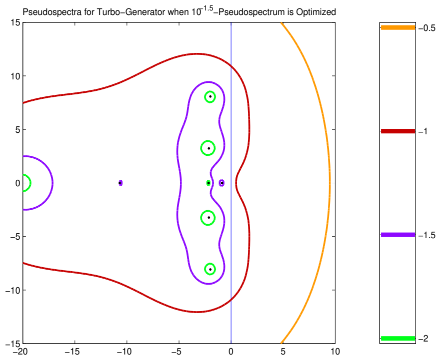

Fig. 3. Pseudospectra for Turbo-Generator when $10^{-1.5}$-Pseudospectrum is Optimized

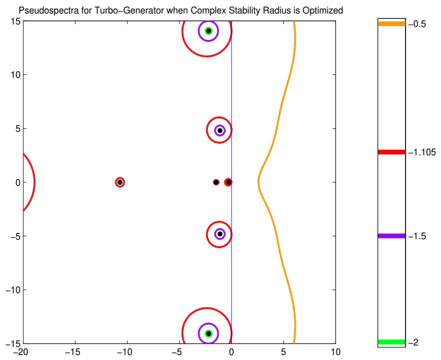

Fig. 4. Pseudospectra for Turbo-Generator when Complex Stability Radius is Optimized

Because the gradient sampling algorithm approximately verifies first-order optimality conditions, we are reasonably confident that the matrices K yielding Figures 2, 3 and 4 are all approximate local minimizers of their respective objectives. We cannot be sure that the minimizers are global, but it seems very likely that they are, based on the fact that repeated optimization runs from different starting points continued to produce approximately the same local minimizer. We do not claim that the matrices K displayed above are accurate approximations to local minimizers; because of the nonsmoothness, the objective value is much more sensitive to some perturbations in K than to others.

We now turn to our second example, a much more difficult stabilization problem. This is a well known model of a Boeing 767 aircraft at a flutter condition (Davison, 1990, Problem No. 90-06). The state space has dimension $n = 55$, but $p = q = 2$, so there are only 4 design parameters in the SOF matrix K. We are not aware of any SOF stabilizing solution published in the literature up until now. Figure 5 shows pseudospectra for

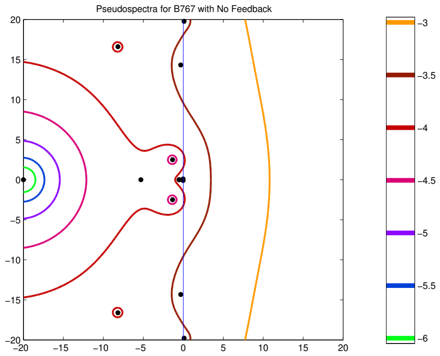

Fig. 5. Pseudospectra for B767 with No Feedback

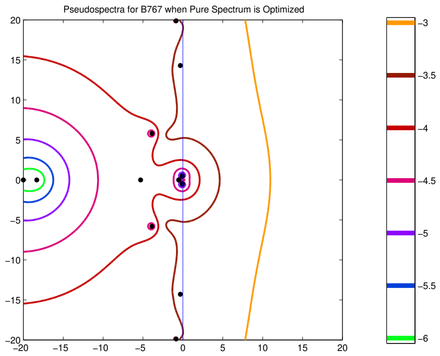

Fig. 6. Pseudospectra for B767 when Pure Spectrum is Locally Optimized for SOF Controller

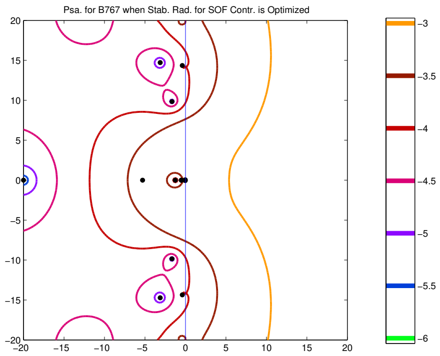

Fig. 7. Pseudospectra for B767 when Complex Stability Radius is Locally Optimized for SOF Controller

the most interesting part of the spectrum of the unstable matrix A, with no feedback. There is a conjugate pair of unstable eigenvalues near the top and bottom of the figure.

We used the gradient sampling method starting from many different initial points to search for local optimizers of the SOF spectral abscissa problem (4) (with $\epsilon = 0$) and the SOF complex stability radius problem (5). We found that it was useful to first search for minimizers of the spectral abscissa and then, once a stabilizing K was obtained, use it to initialize maximization of the complex stability radius, both explicitly and with randomly generated starting points whose

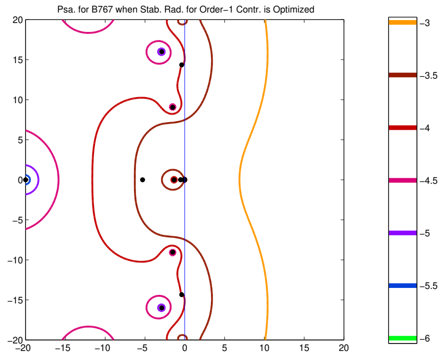

Fig. 8. Pseudospectra for B767 when Complex Stability Radius is Locally Optimized for Order–1 Controller

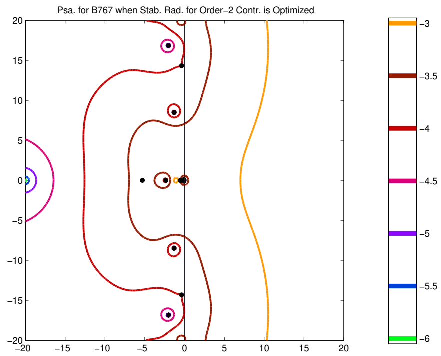

Fig. 9. Pseudospectra for B767 when Complex Stability Radius is Locally Optimized for Order–2 Controller

scaling was determined by the spectral abscissa minimizer.

Figure 6 shows the pseudospectra of $A + BKC$ for the best SOF local optimizer found for (4) with $\epsilon = 0$. The locally optimal spectral abscissa is approximately -7.79×10^{-2}, with approximate minimizer

$$K = \begin{bmatrix} -8.1908e{-}02 & 2.2651e{-}05 \\ -3.8308e{+}00 & 2.2771e{-}04 \end{bmatrix}.$$

Although $A + BKC$ is stable, it is far from being robustly stable. Its complex stability radius is 6.6×10^{-7}. Thus even the 10^{-6}-pseudospectrum crosses the imaginary axis.

Figure 7 shows the pseudospectra of $A+BKC$ for the best SOF local optimizer found for (5). The locally optimal complex stability radius is approximately 7.91×10^{-5}, with approximate maximizer

$$K = \begin{bmatrix} 4.5457e{+}00 & 5.9329e{-}05 \\ 7.2997e{+}00 & 2.1689e{-}04 \end{bmatrix}.$$

Note that the 10^{-4}-pseudospectrum crosses the imaginary axis, but the $10^{-4.5}$-pseudospectrum does not.

Finally, we considered low-order controller design for the same problem, using gradient sampling to maximize the complex stability radius of (1) over

K_1, K_2, K_3, K_4 for order $k = 1$ and $k = 2$. Figure 8 shows the pseudospectra of (1) when the complex stability radius is locally maximized over order–1 controllers. The locally optimal complex stability radius is approximately 9.98×10^{-5}, with approximate maximizer $[K_1\ K_2;\ K_3\ K_4] =$

$$\begin{bmatrix} 3.41e{+}00 & 5.16e{-}05 & 1.07e{-}03 \\ 1.86e{+}00 & 1.24e{-}04 & -9.06e{-}03 \\ -1.09e{-}01 & 2.12e{-}02 & -3.35e{-}02 \end{bmatrix}.$$

Figure 9 shows the pseudospectra of (1) when the complex stability radius is locally maximized over order–2 controllers. The locally optimal complex stability radius is approximately 1.02×10^{-4}, with approximate maximizer $[K_1\ K_2;\ K_3\ K_4] =$

$$\begin{bmatrix} 3.63e{+}00 & 4.99e{-}05 & 8.96e{-}05 & 9.01e{-}03 \\ 1.90e{+}00 & 1.14e{-}04 & -4.78e{-}03 & 1.15e{-}04 \\ -1.13e{-}01 & 2.06e{-}02 & -4.27e{-}02 & 1.17e{-}03 \\ 2.26e{-}03 & 3.51e{-}02 & -1.56e{-}03 & -1.04e{+}00 \end{bmatrix}.$$

In Figures 8 and 9, the 10^{-4}-pseudospectrum is very close to the imaginary axis, actually crossing it in the former but not in the latter.

As with the turbo-generator example, the fact that the gradient sampling algorithm approximately verifies first-order optimality conditions allows us to be reasonably confident that the matrices K yielding Figures 6, 7, 8 and 9 all approximate local optimizers of their respective objectives. However, we cannot conclude that they are global optimizers.

5. CONCLUSIONS

We have demonstrated that the gradient sampling algorithm (Burke *et al.*, 2003*b*) is a powerful and versatile tool for finding local optimizers of nonsmooth, nonconvex optimization problems that arise in robust stabilization. We have also demonstrated the usefulness of pseudospectra in formulating such optimization problems and visualizing the results. Our techniques are applicable to low-order controller design, a problem of great practical importance.

Acknowledgements. We particularly thank Friedemann Leibfritz for kindly providing us with many interesting test examples that he collected from the control literature, including the examples presented in this paper. We thank Bill Helton for bringing the importance of low-order controllers to our attention. We also thank Paul van Dooren, Alexandre Megretski, Carsten Scherer, André Tits and Nick Trefethen for their helpful responses to various questions that arose during this work.

REFERENCES

Blondel, V. and J.N. Tsitsiklis (1997). NP-hardness of some linear control design problems. *SIAM Journal on Control and Optimization* **35**, 2118–2127.

Blondel, V., M. Gevers and A. Lindquist (1995). Survey on the state of systems and control. *European Journal of Control* **1**, 5–23.

Boyd, S. and V. Balakrishnan (1990). A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm. *Systems and Control Letters* **15**, 1–7.

Boyd, S., L. El Ghaoui, E. Feron and V. Balakrishnan (1994). *Linear Matrix Inequalities in System and Control Theory.* SIAM. Philadelphia.

Bruinsma, N.A. and M. Steinbuch (1990). A fast algorithm to compute the H_∞-norm of a transfer function matrix. *Systems Control Letters* **14**, 287–293.

Burke, J.V., A.S. Lewis and M.L. Overton (2001). Optimal stability and eigenvalue multiplicity. *Foundations of Computational Mathematics* **1**, 205–225.

Burke, J.V., A.S. Lewis and M.L. Overton (2002). Two numerical methods for optimizing matrix stability. *Lin. Alg. Appl.* **351-352**, 117–145.

Burke, J.V., A.S. Lewis and M.L. Overton (2003*a*). Optimization and pseudospectra, with applications to robust stability. *SIAM Journal on Matrix Analysis and Applications.* To appear.

Burke, J.V., A.S. Lewis and M.L. Overton (2003*b*). A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. In preparation.

Burke, J.V., A.S. Lewis and M.L. Overton (2003*c*). Robust stability and a criss-cross algorithm for pseudospectra. *IMA J. Numer. Anal.* To appear.

Davison, E.J. (1990). *Benchmark problems for control system design: Report of the IFAC Theory Committee.* IFAC. Laxenberg.

Hinrichsen, D. and A.J. Pritchard (1986). Stability radii of linear systems. *Systems and Control Letters* **7**, 1–10.

Hung, Y. S. and A. G. J. MacFarlane (1982). *Multivariable feedback: A quasi–classical approach.* Springer–Verlag. Berlin, Heidelberg, New York. Lecture Notes in Control and Information Sciences.

Trefethen, L.N. (1990). Approximation theory and numerical linear algebra. In: *Algorithms for Approximation, II (Shrivenham, 1988)* (J. C. Mason and M. G. Cox, Eds.). Chapman and Hall. London. pp. 336–360.

Trefethen, L.N. (1997). Pseudospectra of linear operators. *SIAM Review* **39**, 383–406.