Nonsmooth Optimization for Multiband Frequency Domain Control Design

Pierre Apkarian * Dominikus Noll †

Abstract

Multiband frequency domain synthesis consists in the minimization of a finite family of closed-loop transfer functions on prescribed frequency intervals. This is an algorithmically difficult problem due to its inherent nonsmoothness and nonconvexity. We extend our previous work on nonsmooth H_∞ synthesis to develop a nonsmooth optimization technique to compute local solutions to multiband synthesis problems. The proposed method is shown to perform well on illustrative examples.

Keywords: H_∞-synthesis, multi-channel design, multi-objective optimization, multidisk problems, concurring performance specifications, static output feedback, reduced-order synthesis, decentralized control, PID, NP-hard problems, nonsmooth optimization.

1 Introduction

In this work we present a new algorithmic approach to multi frequency band feedback control synthesis. We consider simultaneous minimization of a finite family of closed-loop performance functions

$$f(K) = \max_{i=1,\dots,N} \|T_{w^i \to z^i}(K)\|_{I_i}, \tag{1}$$

where K stands for the feedback controller, $s \mapsto T_{w^i\to z^i}(K,s)$ is the ith closed-loop performance channel, and $\|T_{w^i\to z^i}(K)\|_{I_i}$ denotes the peak value of the transfer function max singular value norm on a prescribed frequency interval I_i:

$$\|T_{w^i\to z^i}(K)\|_{I_i} = \sup_{\omega\in I_i} \overline{\sigma}\left(T_{w^i\to z^i}(K, j\omega)\right).$$

Typically, I_i is a closed interval $I_i = [\omega_1^i,\, \omega_2^i]$, or more generally, a finite union of intervals:

$$I_i = [\omega_1^i,\, \omega_2^i] \cup \ldots \cup [\omega_{q_i}^i,\, \omega_{q_{i+1}}^i],$$

*ONERA-CERT, Centre d'études et de recherche de Toulouse, Control System Department, 2 av. Edouard Belin, 31055 Toulouse, France - and - Université Paul Sabatier, Institut de Mathématiques, Toulouse, France - Email: `apkarian@cert.fr` - Tel: +33 5.62.25.27.84 - Fax: +33 5.62.25.27.64.

†Université Paul Sabatier, Institut de Mathématiques, 118, route de Narbonne, 31062 Toulouse, France - Email: `noll@mip.ups-tlse.fr` - Tel: +33 5.61.55.86.22 - Fax: +33 5.61.55.83.85.

where right interval tips may take infinite values. For a single channel, $i = 1$ and $I_1 = [0, \infty]$, minimizing $f(K)$ subject to closed-loop stability reduces to a standard H_∞ synthesis problem.

The approach which we present for multiband synthesis was originally laid down in [2, 1, 3] for the standard H_∞ synthesis problem. It leads to efficient resolution algorithms, because a substantial part of the computations is carried out in the frequency domain, where the plant state dimension only mildly affects cputimes. Our method avoids the notorious difficulty of those approaches based on linear or bilinear matrix inequalities, where the presence of Lyapunov variables, whose number grows quadratically with the state-space dimension, quickly leads to large size optimization programs as systems get sizable. We have identified this as the major source of breakdown for most existing codes.

Multiband control design is of great practical interest mainly for two reasons:

- Very often practical design criteria are expressed as frequency domain constraints on limited frequency bands.
- In the traditional approach, weighting functions are used to specify frequency bands. But the search for suitable weighting functions is often a critical task and their use increases the controller order.

Despite its importance, only very few methods for multiband synthesis have been reported in the literature. In [15], the authors develop an extension of the Kalman-Yakubovich-Popov Lemma [17] in order to handle band restricted frequency domain constraints. The resulting problem is nonconvex even in the state-feedback case. The authors propose to solve it by forcing convexity, so that standard SDP solvers can be used. This may lead to a fairly conservative procedure.

There exist classical loop-shaping methods, like for instance the QFT method [14], which may be used to solve related synthesis problems. The QFT method exploits graphical tools and interfaces, but in order to work satisfactory, requires an advanced level of intuition. Moreover, such an approach is no longer suited if additional structural constraints on the controller have to be satisfied.

Similar comments could be made about synthesis methods based on the Youla parametrization, which generally lead to high-order controllers, see [9] and references therein. The idea of band restricted constraints in control design can be traced back to the classical Bode, Nyquist and Nichols plots to design simple structure controllers such as PID and phase-lag controllers [13, 4]. Unfortunately, these tools are mainly limited to single-input single-output systems, even though some multivariable generalizations have been attempted over the years [16].

We believe that this state-of-the-art shows an unexplored domain, which warrants a fresh investigation based on recent progress in optimization for synthesis. As we shall see, an approach based on (1) allows to take multiband constraints into account much more directly and more naturally.

It is important to notice that in contrast with H_∞ synthesis examined in [2] and multidisk synthesis [3], multi band design leads to an additional difficulty. Closed loop stability of the controller K has to be built into a mathematical programming constraint. We discuss and compare two possibilities how this could be done, and then propose a penalty and a barrier model suited for the numerical solution of these models.

The structure of the paper is as follows. In section 2, we provide a precise statement of the multiband frequency domain control design problem. Two different model algorithms with explicit stability constraints are discussed in sections 3 and 4. The necessary ingredients to implement a sequential penalty/barrier method are detailed in section 5. Numerical experiments are presented in section 6.

Notation

Let $\mathbb{R}^{n\times m}$ be the space of $n\times m$ matrices, equipped with the corresponding scalar product $\langle X,Y\rangle = \mathrm{Tr}(X^TY)$, where X^T is the transpose of the matrix X, $\mathrm{Tr}\,X$ its trace. For complex matrices, X^H denotes the transconjugate. For Hermitian or symmetric matrices, $X\succ Y$ means that $X-Y$ is positive definite, $X\succeq Y$ that $X-Y$ is positive semi-definite. For ease of notations, we define the following sets of Hermitian matrices: $\mathbb{B}_m := \{X\in\mathbb{S}_m : X\succeq 0,\ \mathrm{Tr}\,(X)=1\}$. Consider q-tuples of Hermitian matrices $(Y_1,\dots,Y_q)$, we define the set

$$\mathbb{B}_m^q := \left\{(Y_1,\dots,Y_q) : Y_i\in\mathbb{S}_m,\ Y_i\succeq 0,\ \sum_{i=1}^{q}\mathrm{Tr}\,(Y_i)=1\right\}.$$

For short, we shall use $\mathbb{B}$ or $\mathbb{B}^q$ when the dimension needs not be specified. We use the symbols $\lambda_{\max}$ and $\lambda_{\min}$ to denote the maximum and the minimum eigenvalue of a symmetric or Hermitian matrix and $\overline{\sigma}$ and $\underline{\sigma}$ to denote the maximum and the minimum singular value of a general matrix. The Frobenius norm of a matrix M is denoted $\|M\|_F$ and is defined by $\|M\|_F=\sqrt{\mathrm{Tr}\,(M^HM)}$. We shall use concepts from nonsmooth analysis covered by [11]. In particular, for a locally Lipschitz function $f:\mathbb{R}^n\to\mathbb{R}$, $\partial f(x)$ denotes its Clarke subdifferential or generalized gradient at x, $f'(x;d)$ the Clarke directional derivative. In the sequel of the paper, each $T_{w^i\to z^i}$ is a smooth operator defined on the open domain $\mathcal{D}\subset\mathbb{R}^{(m_2+k)\times(p_2+k)}$ of kth order stabilizing feedback controllers

$$K := \begin{bmatrix} A_K & B_K\\ C_K & D_K\end{bmatrix},\ A_K\in\mathbb{R}^{k\times k}$$

with values in the infinite dimensional space RH_∞ of rational stable transfer matrix functions.

2 Multiband frequency domain design

We consider a plant P in state-space form

$$P(s):\begin{bmatrix}\dot{x}\\ y\end{bmatrix}=\begin{bmatrix}A & B\\ C & D\end{bmatrix}\begin{bmatrix}x\\ u\end{bmatrix}$$

together with N concurring performance specifications, represented as a family of plants $P^i(s)$ described in state-space form as

$$P^i(s):\quad \begin{bmatrix}\dot{x}^i\\ z^i\\ y^i\end{bmatrix}=\begin{bmatrix}A^i & B_1^i & B_2^i\\ C_1^i & D_{11}^i & D_{12}^i\\ C_2^i & D_{21}^i & D_{22}^i\end{bmatrix}\begin{bmatrix}x^i\\ w^i\\ u^i\end{bmatrix},\ i=1,\dots,N, \tag{2}$$

where $x^i\in\mathbb{R}^{n^i}$ is the state vector of P^i, $u^i\in\mathbb{R}^{m_2}$ the vector of control inputs, $w^i\in\mathbb{R}^{m_1^i}$ the vector of exogenous inputs, $y^i\in\mathbb{R}^{p_2}$ the vector of measurements and $z^i\in\mathbb{R}^{p_1^i}$ the controlled or performance vector associated with the ith input w^i. The performance channels typically incorporate frequency filters which create new states, so that the matrices A^i contain the original system matrices A, etc. The difference with the usual multi-channel synthesis is that each $T_{w^i\to z^i}$ is only tested on a specific frequency band I_i. Without loss, it is assumed throughout that $D = 0$ and $D_{22}^i = 0$ for all i.

The multiband synthesis problem consists in designing a dynamic output feedback controller $u^i = K(s)y^i$ for the plant family (2) with the following properties:

- **Internal stability:** The controller K exponentially stabilizes the original plant P in closed-loop.
- **Performance:** Among all internally stabilizing controllers, K minimizes the worst case performance function $f(K) = \max_{i=1,\dots,N} \|T_{w^i\to z^i}(K)\|_{I_i}$.

We assume that the controller K has the following frequency domain representation:

$$K(s) = C_K(sI - A_K)^{-1}B_K + D_K, \qquad A_K \in \mathbb{R}^{k\times k}, \tag{3}$$

where k is the order of the controller, and where the case $k = 0$ of a static controller $K(s) = D_K$ is included.

Often practical considerations require additional structural constraints on the controller K, also referred to as control law specifications. For instance, it may be desired to design low-order controllers ($0 \le k \ll n_i$) or controllers with prescribed-pattern, sparse controllers, decentralized controllers, observed-based controllers, PID control structures, synthesis on a finite set of transfer functions, and much else. Formally, the synthesis problem may then be represented as

$$\begin{array}{ll} \text{minimize} & f(K) = \max_{i=1,\dots,N} \|T_{w^i\to z^i}(K)\|_{I_i} \\ \text{subject to} & K \text{ stabilizes } (A, B, C) \\ & K \in \mathcal{K} \end{array} \tag{4}$$

where $K \in \mathcal{K}$ represents a structural constraint on $K(s)$ in (3). In most cases, this takes the more amenable form of an equality constraint $g(K) = 0$, which as a rule can be eliminated explicitly.

Remark. A difficulty in program (4) is that the stability constraint $K \in \mathcal{D}$, where $\mathcal{D}$ is the set of stabilizing controllers, is not a constraint in the usual sense of mathematical programming, because the set $\mathcal{D}$ is open, and an element K on the boundary of this domain is *not* a valid solution. Since an optimization algorithm for (4) will converge to a solution on the boundary of $\mathcal{D}$, we have to modify this constraint in order to avoid numerical failure. Notice that a similar imbroglio arises in more familiar situations like in the bounded real lemma or the KYP-lemma, where we have to replace strict inequalities $\mathcal{B}(x) \prec 0$ by non strict inequalities $\mathcal{B}(x) \preceq -\varepsilon I$ for a suitable threshold $\varepsilon > 0$. The problem is then in all these cases to make the choice of such a threshold physically meaningful.

3 Model I: distance to instability

In this section we present a first systematic way to build a constraint which guarantees closed-loop stability. To simplify the presentation, we work with static controllers.

Let us start by introducing a stabilizing channel $s \mapsto T_{\text{stab}}(K, s) := (sI - (A + BKC))^{-1}$ for P, where P is the original plant without performance channels, defined by (A, B, C) and $D = 0$.

Indeed, K stabilizes P in closed-loop if and only if $T_{\text{stab}}(K)$ is exponentially stable. The stability domain $\mathcal{D}$ in (4) may then be written as

$$\mathcal{D} = \{K \in \mathbb{R}^{m_2 \times p_2} : \| (sI - (A + BKC))^{-1} \|_\infty < +\infty\},$$

where $\| \cdot \|_\infty$ is the H_∞ norm. We then replace $\mathcal{D}$ by the smaller closed set

$$\mathcal{D}_b = \{K \in \mathbb{R}^{m_2 \times p_2} : \| (sI - (A + BKC))^{-1} \|_\infty \leq b\},$$

where $b > 0$ is some large constant. The following mathematical program may then be considered:

$$\begin{array}{ll} \text{minimize} & f(K) = \max\limits_{i=1,\dots,N} \|T_{w^i \to z^i}(K)\|_{I_i} \\ \text{subject to} & g(K) := \|T_{\text{stab}}(K)\|_\infty \leq b \\ & K \in \mathcal{K}. \end{array} \tag{5}$$

When the parameter b is relevant, we refer to this program as $(5)_b$, or as the large constant program.

Is there a natural way to fix the numerical value b? To answer this question we consider the spectral abscissa of a system matrix A,

$$\alpha(A) = \max\{\operatorname{Re}\lambda : \lambda \text{ eigenvalue of } A\},$$

also known as the stability degree of the transfer matrix $H(s) = (sI - A)^{-1}$, [8, p. 138]. For a threshold $\varepsilon > 0$ the pseudo-spectral abscissa is defined as

$$\alpha_\varepsilon(A) = \max\{\operatorname{Re}\lambda : \lambda \text{ eigenvalue of some } A' \text{ with } \|A - A'\|_F \leq \varepsilon\}.$$

While stability of a system A is equivalent to $\alpha(A) < 0$, we can interpret $\alpha_\varepsilon(A) \leq 0$ as some robust form of stability, a point of view put forward in [18] and [10]. In particular, for every $\varepsilon > 0$, $\alpha_\varepsilon(A) \leq 0$ implies $\alpha(A) < 0$.

In order to decide how to choose $\varepsilon > 0$, we consider the distance to instability of a matrix

$$\beta(A) = \inf\{\|X\|_F : A + X \text{ instable}\}.$$

It is easy to see that

$$\alpha_\varepsilon(A) \leq 0 \quad \Leftrightarrow \quad \beta(A) \geq \varepsilon \quad \Leftrightarrow \quad \|(sI - A)^{-1}\|_\infty \leq \frac{1}{\varepsilon}.$$

That means, our natural choice is $b = 1/\beta$, where β is the smallest distance to instability which we grant the closed loop system $A + BKC$.

How do we solve program (5) numerically? This will be explained during the remainder of this section. We introduce the optimization program

$$\begin{aligned} &\min_{K \in \mathbb{R}^{m_2 \times p_2}} \max \left\{ \max_{i=1,\dots,N} \|T_{w^i \to z^i}(K)\|_{I_i}, \mu \|T_{\text{stab}}(K)\|_\infty \right\} \\ &= \min_{K \in \mathbb{R}^{m_2 \times p_2}} \max\{f(K), \mu\, g(K)\} =: \min_{K \in \mathbb{R}^{m_2 \times p_2}} f_\mu(K) \end{aligned} \tag{6}$$

subject to the constraint $K \in \mathcal{K}$. We refer to $(6)_\mu$ as the penalty program, to $\mu > 0$ as the penalty parameter, and to f_μ as the penalty function. During the following, to simplify the discussion, we will assume that the structural constraint has the explicit form $K = \Phi(\kappa)$, so that it may be eliminated. This is usually the case in practice.

LEMMA 1 *Let K_μ be a local minimum of $(6)_\mu$ which is non degenerate in the sense that it is neither a local minimum of f alone, nor a local minimum of g alone. Then K_μ is a Karush-Kuhn-Tucker point of program $(5)_{b(\mu)}$ with $b(\mu) = g(K_\mu) = f(K_\mu)/\mu$.*

Proof. Given the fact that K_μ is non degenerate, the necessary optimality conditions for program $(6)_\mu$ are as follows: There exists $0 < t_\mu < 1$ such that

$$0 \in t_\mu \partial f(K_\mu) + (1 - t_\mu)\mu \partial g(K_\mu) \quad \text{and} \quad f(K_\mu) = \mu g(K_\mu).$$

Let us now write the Karush-Kuhn-Tucker conditions for $(5)_b$: There exists a Lagrange multiplier λ_b such that

$$0 \in \partial f(K_b) + \lambda_b \partial g(K_b), \;\; g(K_b) - b \leq 0, \;\; \lambda_b \geq 0, \;\; \lambda_b(g(K_b) - b) = 0.$$

We then see that the solution K_μ of $(6)_\mu$ solves $(5)_b$ if we set

$$b(\mu) = g(K_\mu), \quad \lambda_{b(\mu)} = \frac{(1 - t_\mu)\mu}{t_\mu}.$$

This proves the claim. □

This result has a natural converse.

LEMMA 2 *Let K_b be a local minimum of program $(5)_b$ which is non degenerate in the sense that it is not a Karush-Kuhn-Tucker point of f alone. Then K_b is a critical point of program $(6)_{\mu(b)}$ with $\mu(b) = \frac{f(K_b)}{g(K_b)}$.*

Proof. We compare once again the necessary optimality conditions. Reading the formulas backwards, we first find that

$$\mu(b) = \frac{f(K_b)}{g(K_b)}.$$

Then reading $\lambda_b = \frac{(1-t_\mu)\mu}{t_\mu}$ backwards leads to

$$t_{\mu(b)} = \frac{\mu(b)}{\lambda_b + \mu(b)} = \frac{f(K_b)}{f(K_b) + \lambda_b g(K_b)} \in (0, 1).$$

□

Remark 3.1 Picking the correct local minimum in each program, we see that there is now at least locally a one-to-one correspondence between both programs, $(5)_b$ and $(6)_\mu$, in the sense that the functions $\mu \mapsto b(\mu)$ and $b \mapsto \mu(b)$ are inverse to each other, and $K_b = K_{\mu(b)}$, $K_\mu = K_{b(\mu)}$. In order to find the solution K_b for the parameter $b = \beta^{-1}$ chosen above, we therefore have to find $\mu(b) = \mu(\beta^{-1})$ and solve the corresponding penalty program $(6)_{\mu(\beta^{-1})}$.

4 Model II: shifting poles

Let us now consider a second possibility to fix a closed subset of $\mathcal{D}$, this time based on the shifted H_∞ norm, [8, p. 100]:

$$\|H(\cdot)\|_{\infty,\alpha} = \|H(\cdot + \alpha)\|_\infty .$$

For $\alpha < 0$, the condition $\|H\|_{\infty,\alpha} < +\infty$ guarantees that the poles of $H(s)$ lie to the left of the line $\mathrm{Re}\, s = \alpha < 0$. That means that for every $\alpha < 0$, the closure of the open domain

$$\mathcal{D}^\alpha = \{K \in \mathbb{R}^{m_2 \times p_2} : \| \left(sI - (A + BKC)\right)^{-1} \|_{\infty,\alpha} < +\infty\}$$

is a possible candidate for a mathematically tractable domain, because $\overline{\mathcal{D}}^\alpha \subset \mathcal{D}$. Indeed, elements K on the boundary of $\mathcal{D}^\alpha$ still have $\mathrm{Re}\, \lambda \leq \alpha < 0$ for the poles λ of $A + BKC$, hence these K are closed-loop stabilizing. In consequence, we consider the following optimization program

$$\begin{array}{ll} \text{minimize} & f(K) = \max\limits_{i=1,\dots,N} \|T_{w^i \to z^i}(K)\|_{I_i} \\ \text{subject to} & K \in \overline{\mathcal{D}}^\alpha \\ & K \in \mathcal{K} \end{array} \tag{7}$$

which we denote as $(7)^\alpha$ if the parameter $\alpha < 0$ matters.

Having prepared the rationale, let us now discuss an algorithmic approach to program (7). The situation is slightly more complicated than in the penalty case, because the feasible domain is not easily represented as a constraint set in the usual sense of nonlinear programming. What we have, though, is a barrier function for the domain $\mathcal{D}^\alpha$. Putting

$$h_\alpha(K) = \|T_{\mathrm{stab}}(K)\|_{\infty,\alpha},$$

where $T_{\mathrm{stab}}(s) := (sI - (A + BKC)^{-1}$ is the stabilizing channel for plant P, we see that

$$\mathcal{D}^\alpha = \{K \in \mathbb{R}^{m_2 \times p_2} : h_\alpha(K) < +\infty\}.$$

We may then consider the following family of programs

$$\min_{K \in \mathbb{R}^{m_2 \times p_2}} \max\{f(K), \mu\, h_\alpha(K)\} =: \min_{K \in \mathbb{R}^{m_2 \times p_2}} f_{\mu,\alpha}(K), \tag{8}$$

subject to the constraint $K \in \mathcal{K}$, denoted as $(8)^{\mu,\alpha}$ when a reference to $\mu > 0$ and $\alpha < 0$ is made. We refer to $f_{\mu,a}(K)$ as the barrier function.

We now have the following result, relating (7) and (8).

Lemma 3 *Let $K^{\mu,\alpha}$ be a local minimum of $(8)^{\mu,\alpha}$ which is non degenerate in the sense that it is neither a critical point of f alone, nor a critical point of h_α alone. Let K^α be an accumulation point of the sequence $K^{\mu,\alpha}$ as $\mu \to 0$. Suppose $\min_{\omega \in \mathbb{R}} \underline{\sigma}(j\omega I - (ABK^a C - \alpha I))$ is attained on a finite set of frequencies. Then K^α is a critical point of program $(7)^\alpha$.*

Proof. 1) Let us start by writing down the necessary optimality condition for program $(7)^\alpha$ at K^α. It says that there exists a subgradient $G \in \partial f(K^\alpha)$ such that $-G$ is in the Clarke normal cone $N_{\overline{\mathcal{D}}^\alpha}(K^\alpha)$ of the set $\overline{\mathcal{D}}^\alpha$ at K^α. See e.g. [6] for this.

2) Let us next write the necessary optimality conditions for program $(8)^{\mu,\alpha}$. They say that there exists $0 < t_{\mu,\alpha} < 1$ such that

$$0 \in t_{\mu,\alpha}\partial f(K^{\mu,\alpha}) + (1-t_{\mu,\alpha})\mu\,\partial h_\alpha(K^{\mu,\alpha}), \qquad f(K^{\mu,\alpha}) = \mu\, h_\alpha(K^{\mu,\alpha}).$$

We introduce the level sets

$$\mathcal{D}^\alpha(\mu) = \{K \in \mathbb{R}^{m_2\times p_2} : h_\alpha(K) \le h_\alpha(K^{\mu,\alpha})\}.$$

There are now two possibilities. Either $h_\alpha(K^{\mu,\alpha}) \to \infty$ as $\mu \to 0$, or there exists a subsequence for which these values are bounded. In the latter case, from the right hand equation above, we then have $f(K^{\mu,\alpha}) \to 0$, hence $f(K^\alpha) = 0$. This case is clearly exceptional, because here K^α is a global minimum of f alone.

3) Let us now assume that $h_\alpha(K^{\mu,\alpha}) \to \infty$, so that the domains $\mathcal{D}^\alpha(\mu)$ grow as $\mu \to 0$. In fact, we now have $\cup_{\mu>0}\mathcal{D}^\alpha(\mu) = \mathcal{D}^\alpha$. ¿From the left hand condition above we see that there exists a subgradient $G_{\mu,\alpha} \in \partial f(K^{\mu,\alpha})$ such that

$$-\frac{(1-t_{\mu,\alpha})\mu}{t_{\mu,\alpha}}\, G_{\mu,\alpha} \in \partial h_\alpha(K^{\mu,\alpha}).$$

In other words, this negative subgradient of f at $K^{\mu,\alpha}$ is a direction in the normal cone $N_{\mathcal{D}^\alpha(\mu)}(K^{\mu,\alpha})$ to the level set $\mathcal{D}^\alpha(\mu)$ at $K^{\mu,\alpha}$. Passing to a subsequence, we may assume that $G_{\mu,\alpha} \to G_\alpha$, and therefore upper semicontinuity of the Clarke subdifferential [11] gives $G_\alpha \in \partial f(K^\alpha)$. We now wish to show that G_α is in the normal cone $N_{\overline{\mathcal{D}}^\alpha}(K^\alpha)$, because then the necessary optimality condition in step 1) is satisfied.

4) Let us introduce the following function

$$\phi_\alpha(K) = \begin{cases} -h_\alpha(K)^{-2} & \text{if } h_\alpha(K) < \infty \\ 0 & \text{else} \end{cases}$$

Then $\mathcal{D}^\alpha = \{K : \phi_\alpha(K) < 0\}$, and $\mathcal{D}^\alpha(\mu) = \{K : \phi_\alpha(K) \le -\frac{1}{h_\alpha(K^{\mu,\alpha})^2}\}$. Notice, however, that $\overline{\mathcal{D}}^\alpha \ne \{K : \phi_\alpha(K) \le 0\}$. That means, we cannot directly conclude with the help of the upper semicontinuity of the Clarke subdifferential of ϕ_α, as we did above for the subdifferential of f. This is what complicates this proof.

Let us show that ϕ_α is locally Lipschitz. Since h_α is locally Lipschitz, this is certainly true in the set $\mathcal{D}^\alpha$. Only points K on the boundary might cause problems. But

$$\begin{aligned}\phi_\alpha(K) = -h_\alpha(K)^{-2} &= -\frac{1}{\max_{\omega\in\mathbb{R}} \overline{\sigma}\left((j\omega I - A - BKC + \alpha I)^{-1}\right)^2} \\ &= -\min_{\omega\in\mathbb{R}} \underline{\sigma}\,(j\omega I - A - BKC + \alpha I)^2 \\ &= -\min_{\omega\in\mathbb{R}} \lambda_{\min}\left([(j\omega+\alpha)I - (A+BKC)][(j\omega+\alpha)I - (A+BKC)]^H\right)\end{aligned}$$

and this is a locally Lipschitz function, because for fixed ω, the minimum eigenvalue of a symmetric matrix function is locally Lipschitz. The min operator does not alter this. This representation also shows that ϕ_α has value 0 even outside the set $\overline{\mathcal{D}}^\alpha$, which is therefore not the level set of ϕ_α at level 0.

Now we use again upper semi-continuity of the Clarke subdifferential. We have

$$\limsup_{\mu\to 0} \partial\phi_\alpha(K^{\mu,\alpha}) \subset \partial\phi_\alpha(K^\alpha).$$

This implies

$$\limsup_{\mu\to 0} N_{\mathcal{D}^\alpha(\mu)}(K^{\mu,\alpha}) \subset \Lambda_\alpha(K^\alpha),$$

where $\Lambda_\alpha(K)$ is the convex cone generated by the compact convex set $\partial\phi_\alpha(K)$, because the normal cone to $\mathcal{D}^\alpha(\mu)$ is generated by the subdifferential of ϕ_α at $K^{\mu,\alpha}$. Recall the difficulty: our proof is not finished because $\Lambda_\alpha(K^\alpha)$ is *not* identical with the Clarke normal cone $N_{\overline{\mathcal{D}}^\alpha}(K^\alpha)$ to $\overline{\mathcal{D}}^\alpha$ at K^α.

Let us show that $\Lambda_\alpha(K^\alpha)$ is pointed, that is, $\Lambda_\alpha(K^\alpha) \cap -\Lambda_\alpha(K^\alpha) = \{0\}$. This follows as soon as we show that $\pm G \in \partial\phi_\alpha(K^\alpha)$ implies $G = 0$.

By hypothesis, the minimum singular value at K^α is attained on a finite set of frequencies. This implies that ϕ_α is Clarke regular at K^α. Hence the Clarke directional derivative coincides with the Dini directional derivative. That means

$$\partial\phi_\alpha(K^\alpha) = \{G : \forall D\ \langle G, D\rangle \le \phi'_\alpha(K^\alpha; D) = \liminf_{t\to 0^+} t^{-1}\left(\phi_\alpha(K^\alpha + tD) - \phi_\alpha(K^\alpha)\right)\}.$$

But $\phi_\alpha(K^\alpha) = 0$ so for fixed $\varepsilon > 0$ we can find $t_\varepsilon > 0$ such that $\langle G, D\rangle \le t_\varepsilon^{-1}\phi_\alpha(K^\alpha + t_\varepsilon D) + \varepsilon \le \varepsilon$, the latter since $\phi_\alpha \le 0$. We have shown $\langle G, D\rangle \le \varepsilon$, and since ε was arbitrary, we have $\langle G, D\rangle \le 0$. Now we use the fact that $-G$ is also a subgradient. Repeating the argument then shows $-\langle G, D\rangle \le 0$. Altogether, $\langle G, D\rangle = 0$, and since D was arbitrary, this gives $G = 0$.

5) Having shown that $\Lambda_\alpha(K^\alpha)$ is pointed, it follows that the convex hull of $\limsup_{\mu\to 0} N_{\mathcal{D}^\alpha(\mu)}(K^{\mu,\alpha})$ is pointed, because by 4) it is contained in $\Lambda_\alpha(K^\alpha)$. Now we use Proposition 4.1 and Theorem 2.3 in [12] to deduce that $\limsup_{\mu\to 0} N_{\mathcal{D}^\alpha(\mu)}(K^{\mu,\alpha}) \subset N_{\overline{\mathcal{D}}^\alpha}(K^\alpha)$. In the terminology of that paper, this property is referred to as normal convergence. □

Remark 4.1 1) Note that the above reasoning and formulas carry over to dynamic controllers if a standard dynamic augmentation of the plant is performed. See [2] for details.

2) Notice that the barrier function in (7) and the penalty function in (5) have almost identical structure. This means that the algorithmic approach discussed in the following sections applies to both programs with minor adaption.

3) Normal convergence of sets as defined in [12] is a suitable concept to describe approximation of mathematical programs. If the constraint set is repesented as the level set of a locally Lipschitz operator, normal convergence is satisfied. However, in our case, the limiting set $\overline{\mathcal{D}}^\alpha$ is not a level set, which complicates the situation. Academic counterexamples where normal convergence fails can be constructed; see [12].

5 Algorithms for multiband control design

In the sequel, we describe the different ingredients for solving the multiband control design problem with algorithm models I and II of sections 3 and 4, respectively.

5.1 Subdifferential of the barrier function

At the core of our algorithm is the computation of the Clarke subdifferential of the barrier function f_μ and the penalty function $f_{\mu,\alpha}$. Precursors of the results here were obtained in [2] for the H_∞ norm. Below we will therefore focus on what is new.

In order to unify the presentation, we introduce a common terminology for both cases. For the penalty function f_μ at fixed $\mu > 0$, we introduce a new closed-loop transfer channel:

$$T_{w^{N+1}\to z^{N+1}}(K) = \mu\, T_{\text{stab}}(K),$$

so that $f_\mu(K) = \max_{i=1,\dots,N+1} \|T_{w^i\to z^i}(K)\|_{I_i}$ when we set $I_{N+1} = [0,\infty]$. Similarly, for the barrier function $f_{\mu,\alpha}$ at fixed $\mu > 0$, $\alpha < 0$, we introduce the $(N+1)$st channel in the form

$$T_{w^{N+1}\to z^{N+1}}(K,s) = \mu\, T_{\text{stab}}(K, s+\alpha),$$

so that again $f_{\mu,\alpha}(K) = \max_{i=1,\dots,N+1} \|T_{w^i\to z^i}(K)\|_{I_i}$ with $I_{N+1} = [0,\infty]$.

Noting that the formulas for the Clarke subdifferential of $f_{\mu,\alpha}$ are easily inferred from those of f_μ, we will restrict the discussion to f_μ from now on. Assuming a static controller, $k = 0$, we introduce the simplifying closed-loop notation in state space

$$\begin{gathered}\mathcal{A}^i(K) := A^i + B_2^i K C_2^i,\ \mathcal{B}^i(K) := B_1^i + B_2^i K D_{21}^i,\ \mathcal{C}^i(K) := C_1^i + D_{12}^i K C_2^i,\\ \mathcal{D}^i(K) := D_{11}^i + D_{12}^i K D_{21}^i\,,\end{gathered} \tag{9}$$

and in frequency domain

$$\begin{bmatrix} T_{w^i\to z^i}(K,s) & G_{12}^i(K,s) \\ G_{21}^i(K,s) & \star \end{bmatrix} := \begin{bmatrix} \mathcal{C}^i(K) \\ C_2^i \end{bmatrix} (sI - \mathcal{A}^i(K))^{-1} \begin{bmatrix} \mathcal{B}^i(K) & B_2^i \end{bmatrix} + \begin{bmatrix} \mathcal{D}^i(K) & D_{12}^i \\ D_{21}^i & \star \end{bmatrix}.$$

Here, for $i = N+1$, we define the plant

$$P^{N+1}(s): \quad \begin{bmatrix} \dot{x}^{N+1} \\ z^{N+1} \\ y^{N+1} \end{bmatrix} = \begin{bmatrix} A & I & B \\ I & 0 & 0 \\ C & 0 & 0 \end{bmatrix} \begin{bmatrix} x^{N+1} \\ w^{N+1} \\ u^{N+1} \end{bmatrix}, i = 1,\dots,N, \tag{10}$$

where $x^{N+1} \in \mathbb{R}^n$, n is the dimension of A, $u^{N+1} \in \mathbb{R}^{m_2}$, $w^{N+1} \in \mathbb{R}^n$, $y^{N+1} \in \mathbb{R}^{p_2}$, and $z^{N+1} \in \mathbb{R}^n$.

Let us now introduce the notion of active frequencies. For a given controller K, active channels or specifications are obtained through the index set

$$I_\mu(K) := \{i \in \{1,\dots,N+1\} : \|T_{w^i\to z^i}(K)\|_{I_i} = f_\mu(K)\}, \tag{11}$$

Moreover, for each $i \in I_\mu(K)$, we consider the set of active frequencies

$$\Omega_\mu^i(K) = \{\omega \in I_i : \overline{\sigma}\left(T_{w^i\to z^i}(K, j\omega)\right) = f_\mu(K)\}.$$

We assume throughout that $\Omega_\mu^i(K)$ is a finite set, indexed as

$$\Omega_\mu^i(K) = \{\omega_\nu^i : \nu = 1,\dots,p^i\},\ i \in I_\mu(K). \tag{12}$$

The set of all active frequencies is denoted as $\Omega_\mu(K)$. Armed with these definitions, we have the following:

Theorem 5.1 *Assume that the controller K is static, $k = 0$, and stabilizes the basic plant P^{N+1} in (10), that is, $K \in \mathcal{D}$. With the notations introduced in (11) and (12), let Q_ν^i be a matrix whose columns form an orthonormal basis of the eigenspace of $T_{w^i\to z^i}(K, j\omega_\nu^i)T_{w^i\to z^i}(K, j\omega_\nu^i)^H$ associated with the largest eigenvalue $\lambda_1\left(T_{w^i\to z^i}(K, j\omega_\nu^i)T_{w^i\to z^i}(K, j\omega_\nu^i)^H\right) = \overline{\sigma}(T_{w^i\to z^i}(K, j\omega_\nu^i))^2$. Then, the Clarke subdifferential of the mapping f_μ at $K \in \mathcal{D}$ is the compact and convex set $\partial f_\mu(K) = \{\Phi_Y : Y \in \mathbb{B}^p\}$, with $p := \sum_{i\in I_\mu(K)} p^i$ and where*

$$\Phi_Y = f_\mu(K)^{-1} \sum_{i\in I_\mu(K)} \sum_{\nu=1,\dots,p^i} \text{Re}\left\{G_{21}^i(K, j\omega_\nu^i)\, T_{w^i\to z^i}(K, j\omega_\nu^i)^H Q_\nu^i Y_\nu^i (Q_\nu^i)^H G_{12}^i(K, j\omega_\nu^i)\right\}^T . \quad (13)$$

The formula also applies to $f_{\mu,\alpha}$ when suitably adapted.

Proof: The proof is based on the representation of the Clarke subdifferential of finite max functions [11] and is omitted for brevity. The reader is referred to [2, 3, 5] for related cases. ■

5.2 Solving the subproblem

In this section we describe an extension of the nonsmooth technique originally developed in [2, 1] for H_∞ synthesis, and in [3] for multidisk problems. The method is convergent and has been tested on a variety of sizable problems.

As before, we consider minimization of f_μ for fixed μ, and minimization of $f_{\mu,\alpha}$ for fixed μ, α. This is referred to as solving the subproblem. Writing

$$f_\mu(K, \omega) := \max_{i=1,\dots,N+1} \{\overline{\sigma}\left(T_{w^i\to z^i}(K, j\omega)\right) : \omega \in I_i\} ,$$

we see that $f_\mu(K) = \max_{\omega\in\mathbb{R}} f_\mu(K, \omega)$, so that minimization of f_μ may be interpreted as a semi-infinite minimization problem involving the infinite family $f_\mu(\cdot, \omega)$. At given K, recall that $\Omega_\mu(K)$ is the set of active frequencies at K. Clearly, $f_\mu(K, \omega) \le f_\mu(K)$ for all $\omega \in \mathbb{R}$ and $f_\mu(K, \omega) = f_\mu(K)$ for $\omega \in \Omega_\mu(K)$. As a consequence of Theorem 5.1, the subdifferential of the function $f_\mu(K, \omega)$ at K is the set of subgradients

$$\Phi_{Y,\omega} := f_\mu(K, \omega)^{-1} \sum_{i\in I_\omega(K)} \text{Re}\left\{G_{21}^i(K, j\omega)\, T_{w^i\to z^i}(K, j\omega)^H Q_\omega^i Y_\omega^i (Q_\omega^i)^H G_{12}^i(K, j\omega)\right\}^T ,$$

where

$$I_\omega(K) = \{i \in \{1, \dots, N+1\} : \omega \in I_i,\ \overline{\sigma}(T_{w^i\to z^i}(K, j\omega)) = f_\mu(K, \omega)\}$$

is the index set of active models at frequency ω. Here the columns of the matrix Q_ω^i form an orthonormal basis of the eigenspace of $T_{w^i\to z^i}(K, j\omega)T_{w^i\to z^i}(K, j\omega)^H$ associated with its largest eigenvalue, and

$$\sum_{i\in I_\omega(K)} \text{Tr}\, Y_\omega^i = 1\,, \quad Y_\omega^i = (Y_\omega^i)^H \succeq 0\,.$$

Let us consider finite extensions $\Omega_{e,\mu}(K)$ of the set of active frequencies $\Omega_\mu(K)$. For any such $\Omega_{e,\mu}(K)$, and for fixed $\delta > 0$, we introduce the optimality function

$$\theta_{e,\mu}(K) := \inf_{H\in\mathbb{R}^{m_2\times p_2}} \sup_{\omega\in\Omega_{e,\mu}(K)} \sup_{\sum_{i\in I_\omega(K)} \text{Tr}\, Y_\omega^i=1,\, Y_\omega^i\succeq 0} -f_\mu(K) + f_\mu(K, \omega) + \langle \Phi_{Y,\omega}, H\rangle + \frac{1}{2}\delta\|H\|_F^2 . \quad (14)$$

When $\Omega_{e,\mu}(K) = \Omega_\mu(K)$, we use the notation $\theta_\mu(K)$. Since $\Omega_\mu(K) \subset \Omega_{e,\mu}(K)$, we have $\theta_\mu(K) \leq \theta_{e,\mu}(K)$ for any of these extensions.

We refer to $\theta_\mu(K)$ and $\theta_{e,\mu}(K)$ as optimality functions, because they share the following property: $\theta_{e,\mu}(K) \leq 0$ for all K, and $\theta_{e,\mu}(K) = 0$ implies that K is a critical point of f_μ [2]. In consequence, optimality functions can be used to generate descent steps. In order to do this, we show that optimality function (14) has a tractable dual form.

Proposition 5.2 *The dual formula for $\theta_{e,\mu}(K)$ is:*

$$\theta_{e,\mu}(K) = \sup_{\sum_{\omega\in\Omega_{e,\mu}(K)} \tau_\omega = 1,\, \tau_\omega \geq 0} \; \sup_{\sum_{i\in I_\omega(K)} \operatorname{Tr} Y_\omega^i = 1,\, Y_\omega^i \succeq 0} \; \sum_{\omega\in\Omega_{e,\mu}(K)} \tau_\omega (f_\mu(K,\omega) - f_\mu(K)) - \frac{1}{2\delta} \| \sum_{\omega\in\Omega_{e,\mu}(K)} \tau_\omega \Phi_{Y,\omega} \|_F^2 \,. \tag{15}$$

The associated optimal descent direction in the controller space is given as

$$H(K) := -\frac{1}{\delta} \sum_{\omega\in\Omega_{e,\mu}(K)} \tau_\omega \Phi_{Y,\omega} \,. \tag{16}$$

Proof. The proof is essentially covered by the results in [3] and is omitted for brevity. ■

Remark 5.3 The appealing feature of program (15) is that it is a small size SDP, or even a convex QP when singular values are simple, which appears to be the rule.

It is also worth noting here that band restricted norms $\|\cdot\|_{I_i}$ and peak frequencies $\omega \in \Omega_\mu(K)$ are easily computed via the bisection algorithm in [7]. We only have to confine the search for peak frequencies to the intervals I_i for $i = 1, \ldots, N+1$.

Proposition 5.2 suggests the following descent scheme for solving the subproblems for given K and μ respectively μ, α .

<u>Nonsmooth descent algorithm for the subproblem</u>

	Fix $\delta > 0$, $0 < \vartheta < 1$, $0 < \rho < 1$.
1.	**Initialization**. Find a controller K which stabilizes the plant P.
2.	**Generate frequencies**. Given the current K, compute $f_\mu(K)$ and obtain active frequencies $\Omega_\mu(K)$. Select a finite enriched set of frequencies $\Omega_{e,\mu}(K)$ containing $\Omega_\mu(K)$.
3.	**Descent direction.** Compute $\theta_{e,\mu}(K)$ and the solution (τ, Y) of SDP or convex QP (15). If $\theta_{e,\mu}(K) = 0$, stop, because $0 \in \partial f_\mu(K)$. Otherwise compute descent direction $H(K)$ given in (16).
4.	**Line search.** Find largest $t = \vartheta^k$ such that $f_\mu(K + t\,H(K)) \leq f_\mu(K) + t\rho\theta_\mu(K)$ and such that $K + tH(K)$ remains stabilizing.
5.	**Step.** Replace K by $K + tH(K)$, increase iteration counter by one, and go back to step 2.

As shown in [2, 3], for fixed μ the descent scheme is guaranteed to converge to a critical point K_μ with $0 \in \partial f_\mu(K_\mu)$, starting from arbitrary $K \in \mathcal{D}$. However, the subproblem becomes increasingly ill-conditioned as μ gets smaller. This results in a large number of iterations and may even lead to failure to reach criticality. Experience with barrier algorithms suggests to start each subproblem with a good approximation of a local solution. This is examined in the next section.

5.3 Enriched sets of frequencies

Choosing an extended set of frequencies $\Omega_{e,\mu}$ in step 2 is a key ingredient for the success of our technique and is beneficial mainly for two reasons:

- It renders the algorithm less dependent on the accuracy within which peak frequencies in Ω_μ are computed. A consequence is that the computed search direction behaves more smoothly.
- It captures more information on the frequency responses $\omega \mapsto \overline{\sigma}\left(T_{w^i\to z^i}(K, j\omega)\right)$ on their associated intervals I_i. This leads to better step lengths.

In our numerical testing, we have used the following simple scheme to compute an enlarged set of frequencies:

<u>Construction of enriched sets of frequencies</u>

Fix $0 < \eta < 1$.

1. Compute $f_\mu(K)$ using a band restricted version of the bisection algorithm [7] applied to each channel $T_{w^i\to z^i}$ and obtain $\Omega_\mu(K)$.
2. Define a cut-off level $\gamma_c := \eta\, f_\mu(K)$.
3. Determine nearly active channels using γ_c. Channel with index i is retained for frequency griding whenever $\|T_{w^i\to z^i}(K)\|_{I_i} \geq \gamma_c$.
4. For each nearly active model i, grid frequency subintervals of I_i where $\overline{\sigma}(T_{w^i\to z^i}(K, j\omega)) \geq \gamma_c$ and add peak frequencies of $f_\mu(K)$ to assure that $\Omega_{e,\mu}(K)$ contains $\Omega_\mu(K)$.
5. If $\Omega_{e,\mu}(K)$ is too large, truncate to retain the first F frequencies with leading singular values.

More sophisticated versions are possible [3], but the above simple scheme was efficient in a number of experiments. Typical values for η used in the experimental section are $\eta = 0.8$ or $\eta = 0.9$. The number of retained frequencies is between $F = 30$ and $F = 300$, which allows to keep control of the computational load to generate search directions.

5.4 Combined algorithm

We have to assemble the elements of the previous sections into a combined algorithm. Here a difference between models (5) and (7) occurs. Namely, in the first case we have to drive μ to the specific value where $b(\mu) = \beta^{-1}$, where $\beta > 0$ is our prior threshold for the distance to instability. On the other hand, in model (7), we fix the threshold $\alpha < 0$ for the poles λ of the closed loop systems, that is $\operatorname{Re}\lambda \leq \alpha < 0$, but have to drive μ to zero. In both cases, however, we will use a homotopy approach and start with a moderate size μ to solve $(6)_\mu$ respectively $(7)^{\mu,\alpha}$. Then we will update μ to μ^+ and use the solution $K^{\mu,\alpha}$ as initial point for the next solution of the subproblem. Different strategies to steer the parameter μ are discussed in the experimental section. What will also have to be discussed is to what precision the early subproblems need to be solved, and how a successive refinement should be organized.

6 Numerical experiments

As a simple example, we consider the double integrator $G(s) = \frac{1}{s^2}$ which is one of the most fundamental plants in control applications. Design specifications in the form of multiband constraints are borrowed from [19] and involve the sensitivity function $S := (I + GK)^{-1}$ and the complementary sensitivity function $T := GK(I + GK)^{-1}$. Multiband constraints are as follows

- disturbance rejection and tracking

$$|S(j\omega)| \leq 0.85, \text{ for } \omega \in I_1 := [0, 0.5] \,.$$

- gain-phase margins

$$|S(j\omega)| \leq 1.30, \text{ for } \omega \in I_2 := [0.5, 2] \,.$$

- bandwidth

$$|T(j\omega)| \leq 0.707, \text{ for } \omega \in I_3 := [2, 4] \,.$$

- roll-off

$$|w(j\omega)\, T(j\omega)| \leq 1.0, \text{ for } \omega \in I_4 := [4, \infty] \,,$$

  where $w(s)$ is the weighing function

$$w(s) := \frac{0.2634s^2 + 1.659s + 5.333}{0.0001s^2 + 0.014s + 1} \,.$$

This problem is cast as a multiband H_∞ synthesis problem in the form (4):

$$\begin{array}{ll} \text{minimize} & f(K) := \max\left\{ \frac{1}{0.85}\|S\|_{I_1}, \frac{1}{1.30}\|S\|_{I_2}, \frac{1}{0.707}\|T\|_{I_3}, \|w(s)T\|_{I_4} \right\} \\ \text{subject to} & K \text{ stabilizes } G(s) \end{array}$$

As explained in section 2, the stability constraint could be represented either as distance to instability constraint, using the penalty function:

$$\text{minimize } f_\mu(K) := \max\left\{ \frac{1}{0.85}\|S\|_{I_1}, \frac{1}{1.30}\|S\|_{I_2}, \frac{1}{0.707}\|T\|_{I_3}, \|w(s)T\|_{I_4}, \mu\|T_{\text{stab}}\|_\infty \right\}$$

(model algorithm I) where μ is the penalty parameter, and where $T_{\text{stab}}(K, s) = (sI - (A + BKC))^{-1}$ is the stabilizing channel for the plant, or as a barrier approach (model algorithm II), where we consider

$$\text{minimize } f_{\mu,\alpha}(K) := \max\left\{ \frac{1}{0.85}\|S\|_{I_1}, \frac{1}{1.30}\|S\|_{I_2}, \frac{1}{0.707}\|T\|_{I_3}, \|w(s)T\|_{I_4}, \mu\|T_{\text{stab}}\|_{\infty,\alpha} \right\}$$

for a threshold $\alpha < 0$, restricting poles λ of the closed-loop system to $\operatorname{Re}\lambda \leq \alpha < 0$, and for the barrier parameter $\mu > 0$. In particular, it will be interesting to see the relationship between β and $-\alpha$.

6.1 Model I: numerical difficulties with a single solve

In order to emphasize the numerical difficulties with small penalty parameters we have conducted a family of experiments for various values of β, assuming that the corresponding penalty values μ are known. All experiments are started from the same stabilizing controller K. The order of the sought controller was set to $k = 1$ in this preliminary study.

Our first experiment shows that it may not be a good idea to solve program (6) directly for the "correct" value μ_β giving $b(\mu_\beta) = \beta^{-1} = b$, because numerical difficulties arise.

β	μ	multiband performance				iter
0.68	10	0.09	0.42	1.43	[1.44]	26
0.35	1	0.17	1.06	1.07	[2.84]	32*
0.07	0.01	[1.44]	1.17	0.25	[1.44]	> 200
1.3e–4	1e–3	0.20	0.51	1.13	[7.57]	5*

Table 1: Numerical difficulties when solving directly for μ_β
* failure to achieve descent

The second column in Table 6.1 displays values of the penalty parameter μ needed to achieve the distance to instability β given in column 1. The third column provides the achieved multiband performances in the order of their introduction. The last column gives the number of inner iterations needed to reach convergence. Boxed values correspond to the final (max) multiband performance.

Note that none of the designs satisfies the performance requirements given in the beginning of the section as a value of the multiband performances below unity is required for this to hold. When β decreases and the penalty parameter μ therefore also takes smaller values, problems become increasingly ill-conditioned and either a large number of iterations are performed (row 4) or failure to achieve descent occurs (rows 2 and 4). Typically, breakdown occurs when the computed search direction is pointing uphill.

The conclusion of these experiments is that a homotopy search in the parameter μ is required. Steering μ directly or too quickly to the correct value μ_β, respectively to 0, causes failure.

6.2 Design with algorithm I and II

To overcome the above difficulties, we consider a sequence of subproblems where the barrier parameter μ is gradually decreased and controllers obtained for a subproblem serve as initial point for the next subproblem. As we are using a first-order nonsmooth method, the barrier parameter must not be decreased too aggressively. In the experiments to follow, we have used the update $\mu \leftarrow \mu/3$. For the same reason, subproblems are stopped whenever coarse criticality is met. Our stopping criterion for the subproblems uses the criticality measure $\theta_{e,\mu}(K) \leq 0$ in (15) and is defined as $\theta_{e,\mu}(K) > -\varepsilon_s$ with the updating rule $\varepsilon_s \leftarrow \max(1e\text{–}4, \varepsilon_s/2)$ and the initialization $\varepsilon_s = 10$. In this form, we require less computational work in the early iterations, while accuracy is increased close to a local solution.

Design with algorithm model I

According to our analysis in section 3, we have set b to a large value, $b = 10^5$, which corresponds to the distance to instability $\beta = 10^{-5}$. The penalty parameter μ is then decreased as long as $\| (sI - (A + BKC))^{-1} \|_\infty < b$. Results are given in Table 6.2.

Design with algorithm model II

Here the strategy is radically different as we require a minimum stability degree using the shifted H_∞ norm in section 4. Based on the philosophy $-\alpha \approx \beta$, we have set $\alpha = -1e{-}5$. The barrier parameter μ is driven to zero with the updating rule mentioned above as long as $\mu > 1e{-}8$. Both algorithms I and II are initialized with the same stabilizing controller, see table 6.2.

	α	β	multiband performance	μ	iter	$K(s)$
initial	-0.76	0.26	0.091 0.283 2.41 [42.37]	-	-	$\frac{31.41s+23.05}{s+2.286}$
model I	$-6.322e{-}5$	$1e{-}5$	[0.846] [0.846] 0.31 [0.846]	$5.51e{-}5$	108	$\frac{2.31s+2.78e-8}{s+4.26}$
model II	$-1.02e{-}5$	$7.32e{-}8$	[0.846] [0.846] 0.31 [0.846]	$8.6e{-}9$	90	$\frac{2.31s+2.35e-5}{s+4.26}$

TABLE 2: Designs with algorithm models I and II. In model I the value $\beta = 10^{-5}$ is fixed, in model II the value $\mathrm{Re}\,\lambda = -\alpha = 10^{-5}$ is imposed.

Column 'α' and 'β' provide the initial and final closed-loop spectral abscissa and distance to instability, respectively. The fourth column displays the achieved multiband performances. Column 'μ' gives the final values of the penalty respectively barrier parameter. Column 'iter' indicates the total number of inner iterations to meet our termination criterion.

Note first that controllers obtained with algorithm models I and II both meet all design requirements since all band restricted performances are below unity. This represents 20 % improvements over the results in [19]. Controllers obtained using both methods are nearly indistinguishable. Algorithm model II appears markedly superior in terms of number of inner iterations, a fact that should be confirmed on a complementary set of numerical experiences.

Figures 1 and 2 display the evolutions of the four band restricted performances versus the subproblem iteration index. Clearly, model algorithms I and II behave very similarly and could be stopped much earlier to reduce the computational overhead. It is also instructive to note that both techniques terminate at a nonsmooth local minimum where 3 over 4 band restricted performances coincide. A phenomenon that is often observed with nonsmooth problems and which has motivated research in nonsmooth optimization.

For illustration purpose, in figures 3 and 4 we show the gain plots of each channel at different stages of the barrier algorithm. The vertical lines indicate the frequency bands I_i. The horizontal lines define the cut-off level γ_c used to construct enriched sets of frequencies. The asterisk symbols correspond to gridded frequencies which have been selected to construct the enriched sets $\Omega_{e,\mu}$ and the bundle of subgradients. Figure 4, which correspond to the last inner iteration, provides a graphical verification of the achieved band restricted performances.

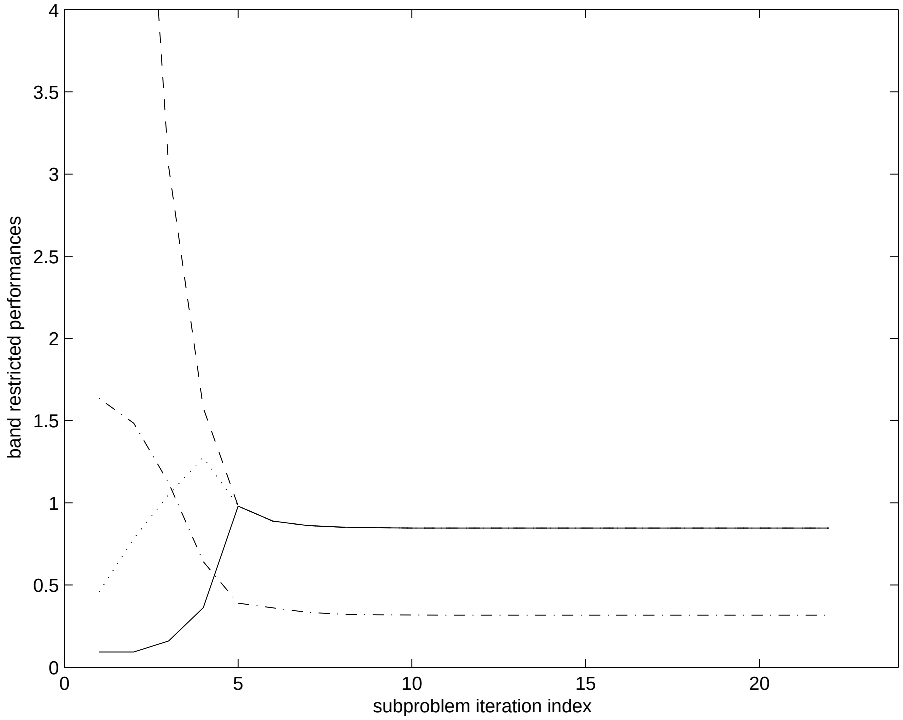

FIGURE 1: Evolution of band restricted performances vs. subproblem index model algorithm I

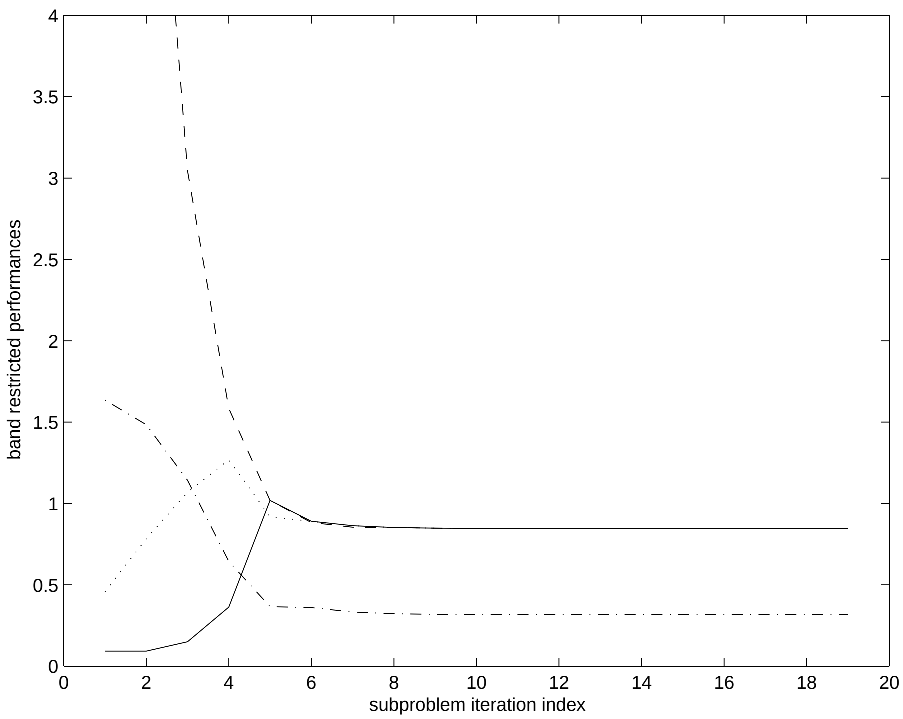

FIGURE 2: Evolution of band restricted performances vs. subproblem index model algorithm II

7 Conclusion

Multiband H_∞ synthesis is a practically important problem for which convincing algorithmic approaches are lacking. We have presented a new way to address this difficult problem using methods from nonsmooth optimization. Two ways to model closed-loop stability as a mathematical programming constraint have been introduced, discussed and compared. Correspondingly, apropriate nonsmooth optimization methods have been developed and their convergence has been established. They work satisfactory on a numerical test example involving the double integrator.

References

[1] P. Apkarian and D. Noll, *Controller design via nonsmooth multi-directional search*, SIAM J. on Control and Optimization, (2005).

[2] ——, *Nonsmooth H_∞ synthesis*, IEEE Trans. Aut. Control, (2005). to appear.

[3] ——, *Nonsmooth optimization for multidisk H_∞ synthesis*, submitted, (2005).

[4] H. W. Bode, *Network Analysis and Feedback Amplifier Design*, Van Nostrand, New York, 1945.

[5] V. Bompart, D. Noll, and P. Apkarian, *Second-order nonsmooth optimization for H_∞ and H_2 syntheses*, in preparation, (2005).

[6] J. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, 2000.

[7] S. Boyd, V. Balakrishnan, and P. Kabamba, *A bisection method for computing the H_∞ norm of a transfer matrix and related problems*, Mathematics of Control, Signals, and Systems, 2 (1989), pp. 207–219.

[8] S. Boyd and C. Barratt, *Linear Controller Design: Limits of Performance*, Prentice-Hall, 1991.

[9] S. Boyd, C. Barratt, and S. Norman, *Linear controller design: Limits of performance via convex optimization*, Proc. IEEE, 78 (1990), pp. 529–574.

[10] J. Burke, A. Lewis, and M. Overton, *Two numerical methods for optimizing matrix stability*, Linear Algebra and its Applications 351-352, (2002), pp. 147–184.

[11] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Math. Soc. Series, John Wiley & Sons, New York, 1983.

[12] B. Cornet and M.-O. Czarnecky, *Smooth normal approximations of epi-Lipschitzian subsets of $\mathbb{R}^n$*, SIAM J. on Control and Optimization, 37 (1999), pp. 710 – 730.

[13] G. F. Franklin, J. D. Powell, and A. Emami-Naeni, *Feedback Control of Dynamic Systems*, Addison-Wesley, 1986.

[14] I. Horowitz, *Quantitative feedback theory*, IEE Proc., 129-D (1982), pp. 215–226.

[15] T. Iwasaki and S. Hara, *Static gain feedback control synthesis with general frequency domain specifications.*

[16] A. G. J. MacFarlane and I. Postlethwaite, *The generalized Nyquist stability criterion and multivariable root loci.*, Int. J. Control, 25 (1977), pp. 81–127.

[17] A. Rantzer, *On the the Kalman-Yacubovich-Popov Lemma*, Syst. Control Letters, 28 (1996), pp. 7–10.

[18] L. Trefethen, *Pseudospectra of linear operators*, SIAM Review, 39 (1997), pp. 383 – 406.

[19] V.I.George, *Optimum design of robust controller: Frequency domain approach using* matlab, Journal IE(I), (2004).

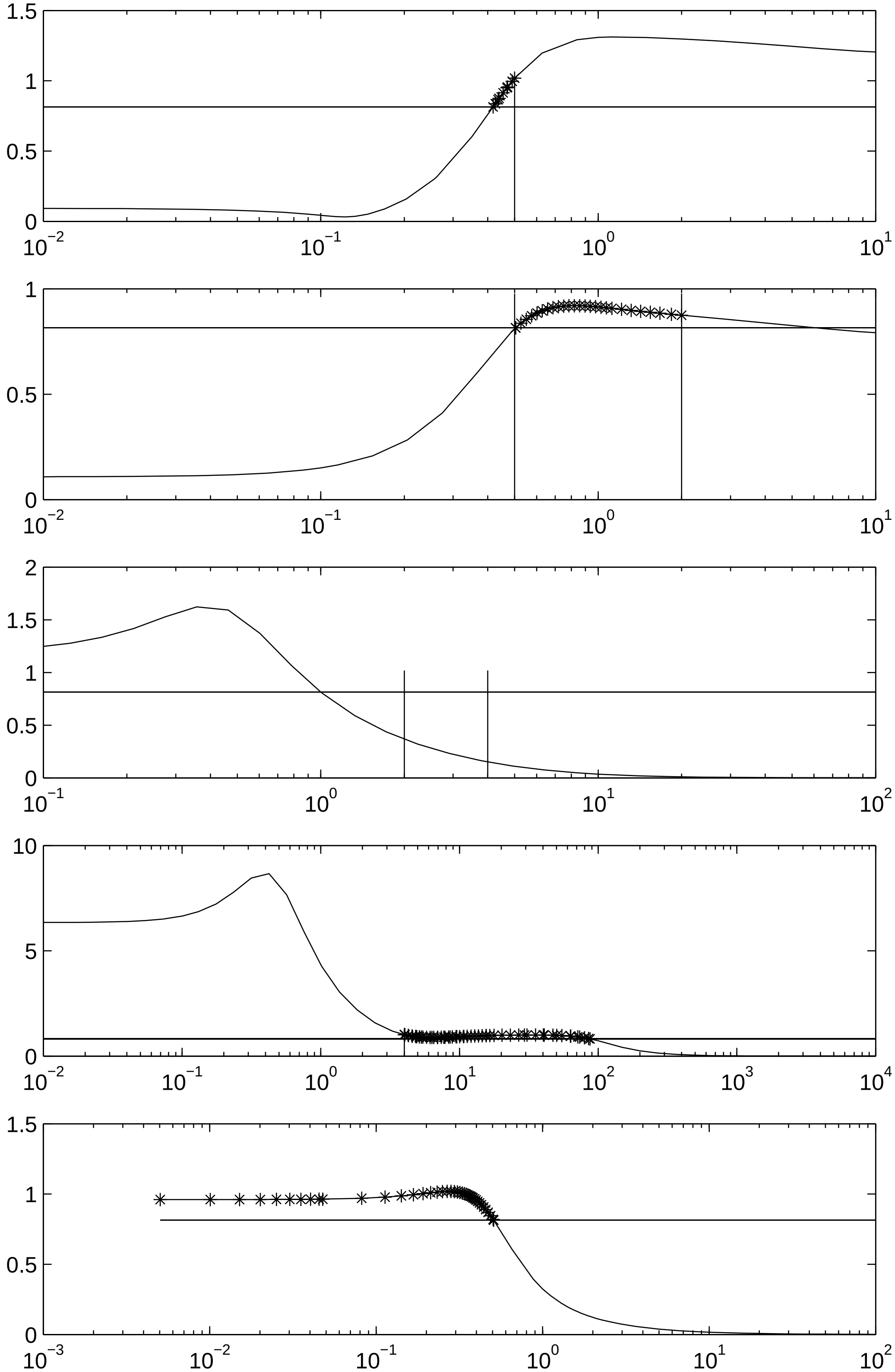

Figure 3: singular values of each specifications vs. frequency (rad/s)
$\mu = 1.23e{-}1$
'*' selected frequencies to form bundle of subgradients
outer iteration 5, inner iteration 6

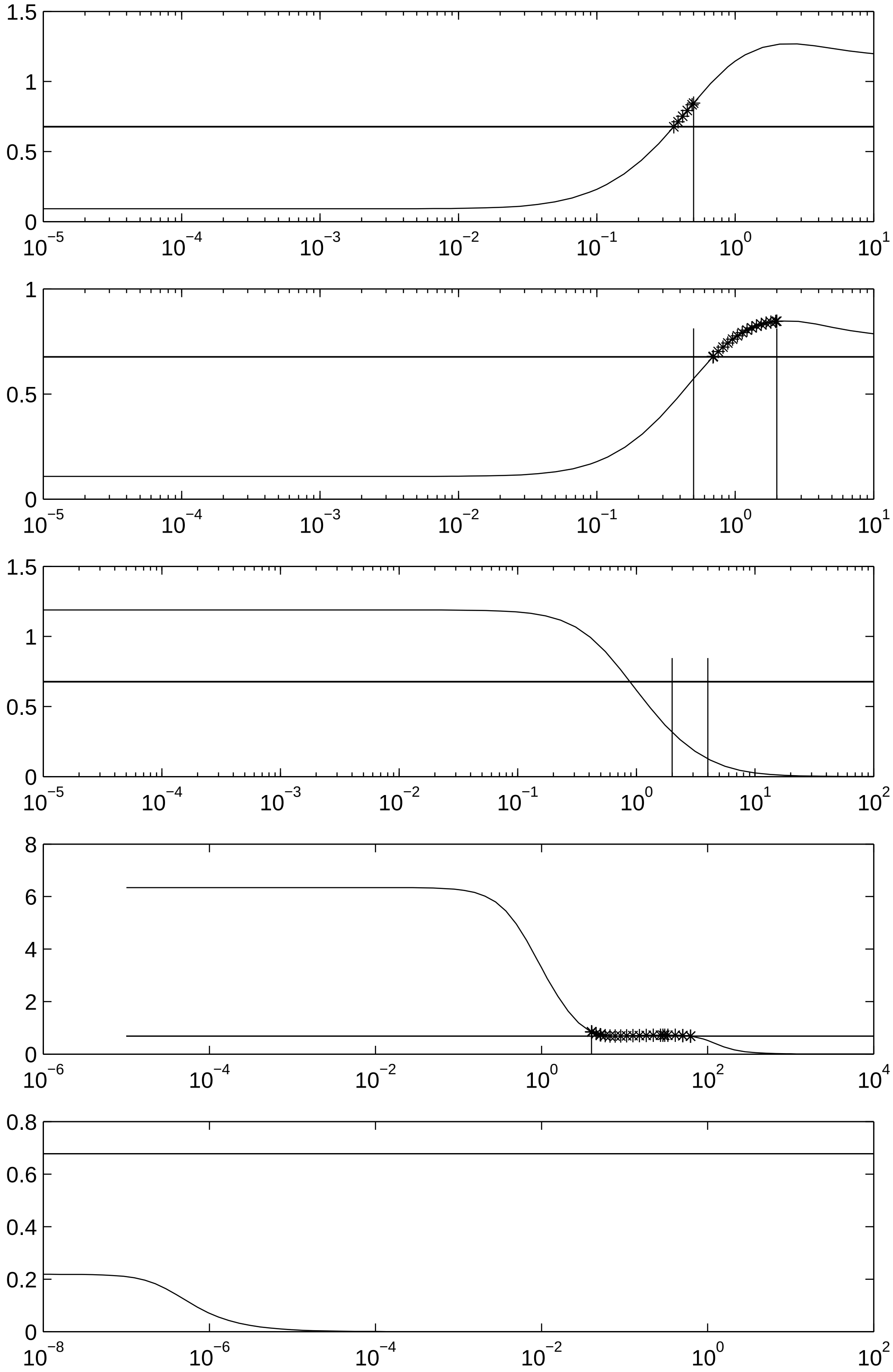

FIGURE 4: singular values of each specifications vs. frequency (rad/s)
$\mu = 8.6e{-}9$
'*' selected frequencies to form bundle of subgradients outer iteration 19, inner iteration 1