Proceedings of the
American Control Conference
San Francisco, California • June 1993

Minimum Control Effort State–Feedback $\mathcal{H}_\infty$-Control

Ewald Schömig[1]
Department of
Electrical Engineering, FT-10
University of Washington
Seattle, Washington 98195

Mario Sznaier[2]
Electrical Engineering
Department
University of Central Florida
Orlando, FL 32816-0450

Uy-Loi Ly[3]
Department of
Aero- and Astronautics, FS-10
University of Washington
Seattle, Washington 98195

Abstract

Optimal $\mathcal{H}_\infty$–controllers may exhibit large gains, resulting in large control efforts. In this paper we consider the problem of designing a *minimum gain* static full state-feedback controller such that the closed–loop transfer function satisfies a $\mathcal{H}_\infty$–constraint. The main result of the paper shows that, by minimising an upper bound for the Frobenius–norm of the feedback–gain matrix and using a parametrisation as in [6], the problem can be cast into a finite–dimensional, convex optimisation problem. Scalar cost–functions for the $\mathcal{H}_\infty$–bound and various other constraints allow the application of gradient–based software packages to these problems. Finally, we illustrate how to apply this theory to the mixed $\mathcal{H}_2/\mathcal{H}_\infty$–control problem with minimum control effort.

1. Introduction

Consider the following linear time–invariant system:

$$\Sigma : \begin{cases} \dot{x} &= Ax + B_1 w + B_2 u \\ z_2 &= C_0 x + D_0 u \\ z_\infty &= C_1 x + D_1 u \\ y &= x \\ u &= Ky \end{cases} \qquad (1)$$

where (A, B_2) is controllable. $x \in R^n$ represents the states, $u \in R^m$ represents the control action, $z_2 \in R^p$ represents variables subject to possible $\mathcal{H}_2$–performance specifications, $w \in R^r$ represents an exogenous disturbance and the transfer function from w to $z_\infty \in R^q$ is subject to the $\mathcal{H}_\infty$–bound. Note that a non–zero direct feedthrough matrix from w to z_∞ can be incorporated into this framework as well. However, without loss of generality we assume this feedthrough matrix to be zero. Given a state–feedback matrix $K \in R^{m \times n}$, the closed–loop system can be expressed as follows:

$$\Sigma_{cl} : \begin{cases} \dot{x}_{cl} &= A_{cl} x_{cl} + B_1 w \\ &= (A + B_2 K) x_{cl} + B_1 w \\ z_2 &= (C_0 + D_0 K) x_{cl} = C_{cl0} x_{cl} \\ z_\infty &= (C_1 + D_1 K) x_{cl} = C_{cl1} x_{cl} \end{cases} \qquad (2)$$

Let $T_2(s)$ denote the closed–loop transfer function from $w(s)$ to $z_2(s)$ and $T_\infty(s)$ the closed–loop transfer function from $w(s)$ to $z_\infty(s)$. Let $Tr(.)$ represent the trace operator, then we can state the design objective of a *minimum control effort* $\mathcal{H}_\infty$–problem as follows:

P1. Find a stabilising state–feedback gain matrix K such that $\|T_\infty(s)\|_\infty < \gamma$ and (an upper bound for) $\|K\|_F = Tr\{KK^T\}$ is minimised.
We also address the problem where $\|K\|_F$ is not actually minimised but bounded from above by a certain prespecified value b_K. Hence we can define an alternate criterion as follows:
P1'. Find a stabilising state–feedback gain matrix K such that $\|T_\infty(s)\|_\infty < \gamma$ and $\|K\|_F < b_K$.

The mixed $\mathcal{H}_2/\mathcal{H}_\infty$–control problem with minimum control effort can be put into the following form:
P2. Find a stabilising state–feedback gain matrix K such that $\|T_\infty(s)\|_\infty < \gamma$ and an upper bound for the weighted sum of $\|T_2(s)\|_2$ and $\|K\|_F$ is minimised.

Similar to objective P1' we can also define design objectives where bounds are imposed on $\|T_2(s)\|_2$ and/or $\|K\|_F$ if so desired. These cases will be outlined later. Note that all problems involving an $\mathcal{H}_\infty$–bound have a solution if and only if the associated pure $\mathcal{H}_\infty$–problem has a solution as shown in [14]. Also, design strategies that include a bound either on $\|K\|_F$ as in P1' or a bound on $\|T_2(s)\|_2$ such that $\|T_2(s)\|_2 < b_2$ may not have a solution even if the corresponding pure $\mathcal{H}_\infty$–bound problem has a solution.
In general all the above objectives will reduce the control effort according to $\|u\|_2 \leq \|K\|_F \|x\|_2$. Without proof we state the following results.

Lemma 1 ([6], [7]) *For a stable system Σ_{cl} the following statements are equivalent:*

[1,3] This work was supported in part by NASA Ames DFRF under grant NAG-2-629.
[2] This work was supported in part by NSF under grant ECS–9211169.

1. $\|T_\infty(s)\|_\infty < \gamma$

2. *ARI: There exists a symmetric positive–definite matrix Y such that*

$$YA_{cl}^T + A_{cl}Y + \gamma^{-2}YC_{cl1}^TC_{cl1}Y + B_1B_1^T < 0 \quad (3)$$

For future reference we define the convex sets

$$\begin{aligned}\Theta &:= \{X \in R^{n\times n} : X = X^T > 0\}\\ \Upsilon &:= \{(X,W) \in \Theta \times R^{m\times n}\}\\ \Omega &:= \{(\tau, X) \in R \times \Theta : \tau > 0\}\\ \Psi &:= \{(\tau, X, W) \in \Omega \times R^{m\times n}\}\end{aligned}$$

Lemma 2 ([6]) *Consider the system defined in Lemma 1 and let $K = WX^{-1}$ with $(X,W) \in \Upsilon$, then the following holds:*
1. The matrix function $Q(X,W) : \Upsilon \to R^{n\times n}$

$$\begin{aligned}Q(X,W) &:= X[A + B_2WX^{-1}]^T + [A + B_2WX^{-1}]X\\ &+ \gamma^{-2}X[C_1 + D_1WX^{-1}]^T[C_1 + D_1WX^{-1}]X\\ &+ B_1B_1^T \qquad (4)\end{aligned}$$

is convex on Υ. Furthermore, there exists a static state–feedback $K = WX^{-1}$ such that $\|T_\infty(s)\|_\infty < \gamma$ if and only if there is a $(X,W) \in \Upsilon$ such that $Q(X,W) < 0$.
Convexity is defined in terms of the usual ordering of symmetric matrices:

$$\begin{aligned}&Q[\alpha(W_1,X_1) + (1-\alpha)(W_2,X_2)] \qquad (5)\\ \le\ &\alpha Q(W_1,X_1) + (1-\alpha)Q(W_2,X_2).\end{aligned}$$

for two pairs of matrices $(W_1,X_1) \in \Upsilon$ and $(W_2,X_2) \in \Upsilon$.
2. The scalar quantity

$$R(X,W) = Tr\{[C_0 + D_0WX^{-1}]X[C_0 + D_0WX^{-1}]^T\}$$

is convex on Υ. Furthermore, if $(X,W) \in \Upsilon$ satisfies $Q(X,W) < 0$, then $R(X,W) > \|T_2(s)\|_2^2$.

Obviously $Q(X,W)$ represents the left–side argument of inequality (3) with a state feedback $K = WX^{-1}$ in place and $Y = X \in \Theta$. Hence the parametrization $K = WX^{-1}$ allows the formulation of the mixed $\mathcal{H}_2/\mathcal{H}_\infty$–problem in a convex setting. It should be noted that a pair $(X,W) \in \Upsilon$ that satisfies $Q(X,W) < 0$ implies that $\|T_\infty(s)\|_\infty < \gamma$. However, $Q(\bar{X},\bar{W}) \in \Psi$ and $\bar{K} = \bar{W}\bar{X}^{-1}$ that satisfies the $\mathcal{H}_\infty$–bound $\|T_\infty(s)\|_\infty < \gamma$ does not necessarily imply $Q(\bar{X},\bar{W}) < 0$. In this case Lemma 1 states that there is a matrix Y such that (3) is satisfied. In order to have $Q(\bar{X},W) < 0$ satisfied, we have to require additionally that $Y = \bar{X}$. Hence $Q(X,W) < 0$ is only a sufficient condition for $\|T_\infty(s)\|_\infty < \gamma$, not a necessary one. Note that Lemma 2 only refers to the existence of a static state–feedback matrix (that satisfies the $\mathcal{H}_\infty$–bound) in terms of $Q(X,W)$. In the next section we will give a convex upper bound for $\|WX^{-1}\|_F$ that allows us to formulate the overall problem P1 as a convex optimization problem.

2. Convex Upper Bounds for $\|WX^{-1}\|_F$

Theorem 1 *Consider the Frobenius norm of the state–feedback gain matrix $\|K\|_F = \|WX^{-1}\|_F$, then*

$$\begin{aligned}J_{B1}(\tau,X,W) &= \frac{1}{2}\tau^2\lambda_{max}(X^{-1}) + \frac{1}{2}Tr(W^TW)\\ J_{B2}(\tau,X,W) &= \frac{1}{2}Tr(\tau^2X^{-1}) + \frac{1}{2}Tr(W^TW)\end{aligned}$$

with $\tau^2X \ge I$ represent upper bounds for $\|WX^{-1}\|_F$ such that

$$\|WX^{-1}\|_F \le J_{B1}(\tau,X,W) \le J_{B2}(\tau,X,W).$$

Furthermore, $J_{B1}(\tau,X,W)$ and $J_{B2}(\tau,X,W)$ are convex on Ψ and

$$\tfrac{1}{\tau^2}I - X \le 0 \ (\Longleftrightarrow \tau^2X \ge I)$$

is a convex constraint on Ψ.

Proof: The following chain of inequalities proofs that $J_{B1}(\tau,X,W)$ and $J_{B2}(\tau,X,W)$ represent upper bounds for $\|WX^{-1}\|_F$.

$$\begin{aligned}\|WX^{-1}\|_F &= \sqrt{Tr(WX^{-1}X^{-1}W^T)} \qquad (6)\\ &\le \sqrt{Tr(\tau^2X^{-1}W^TW)} \qquad (7)\\ &\qquad s.\ t.\ \ \frac{1}{\tau^2}I - X \le 0\\ &\le \sqrt{\tau^2\lambda_{max}(X^{-1})Tr(W^TW)} \qquad (8)\\ &\le \frac{1}{2}\tau^2\lambda_{max}(X^{-1}) + \frac{1}{2}Tr(W^TW) \quad (9)\\ &\le \frac{1}{2}Tr(\tau^2X^{-1}) + \frac{1}{2}Tr(W^TW) \quad (10)\end{aligned}$$

provided that $\frac{1}{\tau^2}I - X \le 0$. Obviously equation (9) is equivalent to $J_{B1}(\tau,X,W)$ and equation (10) represents $J_{B2}(\tau,X,W)$. Equation (7) follows from (6) by the scaling of $\|K\|_F$ with $\tau^2X \ge I$. (8) follows from (7) using Lemma 3 in the Appendix. (9) follows from (8) using the arithmetic–geometric mean inequality with $\alpha = \frac{1}{2}$ and the facts that $\lambda_{max}(\tau^2X^{-1}) \ge 0$ and $Tr(W^TW) \ge 0$ (see Appendix). $J_{B1}(\tau,X,W) \le J_{B2}(\tau,X,W)$ finally follows from $\lambda_{max}(Z) \le Tr(Z)$ for any $Z \in \Theta$. Convexity of $Tr(W^TW)$ is shown in [4] (p. 556, problem 33) and the remaining convexity proofs are provided in the Appendix (see Theorems 4 and 5). As the sum of convex mappings is convex, overall convexity follows. Note that $\frac{1}{\tau^2}I - X \le 0$ is equivalent to $\tau^2X \ge I$. $\frac{1}{\tau^2}I - X \le 0$, however, is a convex constraint on Ω as shown in Theorem 6.

Both bounds are continuous on Ψ and $J_{B1}(\tau,X,W)$ is obviously a tighter bound than $J_{B2}(\tau,X,W)$. However, $J_{B2}(\tau,X,W)$ is differentiable on Ψ while $J_{B1}(\tau,X,W)$ is not differentiable at points where $\lambda_{max}(X^{-1}) = \lambda_i(X^{-1}) = \lambda_j(X^{-1})$, $i \ne j$. This property is important in the numerical solution of the minimisation problem. Design problems corresponding to these objectives can now be stated as follows.

P1: Minimum effort control with an $\mathcal{H}_\infty$-bound:

$$\min_{(\tau,X,W)\in\Psi} J_{P1}(\tau,X,W)$$
$$J_{P1}(\tau,X,W) = J_{Bi}(\tau,X,W),\quad i=1 \ \ or \ \ i=2$$
$$s.\ t.\ \ Q(X,W)<0$$
$$\frac{1}{\tau^2}I - X \le 0$$

P2: Minimum effort mixed $\mathcal{H}_2/\mathcal{H}_\infty$-control:

$$\min_{(\tau,X,W)\in\Psi} J_{P2}(\tau,X,W)$$
$$J_{P2}(\tau,X,W) = \beta J_{Bi}(\tau,X,W) + (1-\beta)R(X,W),$$
$$i=1 \ \ or \ \ i=2$$
$$s.\ t.\ \ Q(X,W)<0$$
$$\frac{1}{\tau^2}I - X \le 0$$

where $\beta \in [0,1]$ is a weighting factor. For $\beta = 0$ only the $\mathcal{H}_2$-performance measure is taken into consideration, with $\beta = 1$ the minimum effort control problem is addressed. Both minimization problems are continuous in all involved parameters and convex on Ψ.

3. A Gradient–Based Formulation

Ellipsoid or Cutting-Plane methods are applicable to this type of problem. For a review of the advantages and disadvantages of these methods and descent methods see [2] and references therein. In many cases, however, descent-methods provide faster convergence rates. In this section, we will show how to characterise the above constraints in terms of differentiable functions that maintain the convexity properties of the original constraints. Hence we arrive at unconstrained optimization problems.

Theorem 2 *Let $(X,W)\in\Upsilon$. We define a scalar measure for the $\mathcal{H}_\infty$-bound as*

$$J_I(X,W,t_f) = Tr\{e^{Q(X,W)t_f}\} \qquad (11)$$

where the scaling factor t_f is introduced for algorithmic reasons (see section 4). $J_I(X,W,t_f)$ has the following properties:

1. *$J_I(X,W,t_f)$ is non-negative.*
   Given a t_f, $J_I(X,W,t_f)$ is continuous, differentiable and convex on Υ.
2. $\lim_{t_f\to\infty}\min_{X,W} J_I(X,W,t_f) = 0 \ \Leftrightarrow\ Q(X,W) < 0$

Proof: Convexity follows from (6), Weyl's Theorem and Lemma 5 (see Appendix). The latter property of $J_I(X,W,t_f)$ follows from the fact that $Q(X,W) < 0$ is equivalent to $Q(X,W)$ being stable as $Q(X,W)$ is Hermitian. As $Tr\{e^{Q(X,W)t_f}\}$ is the sum of the exponential of the eigenvalues of $Q(X,W)t_f$, property *2.* follows directly.

It can be shown that first–order gradients of $J_I(X,W,t_f)$ can be found using the matrix series expansion of the involved matrix exponential and Kleinman's Lemma (see e.g. [13], p.263). After some matrix algebra, the gradient expressions are as follows.

$$\frac{\delta J_I(X,W,t_f)}{\delta W} = 2t_f[B_2^T + D_1^T U]e^{Q(X,W)t_f}$$
$$\frac{\delta J_I(X,W,t_f)}{\delta X} = t_f[Te^{Q(X,W)t_f} + e^{Q(X,W)t_f}T^T]$$

where

$$T = A^T + C_1^T U$$
$$U = \gamma^{-2}(C_1X + D_1W)$$

Other constraints such as $\frac{1}{\tau^2}I - X \le 0$ can be converted to convex scalar functions as well. Also, at this point we want to emphasize, that suboptimal design objectives such as in P1' can be accommodated in the same way by forming a scalar penalty function. As long as the constraint is in the form of a Hermitian matrix inequality or a scalar inequality such as $J_{B1}(\tau,X,W) < b_K$ or $R(X,W) < b_2$, this method will result in a penalty function with the same property as in Theorem 2 retaining the convexity properties of the original constraint.

4. Proposed Algorithm

The problem formulation combines all the peformance costs and the constraint penalty functions into single cost function $J_O(X,W,t_f)$.

$$J_O(X,W,t_f) = J_P(\tau,X,W) + J_C(\tau,X,W,t_f) \qquad (12)$$

where $J_P(\tau,X,W)$ is either $J_{P1}(\tau,X,W)$ or $J_{P2}(\tau,X,W)$ and represents the performance objective. $J_C(\tau,X,Wt_f)$ is the sum of all penalty function terms corresponding to constraints including the $\mathcal{H}_\infty$-penalty function $J_I(X,W,t_f)$. The proposed algorithm starts at a small t_{fs} so that initial guesses W, X and τ that do not satisfy the constraints will not result in numerical overflow problems. In a feasibility stage we optimize on $J_C(\tau,X,W,t_{fs})$ only, trying to find a feasible solution $\bar{W},\bar{X}$ and $\bar{\tau}$ satisfying the relevant constraints. Once a feasible solution is found, the performance part of the overall cost function is optimised. t_f is increased to a large value t_{fl} such that the $J_C(\bar{\tau},\bar{X},W,t_{fl}) << J_P(\bar{\tau},\bar{X},\bar{W},t_{fl})$ (note, that in the limit as $t_f\to\infty$ all exponential terms in $J_C(W,X,\tau,t_f)$ will go to zero if the according constraints are satisfied). In fact, for large but finite t_f, $J_C(\tau,X,W,t_f)$ acts as a barrier function in the optimization process.

5. Example

To illustrate our approach, consider a 4^{th}–order system used in [11]. It represents the scaled subsystem of the lateral dynamics of a BOEING 767 aircraft:

$$A = \begin{pmatrix} -0.0168 & 0.1121 & 0.0003 & -0.5608 \\ -0.0164 & -0.7771 & 0.9945 & 0.0015 \\ -0.0417 & -3.6595 & -0.9544 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -0.0243 \\ -0.0634 \\ -3.6942 \\ 0 \end{pmatrix},$$

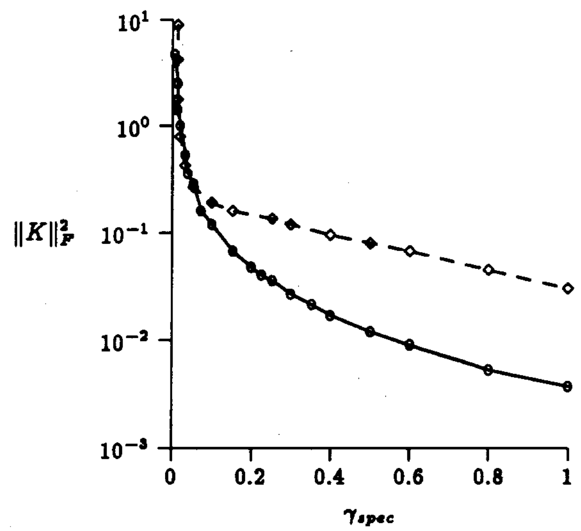
$$C_0 = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}, \quad D_0 = 1,$$

$$C_1 = \begin{pmatrix} 0.01 & 0 & 0.01 & 0 \end{pmatrix}, \quad D_1 = 0.01.$$

The open-loop system is stable, the subsystem $T_\infty(s)$ is non-minimum phase. The minimally achievable $\|T_\infty(s)\|_\infty$ is approximately 0.007. In the following picture we plot 2 curves. Each $\diamond$ represents a point design solving the convex optimisation problem P1 (see (11) with $J_{B1}(\tau, X, W)$ as performance index subject to $Q(X,W) < 0$ with γ_{spec} being the specified $\mathcal{H}_\infty$-bound. Each o represents a point design that solves the following (non-convex) optimisation problem.

$$\min_{(\tau,X,W)\in\Phi} J_{B1}(\tau, X, W)$$

subject to $\|T_\infty(s)\|_\infty < \gamma_{spec}$. Both curves show a typical



behaviour for mixed performance/robustness design objectives. For large γ_{spec}, $\|K\|_F$ is very small. If the overall problem becomes unconstrained in terms of the $\mathcal{H}_\infty$-constraint, that is if γ_{spec} is chosen large enough, $\|K\|_F$ will converge to zero (if the open-loop plant is stable). For small γ_{spec} on the other hand a dramatic increase in the controller gain can be observed for both design curves. The constraint $Q(X,W) < 0$ – as pointed out earlier – is conservative in terms of the $\mathcal{H}_\infty$-bound as it only represents a sufficient condition for $\|T_\infty(s)\|_\infty < \gamma_{spec}$. This fact is reflected in the difference between the two curves. The constraint $\|T_\infty(s)\|_\infty < \gamma_{spec}$ yields better performance than the constraint $Q(X,W) < 0$ for a given γ_{spec}. However, the latter constraint is convex while $\|T_\infty(s)\|_\infty < \gamma_{spec}$ is not.

6. Conclusions

During the last decade a powerful $\mathcal{H}_\infty$-framework has been developed, addressing the issue of robust stability in the presence of norm-bounded plant perturbations. In general, suboptimal controllers are preferred, since optimal $\mathcal{H}_\infty$-controllers may exhibit some undesirable properties, such as very large gains. Since suboptimal controllers are seldom unique, the extra degrees of freedom available can be used to solve a problem of the form *nominal performance with robust stability*. Nominal performance in this paper is characterised by *minimum control effort*. First we have shown that, by using a controller characterisation as in [6], this problem can be cast into a finite–dimensional convex constrained optimisation form. In the second part transformed this optimisation problem into a unconstrained differentiable optimisation problem, amenable to solution by gradient–based methods.

References

[1] E.F. Beckenbach, R. Bellman, "Inequalities," Fourth Printing, Springer Verlag, 1983.

[2] S.P. Boyd, C.H. Barratt, "Linear Controller Design, Limits of Performance," Prentice Hall, 1991.

[3] R.Y. Chiang, M.G. Safonov, "Design of an $\mathcal{H}_\infty$ controller for a Lightly Damped System using a Bilinear Pole Shifting Transform," ACC 1991, pp. 1927–1928.

[4] R. Horn, C.R. Johnson, "Matrix Analysis," Cambridge University Press, 1985.

[5] R. Horn, C.R. Johnson, "Topics in Matrix Analysis," Cambridge University Press, 1991.

[6] P.P. Khargonekar, M.A. Rotea, "Mixed H_2/H_∞ Control: A Convex Optimisation Approach," *IEEE Trans. AC-36,* No. 7, 1991, pp. 824–837.

[7] P.P. Khargonekar, K. Zhou, "An Algebraic Riccati Equation Approach to H^∞ Optimisation," *Systems and Control Letters 11,* 1988, pp. 85–91.

[8] J.R. Magnus, H. Neudecker, "Matrix Differential Calculus With Applications In Statistics And Econometrics," Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, 1988.

[9] M.A. Rotea, P.P. Khargonekar, "H^2-Control with an H_∞-Constraint: The State Feedback Case," *Automatica,* Vol. 27, No. 2, 1991, pp. 307–316.

[10] M.A. Rotea, P.P. Khargonekar, "Simultaneous H^2/H^∞ Optimal Control with State Feedback," ACC 1990, pp. 2380–2384.

[11] E. Schömig, U. Ly, "Mixed H_2/H_∞-Control With An Output-Feedback Compensator Using Parameter Optimisation," ACC 1992, pp. 733–737.

[12] E. Schömig, "Mixed H_2/H_∞-Control with General Output–Feedback Using Parameter Optimisation Methods," Ph.D. Dissertation, Electrical Engineering Department, University of Washington, 1993.

[13] A. Weinmann, "Uncertain Models and Robust Control," Springer-Verlag, Wien New York, 1991.

[14] K. Zhou, J. Doyle, K. Glover, B. Bodenheimer, "Mixed H_2 and H_∞ Control," ACC 1990, pp. 2502–2507.

Appendix

Theorem 3 (Weyl's Theorem, [4], p.181)
Let $G,\ H \in R^{n\times n}$ be Hermitian matrices, let the eigenvalues of G, H and $G+H$ be arranged in the following order
$\lambda_1(G) \le \lambda_2(G) \le \dots \le \lambda_n(G) = \lambda_{max}(G)$
$\lambda_1(H) \le \lambda_2(H) \le \dots \le \lambda_n(H) = \lambda_{max}(H)$
$\lambda_1(G+H) \le \lambda_2(G+H) \le \dots \le \lambda_n(G+H) = \lambda_{max}(G+H)$,
then

$$\lambda_i(G+H) \le \lambda_i(G) + \lambda_{max}(H) \tag{13}$$

for all $i = 1, 2, \dots n$.
In particular we have

$$\lambda_{max}(G+H) \le \lambda_{max}(G) + \lambda_{max}(H). \tag{14}$$

and, for $H \le 0$,

$$\begin{aligned} \lambda_i(G+H) &\le \lambda_i(G) \\ \lambda_{max}(G+H) &\le \lambda_{max}(G). \end{aligned}$$

Lemma 3 ([13], p.630)
Let $G,\ H \in R^{n\times n}$ be Hermitian matrices such that $G \ge 0$ and $H \ge 0$, then

$$Tr(GH) \le \lambda_{max}(G) Tr(H). \tag{15}$$

Lemma 4 ([1])
(Arithmetic–Geometric Mean Inequality)
Let x and y be two non-negative scalars, then

$$x^\alpha y^{(1-\alpha)} \le \alpha x + (1-\alpha) y \tag{16}$$

for every $\alpha \in (0,1)$.

Lemma 5 ([5])
Let Z_1 and Z_2 be Hermitian matrices and $\alpha \in (0,1)$, then

$$\begin{aligned} Tr\{e^{[\alpha Z_1 + (1-\alpha) Z_2] t_f}\} &\le [Tr(e^{Z_1 t_f})]^\alpha [Tr(e^{Z_2 t_f})]^{(1-\alpha)} \\ &\le \alpha Tr(e^{Z_1 t_f}) + (1-\alpha) Tr(e^{Z_2 t_f}). \end{aligned}$$

Theorem 4
The function

$$J(\tau, X) = \tau^2 \lambda_{max}(X^{-1}) \tag{17}$$

is convex on Ω.

Proof: The proof utilises results in [6] and is very similar to that. As $J(\alpha\tau, \alpha X) = \alpha J(\tau, X)$ we only have to show that

$$J(\tau_1+\tau_2, X_1+X_2) \le J(\tau_1, X_1) + J(\tau_2, X_2). \tag{18}$$

Let S be a nonsingular matrix such that

$$\begin{aligned} S^T X_1 S &= \Lambda_1 = diag(\lambda_{1,i}),\ i = 1,2,\dots,n \\ S^T X_2 S &= \Lambda_2 = diag(\lambda_{2,i}),\ i = 1,2,\dots,n \end{aligned}$$

Such a matrix exists for positive-definite matrices X_1 and X_2 (see [6]).

$$\begin{aligned} & J(\tau_1+\tau_2, X_1+X_2) \\ = {} & \lambda_{max}[(\tau_1+\tau_2)^2 (X_1+X_2)^{-1}] \\ = {} & \lambda_{max}[S^{-1}\{diag(\frac{(\tau_1+\tau_2)^2}{\lambda_{1,i}+\lambda_{2,i}})\} S^{-T}] \\ = {} & \lambda_{max}[S^{-1}\{diag(\frac{\tau_1^2}{\lambda_{1,i}} + \frac{\tau_2^2}{\lambda_{2,i}} + \psi_i)\} S^{-T}] \\ = {} & \lambda_{max}[\tau_1^2 S^{-1}\Lambda_1^{-1}S^{-T} + \tau_2^2 S^{-1}\Lambda_2^{-1}S^{-T} \\ & + S^{-1} diag(\psi_i) S^{-T}] \end{aligned}$$

where $\psi_i = -\frac{(\tau_1\lambda_{2,i} - \tau_2\lambda_{1,i})^2}{(\lambda_{1,i}\lambda_{2,i})\lambda_{1,i}\lambda_{2,i}}$, $i = 1, 2, \dots, n$.
Thus

$$J(\tau_1+\tau_2, X_1+X_2) = \lambda_{max}[\tau_1^2 X_1^{-1} + \tau_2^2 X^{-1} + Q] \tag{19}$$

for $Q = S^{-1} diag(\psi_i) S^{-T} \le 0$ and hence

$$\begin{aligned} J(\tau_1+\tau_2, X_1+X_2) &\le \lambda_{max}(\tau_1^2 X_1^{-1} + \tau_2^2 X_2^{-1}) \\ &\le \lambda_{max}(\tau_1^2 X_1^{-1}) + \lambda_{max}(\tau_2^2 X_2^{-1}) \\ &= J(\tau_1, X_1) + J(\tau_2, X_2) \end{aligned}$$

The inequalities follow immediately from the above Lemmas and Weyl's Theorem. An alternative proof can be constructed using Fischer's min–max theorem (see [8]).

Theorem 5
The function

$$J(\tau, X) = Tr(\tau^2 X^{-1}) \tag{20}$$

is convex on Ω.

Proof: A proof can be constructed using the same tools as above. It is essentially equivalent to the proof of Lemma 4.4 in [6] and is omitted here.

Theorem 6
Consider a mapping $f : \Omega \to R^{n\times n}$ and f given by

$$f(\tau, X) = \frac{1}{\tau^2} I - X. \tag{21}$$

Then $f(\tau, X)$ is a real-analytic convex mapping on Ω.

Proof: $f(\tau, X)$ is affine in X. $\frac{1}{\tau^2}$ is a strictly monotonically decreasing function for all $\tau > 0$. Convexity on Ω follows immediately.