H_∞ Filtering and Smoothing

PAULO OLIVEIRA

DEECIST and ISR

Last Revised: November 10, 2007

Ref. No. NLO#1

Theme

- KF only corresponds to the optimal filtering strategy under restrictive assumptions, and for some objectives (functionals)

- The requirement on the knowledge of the power spectral density of the disturbances is too restrictive. The WNG assumption too.

- Unknown multimodal and/or skewed pdfs are common

However

• Optimality and stability still of great importance, in the presence of uncertainty (robustness)

• Other functionals / objectives can be used to formulate estimation problems. Minimization must be feasible

Norms of Signals

L_1 [0,T] norm

$$\|u\|_1 = \int_0^T |u(t)|\,dt < \infty$$

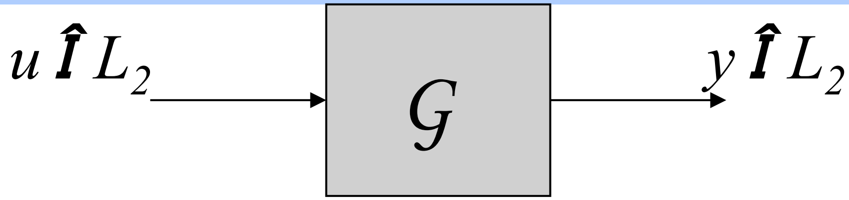
L_2 [0,T] norm
(energy)

$$\|u\|_2 = \left(\int_0^T u(t)^T u(t)\,dt \right)^{\frac{1}{2}} < \infty$$

L_∞ norm
(least upper bound)

$$\|u\|_\infty = \sup_t \left(|u(t)| \right) < \infty$$

Motivation for H_∞ Filtering



For finite energy signals in the input of system $\mathcal{G}$, how much is the minimum energy on the output?

Possible interpretation as a Min-max Nash game in estimation:
Maximum energy in the error is minimized.

For bounded systems, the H_∞ norm is defined as

$$\|\mathcal{G}\|_\infty = \sup_{u \in L_2, \|u\|_2 \neq 0} \frac{\|\mathcal{G}u\|_2}{\|u\|_2}$$

Denominated as the L_2 induced norm.

For LTI systems corresponds to the peak in the Bode diagram.

Norms of Systems

LTI Continuous - time model

$$\Sigma_{\mathcal{G}}:\begin{cases}\dot{x}(t)=Ax(t)+Bu(t)\\ y(t)=Cx(t)\end{cases}$$

Transfer function

$$G(s)=C(sI-A)^{-1}B$$

H_2 norm

$$\|\mathcal{G}\|_2=\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty}tr\left(G(j\omega)G^*(j\omega)\right)d\omega\right)^{1/2}$$

$$=\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty}\sum_i\sigma_i^2(G(j\omega))d\omega\right)^{1/2}$$

H_∞ norm

$$\|\mathcal{G}\|_\infty=\sup_\omega\sigma_{\max}[G(j\omega)]$$

*Input-output Relations**

	Stochastic	$\|y\|_2$	$\|y\|_\infty$
Stochastic	$\|G\|_2$	∞	?
$u(t)=\delta(t)$	?	$\|G\|_2$	$\|G\|_\infty$
$\|u\|_2$	?	$\|G\|_\infty$	$\|G\|_2$
$\|u\|_\infty$	∞	∞	$\|G\|_1$

*See [1] for details

Plant and Sensor Modeling

$$\Sigma_G : \begin{cases} \dot{x}(t) = A(t)x(t) + B(t)w(t) \\ y(t) = C(t)x(t) + D(t)w(t) \end{cases} \quad t \in [0, T] \qquad \begin{matrix} x(t) \in R^n \\ y(t) \in R^p \\ w(t) \in R^m \end{matrix}$$

$A(t), B(t), C(t), D(t)$ piecewise continuous bounded functions.

$z(t) = L(t)x(t)$ quantity of interest to be estimated.

(A, B) is stabilizable

(A, C) is detectable

$$\mathrm{D}(\mathrm{t})\begin{bmatrix} B(t)^T \\ D(t)^T \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

Independent plant/measurement noises

Normalized measurement noise

H_∞ Filtering

Problem statement : For system $\mathcal{G}$, with known (unknown) initial conditions* and using the measurements $y(t)$, obtain an estimate $\hat{z}(t)$ of $z(t)$ that minimizes the (worst case) indeces

$$J_1 = \sup_{0 \neq w \in L_2} \frac{\|z-\hat{z}\|_2^2}{\|w\|_2^2} \quad or \quad J_2 = \sup_{0 \neq w \in L_2} \frac{\|z-\hat{z}\|_2^2}{\|w\|_2^2 + x_0' R x_0}, \qquad R^{**} > 0.$$

Important questions:

- Given $g>0$, does there exist a filter with finite J_1 (or J_2)?
- Under the assumptions , does it verifies $J_1<g^2$ (or $J_2<g^2$)?
- How to find a realization for such filter?

* Considered 0 without loss of generality

** R^{-1} is a covariance matrix

H_∞ Filtering
Finite Horizon, Known Initial Conditions

Theorem[2] : Let the initial conditions be known and $T < \infty$.

1)There exist a filter such that $J_1 < \gamma^2$ if and only if there exists a symmetric matrix $P(t)$ for $t \in [0, T]$ that satisfies

$$\begin{aligned}\dot{P}(t) = {} & A(t)P(t) + P(t)A^T(t) - P(t)C^T(t)C(t)P(t) \\ & + \frac{1}{\gamma^2} P(t)L^T(t)L(t)P(t) + B(t)B^T(t) \qquad \text{with } P(0) = 0.\end{aligned} \tag{2}$$

2)Moreover, if it exists, one filter for $J_1 < \gamma^2$ is given by

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + P(t)C^T(t)\left[y(t) - C(t)\hat{x}(t)\right] \qquad \text{with } \hat{x}(0) = 0. \tag{3}$$

Null initial conditions considered without loss of generality.

Elements of Proof (I)

The value of the functional J_1 can be written as

$$\frac{\|z-\hat{z}\|_2^2}{\|w\|_2^2}=\gamma^2.$$

From $\Sigma_{\mathcal{G}}$, introducing $\hat{z}(t)=L(t)\hat{x}(t)$ and $\tilde{x}(t)=x(t)-\hat{x}(t)$

$$\frac{1}{\gamma^2}\|L(t)\tilde{x}(t)\|_2^2-\|w\|_2^2=0.$$

Using the $L_2[0,T]$ norm definition (1) we can write

$$\int_0^T\left\{\begin{bmatrix}\tilde{x}(t)^T & w(t)^T\end{bmatrix}\begin{bmatrix}\frac{1}{\gamma^2}L(t)^T L(t) & 0\\ 0 & -I\end{bmatrix}\begin{bmatrix}\tilde{x}(t)\\ w(t)\end{bmatrix}\right\}=0 \qquad (4)$$

Systems' Theory Digression

Dissipativity [6] - The system $\mathcal{G}: w \rightarrow z$ with supply rate $s(t)$ is strictly dissipative if there exists a non - negative function $V: x \rightarrow R$ such that

$$V(x(t_0)) + \int_{t_0}^{t_1} s(u(t), y(t))dt > V(x(t_1))$$

for all $t_0 < t_1$ and for all trajectories of the system.

A system is dissipative if can not provide to the environment the same energy that was suplied by the exterior – energy losses.

Examples: electrical circuits, mechanical systems, thermodynamics...

Moreover, if $V(t)$ is differentiable, $\dot{V}(t) < s(t)$ holds.

From $\int_{t_0}^{t_1} - s(u(t), y(t)) + \dot{V}(x(t))dt < 0$, for any $t \in [t_0, t_1]$.

Lyapunov Stability – Second Method

Lyapunov Stability - An equilibrium point $x = 0$ is stable if

$$\forall \varepsilon, t > 0, \exists_{\delta(\varepsilon)>0} : \qquad \|\mathrm{x}(0)\| < \delta \Rightarrow \|\mathrm{x}(\mathrm{t})\| < \varepsilon.$$

Lyapunov theorem (second method) - The equilibrium point $x = 0$ is stable if there exists a Lyapunov function that verifies

$$i) \quad V(0) = 0$$

$$ii) \; V(x) > \alpha \|x\|_2$$

$$iii) \; \dot{V}(x(t)) \leq 0, \text{ along all solutions of } S.$$

- Note that $V(t) \to \infty$ as $\|x\|_2 \to \infty$.

- Stability of dynamic systems can be studied, whitout solving the differential equations. Sufficient conditions.

- No systematic method to find a Lyapunov function exists.

Elements of Proof (II)

Re - interpreting (4) and resorting to dissipativity concepts

the Lyapunov candidate function $V(x) = \tilde{x}^T(t) P^{-1}(t) \tilde{x}(t)$ is used

$$\int_0^T \left\{ \begin{bmatrix} \tilde{x}(t)^T & w(t)^T \end{bmatrix} \begin{bmatrix} \frac{1}{\gamma^2} L(t)^T L(t) & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ w(t) \end{bmatrix} + \frac{d}{dt} V(t) dt \right\} = 0$$

From the definition of $\mathcal{G}$ and (3), the error dynamics is

$$\dot{\tilde{x}} = \left(A - PC^T C\right)\tilde{x} + \left(B - PC^T D\right) w$$

Therefore

$$\begin{aligned} \dot{V}(t) &= \dot{\tilde{x}}^T(t) P^{-1}(t) \tilde{x}(t) + \tilde{x}^T(t) \dot{P}^{-1}(t) \tilde{x}(t) + \tilde{x}^T(t) P^{-1}(t) \dot{\tilde{x}}(t) \\ &= \dot{\tilde{x}}^T(t) P^{-1}(t) \tilde{x}(t) - \tilde{x}^T(t) P^{-1}(t) \dot{P}^{-1}(t) P^{-1}(t) \tilde{x}(t) + \tilde{x}^T(t) P^{-1}(t) \dot{\tilde{x}}(t) \end{aligned}$$

Elements of Proof (III)

Re-arranging the terms results

$$\int_0^T \left\{ \begin{bmatrix} \tilde{x}(t)^T & w(t) \end{bmatrix}^T \bar{P} \begin{bmatrix} \tilde{x}(t) \\ w(t) \end{bmatrix} \right\} dt = 0,$$

where*

$$\bar{P} = \begin{bmatrix} \frac{1}{\gamma^2} L^T L + A^T P^{-1} - 2C^T C - P^{-1}\dot{P}P^{-1} + P^{-1}A & P^{-1}B - C^T D \\ BP - D^T C & -I \end{bmatrix}.$$

Using (2) and cancelling terms

$$\int_0^T \left\{ \begin{bmatrix} \tilde{x}(t)^T & w(t) \end{bmatrix}^T \begin{bmatrix} -C^T C - PBB^T P & P^{-1}B - C^T D \\ BP - D^T C & -I \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ w(t) \end{bmatrix} \right\} dt = 0$$

*Time-dependence omitted for simplicity.

Elements of Proof (IV)

Schur Complements - Given matrices $U \in R^{nxn}$, $V \in R^{nxm}$, $W \in R^{mxn}$, and $Z \in R^{mxm}$, where $Z > 0$ the Schur complements of matrix $\begin{bmatrix} U & V \\ W & Z \end{bmatrix}$ is $U - VZ^{-1}W$.

Can be seen as a generalization to the matrix inversion lemma.

Using Schur complements

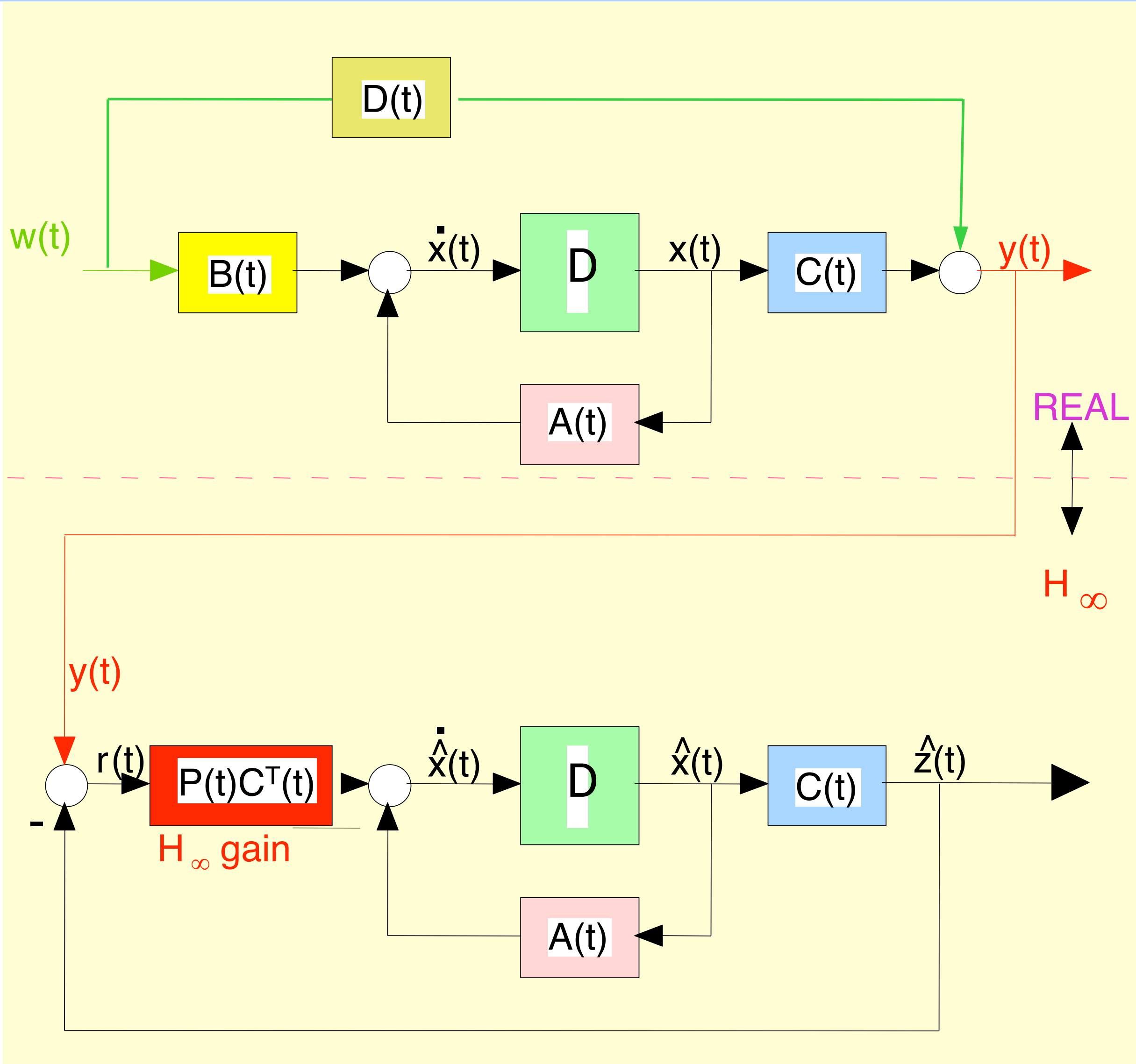
$$-C^T C - PBB^T P^{-1} + \left(B - C^T D\right)\left(BP - D^T C\right) = 0$$

$$-C^T C - PBB^T P^{-1} + PBB^T P^{-1} - P^{-1}BD^T C - C^T DB^T P^{-1} + C^T C = 0$$

using the noises independence and normalization assumptions

$$0 = 0 \quad . \qquad\qquad q.e.d.$$

Visualization of the H_∞ Filter



Discussion

- Optimal structure obtained, similar to LTV Kalman filter

- Unbiased estimator obtained (otherwise $J_1 \rightarrow \infty$, $J_2 \rightarrow \infty$)

- Complete proof is out of scope, but can be obtained

i. using systems' theory [2, 5];

ii. using estimation tools in Krein spaces [3];

- Stationary solutions can also be obtained (finite or infinite horizon cases)

- Modified Riccati equation that

  For $g \rightarrow \infty$ degenerates on the Riccati equation in KF

  Provides more robust solutions, for smaller g

  Unfeasible for $g < g_{\text{min}}$!!

H_∞ Smoothing
Finite Horizon, Known Initial Conditions

Theorem[2] : Let the initial conditions be known and $T < \infty$.

1)There exist a smoother such that $J_1 < \gamma^2$ if and only if there exists a symmetric matrix $X(t)$ for $t \in [0, T]$ that satisfies

$$
-\dot{X}(t) = A^T(t)X(t) + X(t)A(t) - X(t)B(t)B^T(t)X(t) - \frac{1}{\gamma^2}L^T(t)L(t) + C^T(t)C(t)
$$

with $X(T) = 0.$

2)One smoother that minimizes J_1 and verifies $J_1 < \gamma^2$ is

$$
\begin{bmatrix} \dot{\hat{x}}(t) \\ \dot{\lambda}(t) \end{bmatrix} = \begin{bmatrix} A(t) & B(t)B^T(t) \\ C^T(t)C(t) & -A^T(t) \end{bmatrix} \begin{bmatrix} \dot{\hat{x}}(t) \\ \dot{\lambda}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ C(t) \end{bmatrix} y(t)
$$

with $\hat{x}(0) = 0, \lambda(t) = 0.$

Remarks

- Proof is omitted, see [2] for details.
- The H_∞ smoother structure is equal to the H_2!
- Smoothers for all 4 cases are well known.
- Much more recent results than the H_2 solutions
- Other functionals have already been solved, e.g. mixed H_2/H_∞ Also, solutions for nonlinear cases available
- Now a couples of examples from [5] are included to document some of the results outlined

Examples, from [5]

Example 1: In this example, we demonstrate the reduced peak-error-level of an H_∞-filter, and its inherent robustness. We apply H_∞-optimal and L_2-optimal filters on the following second order resonant system

$$\dot{x} = \begin{bmatrix} 0 & w_n \\ -w_n & -2\xi w_n \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w, \quad y = [0 \;\; 1]\, x + n, \quad z = [1 \;\; 0]\, x$$

where w_n and ξ are not certain. The filters were designed for a nominal system with $w_n = 11$ and $\xi = 0.1$. Figure 4.1 depicts the Bode magnitude plot of T_{rd} of the H_∞ and L_2 filters, for the nominal case, and an envelope of T_{rd}, for w_n varying in the range [8.2-13.7] and ξ varying in the range [0.075-0.125].

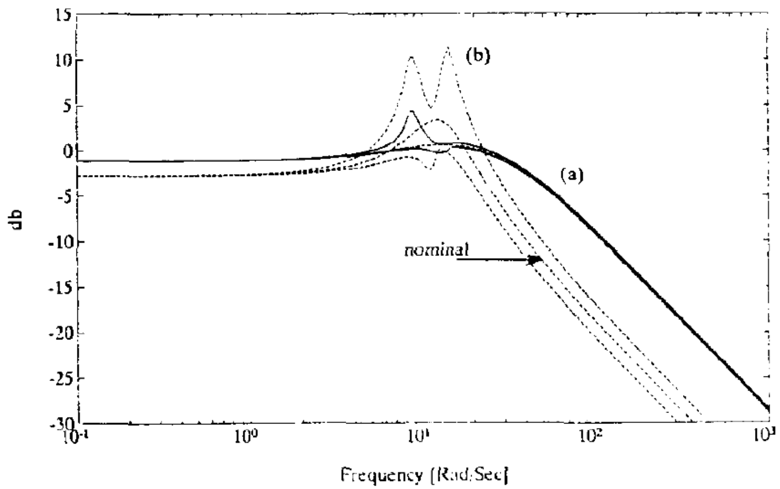


Fig. 4.1: Sensitivity comparison between: (a) the H_∞-filter, and (b) the L_2-filter.

Examples, from [5]

Example 2 (Deconvolution): In this example we demonstrate the tradeoff that exists between the L_2 and H_∞ performance in a continuous-time, steady-state filter design. In the deconvolution problem of Fig. 4.2, we use the noise corrupted measurement of the output of a system, to estimate a regularized version of its input. The regularizing filter is required to make the deconvolution problem well-posed. We look for a filter that achieves $\|T_{rd}\|_\infty < \gamma$ for the following systems:

$$G_s(s) = \frac{100}{s^2 + 0.4s + 100}, \quad G_r(s) = \frac{10^4}{s^2 + 130s + 10^4}, \quad \text{SNR} = 100$$

Recalling that $\gamma \to \infty$ leads to L_2-estimation, we are motivated to check few values of γ. The transfer function T_{rd} for central filters that were designed with different values for γ is depicted in Fig. 4.3. The effect of the design parameter γ on the performance of the above deconvolutor is further emphasized in Fig. 4.4., where the H_∞-norm that is actually achieved is related to the design parameter γ, and the corresponding L_2-norm of T_{rd}. In this typical example, we see that γ is an effective design parameter for values that are near γ_0, where a significant improvement in the L_2 performance can be gained by slightly compromising the H_∞ performance.

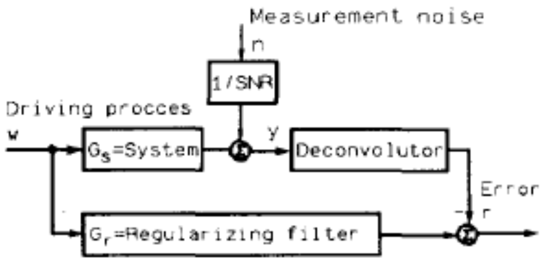


Fig 4.2 The deconvolution scheme

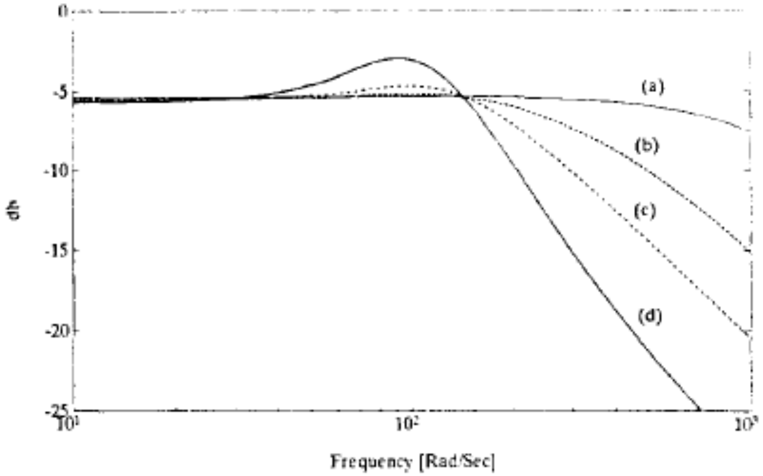


Fig 4.3: The Bode plot of T_{rd} for: (a) $\gamma=\gamma_0$; (b) $\gamma=1.02\gamma_0$; (c) $\gamma=1.1\gamma_0$; (d) $\gamma\to\infty$.

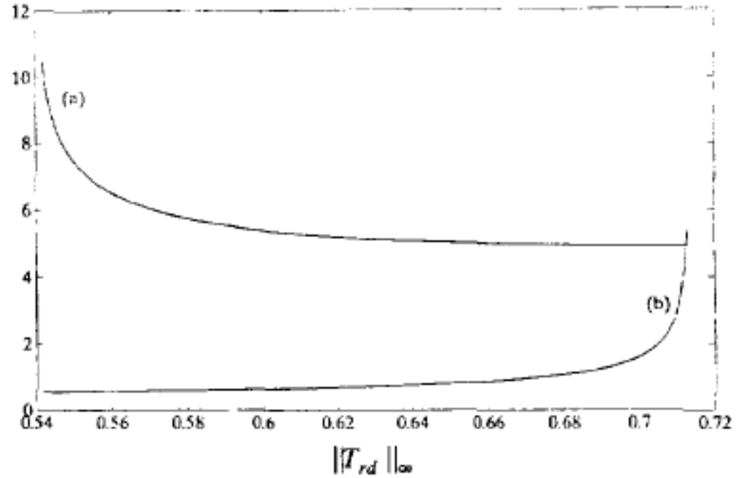


Fig 4.4: The tradeoff between L_2 and H_∞ performance: (a) $\|T_{rd}\|_2$; (b) γ.

References

[1] J. Doyle, B. Francis, and A. Tannenbaum, *Feedback Control Theory,* Macmillan Press, 1990

[2] K. Nagpal and P. Khargonekar, *"Filtering and Smoothing in the* H_∞ *Setting,"* IEEE TAC, vol 36, n2, Feb. 1991

[3] B. Hassibi, A. Sayed, and T, Kailath, *"Linear Estimation in Krein Stpaces – Part I Theory,"* IEEE TAC, vol 41, n1, Jan. 1996

[4] C. Scherer, P. Gahinet, and M. Chilali, *"Multiobjective Output-feedback Control Via LMI Optimization,"* IEEE TAC, vol 42, n7, 1997

[5] U. Shaked and Y. Theodor, *"*H_∞*-Optimal Estimation: A Tutorial," Proc. CDC, pp. 2278-2286Arizona, USA,1992.*

[6] J. Willems, *"Dissipative dynamic systems,"* Arch. Rational Mech. Anal, Vol. 45, pp. 321-393, 1971.