European Journal of Control, vol. 6, no. 1, pp. 89-92, 2000.

Discussion by Ali H. Sayed*

The main result of the article by Bolzern et al. is Theorem 3.[1] In order to put the result into perspective, and in order to comment on its value, let us first summarize its conclusion.

Thus consider the standard state-space model (1)–(2) in the paper, where $\varphi(t)$ denotes a regressor sequence and $\vartheta(t)$ is the state variable. One can pose the problem of estimating the uncorrupted output $z(t) = \varphi(t)'\vartheta(t)$ in a robust manner by seeking an estimator $\hat{z}(t)$ that satisfies a (suboptimal) H_∞ criterion, namely one that guarantees that the following bound is met:

$$\frac{\sum_{t=0}^{T-1}[z(t)-\hat{z}(t)]^2}{\sum_{t=0}^{T-1} q^{-1}w(t)'w(t) \ + \ \sum_{t=0}^{T-1} v(t)^2 \ + \ \left(\vartheta(0)-\hat{\vartheta}_0\right)' P_0^{-1}\left(\vartheta(0)-\hat{\vartheta}_0\right)} \ \leq \ \gamma^2 \tag{1}$$

for some γ, and for any initial guess $\bar{\vartheta}_0$ and any square-summable sequences $\{w, v\}$.

One solution to this problem (also known as the central solution) is given by equations (4),(5),(10), (11) in the paper:

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + K(t)\left(y(t) - \varphi(t)'\hat{\vartheta}(t-1)\right), \quad \vartheta(-1) = \hat{\vartheta}_0, \tag{2}$$

$$\hat{z}(t) = \varphi(t)'\hat{\vartheta}(t), \tag{3}$$

$$K(t) = \frac{\Pi(t)\varphi(t)}{1+\varphi(t)'\Pi(t)\varphi(t)}, \tag{4}$$

$$\Pi(t+1) = \left(\Pi(t)^{-1} + (1-\gamma^{-2})\varphi(t)\varphi(t)'\right)^{-1} \ + \ qI, \ \ \Pi(0) = P_0, \tag{5}$$

if the feasibility condition (12) is satisfied, i.e.,

$$\Pi(t)^{-1} + (1-\gamma^{-2})\varphi(t)\varphi(t)' > 0, \quad t \in [0, T-1].$$

This condition is obviously always satisfied for any $\gamma \geq 1$. Also, since the standard Kalman (or H_2) filter solution satisfies (1) with $\gamma = 2$, it is clear that we need only focus on choices of γ that lie in the interval $[1, 2)$.

Theorem 3 in the paper then states that the above H_∞-filter actually guarantees a tighter bound in (1) above than the desired γ. In particular, the theorem states that it holds that

$$\frac{\sum_{t=0}^{T-1}[z(t)-\hat{z}(t)]^2}{\sum_{t=0}^{T-1} q^{-1}w(t)'w(t) \ + \ \sum_{t=0}^{T-1} v(t)^2 \ + \ \left(\vartheta(0)-\hat{\vartheta}_0\right)' P_0^{-1}\left(\vartheta(0)-\hat{\vartheta}_0\right)} \ \leq \ \left(\frac{2}{1+\gamma^{-2}}\right)^2 \ \leq \gamma^2. \tag{6}$$

*A. H. Sayed is with the Electrical Engineering Department, UCLA, Los Angeles, CA 90095.

[1] Bolzern et al., Tradeoffs between mean-square and worst-case performances in adaptive filtering, *European Journal of Control*, to appear, 2000.

This means the following. If one desires an H_∞ filter with a certain robustness level, say $\bar{\gamma}$, then one could instead design an H_∞-filter as above with the following value of γ,

$$\gamma = \sqrt{\frac{\bar{\gamma}}{2-\bar{\gamma}}} \;\geq\; \bar{\gamma}, \tag{7}$$

which is larger than $\bar{\gamma}$. Then by (6), the filter designed with this γ will have a robustness level that does not exceed

$$\frac{2}{1+\gamma^{-2}} \;=\; \bar{\gamma}.$$

Figure 1 compares the values of γ (solid line) that correspond to values of $\bar{\gamma}$ in the interval $[1,2)$, as given by the transformation (7). The horizontal axis is $\bar{\gamma}$ and the dotted curve is the line $y = \bar{\gamma}$. The solid curve is γ. The figure shows that the difference between γ and $\bar{\gamma}$ becomes pronounced for values $\bar{\gamma} \geq 1.2$. This suggests that the above design procedure of interchanging $\bar{\gamma}$ and γ is most useful for $\bar{\gamma}$ in the interval $1.2 \leq \bar{\gamma} \leq 2$. The question now is why use a larger γ for designing an $H_\infty-$filter of level $\bar{\gamma}$?

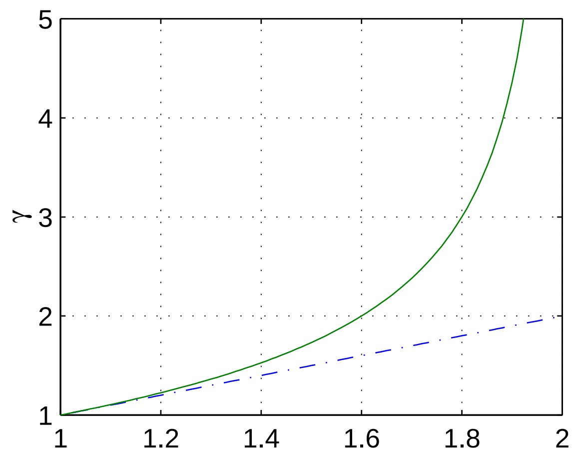

Figure 1: *Comparison of the values of γ and $\bar{\gamma}$ as related by (7). The solid curve shows how γ varies as a function of $\bar{\gamma}$.*

In an attempt to answer the above question, the authors introduce the H_2 (or mean-square) performance that is associated with an H_∞ filter of level γ, and which is given by (cf. Eqs. (6)–(7) in the paper):

$$J_2(\gamma) = \frac{1}{T}\sum_{t=0}^{T-1} \varphi(t)' \left[A(t)P_2(t)A(t)' \;+\; K(t)K(t)'\right]\varphi(t), \tag{8}$$

where $A(t) = I - K(t)\varphi(t)'$ and $P_2(t)$ satisfies

$$P_2(t+1) = A(t)P_2(t)A(t)' \;+\; qI \;+\; K(t)K(t)', \;\; P_2(0) = P_0. \tag{9}$$

Here $K(t)$ is still given by (4) in terms of the Riccati variable $\Pi(t)$. We are writing $J_2(\gamma)$ instead of J_2 in order to emphasize the fact that it is a function of γ (through $\Pi(t)$).

Now the authors' motivation in the paper for proposing the use of a larger γ than the desired $\bar{\gamma}$ is the following. If the above H_2 performance could be shown to decrease with γ, then by using a larger γ one not only guarantees an H_∞-filter with the desired robustness level $\bar{\gamma}$, but one also obtains an H_∞ filter with a better H_2 performance. The validity of such a conjecture, however, is left unanswered in the paper for general models. Right before the beginning of Section 4 the authors state: "It would be tempting to argue that the actual mean-square performance of the central H_∞-filter is monotonic (i.e., decreasing with γ) as well. As a matter of fact, such a property can be proven in the scalar parameter case $n = 1$ or in the case of constant regressor vector, but its general validity is still an open issue." Also, in the last paragraph of the discussion in Section 5, it is stated that "In general, an increase of γ is expected to improve the mean-square-performance at the cost of decreased robustness."

We believe that this issue can be clarified analytically. In any case, let us assume that the authors' conjecture is valid in general so that the H_2-performance of the H_∞-filter (2)–(5) can be expected to decrease with increasing values of γ. The natural question then is whether the improvement can be significant to justify the procedure. Figure 2 plots the H_2-performance of an H_∞-filter for different values of γ in the interval $[1, 2)$. The state-space model used was of dimension 10 and the filter was run for 500 iterations each time (with the same random data for each γ). The noise sequences were Gaussian with variances 0.01. Also $q = 0.1$ and $P_0 = 0.5I$. The regressor vectors were generated randomly. This experiment was repeated extensively several times with random data, always exhibiting a similar behavior.

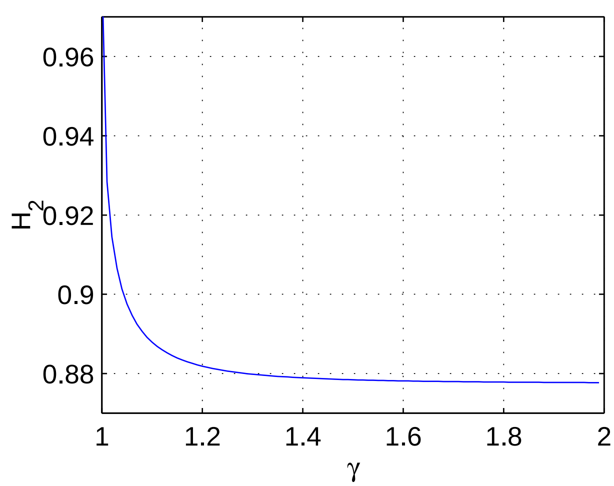

Figure 2: *A plot of the mean-square performance of an H_∞-filter as a function of γ.*

Two remarks follow from this example and from the extensive simulations:

(i) The figure shows that the H_2-performance improves for values of γ that lie approximately in the range $[1, 1.2]$. This is because the H_2-norm seems to decrease rapidly within this interval.

(ii) The figure also shows that for values of γ larger than 1.2 there is minor improvement in the H_2-norm of the filter. This is because for these values of γ, the H_2-norm essentially reaches its steady-state value (which coincides with that of the Kalman filter). Observe that in view of the authors' conjecture that the H_2 performance of an H_∞ filter is generally expected to decrease with γ, and in view of the fact that the Kalman filter achieves the smallest (i.e., optimal) H_2 performace for $\gamma = 2$, it is reasonable to expect (or to conjecture again) that the H_2 performance of an H_∞ filter should in general exhibit the typical behavior shown in Fig. 2. We have observed this behavior in many random simulations. It is of course a worthy and interesting issue to investigate analytically.

While remark (i) suggests that a noticeable improvement in H_2-performance can be obtained for values of γ in the interval $[1, 1.2]$, this observation seems to run against the purpose of the procedure proposed in the paper. We saw in Figure 1 that replacing $\bar{\gamma}$ by a higher value γ according to the transformation (7) could make a difference only for values of γ larger than 1.2. But for this range of γ, Figure 2 above suggests that the H_2-performance is essentially at a steady-state value and is therefore practically insensitive to γ. One way around this difficulty is perhaps to devise a tighter bound than the one presented in Theorem 3 of the paper.

We may also mention that the simulation results in the paper use values $\gamma = 1$ and $\gamma = 1.05$, for which the transformation (7) does not lead to noticeable differences between γ and $\bar{\gamma}$, as indicated by Figure 1. The resulting improvement in H_2 performance that is reported by the authors is not therefore due to the proposed transformation (7) but rather to the inherent sensitivity of the H_2 performance to changes in γ around unity, as indicated in the example of Figure 2. The above discussion suggests that a more convincing example may not exist.

In all, this is a nicely written article that raises some issues that deserve closer examination.