9

Robust Prediction, Filtering and Smoothing

9.1 Introduction

The previously-discussed optimum predictor, filter and smoother solutions assume that the model parameters are correct, the noise processes are Gaussian and their associated covariances are known precisely. These solutions are optimal in a mean-square-error sense, that is they provide the best average performance. If the above assumptions are correct, then the filter's mean-square-error equals the trace of design error covariance. The underlying modelling and noise assumptions are a often convenient fiction. They do, however, serve to allow estimated performance to be weighed against implementation complexity.

In general, robustness means "the persistence of a system's characteristic behaviour under perturbations or conditions of uncertainty" [1]. In an estimation context, robust solutions refer to those that accommodate uncertainties in problem specifications. They are also known as worst-case or peak error designs. The standard predictor, filter and smoother structures are retained but a larger design error covariance is used to account for the presence of modelling error.

Designs that cater for worst cases are likely to exhibit poor average performance. Suppose that a bridge designed for average loading conditions returns an acceptable cost benefit. Then a design that is focussed on accommodating infrequent peak loads is likely to provide degraded average cost performance. Similarly, a worst-case shoe design that accommodates rarely occurring large feet would provide poor fitting performance on average. That is, robust designs tend to be conservative. In practice, a trade-off may be desired between optimum and robust designs.

The material canvassed herein is based on the H_∞ filtering results from robust control. The robust control literature is vast, see [2] – [33] and the references therein. As suggested above, the H_∞ solutions of interest here involve observers having gains that are obtained by solving Riccati equations. This Riccati equation solution approach relies on the Bounded Real Lemma – see the pioneering work by Vaidyanathan [2] and Petersen [3]. The Bounded Real Lemma is implicit with game theory [9] – [19]. Indeed, the continuous-time solutions presented in this section originate from the game theoretic approach of Doyle, Glover, Khargonekar, Francis Limebeer, Anderson, Khargonekar, Green, Theodore and Shaked, see [4], [13], [15], [21]. The discussed discrete-time versions stem from the results of Limebeer, Green, Walker, Yaesh, Shaked, Xie, de Souza and Wang, see [5], [11], [18], [19], [21]. In the parlance of game theory: "a statistician is trying to best estimate a linear combination of the states of a system that is driven by nature; nature is trying to cause the statistician's estimate

---

"On a huge hill, Cragged, and steep, Truth stands, and he that will Reach her, about must, and about must go." *John Donne* to be as erroneous as possible, while trying to minimize the energy it invests in driving the system" [19].

Pertinent state-space H_∞ predictors, filters and smoothers are described in [4] – [19]. Some prediction, filtering and smoothing results are summarised in [13] and methods for accommodating model uncertainty are described in [14], [18], [19]. The aforementioned methods for handling model uncertainty can result in conservative designs (that depart far from optimality). This has prompted the use of linear matrix inequality solvers in [20], [23] to search for optimal solutions to model uncertainty problems.

It is explained in [15], [19], [21] that a saddle-point strategy for the games leads to robust estimators, and the resulting robust smoothing, filtering and prediction solutions are summarised below. While the solution structures remain unchanged, designers need to tweak the scalar within the underlying Riccati equations.

This chapter has two main parts. Section 9.2 describes robust continuous-time solutions and the discrete-time counterparts are presented in Section 9.3. The previously discussed techniques each rely on a trick. The optimum filters and smoothers arise by completing the square. In maximum-likelihood estimation, a function is differentiated with respect to an unknown parameter and then set to zero. The trick behind the described robust estimation techniques is the Bounded Real Lemma, which opens the discussions.

9.2 Robust Continuous-time Estimation

9.2.1 Continuous-Time Bounded Real Lemma

First, consider the unforced system

$$\dot{x}(t) = A(t)x(t) \tag{1}$$

over a time interval $t \in [0, T]$, where $A(t) \in \mathbb{R}^{n\times n}$. For notational convenience, define the stacked vector $x = \{x(t),\ t \in [0, T]\}$. From Lyapunov stability theory [36], the system (1) is asymptotically stable if there exists a function $V(x(t)) > 0$ such that $\dot{V}(x(t)) < 0$. A possible Lyapunov function is $V(x(t)) = x^T(t)P(t)x(t)$, where $P(t) = P^T(t) \in \mathbb{R}^{n\times n}$ is positive definite. To ensure $x \in \mathcal{L}_2$ it is required to establish that

$$\dot{V}(x(t)) = \dot{x}^T(t)P(t)x(t) + x^T(t)\dot{P}(t)x(t) + x^T(t)P(t)\dot{x}(t) < 0\ . \tag{2}$$

Now consider the output of a linear time varying system, $y = \mathcal{G}w$, having the state-space representation

$$\dot{x}(t) = A(t)x(t) + B(t)w(t)\ , \tag{3}$$

"Uncertainty is one of the defining features of Science. Absolute proof only exists in mathematics. In the real world, it is impossible to prove that theories are right in every circumstance; we can only prove that they are wrong. This provisionality can cause people to lose faith in the conclusions of science, but it shouldn't. The recent history of science is not one of well-established theories being proven wrong. Rather, it is of theories being gradually refined." *New Scientist vol. 212 no. 2835*

$$y(t) = C(t)x(t), \tag{4}$$

where $w(t) \in \mathbb{R}^m$, $B(t) \in \mathbb{R}^{n\times m}$ and $C(t) \in \mathbb{R}^{p\times n}$. Assume temporarily that $E\{w(t)w^T(\tau)\} = I\delta(t-\tau)$. The Bounded Real Lemma [13], [15], [21], states that $w \in \mathcal{L}_2$ implies $y \in \mathcal{L}_2$ if

$$\dot{V}(x(t)) + y^T(t)y(t) - \gamma^2 w^T(t)w(t) < 0 \tag{5}$$

for a $\gamma \in \mathbb{R}$. Integrating (5) from $t = 0$ to $t = T$ gives

$$\int_0^T \dot{V}(x(t))\ dt + \int_0^T y^T(t)y(t)\ dt - \gamma^2 \int_0^T w^T(t)w(t)\ dt < 0 \tag{6}$$

and noting that $\int_0^T \dot{V}(x(t))\ dt = x^T(T)P(T)x(T) - x^T(0)P(0)x(0)$, another objective is

$$\frac{x^T(T)P(T)x(T) - x^T(0)P(0)x(0) + \int_0^T y^T(t)y(t)\ dt}{\int_0^T w^T(t)w(t)\ dt} \le \gamma^2 . \tag{7}$$

Under the assumptions $x(0) = 0$ and $P(T) = 0$, the above inequality simplifies to

$$\frac{\left\|y(t)\right\|_2^2}{\left\|w(t)\right\|_2^2} = \frac{\int_0^T y^T(t)y\ (t)dt}{\int_0^T w^T(t)w(t)\ dt} \le \gamma^2 . \tag{8}$$

The ∞-norm of $\mathcal{G}$ is defined as

$$\left\|\mathcal{G}\right\|_\infty = \frac{\left\|y\right\|_2}{\left\|w\right\|_2} = \frac{\left\|\mathcal{G}w\right\|_2}{\left\|w\right\|_2} . \tag{9}$$

The Lebesgue ∞-space is the set of systems having finite ∞-norm and is denoted by $\mathcal{L}_\infty$. That is, $\mathcal{G} \in \mathcal{L}_\infty$, if there exists a $\gamma \in \mathbb{R}$ such that

$$\sup_{\left\|w\right\|_2 \ne 0} \left\|\mathcal{G}\right\|_\infty = \sup_{\left\|w\right\|_2 \ne 0} \frac{\left\|y\right\|_2}{\left\|w\right\|_2} \le \gamma , \tag{10}$$

namely, the supremum (or maximum) ratio of the output and input 2-norms is finite. The conditions under which $\mathcal{G} \in \mathcal{L}_\infty$ are specified below. The accompanying sufficiency proof combines the approaches of [15], [31]. A further five proofs for this important result appear in [21].

Lemma 1: *The continuous-time Bounded Real Lemma [15], [13], [21]: In respect of the above system* $\mathcal{G}$*, suppose that the Riccati differential equation*

$$-\dot{P}(t) = P(t)A(t) + A^T(t)P(t) + C^T(t)C(t) + \gamma^{-2}P(t)B(t)B^T(t)P(t)\ , \tag{11}$$

has a solution on [0, T]. Then $\left\|\mathcal{G}\right\|_\infty \le \gamma$ *for any* $w \in \mathcal{L}_2$.

"Information is the resolution of uncertainty." *Claude Elwood Shannon*

Proof: *From (2) – (5),*

$$\dot{V}(t) + y^T(t)y(t) - \gamma^2 w^T(t)w(t)$$

$$= x^T(t)C^T(t)C(t)x(t) - \gamma^2 w^T(t)w(t) + x^T(t)\dot{P}(t)x(t)$$

$$+(A(t)x(t) + B(t)w(t))^T P(t)x(t) \ +x^T(t)P(t)(A(t)x^T(t) + B(t)w(t))$$

$$= \gamma^{-2}x^T(t)P(t)B(t)B^T(t)P(t)x(t) - \gamma^2 w^T(t)w(t)) \ +w^T(t)B^T(t)P(t)x(t) + x^T(t)P(t)B(t)w(t)$$

$$= -\gamma^2(w(t) - \gamma^{-2}B(t)P(t)x(t))^T(w(t) - \gamma^{-2}B(t)P(t)x(t)),$$

which implies (6) and (7). Inequality (8) is established under the assumptions x(0) = 0 and P(T) = 0.

□

In general, where $E\{w(t)w^T(\tau)\} = Q(t)\delta(t-\tau)$, the scaled matrix $\overline{B}(t) = B(t)Q^{1/2}(t)$ may be used in place of $B(t)$ above. When the plant $\mathcal{G}$ has a direct feedthrough matrix, that is,

$$y(t) = C(t)x(t) + D(t)w(t), \tag{12}$$

$D(t) \in \mathbb{R}^{p\times m}$, the above Riccati differential equation is generalised to

$$\begin{aligned} -\dot{P}_1(t) = P_1(t)(A(t) + B(t)M^{-1}(t)D^T(t)C(t)) \ +(A(t) + B(t)M^{-1}(t)D^T(t)C(t))^T P_1(t) \\ +\gamma^{-2}B(t)M^{-1}(t)B^T(t) + C^T(t)(I + D(t)M^{-1}(t)D^T(t))C(t), \end{aligned} \tag{13}$$

where $M(t) = \gamma^2 I - D^T(t)D(t) > 0$. A proof is requested in the problems.

Criterion (8) indicates that the ratio of the system's output and input energies is bounded above by γ^2 for any $w \in \mathcal{L}_2$, including worst-case w. Consequently, solutions satisfying (8) are often called worst-case designs.

9.2.2 Continuous-Time H∞ Filtering

9.2.2.1 Problem Definition

Now that the Bounded Real Lemma has been defined, the H∞ filter can be set out. The general filtering problem is depicted in Fig. 1. It is assumed that the system $\mathcal{G}_2$ has the state-space realisation

$$\dot{x}(t) = A(t)x(t) + B(t)w(t), \quad x(0) = 0, \tag{14}$$

$$y_2(t) = C_2(t)x(t). \tag{15}$$

Suppose that the system $\mathcal{G}_1$ has the realisation (14) and

$$y_1(t) = C_1(t)x(t). \tag{16}$$

"All exact science is dominated by the idea of approximation." *Earl Bertrand Arthur William*

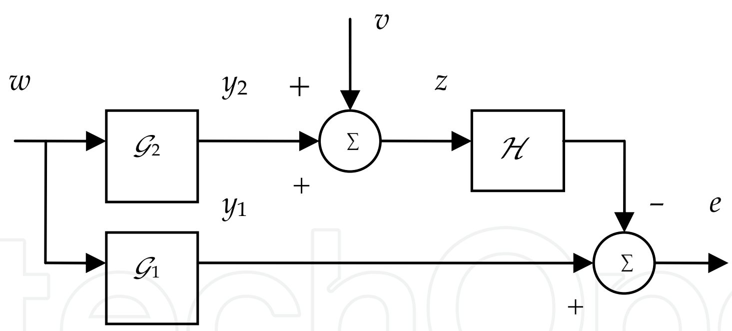


Figure 1. The general filtering problem. The objective is to estimate the output of $\mathcal{G}_1$ from noisy measurements of $\mathcal{G}_2$.

It is desired to find a causal solution $\mathcal{H}$ that produces estimates $\hat{y}_1(t\,|\,t)$ of $y_1(t)$ from the measurements,

$$z(t) = y_2(t) + v(t)\ , \tag{17}$$

at time t so that the output estimation error,

$$e(t\,|\,t) = y_1(t) - \hat{y}_1(t\,|\,t)\ , \tag{18}$$

is in $\mathcal{L}_2$. The error signal (18) is generated by a system denoted by $e = \mathcal{R}_{ei}$, where $i = \begin{bmatrix} v \\ w \end{bmatrix}$ and $\mathcal{R}_{ei} = -[\mathcal{H} \quad \mathcal{H}\mathcal{G}_2 - \mathcal{G}_1]$. Hence, the objective is to achieve $\int_0^T e^T(t\,|\,t)e(t\,|\,t)\ dt$ – $\gamma^2\int_0^T i^T(t)i(t)\ dt$ < 0 for some $\gamma \in \mathbb{R}$. For convenience, it is assumed here that $w(t) \in \mathbb{R}^m$, $E\{w(t)\} = 0$, $E\{w(t)w^T(\tau)\} = Q(t)\delta(t-\tau)$, $v(t) \in \mathbb{R}^p$, $E\{v(t)\} = 0$, $E\{v(t)v^T(\tau)\} = R(t)\delta(t-\tau)$ and $E\{w(t)v^T(\tau)\} = 0$.

9.2.2.2 H$_\infty$ Solution

A parameterisation of all solutions for the H$_\infty$ filter is developed in [21]. A minimum-entropy filter arises when the contractive operator within [21] is zero and is given by

$$\dot{\hat{x}}(t\,|\,t) = \big(A(t) - K(t)C_2(t)\big)\hat{x}(t\,|\,t) + K(t)z(t),\quad \hat{x}(0) = 0, \tag{19}$$

$$\hat{y}_1(t\,|\,t) = C_1(t\,|\,t)\hat{x}(t\,|\,t), \tag{20}$$

where

$$K(t) = P(t)C_2^T(t)R^{-1}(t) \tag{21}$$

"Uncertainty and expectation are the joys of life. Security is an insipid thing, through the overtaking and possessing of a wish discovers the folly of the chase." *William Congreve* is the filter gain and $P(t) = P^T(t) > 0$ is the solution of the Riccati differential equation

$$\begin{aligned}\dot{P}(t) &= A(t)P(t) + P(t)A^T(t) + B(t)Q(t)B^T(t) \\ &\quad - P(t)(C_2^T(t)R^{-1}(t)C_2(t) - \gamma^{-2}C_1^T(t)C_1(t))P(t),\ P(0) = 0.\end{aligned} \tag{22}$$

It can be seen that the H_∞ filter has a structure akin to the Kalman filter. A point of difference is that the solution to the above Riccati differential equation solution depends on $C_1(t)$, the linear combination of states being estimated.

9.2.2.3 Properties

Define $\bar{A}(t) = A(t) - K(t)C_2(t)$. Subtracting (19) – (20) from (14) – (15) yields the error system

$$\begin{aligned}\begin{bmatrix}\dot{\tilde{x}}(t|t) \\ e(t|t)\end{bmatrix} &= \begin{bmatrix}\bar{A}(t) & [-K(t) & B(t)] \\ C_1(t) & [0 & 0]\end{bmatrix}\begin{bmatrix}\tilde{x}(t|t) \\ \begin{bmatrix}v(t) \\ w(t)\end{bmatrix}\end{bmatrix},\ \tilde{x}(0) = 0, \\ &= \mathcal{R}_{ei}i,\end{aligned} \tag{23}$$

where $\tilde{x}(t|t) = x(t) - \hat{x}(t|t)$ and $\mathcal{R}_{ei} = \begin{bmatrix}\bar{A}(t) & [-K(t) & B(t)] \\ C_1(t) & [0 & 0]\end{bmatrix}$. The adjoint of $\mathcal{R}_{ei}$ is given by $\mathcal{R}_{ei}^H = \begin{bmatrix}-\bar{A}^T(t) & C_1^T(t) \\ \begin{bmatrix}K^T(t) \\ -B^T(t)\end{bmatrix} & \begin{bmatrix}0 \\ 0\end{bmatrix}\end{bmatrix}$. It is shown below that the estimation error satisfies the desired performance objective.

Lemma 2: *In respect of the H_∞ problem (14) – (18), the solution (19) – (20) achieves the performance* $x^T(T)P(T)x(T) - x^T(0)P(0)x(0) + \int_0^T e^T(t|t)e(t|t)\,dt - \gamma^2\int_0^T i^T(t)i(t)\,dt < 0$.

Proof: *Following the approach in [15], [21], by applying Lemma 1 to the adjoint of (23), it is required that there exists a positive definite symmetric solution to*

$$\begin{aligned}-\dot{P}(\tau) &= \bar{A}(\tau)P(\tau) + P(\tau)\bar{A}^T(\tau) + B(\tau)Q(\tau)B^T(\tau) + K(\tau)R(\tau)K^T(\tau) \\ &\quad + \gamma^{-2}P(\tau)C_1^T(\tau)C_1(\tau))P(\tau),\quad P(\tau)\big|_{\tau=T} = 0,\end{aligned}$$

"Although economists have studied the sensitivity of import and export volumes to changes in the exchange rate, there is still much uncertainty about just how much the dollar must change to bring about any given reduction in our trade deficit." *Martin Stuart Feldstein* *on [0, T] for some* $\gamma \in \mathbb{R}$*, in which* $\tau = T - t$ *is a time-to-go variable. Substituting* $K(\tau) = P(\tau)C_2^T(\tau)R^{-1}(\tau)$ *into the above Riccati differential equation yields*

$$-\dot{P}(\tau) = A(\tau)P(\tau) + P(\tau)A^T(\tau) + B(\tau)Q(\tau)B^T(\tau) - P(\tau)(C_2^T(\tau)R^{-1}(\tau)(C_2(\tau) - \gamma^{-2}C_1^T(\tau)C_1(\tau))P(\tau),$$

$$P(\tau)\big|_{\tau=T} = 0.$$

Taking adjoints to address the problem (23) leads to (22), for which the existence of a positive define solution implies $x^T(T)P(T)x(T) - x^T(0)P(0)x(0) + \int_0^T e^T(t|t)e(t|t)\,dt - \gamma^2\int_0^T i^T(t)i(t)\,dt < 0$*. Thus, under the assumption* $x(0) = 0$*,* $\int_0^T e^T(t|t)e(t|t)\,dt - \gamma^2\int_0^T i^T(t)i(t)\,dt < -x^T(T)P(T)x(T) < 0$*. Therefore,* $\mathcal{R}_{ei} \in \mathcal{L}_\infty$*, that is,* $w, v \in \mathcal{L}_2 \Rightarrow e \in \mathcal{L}_2$*.* □

9.2.2.4 Trading-Off H∞ Performance

In a robust filter design it is desired to meet an H∞ performance objective for a minimum possible γ. A minimum γ can be found by conducting a search and checking for the existence of positive definite solutions to the Riccati differential equation (22). This search is tractable because $\dot{P}(t)$ is a convex function of γ^2, since $\dfrac{\partial^2 \dot{P}(t)}{\partial^2\gamma^2} = \gamma^{-6}P(t)C_1^T(t)C_1(t))P(t) > 0$.

In some applications it may be possible to estimate *a priori* values for γ. Recall for output estimation problems that the error is generated by $e = \mathcal{R}_{ei}i$, where $\mathcal{R}_{ei} = -[\mathcal{H} \quad (\mathcal{H} - I)\mathcal{G}_1]$. From the arguments of Chapters 1 – 2 and [28], for single-input-single-output plants $\lim_{\sigma_v^2 \to 0}|\mathcal{H}| = 1$ and $\lim_{\sigma_v^2 \to 0}|\mathcal{R}_{ei}| = 1$, which implies $\lim_{\sigma_v^2 \to 0}\left\|\mathcal{R}_{ei}\mathcal{R}_{ei}^H\right\|_\infty = \sigma_v^2$. Since the H∞ filter achieves the performance $\left\|\mathcal{R}_{ei}\mathcal{R}_{ei}^H\right\|_\infty < \gamma^2$, it follows that an *a priori* design estimate is $\gamma = \sigma_v$ at high signal-to-noise-ratios.

When the problem is stationary (or time-invariant), the filter gain is precalculated as $K = PC^TR^{-1}$, where P is the solution of the algebraic Riccati equation

$$0 = AP + PA^T - P(C_2^TR^{-1}C_2 - \gamma^{-2}C_1^TC_1)P + BQB^T. \tag{24}$$

Suppose that $\mathcal{G}_2 = \mathcal{G}_1$ is a time-invariant single-input-single-output system and let $R_{ei}(s)$ denote the transfer function of $\mathcal{R}_{ei}$. Then Parseval's Theorem states that the average total energy of $e(t|t)$ is

$$\left\|e(t|t)\right\|_2^2 = \int_{-\infty}^{-\infty}\left|e(t|t)\right|^2 dt = \int_{-j\infty}^{j\infty} R_{ei}R_{ei}^H(s)\,ds = \int_{-j\infty}^{j\infty}\left|R_{ei}(s)\right|^2 ds, \tag{25}$$

"Life is uncertain. Eat dessert first." *Ernestine Ulmer* which equals the area under the error power spectral density, $R_{ei}R_{ei}^H(s)$. Recall that the optimal filter (in which $\gamma = \infty$) minimises (25), whereas the H∞ filter minimises

$$\sup_{\|i\|_2 \neq 0} \left\| R_{ei} R_{ei}^H \right\|_\infty = \sup_{\|i\|_2 \neq 0} \frac{\|e\|_2^2}{\|i\|_2^2} < \gamma^2 . \tag{26}$$

In view of (25) and (26), it follows that the H∞ filter minimises the maximum magnitude of $R_{ei}R_{ei}^H(s)$. Consequently, it is also called a 'minimax filter'. However, robust designs, which accommodate uncertain inputs tend to be conservative. Therefore, it is prudent to investigate using a larger γ to achieve a trade-off between H∞ and minimum-mean-square-error performance criteria.

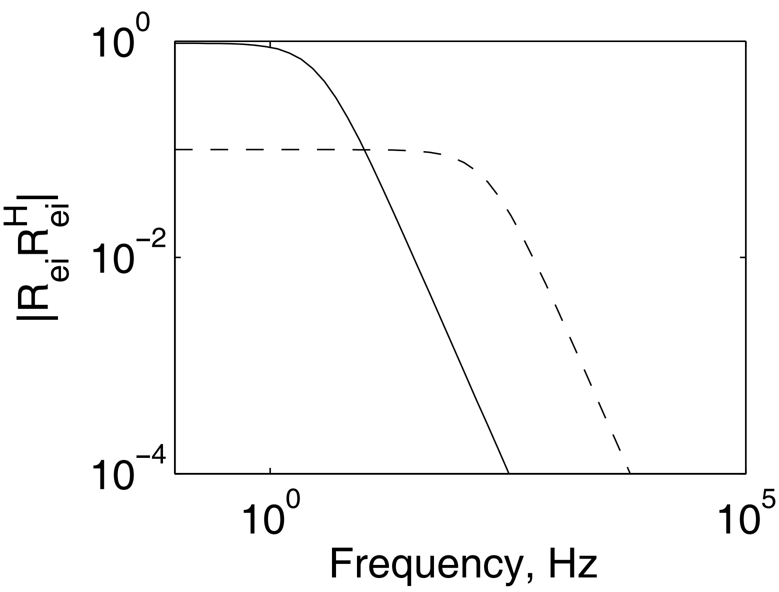


Figure 2. $\left|R_{ei} R_{ei}^H(s)\right|$ versus frequency for Example 1: optimal filter (solid line) and H∞ filter (dotted line).

Example 1. Consider a time-invariant output estimation problem where A = -1, B = C_2 = C_1 = 1, σ_w^2 = 10 and σ_v^2 = 0.1. The magnitude of the error spectrum exhibited by the optimal filter (designed with $\gamma^2 = 10^8$) is indicated by the solid line of Fig. 2. From a search, a minimum of $\gamma^2 = 0.099$ was found such that the algebraic Riccati equation (24) has a positive definite solution, which concurs with the *a priori* estimate of $\gamma^2 \approx \sigma_v^2$. The magnitude of the error spectrum exhibited by the H∞ filter is indicated by the dotted line of Fig. 2. The figure demonstrates that the filter achieves $\left|R_{ei}R_{ei}^H(s)\right| < \gamma^2$. Although the H∞ filter reduces the peak of the error spectrum by 10 dB, it can be seen that the area under the curve is larger, that is, the mean square error increases. Consequently, some intermediate value of γ may need to be considered to trade off peak error (spectrum) and average error performance.

"If the uncertainty is larger than the effect, the effect itself becomes moot." *Patrick Frank*

9.2.3 Accommodating Uncertainty

The above filters are designed for situations in which the inputs $v(t)$ and $w(t)$ are uncertain. Next, problems in which model uncertainty is present are discussed. The described approaches involve converting the uncertainty into a fictitious noise source and solving an auxiliary H_∞ filtering problem.

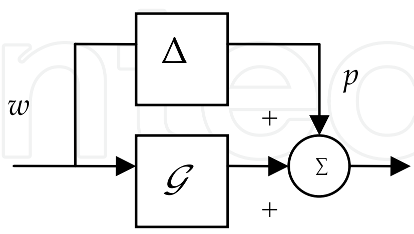


Figure 3. Representation of additive model uncertainty.

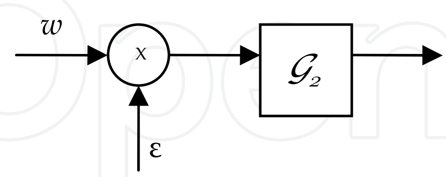


Figure 4. Input scaling in lieu of a problem that possesses an uncertainty.

9.2.3.1 Additive Uncertainty

Consider a time-invariant output estimation problem in which the nominal model is $\mathcal{G}_2 + \Delta$, where $\mathcal{G}_2$ is known and Δ is unknown, as depicted in Fig. 3. The $p(t)$ represents a fictitious signal to account for discrepancies due to the uncertainty. It is argued below that a solution to the H_∞ filtering problem can be found by solving an auxiliary problem in which the input is scaled by $\varepsilon \in \mathbb{R}$ as shown in Fig. 4. In lieu of the filtering problem possessing the uncertainty Δ, an auxiliary problem is defined as

$$\dot{x}(t) = Ax(t) + Bw(t) + Bp(t)\,,\ x(0) = 0\,, \tag{27}$$

$$z(t) = C_2(t)x(t) + v(t)\,, \tag{28}$$

$$e(t\,|\,t) = C_1(t)x(t) - C_1(t)\hat{x}(t)\,, \tag{29}$$

where $p(t)$ is an additional exogenous input satisfying

$$\left\|p\right\|_2^2 < \delta^2 \left\|w\right\|_2^2,\ \ \delta \in \mathbb{R}\,. \tag{30}$$

Consider the scaled H_∞ filtering problem where

$$\dot{x}(t) = Ax(t) + B\varepsilon w(t)\,,\ x(0) = 0\,, \tag{31}$$

$$z(t) = C_2(t)x(t) + v(t)\,, \tag{32}$$

$$e(t\,|\,t) = C_1(t)x(t) - C_1(t)\hat{x}(t)\,, \tag{33}$$

in which $\varepsilon^2 = (1 + \delta^2)^{-1}$.

"A theory has only the alternative of being right or wrong. A model has a third possibility - it may be right but irrelevant." *Manfred Eigen*

Lemma 3 [26]: *Suppose for a $\gamma \neq 0$ that the scaled H_∞ problem (31) – (33) is solvable, that is, $\|e\|_2^2$ < $\gamma^2(\varepsilon^{-2}\|w\|_2^2$ + $\|v\|_2^2)$. Then, this guarantees the performance*

$$\|e\|_2^2 < \gamma^2(\|w\|_2^2 + \|p\|_2^2 + \|v\|_2^2) \tag{34}$$

for the solution of the auxiliary problem (27) – (29).

Proof: *From the assumption that problem (31) – (33) is solvable, it follows that $\|e\|_2^2$ < $\gamma^2(\varepsilon^{-2}\|w\|_2^2$ + $\|v\|_2^2)$. Substituting for ε, using (30) and rearranging yields (34).* □

9.2.4 Multiplicative Uncertainty

Next, consider a filtering problem in which the model is $G(I + \Delta)$, as depicted in Fig. 5. It is again assumed that G and Δ are known and unknown transfer function matrices, respectively. This problem may similarly be solved using Lemma 3. Thus a filter that accommodates additive or multiplicative uncertainty simply requires scaling of an input. The above scaling is only sufficient for a H_∞ performance criterion to be met. The design may well be too conservative and it is worthwhile to explore the merits of using values for δ less than the uncertainty's assumed norm bound.

9.2.5 Parametric Uncertainty

Finally, consider a time-invariant output estimation problem in which the state matrix is uncertain, namely,

$$\dot{x}(t) = (A + \Delta_A)x(t) + Bw(t), \quad x(0) = 0, \tag{35}$$

$$z(t) = C_2(t)x(t) + v(t), \tag{36}$$

$$e(t \mid t) = C_1(t)x(t) - C_1(t)\hat{x}(t), \tag{37}$$

where $\Delta_A \in \mathbb{R}^{n \times n}$ is unknown. Define an auxiliary H_∞ filtering problem by

$$\dot{x}(t) = Ax(t) + Bw(t) + p(t), \quad x(0) = 0, \tag{38}$$

(36) and (37), where $p(t) = \Delta_A x(t)$ is a fictitious exogenous input. A solution to this problem would achieve

$$\|e\|_2^2 < \gamma^2(\|w\|_2^2 + \|p\|_2^2 + \|v\|_2^2) \tag{39}$$

for a $\gamma \neq 0$. From the approach of [14], [18], [19], consider the scaled filtering problem

$$\dot{x}(t) = Ax(t) + \bar{B}\bar{w}(t), \quad x(0) = 0, \tag{40}$$

"Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful." *George Edward Pelham Box*

(36), (37), where $\bar{B} = \begin{bmatrix} B & \varepsilon^{-1} \end{bmatrix}$, $\bar{w}(t) = \begin{bmatrix} w(t) \\ \varepsilon p(t) \end{bmatrix}$ and 0 < ε < 1. Then the solution of this H∞ filtering problem satisfies

$$\|e\|_2^2 < \gamma^2 (\|w\|_2^2 + \varepsilon^2 \|p\|_2^2 + \|v\|_2^2) , \tag{41}$$

which implies (39). Thus, state matrix parameter uncertainty can be accommodated by including a scaled input in the solution of an auxiliary H∞ filtering problem. Similar solutions to problems in which other state-space parameters are uncertain appear in [14], [18], [19].

9.2.6 Continuous-Time H∞ Smoothing

9.2.6.1 Background

There are three kinds of H∞ smoothers: fixed point, fixed lag and fixed interval (see the tutorial [13]). The next development is concerned with continuous-time H∞ fixed-interval smoothing. The smoother in [10] arises as a combination of forward states from an H∞ filter and adjoint states that evolve according to a Hamiltonian matrix. A different fixed-interval smoothing problem to [10] is found in [16] by solving for saddle conditions within differential games. A summary of some filtering and smoothing results appears in [13]. Robust prediction, filtering and smoothing problems are addressed in [22]; the H∞ predictor, filter and smoother require the solution of a Riccati differential equation that evolves forward in time, whereas the smoother additionally requires another to be solved in reverse-time. Another approach for combining forward and adjoint estimates is described [32] where the Fraser-Potter formula is used to construct a smoothed estimate.

Continuous-time, fixed-interval smoothers that differ from the formulations within [10], [13], [16], [22], [32] are reported in [34] – [35]. A robust version of [34] – [35] appears in [33], which is described below.

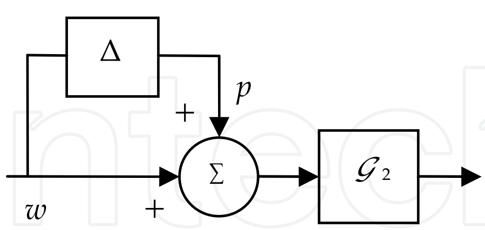


Figure 5. Representation of multiplicative model uncertainty.

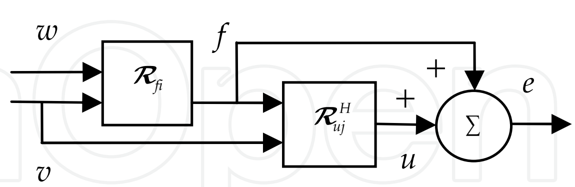


Figure 6. Robust smoother error structure.

"The purpose of models is not to fit the data but to sharpen the questions." *Samuel Karlin*

9.2.6.2 Problem Definition

Once again, it is assumed that the data is generated by (14) – (17). For convenience, attention is confined to output estimation, namely $\mathcal{G}_2$ = $\mathcal{G}_1$ within Fig. 1. Input and state estimation problems can be handled similarly using the solution structures described in Chapter 6. It is desired to find a fixed-interval smoother solution $\mathcal{H}$ that produces estimates $\hat{y}_1(t|T)$ of $y_1(t)$ so that the output estimation error

$$e(t|T) = y_1(t) - \hat{y}_1(t|T) \tag{42}$$

is in $\mathcal{L}_2$. As before, the map from the inputs i = $\begin{bmatrix} v \\ w \end{bmatrix}$ to the error is denoted by $\mathcal{R}_{ei}$ = $-[\mathcal{H} \quad \mathcal{H}\mathcal{G}_1 - \mathcal{G}_1]$ and the objective is to achieve $\int_0^T e^T(t|T)e(t|T)\ dt$ – $\gamma^2 \int_0^T i^T(t)i(t)\ dt$ < 0 for some $\gamma \in \mathbb{R}$.

9.2.6.3 H∞ Solution

The following H∞ fixed-interval smoother exploits the structure of the minimum-variance smoother but uses the gain (21) calculated from the solution of the Riccati differential equation (22) akin to the H∞ filter. An approximate Wiener-Hopf factor inverse, $\hat{\Delta}^{-1}$, is given by

$$\begin{bmatrix} \dot{\hat{x}}(t|t) \\ \alpha(t) \end{bmatrix} = \begin{bmatrix} A(t) - K(t)C(t) & K(t) \\ -R^{-1/2}(t)C(t) & R^{-1/2}(t) \end{bmatrix} \begin{bmatrix} \hat{x}(t|t) \\ z(t) \end{bmatrix}. \tag{43}$$

An inspection reveals that the states within (43) are the same as those calculated by the H∞ filter (19). The adjoint of $\hat{\Delta}^{-1}$, which is denoted by $\hat{\Delta}^{-H}$, has the realisation

$$\begin{bmatrix} -\dot{\xi}(t) \\ \beta(t) \end{bmatrix} = \begin{bmatrix} A^T(t) - C^T(t)K^T(t) & -C^T(t)R^{-1/2}(t) \\ K^T(t) & R^{-1/2}(t) \end{bmatrix} \begin{bmatrix} \xi(t) \\ \alpha(t) \end{bmatrix}. \tag{44}$$

Output estimates are obtained as

$$\hat{y}(t|T) = z(t) - R(t)\beta(t). \tag{45}$$

However, an additional condition requires checking in order to guarantee that the smoother actually achieves the above performance objective; the existence of a solution $P_2(t)$ = $P_2^T(t)$ > 0 is required for the auxiliary Riccati differential equation

$$\begin{aligned} -\dot{P}_2(t) &= \bar{A}(t)P_2(t) + P_2(t)\bar{A}^T(t) + K(t)R^2(t)K^T(t) \\ &\quad + \gamma_2^{-2}P_2(t)C^T(t)R^{-1}(t)\big(C(t)P(t)C^T(t) + R(t)\big)R^{-1}(t)C(t)P_2(t),\ P_2(T) = 0, \end{aligned} \tag{46}$$

where $\bar{A}(t) = A(t) - K(t)C(t)$.

"Certainty is the mother of quiet and repose, and uncertainty the cause of variance and contentions." *Edward Coke*

9.2.7 Performance

It will be shown subsequently that the robust fixed-interval smoother (43) – (45) has the error structure shown in Fig. 6, which is examined below.

Lemma 4 [33]: *Consider the arrangement of two linear systems* $f = \mathcal{R}_{fi} i$ *and* $u = \mathcal{R}_{uj}^H j$ *shown in Fig. 6, in which* $i = \begin{bmatrix} w \\ v \end{bmatrix}$ *and* $j = \begin{bmatrix} f \\ v \end{bmatrix}$. *Let* $\mathcal{R}_{ei}$ *denote the map from i to e. Assume that w and v* $\in \mathcal{L}_2$. *If and only if: (i)* $\mathcal{R}_{fi} \in \mathcal{L}_\infty$ *and (ii)* $\mathcal{R}_{uj}^H \in \mathcal{L}_\infty$, *then (i) f, u, e* $\in \mathcal{L}_2$ *and (ii)* $\mathcal{R}_{ei} \in \mathcal{L}_\infty$.

Proof: *(i) To establish sufficiency, note that* $\|i\|_2 \le \|w\|_2 + \|v\|_2$ => $d \in \mathcal{L}_2$, *which with Condition (i) =>* $f \in \mathcal{L}_2$. *Similarly,* $\|j\|_2 \le \|f\|_2 + \|v\|_2$ => $j \in \mathcal{L}_2$, *which with Condition (ii) =>* $u \in \mathcal{L}_2$. *Also,* $\|e\|_2 \le \|f\|_2 + \|u\|_2$ => $e \in \mathcal{L}_2$. *The necessity of (i) follows from the assumption* $i \in \mathcal{L}_2$ *together with the property* $\mathcal{R}_{fi} \mathcal{L}_2 \subset \mathcal{L}_2$ => $\mathcal{R}_{fi} \in \mathcal{L}_\infty$ *(see [p. 83, 21]). Similarly,* $j \in \mathcal{L}_2$ *together with the property* $\mathcal{R}_{uj}^H \mathcal{L}_2 \subset \mathcal{L}_2$ => $\mathcal{R}_{uj}^H \in \mathcal{L}_\infty$.

(ii) Finally, $i \in \mathcal{L}_2$, $e = \mathcal{R}_{ei} i \in \mathcal{L}_2$ *together with the property* $\mathcal{R}_{ei} \mathcal{L}_2 \subset \mathcal{L}_2$ => $\mathcal{R}_{ei} \in \mathcal{L}_\infty$. □

It is easily shown that the error system, $\mathcal{R}_{ei}$, for the model (14) – (15), the data (17) and the smoother (43) – (45), is given by

$$\begin{bmatrix} \dot{\tilde{x}}(t) \\ -\dot{\xi}(t) \\ e(t|T) \end{bmatrix} = \begin{bmatrix} \bar{A}(t) & 0 & B(t) & -K(t) \\ -C^T(t)R^{-1}(t)C(t) & \bar{A}^T(t) & 0 & -C^T(t)R^{-1}(t) \\ C(t) & R(t)K^T(t) & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ \xi(t) \\ w(t) \\ v(t) \end{bmatrix}, \tilde{x}(0) = 0, \quad (47)$$
$$\xi(T) = 0,$$

where $\tilde{x}(t|t) = x(t) - \hat{x}(t|t)$. The conditions for the smoother attaining the desired performance objective are described below.

Lemma 5 [33]: *In respect of the smoother error system (47), if there exist symmetric positive define solutions to (22) and (46) for* γ, $\gamma_2 > 0$, *then the smoother (43) – (45) achieves* $\mathcal{R}_{ei} \in \mathcal{L}_\infty$, *that is,* $i \in \mathcal{L}_2$ *implies* $e \in \mathcal{L}_2$.

Proof: *Since* $\tilde{x}(t|t)$ *is decoupled from ξ(t),* $\mathcal{R}_{ei}$ *is equivalent to the arrangement of two systems* $\mathcal{R}_{fi}$ *and* $\mathcal{R}_{uj}^H$ *shown in Fig. 6. The* $\mathcal{R}_{fi}$ *is defined by (23) in which* $C_2(t) = C(t)$. *From Lemma 2, the existence of a positive definite solution to (22) implies* $\mathcal{R}_{fi} \in \mathcal{L}_\infty$. *The* $\mathcal{R}_{uj}^H$ *is given by the system*

$$-\dot{\xi}(\tau) = \bar{A}^T(\tau)\xi(\tau) - C^T(\tau)R^{-1}(\tau)\tilde{y}(\tau|\tau) - C^T(\tau)R^{-1}(\tau)v(\tau), \; \xi(T) = 0, \quad (48)$$

$$u(\tau) = R(\tau)K^T(\tau)\xi(\tau). \quad (49)$$

"Doubt is uncomfortable, certainty is ridiculous." *François-Marie Arouet de Voltaire*

For the above system to be in $\mathcal{L}_\infty$, from Lemma 4, it is required that there exists a solution to (46) for which the existence of a positive definite solution implies $\mathcal{R}_{uj}^H \in \mathcal{L}_\infty$. The claim $\mathcal{R}_{ei} \in \mathcal{L}_\infty$ follows from Lemma 4. □

The H_∞ solution can be derived as a solution to a two-point boundary value problem, which involves a trade-off between causal and noncausal processes (see [10], [15], [21]). This suggests that the H_∞ performance of the above smoother would not improve on that of the filter. Indeed, from Fig. 6, $e = f + u$ and the triangle rule yields $\|e\|_2 \le \|f\|_2 + \|u\|_2$, where f is the H_∞ filter error. That is, the error upper bound for the H_∞ fixed-interval smoother (43) – (45) is greater than that for the H_∞ filter (19) – (20). It is observed below that compared to the minimum-variance case, the H_∞ solution exhibits an increased mean-square error.

Lemma 6 [33]: *For the output estimation problem (14) – (18), in which $C_2(t) = C_1(t) = C(t)$, the smoother solution (43) – (45) results in*

$$\left\|\mathcal{R}_{ei}\mathcal{R}_{ei}^H\right\|_{2,\gamma^{-2} > 0} > \left\|\mathcal{R}_{ei}\mathcal{R}_{ei}^H\right\|_{2,\gamma^{-2} = 0} . \tag{50}$$

Proof: *By expanding $\mathcal{R}_{ei}\mathcal{R}_{ei}^H$ and completing the squares, it can be shown that $\mathcal{R}_{ei}\mathcal{R}_{ei}^H = \mathcal{R}_{ei1}\mathcal{R}_{ei1}^H + \mathcal{R}_{ei2}\mathcal{R}_{ei2}^H$, where $\mathcal{R}_{ei2}\mathcal{R}_{ei2}^H = \mathcal{G}_1 Q(t)\mathcal{G}_1^H - \mathcal{G}_1 Q(t)\mathcal{G}_1^H\Delta^{-H}\Delta^{-1}\mathcal{G}_1 Q(t)\mathcal{G}_1^H$ and $\mathcal{R}_{ei1} = \mathcal{H}\Delta - \mathcal{G}_1 Q(t)\mathcal{G}_1^H\Delta^{-H} = [\mathcal{H} - I + R(t)(\Delta\Delta^H)^{-1}]\Delta$. Substituting $\mathcal{H} = I - R(t)(\hat{\Delta}\hat{\Delta}^H)^{-1}$ into $\mathcal{R}_{ei1}$ yields*

$$\mathcal{R}_{ei1} = R(t)[(\Delta\Delta^H)^{-1} - (\hat{\Delta}\hat{\Delta}^H)^{-1}]\Delta , \tag{51}$$

which suggests $\hat{\Delta} = C(t)\mathcal{G}_0 K(t)R^{1/2}(t) + R^{1/2}(t)$, where $\mathcal{G}_0$ denotes an operator having the state-space realization $\begin{bmatrix} A(t) & I \\ I & 0 \end{bmatrix}$. Constructing $\hat{\Delta}\hat{\Delta}^H = C(t)\mathcal{G}_0[K(t)R(t)K^T(t) - P(t)A^T(t) - A(t)P(t)]\mathcal{G}_0^H C^T(t) + R(t)$ and using (22) yields $\hat{\Delta}\hat{\Delta}^H = C(t)\mathcal{G}_0[B(t)Q(t)B^T(t) - \dot{P}(t) + \gamma^{-2}P(t)C^T(t)C(t)P(t)]\mathcal{G}_0^H C^T(t) + R(t)$. Comparison with $\Delta\Delta^H = C(t)\mathcal{G}_0 B(t)Q(t)B^T(t)\mathcal{G}_0^H C^T(t) + R(t)$ leads to $\hat{\Delta}\hat{\Delta}^H = \Delta\Delta^H - C(t)\mathcal{G}_0(\dot{P}(t) + \gamma^{-2}P(t)C^T(t)C(t)P(t))\mathcal{G}_0^H C^T(t)$. Substituting for $\hat{\Delta}\hat{\Delta}^H$ into (51) yields

$$\mathcal{R}_{ei1} = R(t)[(\Delta\Delta^H)^{-1} - (\Delta\Delta^H - C(t)\mathcal{G}_0(\dot{P}(t) - \gamma^{-2}P(t)C^T(t)C(t)P(t))\mathcal{G}_0^H C^T(t))^{-1}]\Delta . \tag{52}$$

The observation (50) follows by inspection of (52). □

Thus, the cost of designing for worst case input conditions is a deterioration in the mean performance. Note that the best possible average performance $\left\|R_{ei}R_{ei}^H\right\|_2 = \left\|R_{ei2}R_{ei2}^H\right\|_2$ can be attained in problems where there are no uncertainties present, $\gamma^{-2} = 0$ and the Riccati equation solution has converged, that is, $\dot{P}(t) = 0$, in which case $\hat{\Delta}\hat{\Delta}^H = \Delta\Delta^H$ and $\mathcal{R}_{ei1}$ is a zero matrix.

"We know accurately only when we know little, with knowledge doubt increases." *Johann Wolfgang von Goethe*

9.2.8 Performance Comparison

It is of interest to compare to compare the performance of (43) – (45) with the H_∞ smoother described in [10], [13], [16], namely,

$$\begin{bmatrix} \dot{\hat{x}}(t\,|\,T) \\ \dot{\xi}(t) \end{bmatrix} = \begin{bmatrix} A(t) & B(t)Q(t)B^T(t) \\ C^T(t)R^{-1}(t)C(t) & -A'(t) \end{bmatrix} \begin{bmatrix} \hat{x}(t\,|\,T) \\ \xi(t) \end{bmatrix} + \begin{bmatrix} 0 \\ -C^T(t)R^{-1}(t) \end{bmatrix} z(t) \,, \tag{53}$$

$$\hat{x}(t\,|\,T) = \hat{x}(t) + P(t)\xi(t) \tag{54}$$

and (22). Substituting (54) and its differential into the first row of (53) together with (21) yields

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + K(t)(z(t) - C(t)\hat{x}(t)) \,, \tag{55}$$

which reverts to the Kalman filter at γ^{-2} = 0. Substituting $\xi(t) = P^{-1}(t)(\hat{x}(t\,|\,T) - \hat{x}(t))$ into the second row of (53) yields

$$\dot{\hat{x}}(t\,|\,T) = A(t)\hat{x}(t) + G(t)(\hat{x}(t\,|\,T) - \hat{x}(t)) \,, \tag{56}$$

where $G(t) = B(t)Q(t)B^T(t)P_2^{-1}(t)$, which reverts to the maximum-likelihood smoother at γ^{-2} = 0. Thus, the Hamiltonian form (53) – (54) can be realised by calculating the filtered estimate (55) and then obtaining the smoothed estimate from (56).

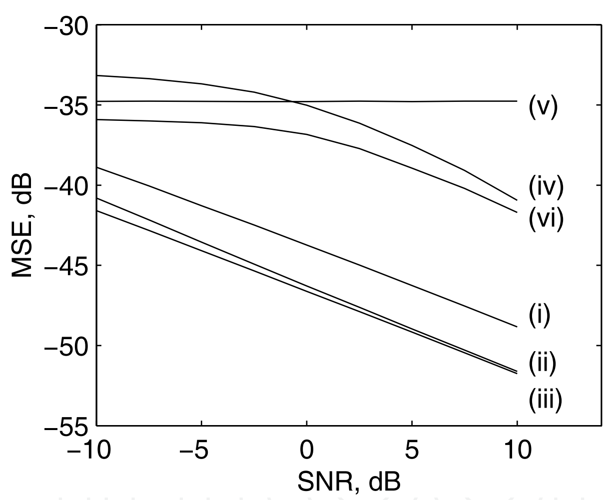


Figure 7. Fixed-interval smoother performance comparison for Gaussian process noise: (i) Kalman filter; (ii) Maximum likelihood smoother; (iii) Minimum-variance smoother; (iv) H_∞ filter; (v) H_∞ smoother [10], [13], [16]; and (vi) H_∞ smoother (43) – (45).

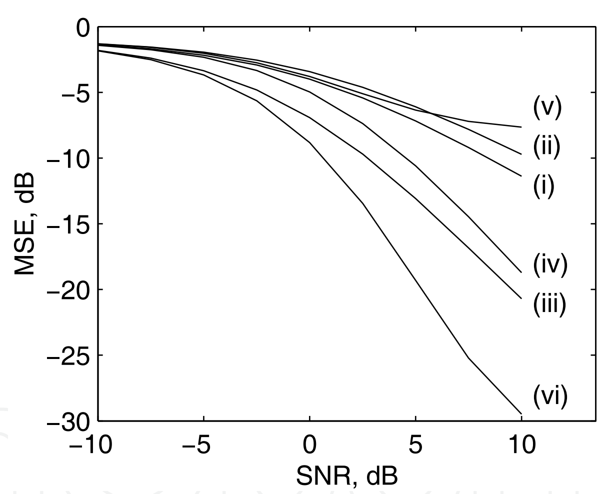


Figure 8. Fixed-interval smoother performance comparison for sinusoidal process noise: (i) Kalman filter; (ii) Maximum likelihood smoother; (iii) Minimum-variance smoother; (iv) H_∞ filter; (v) H_∞ smoother [10], [13], [16]; and (vi) H_∞ smoother (43) – (45).

"Inquiry is fatal to certainty." *William James Durant*

Example: 2 [35]. Let $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $B = C = Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $R = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}$ denote time-invariant parameters for an output estimation problem. Simulations were conducted for the case of $T = 100$ seconds, $\delta t = 1$ millisecond, using 500 realizations of zero-mean, Gaussian process noise and measurement noise. The resulting mean-square-error (MSE) versus signal-to-noise ratio (SNR) are shown in Fig. 7. The H_∞ solutions were calculated using *a priori* designs of $\gamma^{-2} = \sigma_v^{-2}$ within (22). It can be seen from trace (vi) of Fig. 7 that the H_∞ smoothers exhibit poor performance when the exogenous inputs are in fact Gaussian, which illustrates Lemma 6. The figure demonstrates that the minimum-variance smoother out-performs the maximum-likelihood smoother. However, at high SNR, the difference in smoother performance is inconsequential. Intermediate values for γ^{-2} may be selected to realise a smoother design that achieves a trade-off between minimum-variance performance (trace (iii)) and H_∞ performance (trace (v)).

Example 3 [35]. Consider the non-Gaussian process noise signal $w(t) = \sin(t)\sigma_{\sin(t)}^{-1}$, where $\sigma_{\sin(t)}^{2}$ denotes the sample variance of sin(t). The results of a simulation study appear in Fig. 8. It can be seen that the H_∞ solutions, which accommodate input uncertainty, perform better than those relying on Gaussian noise assumptions. In this example, the developed H_∞ smoother (43) – (45) exhibits the best mean-square-error performance.

9.3 Robust Discrete-time Estimation

9.3.1 Discrete-Time Bounded Real Lemma

The development of discrete-time H_∞ filters and smoothers proceeds analogously to the continuous-time case. From Lyapunov stability theory [36], for the unforced system

$$x_{k+1} = A_k x_k , \tag{57}$$

$A_k \in \mathbb{R}^{n \times n}$, to be asymptotically stable over the interval $k \in [1, N]$, a Lyapunov function, $V_k(x_k)$, is required to satisfy $\Delta V_k(x_k) < 0$, where $\Delta V_k(x_k) = V_{k+1}(x_k) - V_k(x_k)$ denotes the first backward difference of $V_k(x_k)$. Consider the candidate Lyapunov function $V_k(x_k) = x_k^T P_k x_k$, where $P_k = P_k^T \in \mathbb{R}^{n \times n}$ is positive definite. To guarantee $x_k \in \ell_2$, it is required that

$$\Delta V_k(x_k) = x_{k+1}^T P_{k+1} x_{k+1} - x_k^T P_k x_k < 0 . \tag{58}$$

Now let $y = \mathcal{G} w$ denote the output of the system

$$x_{k+1} = A_k x_k + B_k w_k , \tag{59}$$

$$y_k = C_k x_k , \tag{60}$$

where $w_k \in \mathbb{R}^m$, $B_k \in \mathbb{R}^{n \times m}$ and $C_k \in \mathbb{R}^{p \times n}$.

"Education is the path from cocky ignorance to miserable uncertainty." *Samuel Langhorne Clemens aka. Mark Twain*

The Bounded Real Lemma [18] states that $w \in \ell_2$ implies $y \in \ell_2$ if

$$x_{k+1}^T P_{k+1} x_{k+1} - x_k^T P_k x_k + y_k^T y_k - \gamma^2 w_k^T w_k < 0 \tag{61}$$

for a $\gamma \in \mathbb{R}$. Summing (61) from k = 0 to k = N – 1 yields the objective

$$-x_0^T P_0 x_0 + \sum_{k=0}^{N-1} y_k^T y_k - \gamma^2 \sum_{k=0}^{N-1} w_k^T w_k < 0 \, , \tag{62}$$

that is,

$$\frac{-x_0^T P_0 x_0 + \sum_{k=0}^{N-1} y_k^T y_k}{\sum_{k=0}^{N-1} w_k^T w_k} < \gamma^2 \, . \tag{63}$$

Assuming that $x_0 = 0$,

$$\left\|\mathcal{G}\right\|_\infty = \frac{\left\|y\right\|_2}{\left\|w\right\|_2} = \frac{\left\|\mathcal{G}w\right\|_2}{\left\|w\right\|_2} = \frac{\sqrt{\sum_{k=0}^{N-1} y_k^T y_k}}{\sqrt{\sum_{k=0}^{N-1} w_k^T w_k}} < \gamma \, . \tag{64}$$

Conditions for achieving the above objectives are established below.

Lemma 7: *The discrete-time Bounded Real Lemma [18]: In respect of the above system $\mathcal{G}$, suppose that the Riccati difference equation*

$$P_k = A_k^T P_{k+1} A_k + \gamma^{-2} A_k^T P_{k+1} B_k (I - \gamma^{-2} B_k^T P_{k+1} B_k)^{-1} B_k^T P_{k+1} A_k + C_k^T C_k \, , \tag{65}$$

with $P_T = 0$, has a positive definite symmetric solution on [0, N]. Then $\left\|\mathcal{G}\right\|_\infty \le \gamma$ for any $w \in \ell_2$.

Proof: *From the approach of Xie et al [18], define*

$$p_k = w_k - \gamma^{-2} (I - \gamma^{-2} B_k^T P_{k+1} B_k)^{-1} B_k^T P_{k+1} A_k x_k \, . \tag{66}$$

It is easily verified that

$$x_{k+1}^T P_{k+1} x_{k+1} - x_k^T P_k x_k + y_k^T y_k - \gamma^2 w_k^T w_k = -\gamma^{-2} p_k^T (I - \gamma^{-2} B_k^T P_{k+1} B_k)^{-1} p_k - x_k^T A_k^T P_{k+1} A_k x_k \, ,$$

which implies (61) – (62) and (63) under the assumption $x_0 = 0$. □

The above lemma relies on the simplifying assumption $E\{w_j w_k^T\} = I\delta_{jk}$. When $E\{w_j w_k^T\} = Q_k \delta_{jk}$, the scaled matrix $\bar{B}_k = B_k Q_k^{1/2}$ may be used in place of B_k above. In the case where $\mathcal{G}$ possesses a direct feedthrough matrix, namely, $y_k = C_k x_k + D_k w_k$, the Riccati difference equation within the above lemma becomes

$$\begin{aligned} P_k &= A_k^T P_{k+1} A_k + C_k^T C_k \\ &\quad + \gamma^{-2} (A_k^T P_{k+1} B_k + C_k^T D_k)(I - \gamma^{-2} B_k^T P_{k+1} B_k - \gamma^{-2} D_k^T D_k)^{-1} (B_k^T P_{k+1} A_k + D_k^T C_k) \, . \end{aligned} \tag{67}$$

"And as he thus spake for himself, Festus said with a loud voice, Paul, thou art beside thyself; much learning doth make thee mad." *Acts 26: 24*

A verification is requested in the problems. It will be shown that predictors, filters and smoothers satisfy a H∞ performance objective if there exist solutions to Riccati difference equations arising from the application of Lemma 7 to the corresponding error systems. A summary of the discrete-time results from [5], [11], [13] and the further details described in [21], [30], is presented below.

9.3.2 Discrete-Time H∞ Prediction

9.3.2.1 Problem Definition

Consider a nominal system $\mathcal{G}_2$

$$x_{k+1} = A_k x_k + B_k w_k, \tag{68}$$

$$y_{2,k} = C_{2,k} x_k , \tag{69}$$

together with a fictitious reference system $\mathcal{G}_1$ realised by (68) and

$$y_{1,k} = C_{1,k} x_k , \tag{70}$$

where A_k, B_k, $C_{2,k}$ and $C_{1,k}$ are of appropriate dimensions. The problem of interest is to find a solution $\mathcal{H}$ that produces one-step-ahead predictions, $\hat{y}_{1,k/k-1}$, given measurements

$$z_k = y_{2,k} + v_k \tag{71}$$

at time k – 1. The prediction error is defined as

$$e_{k/k-1} = y_{1,k} - \hat{y}_{1,k/k-1} . \tag{72}$$

The error sequence (72) is generated by $e = \mathcal{R}_{ei}\, i$, where $\mathcal{R}_{ei} = -[\mathcal{H} \quad \mathcal{H}\mathcal{G}_2 - \mathcal{G}_1]$, $i = \begin{bmatrix} v \\ w \end{bmatrix}$ and the objective is to achieve $\sum_{k=0}^{N-1} e^T_{k/k-1} e_{k/k-1} - \gamma^2 \sum_{k=0}^{N-1} i_k^T i_k < 0$, for some $\gamma \in \mathbb{R}$. For convenience, it is assumed that $w_k \in \mathbb{R}^m$, $E\{w_k\} = 0$, $E\{w_j w_k^T\} = Q_k \delta_{jk}$, $v_k \in \mathbb{R}^p$, $E\{v_k\} = 0$, $E\{v_j v_k^T\} = R_k \delta_{jk}$ and $E\{w_j v_k^T\} = 0$.

9.3.2.2 H∞ Solution

The H∞ predictor has the same structure as the optimum minimum-variance (or Kalman) predictor. It is given by

$$\hat{x}_{k+1/k} = \left(A_k - K_k C_{2,k}\right)\hat{x}_{k/k-1} + K_k z_k, \tag{73}$$

$$\hat{y}_{1,k/k-1} = C_{1,k} \hat{x}_{k/k-1}, \tag{74}$$

"Why waste time learning when ignorance is instantaneous?" *William Boyd Watterson II* where

$$K_k = A_k P_{k/k-1} C_{2,k}^T (C_{2,k} P_{k/k-1} C_{2,k}^T + R_k)^{-1} \tag{75}$$

is the one-step-ahead predictor gain,

$$P_{k/k-1} = (M_k^{-1} - \gamma^{-2} C_{1,k+1}^T C_{1,k+1})^{-1}, \tag{76}$$

and $M_k = M_k^T > 0$ satisfies the Riccati differential equation

$$M_{k+1} = A_k M_k A_k^T + B_k Q_k B_k^T - A_k M_k \begin{bmatrix} C_{1,k}^T & C_{2,k}^T \end{bmatrix} \begin{bmatrix} C_{1,k} M_k C_{1,k}^T - \gamma^2 I & C_{1,k} M_k C_{2,k}^T \\ C_{2,k} M_k C_{1,k}^T & R_k + C_{2,k} M_k C_{2,k}^T \end{bmatrix}^{-1} \begin{bmatrix} C_{1,k} \\ C_{2,k} \end{bmatrix} M_k A_k^T \tag{77}$$

such that $\begin{bmatrix} C_{1,k} M_k C_{1,k}^T - \gamma^2 I & C_{1,k} M_k C_{2,k}^T \\ C_{2,k} M_k C_{1,k}^T & R_k + C_{2,k} M_k C_{2,k}^T \end{bmatrix} > 0$. The above predictor is also known as an *a priori* filter within [11], [13], [30].

9.3.2.3 Performance

Following the approach in the continuous-time case, by subtracting (73) – (74) from (68), (70), the predictor error system is

$$\begin{bmatrix} \tilde{x}_{k+1/k} \\ e_{k/k-1} \end{bmatrix} = \begin{bmatrix} A_k - K_k C_{2,k} & [-K_k \quad B_k] \\ C_{1,k} & [0 \quad 0] \end{bmatrix} \begin{bmatrix} \tilde{x}_{k/k-1} \\ \begin{bmatrix} v_k \\ w_k \end{bmatrix} \end{bmatrix}, \; \tilde{x}_0 = 0, \tag{78}$$

$$= \mathcal{R}_{ei} i ,$$

where $\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1}$, $\mathcal{R}_{ei} = \begin{bmatrix} A_k - K_k C_{2,k} & [-K_k \quad B_k] \\ C_{1,k} & [0 \quad 0] \end{bmatrix}$ and $i = \begin{bmatrix} v \\ w \end{bmatrix}$. It is shown below that the prediction error satisfies the desired performance objective.

Lemma 8 [11], [13], [30]: *In respect of the H∞ prediction problem (68) – (72), the existence of* $M_k = M_k^T > 0$ *for the Riccati differential equation (77) ensures that the solution (73) – (74) achieves the performance objective* $\sum_{k=0}^{N-1} e_{k/k-1}^T e_{k/k-1} - \gamma^2 \sum_{k=0}^{N-1} i_k^T i_k < 0$.

Proof: *By applying the Bounded Real Lemma to* $\mathcal{R}_{ei}^H$ *and taking the adjoint to address* $\mathcal{R}_{ei}$, *it is required that there exists a positive define symmetric solution to*

"Give me a fruitful error any time, full of seeds bursting with its own corrections. You can keep your sterile truth for yourself." *Vilfredo Federico Damaso Pareto*

$$P_{k+1} = (A_k - K_kC_{2,k})P_k(A_k - K_kC_{2,k})^T$$
$$+\gamma^{-2}(A_k - K_kC_{2,k})P_kC_{1,k}^T(I - \gamma^{-2}C_{1,k}P_kC_{1,k}^T)^{-1}C_{2,k}P_k(A_k - K_kC_{2,k})^T + K_kR_kK_k^T + B_kQ_kB_k^T$$
$$= (A_k - K_kC_{2,k})(P_k + \gamma^{-2}P_kC_{1,k}^T(I - \gamma^{-2}C_{1,k}P_kC_{1,k}^T)^{-1}C_{1,k}P_k)(A_k - K_kC_{2,k})^T + K_kR_kK_k^T + B_kQ_kB_k^T$$
$$= (A_k - K_kC_{2,k})(P_k^{-1} + \gamma^{-2}C_{1,k}^TC_{1,k})^{-1}(A_k - K_kC_{2,k})^T + K_kR_kK_k^T + B_kQ_kB_k^T \tag{79}$$

in which use was made of the Matrix Inversion Lemma. Defining $P_{k/k-1} = (P_k^{-1} + \gamma^{-2}C_{1,k}^TC_{1,k})^{-1}$ *leads to*

$$(P_{k+1/k}^{-1} + \gamma^{-2}C_{1,k+1}^TC_{1,k+1})^{-1} = (A_k - K_kC_{2,k})P_{k/k-1}(A_k - K_kC_{2,k})^T + K_kR_kK_k^T + B_kQ_kB_k^T$$
$$= A_kP_{k/k-1}A_k^T + B_kQ_kB_k^T - A_kP_{k/k-1}C_{2,k}^T(R + C_kP_{k/k-1}C_{2,k}^T)^{-1}C_{2,k}P_{k/k-1}A_k^T.$$

and applying the Matrix Inversion Lemma gives

$$(P_{k+1/k}^{-1} + \gamma^{-2}C_{1,k+1}^TC_{1,k+1})^{-1} = A_kP_{k/k-1}A_k^T + B_kQ_kB_k^T - A_kP_{k/k-1}C_{2,k}^T(R + C_{2,k}P_{k/k-1}C_{2,k}^T)^{-1}C_{2,k}P_{k/k-1}A_k^T$$
$$= A_k(P_{k/k-1} + C_{2,k}^TR_k^{-1}C_{2,k})^{-1}A_k^T + B_kQ_kB_k^T .$$

The change of variable (76), namely, $P_{k/k-1}^{-1} = M_k^{-1} - \gamma^{-2}C_{1,k}^TC_{1,k}$, *results in*

$$\begin{aligned} M_{k+1} &= A_k(M_k^{-1} + C_{2,k}^TR_k^{-1}C_{2,k} - \gamma^{-2}C_{1,k}^TC_{1,k})^{-1}A_k^T + B_kQ_kB_k^T \\ &= A_k(M_k^{-1} + \bar{C}_k\bar{R}_k\bar{C}_k^T)^{-1}A_k^T + B_kQ_kB_k^T . \end{aligned} \tag{80}$$

where $\bar{C}_k = \begin{bmatrix} C_{1,k} \\ C_{2,k} \end{bmatrix}$ *and* $\bar{R}_k = \begin{bmatrix} -\gamma^{-2}I & 0 \\ 0 & R_k \end{bmatrix}$. *Applying the Matrix Inversion Lemma within (80) gives*

$$M_{k+1} = A_kMA_k^T - A_kM_k\bar{C}_k^T(\bar{R}_k + \bar{C}_kM_k\bar{C}_k^T)^{-1}\bar{C}_kM_kA_k^T + B_kQ_kB_k^T , \tag{81}$$

Expanding (81) yields (77). The existence of $M_k > 0$ *for the above Riccati differential equation implies* $P_k > 0$ *for (79). Thus, it follows from Lemma 7 that the stated performance objective is achieved.* □

9.3.3 Discrete-Time H∞ Filtering

9.3.3.1 Problem Definition

Consider again the configuration of Fig. 1. Assume that the systems $\mathcal{G}_2$ and $\mathcal{G}_1$ have the realisations (68) – (69) and (68), (70), respectively. It is desired to find a solution $\mathcal{H}$ that operates on the measurements (71) and produces the filtered estimates $\hat{y}_{1,k/k}$. The filtered error sequence,

$$e_{k/k} = y_{1,k} - \hat{y}_{1,k/k} , \tag{82}$$

"Never interrupt your enemy when he is making a mistake." *Napoléon Bonaparte* is generated by $e = \mathcal{R}_{ei} i$, where $\mathcal{R}_{ei} = -[\mathcal{H} \quad \mathcal{H}\mathcal{G}_2 - \mathcal{G}_1]$, $i = \begin{bmatrix} v \\ w \end{bmatrix}$. The H∞ performance objective is to achieve $\sum_{k=0}^{N-1} e_{k/k}^T e_{k/k} - \gamma^2 \sum_{k=0}^{N-1} i_k^T i_k < 0$, for some $\gamma \in \mathbb{R}$.

9.3.3.2 H∞ Solution

As explained in Chapter 4, filtered states can be evolved from

$$\hat{x}_{k/k} = A_{k-1}\hat{x}_{k-1/k-1} + L_k(z_k - C_{2,k}A_{k-1}\hat{x}_{k-1/k-1}), \tag{83}$$

where $L_k \in \mathbb{R}^{n \times p}$ is a filter gain. The above recursion is called an *a posteriori* filter in [11], [13], [30]. Output estimates are obtained from

$$\hat{y}_{1,k-1/k-1} = C_{1,k-1}\hat{x}_{k-1/k-1}. \tag{84}$$

The filter gain is calculated as

$$L_k = M_k C_k^T (C_k M_k C_k^T + R_k)^{-1}, \tag{85}$$

where $M_k = M_k^T > 0$ satisfies the Riccati differential equation

$$\begin{aligned} M_k &= A_{k-1}M_{k-1}A_{k-1}^T + B_{k-1}Q_{k-1}B_{k-1}^T - A_{k-1}M_{k-1}\begin{bmatrix} C_{1,k-1}^T & C_{2,k-1}^T \end{bmatrix} \\ &\quad \times \begin{bmatrix} C_{1,k-1}M_{k-1}C_{1,k-1}^T - \gamma^2 I & C_{1,k-1}M_{k-1}C_{2,k-1}^T \\ C_{2,k-1}M_{k-1}C_{1,k-1}^T & R_{k-1} + C_{2,k-1}M_{k-1}C_{2,k-1}^T \end{bmatrix}^{-1} \begin{bmatrix} C_{1,k-1} \\ C_{2,k-1} \end{bmatrix} M_{k-1}A_{k-1}^T \end{aligned} \tag{86}$$

such that $\begin{bmatrix} C_{1,k-1}M_{k-1}C_{1,k-1}^T - \gamma^2 I & C_{1,k-1}M_{k-1}C_{2,k-1}^T \\ C_{2,k-1}M_{k-1}C_{1,k-1}^T & R_{k-1} + C_{2,k-1}M_{k-1}C_{2,k-1}^T \end{bmatrix} > 0$.

9.3.3.3 Performance

Subtracting from (83) from (68) gives $\tilde{x}_{k/k} = A_{k-1}\hat{x}_{k-1/k-1} + B_{k-1}w_{k-1} - A_{k-1}\hat{x}_{k-1/k-1} + L_k C_{2,k} A_{k-1}\hat{x}_{k-1/k-1} + L_k(C_{2,k}(A_{k-1}x_{k-1} + B_{k-1}w_{k-1}) + v_k)$. Denote $i_k = \begin{bmatrix} v_k \\ w_{k-1} \end{bmatrix}$, then the filtered error system may be written as

$$\begin{aligned} \begin{bmatrix} \tilde{x}_{k/k} \\ e_{k-1/k-1} \end{bmatrix} &= \begin{bmatrix} (I - L_k C_{2,k})A_{k-1} & [-L_k \quad (I - L_k C_{2,k})B_{k-1}] \\ C_{1,k} & [0 \quad 0] \end{bmatrix} \begin{bmatrix} \tilde{x}_{k-1/k-1} \\ i_k \end{bmatrix} \\ &= \mathcal{R}_{ei} i, \end{aligned} \tag{87}$$

"I believe the most solemn duty of the American president is to protect the American people. If America shows uncertainty and weakness in this decade, the world will drift toward tragedy. This will not happen on my watch." *George Walker Bush* with $\tilde{x}_0 = 0$, where $\mathcal{R}_{ei} = \begin{bmatrix} (I - L_k C_{2,k}) A_{k-1} & [-L_k \quad (I - L_k C_{2,k}) B_{k-1}] \\ C_{1,k} & [0 \qquad 0] \end{bmatrix}$. It is shown below that the filtered error satisfies the desired performance objective.

Lemma 9 [11], [13], [30]: *In respect of the H∞ problem (68) – (70), (82), the solution (83) – (84) achieves the performance* $\sum_{k=0}^{N-1} e_{k/k}^T e_{k/k} - \gamma^2 \sum_{k=0}^{N-1} i_k^T i_k < 0$.

Proof: *By applying the Bounded Real Lemma to* $\mathcal{R}_{ei}^H$ *and taking the adjoint to address* $\mathcal{R}_{ei}$, *it is required that there exists a positive define symmetric solution to*

$$\begin{aligned} P_{k+1} &= (I - L_k C_{2,k}) A_{k-1} P_k A_{k-1}^T (I - C_{2,k}^T L_k^T) \\ &\quad + \gamma^{-2} (I - L_k C_{2,k}) A_{k-1} P_k C_{1,k}^T (I - \gamma^{-2} C_{1,k} P_k C_{1,k}^T)^{-1} C_{1,k} P_k A_{k-1}^T (I - L_k C_{2,k})^T \\ &\quad + (I - L_k C_{2,k}) B_{k-1} Q_{k-1} B_{k-1}^T (I - C_{2,k}^T L_k^T) + L_k R_k L_k^T \\ &= (I - L_k C_{2,k}) A_{k-1} (P_k^{-1} - \gamma^{-2} C_{1,k}^T C_{1,k}) A_{k-1}^T (I - C_{2,k}^T L_k^T) \\ &\quad + (I - L_k C_{2,k}) B_{k-1} Q_{k-1} B_{k-1}^T (I - C_{2,k}^T L_k^T) + L_k R_k L_k^T , \end{aligned} \tag{88}$$

in which use was made of the Matrix Inversion Lemma. Defining

$$P_{k/k-1}^{-1} = P_k^{-1} - \gamma^{-2} C_{1,k}^T C_{1,k} , \tag{89}$$

using (85) and applying the Matrix Inversion Lemma leads to

$$\begin{aligned} (P_{k+1/k}^{-1} + \gamma^{-2} C_{1,k}^T C_{1,k}) &= (I - L_k C_{2,k})(A_{k-1} P_{k/k-1} A_{k-1}^T + B_{k-1} Q_{k-1} B_{k-1}^T)(I - C_{2,k}^T L_k^T) + L_k R_k L_k^T \\ &= (I - L_k C_{2,k}) M_k (I - C_{2,k}^T L_k^T) + L_k R_k L_k^T \\ &= M_k - M_k C_{2,k}^T (R_k + C_{2,k} M_k C_{2,k}^T)^{-1} C_{2,k} M_k \\ &= (M_k^{-1} + C_{2,k}^T R_k^{-1} C_{2,k})^{-1} , \end{aligned} \tag{90}$$

where

$$M_k = A_{k-1} P_{k/k-1} A_{k-1}^T + B_{k-1} Q_{k-1} B_{k-1}^T . \tag{91}$$

It follows from (90) that $P_{k+1/k}^{-1} + \gamma^{-2} C_{1,k}^T C_{1,k} = M_k^{-1} + C_{2,k}^T R_k^{-1} C_{2,k}$ *and*

$$\begin{aligned} P_{k/k-1}^{-1} &= M_{k-1}^{-1} + C_{2,k-1}^T R_{k-1}^{-1} C_{2,k-1} - \gamma^{-2} C_{1,k-1}^T C_{1,k-1} \\ &= M_{k-1}^{-1} + \bar{C}_{k-1}^T \bar{R}_{k-1}^{-1} \bar{C}_{k-1} , \end{aligned} \tag{92}$$

where $\bar{C}_k = \begin{bmatrix} C_{1,k} \\ C_{2,k} \end{bmatrix}$ *and* $\bar{R}_k = \begin{bmatrix} -\gamma^{-2} I & 0 \\ 0 & R_k \end{bmatrix}$. *Substituting (92) into (91) yields*

"Hell, there are no rules here – we're trying to accomplish something." *Thomas Alva Edison*

$$M_k = A_{k-1}(M_{k-1}^{-1} + \bar{C}_{k-1}\bar{R}_{k-1}\bar{C}_{k-1}^T)^{-1}A_{k-1}^T + B_{k-1}Q_{k-1}B_{k-1}^T, \tag{93}$$

which is the same as (86). The existence of $M_k > 0$ for the above Riccati difference equation implies the existence of a $P_k > 0$ for (88). Thus, it follows from Lemma 7 that the stated performance objective is achieved. □

9.3.4 Solution to the General Filtering Problem

Limebeer, Green and Walker express Riccati difference equations such as (86) in a compact form using J-factorisation [5], [21]. The solutions for the general filtering problem follow immediately from their results. Consider

$$\begin{bmatrix} x_{k+1} \\ e_{k/k} \\ z_k \end{bmatrix} = \begin{bmatrix} A_k & B_{1,1,k} & 0 \\ C_{1,1,k} & D_{1,1,k} & D_{1,2,k} \\ C_{2,1,k} & D_{2,1,k} & 0 \end{bmatrix} \begin{bmatrix} x_k \\ i_k \\ -\hat{y}_{1,k/k} \end{bmatrix}. \tag{94}$$

Let $J_k = \begin{bmatrix} E\{i_k i_k^T\} & 0 \\ 0 & -\gamma^2 \end{bmatrix}$, $\bar{C}_k = \begin{bmatrix} C_{1,1,k} \\ C_{2,1,k} \end{bmatrix}$, $\bar{D}_k = \begin{bmatrix} D_{1,1,k} & D_{1,2,k} \\ D_{2,1,k} & 0 \end{bmatrix}$ and $\bar{B}_k = \begin{bmatrix} B_{1,1,k} & 0 \end{bmatrix}$. From the approach of [5], [21], the Riccati difference equation corresponding to the H_∞ problem (94) is

$$\begin{aligned} M_{k+1} &= A_k M_k A_k^T + \bar{B}_k J_k \bar{B}_k^T \\ &\quad - (A_k M_k \bar{C}_k^T + \bar{B}_k J_k \bar{D}_k^T)(\bar{C}_k M_k \bar{C}_k^T + \bar{D}_k J_k \bar{D}_k^T)^{-1}(A_k M_k \bar{C}_k^T + \bar{B}_k J_k \bar{D}_k^T)^T. \end{aligned} \tag{95}$$

Suppose in a general filtering problem that $\mathcal{G}_2$ is realised by (68), $y_{2,k} = C_{2,k}x_k + D_{2,k}w_k$, $\mathcal{G}_1$ is realised by (68) and $y_{1,k} = C_{1,k}x_k + D_{1,k}w_k$. Then substituting $B_{1,1,k} = \begin{bmatrix} 0 & B_k \end{bmatrix}$, $C_{1,1,k} = C_{1,k}$, $C_{2,1,k} = C_{2,k}$, $D_{1,1,k} = \begin{bmatrix} 0 & D_{1,k} \end{bmatrix}$, $D_{1,2,k} = I$ and $D_{2,1,k} = \begin{bmatrix} I & D_{2,k} \end{bmatrix}$ into (95) yields

$$\begin{aligned} M_{k+1} &= A_k M_k A_k^T - \begin{bmatrix} A_k M_k C_{1,k}^T + B_k Q_k D_{1,k}^T \\ A_k M_k C_{2,k}^T + B_k Q_k D_{2,k}^T \end{bmatrix} \\ &\quad \times \begin{bmatrix} C_{1,k} M_k C_{1,k}^T + D_{1,k} Q_k D_{1,k}^T - \gamma^2 I & C_{1,k} M_k C_{2,k}^T + D_{1,k} Q_k D_{2,k}^T \\ C_{2,k} M_k C_{1,k}^T + D_{2,k} Q_k D_{1,k}^T & R_k + C_{2,k} M_k C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} C_{1,k} M_k A_k^T + D_{1,k} Q_k B_{1,k}^T & C_{2,k} M_k A_k^T + D_{2,k} Q_k B_k^T \end{bmatrix} + B_k Q_k B_k^T. \end{aligned} \tag{96}$$

The filter solution is given by

$$\hat{x}_{k+1/k} = A_k \hat{x}_{k/k-1} + K_k(z_k - C_{2,1,k}\hat{x}_{k/k-1}), \tag{97}$$

$$\hat{y}_{1,k/k} = C_{1,1,k}\hat{x}_{k/k-1} + L_k(z_k - C_{2,1,k}\hat{x}_{k/k-1}), \tag{98}$$

where $K_k = (A_k M_k C_{2,k}^T + B_k Q_k D_{2,k}^T)\Omega_k^{-1}$, $L_k = (C_{1,k} M_k C_{2,k}^T + D_{1,k} Q_k D_{2,k}^T)\Omega_k^{-1}$ and $\Omega_k = C_{2,k} M_k C_{2,k}^T + D_{2,k} Q_k D_{2,k}^T + R_k$.

"If we knew what it is we were doing, it would not be called research, would it?" *Albert Einstein*

9.3.5 Discrete-Time H∞ Smoothing

9.3.5.1 Problem Definition

Suppose that measurements (72) of a system (68) – (69) are available over an interval $k \in [1, N]$. The problem of interest is to calculate smoothed estimates $\hat{y}_{k/N}$ of y_k such that the error sequence

$$e_{k/N} = y_k - \hat{y}_{k/N} \tag{99}$$

is in ℓ_2.

9.3.5.2 H∞ Solution

The following fixed-interval smoother for output estimation [28] employs the gain for the H∞ predictor,

$$K_k = A_k P_{k/k-1} C_{2,k}^T \Omega_k^{-1}, \tag{100}$$

where $\Omega_k = C_{2,k} P_{k/k-1} C_{2,k}^T + R_k$, in which $P_{k/k-1}$ is obtained from (76) and (77). The gain (100) is used in the minimum-variance smoother structure described in Chapter 7, *viz.*,

$$\begin{bmatrix} \hat{x}_{k+1/k} \\ \alpha_k \end{bmatrix} = \begin{bmatrix} A_k - K_k C_{2,k} & K_k \\ -\Omega_k^{-1/2} C_{2,k} & \Omega_k^{-1/2} \end{bmatrix} \begin{bmatrix} \hat{x}_{k/k-1} \\ z_k \end{bmatrix}, \tag{101}$$

$$\begin{bmatrix} \xi_{k-1} \\ \beta_k \end{bmatrix} = \begin{bmatrix} A_k^T - C_{2,k}^T K_k^T & C_{2,k}^T \Omega_k^{-1/2} \\ -K_k^T & \Omega_k^{-1/2} \end{bmatrix} \begin{bmatrix} \xi_k \\ \alpha_k \end{bmatrix}, \xi_N = 0, \tag{102}$$

$$\hat{y}_{k/N} = z_k - R_k \beta_k . \tag{103}$$

It is argued below that this smoother meets the desired H∞ performance objective.

9.3.5.3 H∞ Performance

It is easily shown that the smoother error system is

$$\begin{aligned} \begin{bmatrix} \tilde{x}_{k+1/k} \\ \lambda_{k-1} \\ e_{k/N} \end{bmatrix} &= \begin{bmatrix} \begin{bmatrix} A_k - K_k C_{2,k} & 0 \\ C_{2,k}^T \Omega_k^{-1} C_{2,k} & A_k^T - C_{2,k}^T K_k^T \end{bmatrix} & \begin{bmatrix} -K_k & B_k \\ C_{2,k}^T \Omega_k^{-1} & 0 \end{bmatrix} \\ \begin{bmatrix} R_k \Omega_k^{-1} C_{2,k} & -R_k K_k^T \end{bmatrix} & \begin{bmatrix} R_k \Omega_k^{-1} - I & 0 \end{bmatrix} \end{bmatrix} \begin{bmatrix} \tilde{x}_{k/k-1} \\ \lambda_k \\ v_k \\ w_k \end{bmatrix}, \\ &= \boldsymbol{\mathcal{R}}_{ei} i , \end{aligned} \tag{104}$$

with $\tilde{x}_0 = 0$, where $\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1}$, $i = \begin{bmatrix} v \\ w \end{bmatrix}$ and

"I have had my results for a long time: but I do not yet know how I am to arrive at them." *Karl Friedrich Gauss*

$$\mathcal{R}_{ei} = \begin{bmatrix} \begin{bmatrix} A_k - K_k C_{2,k} & 0 \\ C_{2,k}^T \Omega_k^{-1} C_{2,k} & A_k^T - C_{2,k}^T K_k^T \end{bmatrix} & \begin{bmatrix} -K_k & B_k \\ C_{2,k}^T \Omega_k^{-1} & 0 \end{bmatrix} \\ \begin{bmatrix} R_k \Omega_k^{-1} C_{2,k} & -R_k K_k^T \end{bmatrix} & \begin{bmatrix} R_k \Omega_k^{-1} - I & 0 \end{bmatrix} \end{bmatrix}.$$

Lemma 10: *In respect of the smoother error system (104), if there exists a symmetric positive definite solutions to (77) for γ > 0, then the smoother (101) – (103) achieves $\mathcal{R}_{ei} \in \ell_\infty$, that is, $i \in \ell_2$ implies $e \in \ell_2$.*

Outline of Proof: *From Lemma 8, $\tilde{x} \in \ell_2$, since it evolves within the predictor error system. Therefore, $\lambda \in \ell_2$, since it evolves within the adjoint predictor error system. Then $e \in \ell_2$, since it is a linear combination of $\tilde{x}$, λ and $i \in \ell_2$.* □

9.3.5.4 Performance Comparison

Example 4 [28]. A voiced speech utterance "a e i o u" was sampled at 8 kHz for the purpose of comparing smoother performance. Simulations were conducted with the zero-mean, unity-variance speech sample interpolated to a 16 kHz sample rate, to which 200 realizations of Gaussian measurement noise were added and the signal to noise ratio was varied from -5 to 5 dB. The speech sample is modelled as a first-order autoregressive process

$$x_{k+1} = Ax_k + w_k , \tag{105}$$

where $A \in \mathbb{R}$, $0 < A < 1$. Estimates for σ_w^2 and A were calculated at 20 dB SNR using an EM algorithm, see Chapter 8.

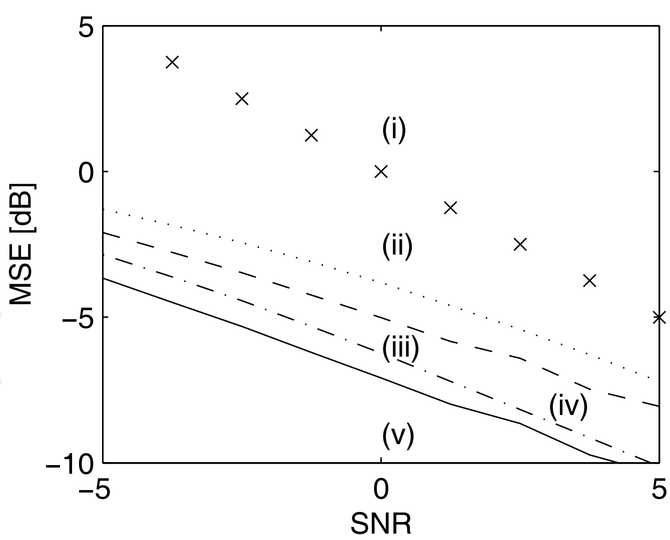


Fig. 9. Speech estimate performance comparison: (i) data (crosses), (ii) Kalman filter (dotted line), (iii) H $_\infty$ filter (dashed line), (iv) minimum-variance smoother (dot-dashed line) and (v) H $_\infty$ smoother (solid line).

"If I have seen further it is only by standing on the shoulders of giants." *Isaac Newton*

Simulations were conducted in which a minimum-variance filter and a fixed-interval smoother were employed to recover the speech message from noisy measurements. The results are provided in Fig. 9. As expected, the smoother out-performs the filter. Searches were conducted for minimum values of γ such that solutions to the design Riccati difference equations were positive definite for each noise realisation. The performance of the resulting H_∞ filter and smoother are indicated by the dashed line and solid line of the figure. It can be seen for this example that the H_∞ filter out-performs the Kalman filter. The figure also indicates that the robust smoother provides the best performance and exhibits about 4 dB reduction in mean-square-error compared to the Kalman filter at 0 dB SNR. This performance benefit needs to be reconciled against the extra calculation cost of combining robust forward and backward state predictors within (101) – (103).

9.3.5.5 High SNR and Low SNR Asymptotes

An understanding of why robust solutions are beneficial in the presence of uncertainties can be gleaned by examining single-input-single-output filtering and equalisation. Consider a time-invariant plant having the canonical form

$$A = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & & \ddots & 0 \\ \vdots & & & \\ -a_0 & -a_1 & & -a_{n-1} \end{bmatrix},\ B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},\ C_2 = \begin{bmatrix} c & 0 & \cdots & 0 \end{bmatrix},$$

$a_0, \ldots a_{n-1}, c \in \mathbb{R}$. Since the plant is time-invariant, the transfer function exists and is denoted by $G(z)$. Some notation is defined prior to stating some observations for output estimation problems. Suppose that an H_∞ filter has been constructed for the above plant. Let the H_∞ algebraic Riccati equation solution, predictor gain, filter gain, predictor, filter and smoother transfer function matrices be denoted by $P^{(\infty)}$, $K^{(\infty)}$, $L^{(\infty)}$, $H_P^{(\infty)}(z)$, $H_F^{(\infty)}(z)$ and $H_S^{(\infty)}(z)$ respectively. The H_∞ filter transfer function matrix may be written as $H_F^{(\infty)}(z) = L^{(\infty)} + (I - L^{(\infty)})H_P^{(\infty)}(z)$ where $L^{(\infty)} = I - R(\Omega^{(\infty)})^{-1}$. The transfer function matrix of the map from the inputs to the filter output estimation error is

$$R_{ei}^{(\infty)}(z) = -[H_F^{(\infty)}(z)\sigma_v \quad (H_F^{(\infty)}(z) - I)G(z)\sigma_w]\,. \tag{106}$$

The H_∞ smoother transfer function matrix can be written as $H_S^{(\infty)}(z) = I - R(I - (H_P^{(\infty)}(z))^H)(\Omega^{(\infty)})^{-1}(I - H_P^{(\infty)}(z))$. Similarly, let $P^{(2)}$, $K^{(2)}$, $L^{(2)}$, $H_F^{(2)}(z)$ and $H_S^{(2)}(z)$ denote the minimum-variance algebraic Riccati equation solution, predictor gain, filter gain, filter and smoother transfer function matrices respectively.

"In computer science, we stand on each other's feet." *Brian K. Reid*

Proposition 1 [28]: *In the above output estimation problem:*

(i)

$$\lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| H_F^{(\infty)}(e^{j\omega}) \right| = 1. \tag{107}$$

(ii)

$$\lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| H_F^{(2)}(e^{j\omega}) \right| < \lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| H_F^{(\infty)}(e^{j\omega}) \right|. \tag{108}$$

(iii)

$$\lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| R_{ei}^{(\infty)}(R_{ei}^{(\infty)})^H(e^{j\omega}) \right| = \lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| H_F^{(\infty)}(H_F^{(\infty)})^H(e^{j\omega}) \right| \sigma_v^2. \tag{109}$$

(iv)

$$\lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| H_S^{(\infty)}(e^{j\omega}) \right| = 1. \tag{110}$$

(v)

$$\lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| H_S^{(2)}(e^{j\omega}) \right| < \lim_{\sigma_v^2 \to 0} \sup_{\omega \in \{-\pi,\pi\}} \left| H_S^{(\infty)}(e^{j\omega}) \right|. \tag{111}$$

Outline of Proof: *(i) Let* $p_{(1,1)}^{(\infty)}$ *denote the (1,1) component of* $P^{(\infty)}$. *The low measurement noise observation (107) follows from* $L^{(\infty)} = 1 - \sigma_v^2(c^2 p_{(1,1)}^{(\infty)} + \sigma_v^2)^{-1}$ *which implies* $\lim_{\sigma_v^2 \to 0} L^{(\infty)} = 1$.

(ii) Observation (108) follows from $p_{(1,1)}^{(\infty)} > p_{(1,1)}^{(2)}$, *which results in* $\lim_{\sigma_v^2 \to 0} L^{(\infty)} > \lim_{\sigma_v^2 \to 0} L^{(2)}$.

(iii) Observation (109) follows immediately from the application of (107) in (106).

(iv) Observation (110) follows from $\lim_{\sigma_v^2 \to 0} \sigma_v^2(c^2 p_{(1,1)}^{(\infty)} + \sigma_v^2)^{-1} = 0$.

(v) Observation (111) follows from $p_{(1,1)}^{(\infty)} > p_{(1,1)}^{(2)}$ *which results in* $\lim_{\sigma_v^2 \to 0} \Omega_3^{(\infty)} > \lim_{\sigma_v^2 \to 0} \Omega^{(2)}$. □

An interpretation of (107) and (110) is that the maximum magnitudes of the filters and smoothers asymptotically approach a short circuit (or zero impedance) when $\sigma_v^2 \to 0$. From (108) and (111), as $\sigma_v^2 \to 0$, the maximum magnitudes of the H∞ solutions approach the short circuit asymptote closer than the optimal minimum-variance solutions. That is, for low measurement noise, the robust solutions accommodate some uncertainty by giving greater weighting to the data. Since $\lim_{\sigma_v^2 \to 0} \left\| R_{ei} \right\|_\infty \to \sigma_v$ and the H∞ filter achieves the performance $\left\| R_{ei} \right\|_\infty < \gamma$, it follows from (109) that an *a priori* design estimate is $\gamma = \sigma_v$.

"All programmers are optimists." *Frederick P. Brooks, Jr*

Suppose now that a time-invariant plant has the transfer function $G(z) = C(zI - A)^{-1}B + D$, where A, B and C are defined above together with $D \in \mathbb{R}$. Consider an input estimation (or equalisation) problem in which the transfer function matrix of the causal H_∞ solution that estimates the input of the plant is

$$H_F^{(\infty)}(z) = QD^T\left(\Omega^{(\infty)}\right)^{-1} - QD^T\left(\Omega^{(\infty)}\right)^{-1}H_P^{(\infty)}(z). \tag{112}$$

The transfer function matrix of the map from the inputs to the input estimation error is

$$R_{ei}^{(\infty)}(z) = -[H_F^{(\infty)}(z)\sigma_v \quad (H_F^{(\infty)}G(z) - I)\sigma_w]. \tag{113}$$

The noncausal H_∞ transfer function matrix of the input estimator can be written as $H_S^{(\infty)}(z)$ = $QG^{-H}(z)(I - (H_P^{(\infty)}(z))^H)(\Omega_3^{(\infty)})^{-1}(I - H_P^{(\infty)}(z))$.

Proposition 2 [28]: *For the above input estimation problem:*

(i)

$$\lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|H_F^{(\infty)}(e^{j\omega})\right| = 0. \tag{114}$$

(ii)

$$\lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|H_F^{(2)}(e^{j\omega})\right| > \lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|H_F^{(\infty)}(e^{j\omega})\right|. \tag{115}$$

(iii)

$$\begin{aligned}&\lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|R_{ei}^{(\infty)}(R_{ei}^{(\infty)})^H(e^{j\omega})\right| \\ &\quad = \lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|(H_F^{(\infty)}G^{(\infty)}(e^{j\omega}) - 1)(H_F^{(\infty)}G^{(\infty)}(e^{j\omega}) - 1)^H\right|\sigma_w^2.\end{aligned} \tag{116}$$

(iv)

$$\lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|H_S^{(\infty)}(e^{j\omega})\right| = 0. \tag{117}$$

(v)

$$\lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|H_S^{(2)}(e^{j\omega})\right| > \lim_{\sigma_v^{-2}\to 0}\sup_{\omega\in\{-\pi,\pi\}}\left|H_S^{(\infty)}(e^{j\omega})\right|. \tag{118}$$

Outline of Proof: *(i) and (iv) The high measurement noise observations (114) and (117) follow from* $\Omega^{(\infty)} = c^2 p_{(1,1)}^{(\infty)} + D\sigma_w^2 + \sigma_v^2$ *which implies* $\lim_{\sigma_v^{-2}\to 0}(\Omega_3^{(\infty)})^{-1}$ *= 0.*

(ii) and (v) The observations (115) and (118) follow from $p_{(1,1)}^{(\infty)} > p_{(1,1)}^{(2)}$*, which results in* $\lim_{\sigma_v^{2}\to 0}\Omega_3^{(\infty)}$ > $\lim_{\sigma_v^{2}\to 0}\Omega^{(2)}$.

"Always code as if the guy who ends up maintaining your code will be a violent psychopath who knows where you live." *Damian Conway* *(iii) The observation (116) follows immediately from the application of (114) in (113).* □

An interpretation of (114) and (117) is that the maximum magnitudes of the equalisers asymptotically approach an open circuit (or infinite impedance) when $\sigma_v^{-2} \to 0$. From (115) and (118), as $\sigma_v^{-2} \to 0$, the maximum magnitude of the H_∞ solution approaches the open circuit asymptote closer than that of the optimum minimum-variance solution. That is, under high measurement noise conditions, robust solutions accommodate some uncertainty by giving less weighting to the data. Since $\lim_{\sigma_v^{-2} \to 0} \left\| R_{ei} \right\|_\infty \to \sigma_w$, the H_∞ solution achieves the performance $\left\| R_{ei} \right\|_\infty < \gamma$, it follows from (116) than an *a priori* design estimate is $\gamma = \sigma_w$.

Proposition 1 follows intuitively. Indeed, the short circuit asymptote is sometimes referred to as the singular filter. Proposition 2 may appear counter-intuitive and warrants further explanation. When the plant is minimum phase and the measurement noise is negligible, the equaliser inverts the plant. Conversely, when the equalisation problem is dominated by measurement noise, the solution is a low gain filter; that is, the estimation error is minimised by giving less weighting to the data.

9.4 Conclusion

Uncertainties are invariably present within the specification of practical problems. Consequently, robust solutions have arisen to accommodate uncertain inputs and plant models. The H_∞ performance objective is to minimise the ratio of the output energy to the input energy of an error system, that is, minimise

$$\sup_{\|i\|_2 \neq 0} \left\| \mathcal{R}_{ei} \right\|_\infty = \sup_{\|i\|_2 \neq 0} \frac{\|e\|_2}{\|i\|_2} \le \gamma$$

for some $\gamma \in \mathbb{R}$. In the time-invariant case, the objective is equivalent to minimising the maximum magnitude of the error power spectrum density.

Predictors, filters and smoothers that satisfy the above performance objective are found by applying the Bounded Real Lemma. The standard solution structures are retained but larger design error covariances are employed to account for the presence of uncertainty. In continuous time output estimation, the error covariance is found from the solution of

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) - P(t)(C^T(t)R^{-1}(t)C\ (t) - \gamma^{-2}C^T(t)C\ (t))P(t) + B(t)Q(t)B^T(t) .$$

Discrete-time predictors, filters and smoothers for output estimation rely on the solution of

$$P_{k+1} = A_k P_k A_k^T - A_k P_k \begin{bmatrix} C_k^T & C_k^T \end{bmatrix} \begin{bmatrix} C_k P_k C_k^T - \gamma^2 I & C_k P_k C_k^T \\ C_k P_k C_k^T & R_k + C_k P_k C_k^T \end{bmatrix}^{-1} \begin{bmatrix} C_k \\ C_k \end{bmatrix} P_k A_k^T + B_k Q_k B_k^T .$$

"Your most unhappy customers are your greatest source of learning." *William Henry (Bill) Gates III*

It follows that the H_∞ designs revert to the optimum minimum-variance solutions as $\gamma^{-2} \rightarrow 0$. Since robust solutions are conservative, the art of design involves finding satisfactory trade-offs between average and worst-case performance criteria, namely, tweaking the γ.

A summary of suggested approaches for different linear estimation problem conditions is presented in Table 1. When the problem parameters are known precisely then the optimum minimum-variance solutions cannot be improved upon. However, when the inputs or the models are uncertain, robust solutions may provide improved mean-square-error performance. In the case of low measurement noise output-estimation, the benefit arises because greater weighting is given to the data. Conversely, for high measurement noise input estimation, robust solutions accommodate uncertainty by giving less weighting to the data.

PROBLEM CONDITIONS	SUGGESTED APPROACHES
Gaussian process and measurement noises, known 2nd-order statistics. Known system model parameters.	1. Optimal minimum-variance (or Kalman) filter. 2. Fixed-lag smoothers, which improve on filter performance (see Lemma 3 and Example 1 of Chapter 7). They suit on-line applications and have low additional complexity. A sufficiently large smoothing lag results in optimal performance (see Example 3 of Chapter 7). 3. Maximum-likelihood (or Rauch-Tung-Striebel) smoothers, which also improve on filter performance (see Lemma 6 of Chapter 6 and Lemma 4 of Chapter 7). They can provide close to optimal performance (see Example 5 of Chapter 6). 4. The minimum-variance smoother provides the best performance (see Lemma 12 of Chapter 6 and Lemma 8 of Chapter 7) at the cost of increased complexity (see Example 5 of Chapter 6 and Example 2 of Chapter 7).
Uncertain process and measurement noises, known 2nd-order statistics. Known system model parameters.	1. Optimal minimum-variance filter, which does not rely on Gaussian noise assumptions. 2. Optimal minimum-variance smoother, which similarly does not rely on Gaussian noise assumptions (see Example 6 of Chapter 6). 3. Robust filter which trades off H_∞ performance (see Lemmas 2, 9) and mean-square-error performance (see Example 3). 4. Robust smoother which trades off H_∞ performance (see Lemmas 5, 10) and mean-square-error performance (see Example 3).
Uncertain processes and measurement noises. Uncertain system model parameters.	1. Robust filter (see Example 4). 2. Robust smoother (see Example 4). 3. Robust filter or smoother with scaled inputs (see Lemma 3).

Table 1. Suggested approaches for different linear estimation problem conditions.

"A computer lets you make more mistakes than almost any invention in history, with the possible exceptions of tequila and hand guns." *Mitch Ratcliffe*

9.5 Problems

Problem 1 [31].

(i) Consider a system $\mathcal{G}$ having the state-space representation $\dot{x}(t) = Ax(t) + Bw(t)$, $y(t) = Cx(t)$. Show that if there exists a matrix $P = P^T > 0$ such that $\begin{bmatrix} A^TP + PA + C^TC & PB \\ B^TP & -\gamma^2 I \end{bmatrix} < 0$ then $x^T(T)Px(T) - x^T(0)Px(0) + \int_0^T y^T(t)y(t)dt \le \gamma^2 \int_0^T w^T(t)w(t)\ dt$.

(ii) Generalise (i) for the case where $y(t) = Cx(t) + Dw(t)$.

Problem 2. Consider a system $\mathcal{G}$ modelled by $\dot{x}(t) = A(t)x(t) + B(t)w(t)$, $y(t) = C(t)x(t) + D(t)w(t)$. Suppose that the Riccati differential equation

$$-\dot{P}(t) = P(t)(A(t) + B(t)M^{-1}(t)D^T(t)C(t)) + (A(t) + B(t)M^{-1}(t)D^T(t)C(t))^T P(t)$$
$$+\gamma^{-2}B(t)M^{-1}(t)B^T(t) + C^T(t)(I + D(t)M^{-1}(t)D^T(t))C(t),$$

$M(t) = \gamma^2 I - D^T(t)D(t) > 0$, has a solution on $[0, T]$. Show that $\|\mathcal{G}\|_\infty \le \gamma$ for any $w \in \mathcal{L}_2$. (Hint: define $V(x(t)) = x^T(t)P(t)\,x(t)$ and show that $\dot{V}(x(t)) + y^T(t)y(t) - \gamma^2 w^T(t)w(t) < 0$.)

Problem 3. For measurements $z(t) = y(t) + v(t)$ of a system realised by $\dot{x}(t) = A(t)x(t) + B(t)w(t)$, $y(t) = C(t)x(t)$, show that the map from the inputs $i = \begin{bmatrix} v \\ w \end{bmatrix}$ to the H_∞ fixed-interval smoother error $e(t|T)$ is

$$\begin{bmatrix} \dot{\tilde{x}}(t) \\ -\dot{\xi}(t) \\ e(t|T) \end{bmatrix} = \begin{bmatrix} A(t) - K(t)C(t) & 0 & B(t) & -K(t) \\ -C^T(t)R^{-1}(t)C(t) & A^T(t) - C^T(t)K^T(t) & 0 & -C^T(t)R^{-1}(t) \\ C(t) & R(t)K^T(t) & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ \xi(t) \\ w(t) \\ v(t) \end{bmatrix}.$$

Problem 4 [18].

(i) For a $\mathcal{G}$ modelled by $x_{k+1} = A_k x_k + B_k w_k$, $y_k = C_k x_k\ D_k w_k$, show that the existence of a solution to the Riccati difference equation

$$P_k = A_k^T P_{k+1} A_k + \gamma^{-2} A_k^T P_{k+1} B_k (I - \gamma^{-2} B_k^T P_{k+1} B_k)^{-1} B_k^T P_{k+1} A_k + C_k^T C_k$$

is sufficient for $x_k^T P_k x_k - x_{k+1}^T P_{k+1} x_{k+1} + y_k^T y_k - \gamma^2 w_k^T w_k < 0$. Hint: construct $x_{k+1}^T P_{k+1} x_{k+1} - x_k^T P_k x_k$ and show that

$$x_{k+1}^T P_{k+1} x_{k+1} - x_k^T P_k x_k + y_k^T y_k - \gamma^2 w_k^T w_k = -\gamma^{-2} p_k^T (I - \gamma^{-2} B_k^T P_{k+1} B_k)^{-1} p_k - x_k^T A_k^T P_{k+1} A_k x_k,$$

"A computer once beat me at chess, but it was no match for me at kick boxing." *Emo Philips* where $p_k = w_k - \gamma^{-2}(I - \gamma^{-2}B_k^T P_{k+1} B_k)^{-1} B_k^T P_{k+1} A_k x_k$.

(ii) Show that $-x_0^T P_0 x_0 + \sum_{k=0}^{N-1} y_k^T y_k - \gamma^2 \sum_{k=0}^{N-1} w_k^T w_k < 0$.

Problem 5. Now consider the model $x_{k+1} = A_k x_k + B_k w_k$, $y_k = C_k x_k + D_k w_k$ and show that the existence of a solution to the Riccati difference equation

$$P_k = A_k^T P_{k+1} A_k + \gamma^{-2}(A_k^T P_{k+1} B_k + C_k^T D_k)(I - \gamma^{-2}B_k^T P_{k+1} B_k - \gamma^{-2}D_k^T D_k)^{-1}(B_k^T P_{k+1} A_k + D_k^T C_k) + C_k^T C_k$$

is sufficient for $x_k^T P_k x_k - x_{k+1}^T P_{k+1} x_{k+1} + y_k^T y_k - \gamma^2 w_k^T w_k < 0$. Hint: define

$$p_k = w_k - \gamma^{-2}(I - \gamma^{-2}B_k^T P_{k+1} B_k - \gamma^{-2}D_k^T D_k)^{-1}(B_k^T P_{k+1} A_k + D_k^T C_k) .$$

Problem 6. Suppose that a predictor attains a H_∞ performance objective, that is, the conditions of Lemma 8 are satisfied. Show that using the predicted states to construct filtered output estimates $\hat{y}_{k/k}$ results in $\tilde{y}_{k/k} = y - \hat{y}_{k/k} \in \ell_2$.

9.6 Glossary

$\mathcal{L}_\infty$	The Lebesgue ∞-space defined as the set of continuous-time systems having finite ∞-norm.
$\mathcal{R}_{ei} \in \mathcal{L}_\infty$	The map $\mathcal{R}_{ei}$ from the inputs $i(t)$ to the estimation error $e(t)$ satisfies $\int_0^T e^T(t)e(t)dt - \gamma^2 \int_0^T i^T(t)i(t)dt < 0$. Therefore, $i \in \mathcal{L}_2$ implies $e \in \mathcal{L}_2$.
ℓ_∞	The Lebesgue ∞-space defined as the set of discrete-time systems having finite ∞-norm.
$\mathcal{R}_{ei} \in \ell_\infty$	The map $\mathcal{R}_{ei}$ from the inputs i_k to the estimation error e_k satisfies $\sum_{k=0}^{N-1} e_k^T e_k - \gamma^2 \sum_{k=0}^{N-1} i_k^T i_k < 0$. Therefore, $i \in \ell_2$ implies $e \in \ell_2$.

9.7 References

[1] J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle and J. Doyle, "Robustness of Cellular Functions", *Cell*, vol. 118, pp. 675 – 685, Dep. 2004.

[2] P. P. Vaidyanathan, "The Discrete-Time Bounded-Real Lemma in Digital Filtering", *IEEE Transactions on Circuits and Systems*, vol. 32, no. 9, pp. 918 – 924, Sep. 1985.

[3] I. Petersen, "A Riccati equation approach to the design of stabilizing controllers and observers for a class of uncertain linear systems", *IEEE Transactions on Automatic Control*, vol. 30, iss. 9, pp. 904 – 907, 1985.

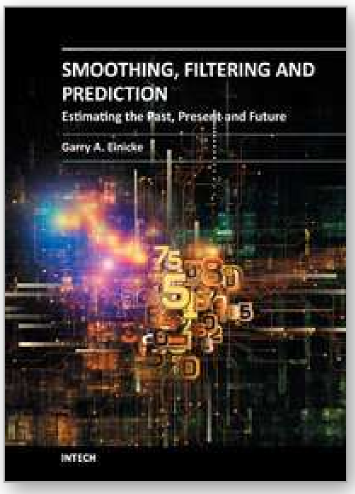
"On two occasions I have been asked, 'If you put into the machine wrong figures, will the right answers come out?' I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question." *Charles Babbage*

[4] J. C. Doyle, K. Glover, P. P. Khargonekar and B. A. Francis, "State-Space Solutions to Standard H_2 and H_∞ Control Problems", *IEEE Transactions on Automatic Control*, vol. 34, no. 8, pp. 831 – 847, Aug. 198
[5] D. J. N. Limebeer, M. Green and D. Walker, "Discrete-Time H_∞ Control, *Proceedings 28th IEEE Conference on Decision and Control*, Tampa, Florida, pp. 392 – 396, Dec. 198
[6] T. Basar, "Optimum performance levels for minimax filters, predictors and smoothers" *Systems and Control Letters*, vol. 16, pp. 309 – 317, 198
[7] D. S. Bernstein and W. M. Haddad, "Steady State Kalman Filtering with an H_∞ Error Bound", *Systems and Control Letters*, vol. 12, pp. 9 – 16, 198
[8] U. Shaked, "H_∞-Minimum Error State Estimation of Linear Stationary Processes", *IEEE Transactions on Automatic Control*, vol. 35, no. 5, pp. 554 – 558, May, 1990.
[9] T. Başar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, Series in Systems & Control: Foundations & Applications, Birkhäuser, Boston, 1991.
[10] K. M. Nagpal and P. P. Khargonekar, "Filtering and Smoothing in an H_∞ Setting", *IEEE Transactions on Automatic Control*, vol. 36, no. 2, pp 152 – 166, Feb. 1991.
[11] I. Yaesh and U. Shaked, "H_∞-Optimal Estimation – The Discrete Time Case", *Proceedings of the MTNS*, pp. 261 – 267, Jun. 1991.
[12] A. Stoorvogel, *The H_∞ Control Problem*, Series in Systems and Control Engineering, Prentice Hall International (UK) Ltd, Hertfordshire, 1992.
[13] U. Shaked and Y. Theodor, "H_∞ Optimal Estimation: A Tutorial", *Proceedings 31st IEEE Conference on Decision and Control*, pp. 2278 – 2286, Tucson, Arizona, Dec. 1992.
[14] C. E. de Souza, U. Shaked and M. Fu, "Robust H_∞ Filtering with Parametric Uncertainty and Deterministic Input Signal", *Proceedings 31st IEEE Conference on Decision and Control*, Tucson, Arizona, pp. 2305 – 2310, Dec. 1992.
[15] D. J. N. Limebeer, B. D. O. Anderson, P. P. Khargonekar and M. Green, "A Game Theoretic Approach to H^∞ Control for Time-Varying Systems", *SIAM Journal on Control and Optimization*, vol. 30, no. 2, pp. 262 – 283, Mar. 1992.
[16] I. Yaesh and Shaked, "Game Theory Approach to Finite-Time Horizon Optimal Estimation", *IEEE Transactions on Automatic Control*, vol. 38, no. 6, pp. 957 – 963, Jun. 1993.
[17] B. van Keulen, *H_∞ Control for Distributed Parameter Systems: A State-Space Approach*, Series in Systems & Control: Foundations & Applications, Birkhäuser, Boston, 1993.
[18] L. Xie, C. E. De Souza and Y. Wang, "Robust Control of Discrete Time Uncertain Dynamical Systems", *Automatica*, vol. 29, no. 4, pp. 1133 – 1137, 1993.
[19] Y. Theodore, U. Shaked and C. E. de Souza, "A Game Theory Approach To Robust Discrete-Time H_∞-Estimation", *IEEE Transactions on Signal Processing*, vol. 42, no. 6, pp. 1486 – 1495, Jun. 1994.
[20] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, vol. 15, SIAM, Philadelphia, 1994.
[21] M. Green and D. J. N. Limebeer, *Linear Robust Control*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1995.

"Never trust a computer you can't throw out a window." *Stephen Gary Wozniak*

[22] S. O. R. Moheimani, A. V. Savkin and I. R. Petersen, "Robust Filtering, Prediction, Smoothing and Observability of Uncertain Systems", *IEEE Transactions on Circuits and Systems*, vol. 45, no. 4, pp. 446 – 457, Apr. 1998.
[23] J. C. Geromel, "Optimal Linear Filtering Under Parameter Uncertainty", *IEEE Transactions on Signal Processing,* vol. 47, no. 1, pp. 168 – 175, Jan. 199
[24] T. Başar and G. J. Oldsder, *Dynamic Noncooperative Game Theory*, Second Edition, SIAM, Philadelphia, 199
[25] B. Hassibi, A. H. Sayed and T. Kailath, *Indefinite-Quadratic Estimation and Control: A Unified Approach to H^∞ and H^2 Theories*, SIAM, Philadelphia, 199
[26] G. A. Einicke and L. B. White, "Robust Extended Kalman Filtering", *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2596 – 2599, Sep., 199
[27] Z. Wang and B. Huang, "Robust H_2/H_∞ Filtering for Linear Systems with Error Variance Constraints", *IEEE Transactions on Signal Processing,* vol. 48, no. 9, pp. 2463 – 2467, Aug. 2000.
[28] G. A. Einicke, "Optimal and Robust Noncausal Filter Formulations", *IEEE Transactions on Signal Processing,* vol. 54, no. 3, pp. 1069 - 1077, Mar. 2006.
[29] A. Saberi, A. A. Stoorvogel and P. Sannuti, *Filtering Theory With Applications to Fault Detection, Isolation, and Estimation*, Birkhäuser, Boston, 2007.
[30] F. L. Lewis, L. Xie and D. Popa, *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*, Second Edition, Series in Automation and Control Engineering, Taylor & Francis Group, LLC, 2008.
[31] J.-C. Lo and M.-L. Lin, "Observer-Based Robust H_∞ Control for Fuzzy Systems Using Two-Step Procedure", *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 3, pp. 350 – 359, Jun. 2004.
[32] E. Blanco, P. Neveux and H. Thomas, "The H_∞ Fixed-Interval Smoothing Problem for Continuous Systems", *IEEE Transactions on Signal Processing,* vol. 54, no. 11, pp. 4085 – 4090, Nov. 2006.
[33] G. A. Einicke, "A Solution to the Continuous-Time H-infinity Fixed-Interval Smoother Problem", *IEEE Transactions on Automatic Control,* vol. 54, no. 12, pp. 2904 – 2908, Dec. 200
[34] G. A. Einicke, "Asymptotic Optimality of the Minimum-Variance Fixed-Interval Smoother", *IEEE Transactions on Signal Processing,* vol. 55, no. 4, pp. 1543 – 1547, Apr. 2007.
[35] G. A. Einicke, J. C. Ralston, C. O. Hargrave, D. C. Reid and D. W. Hainsworth, "Longwall Mining Automation, An Application of Minimum-Variance Smoothing", *IEEE Control Systems Magazine*, vol. 28, no. 6, pp. 28 – 37, Dec. 2008.
[36] K. Ogata, *Discrete-time Control Systems*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1987.

"The most amazing achievement of the computer software industry is its continuing cancellation of the steady and staggering gains made by the computer hardware industry." *Henry Petroski*



Smoothing, Filtering and Prediction - Estimating The Past, Present and Future
Edited by

ISBN 978-953-307-752-9
Hard cover, 276 pages
Publisher InTech
Published online 24, February, 2012
Published in print edition February, 2012

This book describes the classical smoothing, filtering and prediction techniques together with some more recently developed embellishments for improving performance within applications. It aims to present the subject in an accessible way, so that it can serve as a practical guide for undergraduates and newcomers to the field. The material is organised as a ten-lecture course. The foundations are laid in Chapters 1 and 2, which explain minimum-mean-square-error solution construction and asymptotic behaviour. Chapters 3 and 4 introduce continuous-time and discrete-time minimum-variance filtering. Generalisations for missing data, deterministic inputs, correlated noises, direct feedthrough terms, output estimation and equalisation are described. Chapter 5 simplifies the minimum-variance filtering results for steady-state problems. Observability, Riccati equation solution convergence, asymptotic stability and Wiener filter equivalence are discussed. Chapters 6 and 7 cover the subject of continuous-time and discrete-time smoothing. The main fixed-lag, fixed-point and fixed-interval smoother results are derived. It is shown that the minimum-variance fixed-interval smoother attains the best performance. Chapter 8 attends to parameter estimation. As the above-mentioned approaches all rely on knowledge of the underlying model parameters, maximum-likelihood techniques within expectation-maximisation algorithms for joint state and parameter estimation are described. Chapter 9 is concerned with robust techniques that accommodate uncertainties within problem specifications. An extra term within Riccati equations enables designers to trade-off average error and peak error performance. Chapter 10 rounds off the course by applying the afore-mentioned linear techniques to nonlinear estimation problems. It is demonstrated that step-wise linearisations can be used within predictors, filters and smoothers, albeit by forsaking optimal performance guarantees.

InTech Europe
University Campus STeP Ri

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821