Les Cahiers du GERAD ISSN: 0711–2440

A Kalman Filter Design Based on the Performance/Robustness Tradeoff

Xu, H.
Mannor, S.

G–2008–53

July 2008

La publication de ces rapports de recherche bénéficie d'une subvention du Fonds québécois de la recherche sur la nature et les technologies.

A Kalman Filter Design Based on the Performance/Robustness Tradeoff

Huan Xu
Shie Mannor

GERAD and Department of Electrical and Computer Engineering
McGill University
Montreal (Quebec) Canada

xuhuan@cim.mcgill.ca, shie.mannor@mcgill.ca

July 2008

Les Cahiers du GERAD

G–2008–53

Abstract

We consider filter design of a linear system with parameter uncertainty. In contrast to the robust Kalman filter which focuses on a worst case analysis, we propose a design methodology based on iteratively solving a tradeoff problem between nominal performance and robustness to the uncertainty. Our proposed filter can be computed online efficiently, is steady-state stable, and is less conservative than the robust filter.

Résumé

Nous étudions la conception de filtres pour systèmes linéaires à paramètres incertains. Alors que le filtre de Kalman dit robuste est fondé sur l'analyse du pire cas, nous proposons ici une alternative fondée sur la solution répétée de problèmes de filtrage où une mesure combinée d'erreur d'estimation nominale et de robustesse à l'égard des incertitudes de paramètres est minimisée. Le filtre proposé peut être calculé en ligne efficacement, est stable et est moins conservateur que le filtre robuste.

Acknowledgments: The authors would like to thank Constantine Caramanis for interesting discussions and comments. The authors are especially thankful to the anonymous reviewers whose comments led to substantial improvements of the paper.

1 Introduction

The Kalman filter addresses the estimation problem for linear systems, and is widely used in many fields including control, finance, communication etc (e.g.,[1, 2]). One central assumption of the Kalman filter is that the underlying state-space model is exactly known. In practice, this assumption is often violated, i.e., the parameters we use as the system dynamics (referred as *nominal parameters* hereafter) are only guesses of the unknown true parameters. It is reported (e.g.,[3, 4, 5]) that in this case, the performance of the Kalman filter can deteriorate significantly. In [6], Sayed proposed a filtering framework based on a worst-case analysis (hereafter referred to as the *robust filter*), i.e., instead of iteratively minimizing the regularized residual norm as the standard Kalman filter does, the robust filter minimizes the worst-possible regularized residual norm over the set of admissible uncertainty.

Empirical studies show that the Kalman filter and the robust filter perform well in different setups: the performance (measured by the steady-state error variance) of the robust filter is significantly better than the Kalman filter when the uncertainty is large; but under small uncertainty, its performance is not satisfactory, indicating over-conservativeness comparing to the standard Kalman filter. Furthermore, the robust filter usually has a slower transient response. Therefore, a filter that exhibits a similar performance to the better filter under all cases is desirable.

In this paper, we present a new filter design approach to achieve this goal by interpolating the standard Kalman filter and the robust filter. To be more specific, in each iteration, the proposed filter finds a *Pareto efficient* filtered estimation by minimizing the convex combination of the nominal regularized residue (the criterion of the Kalman filter) and the worst-case regularized residue (the criterion of the robust filter). This approach leads to an optimization problem that can be solved recursively similarly to the Kalman filter and hence can be applied on-line. The proposed filter is stable and achieves bounded error-variance. Simulation results show that the proposed filter exhibits a similar performance to the better one between the Kalman filter and the robust filter. That is, the performance of proposed filter is similar to the Kalman filter under small uncertainty, and is comparable to the robust filter under large uncertainty. Therefore, the proposed filter is suitable for a wider range of problem setups.

We need to point out that the proposed filter achieves good tradeoff because it is the only interpolating method that achieves Pareto efficiency between the *nominal performance* given by the nominal residue and the *robustness* given by the worst residue. There are several other "robust" filters designs based on $\mathcal{H}_2/\mathcal{H}_\infty$ robust control (e.g., [7, 8, 9, 10, 11, 12]), set-inclusive robust optimization (e.g., [13, 14]), and guaranteed error variance minimization (e.g., [12, 15, 16]). The main difference is that these methods performs de-regularization, and hence need to check certain existence condition in each iteration. If the existence condition fails at some step, the robustness of the filter is not valid anymore. Furthermore, de-regularization leads to a computationally expensive algorithm, and hence is often not suitable in on-line application. See [6] for a more detailed comparison among different robust filter design methodologies.

The paper is organized as follows. We formulate the filtering design as an optimization problem in Section 2, and show how to solve it in Section 3, which leads to the recursive formula of the proposed filter in Section 4. In Section 5 and Section 6 we investigate the theoretical and empirical behavior of the proposed filter respectively. Some concluding remarks are given in Section 7.

Notations: We use capital letters and boldface letters to denote matrices and column vectors respectively. Without further explanations, $\|\cdot\|$ stands for Euclidean norm for vectors, and largest singular value for matrices. The notation $col\{\mathbf{a},\mathbf{b}\}$ stands for a column vector with entries $\mathbf{a}$ and $\mathbf{b}$, and $diag\{A,B\}$ denotes a block diagonal matrix with entries A and B. Given a column vector $\mathbf{z}$ and a positive definite matrix W, $\|\mathbf{z}\|_W^2$ stands for $\mathbf{z}^\top W\mathbf{z}$.

2 Filter formulation

We consider the following system:

$$\begin{aligned}\mathbf{x}_{i+1} =&(F_i + M_i\Delta_i E_{f,i})\mathbf{x}_i + (G_i + M_i\Delta_i E_{g,i})\mathbf{u}_i,\\ \mathbf{y}_i =&H_i\mathbf{x}_i + \mathbf{v}_i, \quad i = 0, 1, \cdots .\end{aligned} \tag{2.1}$$

Here, F_i , G_i, M_i, $E_{f,i}$ and $E_{g,i}$ are known matrices and Δ_i are unknown matrix with $\|\Delta_i\| \leq 1$. The variance of the initial state $\mathbf{x}_0$ is Π_0, and the driving noises $\mathbf{u}_i$ and $\mathbf{v}_i$ are white, zero mean and uncorrelated, with variance Q_i and R_i respectively. This formulation is standard in robust filter design [6, 12]. We denote the estimation of $\mathbf{x}_i$ given observation $\{\mathbf{y}_0, \cdots , \mathbf{y}_j\}$ by $\hat{\mathbf{x}}_{i|j}$, and denote its error variance by $P_{i|j}$. Furthermore, $\hat{\mathbf{x}}_i$ and P_i denote $\hat{\mathbf{x}}_{i|i-1}$ and $P_{i|i-1}$ respectively. We assume $P_{i|i}$ to be invertible, which can be relaxed because the final recursion form is independent of $P_{i|i}^{-1}$.

Both the Kalman filter and the Robust filter iteratively find the optimal/robust smoothing estimation and propagate them respectively (e.g., [1, 2, 6]), i.e.,

$$\begin{aligned}&\text{Kalman Filter:}\\ &(\hat{\mathbf{x}}_{i|i+1}, \hat{\mathbf{u}}_{i|i+1}) := \arg\min_{\mathbf{x}_i,\mathbf{u}_i} \Big\{\|\mathbf{x}_i - \hat{\mathbf{x}}_{i|i}\|^2_{P_{i|i}^{-1}} + \|\mathbf{u}_i\|^2_{Q_i^{-1}}\\ &\qquad\qquad + \|\mathbf{y}_{i+1} - H_{i+1}\mathbf{x}_{i+1}\|^2_{R_{i+1}^{-1}} \big|\Delta_i = 0\Big\},\\ &\hat{\mathbf{x}}_{i+1|i+1} := F_i\hat{\mathbf{x}}_{i|i+1} + G_i\hat{\mathbf{u}}_{i|i+1};\\ &\text{Robust Filter:}\\ &(\hat{\mathbf{x}}_{i|i+1}, \hat{\mathbf{u}}_{i|i+1}) := \arg\min_{\mathbf{x}_i,\mathbf{u}_i}\max_{\|\Delta_i\|\leq 1} \Big\{\|\mathbf{x}_i - \hat{\mathbf{x}}_{i|i}\|^2_{P_{i|i}^{-1}}\\ &\qquad\qquad + \|\mathbf{u}_i\|^2_{Q_i^{-1}} + \|\mathbf{y}_{i+1} - H_{i+1}\mathbf{x}_{i+1}\|^2_{R_{i+1}^{-1}}\Big\},\\ &\hat{\mathbf{x}}_{i+1|i+1} := F_i\hat{\mathbf{x}}_{i|i+1} + G_i\hat{\mathbf{u}}_{i|i+1}.\end{aligned}$$

Notice here, the cost function for the Kalman filter is the error variance under the nominal parameters, whereas the cost function for the robust filter is the worst case error variance. Hence the former criterion stands for the nominal performance of the smoothed estimation, and the latter represents how robust the smoothed estimation is. Ideally, a good estimation should perform well (in the sense of Pareto efficiency) for both criteria. This is equivalent to a minimizer of their convex combination, which leads to the proposed filter:

$$
\begin{aligned}
&\text{Proposed Filter: Fix } \alpha \in (0,1)\\
(\hat{\mathbf{x}}_{i|i+1}, \hat{\mathbf{u}}_{i|i+1}) &:= \arg\min_{\mathbf{x}_i,\mathbf{u}_i} \Big\{ \alpha \big[\|\mathbf{x}_i - \hat{\mathbf{x}}_{i|i}\|^2_{P^{-1}_{i|i}} + \|\mathbf{u}_i\|^2_{Q_i^{-1}} \\
&\quad + \|\mathbf{y}_{i+1} - H_{i+1}\mathbf{x}_{i+1}\|^2_{R^{-1}_{i+1}} \big| \Delta_i = 0 \big] \\
&\quad + (1-\alpha) \max_{\|\Delta_i\| \le 1} \big[\|\mathbf{x}_i - \hat{\mathbf{x}}_{i|i}\|^2_{P^{-1}_{i|i}} \\
&\quad + \|\mathbf{u}_i\|^2_{Q_i^{-1}} + \|\mathbf{y}_{i+1} - H_{i+1}\mathbf{x}_{i+1}\|^2_{R^{-1}_{i+1}} \big] \Big\}, \\
\hat{\mathbf{x}}_{i+1|i+1} &:= F_i \hat{\mathbf{x}}_{i|i+1} + G_i \hat{\mathbf{u}}_{i|i+1}.
\end{aligned}
\tag{2.2}
$$

Notice that, since both criteria are convex functions, not only any minimizer of the convex combination is Pareto efficient, but any Pareto efficient solution must minimize the convex combination for some α. Hence, this formulation computes all the solutions that achieve good tradeoff between the nominal performance and the robustness. This is different from other interpolation such as shrinking the uncertainty set, where the Pareto efficiency is not guaranteed.

3 Solving the Minimization Problem

To minimize Problem (2.2), we denote

$$
\begin{aligned}
&\mathbf{z} \triangleq col\{\mathbf{x}_i - \hat{\mathbf{x}}_{i|i},\, \mathbf{u}_i\}; \quad \mathbf{b} \triangleq \mathbf{y}_{i+1} - H_{i+1}F_i\hat{\mathbf{x}}_{i|i};\\
&A \triangleq H_{i+1}[F_i,\, G_i]; \quad T \triangleq diag\{P^{-1}_{i|i},\, Q_i^{-1}\};\\
&W \triangleq R^{-1}_{i+1}; \quad D \triangleq H_{i+1}M_i; \quad E_a \triangleq [E_{f,i},\, E_{g,i}];\\
&\mathbf{t} \triangleq -E_{f,i}\hat{\mathbf{x}}_{i|i}; \quad \phi(\mathbf{z}) \triangleq \|E_a\mathbf{z} - \mathbf{t}\|.
\end{aligned}
$$

We can rewritten Problem (2.2) as

$$
\begin{aligned}
\arg\min_{\mathbf{z}} : C(\mathbf{z}) \triangleq \mathbf{z}^\top T\mathbf{z} + \alpha(A\mathbf{z} - \mathbf{b})^\top W(A\mathbf{z} - \mathbf{b})&\\
+ (1-\alpha) \max_{\|\mathbf{y}\| \le \phi(\mathbf{z})} \|A\mathbf{z} - \mathbf{b} + D\mathbf{y}\|^2_W&,
\end{aligned}
\tag{3.1}
$$

Problem (3.1) is a bilevel optimization problem which is generally NP-hard. However, following a similar argument as [17], we show this special problem can be efficiently solved by converting into a *unimodal* scalar optimization problem. Before giving the main result of this section, we need to define the following functions of $\lambda \in \big[\|D^\top WD\|, +\infty\big)$:

$$
\begin{aligned}
\overline{W}(\lambda) &\triangleq W + (1-\alpha) W D (\lambda I - D^\top W D)^\dagger D^\top W, \\
\mathbf{z}^o(\lambda) &\triangleq \arg\min_{\mathbf{z}} \Big\{ \mathbf{z}^\top T \mathbf{z} + (A\mathbf{z}-\mathbf{b})^\top \overline{W}(\lambda)(A\mathbf{z}-\mathbf{b}) \\
&\qquad + (1-\alpha)\lambda\phi^2(\mathbf{z}) \Big\}, \\
G(\lambda) &\triangleq \min_{\mathbf{z}} \Big\{ \mathbf{z}^\top T \mathbf{z} + (A\mathbf{z}-\mathbf{b})^\top \overline{W}(\lambda)(A\mathbf{z}-\mathbf{b}) \\
&\qquad + (1-\alpha)\lambda\phi^2(\mathbf{z}) \Big\} \\
&= \mathbf{z}^{o\top}(\lambda) T \mathbf{z}^o(\lambda) + \big(A\mathbf{z}^o(\lambda)-\mathbf{b}\big)^\top \overline{W}(\lambda)\big(A\mathbf{z}^o(\lambda)-\mathbf{b}\big) \\
&\qquad + (1-\alpha)\lambda\phi^2\big(\mathbf{z}^o(\lambda)\big).
\end{aligned}
$$

Here, $(\cdot)^\dagger$ stands for the pseudo inverse of a matrix. Note that $T > 0$, $\phi(\cdot)$ is convex, and $\lambda \geq \|D^\top W D\|$ implies $\overline{W}(\lambda) \geq 0$, hence the definitions of $\mathbf{z}^o(\lambda)$ and $G(\lambda)$ are valid, because the part in the curled bracket is strictly convex on $\mathbf{z}$. Therefore, for any given λ we can evaluate $\mathbf{z}^o(\lambda)$ and $G(\lambda)$. The next theorem shows that the optimal $\mathbf{z}$ for Problem (3.1) can be evaluated by minimizing $G(\lambda)$ using line search and substituting the minimizer into $\mathbf{z}^o(\cdot)$.

Theorem 1

1. *Let* $\lambda^o \triangleq \arg\min_{\lambda \geq \|D^\top W D\|} G(\lambda)$, *we have*

$$
\arg\min_{\mathbf{z}} C(\mathbf{z}) = \mathbf{z}^o(\lambda^o); \quad \min_{\mathbf{z}} C(\mathbf{z}) = G(\lambda^o).
$$

2. *On* $\lambda \geq \|D^\top W D\|$, $G(\lambda)$ *has only one local minimum, which is also its global minimum.*

Proof. Define $R(\mathbf{z},\mathbf{y}) \triangleq (A\mathbf{z} - \mathbf{b} + H\mathbf{y})^\top W (A\mathbf{z} - \mathbf{b} + H\mathbf{y})$ and $\hat{W}(\lambda) \triangleq W + WD(\lambda I - D^\top W D)^\dagger D^\top W$, for $\lambda \in \big[\|D^\top W D\|, +\infty\big)$. Hence $\overline{W}(\lambda) = \alpha W + (1-\alpha)\hat{W}(\lambda)$. Lemma 1 describes the property of $R(\mathbf{z},\mathbf{y})$; its proof can be found in [17].

Lemma 1

(a) *Function* $\max_{\|\mathbf{y}\| \leq \phi(\mathbf{z})} R(\mathbf{z},\mathbf{y})$ *is convex on* $\mathbf{z}$.

(b) *For all* $\mathbf{z}$,

$$
\max_{\|\mathbf{y}\| \leq \phi(\mathbf{z})} R(\mathbf{z},\mathbf{y}) = \\
\min_{\lambda \geq \|D^\top W D\|} (A\mathbf{z}-\mathbf{b})^\top \hat{W}(\lambda)(A\mathbf{z}-\mathbf{b}) + \lambda\phi^2(\mathbf{z}).
$$

(c) $\lambda^o(\mathbf{z}) \triangleq \arg\min_{\lambda \geq \|D^\top W D\|} (A\mathbf{z}-\mathbf{b})^\top \hat{W}(\lambda)(A\mathbf{z}-\mathbf{b}) + \lambda\phi^2(\mathbf{z})$ *is well defined and continuous.*

Therefore,

$$\begin{aligned}\min_{\mathbf{z}} C(\mathbf{z}) &= \min_{\mathbf{z}} \Big\{ \mathbf{z}^\top T\mathbf{z} + \alpha(A\mathbf{z}-\mathbf{b})^\top W(A\mathbf{z}-\mathbf{b}) \\ &\quad + (1-\alpha) \max_{\|\mathbf{y}\|\le\phi(\mathbf{z})} R(\mathbf{z},\mathbf{y}) \Big\} \\ &= \min_{\mathbf{z}} \Big\{ \mathbf{z}^\top T\mathbf{z} + \alpha(A\mathbf{z}-\mathbf{b})^\top W(A\mathbf{z}-\mathbf{b}) + (1-\alpha) \\ &\quad \times \min_{\lambda\ge\|D^\top WD\|} \big[(A\mathbf{z}-\mathbf{b})^\top \hat{W}(\lambda)(A\mathbf{z}-\mathbf{b}) + \lambda\phi^2(\mathbf{z})\big] \Big\} \\ &= \min_{\lambda\ge\|D^\top WD\|} \min_{\mathbf{z}} \Big\{ \mathbf{z}^\top T\mathbf{z} + (A\mathbf{z}-\mathbf{b})^\top \overline{W}(\lambda)(A\mathbf{z}-\mathbf{b}) \\ &\quad + (1-\alpha)\lambda\phi^2(\mathbf{z}) \Big\} = \min_{\lambda\ge\|D^\top WD\|} G(\lambda).\end{aligned}$$

We now show that $G(\cdot)$ is unimodal. Denote $H(\mathbf{z},\lambda) \triangleq \mathbf{z}^\top T\mathbf{z} + (A\mathbf{z}-\mathbf{b})^\top \overline{W}(\lambda)(A\mathbf{z}-\mathbf{b}) + (1-\alpha)\lambda\phi^2(\mathbf{z})$. Observe that $C(\mathbf{z}) = \min_{\lambda\ge\|D^\top WD\|} H(\mathbf{z},\lambda)$ and

$$\begin{aligned}\lambda^o(\mathbf{z}) &= \arg\min_{\lambda\ge\|D^\top WD\|} (A\mathbf{z}-\mathbf{b})^\top \hat{W}(\lambda)(A\mathbf{z}-\mathbf{b}) + \lambda\phi^2(\mathbf{z}) \\ &= \arg\min_{\lambda\ge\|D^\top WD\|} \Big\{ \mathbf{z}^\top T\mathbf{z} + \alpha(A\mathbf{z}-\mathbf{b})^\top W(A\mathbf{z}-\mathbf{b}) \\ &\quad + (1-\alpha)\big[(A\mathbf{z}-\mathbf{b})^\top \hat{W}(\lambda)(A\mathbf{z}-\mathbf{b}) + \lambda\phi^2(\mathbf{z})\big] \Big\} \\ &= \arg\min_{\lambda\ge\|D^\top WD\|} H(\mathbf{z},\lambda).\end{aligned}$$

Hence $G(\lambda) = \min_{\mathbf{z}} H(\mathbf{z},\lambda)$. Note that $C(\mathbf{z})$ is strictly convex and goes to infinity whenever $\|\mathbf{z}\| \uparrow \infty$, which implies $C(\mathbf{z})$ is unimodal and has a unique global minimum. Also note, $H(\mathbf{z},\lambda)$ has the following property: fix one variable, then it is a unimodal function of the other variable and achieves unique minimum on its domain. This, combined with the continuity of $\lambda^o(\mathbf{z})$, establishes the unimodality of $G(\cdot)$ by applying Lemma C.2 in [17]. □

Notice that, $\phi(\mathbf{z}) = \|E_a\mathbf{z} - \mathbf{t}\|$ yields a closed form for $\mathbf{z}^o(\cdot)$:

$$\begin{aligned}\mathbf{z}^o(\lambda) &= \Big(T + A^\top \overline{W}(\lambda)A + (1-\alpha)\lambda E_a^\top E_a\Big)^{-1} \\ &\quad \times \Big(A^\top \overline{W}(\lambda)\mathbf{b} + (1-\alpha)\lambda E_a^\top \mathbf{t}\Big).\end{aligned} \tag{3.2}$$

4 Recursion formula of the filter

Substituting Equation (3.2) into Problem (2.2) and with some algebra, we obtain the recursion formula of the proposed filter. We present the prediction form which propagates $\{\hat{\mathbf{x}}_i, P_i\}$, whereas the measurement-Update form which propagates $\{\hat{\mathbf{x}}_{i|i}, P_{i|i}\}$ can be found in the Appendix. The recursion formula of the proposed filter is a modified version of the Robust filter, where $*$ are the modifications. In addition, $G(\lambda)$ and hence λ^o are also different.

Algorithm 1 Prediction form

1. *Initialize:* $\hat{\mathbf{x}}_0 := 0, \qquad P_0 := \Pi_0, \qquad \hat{R}_0 := R_0.$
2. *Given* $\hat{R}_i, H_i, P_i$, *calculate:*

$$\begin{aligned} P_{i|i} &:= (P_i^{-1} + H_i^\top \hat{R}_i^{-1} H_i)^{-1} \\ &= P_i - P_i H_i^\top (\hat{R}_i + H_i P_i H_i^\top)^{-1} H_i P_i. \end{aligned}$$

3. *Recursion: Construct and minimize* $G(\lambda)$ *over* $(\|M_i^\top H_{i+1}^\top R_{i+1}^{-1} H_{i+1} M_i\|, +\infty)$. *Let the optimal value be* λ_i^o. *Computing the following values:*

$$\begin{aligned} \hat{\lambda}_i :=& (1-\alpha)\lambda_i^o \qquad * \\ \overline{R}_{i+1} :=& R_{i+1} - \lambda^{o-1} H_{i+1} M_i M_i^\top H_{i+1}^\top \\ \hat{R}_{i+1}^{-1} :=& \alpha R_{i+1}^{-1} + (1-\alpha)\overline{R}_{i+1}^{-1} \qquad * \\ \hat{Q}_i^{-1} :=& Q_i^{-1} + \hat{\lambda}_i E_{g,i}^\top \big[I + \hat{\lambda}_i E_{f,i} P_{i|i} E_{f,i}^\top\big]^{-1} E_{g,i} \\ \hat{P}_{i|i} :=& (P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^\top E_{f,i})^{-1} \\ =& P_{i|i} - P_{i|i} E_{f,i}^\top (\hat{\lambda}_i^{-1} I + E_{f,i} P_{i|i} E_{f,i}^\top)^{-1} E_{f,i} P_{i|i} \\ \hat{G}_i :=& G_i - \hat{\lambda}_i F_i \hat{P}_{i|i} E_{f,i}^\top E_{g,i} \\ \hat{F}_i :=& (F_i - \hat{\lambda}_i \hat{G}_i \hat{Q}_i E_{g,i}^\top E_{f,i})(I - \hat{\lambda}_i \hat{P}_{i|i} E_{f,i}^\top E_{f,i}) \\ \overline{H}_i^\top :=& \Big[H_i^\top \hat{R}_i^{-\top/2},\ \sqrt{\hat{\lambda}_i}\,\Big] \\ \overline{R}_{e,i} :=& I + \overline{H}_i P_i \overline{H}_i^\top \\ \overline{K}_i :=& F_i P_i \overline{H}_i^\top \\ P_{i+1} :=& F_i P_i F_i^\top - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top \\ \mathbf{e}_i :=& \mathbf{y}_i - H_i \hat{\mathbf{x}}_i \\ \hat{\mathbf{x}}_{i+1} :=& \hat{F}_i \hat{\mathbf{x}}_i + \hat{F}_i P_{i|i} H_i^\top \hat{R}_i^{-1} \mathbf{e}_i \\ =& \hat{F}_i \hat{\mathbf{x}}_i + \hat{F}_i P_i H_i^\top R_{e,i}^{-1} \mathbf{e}_i. \end{aligned}$$

5 Steady-state Analysis

In this section we studies steady-state characteristics of the proposed filter, namely closed-loop stability and bounded error-variance. Similarly to [6], we restrict our discussion to uncertainty models where all parameters are stationary, except Δ_i, and drop the subscript i. Further assume the uncertainty only appears in F matrix. Hence, we have $\hat{Q} = Q$ and $\hat{G} = G$. In addition, we approximate λ^o by setting $\lambda^o := (1+\beta)\|M^\top H^\top R^{-1} H M\|$ for some $\beta > 0$. The next theorem shows that the proposed filter converges to a stable steady-state filter.

Theorem 2 *Assume that* $\{F, \overline{H}\}$ *is detectable and* $\{F, GQ^{1/2}\}$ *is stabilizable. Then, for any initial condition* $\Pi_0 > 0$, *the Riccati variable* P_i *converges to the unique solution of*

$$P = FPF^\top - FP\overline{H}^\top (I + \overline{H} P \overline{H}^\top)^{-1} \overline{H} P F^\top + GQG^\top. \tag{5.1}$$

Furthermore, the solution P is semi-definite positive, and the steady state closed loop matrix $F_p \triangleq \hat{F}[I - PH^\top R_e^{-1} H]$ is stable.

Proof. The closed loop formula for $\hat{\mathbf{x}}$ is

$$\begin{aligned}\hat{\mathbf{x}}_{i+1} &= \hat{F}_i \hat{\mathbf{x}}_i + \hat{F}_i P_i H^\top R_{e,i}^{-1}[\mathbf{y}_i - H\hat{\mathbf{x}}_i] \\ &= \hat{F}_i[I - P_i H^\top R_{e,i}^{-1} H]\hat{\mathbf{x}}_i + \hat{F}_i P_i H^\top R_{e,i}^{-1}\mathbf{y}_i.\end{aligned}$$

Notice that

$$\begin{aligned}&F\left[I - P_i \overline{H}^\top \overline{R}_{e,i}^{-1} \overline{H}\right] \\ &= F\left[P_i - P_i \overline{H}^\top (I + \overline{H} P_i \overline{H}^\top)^{-1} \overline{H} P_i\right] P_i^{-1} \\ &= F(P_i^{-1} + \overline{H}^\top \overline{H})^{-1} P_i^{-1}.\end{aligned}$$

Now consider the closed loop gain

$$\begin{aligned}F_{p,i} &\triangleq \hat{F}_i[I - P_i H^\top R_{e,i}^{-1} H] \\ =&F\left[I - \hat{\lambda}(P_i^{-1} + \overline{H}^\top \overline{H})^{-1} E_f^\top E_f\right]\left[I - P_i H^\top R_{e,i}^{-1} H\right] \\ =&F(P_i^{-1} + \overline{H}^\top \overline{H})^{-1}\left[P_i^{-1} + \overline{H}^\top \overline{H} - \hat{\lambda} E_f^\top E_f\right] \\ &\times \left[I - P_i H^\top R_{e,i}^{-1} H\right] \\ =&F(P_i^{-1} + \overline{H}^\top \overline{H})^{-1}(P_i^{-1} + H^\top \hat{R}^{-1} H) \\ &\times \left[P_i - P_i H^\top R_{e,i}^{-1} H P_i\right] P_i^{-1} \\ =&F(P_i^{-1} + \overline{H}^\top \overline{H})^{-1}(P_i^{-1} + H^\top \hat{R}^{-1} H) \\ &\times \left[P_i - P_i H^\top (\hat{R}_i + H P_i H^\top)^{-1} H P_i\right] P_i^{-1} \\ =&F(P_i^{-1} + \overline{H}^\top \overline{H})^{-1} P_i^{-1} = F\left[I - P_i \overline{H}^\top \overline{R}_{e,i}^{-1} \overline{H}\right].\end{aligned}$$

The positive definiteness of $\hat{R}$ guarantees that $\overline{H}$ is well defined. Hence, detectability of $\{F, \overline{H}\}$ and the stablizability of $\{F, GQ^{1/2}\}$ guarantee that P_i converges to the unique positive semi-definite solution P of Equation (5.1), which stabilizes the matrix $F[I - P\overline{H}^\top (I + \overline{H} P \overline{H}^\top)^{-1} \overline{H}]$. The stability follows for this matrix equals to the steady state closed loop gain F_p. □

Further assume that the system is quadratically stable, i.e, there exists a matrix $V > 0$ such that

$$V - [F + M\Delta E_f]^\top V[F + M\Delta E_f] > 0, \quad \forall \|\Delta\| \leq 1.$$

which is equivalent to a stable F and a bounded norm $\|E_f(zI - F)^{-1}M\|_\infty < 1$. Denote

$$\begin{aligned}\mathcal{F} &\triangleq \begin{bmatrix} F - F_p P H^\top \hat{R}^{-1} H & F - F_p - F_p P H^\top \hat{R}^{-1} H \\ F_p P H^\top \hat{R}^{-1} H & F_p + F_p P H^\top \hat{R}^{-1} H \end{bmatrix}, \\ \mathcal{G} &\triangleq \begin{bmatrix} G & -F_p P H^\top \hat{R}^{-1} H \\ 0 & F_p P H^\top \hat{R}^{-1} H \end{bmatrix}.\end{aligned}$$

The following theorem shows that the error-variance is uniformly bounded, which is equivalent to saying that the extended system is stable and has a $\mathcal{H}_\infty$ norm less than 1.

Theorem 3 *Let $\tilde{x}_i$ be the estimation error, for any $\mathcal{P} > 0$ such that*

$$\mathcal{P} - \left\{\mathcal{F} + \begin{bmatrix} M \\ 0 \end{bmatrix} \Delta[E_f \;\; E_f]\right\} \mathcal{P} \left\{\mathcal{F} + \begin{bmatrix} M \\ 0 \end{bmatrix} \Delta[E_f \;\; E_f]\right\}^\top - \mathcal{G} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \mathcal{G}^\top \geq 0, \quad \forall \|\Delta\| \leq 1;$$

the error variance satisfies $\lim_{i\to\infty} \mathbb{E}\tilde{x}_i\tilde{x}_i^\top \leq \mathcal{P}_{11}$, where $\mathcal{P}_{11}$ is the $(1,1)$ block entries of $\mathcal{P}$. Furthermore, such $\mathcal{P}$ is guaranteed to exist.

Proof. Define estimation error $\tilde{\mathbf{x}}_i \triangleq \mathbf{x}_i - \hat{\mathbf{x}}_i$, and

$$\delta\mathcal{F}_i \triangleq \begin{bmatrix} M\Delta_i E_f & M\Delta_i E_f \\ 0 & 0 \end{bmatrix}.$$

Hence the extended state equation holds:

$$\begin{bmatrix} \tilde{\mathbf{x}}_{i+1} \\ \hat{\mathbf{x}}_{i+1} \end{bmatrix} = (\mathcal{F} + \delta\mathcal{F}_i) \begin{bmatrix} \tilde{\mathbf{x}}_i \\ \hat{\mathbf{x}}_i \end{bmatrix} + \mathcal{G} \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix}. \tag{5.2}$$

Introduce a similarity transformation:

$$\mathcal{T} \triangleq \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}, \qquad \mathcal{T}^{-1} = \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}.$$

We have,

$$\mathcal{T}(\mathcal{F} + \delta\mathcal{F}_i)\mathcal{T}^{-1} = \begin{bmatrix} F & 0 \\ F_p P H^\top \hat{R}^{-1} H & F_p \end{bmatrix} + \begin{bmatrix} M\Delta_i E_f & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence the first part (i.e., the nominal matrix, denote as $\tilde{\mathcal{F}}$) is stable since F and F_p are stable.

Furthermore, the following equality

$$E_f(zI - F)^{-1}M = [E_f \;\; 0](zI - \tilde{\mathcal{F}})^{-1} \begin{bmatrix} M \\ 0 \end{bmatrix},$$

shows that the extended system has a same $\mathcal{H}_\infty$-norm as the original system. Hence the extended system is quadratically stable. Thus, there exists a positive definite matrix $\mathcal{V}$ such that

$$\mathcal{V} - (\mathcal{F} + \delta\mathcal{F}_i)\mathcal{V}(\mathcal{F} + \delta\mathcal{F}_i)^\top > 0.$$

By scaling $\mathcal{V}$ large enough, we can find a positive $\mathcal{P}$ such that

$$\mathcal{P} \geq (\mathcal{F} + \delta\mathcal{F}_i)\mathcal{P}(\mathcal{F} + \delta\mathcal{F}_i)^\top + \mathcal{G} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \mathcal{G}^\top. \tag{5.3}$$

Let

$$\mathcal{M}_i \triangleq \mathbb{E}\left\{ \begin{bmatrix} \tilde{\mathbf{x}}_i \\ \hat{\mathbf{x}}_i \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_i \\ \hat{\mathbf{x}}_i \end{bmatrix}^\top \right\},$$

then the following recursion formula holds

$$\mathcal{M}_{i+1} = (\mathcal{F} + \delta\mathcal{F}_i)\mathcal{M}_i(\mathcal{F} + \delta\mathcal{F}_i)^\top + \mathcal{G} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \mathcal{G}^\top. \tag{5.4}$$

Subtracting Equation (5.4) from Equation (5.3) we get

$$\mathcal{P} - \mathcal{M}_{i+1} = (\mathcal{F} + \delta\mathcal{F}_i)(\mathcal{P} - \mathcal{M}_i)(\mathcal{F} + \delta\mathcal{F}_i)^\top + \mathcal{Q}_i,$$

for some $\mathcal{Q}_i \geq 0$. The quadratic stability of $\mathcal{F} + \delta\mathcal{F}_i$ implies that $\mathcal{P} - \mathcal{M}_\infty \geq 0$. □

6 Simulation Study

In this section, we investigate the empirical performance of the proposed filter in three parameter setups that differ in the relative magnitude of the uncertainty. The following numerical example is frequently used in robust filtering design (e.g.,[6],[12]):

$$\mathbf{x}_{i+1} = \begin{bmatrix} 0.9802 & 0.0196 + 0.099\Delta_i \\ 0 & 0.9802 \end{bmatrix} \mathbf{x}_i + \mathbf{u}_i,$$

$$\mathbf{y}_i = [1 \;\; -1]\mathbf{x}_i + \mathbf{v}_i,$$

$$\text{where} \quad Q = \begin{bmatrix} 1.9608 & 0.0195 \\ 0.0195 & 1.9608 \end{bmatrix} \quad R = 1, \quad \mathbf{x}_0 \sim N(\mathbf{0}, I).$$

We note that the uncertainty only affects the $F_{1,2}$. and the magnitude of the nominal parameter and the uncertainty are of the same order. The tradeoff parameter α is set to 0.8. The error variance is averaged from 500 trajectories.

In Figure 1(a), the uncertainty Δ is generated according to a uniform distribution in $[-1, \ 1]$, and is fixed for the whole trajectory. In Figure 1(b), the uncertainty is re-generated in each step. In both cases, the proposed filter exhibits a similar steady-state performance to the robust filter, and a faster transient response (i.e., smaller error in the transient stages). We also observe that, for the non-stationary case, the robust filter preforms worse probably due to the fact that time varying uncertainties cancel out.

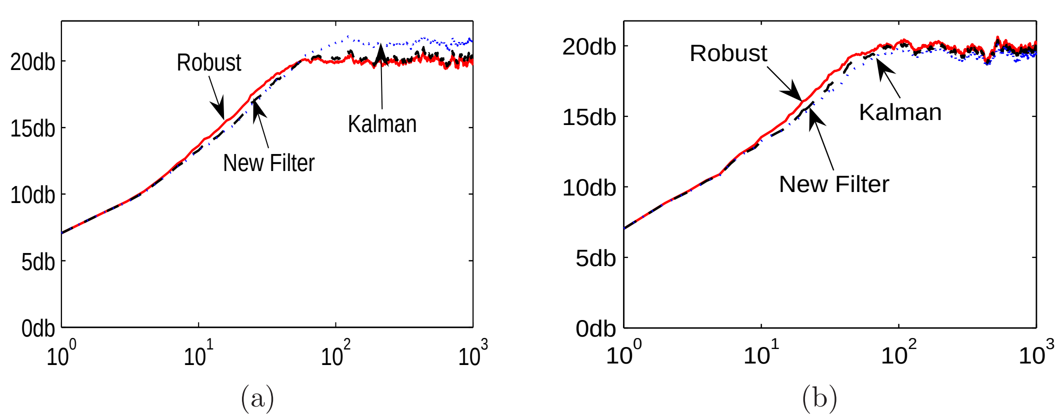
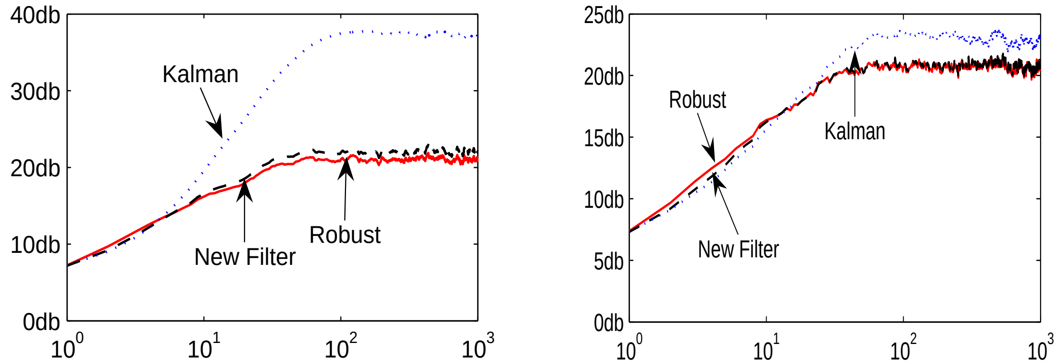
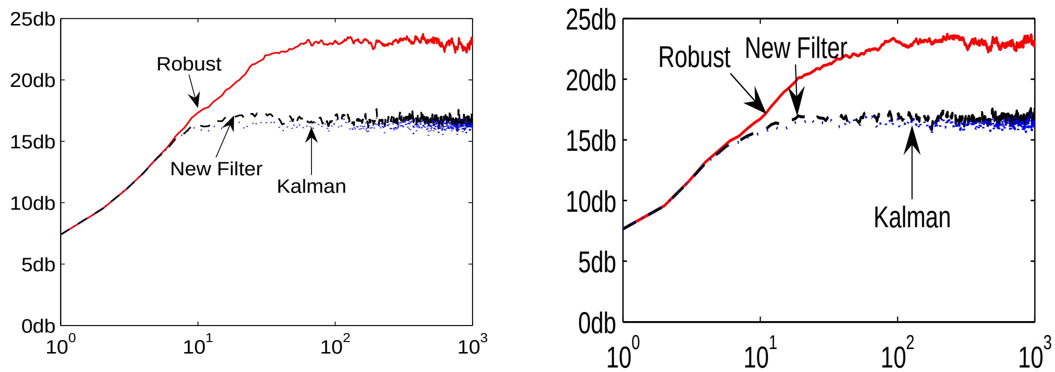


Figure 1: Error variance curves: (a) fixed uncertainty; (b) time-varying uncertainty.



(a) (b)

Figure 2: Error variance curves for large uncertainty: (a) fixed uncertainty; (b) time-varying uncertainty.



(a) (b)

Figure 3: Error variance curves for large nominal value: (a) fixed uncertainty; (b) time-varying uncertainty.

In Figure 2, we depict the case with large uncertainty by setting $F_{1,2} = 0.0196 + 0.99\Delta_i$. In such situation, the performance of the Kalman filter degrades significantly. In contrast, the steady-state error of the proposed filter is only $1dB$ worse than the robust filter in the fixed uncertainty case, and is comparable to the robust filter in the time-varying case. This shows that the proposed filter achieves a comparable robustness as the robust filter.

In Figure 3, we investigate the small uncertainty case by enlarging nominal parameters, i.e., $F_{1,2} = 0.3912 + 0.099\Delta_i$. The robust filter achieves a steady-state error variance around $23dB$, while both the Kalman filter and the proposed filter achieve a steady-state error around $16dB$. This shows that the robust filter could be overly conservative when the uncertainty is comparatively small, whereas the proposed filter does not suffer from such conservativeness.

We further simulate the steady-state error-variance for different α under different uncertainty ratio. Here, $\alpha = 0$ and $\alpha = 1$ are the robust filter and the Kalman filter, respectively; $\gamma = 1$ is the original example. We increase the uncertainty when $\gamma > 1$, and increase the nominal parameter when $\gamma < 1$. Figure 4 shows that when γ is small, (i.e., uncertainty is rel-

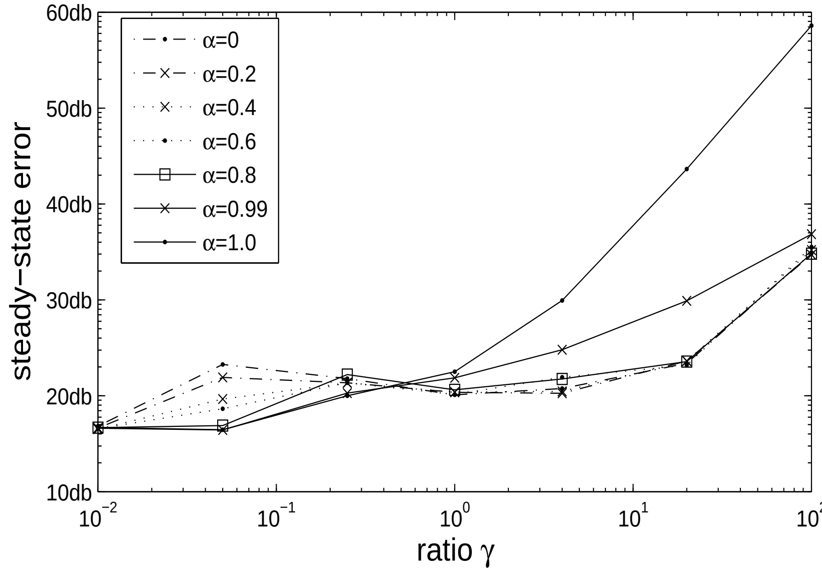


Figure 4: Effect of α on steady-state error.

atively small), larger α achieves better performance. That is, for small uncertainty, focusing on robustness can degrade the performance. On the other hand, for large uncertainty, the steady-state error for the Kalman filter is large. In contrast, even for $\alpha = 0.99$ which means the robust part has a small effect, the proposed filter achieves a much better performance. The overall most-balanced filter in this example is achieved by taking $\alpha = 0.8$, which is also our suggestion for the tradeoff parameter. The exact value of α is not sensitive, for example, choosing $\alpha = 0.6$ instead works well too.

To summarize, the simulation study shows that both the Kalman filter and the robust filter are sensitive to the relative magnitude of the uncertainty. In contrast, in all three cases, the proposed filter exhibits a performance comparable to the better one, and therefore is suitable for a wider range of problems.

7 Concluding Remarks

In this paper, we presented a new algorithm for state estimation of a linear system with uncertainty in the parameters. This filter iteratively finds a smoothed estimation that is Pareto efficient between the nominal performance and the worst performance. The resulting recursive form has a computational cost comparable to the standard Kalman filter, hence can be easily implemented on-line. Under certain technical conditions, the proposed filter converges to a stable steady-state estimator and achieves bounded error-variance. Simulation studies show that the proposed filter overcomes both the sensitivity of the Kalman filter and the overly conservativeness of the robust filter, and hence achieves satisfactory performance under a wider range of parameters.

The main motivation of the proposed approach is obtaining more flexibility in filter design while retaining the computational efficiency. As the simulation study showed, the performance of both the Kalman filter and the robust filter depend on the parameter settings. That is, each of the filters can perform rather poorly under unsuitable parameters. Whether a problem setting is suitable for these filters may not be known beforehand, except a general guideline that small uncertainty favors the standard Kalman filter and large uncertainty favors the robust filter. Moreover, the problem parameters can be time varying. The proposed filter therefore facilitates flexibility since the quality of its performance does not vary dramatically if the magnitude of the uncertainty is not specified perfectly.

Appendix

A Derivation of the Prediction Form:

We derive the prediction form based on solving Problem (2.2). By Theorem 1 and Equation (3.2), we have

$$\begin{aligned} &col(\hat{\mathbf{x}}_{i|i+1} - \hat{\mathbf{x}}_{i|i}, \hat{\mathbf{u}}_{i|i+1}) = \mathbf{z}^o(\lambda^o) \\ =&\Big(T + A^\top \overline{W}(\lambda^o) A + (1-\alpha)\lambda^o E_a^\top E_a\Big)^{-1} \Big(A^\top \overline{W}(\lambda^o)\mathbf{b} + (1-\alpha)\lambda^o E_a^\top \mathbf{t}\Big), \end{aligned} \tag{A.1}$$

where λ^o is the minimizer of function $G(\lambda)$ over $\big[\|D^\top W D\|,\, +\infty\big)$. Since we are using a line search to find out λ^o, we exclude the boundary point $\|D^\top W D\|$. Hence, $\lambda^o I - D^\top W D$ is invertible. Denote

$$\begin{aligned} \overline{R}_{i+1} \triangleq& \hat{W}(\lambda^o)^{-1} \\ =&\big\{W + WD(\lambda^o I - D^\top W D)^{-1} D^\top W\big\}^{-1} \\ =&W^{-1} - (\lambda^o)^{-1} D D^\top = R_{i+1} - (\lambda^o)^{-1} H_{i+1} M_i M_i^\top H_{i+1}^\top. \end{aligned} \tag{A.2}$$

The second equality holds due to the matrix inversion lemma, and the last equality holds by substituting the definition of D and W. Next, define

$$\hat{R}_{i+1} \triangleq \overline{W}(\lambda^o)^{-1} = \big[\alpha W + (1-\alpha)\hat{W}(\lambda^o)\big]^{-1} = \big[\alpha R_{i+1}^{-1} + (1-\alpha)\overline{R}_{i+1}^{-1}\big]^{-1}. \tag{A.3}$$

Notice this definition makes sense since $\hat{W}^(\lambda)$ is positive for $\lambda > \|D^\top W D\|$ and W is also positive. With the following two definition

$$\hat{\lambda}_i \triangleq (1-\alpha)\lambda^o; \quad \hat{T} \triangleq \begin{pmatrix} P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^\top E_{f,i} & \hat{\lambda}_i E_{f,i}^\top E_{g,i} \\ \hat{\lambda}_i E_{g,i}^\top E_{f,i} & Q_i^{-1} + \hat{\lambda}_i E_{g,i}^\top E_{g,i} \end{pmatrix};$$

we rewrite the first term of Equation (A.1):

$$\begin{aligned} &T + A^\top \overline{W}(\lambda^o) A + (1-\alpha)\lambda^o E_a^\top E_a \\ =&\begin{pmatrix} P_{i|i}^{-1} & 0 \\ 0 & Q_i^{-1} \end{pmatrix} + \hat{\lambda}_i [E_{f,i},\, E_{g,i}]^\top [E_{f,i},\, E_{g,i}] + A^\top \overline{W}(\lambda^o) A \\ =&\hat{T} + A^\top \overline{W}(\lambda^o) A = \hat{T} + A^\top \hat{R}_{i+1}^{-1} A. \end{aligned} \tag{A.4}$$

Notice $(1,1)$ block of $\hat{T}$ is strictly positive, by block matrix inversion we have

$$\hat{T}^{-1} = \begin{pmatrix} \hat{P}_{i|i} + \hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i\hat{P}_{i|i} & -\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i \\ -\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i\hat{P}_{i|i} & \hat{Q}_i \end{pmatrix}, \tag{A.5}$$

where $\hat{P}_{i|i} \triangleq \left(P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^\top E_{f,i}\right)^{-1}$ is the inversion of the $(1,1)$ block of the matrix $\hat{T}$ and $\hat{Q}_i \triangleq \left(Q_i^{-1} + \hat{\lambda}_i E_{g,i}^\top E_{g,i} - \hat{\lambda}_i E_{g,i}^\top E_{f,i} \hat{P}_{i|i} E_{f,i}^\top E_{g,i} \hat{\lambda}_i\right)^{-1}$ is the inversion of the Schur complement.

We next simplify $(\hat{T} + A^\top \overline{W}(\lambda^o) A)^{-1}$ by first prove a useful equation:

$$
\begin{aligned}
&[F_i,\, G_i]\hat{T}^{-1}[F_i,\, G_i]^\top \\
=&[F_i,\, G_i]\begin{pmatrix} \hat{P}_{i|i} + \hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i \hat{P}_{i|i} & -\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i \\ -\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i \hat{P}_{i|i} & \hat{Q}_i \end{pmatrix}[F_i,\, G_i]^\top \\
=&F_i(\hat{P}_{i|i} + \hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i \hat{P}_{i|i})F_i^\top - F_i(\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i)G_i^\top \\
&\quad - G_i(\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i \hat{P}_{i|i})F_i^\top + G_i(\hat{Q}_i)G_i^\top \\
=&F_i\hat{P}_{i|i}F_i^\top + F_i\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i \hat{P}_{i|i}F_i^\top - G_i\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i \hat{P}_{i|i}F_i^\top \\
&\quad - F_i\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i G_i^\top + G_i\hat{Q}_i G_i^\top \\
=&F_i\hat{P}_{i|i}F_i^\top - (G_i - F_i\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i})\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i \hat{P}_{i|i}F_i^\top + (G_i - F_i\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i})\hat{Q}_i G_i^\top H_{i+1}^\top \\
=&H_{i+1}F_i\hat{P}_{i|i}F_i^\top + (G_i - \hat{\lambda}_i F_i\hat{P}_{i|i}E_{f,i}E_{g,i})\hat{Q}_i(G_i - \hat{\lambda}_i F_i\hat{P}_{i|i}E_{f,i}E_{g,i})^\top \\
=&F_i\hat{P}_{i|i}F_i^\top + \hat{G}_i\hat{Q}_i\hat{G}_i^\top = P_{i+1},
\end{aligned}
\tag{A.6}
$$

where

$$
\hat{G}_i \triangleq G_i - \hat{\lambda}_i F_i \hat{P}_{i|i} E_{f,i} E_{g,i}; \quad P_{i+1} \triangleq F_i \hat{P}_{i|i} F_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top. \tag{A.7}
$$

Hence we can simplify $A\hat{T}^{-1}A^\top$ as

$$
\begin{aligned}
A\hat{T}^{-1}A^\top &= H_{i+1}[F_i,\, G_i]\hat{T}^{-1}[F_i,\, G_i]^\top H_{i+1}^\top \\
=&H_{i+1}(F_i\hat{P}_{i|i}F_i^\top + \hat{G}_i\hat{Q}_i\hat{G}_i^\top)H_{i+1}^\top = H_{i+1}P_{i+1}H_{i+1}^\top.
\end{aligned}
\tag{A.8}
$$

Define

$$
R_{e,i+1} \triangleq \hat{R}_{i+1} + H_{i+1}P_{i+1}H_{i+1}^\top = \hat{R}_{i+1} + A\hat{T}^{-1}A^\top. \tag{A.9}
$$

Hence

$$
\begin{aligned}
&\left(T + A^\top \overline{W}(\lambda^o)A + (1-\alpha)\lambda^o E_a^\top E_a\right)^{-1} \\
=&(\hat{T} + A^\top \overline{W}(\lambda^o)A)^{-1} = (\hat{T} + A^\top \hat{R}_{i+1}^{-1}A)^{-1} \\
=&\hat{T}^{-1} - \hat{T}^{-1}A^\top(\hat{R}_{i+1} + A\hat{T}^{-1}A^\top)^{-1}A\hat{T}^{-1} \\
=&\hat{T}^{-1} - \hat{T}^{-1}A^\top R_{e,i+1}^{-1}A\hat{T}^{-1} \\
=&\hat{T}^{-1} - \hat{T}^{-1}[F_i,\, G_i]^\top H_{i+1}^\top R_{e,i+1}^{-1}H_{i+1}[F_i,\, G_i]\hat{T}^{-1} \\
=&\hat{T}^{-1} - \hat{T}^{-1}\begin{pmatrix} F_i^\top H_{i+1}^\top R_{e,i+1}^{-1}H_{i+1}F_i & F_i^\top H_{i+1}^\top R_{e,i+1}^{-1}H_{i+1}G_i \\ G_i^\top H_{i+1}^\top R_{e,i+1}^{-1}H_{i+1}F_i & G_i^\top H_{i+1}^\top R_{e,i+1}^{-1}H_{i+1}G_i \end{pmatrix}\hat{T}^{-1}.
\end{aligned}
\tag{A.10}
$$

The equations hold from (A.4), (A.3), (A.9), matrix inversion lemma and definition of A respectively.

Now consider the second term of Equation (A.1), by definition of $\hat{\lambda}_i$, $\hat{R}_{i+1}$ and substituting A, $\mathbf{b}$, E_a and $\mathbf{t}$, we have:

$$\begin{aligned}&A^\top\overline{W}(\lambda^o)\mathbf{b}+(1-\alpha)\lambda^o E_a^\top\mathbf{t}=A^\top\hat{R}_{i+1}^{-1}\mathbf{b}+\hat{\lambda}_i E_a^\top\mathbf{t}\\ =&[F_i,\,G_i]^\top H_{i+1}^\top\hat{R}^{-1}(\mathbf{y}_{i+1}-H_{i+1}F_i\hat{\mathbf{x}}_{i|i})+\hat{\lambda}_i[E_{f,i},\,E_{g,i}]^\top(-E_{f,i}\hat{\mathbf{x}}_{i|i})\\ =&\begin{pmatrix}F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\\ G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\end{pmatrix}\mathbf{y}_{i+1}+\begin{pmatrix}-F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i-\hat{\lambda}_i E_{f,i}^\top E_{f,i}\\ -G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i-\hat{\lambda}_i E_{g,i}^\top E_{f,i}\end{pmatrix}\hat{\mathbf{x}}_{i|i}.\end{aligned}\tag{A.11}$$

Substitute Equation (A.4), Equation (A.16) into Equation (A.1) yields

$$\begin{aligned}&\begin{pmatrix}\hat{\mathbf{x}}_{i|i+1}-\hat{\mathbf{x}}_{i|i}\\ \hat{\mathbf{u}}_{i|i+1}\end{pmatrix}\\ =&(\hat{T}+A^\top\hat{R}_{i+1}^{-1}A)^{-1}\left\{\begin{pmatrix}F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\\ G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\end{pmatrix}\mathbf{y}_{i+1}+\begin{pmatrix}-F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i-\hat{\lambda}_i E_{f,i}^\top E_{f,i}\\ -G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i-\hat{\lambda}_i E_{g,i}^\top E_{f,i}\end{pmatrix}\hat{\mathbf{x}}_{i|i}\right\},\\ &\quad\Downarrow\\ &\begin{pmatrix}\hat{\mathbf{x}}_{i|i+1}\\ \hat{\mathbf{u}}_{i|i+1}\end{pmatrix}\\ =&(\hat{T}+A^\top\hat{R}_{i+1}^{-1}A)^{-1}\left\{\begin{pmatrix}F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\\ G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\end{pmatrix}\mathbf{y}_{i+1}\right.\\ &\left.+\left[\begin{pmatrix}-F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i-\hat{\lambda}_i E_{f,i}^\top E_{f,i}\\ -G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i-\hat{\lambda}_i E_{g,i}^\top E_{f,i}\end{pmatrix}+(\hat{T}+A^\top\hat{R}_{i+1}^{-1}A)\begin{pmatrix}1\\0\end{pmatrix}\right]\hat{\mathbf{x}}_{i|i}\right\},\end{aligned}\tag{A.12}$$

notice

$$\begin{aligned}&(\hat{T}+A^\top\hat{R}_{i+1}^{-1}A)\begin{pmatrix}1\\0\end{pmatrix}\\ =&\left\{\begin{pmatrix}P_{i|i}^{-1}+\hat{\lambda}_i E_{f,i}^\top E_{f,i} & \hat{\lambda}_i E_{f,i}^\top E_{g,i}\\ \hat{\lambda}_i E_{g,i}^\top E_{f,i} & Q_i^{-1}+\hat{\lambda}_i E_{g,i}^\top E_{g,i}\end{pmatrix}+\begin{bmatrix}F_i^\top\\ G_i^\top\end{bmatrix}H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}[F_i,\,G_i]\right\}\begin{pmatrix}1\\0\end{pmatrix}\\ =&\begin{pmatrix}P_{i|i}^{-1}+\hat{\lambda}_i E_{f,i}^\top E_{f,i}+F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i\\ \hat{\lambda}_i E_{g,i}^\top E_{f,i}+G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}H_{i+1}F_i\end{pmatrix}.\end{aligned}$$

Substitute it back into Equation (A.12) we have

$$\begin{pmatrix}\hat{\mathbf{x}}_{i|i+1}\\ \hat{\mathbf{u}}_{i|i+1}\end{pmatrix}=(\hat{T}+A^\top\hat{R}_{i+1}^{-1}A)^{-1}\left\{\begin{pmatrix}F_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\\ G_i^\top H_{i+1}^\top\hat{R}_{i+1}^{-1}\end{pmatrix}\mathbf{y}_{i+1}+\begin{pmatrix}P_{i|i}^{-1}\\ 0\end{pmatrix}\hat{\mathbf{x}}_{i|i}\right\}.\tag{A.13}$$

Substitute this into Problem (2.2) we have

$$
\begin{aligned}
&\hat{\mathbf{x}}_{i+1|i+1} = F_i\hat{\mathbf{x}}_{i|i+1} + G_i\hat{\mathbf{u}}_{i|i+1} \\
=&[F_i,\, G_i]\begin{pmatrix} \hat{\mathbf{x}}_{i|i+1} \\ \hat{\mathbf{u}}_{i|i+1} \end{pmatrix} = [F_i,\, G_i](\hat{T} + A^\top \hat{R}_{i+1}^{-1}A)^{-1}\left\{\begin{pmatrix} F_i^\top H_{i+1}^\top \hat{R}_{i+1}^{-1} \\ G_i^\top H_{i+1}^\top \hat{R}_{i+1}^{-1} \end{pmatrix}\mathbf{y}_{i+1} + \begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix}\hat{\mathbf{x}}_{i|i}\right\} \\
=&[F_i,\, G_i](\hat{T} + A^\top \hat{R}_{i+1}^{-1}A)^{-1}\begin{pmatrix} F_i^\top H_{i+1}^\top \hat{R}_{i+1}^{-1} \\ G_i^\top H_{i+1}^\top \hat{R}_{i+1}^{-1} \end{pmatrix}\mathbf{y}_{i+1} + [F_i,\, G_i](\hat{T} + A^\top \hat{R}_{i+1}^{-1}A)^{-1}\begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix}\hat{\mathbf{x}}_{i|i}
\end{aligned}
\tag{A.14}
$$

We compute the two term separately, the coefficient of $\mathbf{y}_{i+1}$ can be written as

$$
\begin{aligned}
&[F_i,\, G_i](\hat{T} + A^\top \hat{R}_{i+1}^{-1}A)^{-1}\begin{pmatrix} F_i^\top H_{i+1}^\top \hat{R}_{i+1}^{-1} \\ G_i^\top H_{i+1}^\top \hat{R}_{i+1}^{-1} \end{pmatrix} \\
=&[F_i,\, G_i]\left[\hat{T}^{-1} - \hat{T}^{-1}[F_i,\, G_i]^\top H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}[F_i,\, G_i]\hat{T}^{-1}\right]\begin{pmatrix} F_i^\top \\ G_i^\top \end{pmatrix} H_{i+1}^\top \hat{R}_{i+1}^{-1} \\
=&\left\{[F_i,\, G_i]\hat{T}^{-1}[F_i,\, G_i]^\top - [F_i,\, G_i]\hat{T}^{-1}[F_i,\, G_i]^\top H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}[F_i,\, G_i]\hat{T}^{-1}[F_i,\, G_i]^\top\right\} H_{i+1}^\top \hat{R}_{i+1}^{-1} \\
=&(P_{i+1} - P_{i+1}H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}P_{i+1})H_{i+1}^\top \hat{R}_{i+1}^{-1}.
\end{aligned}
\tag{A.15}
$$

The first equality holds from Equation (A.10), and the last equality holds from Equation (A.6).

The coefficient of $\hat{\mathbf{x}}_{i|i}$ can be written as:

$$
\begin{aligned}
&[F_i,\, G_i](\hat{T} + A^\top \hat{R}_{i+1}^{-1}A)^{-1}\begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix} \\
=&[F_i,\, G_i]\left[\hat{T}^{-1} - \hat{T}^{-1}[F_i,\, G_i]^\top H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}[F_i,\, G_i]\hat{T}^{-1}\right]\begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix} \\
=&\left[I - [F_i,\, G_i]\hat{T}^{-1}[F_i,\, G_i]^\top H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}\right][F_i,\, G_i]\hat{T}^{-1}\begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix} \\
=&\left[I - P_{i+1}H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}\right][F_i,\, G_i]\hat{T}_i^{-1}\begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix}.
\end{aligned}
\tag{A.16}
$$

Notice by Equation (A.5), we have

$$
\begin{aligned}
&[F_i,\, G_i]\hat{T}^{-1}\begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix} \\
=&[F_i,\, G_i]\begin{pmatrix} \hat{P}_{i|i} + \hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i\hat{P}_{i|i} & -\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i \\ -\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i\hat{P}_{i|i} & \hat{Q}_i \end{pmatrix}\begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix} \\
=&\big(F_i\hat{P}_{i|i} + F_i\hat{P}_{i|i}\hat{\lambda}_i E_{f,i}^\top E_{g,i}\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i\hat{P}_{i|i} - G_i\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i\hat{P}_{i|i}\big)P_{i|i}^{-1} \\
=&\big(F_i - \hat{G}_i\hat{Q}_i E_{g,i}^\top E_{f,i}\hat{\lambda}_i\big)\hat{P}_{i|i}P_{i|i}^{-1} = \tilde{F}_i\hat{P}_{i|i}P_{i|i}^{-1},
\end{aligned}
\tag{A.17}
$$

where

$$\tilde{F}_i \triangleq F_i - \hat{\lambda}_i \hat{G}_i \hat{Q}_i E_{g,i}^\top E_{f,i}, \tag{A.18}$$

and the second last equality holds from Equation (A.7).

Recall definition of $\hat{P}_{i|i}$, we have

$$\begin{aligned}
&\hat{P}_{i|i} = \left(P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^\top E_{f,i}\right)^{-1} \\
\Rightarrow\ & P_{i|i}^{-1} = \hat{P}_{i|i}^{-1} - \hat{\lambda}_i E_{f,i}^\top E_{f,i} \\
\Rightarrow\ & \hat{P}_{i|i} P_{i|i}^{-1} = \hat{P}_{i|i}(\hat{P}_{i|i}^{-1} - \hat{\lambda}_i E_{f,i}^\top E_{f,i}) = I - \hat{\lambda}_i \hat{P}_{i|i} E_{f,i}^\top E_{f,i} \\
\Rightarrow\ & \hat{F}_i \triangleq (F_i - \hat{\lambda}_i \hat{G}_i \hat{Q}_i E_{g,i}^\top E_{f,i})(I - \hat{\lambda}_i \hat{P}_{i|i} E_{f,i}^\top E_{f,i}) = \tilde{F}_i \hat{P}_{i|i} P_{i|i}^{-1}.
\end{aligned} \tag{A.19}$$

Substitute Equation (A.17) and Equation (A.19) into Equation (A.16), we have

$$[F_i,\, G_i](\hat{T} + A^\top \hat{R}_{i+1}^{-1} A)^{-1} \begin{pmatrix} P_{i|i}^{-1} \\ 0 \end{pmatrix} = \left[I - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}\right] \hat{F}_i. \tag{A.20}$$

Now substitute Equation (A.15) and Equation (A.20) into Equation (A.14), and denote

$$\hat{\mathbf{x}}_{i+1} \triangleq \hat{F}_i \hat{\mathbf{x}}_{i|i}; \quad \mathbf{e}_{i+1} \triangleq \mathbf{y}_{i+1} - H_{i+1}\hat{\mathbf{x}}_{i+1}; \quad P_{i+1|i+1} \triangleq P_{i+1} - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1} P_{i+1}; \tag{A.21}$$

we get

$$\begin{aligned}
&\hat{\mathbf{x}}_{i+1|i+1} = (P_{i+1} - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1} P_{i+1}) H_{i+1}^\top \hat{R}_{i+1}^{-1} \mathbf{y}_{i+1} + \left[I - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}\right] \hat{F}_i \hat{\mathbf{x}}_{i|i} \\
=&(P_{i+1} - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1} P_{i+1}) H_{i+1}^\top \hat{R}_{i+1}^{-1} (\mathbf{e}_{i+1} + H_{i+1} \hat{F}_i \hat{\mathbf{x}}_{i|i}) + \left[I - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}\right] \hat{F}_i \hat{\mathbf{x}}_{i|i} \\
=&P_{i+1|i+1} H_{i+1}^\top \hat{R}_{i+1}^{-1} \mathbf{e}_{i+1} + \left[I - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1}\right] \left[I + P_{i+1} H_{i+1}^\top \hat{R}_{i+1}^{-1} H_{i+1}\right] \hat{F}_i \hat{\mathbf{x}}_{i|i} \\
=&P_{i+1|i+1} H_{i+1}^\top \hat{R}_{i+1}^{-1} \mathbf{e}_{i+1} + \left[I - P_{i+1} H_{i+1}^\top (\hat{R}_{i+1} + H_{i+1} P_{i+1} H_{i+1}^\top)^{-1} H_{i+1}\right] \left[I + P_{i+1} H_{i+1}^\top \hat{R}_{i+1}^{-1} H_{i+1}\right] \hat{F}_i \hat{\mathbf{x}}_{i|i} \\
=&P_{i+1|i+1} H_{i+1}^\top \hat{R}_{i+1}^{-1} \mathbf{e}_{i+1} + \hat{F}_i \hat{\mathbf{x}}_{i|i} = P_{i+1|i+1} H_{i+1}^\top \hat{R}_{i+1}^{-1} \mathbf{e}_{i+1} + \hat{\mathbf{x}}_{i+1}.
\end{aligned} \tag{A.22}$$

The third equality follows form Equation (A.9), and the fourth equality holds from matrix inversion lemma.

Combine all the definitions and Equation (A.22), we get the following measurement-update form.

Algorithm 2 Measurement-Update form

1. *Initialize:*

$$\begin{aligned}
P_{0|0} :=& (\Pi_0^{-1} + H_0^\top R_0^{-1} H_0)^{-1} \\
\hat{\mathbf{x}}_{0|0} :=& P_{0|0} H_0^\top R_0^{-1} \mathbf{y}_0.
\end{aligned}$$

2. *Recursion:*
   Construct and minimize $G(\lambda)$ *over* $(\|M_i^\top H_{i+1}^\top R_{i+1}^{-1} H_{i+1} M_i\|, +\infty)$. *Let the optimal value be* λ_i^o. *Compute the following values:*

$$\begin{aligned}
\hat{\lambda}_i :=& (1-\alpha)\lambda_i^o \\
\overline{R}_{i+1} :=& R_{i+1} - \lambda^{o-1} H_{i+1} M_i M_i^\top H_{i+1}^\top \\
\hat{R}_{i+1}^{-1} :=& \alpha R_{i+1}^{-1} + (1-\alpha)\overline{R}_{i+1}^{-1} \\
\hat{Q}_i^{-1} :=& Q_i^{-1} + \hat{\lambda}_i E_{g,i}^\top \big[I + \hat{\lambda}_i E_{f,i} P_{i|i} E_{f,i}^\top\big]^{-1} E_{g,i} \\
\hat{P}_{i|i} :=& (P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^\top E_{f,i})^{-1} = P_{i|i} - P_{i|i} E_{f,i}^\top (\hat{\lambda}_i^{-1} I + E_{f,i} P_{i|i} E_{f,i}^\top)^{-1} E_{f,i} P_{i|i} \\
\hat{G}_i :=& G_i - \hat{\lambda}_i F_i \hat{P}_{i|i} E_{f,i}^\top E_{g,i} \\
\hat{F}_i :=& (F_i - \hat{\lambda}_i \hat{G}_i \hat{Q}_i E_{g,i}^\top E_{f,i})(I - \hat{\lambda}_i \hat{P}_{i|i} E_{f,i}^\top E_{f,i}) \\
P_{i+1} :=& F_i \hat{P}_{i|i} F_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top \\
R_{e,i+1} :=& \hat{R}_{i+1} + H_{i+1} P_{i+1} H_{i+1}^\top \\
P_{i+1|i+1} :=& P_{i+1} - P_{i+1} H_{i+1}^\top R_{e,i+1}^{-1} H_{i+1} P_{i+1} \\
\hat{\mathbf{x}}_{i+1} :=& \hat{F}_i \hat{\mathbf{x}}_{i|i} \\
\mathbf{e}_{i+1} :=& \mathbf{y}_{i+1} - H_{i+1} \hat{\mathbf{x}}_{i+1} \\
\hat{\mathbf{x}}_{i+1|i+1} :=& \hat{\mathbf{x}}_{i+1} + P_{i+1|i+1} H_{i+1}^\top \hat{R}_{i+1}^{-1} \mathbf{e}_{i+1}.
\end{aligned}$$

To derive the prediction form from the measurement-update form, we need the following two lemma.

Lemma 2

$$P_{i+1} = F_i P_i F_i^\top - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top, \tag{A.23}$$

where

$$\overline{K}_i \triangleq F_i P_i \overline{H}_i^\top, \qquad \overline{R}_{e,i} \triangleq I + \overline{H}_i P_i \overline{H}_i^\top, \qquad \overline{H}_i \triangleq \begin{bmatrix} \hat{R}_i^{-1/2} H_i \\ \sqrt{\hat{\lambda}_i} E_{f,i} \end{bmatrix}.$$

Proof. First note

$$\begin{aligned}
P_{i|i}^{-1} =& (P_i - P_i H_i^\top R_{e,i}^{-1} H_i P_i)^{-1} = \Big(P_i - P_i H_i^\top (\hat{R}_i + H_i P_i H_i^\top)^{-1} H_i P_i\Big)^{-1} \\
=& \Big((P_i^{-1} + H_i^\top \hat{R}_i^{-1} H_i)^{-1}\Big)^{-1} = P_i^{-1} + H_i^\top \hat{R}_i^{-1} H_i.
\end{aligned} \tag{A.24}$$

Hence we have

$$\begin{aligned}
P_{i+1} =& F_i \hat{P}_{i|i} F_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top = F_i \big(P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^\top E_{f,i}\big)^{-1} F_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top \\
=& F_i \big(P_i^{-1} + H_i^\top \hat{R}_i^{-1} H_i + \hat{\lambda}_i E_{f,i}^\top E_{f,i}\big)^{-1} F_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top = F_i \big(P_i^{-1} + \overline{H}_i^\top \overline{H}_i\big)^{-1} F_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i \\
=& F_i P_i F_i^\top - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^\top + \hat{G}_i \hat{Q}_i \hat{G}_i^\top.
\end{aligned}$$

□

Lemma 3

$$P_{i|i}H_i^\top \hat{R}_i^{-1} = P_i H_i^\top R_{e,i}^{-1}.$$

Proof. From Equation (A.24) we have

$$P_{i|i}^{-1}(P_i H_i^\top R_{e,i}^{-1}) = (P_i^{-1} + H_i^\top \hat{R}_i^{-1} H_i)(P_i H_i^\top R_{e,i}^{-1}) = H_i^\top (I + \hat{R}_i^{-1} H_i P_i H_i^\top) R_{e,i}^{-1}$$
$$= H_i^\top \hat{R}_i^{-1}(\hat{R}_i + H_i P_i H_i^\top) R_{e,i}^{-1} = H_i^\top \hat{R}_i^{-1}.$$

By left multiplying $P_{i|i}$ on both sides, the lemma follows. □

Substituting these two Lemma into the measurement-update form, we get the recursion formula of the prediction form.

References

[1] A. Bryson Jr. and Y. C. Ho, *Applied Optimal Control: Optimization, Estimation and Control*. John Willey & Sons, 1975.

[2] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice-Hall, 2000.

[3] A. Gelb, Ed., *Applied Optimal Estimation*. MIT Press, 1974.

[4] H. W. Sorenson, Ed., *Kalman Filtering: Theory and Application*. IEEE Press, 1985.

[5] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice*. Prentice-Hall, 1993.

[6] A. H. Sayed, "A framework for state-space estimation with uncertain models," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 998–1013, July 2001.

[7] P. P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an h_∞-setting," *IEEE Transactions on Automatic Control*, vol. 36, pp. 151–166, 1991.

[8] T. Basar and P. Bernhard, *H_∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach.* Birkhauser, 1991.

[9] U. Shaked and Y. Theodor, "$\mathcal{H}_\infty$-optimal estimation: A tutorial," in *Proceedings: IEEE Conference on Decision and Control*, 1992.

[10] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.

[11] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control: A Unified Approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Theories*. SIAM, Philadelphia, 1999.

[12] I. R. .Petersen and A. V. Savkin, *Robust Kalman Filtering for Signals and Syustems with Large Uncertainties*. Birkauser, 1999.

[13] D. P. Bertsekas and I. B. Rhodes, "Recursive state estimation for a set-membership description of uncertainties," *IEEE Transactions on Automatic Control*, vol. 16, pp. 117–128, 1971.

[14] L. El Ghaoui and G. Calafiore, "Robust filtering for discrete-time systems with bounded noise and parametric uncertaitny," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 1084–1089, 2001.

[15] I. R. Petersen and D. C. McFarlane, "Robust state estimation for uncertain systems," in *Proceedings: IEEE Conference on Decision and Control*, 1991.

[16] L. Xie, Y. C. Soh, and C. de Souza, "Robust kalman filtering for uncertain discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 39, pp. 1310–1314, 1994.

[17] A. H. Sayed, V. H. Nascimento, and F. A. Cipparrone, "A regularized robust design criterion for uncertain data," *SIAM Journal on Matrix Analysis and Its Applications*, vol. 23, no. 4, pp. 1120–1142, 2002.