arXiv:2304.01727v1 [eess.SP] 4 Apr 2023

随机延迟测量的风险敏感滤波 ⋆

Ranjeet Kumar Tiwari* [a], Shovan Bhaumik [a]

[a] *Department of Electrical Engg., Indian Institute of Technology Patna, Pat*^{na}*, India*

摘要

传统的贝叶斯估计需要系统的精确随机模型。然而，在许多实际情况下，这一要求并不总能满足，因为系统可能并未被完全了解，或可能与假设模型存在差异。对于此类系统，我们考虑这样一种情形：测量值通过公共通信网络传输至远程位置，因此接收测量值时会产生时延。本文所考虑的时延是随机的，且在任一给定时刻最多为一步。针对这一情形，本文通过最小化风险敏感误差准则，为线性高斯系统设计了一种鲁棒估计器；该准则定义为累积指数二次误差的期望。本文推导了风险敏感卡尔曼滤波器（RSKF）的稳定性判据，并据此研究所设计滤波器的稳定性。此外，假设与时延相关的时延概率未知，并通过最大化似然函数对其进行估计。仿真结果表明，所提滤波器在标称条件下表现良好；在存在模型不确定性和随机时延的情况下，其性能优于针对随机延迟测量的卡尔曼滤波器和RSKF。

Key words: 风险敏感滤波；随机延迟测量；最优滤波。

1 引言

一种广为人知的最优解是卡尔曼滤波器（KF），可用于估计线性高斯系统的状态 [1]。然而，如果测量值以随机时延到达，则需要对其进行修改 [19]。因此，文献中已有大量研究工作探讨如何估计具有随机时延测量值的线性高斯系统的状态 [24,16,15]。 [19] 讨论了测量值存在一步随机时延时，线性高斯系统的无偏最小方差滤波解。对于具有有界测量延迟和数据包丢失的系统，[22]提出了一种采用状态增广方法的最小方差滤波器。[11] 的作者采用线性时间编码技术，为存在数据包丢弃事件的网络化系统设计了一种最优线性估计器。 [28] 和 [23] 的研究处理了带时间戳的随机时延测量值。 [28] 的作者采用测量重组方法设计最优估计器，而 [23] 采用了状态增广方法。值得注意的是，上述研究均未考虑过程动态中的模型不确定性。

在实际应用中，系统往往并不遵循假设的模型，其模型参数值也会偏离标称值。针对这类系统，文献中已有基于不同代价函数的多种鲁棒估计器 [9,10,27,7,3,20]。文献 [13] 研究了具有指数性能准则的线性高斯系统的状态估计，以应对确定性模型不确定性。文献 [21] 提出了采用相同性能准则的动态规划方法，以实现所需的鲁棒性。文献 [8] 和 [4] 的作者考虑了用于处理系统中此类不确定性的风险敏感滤波解。然而，这些研究在设计状态估计器时，均未考虑测量中的随机延迟与系统模型不确定性同时存在的情况。

本研究所考虑的模型不确定性是被控对象动态特性的一部分，这意味着一个或多个过程参数的取值偏离了原定值。

⋆ 本文未在任何 IFAC 会议上发表。通讯作者：R. K. Tiwari。*Email addresses:* ranjeet.pee16@iitp.ac.in（Ranjeet Kumar Tiwari*），shovan.bhaumik@iitp.ac.in（Shovan Bhaumik）。

这些值是标称值，且并未被准确获知。这些数值上的不确定性是确定的、任意的且未知的。这不应与过程噪声混淆，后者是具有已知分布的随机序列。在给定时刻，测量值会随机延迟，且最多延迟一个时间步，并借助伯努利随机变量对其建模。本研究将过去和当前估计误差平方的指数作为代价函数，并提出一个用于求得通用解的框架。随后，针对具有上述描述的线性高斯系统，最小化该代价函数，并得到一个闭式递归解。

此外，假设随机时延的时延概率未知，并提出一种基于接收测量值联合概率密度的估计方法。本文还利用一致完全可控性与可观测性条件，证明了 RSKF 的稳定性。随后进一步利用这一结果分析所提方法的稳定性，并以猜想的形式给出。所开发的估计器在测量中不存在随机时延时收敛到 RSKF；当随机时延和模型不确定性均不存在时，则变为卡尔曼滤波器。

所提滤波器应用于两个线性估计问题，并将其性能与 RSKF [4] 以及针对随机时延测量的卡尔曼滤波器（KF-RD） [19]进行比较。仿真结果表明，在标称条件下，所提滤波器的性能可以接受（与卡尔曼滤波器相当）；当系统偏离标称属性时，其表现优于 KF-RD 和 RSKF。

本文相较于现有研究的具体贡献可概括如下：

(i) 针对过程动态中存在模型不确定性且测量值存在单步随机时延的情况，推导了线性高斯系统的闭式解。(ii) 通过最大化接收测量的似然，估计测量模型中的时延概率。(iii) 此外，建立了 RSKF 的稳定性准则，并利用该结果证明所提滤波器的稳定性。

本文其余部分安排如下。第 2 节介绍一种递归贝叶斯框架，该框架利用信息状态和风险敏感误差准则，针对存在模型不确定性和测量中随机延迟的系统进行处理。第 3 节运用上一节建立的框架，推导所提滤波器的递归算法。第 4 节讨论 RSKF 和所提滤波器的稳定性。随机时延的时延参数估计。测量值在第 5 节中给出。第 6 节列出了仿真结果。最后，论文以简短的结论收尾。

2 随机延迟测量的风险敏感滤波

考虑一个离散时间非线性系统，其过程方程和测量方程建模如下：

$$x_k = f_{k-1}(x_{k-1}) + w_{k-1}, \tag{1}$$

$$z_k = h_k(x_k) + v_k, \tag{2}$$

其中，状态为 $x_k \in \Re^{n_x}$，测量为 $z_k \in \Re^{n_z}$，$f_k : \Re^{n_x} \to \Re^{n_x}$，且 $h_k : \Re^{n_x} \to \Re^{n_z}$。过程噪声 w_{k-1} 和测量噪声 v_k 是相互独立且同分布（i.i.d.）的随机过程，其概率密度函数（pdf）任意但已知。我们认为实际系统与假设模型不同，并遵循以下过程动态：

$$x_k = f_{k-1}(x_{k-1}) + \Delta f_{k-1}(x_{k-1}) + w_{k-1}, \tag{3}$$

其中，$\Delta f_{k-1}(x_{k-1})$表示任意、确定且未知的过程建模误差。

假设测量值通过公共通信信道传输到远程估计中心。由于信道带宽有限，传输过程中会引入时延。我们假设时延是随机的，且最大时延为一个时间步。接收测量值 y_k 可建模为 [19]

$$y_k = (1 - \beta_k)z_k + \beta_k z_{k-1}, \tag{4}$$

其中，β_k 是服从伯努利分布的独立同分布随机序列，且 $\mathbb{E}[\beta_k] = \alpha_k$，$\mathbb{E}[\beta_i\beta_j^\top] = \mathbb{E}[\beta_i]\mathbb{E}[\beta_j]$，$\forall i \neq j$。

2.1 方法

我们的目标是根据远程接收的测量值 $y_{1:k}$，递归地求出式（1-4）所述底层系统的最优后验估计$\hat{x}^*_{k|k}$。我们考虑如下风险敏感代价准则：

$$J_k(\hat{x}_{k|k}|y_{1:k}) = \mathbb{E}_\beta \mathbb{E}_x \Big[\exp\Big(\sum_{i=1}^{k-1} \mu_{1,i} e^{*\top}_{i|i} e^*_{i|i} + \mu_{2,k} e^\top_{k|k} e_{k|k} \Big) \Big], \tag{5}$$

其中，后验估计误差为 $e^*_{i|i} = (x_i - \hat{x}^*_{i|i})$。$\hat{x}_{i|i}$ 表示时间步 i 的后验估计状态。

$\mu_{1,\ i} \geq 0$ 和 $\mu_{2,\ i} > 0$ 是两个随时间变化的风险参数，分别用于缩放过去误差和当前误差。$\mathbb{E}_x$ 和 $\mathbb{E}_\beta$ 分别表示关于 x_k 的后验密度的期望以及 β_k 的统计量。下文为简便起见，将 $\mathbb{E}_\beta\mathbb{E}_x[.]$ 写作 $\mathbb{E}[.]$。我们寻求每个时间步使式（5）中所表达的代价函数最小化的状态估计，即

$$\hat{x}^*_{k|k} = \arg\ \min_{\hat{x}_{k|k}} J_k(\hat{x}_{k|k}|y_{1:k}). \tag{6}$$

注释 1. 请注意，状态的后验概率密度为$p(x_k|y_{1:k}) \propto p(y_k|x_k, x_{k-1})p(x_k|x_{k-1})$，其中概率密度 $p(y_k|x_k, x_{k-1})$和 $p(x_k|x_{k-1})$是利用过程方程和测量方程以及 v_k 和 w_{k-1} 的概率密度函数构造的。

2.2 通用框架

在本小节中，我们利用风险敏感误差准则的两步贝叶斯框架，推导出求解（6）的通用框架。给定延迟测量，$y_{1:k}$ 的联合后验概率密度可写为：

$$\begin{aligned} p(x_{0:k}|y_{1:k}) &= \frac{p(x_{0:k}, y_{1:k})}{p(y_{1:k})} \\ &= \frac{p(y_k|x_{0:k}, y_{1:k-1})p(x_{0:k}, y_{1:k-1})}{p(y_k|y_{1:k-1})p(y_{1:k-1})} \\ &= \frac{p(y_k|x_{0:k}, y_{1:k-1})p(x_{0:k}|y_{1:k-1})p(y_{1:k-1})}{p(y_k|y_{1:k-1})p(y_{1:k-1})} \\ &= \gamma_k p(y_k|x_{0:k}, y_{1:k-1})p(x_{0:k}|y_{1:k-1}), \end{aligned} \tag{7}$$

其中 $\gamma_k = 1/p(y_k|y_{1:k-1})$，且 $p(y_k|y_{1:k-1}) = \int p(y_k|x_{0:k}, y_{1:k-1})p(x_{0:k}|y_{1:k-1})dx_{0:k}$ 是一个归一化常数。由 (2) 和 (4) 可知，当前测量 y_k 与当前状态x_k 以及前一状态 x_{k-1}相关。因此，假设 y_k，在给定 x_k 和 x_{k-1}的条件下，与先前的测量值、 $y_{1:k-1}$以及状态 $x_{0:k-1}$相互独立，我们可以写成 $p(y_k|x_{0:k}, y_{1:k-1}) = p(y_k|x_k, x_{k-1})$。假设在 (7) 中的时间步 k时，$\hat{x}^*_{0|0}, \cdots, \hat{x}^*_{k-1|k-1}$ 已知，我们将状态的边缘密度写为

$$p(x_k|y_{1:k}) = \gamma_k p(y_k|x_k, x_{k-1})p(x_k|y_{1:k-1}). \tag{8}$$

再次利用$p(x_k|y_{1:k-1})$的查普曼–科尔莫戈罗夫积分，方程 (8) 可重写为

$$\begin{aligned} p(x_k|y_{1:k}) =& \gamma_k p(y_k|x_k, x_{k-1}) \int p(x_k|x_{k-1}) \\ & \times p(x_{k-1}|y_{1:k-1})dx_{k-1}. \end{aligned} \tag{9}$$

信息状态：假设在任意时间步都可以获得一组信息，$\mathcal{I}_k = \{y_{1:k}, e_{1|1}, \cdots, e_{k-1|k-1}\}$ [2]

k。信息状态定义为

$$\sigma_k \triangleq p(x_k|I_k) = \exp\Big(\sum_{i=0}^{k-1} \mu_{1,i} e^\top_{i|i} e_{i|i}\Big) p(x_k|y_{1:k}), \tag{10}$$

其中 $\sigma_0 = p(x_0)$。现在，将 (9) 代入 (10) 得到

$$\begin{aligned} \sigma_k &= \gamma_k p(y_k|x_k, x_{k-1}) \\ &\quad \times \int p(x_k|x_{k-1}) \exp\left(\mu_{1,k-1} e^\top_{k-1|k-1} e_{k-1|k-1}\right) \\ &\quad \times \exp\Big(\sum_{i=0}^{k-2} \mu_{1,i} e^\top_{i|i} e_{i|i}\Big) p(x_{k-1}|y_{1:k-1}) dx_{k-1}, \\ &= \gamma_k p(y_k|x_k, x_{k-1}) \int p(x_k|x_{k-1}) \\ &\quad \times \exp\left(\mu_{1,k-1} e^\top_{k-1|k-1} e_{k-1|k-1}\right) \sigma_{k-1} dx_{k-1}. \end{aligned} \tag{11}$$

我们将预测信息状态密度 $p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1})$ 定义为

$$\begin{aligned} p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1}) &= \int p(x_k|x_{k-1}) \\ &\quad \times \exp\left(\mu_{1,k-1} e^\top_{k-1|k-1} e_{k-1|k-1}\right) \sigma_{k-1} dx_{k-1}. \end{aligned} \tag{12}$$

根据 (11)，信息状态的后验密度可写为

$$\sigma_k \triangleq p(x_k|\mathcal{I}_k) = \gamma_k p(y_k|x_k, x_{k-1}) p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1}). \tag{13}$$

此外，(5) 中定义的代价函数可写为

$$\begin{aligned} J_k(\hat{x}_{k|k}|y_{1:k}) = \int \exp\left(\sum_{i=1}^{k-1} \mu_{1,i} e^\top_{i|i} e_{i|i} + \mu_{2,k} e^\top_{k|k} e_{k|k}\right) \\ \times p(x_k|y_{1:k}) dx_k, \end{aligned}$$

根据 (10)，上式可化简为

$$J_k(\hat{x}_{k|k}|y_{1:k}) = \int \exp\left(\mu_{2,k} e_{k|k} e^\top_{k|k}\right) \sigma_k dx_k. \tag{14}$$

借助 (12)-(14) 递归求解方程 (6)，可得到随机时延测量的风险敏感估计。注意，方程 (14) 是一个指数二次代价函数，它显式地只考虑当前误差。如果底层系统是线性的且噪声服从高斯分布，则可以得到闭式解。然而，对于非线性系统，后验概率密度往往会变得难以进行数值处理，因此通常采用近似解。

注释 2. 尽管 $p(x_k|\mathcal{I}_k)$ 是未归一化的概率密度，但这不会改变使 (14) 达到最小值的估计值 $\hat{x}_{k|k}$ [4]。

注释 3. 可以观察到，对于风险中性情形，$\mu_{1,\ k-1} = 0$，且$\mu_{2,\ k} > 0$，$p(x_k|\mathcal{I}_k)$ 等于$p(x_k|y_{1:k})$，并且代价函数 $J_k(\hat{x}_{k|k}|y_{1:k})$ 退化为标准指数二次函数。

3 具有随机延迟测量的线性高斯系统风险敏感滤波器

在本节中，我们将前一节提出的通用框架用于具有随机时延测量的线性高斯系统，并推导出估计值的闭式解。具有模型不确定性的离散时间线性系统表示为：

$$x_k = (A_{k-1} + \Delta A_{k-1})x_{k-1} + w_{k-1}, \tag{15}$$

$$z_k = C_k x_k + v_k, \tag{16}$$

其中，w_{k-1} 和 v_k 是均值为零、白噪声且相互独立的高斯噪声序列，其协方差分别为Q_{k-1} 和 R_k。延迟测量 y_k与式（4）中的定义相同。还假设初始状态 x_0服从高斯分布，且与 w_{k-1}、v_k 和 β_k相互独立。A_{k-1} 是将 x_{k-1} 状态转移至 x_k 的矩阵，并假设其可逆。ΔA_{k-1} 是任意的、确定性的未知建模不确定性。我们进一步假设，若 A_k^{-1} 存在，则对于所有允许的 ΔA_k取值，$(A_k + \Delta A_k)^{-1}$ 也存在。此外，假设矩阵 A_k、$A_k + \Delta A_k$ 和 C_k 对所有 $k \geq 0$ 均有界，即满足以下不等式：

$$\begin{aligned} 0 \leq &(A_k + \Delta A_k)^\top (A_k + \Delta A_k) \leq \kappa_1 \mathbf{I}, \\ &\text{and } 0 \leq C_k^\top C_k \leq \kappa_2 \mathbf{I}, \end{aligned} \tag{17}$$

其中 κ_1 和 κ_2 为实数正常数 s.

注释 4。ΔA_{k-1} 表示过程模型相对于其标称模型的偏差。注意，由于估计器不知道ΔA_{k-1}，因此估计器仅使用标称过程动态（即 A_{k-1}）来进行状态转移。

3.1 Risk sensitive estimate with delayed measurements

为了获得线性高斯系统的状态估计 $\hat{x}_{k|k}$，需实现式 (12) 和式 (13)，并使代价函数 $J_k(\hat{x}_{k|k}|y_{1:k})$ 最小化。我们假设 σ_k 是未归一化的高斯分布，前提是 $\mu_{1,\ k-1}$ 是一个足够小的非负数 [4]。这一假设是合理的，因为 $\sigma_0 = p(x_0)$ 服从高斯分布；本小节末尾将证明，若 σ_{k-1}，则 σ_k 服从高斯分布取为高斯密度。下面推导由式 (15) 和式 (4) 描述的线性系统的先验风险敏感估计。

定理1 方程（15）和（4）所描述的系统，其预测均值和误差协方差为

$$\begin{aligned} \hat{x}_{k|k-1} &= A_{k-1}\hat{x}_{k-1|k-1}, \\ \Sigma_{k|k-1} &= A_{k-1}(\Sigma_{k-1|k-1}^{-1} - 2\mu_{1,k-1}\mathbf{I})^{-1}A_{k-1}^\top + Q_{k-1}, \end{aligned} \tag{18}$$

其中 $\Sigma_{k-1|k-1}$ 是时间步 $k-1$处的后验误差协方差。

证明。为计算预测估计，我们使用方程（12）中给出的预测信息状态概率密度框架。考虑到 σ_{k-1} 是高斯概率密度，我们有

$$\begin{aligned} \sigma_{k-1} = \gamma_{k-1|k-1}(2\pi)^{-n_x/2}|\Sigma_{k-1|k-1}|^{-1/2} \exp\Big(-\frac{1}{2}\times \\ (x_{k-1} - \hat{x}_{k-1|k-1})^\top \Sigma_{k-1|k-1}^{-1}(x_{k-1} - \hat{x}_{k-1|k-1})\Big). \end{aligned} \tag{19}$$

将（19）代入（12）得

$$\begin{aligned} &p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1}) \\ &= \gamma_{k-1|k-1}(2\pi)^{-n_x/2}|\Sigma_{k-1|k-1}|^{-1/2} \int p(x_k|x_{k-1}) \\ &\times \exp\Big[-\frac{1}{2}\Big((x_{k-1} - \hat{x}_{k-1|k-1})^\top \Sigma_{k-1|k-1}^{-1} \\ &\times (x_{k-1} - \hat{x}_{k-1|k-1}) + (x_{k-1} - \hat{x}_{k-1|k-1})^\top (-2\mu_{1,k-1}\mathbf{I}) \\ &\times (x_{k-1} - \hat{x}_{k-1|k-1})\Big)\Big] dx_{k-1}. \end{aligned}$$

利用矩阵乘法的分配律处理 exp[·] 内部的各项，我们可以将上述方程改写为

$$\begin{aligned} &p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1}) \\ &= \gamma_{k-1|k-1}(2\pi)^{-n_x/2}|\Sigma_{k-1|k-1}|^{-1/2} \int p(x_k|x_{k-1}) \\ &\times \exp[-\frac{1}{2}\big((x_{k-1} - \hat{x}_{k-1|k-1})^\top (\Sigma_{k-1|k-1}^{-1} - 2\mu_{1,k-1}\mathbf{I}) \\ &\times (x_{k-1} - \hat{x}_{k-1|k-1})\big)] dx_{k-1}, \end{aligned} \tag{20}$$

其中，$\mu_{1,k-1}$ 是一个非负实数，满足$|\Sigma_{k-1|k-1}^{-1} - 2\mu_{1,\ k-1}\mathbf{I}| > 0$ 或$2\mu_{1,k-1}\Sigma_{k-1|k-1} < \mathrm{I}$ 对每个 k均成立。显然，式 (20) 中的指数部分表示一个高斯分布，而使其成为归一化分布所需的任何因子都可以调整到其中。

积分号外的常数。方程（20）可以改写为

$$p(x_k|I_{k-1}, e_{k-1|k-1}) = \gamma_{k|k-1} \int p(x_k|x_{k-1}) \times \mathcal{N}(x_{k-1}; \hat{x}_{k-1|k-1}, (\Sigma^{-1}_{k-1|k-1} - 2\mu_{1,k-1}\mathbf{I})^{-1}) dx_{k-1}. \tag{21}$$

根据式 (15) 中给定的假设过程动态，我们有 $p(x_k|x_{k-1}) = \mathcal{N}(x_k;\ A_{k-1}x_{k-1}, Q_{k-1})$ [12]。现在，将其代入式 (21)，并应用高斯乘积定理（参见 [5] 的定理 2.1），可写为

$$p(x_k|I_{k-1}, e_{k-1|k-1}) = \gamma_{k|k-1} \int \mathcal{N}(x_k; \mathbf{M}, \mathbf{S}) \times \mathcal{N}(x_{k-1}; \mathbf{M}_1, \mathbf{S}_1) dx_{k-1}, \tag{22}$$

其中

$$\begin{aligned}
\mathbf{M} &= A_{k-1}\hat{x}_{k-1|k-1}, \\
\mathbf{S} &= A_{k-1}(\Sigma^{-1}_{k-1|k-1} - 2\mu_{1,k-1}\mathbf{I})^{-1}A^\top_{k-1} + Q_{k-1}, \\
G &= (\Sigma^{-1}_{k-1|k-1} - 2\mu_{1,k-1}\mathbf{I})^{-1}A^\top_{k-1}\mathbf{S}^{-1}, \\
\mathbf{M}_1 &= \hat{x}_{k-1|k-1} + G(x_k - A_{k-1}\hat{x}_{k-1|k-1}), \\
\mathbf{S}_1 &= (\Sigma^{-1}_{k-1|k-1} - 2\mu_{1,k-1}\mathbf{I})^{-1} \\
&\quad - GA_{k-1}(\Sigma^{-1}_{k-1|k-1} - 2\mu_{1,k-1}\mathbf{I})^{-1}.
\end{aligned}$$

高斯分布 $\mathcal{N}(x_k;\ \cdot)$ 是 x_k 的函数，可以置于积分号外。利用归一化分布的性质，式 (22) 中的 $\int \mathcal{N}(x_{k-1};\ \cdot)\mathrm{d}x_{k-1} = 1$，$p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1})$ 可在忽略归一化常数的情况下表示为

$$p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1}) \sim \mathcal{N}\big(x_k;\ A_{k-1}\hat{x}_{k-1|k-1}, A_{k-1}(\Sigma^{-1}_{k-1|k-1} - 2\mu_{1,k-1}\mathbf{I})^{-1}A^\top_{k-1} + Q_{k-1}\big). \tag{23}$$

上述概率分布的均值和协方差给出了式 (18) 中的表达式。□

为了计算后验信息状态，我们需要推导条件期望 $\mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}]$、协方差 Σ^{yy}_k 以及交叉协方差 Σ^{xy}_k。为进行这些推导，下面给出一些有用的关系式：

(i) $\mathbb{E}[w_{k-1}w^\top_{k-1}|\mathcal{I}_k] = \mathbb{E}[w^\top_{k-1}w_{k-1}] = Q_{k-1}$, $\mathbb{E}[v_k v^\top_k|\mathcal{I}_k] = \mathbb{E}[v_k v^\top_k] = R_k$ and $\mathbb{E}[w_i v^\top_j|\mathcal{I}_k] = 0$. (ii) $\mathbb{E}[(x_k - \hat{x}_{k|k-1})] = 0$. (iii) $\mathbb{E}_\beta[\beta^2_k] = \mathrm{var}(\beta_k) + (\mathbb{E}_\beta[\beta_k])^2 = \alpha_k$, $\mathbb{E}_\beta[(1-\beta_k)^2]$ (iv) $\mathbb{E}_\beta[(\beta_k - \alpha_k)] = 0$, $\mathbb{E}_\beta[(\beta_k - \alpha_k)^2] = \mathbb{E}_\beta[(1-\beta_k)(1-\beta_k)] = \alpha_k(1-\alpha_k)$. (v) $\mathbb{E}[e_{k|k-1}w^\top_{k-1}|\mathcal{I}_k] = \mathbb{E}[(A_{k-1}x_{k-1} + w_{k-1} - \hat{x}_{k|k-1})w^\top_{k-1}|\mathcal{I}_k] = Q_{k-1}$.

引理 2 测量值的条件期望，y_k，可表示为

$$\mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}] = (1-\alpha_k)C_k\hat{x}_{k|k-1} + \alpha_k C_{k-1}\hat{x}_{k-1|k-1}, \tag{24}$$

而 y_k, Σ^{yy}_k 的条件协方差表示为

$$\begin{aligned}
\Sigma^{yy}_k &= (1-\alpha_k)C_k\Sigma_{k|k-1}C^\top_k + \alpha_k\theta_{k-1}\Sigma_{k|k-1}\theta^\top_{k-1} \\
&\quad + \alpha_k R_{k-1} + (1-\alpha_k)R_k - \alpha_k\theta_{k-1}Q_{k-1}\theta^\top_{k-1} + \alpha_k \\
&\quad \times (1-\alpha_k)(\theta_{k-1} - C_k)\hat{x}_{k|k-1}\hat{x}^\top_{k|k-1}(\theta_{k-1} - C_k)^\top,
\end{aligned} \tag{25}$$

其中 $\theta_{k-1} = C_{k-1}A^{-1}_{k-1}$。

证明。测量 y_k 与过去误差 $e_{1|1}, \ldots, e_{k-1|k-1}$ 相互独立。利用式 (4)，可写为

$$\begin{aligned}
&\mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}] \\
&\quad = \mathbb{E}[((1-\beta_k)(C_k x_k + v_k) \\
&\qquad + \beta_k(C_{k-1}x_{k-1} + v_{k-1}))|y_{1:k-1}] \\
&\quad = \mathbb{E}_\beta[(1-\beta_k)]\mathbb{E}_x[C_k x_k + v_k] \\
&\qquad + \mathbb{E}_\beta[\beta_k]\mathbb{E}_x[C_{k-1}x_{k-1} + v_{k-1}] \\
&\quad = (1-\alpha_k)C_k\hat{x}_{k|k-1} + \alpha_k C_{k-1}\hat{x}_{k-1|k-1}.
\end{aligned}$$

接下来，协方差 Σ^{yy}_k 可计算为

$$\begin{aligned}
\Sigma^{yy}_k &= \mathbb{E}\Big[\Big(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}]\Big) \\
&\quad \times \Big(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}]\Big)^\top\Big] \\
&= \mathbb{E}\Big[\Big((1-\beta_k)(C_k x_k + v_k) + \beta_k(C_{k-1}x_{k-1} + v_{k-1}) \\
&\quad - (1-\alpha_k)C_k\hat{x}_{k|k-1} - \alpha_k C_{k-1}\hat{x}_{k-1|k-1}\Big)\Big(*\Big)^\top\Big],
\end{aligned} \tag{26}$$

其中，$(*)$ 表示其左侧括号中给出的相同项。现在，利用状态的反向演化 $x_{k-1} = A^{-1}_{k-1}(x_k - w_{k-1})$，并重新排列方程 (26) 中的各项，可得

$$\begin{aligned}
&\Sigma^{yy}_k \\
&= \mathbb{E}\Big[\Big((1-\beta_k)C_k(x_k - \hat{x}_{k|k-1}) + (1-\beta_k)C_k\hat{x}_{k|k-1} \\
&+ \beta_k C_{k-1}A^{-1}_{k-1}(x_k - \hat{x}_{k|k-1}) + \beta_k C_{k-1}A^{-1}_{k-1}\hat{x}_{k|k-1} \\
&- \beta_k C_{k-1}A^{-1}_{k-1}w_{k-1} + \beta_k v_{k-1} + (1-\beta_k)v_k - (1-\alpha_k) \\
&\quad \times C_k\hat{x}_{k|k-1} - \alpha_k C_{k-1}A^{-1}_{k-1}\hat{x}_{k|k-1}\Big)\Big(*\Big)^\top\Big],
\end{aligned}$$

or, Σ_k^{yy}

$$= \mathbb{E}\Big[\Big((1-\beta_k)C_k e_{k|k-1} + \beta_k C_{k-1}A_{k-1}^{-1}e_{k|k-1} + \beta_k v_{k-1} + (1-\beta_k)v_k - \beta_k C_{k-1}A_{k-1}^{-1}w_{k-1} + (\beta_k-\alpha_k) \times (C_{k-1}A_{k-1}^{-1} - C_k)\hat{x}_{k|k-1}\Big)\Big(*\Big)^\top\Big]. \tag{27}$$

现在，将两个括号中的各项相乘，并借助引理 2 前面提到的关系 (i)–(v) 应用期望算子，方程 (27) 化简为 (25)。□

引理 3 交叉协方差 Σ_k^{xy} 可表示为

$$\Sigma_k^{xy} = \Sigma_{k|k-1}[(1-\alpha_k)C_k + \alpha_k\theta_{k-1}]^\top - \alpha_k Q_{k-1}\theta_{k-1}^\top. \tag{28}$$

证明。根据定义，可以写为

$$\Sigma_k^{xy} = \mathbb{E}\Big[\Big(x_k - \hat{x}_{k|k-1}\Big)\Big(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}]\Big)^\top\Big].$$

对上述期望的第二项进行与引理 2 类似的运算，可得

$$\Sigma_k^{xy} = \mathbb{E}\Big[\Big(e_{k|k-1}\Big)\Big((1-\beta_k)C_k e_{k|k-1} + \beta_k C_{k-1}A_{k-1}^{-1} \times e_{k|k-1} + \beta_k v_{k-1} + (1-\beta_k)v_k - \beta_k C_{k-1}A_{k-1}^{-1}w_{k-1} + (\beta_k-\alpha_k)(C_{k-1}A_{k-1}^{-1} - C_k)\hat{x}_{k|k-1}\Big)^\top\Big]. \tag{29}$$

再次整理各项并利用关系式 $\mathbb{E}[e_{k|k-1}w_{k-1}^\top|\mathcal{I}_k] = Q_{k-1}$，方程 (29) 可进一步化简为

$$\begin{aligned}\Sigma_k^{xy} &= (1-\alpha_k)\Sigma_{k|k-1}C_k^\top + \alpha_k\Sigma_{k|k-1}(C_{k-1}A_{k-1}^{-1})^\top - \alpha_k Q_{k-1}\theta_{k-1}^\top \\ &= \Sigma_{k|k-1}[(1-\alpha_k)C_k + \alpha_k\theta_{k-1}]^\top - \alpha_k Q_{k-1}\theta_{k-1}^\top.\end{aligned}$$

□

定理 4 系统 (15) 和 (4) 所描述的系统的后验估计和误差协方差可表示为

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + \Sigma_k^{xy}(\Sigma_k^{yy})^{-1}\big(y_k - (1-\alpha_k)C_k\hat{x}_{k|k-1} - \alpha_k C_{k-1}\hat{x}_{k-1|k-1}\big), \\ \Sigma_{k|k} &= \Sigma_{k|k-1} - \Sigma_k^{xy}(\Sigma_k^{yy})^{-1}\Sigma_k^{xy\top}.\end{aligned} \tag{30}$$

证明。后验信息状态 $p(x_k|\mathcal{I}_k)$ 可以根据预测信息状态 $p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1})$ 和当前测量 y_k 计算得到。由 (13) 可得

$$\begin{aligned}p(x_k|\mathcal{I}_k) &= \gamma_k p(x_k, y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}, x_{k-1}) \\ &= \gamma_k p(x_k, y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}),\end{aligned} \tag{31}$$

其中，若已知 $e_{k-1|k-1}$，则 x_{k-1} 为冗余信息。假设 $p(y_k|\mathcal{I}_{k-1}, e_{k-1|k-1})$ 服从高斯分布，且给定为 $p(y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}) \sim \mathcal{N}(y_k; \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}], \Sigma_k^{yy})$，其中 $\mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}]$ 在引理 2 中推导得出。此外，由式 (18) 和 (23)，我们可以写作 $p(x_k|\mathcal{I}_{k-1}, e_{k-1|k-1}) \sim \mathcal{N}(x_k; \hat{x}_{k|k-1}, \Sigma_{k|k-1})$。因此，将联合高斯概率密度写作 $p(x_k, y_k|\mathcal{I}_{k-1}, e_{k-1|k-1})$，可得

$$p(x_k|\mathcal{I}_k) = \gamma_k\mathcal{N}\left(\begin{bmatrix}y_k \\ x_k\end{bmatrix}; \begin{bmatrix}\mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}] \\ \hat{x}_{k|k-1}\end{bmatrix}, \begin{bmatrix}\Sigma_k^{yy} & \Sigma_k^{yx} \\ \Sigma_k^{xy} & \Sigma_k^{xx}\end{bmatrix}\right), \tag{32}$$

其中协方差，$\Sigma_k^{xx} = \Sigma_{k|k-1}$。重新整理式 (32) 中的各项，并按照附录 A 中给出的方法进行平方运算，以确定 $p(x_k|\mathcal{I}_k)$，我们得到

$$\begin{aligned}p(x_k|\mathcal{I}_k) = \gamma_{k|k}\exp\Big[-\frac{1}{2}\Big(&x_k - \hat{x}_{k|k-1} - \Sigma_k^{xy}(\Sigma_k^{yy})^{-1} \times (y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}])\Big)^\top\Big(\Sigma_k^{xx} - \Sigma_k^{xy}(\Sigma_k^{yy})^{-1} \\ &\times (\Sigma_k^{xy})^\top\Big)^{-1}\Big(x_k - \hat{x}_{k|k-1} - \Sigma_k^{xy}(\Sigma_k^{yy})^{-1} \times (y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}])\Big)\Big].\end{aligned} \tag{33}$$

方程 (33) 是一个高斯密度，前提是所有协方差矩阵均可逆，且 $\mu_{1,\,k-1}$ 是一个充分小的正数，并且满足 $2\mu_{1,\,k-1}\Sigma_{k-1|k-1} < \mathrm{I}$。因此，后验估计和误差协方差分别是该概率密度的均值和协方差，$p(x_k|\mathcal{I}_k)$。最后，方程 (24) 和 (33) 推导出 (30)。□

注释 5. 如果 $\mu_{2,k} > 0$ 且 $\mu_{1,k-1} \geq 0$ 满足 $2\mu_{1,\,k-1}\Sigma_{k-1|k-1} < \mathrm{I}$，$\hat{x}_{k|k}$ 是相对于代价函数 $J_k(\hat{x}_{k|k}|y_{1:k})$ 的最优估计。

注释 6. 很容易证明，若测量没有时延（即 $\alpha_k = 0$, $\forall k$），所提出的滤波器将收敛到 [4], [8] 和[17] 中给出的 RSKF。

注释 7. 注意，如果 $\mu_{1,\,k-1} = 0$ 和 $\alpha_k = 0$, $\forall k$，则代价函数 (14) 变为不含过去误差的标准指数二次函数，并且所提滤波器与卡尔曼滤波器一致。

注释 8. 如果我们将后验估计重写为 $\hat{x}_{k|k} = L_k\hat{x}_{k|k-1} + K_k y_k$，其中 $L_k = \mathrm{I} - K_k\big((1-\alpha_k)C_k + \alpha_k C_{k-1}A_{k-1}^{-1}\big)$ 且 $K_k = \Sigma_k^{xy}(\Sigma_k^{yy})^{-1}$，那么利用 [19], 的命题1，可以证明所提出的估计器是无偏的，即 $\mathbb{E}[x_k - \hat{x}_{k|k}] = \mathbb{E}[e_{k|k}] = 0$，前提是 $\mu_{2,k} > 0$，且 $\mu_{1,k-1}$ 是一个较小的非负实数，其中 $2\mu_{1,k-1}\Sigma_{k-1|k-1} < \mathrm{I} \ \ \forall k$。

3.2 风险敏感参数的选择

在以往的研究中， [8], [4], 假设风险敏感参数不随时间变化；然而，这种限制并非必要，因此我们放宽了这一假设。只要风险敏感参数 $\mu_{2,\ k}$为正实数，它就不会影响代价函数 (14) 的优化。此外，后验估计和协方差的表达式与 $\mu_{2,\ k}$无关。不过，风险敏感参数 $\mu_{1,\ k-1}$的选取至关重要，因为它决定了计算当前估计时过去误差的贡献。由方程 (20) 可知，若要求预测协方差 $\Sigma_{k|k-1}$为正定，则参数 $\mu_{1,\ k-1}$必须满足 $|\Sigma_{k-1|k-1}^{-1} - 2\mu_{1,\ k-1}\mathbf{I}| < 0$。进一步求解该不等式，可得 $\mu_{1,\ k-1}$ 的上限；随后选取其值时，应与上限保持足够的安全裕度，以避免所得协方差出现病态。需要注意的是，在每个时间步计算 $\mu_{1,\ k-1}$ 会略微增加计算成本。

4 延迟风险敏感卡尔曼滤波器的稳定性

首先，我们建立均方意义下 RSKF 的稳定性准则，然后利用该准则评述所提滤波器的稳定性。

4.1 风险敏感卡尔曼滤波器的稳定性

RSKF 是一种最优滤波器 [4], 然而，最优性并不意味着稳定性 [14]。我们假设标称系统的状态有界，并将随机系统 (15)、(16) 的可观测性矩阵和可控性矩阵定义为 [14]

$$\begin{aligned}\mathcal{O}_{k,k-l} &= \sum_{i=k-l}^{k}(A_{i,k}+\Delta A_{i,k})^{\top}C_i^{\top}R_i^{-1}C_i(A_{i,k}+\Delta A_{i,k})\\ \mathcal{C}_{k,k-l} &= \sum_{i=k-l}^{k-1}(A_{k,i+1}+\Delta A_{k,i+1})^{\top}Q_i\\ &\qquad\times(A_{k,i+1}+\Delta A_{k,i+1});\ \forall k\geq l,\end{aligned}\tag{34}$$

其中， l 是正整数。$(A_{i,\ k}+\Delta A_{i,\ k})$ 是状态从时间步 k 转移到 i 的后向转移矩阵，而 $(A_{k,\ i}+\Delta A_{k,\ i})$ 是状态从时间步 i 转移到 k 的前向转移矩阵。实际系统 (15) 的转移矩阵定义如下:

$$\begin{aligned}(A_{i,k}+\Delta A_{i,k}) &= \prod_{j=1}^{k-i}(A_{k-j}+\Delta A_{k-j})^{-1},\\ (A_{k,i}+\Delta A_{k,i}) &= \prod_{j=i}^{k-1}(A_j+\Delta A_j);\quad 0\leq i<k,\end{aligned}\tag{35}$$

其中 $A_{k-1,k} = A_k^{-1}{}_{,k-1} = A_{k-1}^{-1}$, $\Delta A_{i,k} = \prod_{j=1}^{k-i}(A_{k-j}+\Delta A_{k-j})^{-1} - \prod_{j=1}^{k-i}A_{k-i}^{-1}$, $\Delta A_{k,i} = \prod_{j=i}^{k-1}(A_j+\Delta A_j) - \prod_{j=i}^{k-1}A_j$, 以及 $A_{k,k} = (A_{k,k}+\Delta A_{k,k}) = \mathrm{I}$。现在，如果由 (15) 和 (16) 描述的系统是 一致完全可观测且一致完全可控的，则称其具有一致完全可观测性和一致完全可控性，其中可观测性矩阵 $\mathcal{O}_{k,k-l}$ 和可控性矩阵 $\mathcal{C}_{k,k-l}$ 均为正定且有界[14], 即

$$\begin{aligned}0<\kappa_3\mathbf{I}\leq\mathcal{O}_{k,k-l}\leq\kappa_4\mathbf{I},\\ 0<\kappa_5\mathbf{I}\leq\mathcal{C}_{k,k-l}\leq\kappa_6\mathbf{I},\end{aligned}\tag{36}$$

其中 κ_3、 κ_4、 κ_5 和 κ_6 均为正实常数。

注释 9. 由于标称系统是实际系统的特例（当 $\Delta A_k = 0$ 时），标称系统的可观测性矩阵和可控性矩阵，分别表示为 $O_{k,\ k-l}$ 和 $\cup_{k,\ k-l}$，在 $\mathcal{O}_{k,\ k-l}$ 和 $\mathcal{C}_{k,\ k-l}$正定且有界时，也正定且有界。

现在，我们用 $\mathrm{P}^{k|k}$ 表示 RSKF 的后验误差协方差，不言而喻，只要 $\mathrm{P}_0 > 0$ 且 $2\mu_{1,\ k-1}\mathrm{P}_{k-1|k-1} < \mathrm{I}$，它将保持正定 [14], ，即 $\mathrm{P}_{k|k} > 0,\ k \geq 0$。稳定性分析的结果将在以下定理中给出。

定理5 如果系统 (15), (16) 一致完全可观测且一致完全可控，并且如果 $\mathrm{P}_0 > 0$, and $2\mu_{1,k-1}\mathrm{P}_{k-1|k-1} < \mathrm{I}$, then $\mathrm{P}^{k|k}$ 对于所有 $k \geq l$ 一致有界，前提是–$\mathrm{I} < O_k^{-1}{}_{,k-l}\Delta O_{k,k-l} < \mathrm{I}$, 其中

$$\Delta O_{k,k-l} = \sum_{i=k-l}^{k}A_{i,k}^{\top}C_i^{\top}R_i^{-1}C_i\Delta A_{i,k}.$$

证明。请参见附录 B。□

注释 10. 当 $\Delta O_{k,k-l}$ 的幅值大于或等于 $O_{k,k-l}$ 时，$^{k|k}$ 可能趋于无穷大（这意味着 RSKF 变得不稳定）。这表示状态转移中的不确定性

矩阵 $\Delta A_{j,k} \geq A_{j,k}$ 对于 $\Delta A_{j,k} > 0$，或者 $\Delta A_{j,k} \leq -A_{j,k}$ 对于 $\Delta A_{j,k} < 0$。

注释11. 如果模型中不存在不确定性，即 $\Delta O_{k,k-l} = 0$，则 (B.6) 中的不等式化简为 $\mathrm{P}^{k|k} \leq O_k^{-1}{}_{,k-l} + \cup_{k,k-l}$，正如 [14] 中针对卡尔曼滤波器所证明的那样。

4.2 随机延迟测量下风险敏感卡尔曼滤波器的稳定性

对于测量无时延的系统，即 α_k =0，所提滤波器化简为 RSKF。在定理5中，我们证明了：如果系统一致完全可控且可观测，那么在某些条件下该滤波器是稳定的。显然，测量中的时延不会改变系统的可控性，因此，如果系统在存在随机测量时延时仍保持可观测，则延迟风险敏感滤波器将是稳定的。

现在，我们将系统 (15) 的状态与前一时刻的状态向量增广，即 $X_k = [x_k^\top \ x_{k-1}^\top]^\top$，得到增广系统

$$\begin{aligned} X_k &= (\phi_{k,k-1} + \Delta\phi_{k,k-1})X_{k-1} + W_{k-1}, \\ y_k &= \mathbf{C}_k + v'_k, \end{aligned} \tag{37}$$

where

$$\phi_{k,k-1} + \Delta\phi_{k,k-1} = \begin{bmatrix} A_{k-1} + \Delta A_{k-1} & 0 \\ 0 & A_{k-2} + \Delta A_{k-2} \end{bmatrix},$$

$W_{k-1} = [w_{k-1}^\top w_{k-2}^\top]^\top$, $\mathbf{C}_k = [(1-\beta_k)C_k \ \ \beta_k C_{k-1}]$, 以及 $v'_k = (1-\beta_k)v_k + \beta_k v_{k-1}$。修正噪声的协方差可计算为 $R'_k = \mathbb{E}[v'_k v'^\top_k] = (1-\alpha_k)R_k + \alpha_k R_{k-1}$. 显然，对于任意 $\alpha_k \in [0,1]$，R'_k 的值介于 R_{k-1} 和 R_k 之间。请注意，在给定时间步 k，v'_k 不为白噪声的概率为 $\alpha_{k-1}(1-\alpha_k)$；当滤波器在连续两个时间步使用相同的测量数据时，它便会变为白噪声。在这种情况下，式（34）中定义的可观测性概念不再成立，因此很难对所提滤波器的稳定性给出严格证明。因此，我们以猜想的形式论证其稳定性，并假设噪声 v'_k为白噪声。

猜想6 如果系统 (4)、(15) 的测量不存在时延（α_k =0）且均匀完全可观测，那么对于 $\alpha_k \in [0, 1-\epsilon]$，底层系统也均匀完全可观测，其中 $0 < \epsilon \leq 1$。

论证。假设修正噪声 v'_k 为白噪声，则系统 (4)、(15) 的可观测性矩阵定义为

$$\mathbf{O}_{k,k-l} = \sum_{i=k-l}^{k} \mathbb{E}_\beta[(\phi_{i,k} + \Delta\phi_{i,k})^\top \mathbf{C}_i^\top R_i'^{-1} \mathbf{C}_i \times (\phi_{i,k} + \Delta\phi_{i,k})]. \tag{38}$$

现在，利用 (37) 以及第3 节中给出的关系 (iii)，我们写作 $\mathrm{O}_{k,\ k-l} = \begin{bmatrix} \mathrm{O}_{11} & 0 \\ 0 & \mathrm{O}_{22} \end{bmatrix}$，其中

$$\begin{aligned} \mathbf{O}_{11} &= \sum_{i=k-l}^{k} (1-\alpha_i)(A_{i,k} + \Delta A_{i,k})^\top C_i^\top R_i'^{-1} C_i \\ &\quad \times (A_{i,k} + \Delta A_{i,k}), \\ \mathbf{O}_{22} &= \sum_{i=k-l}^{k} \alpha_i (A_{i-1,k} + \Delta A_{i-1,k})^\top C_{i-1}^\top R'^{-1}_{i-1} \\ &\quad \times C_{i-1}(A_{i-1,k} + \Delta A_{i-1,k}). \end{aligned}$$

由于我们旨在求出增广状态 $X_k = [x_k^\top \ x_{k-1}^\top]^\top$中 x_k 的可观测性，因此只需证明 O_{11} of $\mathrm{O}_{k,k-l}$是一个正定且有限的矩阵 [6]。鉴于 $\alpha_j \in [0, 1-\epsilon]$，$\mathrm{O}_{11}$的界可表示为

$$\epsilon \mathcal{O}'_{k,k-l} \leq \mathbf{O}_{11} \leq \mathcal{O}'_{k,k-l}, \tag{39}$$

其中 $\mathcal{O}'_{k,k-l}$ 是(34)的可观测性矩阵，其协方差为 R'_i。由于 R'_i 始终位于 R_{i-1} 和 R_i 之间，可观测性矩阵 $\mathcal{O}'_{k,k-l}$ 满足(36)中给出的相同界。因此，O_{11} 对于 $0 < \epsilon \leq 1$ 是一个正定且有界的矩阵，并且系统(4)、(15)是一致完全可观测的。□

注释 12. 如果 $\epsilon = 0$ （即 $\alpha_k = 1$），这意味着所有测量值都延迟了一步，此时可观测性矩阵 O_{11}, 变为半正定，因此该系统不是一致完全可观测的。

5 时延概率的估计

在实践中，对于给定的系统，用户可能不知道时延概率 α_k。在这种情况下，必须先对其进行识别，然后才能进行估计。本节介绍一种方法，假设时延参数为平稳参数，即 $\alpha_k = \alpha, \forall k$，利用接收到的测量值并根据最大似然（ML）准则估计时延参数。该方法通过关于时延参数 α 最大化联合概率密度 $p_\alpha(y_{1:m})$ 来实现，可表示为 [29,25]

$$\hat{\alpha} = \arg \max_{\alpha \in [0,1]} p_\alpha(y_1, \cdots, y_m),$$

其中， m 是用于估计时延参数的测量值总数。利用链式法则，上述联合概率密度可改写为

$$p_\alpha(y_1,\cdots,y_m)=p(y_1)\prod_{k=2}^{m}p_\alpha(y_k|y_{1:k-1}), \tag{40}$$

其中，第一个接收测量值 y_1 被视为无时延，且与 α 相互独立。为简化计算，我们可以对式 (40) 取对数，如下所示:

$$\begin{aligned}L_\alpha(y_{1:m})&=\log p_\alpha(y_{1:m})\\&=\log p(y_1)+\sum_{k=2}^{m}\log p_\alpha(y_k|y_{1:k-1}).\end{aligned} \tag{41}$$

考虑到当前测量 y_k 与当前状态 x_k 以及前一状态 x_{k-1} 均相关，并利用贝叶斯定理对式 (41) 中的似然 $p_\alpha(y_k|y_{1:k-1})$ 进行处理，可得

$$\begin{aligned}p_\alpha(y_k|y_{1:k-1})=\int\int p_\alpha(y_k|x_k,x_{k-1})p(x_k|x_{k-1})\\\times\, p_\alpha(x_{k-1}|y_{1:k-1})dx_k dx_{k-1}.\end{aligned} \tag{42}$$

再次将接收测量值改写为

$$\begin{aligned}y_k&=(1-\beta_k)h_k(x_k)+\beta_k h_{k-1}(x_{k-1})\\&\quad+(1-\beta_k)v_k+\beta_k v_{k-1}\\&=\psi_k(x_k,x_{k-1})+v'_k,\end{aligned} \tag{43}$$

其中 $\psi_k(.)=(1-\beta_k)h_k(x_k)+\beta_k h_{k-1}(x_{k-1})$，且 $v'_k=(1-\beta_k)v_k+\beta_k v_{k-1}$。考虑到在给定 v_k 和 v_{k-1} 的条件下，修正后的测量噪声 v'_k 随时间构成独立序列，并利用式 (43)，状态似然 $p_\alpha(y_k|x_k,x_{k-1})$ 可定义为

$$p_\alpha(y_k|x_k,x_{k-1})=p_{v'_k}(y_k-\psi_k(x_k,x_{k-1})), \tag{44}$$

其中，$p_{v'_k}(\cdot)$ 是修正后测量噪声的概率密度函数。通过对 v'_k 和 β_k 的联合概率密度函数进行边缘化，可构造 $p_{v'_k}(\cdot)$，如下所示:

$$\begin{aligned}p(v'_k)&=p(v'_k,\beta_k=0)+p(v'_k,\beta_k=1)\\&=p(v'_k|\beta_k=0)p(\beta_k=0)+p(v'_k|\beta_k=1)p(\beta_k=1).\end{aligned} \tag{45}$$

由方程 (43)，我们可以写出 $p(v'_k|\beta_k=0)=p(v_k)$ 和 $p(v'_k|\beta_k=1)=p(v_{k-1})$。因此，方程 (45) 化简为

$$p(v'_k)=(1-\alpha_k)p_{v_k}+\alpha_k p_{v_{k-1}}. \tag{46}$$

将 (46) 代入 (44)，我们可以将该方程改写为

$$\begin{aligned}p_\alpha(y_k|x_k,x_{k-1})&=(1-\alpha_k)p_{v_k}(y_k-h_k(x_k))+\alpha_k p_{v_{k-1}}\\&\quad\times(y_k-h_{k-1}(x_{k-1})).\end{aligned} \tag{47}$$

现在，方程 (42) 中的似然函数可改写为:

$$\begin{aligned}p_\alpha(y_k|y_{1:k-1})&=\int\int p_\alpha(y_k|x_k,x_{k-1})\\&\quad\times p_\alpha(x_k,x_{k-1}|y_{1:k-1})dx_k dx_{k-1}\\&=\mathbb{E}[p_\alpha(y_k|x_k,x_{k-1})].\end{aligned} \tag{48}$$

将 (48) 代入 (41) 中的对数似然表达式，并利用 (47)，可得:

$$\begin{aligned}L_\alpha(y_{1:m})=\sum_{k=2}^{m}\log\big[\mathbb{E}[(1-\alpha_k)p_{v_k}(y_k-h_k(x_k))\\+\alpha_k p_{v_{k-1}}(y_k-h_{k-1}(x_{k-1}))]\big],\end{aligned} \tag{49}$$

由于测量 y_1 与参数 α 无关，因此将其忽略。由于在 (49) 中，时延参数和状态均未知，对数似然的解析最大化显然较为复杂，因此需要进行近似。可通过数值方法对 $\alpha\in[0,1]$ 最大化方程 (49)，同时采用序贯蒙特卡洛（SMC）近似计算对数似然。

5.1 对数似然的计算

考虑使用序贯蒙特卡洛近似计算期望，可将线性系统的对数似然函数写为:

$$\begin{aligned}L_\alpha(y_{1:m})=\sum_{k=2}^{m}\log\big[\frac{1}{N}\sum_{i=1}^{N}(1-\alpha_k)p_{v_k}(y_k-C_k x_k^i)\\+\alpha_k p_{v_{k-1}}(y_k-C_{k-1}x_{k-1}^i)\big],\end{aligned} \tag{50}$$

其中，x_{k-1}^i 从概率密度中采样，$p_\alpha(x_{k-1}|y_{1:k-1})=\mathcal{N}(x_{k-1};\hat{x}_{k-1|k-1},\Sigma_{k-1|k-1})$，$x_k^i$ 从概率密度中采样，$p(x_k|x_{k-1})=\mathcal{N}(x_k;A_{k-1}\hat{x}_{k-1|k-1},A_{k-1}\times(\Sigma^{-1}_{k-1|k-1}-2\mu_{1,k-1}\mathbf{I})^{-1}A^\top_{k-1}+Q_{k-1})$，$N$ 为样本总数。算法1概述了可用于估计时延参数的步骤。

6 仿真结果

本节通过模拟两个数值问题，展示所提滤波器相较于现有的 RSKF [4] 和 KF-RD [19] 的有效性。考虑测量中的随机时延的平稳统计特性，首先在

算法1 延迟参数估计

(1) 选择步长（l）的值，以及数量的测量值 (m)。
(2) 计算 $\alpha = 0$ 的对数似然 L_α，并设置 $L_{max} = L_0$.
(3) 对于 $\alpha = 0$: l : 1
- 对于 $k = 2$: m
  - 计算后验估计 $\hat{x}_{k-1|k-1}$ 以及协方差，$\Sigma_{k-1|k-1}$ 使用方程 (18) 和 (30)。
  - 采样 $x^i_{k-1} \sim \mathcal{N}(x_{k-1}; \hat{x}_{k-1|k-1}, \Sigma_{k-1|k-1})$
  - 根据其似然对 x^i_{k-1} 重新采样并得到 $\bar{x}^i_{k-1|k-1}$.
  - 计算 $x^i_k = A_{k-1}\bar{x}_{k-1|k-1}$。
  - 计算 the 状态 似然 ，$p_\alpha(y_k|x_k, x_{k-1})$，使用序贯蒙特卡洛近似法-tion 和（47）。
  - 计算 $L_\alpha = L_\alpha + \log(p_\alpha(y_k|x_k, x_{k-1}))$
- 结束循环
- 如果 $L_\alpha > LL$
  - $L_{max} = L_\alpha, \hat{\alpha} = \alpha$
- 如果结束

（4）结束循环

测量值中，通过最大化方程（50）所描述的似然来估计时延概率$\alpha \in [0, 1]$。随后，将估计得到的时延概率 α 用于针对给定问题实现所提滤波器。

6.1 问题：1

考虑一个二维线性随机系统 [27]

$$x_k = \begin{bmatrix} 0 & -0.5 \\ 1 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} -6 \\ 1 \end{bmatrix} w_{k-1}, \quad (51)$$
$$z_k = [-10\ 1]x_k + v_k,$$

其中，w_{k-1} 是均值为零、协方差为单位阵的白高斯序列，v_k 是均值为零、协方差为 $R = 3.6$的白高斯序列。w_{k-1} 和 v_k是不相关序列。系统建模中的不确定性由 $\Delta A = \begin{bmatrix} 0 & 0 \\ 0 & \delta \end{bmatrix}$. 表示。真实值初始化为 $x_0 \sim \mathcal{N}(0, \Sigma_0)$, 其中 $\Sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$。

假设测量时延统计量是平稳的，即 $\mathbb{E}[\beta_k] = \alpha, \forall k$, 则对每个集合进行时延参数估计（借助算法1），并绘于图 1。从图中可以看出，每个集合的估计值都接近其真实值。50个集合的估计值平均为0.291，而真实时延。

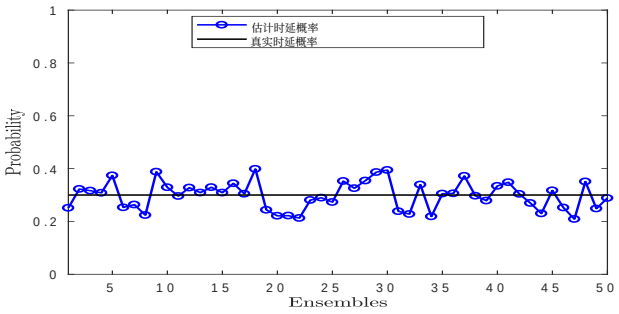


图 1. 估计时延概率。

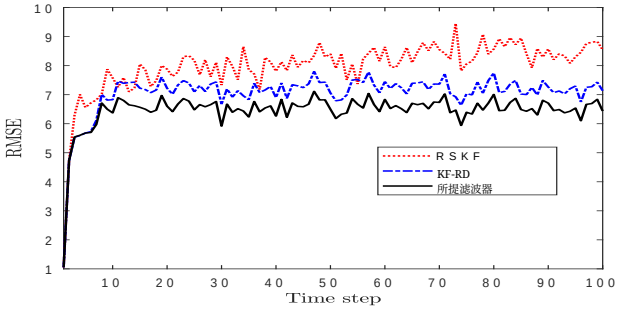


图2. 使用 $\delta = 0.35$ 和 $\alpha = 0.60$时状态1的均方根误差。

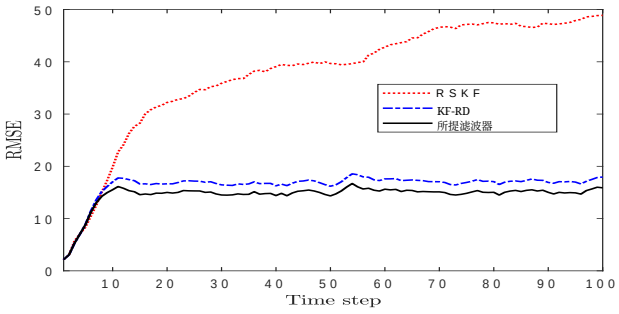


图3. 使用 $\delta = 0.35$ 和 $\alpha = 0.60$时状态2的均方根误差。

所考虑的参数为 0.3。所提滤波器与时延概率的估计值一同实现，并将其性能与 RSKF 和 KF-RD 进行比较。用于评估不同滤波器性能的指标是均方根误差（RMSE）和时间平均均方误差（平均均方误差），它们基于 500 次蒙特卡洛运行计算得出。风险敏感参数,μ_1, $_{k-1}$，选取为小于方程 $|\Sigma^{-1}_{k-1|k-1} - 2\mu_1,\ _{k-1}\mathbf{I}| = 0$ 的正实根。图 2 和图 3 显示了不确定性取值为 $\delta = 0.35$、真实时延概率为 $\alpha = 0.5$ 时的状态均方根误差。可以看出，在存在建模不确定性和测量值随机时延的情况下，所提滤波器的表现优于另外两种滤波器。

通过改变不确定性参数 δ 和概率 α 进行参数化研究。表1展示了在不同参数组合下，各种滤波器对状态1的平均均方误差。类似的结果是

表1 不同滤波器在各种 α 和 δ下的平均均方误差。

δ		α				
		0	0.20	0.40	0.60	0.80
0	RSKF	0.072	17.872	34.025	49.263	64.474
	KF-RD	0.068	13.217	23.102	31.241	36.889
	所提滤波器	0.072	13.294	23.332	31.636	37.616
0.15	RSKF	0.724	18.282	34.322	50.109	64.203
	KF-RD	1.008	13.393	24.201	33.102	40.001
	所提滤波器	0.724	13.171	23.710	32.510	39.141
0.25	RSKF	3.130	20.901	37.007	53.416	67.712
	KF-RD	4.880	15.152	26.910	37.713	48.360
	所提滤波器	3.130	14.105	25.001	35.130	43.925
0.35	RSKF	13.802	30.601	48.404	64.205	80.101
	KF-RD	23.401	20.681	35.593	52.791	71.702
	所提滤波器	13.802	16.901	29.595	51.893	55.796

针对状态2获得的结果未在此处展示。通过这项参数化研究，可以得出以下一些观察结论：

i. 在过程动态不存在任何不确定性（δ =0）的情况下，具有非零风险参数的所提滤波器比 RSKF 表现更好，且与 KF-RD 相当。这是因为 RSKF 中的风险敏感参数 $\mu_{1,\ k-1}$ 用于缩放累积过去误差，但即使无需最小化过去误差的指数，该参数也没有设为零；而 KF-RD 则以风险中性的方式运行。

ii. 随着 α 增大，所有滤波器的平均均方误差都会增大。

iii. 还应注意到，在测量值存在随机时延的情况下，专为处理随机时延而设计的 KF-RD 比 RSKF 更为有效。

iv. 在过程模型存在不确定性的情况下，无论时延概率 α 取何值，所提滤波器的表现始终优于另外两个滤波器。此外，当实际过程动态偏离标称过程动态越多，且测量值发生随机时延的可能性越大时，所提出的估计器相较于其他滤波器的性能提升就越显著。

6.2 问题：2

考虑一架飞机采用恒定转弯率模型，在二维平面内以固定但不确定的转弯率 Ω 进行机动转弯。飞机运动学的四维状态向量表示为 $x_k = [\eta_k\ \nu_k\ \dot{\eta}_k\ \dot{\nu}_k]^\top$，其中 η_k 和 ν_k 表示位置，$\dot{\eta}_k$ 和 $\dot{\nu}_k$ 分别表示沿 X 和 Y 坐标轴的速度。离散时间系统模型为

$$x_k = A_{k-1}x_{k-1} + Bw_{k-1}$$
$$z_k = C_k x_k + v_k,$$

其中

$$A_{k-1} = \begin{bmatrix} 1 & 0 & \frac{\sin \Omega T}{\Omega} & \frac{\cos \Omega T - 1}{\Omega} \\ 0 & 1 & \frac{1-\cos \Omega T}{\Omega} & \frac{\sin \Omega T}{\Omega} \\ 0 & 0 & \cos \Omega T & -\sin \Omega T \\ 0 & 0 & \sin \Omega T & \cos \Omega T \end{bmatrix},$$

T 为采样周期，噪声增益为 $B = aT^2[\frac{1}{2}\mathrm{I}_{2\times 2}\ \mathrm{I}_{2\times 2}]^\top$。测量值由受到噪声干扰的 X 和 Y 坐标给出，因此，$C_k = [\mathrm{I}_{2\times 2}\ \ 0_{2\times 2}]$。$w_{k-1}$ 和 v_k 为不相关的零均值白高斯序列，其协方差分别为 $Q = \mathrm{diag}([0.3^2\ \ 0.3^2\ \ 0.05^2])$和$R = \mathrm{diag}([12\ 12])$。初始值取为 $x_0 = [200\ 200\ 15\ 15]^\top$和 $\Sigma_0 = \mathrm{diag}([10^2\ 10^2\ 4^2\ 4^2])$。参数值选择为 $a = 1$，采样周期为 $T = 0.2$s，转弯率的标称值为 $\Omega = 3^o/\mathrm{s}$。在此问题中，所假设过程模型中的不确定性通过转弯率值引入。

每个集合的时延概率均使用算法1估计，并绘制于图4中。估计的时延概率的平均值为 0.481，而 α 的真实值取为 0.50。用于比较不同滤波器在该问题中表现的性能指标是位置和速度的均方根误差及平均均方误差。任意时间步 k的位置均方根误差可定义为

$$RMSE_{pos_k} = \sqrt{\sum_{m=1}^{M}(\{(\eta_k - \hat{\eta}_k)^2\}_m + \{(\nu_k - \hat{\nu}_k)^2\}_m)},$$

其中 M表示蒙特卡洛运行的总次数。使用 500 次蒙特卡洛运行来计算位置和速度的均方根误差及平均均方误差。如前所述，风险敏感参数 $\mu_{1,\ k}$ 在每个时间步进行计算，并收敛至 0.0013。

位置和速度的均方根误差分别绘制在图5和图6中，并考虑转弯率的不确定性，$\delta\Omega = 2\Omega/3$，其中未知延迟概率取为 $\alpha = 0.6$。从图中可以看出，所提滤波器优于 the

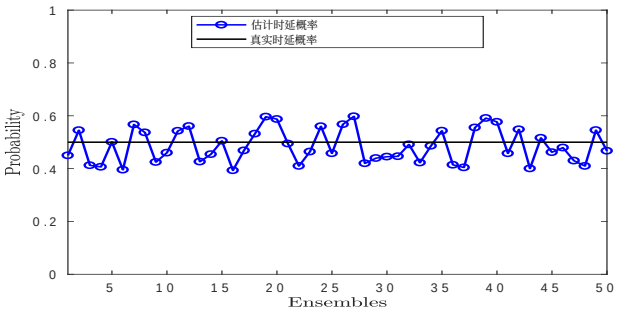


图4. 估计时延概率。

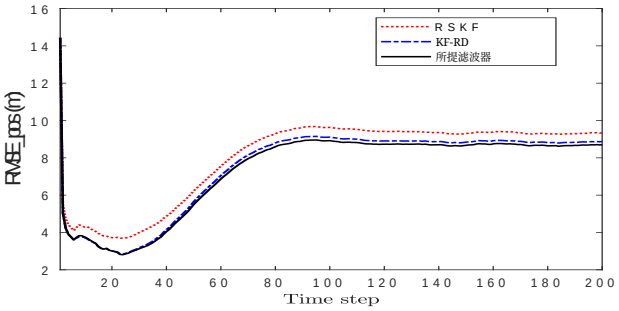


图5. 位置的均方根误差，使用 $\delta\Omega = 2\Omega/3$ 和 $\alpha = 0.60$。

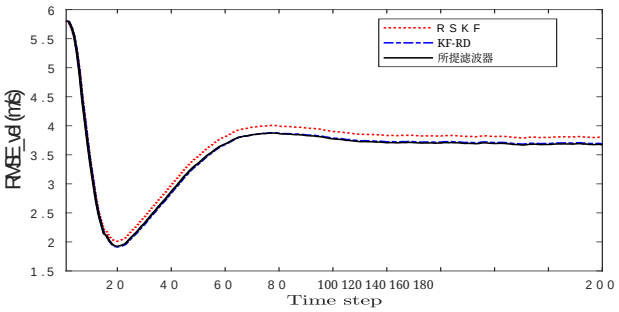


图 6. $\delta\Omega = 2\Omega/3$ 和 $\alpha = 0.60$条件下速度的均方根误差。

另外两种现有滤波器相比，KF-RD 的性能优于 RSKF。还应注意，所有曲线的变化趋势相似，这是因为所有实现的滤波器都是卡尔曼滤波器的变体，且系统为线性系统。当滤波器稳定在某些数值附近后，由于仿真中考虑的系统参数不随时间变化，因此滤波器的数值几乎会一直保持在这些数值附近。

从表 2 所示的参数化研究可以看出，在转弯率不存在不确定性且测量中没有随机时延的情况下，所提滤波器的性能与卡尔曼滤波器相当；而在存在相同情况时，其性能优于 KF-RD 和 RSKF。还可以注意到，在测量中不存在随机时延（$\alpha_k = 0$）时，所提滤波器与 RSKF 的平均均方误差值相等。

表 2 不同滤波器在各种 α 和 $\delta\Omega$条件下的平均均方误差。

$\delta\Omega$		α				
		0	0.20	0.40	0.60	0.80
0	RSKF	3.60	5.19	7.84	11.48	16.35
	KF-RD	3.59	4.50	5.02	5.13	4.92
	所提滤波器	3.60	4.52	5.05	5.15	4.94
$\frac{\Omega}{3}$	RSKF	15.70	18.16	22.04	27.03	32.68
	KF-RD	16.22	17.37	18.51	19.35	19.32
	所提滤波器	15.70	16.85	17.97	18.79	18.74
$\frac{\Omega}{2}$	RSKF	30.79	34.48	39.15	44.52	51.61
	KF-RD	31.96	34.07	35.68	36.41	37.36
	所提滤波器	30.79	32.86	34.40	35.10	35.99
$\frac{2\Omega}{3}$	RSKF	51.99	56.38	62.05	68.99	76.93
	KF-RD	54.12	56.57	58.96	61.04	62.66
	所提滤波器	51.99	54.38	56.69	58.67	60.18

7 结论

本文提出了一个针对存在模型不确定性和测量中随机时延的系统的贝叶斯估计通用框架。遵循所推导的通用框架，可得到线性高斯系统的闭式解。由于实际中估计器可能不知道与随机时延相关的概率，因此本文介绍了一种基于最大化接收测量值似然的方法，用于估计时延概率。此外，本文还利用一致完全可观测性和可控性判据，研究了测量中的随机时延对所提出的风险敏感估计器稳定性的影响。仿真结果证实，在标称条件下，所提滤波器的性能与卡尔曼滤波器相当；当系统存在模型不确定性和一步随机延迟时，其性能优于 RSKF 和 KF-RD。简言之，当底层系统可能同时存在模型不确定性和随机时延时，所提滤波器是一个合适的选择。

致谢

作者感谢加拿大麦克马斯特大学电气与计算机工程系的 Thia Kirubarajan 教授在开展这项工作过程中提出的宝贵建议，

作者还感谢匿名审稿人提出的宝贵意见，这些意见显著提升了本文的质量。

方程 (32) 的简化

考虑以下矩阵等价关系（参见 [26]的附录 A）：

$$\begin{bmatrix}\Sigma_k^{yy} & \Sigma_k^{yx}\\ \Sigma_k^{xy} & \Sigma_k^{xx}\end{bmatrix}^{-1} = \begin{bmatrix}V_k^{yy} & V_k^{yx}\\ V_k^{xy} & V_k^{xx}\end{bmatrix}, \tag{A.1}$$

其中

$$\begin{aligned}
V_k^{yy} &= (\Sigma_k^{yy} - \Sigma_k^{yx}(\Sigma_k^{xx})^{-1}\Sigma_k^{xy})^{-1},\\
V_k^{yx} &= -(\Sigma_k^{yy} - \Sigma_k^{yx}(\Sigma_k^{xx})^{-1}\Sigma_k^{xy})^{-1}\Sigma_k^{yx}(\Sigma_k^{xx})^{-1},\\
V_k^{xy} &= -(\Sigma_k^{xx})^{-1}\Sigma_k^{xy}(\Sigma_k^{yy} - \Sigma_k^{yx}(\Sigma_k^{yy})^{-1}\Sigma_k^{xy})^{-1},\\
V_k^{xx} &= (\Sigma_k^{xx} - \Sigma_k^{xy}(\Sigma_k^{yy})^{-1}\Sigma_k^{yx})^{-1},
\end{aligned}$$

并假设所有矩阵均可逆。利用 (A.1)，将 (32) 展开如下：

$$\begin{aligned}
&p(x_k, y_k|\mathcal{I}_{k-1}, e_{k-1|k-1})\\
&= \exp-\big[\frac{1}{2}(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}])^\top V_k^{yy}(y_k -\\
&\mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}]) + \frac{1}{2}(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}])^\top\\
&\times V_k^{yx}(x_k - \hat{x}_{k|k-1}) + \frac{1}{2}(x_k - \hat{x}_{k|k-1})^\top V_k^{xy}\times\\
&(y_k - C_k\hat{x}_{k|k-1}) + \frac{1}{2}(x_k - \hat{x}_{k|k-1})^\top V_k^{xx}(x_k - \hat{x}_{k|k-1})\big].
\end{aligned} \tag{A.2}$$

对于任意对称矩阵 F 以及向量 g、D 和 c，考虑以下恒等式 [18]:

$$\begin{aligned}
\frac{1}{2}g^\top Fg + g^\top D + c =& \frac{1}{2}(g + F^{-1}D)^\top F(g + F^{-1}D)\\
&+ c - \frac{1}{2}D^\top F^{-1}D.
\end{aligned} \tag{A.3}$$

现在，取 $g = x_k - \hat{x}_{k|k-1}$，$F = V_k^{xx}$，$D = V_k^{xy}(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}])$ $c = \frac{1}{2}(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}])^\top \times V_k^{yy}(y_k - \mathbb{E}[y_k|\mathcal{I}_{k-1}, e_{k-1|k-1}])$，并按照 (A.3) 处理 (A.2)。将不含 x_k 的项并入归一化常数 $\gamma_{k|k}$，并利用 (A.1) 中矩阵的等价关系，可得 (33)。注意，在最终表达式中，Σ_k^{yx} 被替换为 $\Sigma_k^{xy^\top}$ 。

B 定理 5 的证明

忽略过程噪声并考虑标称过程模型后，基于近期 l 次观测得到的 x_k 的最小二乘估计可表示为（见 [14]的引理 7.1）：

$$\bar{x}_{k|k-l:k} = O_{k,k-l}^{-1}\sum_{i=k-l}^{k} A_{i,k}^\top C_i^\top R_i^{-1} z_i, \quad k \geq l. \tag{B.1}$$

估计值 $\bar{x}_{k|k-l:k}$ 并非最优 [14]，因此，

$$\mathbf{P}_{k|k} \leq \mathbb{E}[(x_k - \bar{x}_{k|k-l:k})(x_k - \bar{x}_{k|k-l:k})^\top]. \tag{B.2}$$

根据式 (15) 和 (35)，计算式 (B.2) 中的协方差，可得

$$\begin{aligned}
x_i =& (A_{i,k} + \Delta A_{i,k})x_k - (A_{i,k} + \Delta A_{i,k})\\
&\times \sum_{j=i}^{k-1}(A_{k,j+1} + \Delta A_{k,j+1})w_j,
\end{aligned} \tag{B.3}$$

将由式 (B.3) 和 (16) 得到的 z_i 表达式代入式 (B.1)，可得

$$\begin{aligned}
&(x_k - \bar{x}_{k|k-l:k}) = x_k - O_{k,k-l}^{-1}\sum_{i=k-l}^{k} A_{i,k}^\top C_i^\top R_i^{-1}C_i\\
&\times (A_{i,k} + \Delta A_{i,k})x_k + O_{k,k-l}^{-1}\sum_{i=k-l}^{k} A_{i,k}^\top C_i^\top R_i^{-1}C_i\\
&\times (A_{i,k} + \Delta A_{i,k})\sum_{j=i}^{k-1}(A_{k,j+1} + \Delta A_{k,j+1})w_j\\
&- O_{k,k-l}^{-1}\sum_{i=k-l}^{k} A_{i,k}^\top C_i^\top R_i^{-1}v_i.
\end{aligned}$$

根据 $\Delta O_{k,k-l}$ 的定义，上述表达式可化简为

$$\begin{aligned}
&(x_k - \bar{x}_{k|k-l:k}) = -O_{k,k-l}^{-1}\Delta O_{k,k-l}x_k\\
&+ O_{k,k-l}^{-1}\sum_{i=k-l}^{k} A_{i,k}^\top C_i^\top R_i^{-1}C_i(A_{i,k} + \Delta A_{i,k})\times\\
&\sum_{j=i}^{k-1}(A_{k,j+1} + \Delta A_{k,j+1})w_j - O_{k,k-l}^{-1}\sum_{i=k-l}^{k} A_{i,k}^\top C_i^\top R_i^{-1}v_i,
\end{aligned}$$

and

$$
\begin{aligned}
&\mathbb{E}[(x_k-\bar{x}_{k|k-l:k})(x_k-\bar{x}_{k|k-l:k})^\top]\\
&=O^{-1}_{k,k-l}\Delta O_{k,k-l}\mathbb{E}[x_kx_k^\top]\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}\\
&\quad+\mathbb{E}\Big[\Big(O^{-1}_{k,k-l}\sum_{i=k-l}^{k}A^\top_{i,k}C_i^\top R_i^{-1}C_i(A_{i,k}+\Delta A_{i,k})\\
&\quad\times\sum_{j=i}^{k-1}(A_{k,j+1}+\Delta A_{k,j+1})w_j\Big)\Big(*\Big)^\top\Big]\\
&\quad+\mathbb{E}\Big[\Big(O^{-1}_{k,k-l}\sum_{i=k-l}^{k}A^\top_{i,k}C_i^\top R_i^{-1}v_i\Big)\Big(*\Big)^\top\Big]\\
&\le O^{-1}_{k,k-l}\Delta O_{k,k-l}\mathbb{E}[x_kx_k^\top]\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}\\
&\quad+\mathbb{E}\Big[\Big(O^{-1}_{k,k-l}\sum_{i=k-l}^{k}A^\top_{i,k}C_i^\top R_i^{-1}C_i(A_{i,k}+\Delta A_{i,k})\\
&\quad\times\sum_{j=k-l}^{k-1}(A_{k,j+1}+\Delta A_{k,j+1})w_j\Big)\Big(*\Big)^\top\Big]\\
&\quad+\mathbb{E}\Big[\Big(O^{-1}_{k,k-l}\sum_{i=k-l}^{k}A^\top_{i,k}C_i^\top R_i^{-1}v_i\Big)\Big(*\Big)^\top\Big],
\end{aligned}
$$

其中，(∗) 表示其左侧括号中给出的相同项。现在，利用关系 $\mathbb{E}[x_kx_k^\top]=\mathbf{P}_{k|k}+\hat{x}_{k|k}\hat{x}^\top_{k|k}$ 和 (B.2)，并重新排列各项，我们得到

$$
\begin{aligned}
&\mathbf{P}_{k|k}-O^{-1}_{k,k-l}\Delta O_{k,k-l}\mathbf{P}_{k|k}\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}\\
&\le O^{-1}_{k,k-l}\Delta O_{k,k-l}\hat{x}_{k|k}\hat{x}^\top_{k|k}\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}\\
&\quad+(\mathbf{I}+O^{-1}_{k,k-l}\Delta O_{k,k-l})\mathcal{C}_{k,k-l}(\mathbf{I}+O^{-1}_{k,k-l}\Delta O_{k,k-l})^\top\\
&\quad+O^{-1}_{k,k-l}.
\end{aligned}
\tag{B.4}
$$

对式 (B.4) 两边取欧几里得范数（||.||），并将其左侧的关系式写出，可得

$$
\begin{aligned}
&||\mathbf{P}_{k|k}||-||O^{-1}_{k,k-l}\Delta O_{k,k-l}\mathbf{P}_{k|k}\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}||\\
&\le||\mathbf{P}_{k|k}-O^{-1}_{k,k-l}\Delta O_{k,k-l}\mathbf{P}_{k|k}\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}||.
\end{aligned}
\tag{B.5}
$$

再次利用该关系

$$
\begin{aligned}
&||O^{-1}_{k,k-l}\Delta O_{k,k-l}\mathbf{P}_{k|k}\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}||\\
&\le||O^{-1}_{k,k-l}\Delta O_{k,k-l}||\;||\mathbf{P}_{k|k}||\;||\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}||
\end{aligned}
$$

将其代入式 (B.5)，并在式 (B.4) 中代入范数，可写为

$$
\begin{aligned}
||\mathbf{P}_{k|k}||&\le||O^{-1}_{k,k-l}\Delta O_{k,k-l}\hat{x}_{k|k}\hat{x}^\top_{k|k}\Delta O^\top_{k,k-l}O^{-1^\top}_{k,k-l}\\
&\quad+(\mathbf{I}+O^{-1}_{k,k-l}\Delta O_{k,k-l})\mathcal{C}_{k,k-l}(\mathbf{I}+O^{-1}_{k,k-l}\Delta O_{k,k-l})^\top\\
&\quad+O^{-1}_{k,k-l}||/(1-||O^{-1}_{k,k-l}\Delta O_{k,k-l}||^2).
\end{aligned}
\tag{B.6}
$$

回顾状态有界这一假设，并利用方程 (36)，显然，若过程模型中的不确定性有限且满足条件 $-\mathbf{I}<O^{-1}_{k-l}\Delta O_{k,k-l}<\mathbf{I}$，则对于所有 $k\ge l$，方程 (B.6) 中的 $\mathbf{P}_{k|k}$ in 有界。

参考文献

[1] Brian DO Anderson 和 John B Moore。《 *Optimal filtering*》。Courier Corporation，2012 年。[2]Yaakov Bar-Shalom、X Rong Li 和 Thiagalingam Kirubarajan。《 *Estimation with applications to tracking and navigation: theory algorithms and software*》。John Wiley & Sons，2004 年。[3] S Bhaumik、S Sadhu 和 TK Ghoshal。无迹卡尔曼滤波器的风险敏感公式。 *IET control theory & applications*，3(4)：375–382，2009 年。[4] Rene K Boel、Matthew R James 和 Ian R Petersen。鲁棒性与风险敏感滤波。 *IEEE Transactionson Automatic Control*，47(3)：451–461，2002 年。[5] Subhash Challa、Mark R Morelande、Darko Mu˘sicki 和 Robin J Evans。《 *Fundamentals of object tracking*》。Cambridge University Press，2011 年。[6] Chi-Tsong Chen。《 *Linear System Theory and Design*》。Oxford University Press，1999 年。[7] Subhrakanti Dey。鲁棒非线性估计与控制专题。1995 年。[8] Subhrakanti Dey 和 John B Moore。通过参考概率方法进行风险敏感滤波与平滑。 *IEEE Transactions on Automatic Control*，42(11)：1587–1591，1997 年。[9] Wassim M Haddad 和 Dennis S Bernstein。乘性白噪声线性系统降阶离散时间状态估计的最优投影方程。 *System& control letters*，8(4)：381–388，1987 年。[10] Wassim M Haddad、Dennis S Bernstein 和 Dennis Mustafa。混合范数 $H2/H_\infty$ 调节与估计：离散时间情形。 *System & Control Letters*，16(4)：235–247，1991 年。[11] Lidong He、Dongfang Han、Xiaofan Wang 和 Ling Shi。利用线性时序编码在数据包丢失网络上进行最优线性状态估计。 *Automatica*，49(4)：1075–1082，2013 年。[12]YC Ho 和 Robert Lee。随机估计与控制问题的贝叶斯方法。 *IEEE Transactions onAutomatic Control*，9(4)：333–339，1964 年。[13]David Jacobson。具有指数性能准则的最优随机线性系统及其与确定性微分博弈的关系。*IEEE Transactions on Automatic Control*，18(2)：124–131，1973 年。[14]Andrew H Jazwinski。《*Stochastic processes and filteringtheory*》。Courier Corporation，2007 年。

[15]Jing Ma 和 Shuli Sun。具有随机传感器延迟、多个数据包丢失和不确定观测的系统的最优线性估计器。 *IEEE Transactions on Signal Processing* 59 11 5181 5192, 2011 , ( ): –。

[16] Jing Ma 和 Shuli Sun。具有传输延迟和数据包丢失的网络化随机不确定系统的分布式融合滤波器。 *Signal Processing*，130：268–278，2017 年。[17] John B Moore、Robert J Elliott 和 Subhrakanti Dey。最小方差估计与控制的风险敏感推广。载于 *Nonlinear Control Systems Design 1995*，第 423–428 页。爱思唯尔，1995 年。[18]Kaare Brandt Petersen、Michael Syskind Pedersen 等。《矩阵烹饪书》。 *Technical University of Denmark*，7 15( )：510，2008 年。[19]A Ray、LW Liou 和 JH Shen。利用随机延迟测量进行状态估计。 *Journal of Dynamic Syst Measurement and Control* 115 1 19 , ( ): –26，1993 年。[20] Smita Sadhu、Shovan Bhaumik、Arnaud Doucet 和 Tapan Kumar Ghoshal。风险敏感滤波器的基于粒子方法的公式。 *Signal Processing*，89(3)：314–319，2009 年。[21] Jason L Speyer、C-H Fan 和 Ravi N Banavar。具有指数代价准则的最优随机估计。载于 *[1992] Proc of the 31st IEEE Conf on Decision and Control*，第 2293–2299 页。IEEE，1992 年。[22] Shuli Sun。具有有界随机测量延迟和数据包丢失的系统的线性最小方差估计器。 *Signal processing*，89(7)：1457–1466，2009 年。[23]Shuli Sun。具有或不具有时间戳的随机延迟和丢失测量离散时间系统的最优线性滤波器。 *IEEE Transactions on Automatic Control*，58 6 1551( )：–1556，2012 年。[24] Shuli Sun 和 Guanghui Wang。具有多个随机传输延迟和数据包丢失的网络化系统的建模与估计。 *Systems & Control Letters*，73：6–16，2014 年。[25]Ranjeet Kumar Tiwari、Shovan Bhaumik、Thiagalingam Kirubarajan 等。具有未知时延概率的随机延迟测量粒子滤波器。 *Sensors*，20 19( )：5689，2020 年。[26]Xiaoxu Wang、Yan Liang、Quan Pan 和 Chunhui Zhao。具有一步随机延迟测量的非线性系统高斯滤波器。 *Automatica*，49(4)：976–986，2013 年。[27]Lihua Xie、Yeng Chai Soh 和 Carlos E De Souza。不确定离散时间系统的鲁棒卡尔曼滤波。 *IEEE Transactions on Automatic Control* 39 6 1310 1314, 1994 , ( ): –。

[28]Huanshui Zhang、Gang Feng 和 Chunyan Han。随机时延系统的线性估计。 *Systems & Control Letters* 60 7 450 459, 2011 , ( ): –。

[29]Yonggang Zhang、Yulong Huang、Ning Li 和 Lin Zhao。具有一步随机延迟测量和未知时延概率的粒子滤波器。 *International Journal of Systems Science* 47 1 209 221, 2016 , ( ): –。