Proceedings of the 3rd IFAC Workshop on
Distributed Estimation and Control in Networked Systems
September 14-15, 2012. Santa Barbara, CA, USA

Multi-Sensor Linear State Estimation Under High Rate Quantization

Alex S. Leong * **Subhrakanti Dey** * **Girish N. Nair** *

* *Department of Electrical and Electronic Engineering, University of Melbourne, Vic. 3010, Australia (e-mails: asleong@unimelb.edu.au, sdey@unimelb.edu.au, gnair@unimelb.edu.au).*

Abstract: In this paper we consider state estimation of a discrete time linear system using multiple sensors, where the sensors quantize their individual innovations, which are then combined at the fusion center to form a global state estimate. We obtain an asymptotic approximation for the error covariance matrix that relates the system parameters and quantization levels used by the different sensors. Numerical results show close agreement with the true error covariance for quantization at high rates. An optimal rate allocation problem amongst the different sensors is also considered.

1. INTRODUCTION

Linear state estimation using multiple sensors is a commonly performed task in areas such as radar tracking and industrial monitoring. Nowadays, much of the communication systems used in practice are digital in nature. Therefore, analog measurements made by sensors will need to be quantized before transmission to a central processor or fusion center over a bandwidth limited wireless channel. Characterizing the performance loss due to quantization, for a linear state estimation problem, is the focus of this paper. This can be seen as a first step towards achieving a quantization rate versus state estimation error trade-off for linear dynamical systems, which is largely unavailable in the current literature.

We consider a discrete time linear system. A number of sensors take measurements, perform some local processing before transmitting a processed signal to a fusion center, that then combines these signals to form a global state estimate. At the sensor level, each sensor will quantize their innovations.[1] This is motivated by the fact that for unstable systems, while the state will become unbounded (leading to possible saturation of the quantizer), the innovations process remains of bounded variance (Anderson and Moore (1979)). These quantized innovations are then sent to a fusion center to form a global state estimate, using a modification of the decentralized scheme for unquantized Kalman filtering in Hashemipour et al. (1988).

The work of Nair and Evans (1998) gave structural results on optimal coding for state estimation with measurements obtained over a finite rate digital link, though the focus is more on determining minimum bit rates required for stability. For a linear quadratic control problem with quantized state feedback, the performance with high rate quantization has been studied in Gupta et al. (2006). The idea of quantizing innovations has also been considered in Msechu et al. (2008); You et al. (2011) with different filtering equations from ours. However You et al. (2011) only considers the case of a single sensor, while the multi-sensor setup in Msechu et al. (2008) does not involve a fusion center but instead requires sensors to broadcast their quantized innovations to all other sensors. In Sukhavasi and Hassibi (2011) a filter which involves quantizing the true innovations at the sensor (rather than the approximation to the true innovations considered here and in Msechu et al. (2008); You et al. (2011)) is given, but it is shown that for unstable systems the mean squared error always becomes unbounded with this scheme. Particle filtering schemes are also considered in Sukhavasi and Hassibi (2011), though such schemes are difficult to analyze theoretically.

The paper is organized as follows. We first consider the single sensor case in order to motivate our choice of quantization method, filtering equations, and asymptotic analysis techniques for high rate quantization. We then consider the multi-sensor case. We obtain an asymptotic approximation for the error covariance in terms of the number of quantization levels used by the different sensors, as well as the system parameters. Numerical comparisons are made between the asymptotic expression and Monte Carlo simulations of the true error covariance matrix. While our asymptotic expressions are derived assuming high rate quantization, numerical results suggest that they are quite accurate even for rates as low as 3 bits per sample. We also solve a rate allocation problem in the multi-sensor case for minimizing the trace of the error covariance matrix at the fusion centre when the total rate across the sensors is limited. The special case of scalar systems has been analyzed using different techniques in Leong et al. (2012).

2. SINGLE SENSOR

The system is a discrete time vector linear system

$$x_{k+1} = Ax_k + w_k$$

where $x_k \in \mathbb{R}^n$ and w_k is i.i.d. zero mean Gaussian with covariance matrix $\Sigma_w \geq 0$. The sensor makes a vector measurement

$$y_k = Cx_k + v_k$$

[1] To be more specific, we quantize an approximation to the true innovations due to the nonlinear effect of quantization

978-3-902823-22-9 /12/$20.00 © 2012 IFAC

10.3182/20120914-2-US-4030.00020

where $y_k \in \mathbb{R}^m$, and v_k is i.i.d. zero mean Gaussian with covariance matrix $\Sigma_v > 0$. We assume that $\{w_k\}$ and $\{v_k\}$ are mutually independent, and that the pair $(A, \Sigma_w^{1/2})$ is stabilizable and the pair (A, C) is detectable.

2.1 Kalman filter

We briefly review a few properties of the Kalman filter. Define the state estimates and error covariances[2] $\hat{x}^{kf}_{k|k-1} = \mathbb{E}[x_k|y_0,\dots,y_{k-1}]$, $\hat{x}^{kf}_{k|k} = \mathbb{E}[x_k|y_0,\dots,y_k]$, $P^{kf}_{k|k-1} = \mathbb{E}[(x_k - \hat{x}^{kf}_{k|k-1})(x_k - \hat{x}^{kf}_{k|k-1})^T|y_0,\dots,y_{k-1}]$, $P^{kf}_{k|k} = \mathbb{E}[(x_k - \hat{x}^{kf}_{k|k})(x_k - \hat{x}^{kf}_{k|k})^T|y_0,\dots,y_k]$. The innovations process is

$$\tilde{y}^{kf}_k = y_k - \mathbb{E}[y_k|y_0,\dots,y_{k-1}] = y_k - C\hat{x}^{kf}_{k|k-1}$$

It is well-known (see e.g. Anderson and Moore (1979)) that

$$\tilde{y}^{kf}_k \sim N(0, CP^{kf}_{k|k-1}C^T + \Sigma_v)$$

The Kalman filtering equations (no quantization) are:

$$\begin{aligned}
\hat{x}^{kf}_{k|k-1} &= A\hat{x}^{kf}_{k-1|k-1}\\
\hat{x}^{kf}_{k|k} &= \hat{x}^{kf}_{k|k-1} + K^{kf}_k(y_k - C\hat{x}^{kf}_{k|k-1}) = \hat{x}^{kf}_{k|k-1} + K^{kf}_k\tilde{y}^{kf}_k\\
K^{kf}_k &= P^{kf}_{k|k-1}C^T(CP^{kf}_{k|k-1}C^T + \Sigma_v)^{-1}\\
P^{kf}_{k|k-1} &= AP^{kf}_{k-1|k-1}A^T + \Sigma_w\\
P^{kf}_{k|k} &= P^{kf}_{k|k-1} - P^{kf}_{k|k-1}C^T(CP^{kf}_{k|k-1}C^T + \Sigma_v)^{-1}CP^{kf}_{k|k-1}
\end{aligned} \tag{1}$$

Under the stabilizability and detectability assumptions, as $k \to \infty$, $P^{kf}_{k|k-1}$ converges to a steady state value P^{kf}_∞ that satisfies the algebraic Riccati equation:

$$P^{kf}_\infty = AP^{kf}_\infty A^T + \Sigma_w - AP^{kf}_\infty C^T(CP^{kf}_\infty C^T + \Sigma_v)^{-1}CP^{kf}_\infty A^T \tag{2}$$

2.2 Quantized filtering scheme

In this paper we consider a suboptimal quantized filtering scheme where we run a slightly modified version of the unquantized filtering equations given in (1):

$$\begin{aligned}
\hat{x}_{k|k-1} &= A\hat{x}_{k-1|k-1}\\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k q_k(y_k - C\hat{x}_{k|k-1})\\
K_k &= P_{k|k-1}C^T(CP_{k|k-1}C^T + \Sigma_v + \Sigma_{n,k})^{-1}\\
P_{k|k-1} &= AP_{k-1|k-1}A^T + \Sigma_w\\
P_{k|k} &= P_{k|k-1} - P_{k|k-1}C^T(CP_{k|k-1}C^T + \Sigma_v + \Sigma_{n,k})^{-1}CP_{k|k-1}
\end{aligned} \tag{3}$$

where $q_k(y_k - C\hat{x}_{k|k-1})$ is the quantization of the vector $y_k - C\hat{x}_{k|k-1}$, and the matrix $\Sigma_{n,k}$ is a term to account for the quantization noise. Note that due to quantization $\hat{x}_k$, P_k, and $y_k - C\hat{x}_{k|k-1}$ are not the true conditional mean, error covariance matrix and innovations respectively, but for high rate quantization the approximations should be quite accurate.

Let $P_k \triangleq P_{k|k-1}$. Under high rate quantization, we assume that the quantity $y_k - C\hat{x}_{k|k-1}$ is approximately $N(0, CP_kC^T + \Sigma_v)$. Since $y_k - C\hat{x}_{k|k-1}$ is a vector, we will use vector quantizers with N quantization values. In general, optimal vector quantization (optimal in terms of minimizing the distortion) is a difficult problem where many open questions remain. The LBG algorithm (Gersho and Gray (1992)) can be used to find locally optimal vector quantizers but requires numerical methods to compute, and the resulting quantizers often lack structure. We thus consider mostly the case of lattice quantizers, whose regular structure makes for efficient encoding and implementation.[3]

We will first diagonalize $CP_kC^T + \Sigma_v$ as

$$CP_kC^T + \Sigma_v = U_k\Lambda_kU_k^T$$

where U_k is a unitary (in fact orthogonal) matrix of eigenvectors and Λ_k is a diagonal matrix of eigenvalues (we recall that every real symmetric matrix is diagonalizable, and the eigenvalues of a positive definite matrix are positive). Then the distribution

$$N(0, CP_kC^T + \Sigma_v) = U_k\Lambda_k^{1/2}N(0, I).$$

Now for zero mean multivariate Gaussian distributions with i.i.d. components, asymptotically optimal lattice quantizers have been considered in Moo (1998), with analytical expressions derived for the distortion and sizes of the cells in the lattice quantizer. Thus one way to vector quantize $y_k - C\hat{x}_{k|k-1}$ is to first multiply it by $(U_k\Lambda_k^{1/2})^{-1}$ to transform into (approximately) $N(0, I)$ random vectors, quantizing this using the asymptotically optimal lattice quantizers from Moo (1998), and then multiplying the quantized vector by $U_k\Lambda_k^{1/2}$, i.e.

$$q_k(y_k - C\hat{x}_{k|k-1}) = U_k\Lambda_k^{1/2}\tilde{q}((U_k\Lambda_k^{1/2})^{-1}(y_k - C\hat{x}_{k|k-1}))$$

where $\tilde{q}$ is the lattice quantizer of Moo (1998) (note that multiplication by $U_k\Lambda_k^{1/2}$ is a linear transformation which preserves the number of values in the quantizer codebook). For asymptotically optimal lattice quantization of a Gaussian random vector with i.i.d. components, each having variance σ^2, the distortion per dimension $D_N \triangleq \frac{1}{m}\mathbb{E}[(x - \tilde{q}(x))^T(x - \tilde{q}(x))]$ is given by (see Moo (1998)):

$$D_N \sim \frac{M(S_0)V^{2/m}}{\eta^2}\frac{\frac{2}{m}\ln N}{N^{2/m}} \triangleq \delta_N$$

where m represents the dimension of the vector to be quantized, N the number of quantization values, $\eta = \frac{1}{\sigma}\sqrt{\frac{\Gamma(3/2)}{\Gamma(1/2)}} = \frac{1}{\sigma}\sqrt{\frac{1}{2}}$, $V = \frac{(\Gamma(1/2))^m}{\Gamma(m/2+1)} = \frac{\pi^{m/2}}{\Gamma(m/2+1)}$, $M(S_0) = \frac{\frac{1}{k}\int_{S_0}||x-y||_2^2dx}{v(S_0)^{1+2/m}}$, S_0 is a Voronoi cell of the lattice, $v(S_0)$ is the volume of S_0, and $M(S_0)$ is the normalized moment of inertia of S_0. The asymptotically optimal scaling a_N of the fundamental cells of the lattice is given by:

$$a_N \sim \sqrt{\frac{2}{m}}\frac{1}{\eta}\left(\frac{V}{v(S_0)}\right)^{1/m}\frac{\sqrt{\ln N}}{N^{1/m}}$$

Since the components are i.i.d., if we assume that the quantization errors are spread evenly amongst all components, then $\mathbb{E}[(x - \tilde{q}(x))(x - \tilde{q}(x))^T] \approx \delta_N I$, and the term $\Sigma_{n,k}$ in (3) is then defined as

$$\Sigma_{n,k} \triangleq U_k\Lambda_k^{1/2}(\delta_N I)\Lambda_k^{1/2}U_k^T = \delta_N(CP_kC^T + \Sigma_v). \tag{4}$$

[2] Similar to Sukhavasi and Hassibi (2011), we use the superscript "*kf*" to denote the true Kalman filtering quantities.

[3] For scalar measurements, lattice quantization reduces to the case of the uniform scalar quantizer. For scalar measurements we will also consider the case of optimal quantization, see Section 2.2.1.

Remark: The expressions for the asymptotic distortion and asymptotically optimal scaling clearly depends on the choice of fundamental cell S_0. However, the optimal shapes for S_0 are generally not known. [4] As an example, suppose S_0 is an m-dimensional cube of length 1. Then one can easily compute that $v(S_0) = 1$ and $M(S_0) = 1/12$. Therefore in this case we have (with $\sigma^2 = 1$):

$$D_N \sim \frac{\pi}{3m(\Gamma(m/2+1))^{2/m}} \frac{\ln(N)}{N^{2/m}},$$
$$a_N \sim \frac{2\sqrt{\pi}}{\sqrt{m}(\Gamma(m/2+1))^{1/m}} \frac{\sqrt{\ln N}}{N^{1/m}}.$$

For dimension $m = 1$, these expressions further simplify to $D_N \sim \frac{4\ln N}{3N^2}$ and $a_N \sim \frac{4\sqrt{\ln N}}{N}$, see below.

The case of scalar measurements In the special case of scalar measurements, the quantity to be quantized is scalar, and the lattice quantizer reduces to the uniform quantizer. The asymptotically optimal step size for uniform quantization of Gaussian variables has also been derived in Hui and Neuhoff (2001). Under high rate quantization, the step size Δ_N is asymptotically $\Delta_N \sim \frac{4\sqrt{\ln N}}{N}\sigma$ and the resulting squared error distortion is asymptotically $D_N \sim \frac{4\ln N}{3N^2}\sigma^2$ where σ^2 is the variance of the Gaussian random variable that is to be quantized.

Furthermore, for Gaussian random variables the optimal quantizers have also been tabulated in Max (1960) for N up to 36, and can be computed for other values of N relatively easily. For N large, it is known that the resulting squared error distortion satisfies the Panter-Dite formula, so that $D_N \sim \frac{\pi\sqrt{3}}{2N^2}\sigma^2$. We therefore have (note that here $CP_kC^T + \Sigma_v$ is scalar) $\Sigma_{n,k} = \delta_N(CP_kC^T + \Sigma_v)$, where

$$\delta_N = \begin{cases} \frac{\pi\sqrt{3}}{2N^2}, & \text{optimal quantization} \\ \frac{4\ln N}{3N^2}, & \text{optimal uniform quantization} \end{cases} \tag{5}$$

2.3 Asymptotic analysis

We wish to determine the asymptotic behaviour of $\text{tr}(P_\infty)$ for large N, where P_∞ is the limit of P_k as $k \to \infty$ that from (3) and (4) satisfies the equation

$$P_\infty = AP_\infty A^T + \Sigma_w - \frac{AP_\infty C^T(CP_\infty C^T + \Sigma_v)^{-1}CP_\infty A^T}{1+\delta_N} \tag{6}$$

In the scalar case an analytical expression for P_∞ can be derived and analyzed (see Leong et al. (2012)) to find asymptotic approximations. However, in the vector case we do not have a closed form expression for P_∞ or $\text{tr}(P_\infty)$. Instead we will use a different technique, which is based on the method used to find asymptotic solutions to algebraic equations in perturbation theory (see e.g. Holmes (1995)), but extended to matrices. With this technique, we can in fact derive an asymptotic expression for the whole matrix P_∞, and not just its trace.

Notation: We will call a matrix $O(\mathbb{1})$ if all its entries are $O(1)$, and call a matrix $O(\epsilon\mathbb{1})$ if all its entries are $O(\epsilon)$.

[4] Even for lattice quantization of uniformly distributed random vectors, the optimal cell shapes are only known for dimensions $m = 1, 2, 3$.

Motivated by the asymptotic result in the scalar case (Leong et al. (2012)), where it is shown that $P_\infty = P_\infty^{kf} + \kappa\delta_N + O(\delta_N^2)$ for some constant κ, we assume that P_∞ can be written in the form

$$P_\infty = \Phi_0 + \delta_N\Phi_1 + \delta_N^2\Phi_2 + \ldots \tag{7}$$

where $\Phi_0, \Phi_1, \ldots$ etc. are matrices not dependent on N. Substituting the form (7) into equation (6), we have

$$\begin{aligned}
&\Phi_0 + \delta_N\Phi_1 + \cdots = A(\Phi_0 + \delta_N\Phi_1 + \ldots)A^T + \Sigma_w \\
&-A(\Phi_0+\delta_N\Phi_1+\ldots)C^T(C(\Phi_0+\delta_N\Phi_1+\ldots)C^T+\Sigma_v)^{-1} \\
&\times C(\Phi_0 + \delta_N\Phi_1 + \ldots)A^T\frac{1}{1+\delta_N} \\
&= A(\Phi_0+\delta_N\Phi_1+\ldots)A^T+\Sigma_w - A(\Phi_0+\delta_N\Phi_i+\ldots)C^T \\
&\times [(C\Phi_0C^T + \Sigma_v)^{-1} - \delta_N(C\Phi_0C^T + \Sigma_v)^{-1}C\Phi_1C^T \\
&\times (C\Phi_0C^T + \Sigma_v)^{-1} + \ldots] \\
&\times C(\Phi_0 + \delta_N\Phi_1 + \ldots)A^T(1 - \delta_N + \ldots)
\end{aligned} \tag{8}$$

where the second equality follows from the following generalization of a result from p.26 of Holmes (1995):

Proposition 1. Suppose $||\sum_{i=1}^{M} \epsilon_i A^{-1}B_i|| < 1$ and A is invertible. Then as $\epsilon_i \to 0, i = 1, \ldots, M$,

$$(A + \sum_{i=1}^{M}\epsilon_iB_i)^{-1} = A^{-1} - \sum_{i=1}^{M}\epsilon_iA^{-1}B_iA^{-1} + \sum_{i,j}O(\epsilon_i\epsilon_j\mathbb{1})$$

Due to space constraints, the proof is omitted.

Similar to the asymptotic technique in Holmes (1995), we can derive an asymptotic expression for P_∞ by successively solving for Φ_0, Φ_1, etc. Equating the $O(\mathbb{1})$ terms in (8) we obtain the equation

$$\Phi_0 = A\Phi_0A^T + \Sigma_w - A\Phi_0C^T(C\Phi_0C^T + \Sigma_v)^{-1}C\Phi_0A^T$$

which can be used to compute Φ_0. Comparing with (2), we see that $\Phi_0 = P_\infty^{kf}$.

Equating the $O(\delta_N\mathbb{1})$ terms in (8) we obtain the equation

$$\begin{aligned}
\Phi_1 &= A\Phi_1A^T - A\Phi_1C^T(C\Phi_0C^T + \Sigma_v)^{-1}C\Phi_0A^T \\
&- A\Phi_0C^T(C\Phi_0C^T + \Sigma_v)^{-1}C\Phi_1A^T \\
&+ A\Phi_0C^T(C\Phi_0C^T+\Sigma_v)^{-1}C\Phi_1C^T(C\Phi_0C^T+\Sigma_v)^{-1}C\Phi_0A^T \\
&+ A\Phi_0C^T(C\Phi_0C^T + \Sigma_v)^{-1}C\Phi_0A^T \\
&= \left(A - A\Phi_0C^T(C\Phi_0C^T + \Sigma_v)^{-1}C\right)\Phi_1 \\
&\quad \times \left(A - A\Phi_0C^T(C\Phi_0C^T + \Sigma_v)^{-1}C\right)^T \\
&\quad + A\Phi_0C^T(C\Phi_0C^T + \Sigma_v)^{-1}C\Phi_0A^T
\end{aligned} \tag{9}$$

which is a Lyapunov equation.

Thus, asymptotically we have $P_\infty = P_\infty^{kf} + \Phi_1\delta_N + O(\delta^2(N)\mathbb{1})$, where δ_N decays to zero at the rate $\frac{\ln(N)}{N^{2/m}}$ for the lattice quantizers of Moo (1998), P_∞^{kf} can be found by solving numerically the algebraic Riccati equation (2), and Φ_1 can be found by solving numerically the Lyapunov equation (9). For the special case of scalar measurements and optimal quantization, δ_N decays at the rate $\frac{1}{N^2}$.

3. MULTIPLE SENSORS

The system is still the vector linear system $x_{k+1} = Ax_k + w_k$, with $x_k \in \mathbb{R}^n$, but now with M different sensors each making measurements:

$$y_{i,k} = C_i x_k + v_{i,k}, i = 1, \dots, M$$

where $y_{i,k} \in \mathbb{R}^{m_i}$, $w_k \sim N(0, \Sigma_w)$ and $v_{i,k} \sim N(0, \Sigma_{i,v})$. We assume that $\{w_k\}$ and $\{v_{i,k}\}, \forall i$ are mutually independent, and that the pair $(A, \Sigma_w^{1/2})$ is stabilizable and the pairs (A, C_i) are detectable for each i.

It is assumed that the individual sensors can perform some local processing, with a fusion center then using an appropriate fusion rule to compute a global estimate of the state x_k. See Fig. 1 for a diagram of the system model.

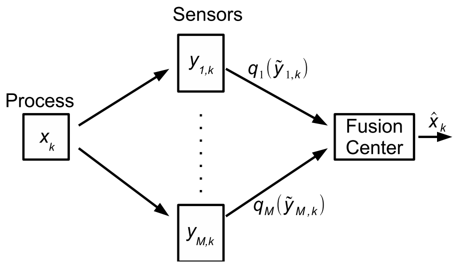

Fig. 1. System model: Multi-sensor

3.1 Decentralized Kalman filter

In Hashemipour et al. (1988), it is shown that in the case where there is no quantization, each sensor can run its own individual Kalman filter to obtain local state estimates, which can then be combined at the fusion center to obtain a global state estimate, that is the same as if the fusion center had access to the individual measurements. We summarize the equations below.

Define the local estimates and error covariances: $\hat{x}^{kf}_{i,k|k-1} = \mathbb{E}[x_k|y_{i,0}, \dots, y_{i,k-1}]$, $\hat{x}^{kf}_{i,k|k} = \mathbb{E}[x_k|y_{i,0}, \dots, y_{i,k}]$, $P^{kf}_{i,k|k-1} = \mathbb{E}[(x_k - \hat{x}^{kf}_{i,k|k-1})(x_k - \hat{x}^{kf}_{i,k|k-1})^T|y_{i,0}, \dots, y_{i,k-1}]$, $P^{kf}_{i,k|k} = \mathbb{E}[(x_k - \hat{x}^{kf}_{i,k|k})(x_k - \hat{x}^{kf}_{i,k|k})^T|y_{i,0}, \dots, y_{i,k}]$, and the global quantities: $\hat{x}^{kf}_{k|k-1} = \mathbb{E}[x_k|\mathbf{y}_0, \dots, \mathbf{y}_{k-1}]$,
$\hat{x}^{kf}_{k|k} = \mathbb{E}[x_k|\mathbf{y}_0, \dots, \mathbf{y}_k]$,
$P^{kf}_{k|k-1} = \mathbb{E}[(x_k - \hat{x}^{kf}_{k|k-1})(x_k - \hat{x}^{kf}_{k|k-1})^T|\mathbf{y}_0, \dots, \mathbf{y}_{k-1}]$,
$P^{kf}_{k|k} = \mathbb{E}[(x_k - \hat{x}^{kf}_{k|k})(x_k - \hat{x}^{kf}_{k|k})^T|\mathbf{y}_0, \dots, \mathbf{y}_k]$, where $\mathbf{y}_k \triangleq (y^T_{1,k}, \dots, y^T_{M,k})^T$.

The sensors run their individual Kalman filtering equations, for $i = 1, \dots, M$, whose equations take the form (1) but replacing y_k with $y_{i,k}$, C with C_i, Σ_v with $\Sigma_{i,v}$ etc. The fusion center makes use of the local estimates $\hat{x}^{kf}_{i,k|k-1}$ and $\hat{x}^{kf}_{i,k|k}$ and local error covariances $P^{kf}_{i,k|k-1}$ and $P^{kf}_{i,k|k}$ to compute global estimates as follows:

$$\hat{x}^{kf}_{k|k-1} = A\hat{x}^{kf}_{k-1|k-1}$$

$$\hat{x}^{kf}_{k|k} = P^{kf}_{k|k}\Big(P^{kf^{-1}}_{k|k-1}\hat{x}^{kf}_{k|k-1} + \sum_{i=1}^{M}\Big\{P^{kf^{-1}}_{i,k|k}\hat{x}^{kf}_{i,k|k} - P^{kf^{-1}}_{i,k|k-1}\hat{x}^{kf}_{i,k|k-1}\Big\}\Big)$$

$$P^{kf}_{k|k-1} = AP^{kf}_{k-1|k-1}A^T + \Sigma_w$$

$$P^{kf}_{k|k} = P^{kf}_{k|k-1} - P^{kf}_{k|k-1}\mathbf{C}^T(\mathbf{C}P^{kf}_{k|k-1}\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}P^{kf}_{k|k-1}$$

where $\mathbf{C} = \left[\, C_1^T \mid \dots \mid C_M^T \,\right]^T$ and Σ_v is a block diagonal matrix given by $\Sigma_v = \text{diag}(\Sigma_{1,v}, \dots, \Sigma_{M,v})$. Note that instead of the sensors sending their local estimates and error covariances, the local innovations $\tilde{y}^{kf}_{i,k} = y_{i,k} - C_i\hat{x}^{kf}_{i,k|k-1}$ can be sent to the fusion center instead, since the fusion center can reconstruct $\hat{x}^{kf}_{i,k|k}$, $\hat{x}^{kf}_{i,k+1|k}$, $P^{kf}_{i,k|k}$ and $P^{kf}_{i,k+1|k}$ from $\tilde{y}_{i,k}$ provided it has knowledge of all the sensor parameters C_i and $\Sigma_{i,v}$, $i = 1, \dots, M$.

As $k \to \infty$, the local error covariance matrices $P^{kf}_{i,k|k-1}$ have steady state values $P^{kf}_{i,\infty}$ satisfying the algebraic Riccati equations:

$$P^{kf}_{i,\infty} = AP^{kf}_{i,\infty}A^T + \Sigma_w$$
$$- AP^{kf}_{i,\infty}C_i^T(C_iP^{kf}_{i,\infty}C_i^T + \Sigma_{i,v})^{-1}C_iP^{kf}_{i,\infty}A^T, i = 1, \dots, M,$$

and the global error covariance matrix $P^{kf}_{k|k-1}$ has steady state value P^{kf}_∞ that satisfies the algebraic Riccati equation

$$P^{kf}_\infty = AP^{kf}_\infty A^T + \Sigma_w - AP^{kf}_\infty\mathbf{C}^T(\mathbf{C}P^{kf}_\infty\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}P^{kf}_\infty A^T \quad (10)$$

3.2 Quantized filtering scheme

As in the single sensor case, we can consider a suboptimal scheme which are a slightly modified version of the unquantized decentralized Kalman filtering equations. The individual sensors run the following equations, for $i = 1, \dots, M$:

$$\hat{x}_{i,k|k-1} = A\hat{x}_{i,k-1|k-1}$$
$$\hat{x}_{i,k|k} = \hat{x}_{i,k|k-1} + K_{i,k}q_{i,k}(y_{i,k} - C_i\hat{x}_{i,k|k-1})$$
$$K_{i,k} = P_{i,k|k-1}C_i^T(C_iP_{i,k|k-1}C_i^T + \Sigma_{i,v} + \Sigma_{i,n,k})^{-1}$$
$$P_{i,k|k-1} = AP_{i,k-1|k-1}A^T + \Sigma_w$$
$$P_{i,k|k} = P_{i,k|k-1} - P_{i,k|k-1}C_i^T \times (C_iP_{i,k|k-1}C_i^T + \Sigma_{i,v} + \Sigma_{i,n,k})^{-1}C_iP_{i,k|k-1}$$

while the fusion center runs the following equations:

$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1}$$

$$\hat{x}_{k|k} = P_{k|k}\Big(P^{-1}_{k|k-1}\hat{x}_{k|k-1} + \sum_{i=1}^{M}\Big\{P^{-1}_{i,k|k}\hat{x}_{i,k|k} - P^{-1}_{i,k|k-1}\hat{x}_{i,k|k-1}\Big\}\Big)$$

$$P_{k|k-1} = AP_{k-1|k-1}A^T + \Sigma_w$$

$$P_{k|k} = P_{k|k-1} - P_{k|k-1}\mathbf{C}^T(\mathbf{C}P_{k|k-1}\mathbf{C}^T + \Sigma_v + \Sigma_{n,k})^{-1}\mathbf{C}P_{k|k-1}$$

where $\Sigma_{n,k} = \text{diag}(\Sigma_{1,n,k}, \dots, \Sigma_{M,n,k})$ is a block diagonal matrix, and $q_{i,k}(y_{i,k} - C_i\hat{x}_{i,k|k-1})$ is the quantization of $y_{i,k} - C_i\hat{x}_{i,k|k-1}$, with corresponding term $\Sigma_{i,n,k}$ to account for the quantization noise. The values $q_{i,k}(y_{i,k} - C_i\hat{x}_{i,k|k-1})$ are the quantities that are sent to the fusion center. Similar to the previous subsection, the fusion center can reconstruct $\hat{x}_{i,k|k}$, $\hat{x}_{i,k+1|k}$, $P_{i,k|k}$ and $P_{i,k+1|k}$ from $q_{i,k}(y_{i,k} - C_i\hat{x}_{i,k|k-1})$ and knowledge of the sensor parameters.

Call $P_{i,k} \triangleq P_{i,k|k-1}$. We again use the lattice vector quantizers of Moo (1998), with N_i quantizer values for each sensor i. At high rates, assuming that $y_{i,k} - C_i\hat{x}_{i,k|k-1}$ is approximately $N(0, C_iP_{i,k}C_i^T + \Sigma_{i,v})$, we obtain similar to the single sensor case the quantization

$$q_{i,k}(y_{i,k} - C_i\hat{x}_{i,k|k-1}) = U_{i,k}\Lambda_{i,k}^{1/2}\tilde{q}((U_{i,k}\Lambda_{i,k}^{1/2})^{-1}(y_{i,k} - C_i\hat{x}_{i,k|k-1}))$$

where $U_{i,k}$ and $\Lambda_{i,k}$ come from the diagonalization

$$C_iP_{i,k}C_i^T + \Sigma_{i,v} = U_{i,k}\Lambda_{i,k}U_{i,k}^T$$

and $\tilde{q}$ is the lattice quantizer of Moo (1998). We also have

$$\Sigma_{i,n,k} \triangleq \delta_{i,N_i}(C_iP_{i,k}C_i^T + \Sigma_{i,v})$$

where $\delta_{i,N_i} = \frac{M_i(S_{i,0})V_i^{2/m_i}}{\eta^2}\frac{2\ln N_i}{m_iN_i^{2/m_i}}$. In the special case of sensors having scalar measurements, we have

$$\delta_{i,N_i} = \begin{cases} \frac{\pi\sqrt{3}}{2N_i^2}, & \text{optimal quantization} \\ \frac{4\ln N_i}{3N_i^2}, & \text{optimal uniform quantization} \end{cases} \quad (11)$$

3.3 Asymptotic analysis

Let $P_{i,\infty}$ be the steady state value of $P_{i,k}$ that satisfies

$$P_{i,\infty} = AP_{i,\infty}A^T + \Sigma_w - AP_{i,\infty}C_i^T(C_iP_{i,\infty}C_i^T + \Sigma_{i,v} + \Sigma_{i,n})^{-1}C_iP_{i,\infty}A^T$$

and P_∞ be the steady state value of P_k that satisfies

$$P_\infty = AP_\infty A^T + \Sigma_w - AP_\infty\mathbf{C}^T(\mathbf{C}P_\infty\mathbf{C}^T + \Sigma_v + \Sigma_n)^{-1}\mathbf{C}P_\infty A^T. \quad (12)$$

where $\Sigma_{i,n} = \delta_{i,N_i}(C_iP_{i,\infty}C_i^T + \Sigma_{i,v})$ and $\Sigma_n = \text{diag}(\Sigma_{1,n}, \dots, \Sigma_{M,n})$. We now determine the asymptotic behaviour of P_∞ as $N_i \to \infty, \forall i$. Motivated by the asymptotic result for scalar systems in Leong et al. (2012), where it is shown that $P_\infty = P_\infty^{kf} + \sum_{i=1}^M \kappa_i\delta_{i,N_i} + \sum_{i,j} O(\delta_{i,N_i}\delta_{j,N_j})$ with κ_i being constants, we assume that P_∞ takes the form

$$P_\infty = \Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \sum_{i,j} O(\delta_{i,N_i}\delta_{j,N_j}\mathbb{1}) \quad (13)$$

where $\Phi_0, \Phi_{1,i}, i = 1, \dots, M$ are matrices not dependent on N_i. Substituting (13) into (12) we obtain

$$\begin{aligned}\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \cdots &= A(\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \dots)A^T \\ &+ \Sigma_w - A(\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \dots)\mathbf{C}^T \\ &\times (\mathbf{C}(\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \dots)\mathbf{C}^T + \Sigma_v + \Sigma_n)^{-1} \\ &\times \mathbf{C}(\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \dots)A^T\end{aligned} \quad (14)$$

We will need to further simplify (14) before we can solve for Φ_0 and $\Phi_{1,i}, i = 1, \dots, M$. First, from the analysis of the single sensor case in Section 2.3, we have $P_{i,\infty} = P_{i,\infty}^{kf} + O(\delta_{i,N_i}\mathbb{1})$, and hence

$$\begin{aligned}\Sigma_n &= \text{diag}(\Sigma_{1,n}, \dots, \Sigma_{M,n}) \\ &= \text{diag}(\delta_1(N_1)(C_1P_{1,\infty}^{kf}C_1^T + \Sigma_{1,v}) + O(\delta_1^2(N_1)), \dots, \\ &\qquad \delta_M(N_M)(C_MP_{M,\infty}^{kf}C_M^T + \Sigma_{M,v}) + O(\delta_M^2(N_M))) \\ &= \sum_{i=1}^M \delta_{i,N_i}F_i + \sum_{i=1}^M O(\delta_i^2(N_i)\mathbb{1})\end{aligned}$$

where F_i is a block diagonal matrix with i-th diagonal submatrix equal to $C_iP_{i,\infty}^{kf}C_i^T + \Sigma_{i,v}$, and zeros elsewhere.

After some algebraic manipulations and an application of Proposition 1 we can then rewrite (14) as

$$\begin{aligned}\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \cdots &= A(\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \dots)A^T \\ &+ \Sigma_w - A(\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \dots)\mathbf{C}^T[(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1} \\ &- \sum_{i=1}^M \delta_{i,N_i}(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}(\mathbf{C}\Phi_{1,i}\mathbf{C}^T + F_i)(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1} \\ &+ \dots]\mathbf{C}(\Phi_0 + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \dots)A^T\end{aligned} \quad (15)$$

Equating the $O(\mathbb{1})$ terms in (15), we obtain

$$\Phi_0 = A\Phi_0A^T + \Sigma_w - A\Phi_0\mathbf{C}^T(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\Phi_0A^T$$

This is the same equation as (10) satisfied by P_∞^{kf}, thus $\Phi_0 = P_\infty^{kf}$.

Equating the $O(\delta_{i,N_i}\mathbb{1})$ terms in (15), we have for each i:

$$\begin{aligned}\Phi_{1,i} &= A\Phi_{1,i}A^T - A\Phi_{1,i}\mathbf{C}^T(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\Phi_0A^T \\ &- A\Phi_0\mathbf{C}^T(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\Phi_{1,i}A^T + A\Phi_0\mathbf{C}^T \\ &\times(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}(\mathbf{C}\Phi_{1,i}\mathbf{C}^T + F_i)(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\Phi_0A^T \\ &= \left(A - A\Phi_0\mathbf{C}^T(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\right)\Phi_{1,i} \\ &\times \left(A - A\Phi_0\mathbf{C}^T(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\right)^T \\ &+ A\Phi_0\mathbf{C}^T(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}F_i(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\Phi_0A^T\end{aligned} \quad (16)$$

Hence, asymptotically P_∞ behaves like $P_\infty = P_\infty^{kf} + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i} + \sum_{i,j} O(\delta_{i,N_i}\delta_{j,N_j}\mathbb{1})$, where P_∞^{kf} is the unquantized steady state error covariance that can be found numerically by solving the algebraic Riccati equation (10), and $\Phi_{1,i}, i = 1, \dots, M$ can be found numerically by solving the Lyapunov equations (16).

3.4 A rate allocation problem

Suppose we are given R_{tot}, where R_{tot} is large. We want to determine how this total rate is to be allocated amongst the sensors. The rate of each sensor R_i is defined as $R_i = \log_2(N_i)$. One way to allocate the rates is to minimize the trace of the asymptotic expression $P_\infty = P_\infty^{kf} + \sum_{i=1}^M \delta_{i,N_i}\Phi_{1,i}$. We then have for lattice quantization the integer program:

$$\min_{R_1,\dots,R_M\in\mathbb{Z}^+} \text{tr}(P_\infty^{kf}) + \sum_{i=1}^M \frac{e_iR_i}{2^{2R_i/m_i}} \text{ s.t. } \sum_{i=1}^M R_i = R_{tot} \quad (17)$$

where $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ and $e_i \triangleq \frac{M_i(S_{i,0})V_i^{2/m_i}}{\eta^2}\frac{2\ln 2}{m_i}\text{tr}(\Phi_{1,i})$.

In the case of scalar measurements and optimal quantization, the integer program is instead:

$$\min_{R_1,\ldots,R_M \in \mathbb{Z}^+} \operatorname{tr}(P_\infty^{kf}) + \sum_{i=1}^{M} \frac{f_i}{2^{2R_i}} \text{ s.t. } \sum_{i=1}^{M} R_i = R_{tot} \quad (18)$$

where $f_i \triangleq \frac{\pi\sqrt{3}}{2}\operatorname{tr}(\Phi_{1,i})$. An efficient suboptimal solution to problem (18) can also be derived. Let $R_i = \alpha_i R_{tot}$ where $0 \leq \alpha_i \leq 1$, and R_i is not constrained to be integer valued. We then have the problem:

$$\min_{\alpha_1,\ldots,\alpha_M} \operatorname{tr}(P_\infty^{kf}) + \sum_{i=1}^{M} \frac{f_i}{2^{2\alpha_i R_{tot}}}, \text{ s.t. } \sum_{i=1}^{M} \alpha_i = 1, \alpha_i \geq 0 \quad (19)$$

We have the following result:

Lemma 1. The optimization problem (19), where $f_i \geq 0$ are constants, has solution

$$\alpha_i^* = \frac{1}{M} + \frac{1}{2R_{tot}} \log_2 \frac{f_i}{\left(\prod_{j=1}^{M} f_j\right)^{1/M}} \quad (20)$$

Proof The optimal solution follows from analyzing the Karush-Kuhn-Tucker conditions. The derivation is omitted. ■

A suboptimal solution to problem (18) can then be obtained by rounding the solutions obtained from problem (19) to the nearest integer.

Remark: For problem (17), even if we don't constrain the rates to be integer valued, the resulting optimization problem will still be non-convex.

4. NUMERICAL STUDIES

We consider a vector system with parameters $A = \begin{bmatrix} 1.2 & 0.5 \\ 0 & 1.1 \end{bmatrix}$ and $\Sigma_w = I$.

We first consider the case of a single sensor with scalar measurements, and parameters $C_1 = [\,1\ 1\,]$, $\Sigma_{1,v} = 1$. In Fig. 2 we plot the results from Monte Carlo simulations of the trace of the true error covariance $\operatorname{tr}\mathbb{E}[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T]$, together with $\operatorname{tr}(P_\infty)$ and the asymptotic expression for $\operatorname{tr}(P_\infty)$ derived in Section 2.3, for different values of $R_1 = \log_2(N_1)$. We use uniform quantization (similar results can be obtained for optimal quantization but are omitted due to space constraints). We see that the asymptotic expression is very close to the Monte Carlo simulations for rates of 3 bits per sample and above.

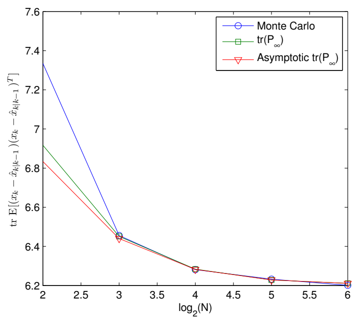

Fig. 2. Error covariance and asymptotic expression: Single sensor

We next add an additional sensor, with parameters $C_2 = [\,1\ 1\,]$, $\Sigma_{2,v} = 0.2$. We consider the rate allocation problem (18) with $R_{tot} = 8$, where we now use optimal quantization. In Table 1 we tabulate the results for some integer combinations of $R_1 = \log_2(N_1)$ and $R_2 = \log_2(N_2)$, with $R_1 + R_2 = 8$. We again present the results from Monte Carlo simulations, together with $\operatorname{tr}(P_\infty)$ and the asymptotic expression for $\operatorname{tr}(P_\infty)$ derived in Section 3.3. We see that in this example $R_1 = 3$, $R_2 = 5$ gives the best performance. Solving the problem (19) gives the solution $\alpha_1^* = 0.3798$, $\alpha_2^* = 0.6202$, corresponding to rates $R_1^* = 3.0386$, $R_2^* = 4.9614$.

Table 1. Error covariance and asymptotic expression: Two sensors

R_1	R_2	Monte Carlo	$\operatorname{tr}(P_\infty)$	Asymptotic $\operatorname{tr}(P_\infty)$
2	6	5.474	5.2213	5.2321
3	5	5.219	5.2119	5.2124
4	4	5.240	5.2290	5.2289
5	3	5.315	5.3136	5.3185
6	2	6.306	5.5996	5.6829

REFERENCES

Anderson, B.D.O. and Moore, J.B. (1979). *Optimal Filtering*. Prentice Hall, New Jersey.

Gersho, A. and Gray, R.M. (1992). *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston.

Gupta, V., Dana, A.F., Murray, R.M., and Hassibi, B. (2006). On the effect of quantization on performance at high rates. In *Proc. American Contr. Conf.*, 1364–1369. Minneapolis, MN.

Hashemipour, H.R., Roy, S., and Laub, A.J. (1988). Decentralized structures for parallel Kalman filtering. *IEEE Trans. Autom. Control*, 33(1), 88–94.

Holmes, M.H. (1995). *Introduction to Perturbation Methods*. Springer-Verlag, New York.

Hui, D. and Neuhoff, D.L. (2001). Asymptotic analysis of optimal fixed-rate uniform scalar quantization. *IEEE Trans. Inf. Theory*, 47(3), 957–977.

Leong, A.S., Dey, S., and Nair, G.N. (2012). On multi-sensor linear state estimation under high rate quantization. In *Proc. ISCCSP*. Rome, Italy.

Max, J. (1960). Quantizing for minimum distortion. *IRE Trans. Inform. Theory*, 6(1), 7–12.

Moo, P.W. (1998). *Asymptotic Analysis of Lattice-Based Quantization*. Ph.D. thesis, University of Michigan.

Msechu, E.J., Roumeliotis, S.I., Ribeiro, A., and Giannakis, G.B. (2008). Decentralized quantized Kalman filtering with scalable communication cost. *IEEE Trans. Signal Process.*, 56(8), 3727–3741.

Nair, G.N. and Evans, R.J. (1998). State estimation under bit-rate constraints. In *Proc. IEEE Conf. Decision and Control*, 251–256. Tampa, FL.

Sukhavasi, R.T. and Hassibi, B. (2011). The Kalman like particle filter: Optimal estimation with quantized innovations/measurements. Submitted for publication.

You, K., Xie, L., Sun, S., and Xiao, W. (2011). Quantized filtering of linear stochastic systems. *Trans. Institute of Measurement and Control*, 33(6), 683–698.