Linear State Estimation Via Multiple Sensors Over Rate-Constrained Channels

Subhrakanti Dey
Joint work with Alex Leong and Girish Nair

THE UNIVERSITY OF MELBOURNE

Department of Electrical and Electronic Engineering
THE UNIVERSITY OF MELBOURNE, AUSTRALIA

THE UNIVERSITY OF MELBOURNE

Outline

- Introduction & Motivation
- Multi-terminal estimation problems
- Single Sensor
- Multiple Sensors
- Numerical Studies
- Remarks and Conclusions

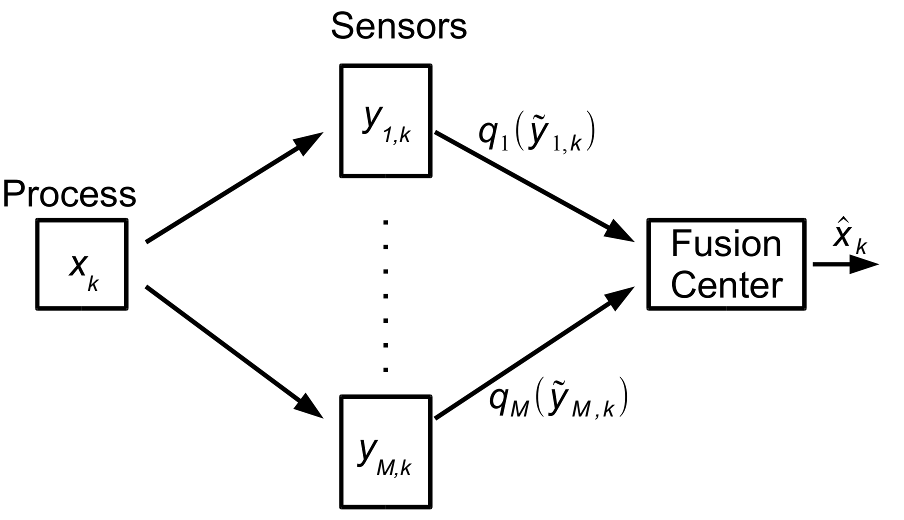
Introduction

- Linear state estimation using multiple sensors is a commonly performed task in e.g. radar tracking, industrial monitoring, remote sensing, wireless control systems, mobile robotics

- Many systems nowadays use digital communications
  - Analog signals need to be quantized
- Wireless channels are bandwidth limited
  - Sensor network applications: severe bandwidth limitations

- Characterize the trade-off between estimation performance and quantization rate (extension of the traditional rate-distortion theory)

Introduction



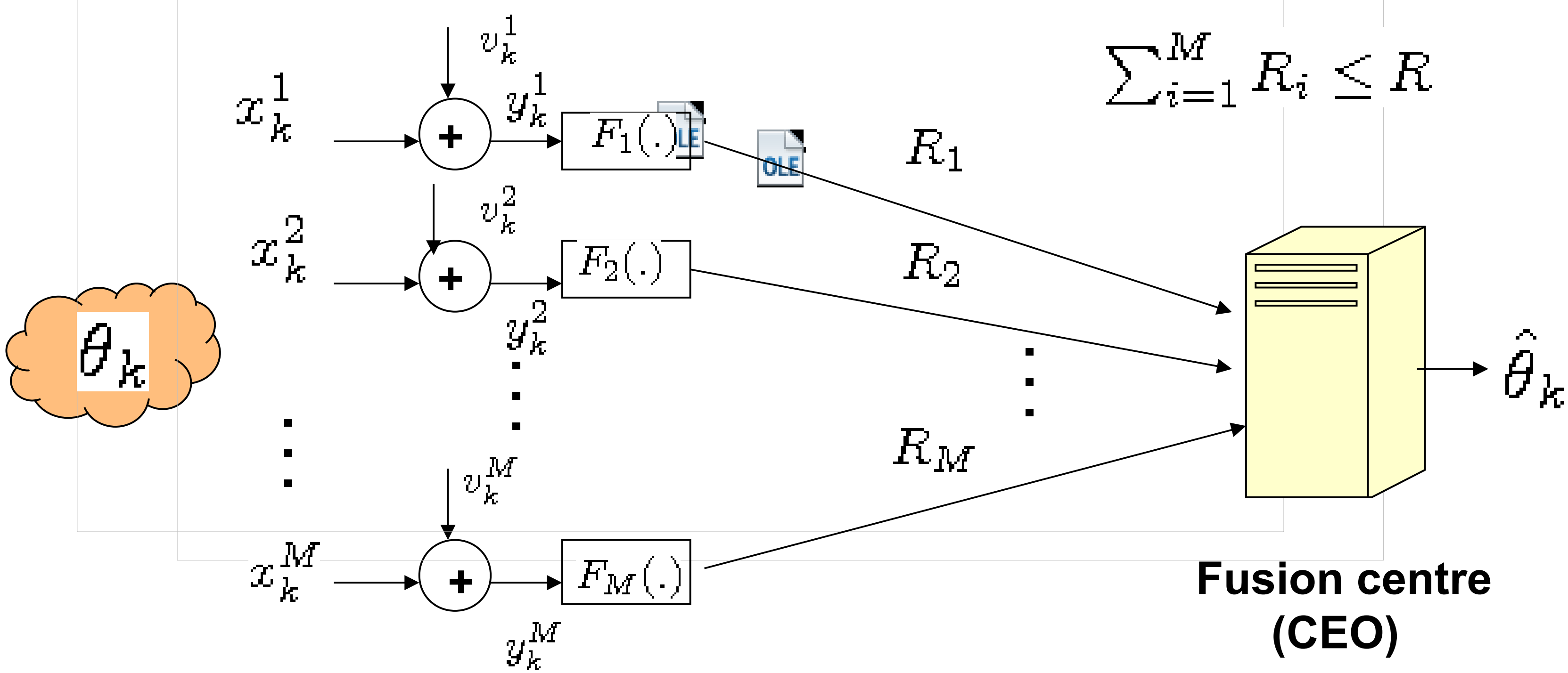
- Estimate a discrete time linear system
- Sensors transmit quantized innovations
  - For unstable systems, states become unbounded while innovations remains of bounded variance
- In our work, we establish a relationship between quantization rate and estimation error for linear dynamical system in a multi-terminal setting, in the case of high rate quantization

Introduction – Related Work

- Similar ideas of quantizing the innovations have been previously considered
  - [Nair&Evans,04] - Single sensor, stable scheme but performance difficult to analyze
  - [Msechu et al. 2008], [You et al. 2011] – Estimator not stable for unstable systems
  - [Sukhavasi and Hassibi, 2011] – Single sensor, particle filter based scheme, performance difficult to analyze
  - [Fu and deSouza, 2009] – Single sensor, logarithmic quantizer, proof of stability for bounded noise
  - Information theoretic multi-terminal estimation: CEO problem

The CEO Problem

- *Simplified single-hop setup:* multiple sensors communicating with a fusion centre over bandwidth constrained channels



The CEO Problem

- Original Results: Viswanathan and Berger [1996] for an i.i.d. scalar Gaussian source
- Rate distortion region: Oohama [1998] for an i.i.d. scalar Gaussian source
- Recent extensions to vector sources and correlated noise across sensors
- Most of these results apply to memoryless sources (at most stationary) and require source coding over asymptotically large block lengths
- Cannot be applied to linear dynamical systems (which have memory and may be unstable) or systems where coding over large numbers of blocks may not be feasible (delay-sensitive applications e.g. wireless control)

Multi-terminal state estimation for linear dynamical systems with rate constraints

- Basic ideas: quantize the innovations (requires smart sensors who can perform their own Kalman filtering) at each sensor
- Apply high rate quantization theory (although in theory this only applies at high rates, performance is quite good at moderate rates (3-4 bits per sample))
- We will study the single sensor case first, followed by multiple sensors
- Difficulty: static quantization may not result in a stable estimate for unstable systems, need to use dynamic quantization
- Assumption: Fusion centre has knowledge of system parameters

Single Sensor

- Vector system $x_{k+1} = Ax_k + w_k$
- Scalar sensor measurement $y_k = Cx_k + v_k$

- Without quantization, optimal estimation given by Kalman filter

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

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

Single Sensor – Quantized Filtering Scheme

- Quantized filtering scheme (at both sensor and fusion centre)

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

- $l_k q(\frac{y_k - c\hat{x}_{k|k-1}}{l_k})$ is quantization of the "innovations"
- l_k is scaling factor for adaptive "zooming" quantizers
  - if quantizer saturates, can "zoom out"
  - used to prove stability for unbounded (Gaussian) noise
- $\Sigma_{n,k}$ is an extra term to account for quantization noise variance

Single Sensor – Quantized Filtering Scheme

- Use shorthand $P_k = P_{k|k-1}$
- Assume $y_k - C\hat{x}_{k|k-1}$ is approximately $N(0, CP_kC^T + \Sigma_v)$
- Can use a uniform quantizer of *N* levels
  - Asymptotically optimal quantizer range and distortion given in [Hui&Neuhoff,2001], can then obtain $\Sigma_{n,k} = \delta_N(CP_kC^T + \Sigma_v)$

    where $\delta_N = \frac{4 \ln N}{3N^2}$
  - Can generalize to lattice vector quantizers
- Can also use an "optimal" Lloyd-Max quantizer of *N* levels (optimal for Gaussian distribution)
  - Can obtain $\Sigma_{n,k} = \delta_N(CP_kC^T + \Sigma_v)$ where $\delta_N = \frac{\pi\sqrt{3}}{2N^2}$
  - Difficult to generalize to vector quantizers (optimal quantizers not known in general)
- Quantizer of [Nair&Evans,04] – Can be used but performance difficult to analyze

Single Sensor - Stability

- Choose

$$l_k = ||C||\tilde{l}_k + d_v$$

$$\tilde{l}_k = ||A(I - K_k C)||\tilde{l}_{k-1} + d_w + ||AK_k||d_v + ||AK_k||(||C||\tilde{l}_{k-1} + d_v)\kappa(\omega_{k-1})$$

with $d_w > 0$ and $d_v > 0$ being constants, and

$$\kappa(\omega_k) = \begin{cases} \frac{2\sqrt{\ln N}}{N} & , \text{ quantizer not saturated} \\ \sqrt{\ln(N)} & , \text{ quantizer saturated} \end{cases}$$

- Define $f_k = x_k - \hat{x}_{k|k-1}$

- Theorem:

$\mathbb{E}[||f_k||^2]$ is bounded $\forall k$ for sufficiently large N.

Single Sensor – Proof of Stability

- Sketch of proof

- Similar to [Nair&Evans,04], consider an upper bound to $\mathbb{E}[||f_k||^2]$ given by

$$||f_k, L||_* \triangleq \sqrt{\mathbb{E}[L^2 + |f_k|^{2+\epsilon} L^{-\epsilon}}$$

  for some random variable L>0 and some $\epsilon > 0$

- Can then show the following Lemma:

$$||X - Lq\left(\frac{X}{L}\right), L\kappa(\Omega)||_* \leq \frac{\zeta}{(\ln N)^{\epsilon/2}} ||X, L||_*$$

  where ζ is a constant that depends only on ϵ and N

Single Sensor – Proof of Stability

- Using the lemma and similar arguments from [Gurt&Nair,09], can then derive the recursive relationship

$$||f_{k+1}, \tilde{l}_{k+1}||_* \leq \left(||A(I - K_k C)|| + ||AK_k||.||C|| \frac{\zeta}{(\ln N)^{\epsilon/2}} \right) ||f_k, \tilde{l}_k||_*$$
$$+ ||w_k, d_w||_* + ||AK_k|| \left(1 + \frac{\zeta}{(\ln N)^{\epsilon/2}} \right) ||v_k, d_v||_*$$

- $||w_k, d_w||_*$ and $||v_k, d_v||_*$ can be upper bounded by constants
- Since $||A(I - K_k C)|| < 1$, and $K_k \to K$, we have

$$\left(||A(I - K_k C)|| + ||AK_k||.||C|| \frac{\zeta}{(\ln N)^{\epsilon/2}} \right) < 1$$

for *N* sufficiently large, which proves that $||f_k, \tilde{l}_k||_*$, and hence $\mathbb{E}[||f_k||^2]$, is bounded for all *k*

Single Sensor – Choice of scaling factors

- Recall

$$l_k = ||C||\tilde{l}_k + d_v$$

$$\tilde{l}_k = ||A(I - K_k C)||\tilde{l}_{k-1} + d_w + ||AK_k||d_v + ||AK_k||(||C||\tilde{l}_{k-1} + d_v)\kappa(\omega_{k-1})$$

- Choice of *dv* and *dw* can affect performance
- If we choose $d_v = \frac{1-||A(I-KC)||-||AK||.||C||\kappa_{min}-||C||d_w}{1-||A(I-KC)||+||AK||.||C||}$

  where *K* is the steady state value of *Kk* and $\kappa_{min} = \frac{2\sqrt{\ln N}}{N}$, then $l_k \approx 1$ for large *N*
  - Reason: For large *N*, quantizer saturation is rare. Choice of *dv* ensures that $l_k \to 1$ when saturation doesn't occur.

Single Sensor – Asymptotic Analysis

- *Pk* is an approximation to the mean squared error
- As $k \to \infty, \quad P_k \to P_\infty$ satisfying

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

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

- Assume high rate quantization (or large *N)* and analyze behaviour of P_∞ with *N*

- Difficulty - no closed form expression for P_∞ in vector systems

Single Sensor – Asymptotic Analysis

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

- Technique used - Extend method for finding asymptotic solutions to algebraic equations in perturbation theory to matrices
- Write P_∞ as $P_\infty = \Phi_0 + \delta_N \Phi_1 + \delta_N^2 \Phi_2 + \ldots$
  where $\Phi_0, \Phi_1, \ldots$ are matrices not dependent on *N*
- Substitute $P_\infty = \Phi_0 + \delta_N \Phi_1 + \delta_N^2 \Phi_2 + \ldots$ into equation above

Single Sensor – Asymptotic Analysis

- Obtain

$$\Phi_0 + \delta_N \Phi_1 + \cdots = A(\Phi_0 + \delta_N \Phi_1 + \dots)A^T + \Sigma_w$$
$$-A(\Phi_0+\delta_N\Phi_1+\dots)C^T(C(\Phi_0+\delta_N\Phi_1+\dots)C^T+\Sigma_v)^{-1}$$
$$\times C(\Phi_0 + \delta_N \Phi_1 + \dots)A^T \frac{1}{1+\delta_N}$$
$$= A(\Phi_0+\delta_N\Phi_1+\dots)A^T+\Sigma_w - A(\Phi_0+\delta_N\Phi_i+\dots)C^T$$
$$\times [(C\Phi_0C^T + \Sigma_v)^{-1} - \delta_N(C\Phi_0C^T + \Sigma_v)^{-1}C\Phi_1C^T(C\Phi_0C^T + \Sigma_v)^{-1} + \dots]$$
$$\times C(\Phi_0 + \delta_N \Phi_1 + \dots)A^T(1 - \delta_N + \dots)$$

- Collect terms of same order to solve for $\Phi_0, \Phi_1, \dots$

Single Sensor – Asymptotic Analysis

- Collecting "constant" terms:
  $$\Phi_0 = A\Phi_0 A^T + \Sigma_w - A\Phi_0 C^T (C\Phi_0 C^T + \Sigma_v)^{-1} C\Phi_0 A^T$$
- Algebraic Riccati equation, can solve for Φ_0
- Same equation as satisfied by P_∞^{kf}, the steady state error covariance in the case of no quantization

- Collecting $O(\delta_N)$ terms:

$$\Phi_1 = \left(A - A\Phi_0 C^T (C\Phi_0 C^T + \Sigma_v)^{-1} C\right) \Phi_1 \left(A - A\Phi_0 C^T (C\Phi_0 C^T + \Sigma_v)^{-1} C\right)^T + A\Phi_0 C^T (C\Phi_0 C^T + \Sigma_v)^{-1} C\Phi_0 A^T$$

- Lyapunov equation, can solve for Φ_1

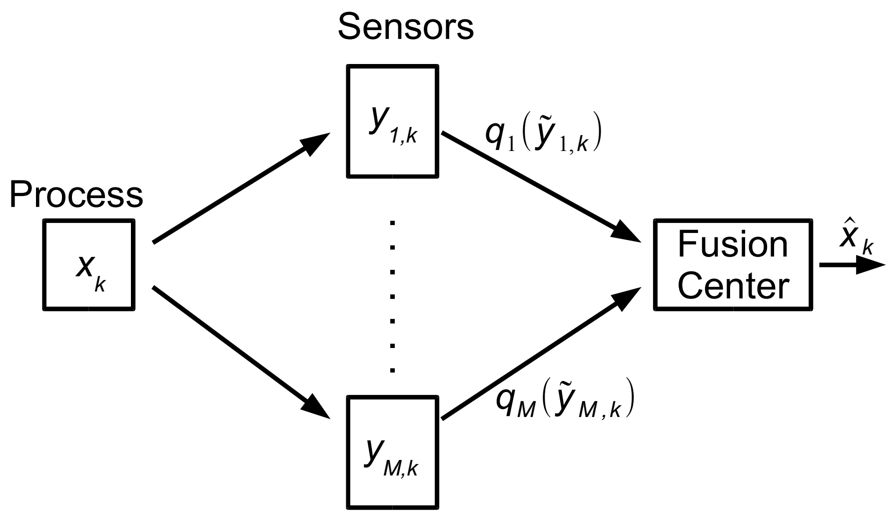
Single Sensor – Asymptotic Analysis

- Therefore

$$P_\infty = P_\infty^{kf} + \delta_N \Phi_1 + \dots$$

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

Multiple Sensors



- Vector system $x_{k+1} = Ax_k + w_k$
- *M* sensors with scalar measurements

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

- Detectability at all sensors assumed (without this, the problem is much harder and currently under investigation)

Multiple Sensors – Decentralized Kalman Filter

- In the case with no quantization, [Hashemipour et al. 1988]
  - Sensors run individual Kalman filters using local information
  - Fusion centre combines local estimates to form global estimate
  - Global estimate same as fusion centre having access to individual sensor measurements

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

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

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

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

- $\hat{x}^{kf}_{i,k|k+1}, \hat{x}^{kf}_{i,k|k}, P^{kf}_{i,k|k+1}, P^{kf}_{i,k|k}$ are local quantities computed at individual sensors
  - Can be reconstructed at fusion centre if sensors send local innovations

Multiple Sensors - Quantized Filtering Scheme

- Modify the scheme of [Hashemipour et al. 1988]
- Individual sensors run:

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

$$\hat{x}_{i,k|k} = \hat{x}_{i,k|k-1} + K_{i,k} l_{i,k} q_{i,k} \left(\frac{y_{i,k} - C_i \hat{x}_{i,k|k-1}}{l_{i,k}} \right)$$

$$K_{i,k} = P_{i,k|k-1} C_i^T (C_i P_{i,k|k-1} C_i^T + \Sigma_{i,v} + \Sigma_{i,n,k})^{-1}$$

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

$$P_{i,k|k} = P_{i,k|k-1} - P_{i,k|k-1} C_i^T (C_i P_{i,k|k-1} C_i^T + \Sigma_{i,v} + \Sigma_{i,n,k})^{-1} C_i P_{i,k|k-1}$$

- Fusion centre runs:

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

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

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

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

Multiple Sensors - Quantized Filtering Scheme

- Sensor *i* uses either asymptotically optimal uniform quantizer of *Ni* quantization levels or "optimal" quantizer of *Ni* quantization levels

- We have $\Sigma_{i,n,k} = \delta_{i,N_i}(C_i P_{i,k} C_i^T + \Sigma_{i,v})$
  where

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

- *li,k* are updated as in single sensor case
- Provided that *Ni* is sufficiently large that the filter is stable when restricted to any single sensor, then stability of the quantized filtering scheme for multiple sensors will also hold.

Multiple Sensors – Asymptotic Analysis

- Study the behaviour of P_∞ as $N_i \to \infty, \forall i$
- From analysis of single sensor case, we have

$$P_{i,\infty} = P_{i,\infty}^{kf} + O\left(\delta_{i,N_i}\right)$$

- Making use of this result and similar techniques to single sensor case, can find that

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

where $\Phi_{1,i}$ satisfy Lyapunov equations

$$\Phi_{1,i} = \left(A - A\Phi_0 \mathbf{C}^T (\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\right) \Phi_{1,i} \left(A - A\Phi_0 \mathbf{C}^T (\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\right)^T$$
$$+ A\Phi_0\mathbf{C}^T(\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1} F_i (\mathbf{C}\Phi_0\mathbf{C}^T + \Sigma_v)^{-1}\mathbf{C}\Phi_0 A^T$$

Multiple Sensors – Rate Allocation

- Want to allocate a total rate R_{tot} amongst the sensors
- Sensor *i* has rate $R_i = \log_2(N_i)$

- One possible formulation is to minimize trace of asymptotic expression for P_∞ subject to

$$\sum_{i=1}^{M} R_i = R_{tot}$$

- Will obtain discrete optimization problems

Multiple Sensors – Rate Allocation

- For uniform quantization, the discrete optimization problem is

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

where $e_i = \frac{4 \ln 2}{3} \operatorname{tr}(\Phi_{1,i})$

- If we relax assumption that *Ri* is integer, have the problem

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

- However, this relaxed problem is still non-convex

Multiple Sensors – Rate Allocation

- For optimal quantization, the discrete optimization problem is

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

where now $e_i = \frac{\pi\sqrt{3}}{2}\mathrm{tr}(\Phi_{1,i})$

- If we relax assumption that *Ri* is integer, have the problem

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

- Lemma: The optimal solution to relaxed problem is

$$\alpha_i^* = \frac{1}{M} + \frac{1}{2R_{tot}} \log_2 \frac{e_i}{\left(\prod_{j=1}^{M} e_j\right)^{1/M}}$$

THE UNIVERSITY OF MELBOURNE

Numerical Studies

- System parameters:

$$A = \begin{bmatrix} 1.2 & 0.5 \\ 0 & 1.1 \end{bmatrix}, \quad \Sigma_w = I$$

- Single sensor case:

$$C = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad \Sigma_v = 1$$

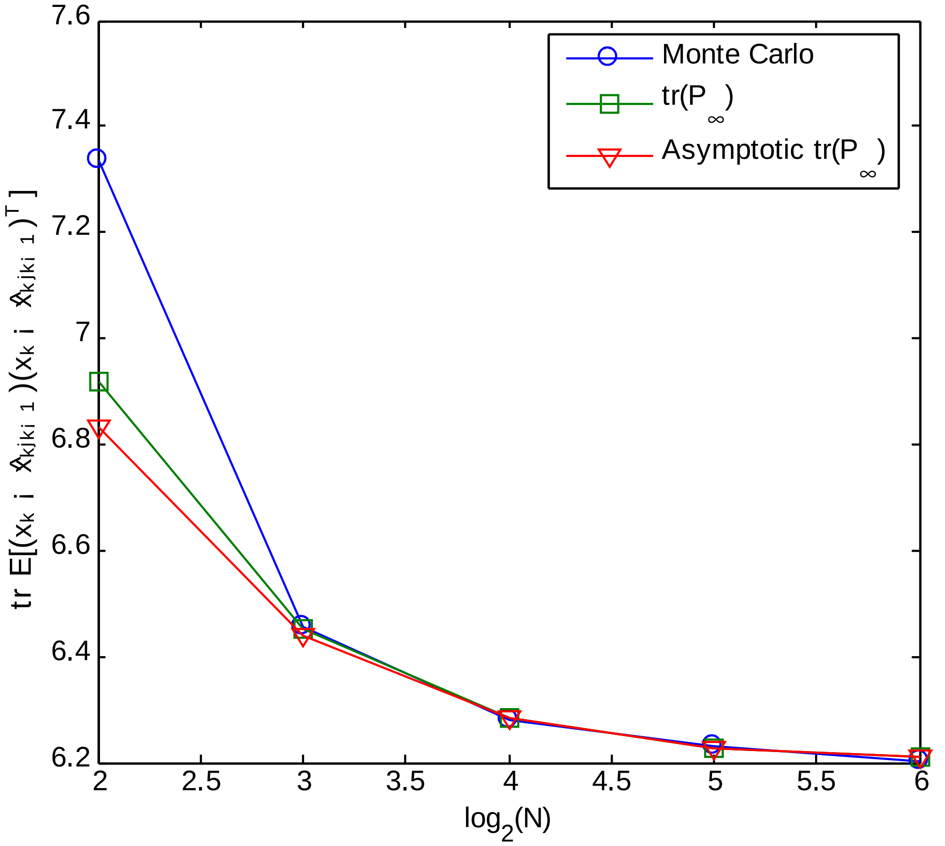
- Two sensors case:

$$C_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad \Sigma_{1,v} = 1$$

$$C_2 = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad \Sigma_{2,v} = 0.2$$

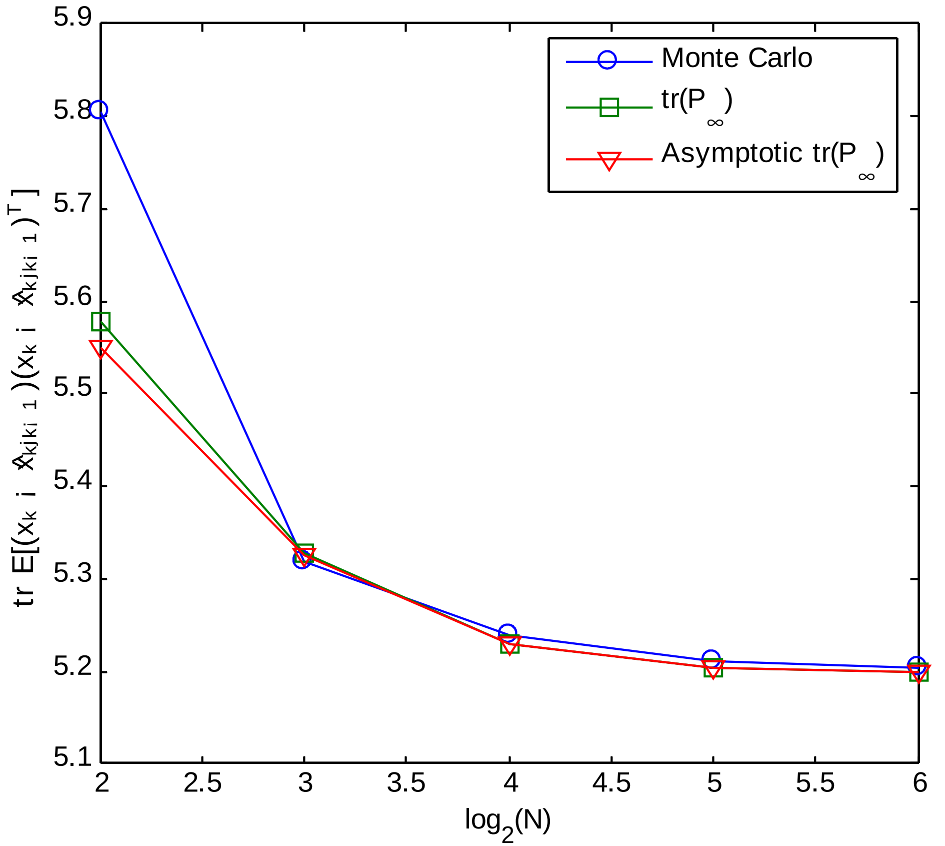
Numerical Studies

- Single sensor, uniform quantizer



Numerical Studies

- Two sensors, optimal quantizer, *N1=N2=N*



Numerical Studies

- Two sensors, uniform quantization
- Rate allocation, $R_{tot} = 8$

R_1	R_2	Monte Carlo	$\text{tr}(P_\infty)$	Asymptotic $\text{tr}(P_\infty)$
2	6	5.242	5.2181	5.2232
3	5	5.237	5.2177	5.2179
4	4	5.247	5.2408	5.2405
5	3	5.322	5.3166	5.3212
6	2	5.585	5.4886	5.5271

Numerical Studies

- Two sensors, optimal quantization
- Rate allocation, $R_{tot} = 8$

R_1	R_2	Monte Carlo	$\mathrm{tr}(P_\infty)$	Asymptotic $\mathrm{tr}(P_\infty)$
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

- Solving the relaxed problem gives $\alpha_1^* = 0.3798, \alpha_2^* = 0.6202$, corresponding to rates $R_1^* = 3.0386, R_2^* = 4.9614$

Conclusions and further work

- Derived asymptotic expression relating estimation error with quantization rates of sensors
- Sketched a proof of stability of the scheme
- Considered a rate allocation problem

- Further areas of investigation
  - Packet loss and high rate quantization
  - Vector measurements: dynamic quantization for lattice vector quantizers
  - Detectability at all sensors a strong assumption
  - Low data rates?
    - Proof of stability here holds for sufficiently high bit rates
    - May need different schemes to achieve stability for lower bit rates
    - Tradeoff between estimation performance and data rate for rates close to minimum bit rates