23rd International Symposium on Mathematical Theory of Networks and Systems
Hong Kong University of Science and Technology, Hong Kong, July 16-20, 2018

Results on Transmission Power Control for Remote State Estimation

Junfeng Wu, Yuzhe Li, Daniel E. Quevedo, Ling Shi

***Abstract*—We consider a sensor transmission power control problem for remote state estimation. In this problem, a sensor sends its local estimate to a remote estimator over a wireless packet-dropping communication channel. The transmission power is determined by a recently proposed algorithm which uses the innovative information contained in the measurement. In the current paper, we focus on parameter optimization arising from the selection of design parameters for this power controller. The existing work obtained a suboptimal solution to the parameter optimization problem, while by using a vector rearrangement inequality argument and the vector majorization, we now show that there exists an optimal solution within a subset of the whole feasible set. By leveraging this property, we obtain an optimal solution via solving a convex optimization problem. Key words: Kalman filtering; transmission power control; packet loss; majorization; convex optimization**

***Index Terms*—Kalman filtering; transmission power control; packet loss; majorization; convex optimization**

I. Notations

$\mathbb{S}_+^n$ is the set of n by n positive semi-definite matrices. For a square matrix X, by abuse of notations, we use $\det(X)$ and X^{-1} in case of a singular matrix X, to denote the pseudo-determinant and the Moore-Penrose pseudo-inverse. For a vector $x \in \mathbb{R}^n$, we use $x^{\downarrow}$ and $x^{\uparrow}$ to represent the vectors with the same entries, but re-ordered in decreasing and increasing order respectively. The symbol $\mathcal{N}(x, \Sigma)$ denotes a Gaussian distribution with mean x and covariance Σ. We introduce an operator $h : \mathbb{S}_+^n \to \mathbb{S}_+^n$, where $h(X) \triangleq AXA' + W,\ W \in \mathbb{S}_+^n$.

II. System Model

Consider a discrete-time linear time-invariant (LTI) system measured by a sensor:

$$x_{k+1} = Ax_k + w_k,$$
$$y_k = Cx_k + v_k,$$

where $A \in \mathbb{R}^{n\times n}$ and $k \in \mathbb{N}$, $x_k \in \mathbb{R}^n$ is the system state, $y_k \in \mathbb{R}^m$ is the sensor's measurement, the state noise $w_k \in \mathbb{R}^n$ and observation noise $v_k \in \mathbb{R}^m$ are zero-mean i.i.d. Gaussian with $\mathbb{E}[w_k w_j'] = \delta_{kj} W$ ($W \geq 0$), $\mathbb{E}[v_k (v_j)'] = \delta_{kj} R$ ($R \succ 0$), $\mathbb{E}[w_k (v_j)'] = 0\ \forall j, k \in \mathbb{N}$. The initial state x_0 is a zero-mean Gaussian random vector, uncorrelated with w_k and v_k. The pair (A, C) is assumed to be detectable and (A, W) stabilizable.

J. Wu is with the College of Control Science and Engineering, Zhejiang University, 310027, Hangzhou, P. R. China Email: jjfwu@zju.edu.cn. Y. Li is Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 1H9. Email: yuzhe2@ualberta.ca. Daniel E. Quevedo is with Department of Electrical Engineering (EIM-E), Paderborn University, Germany. Email: dquevedo@ieee.org. L. Shi is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong Email: eesling@ust.hk.

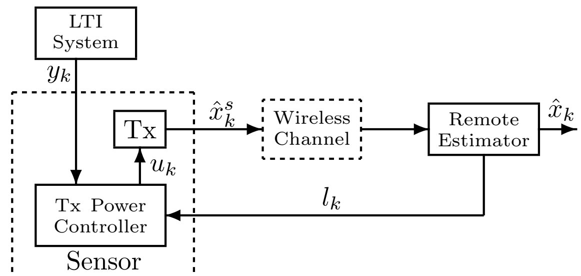

Fig. 1: Remote state estimation scheme.

As shown in Fig. 1, the sensor locally runs a Kalman filter and generates a local MMSE estimate. Then it transmits the local estimate to a remote estimator using power level u_k to be designed. Denote the sensor's local estimate and error covariance by $\hat{x}_k^s$ and P_k^s respectively, i.e., $\hat{x}_k^s \triangleq \mathbb{E}[x_k | y_1, ..., y_k]$ and $P_k^s \triangleq \mathbb{E}[(x_k - \hat{x}_k^s)(x_k - \hat{x}_k^s)' | y_1, ..., y_k]$. We assume that this local Kalman filter has entered steady state, that is, $P_k^s = \overline{P} \geq 0,\ \forall k \in \mathbb{N}$.

The sensor sends data to a remote estimator over an additive white Gaussian noise (AWGN) channel suffering from channel fading [1]. The details and assumptions for the communication channel are provides in [2].

We use a random binary process $\{\gamma_k\}_{k\in\mathbb{N}}$ to describe communication success as follows:

$$\gamma_k = \begin{cases} 1, & \text{if } \hat{x}_k^s \text{ arrives error-free at time } k, \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

initialized with $\gamma_0 = 1$. Let $u_k \in [0, +\infty)$ be the transmission power for the QAM symbol at time k. From [1], the packet loss probability can be approximated by

$$\Pr(\gamma_k = 0 | u_k, h_k) \approx \exp\left(\frac{-\alpha h_k u_k}{N_0 B}\right),$$

where N_0 is the AWGN power spectral density, B is the channel bandwidth, h_k is the channel power gain, and α is a constant depending on the specific modulation scheme used. Throughout this paper we will adopt (1) with equality.

III. Transmission Power Control and Remote State Estimation

We restrict our attention to one type of transmission power controllers that render the estimation problem linear and tractable. The benefit of a linear estimation process is that a closed-form recursive MMSE estimator can be derived. See [2] for the idea of preserving the linearity of the estimation process described below.

Let the incremental innovation contained in the sensor's local estimate compared to the latest reception instant be defined as follows: $z_k = \hat{x}_k^s - A^{\tau_k}\hat{x}_{k-\tau_k}^s$, where $\tau_k \triangleq k - \max_{1\leqslant t\leqslant k-1}\{t : \gamma_t = 1\}$. Consider a transmission power controller $f : \mathbb{R}^n \mapsto [0, \infty)$ of the following form:

$$u_k = f_k(z_k) \triangleq \frac{N_0 B}{2\alpha h_k} z_k' Q_k z_k + \varphi_k, \tag{2}$$

where φ_k is a baseline power level independent of z_k, and $Q_k \in \mathbb{S}_+^n$ and $\varphi_k \geqslant 0$ are two parameters to be designed.

We denote by $\mathcal{I}_k$ the information available to the remote estimator up to time k, i.e.,

$$\mathcal{I}_k = \{\gamma_1\hat{x}_1^s, ..., \gamma_k\hat{x}_k^s\} \cup \{\gamma_1, \dots, \gamma_k\} \cup \{h_1, \dots, h_k\}.$$

The remote estimator generates the MMSE estimate, $\hat{x}_k$, based on $\mathcal{I}_k$, where $\hat{x}_k = \mathbb{E}_{f_{1:k}}[x_k|\mathcal{I}_k]$ (cf., [3]). The corresponding estimation error covariance is defined as $P_k \triangleq \mathbb{E}_{f_{1:k}}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)'|\mathcal{I}_k]$, where expectations are taken under fixed $f_{1:k} \triangleq (f_1, \dots, f_k)$.

Before presenting our main results on how to design the power controller parameters, we will first review some selected results from [2]. Readers are referred to [2] for the proofs.

Lemma 1 ([2, Lemma 4.4]): Consider the transmission power controller (2). The conditional probability $p(z_k; z|\mathcal{I}_{k-1}) \sim \mathcal{N}(0, \Sigma_k)$, where Σ_k is given by the following recursion:

$$\Sigma_k = \begin{cases} A\Psi_{k-1}A' + h(\overline{P}) - \overline{P}, & \text{if } \gamma_{k-1} = 0 \\ h(\overline{P}) - \overline{P}, & \text{if } \gamma_{k-1} = 1, \end{cases}$$

with $\Psi_{k-1} = \begin{cases} 0, & \text{if } k = 1 \\ (Q_{k-1} + {\Sigma_{k-1}}^{-1})^{-1}, & \text{if } k = 2, \dots. \end{cases}$

Moreover, $p(z_k; z|\mathcal{I}_{k-1}, \gamma_k = 0) \sim \mathcal{N}(0, \Psi_k)$.

In general, transmission power controllers depending on sensor measurements lead to complex nonlinear filtering problems. Interestingly, with power controller as in (2), the optimal estimates are easy to find. The following lemma shows how the estimate and the estimation error are explicitly computed.

Lemma 2 ([2, Theorem 4.8]): Consider the transmission power controller (2). The MMSE estimate of x_k is given by

$$\hat{x}_k = \begin{cases} \hat{x}_k^s, & \text{if } \gamma_k = 1, \\ A^{\tau_k}\hat{x}_{k-\tau_k}, & \text{if } \gamma_k = 0. \end{cases} \tag{3}$$

The estimation error covariance at the estimator is given by

$$P_k = \begin{cases} \overline{P}, & \text{if } \gamma_k = 1, \\ \Psi_k + \overline{P}, & \text{if } \gamma_k = 0. \end{cases}$$

The packet loss probability and the expected transmission power can be characterized in terms of h_k, φ_k, Ψ_k and Σ_k as follows.

Lemma 3 ([2, Proposition 4.5]): Consider the transmission power controller (2). The packet loss probability is given by

$$\mathbb{P}(\gamma_k = 0|\mathcal{I}_{k-1}, h_k) = \frac{1}{\sqrt{\det(\Sigma_k)\det(\Psi_k^{-1})}} \exp\left(-\beta h_k \varphi_k\right).$$

The expected transmission power is

$$\mathbb{E}[u_k|\mathcal{I}_{k-1}, h_k] = \frac{1}{2\beta h_k}\left(\mathrm{Tr}(\Sigma_k\Psi_k^{-1}) - n_{\tau_k}\right) + \varphi_k,$$

where n_{τ_k} is the rank of $h^{\tau_k}(\overline{P}) - \overline{P}$, which is known given $\mathcal{I}_{k-1}$.

IV. Main Results

In this section, we design controller parameters Q_k and φ_k which minimize the expected estimation error per time step, subject to transmission energy constraints. We thus focus on the following problem:

Problem 1:

$$\begin{aligned} \underset{Q_k,\varphi_k}{\text{minimize}} \quad & \mathbb{E}[\mathrm{Tr}(P_k)|\mathcal{I}_{k-1}, h_k] \\ \text{subject to} \quad & \mathbb{E}[u_k|\mathcal{I}_{k-1}, h_k] \le \bar{u}, \quad \bar{u} > 0, \end{aligned}$$

where the expectation is taken over binary variable γ_k, and $\bar{u}$ is the energy constraint. To the best of our knowledge, it is rather difficult to optimize the performance over a long horizon under transmission power controller (2). Problem 1 was originally proposed and studied in [2], where only suboptimal solutions were given. By Lemmas 2 and 3, we can break down $\mathbb{E}[\mathrm{Tr}(P_k)|\mathcal{I}_{k-1}, h_k]$ according to $\gamma_k = 0$ or $\gamma_k = 1$, i.e.,

$$\mathbb{E}[\mathrm{Tr}(P_k)|\mathcal{I}_{k-1}, h_k] = \overline{P} + \mathbb{P}(\gamma_k = 0|\mathcal{I}_{k-1}, h_k)\mathrm{Tr}(\Psi_k).$$

Then, we equivalently recast Problem 1 as follows:

Problem 2: Given $\mathcal{I}_{k-1}$ and h_k,

$$\begin{aligned} \underset{\Psi_k,\varphi_k}{\text{minimize}} \quad & \frac{\exp\left(-\beta h_k\varphi_k\right)}{\sqrt{\det(\Sigma_k)\det(\Psi_k^{-1})}}\mathrm{Tr}(\Psi_k), \\ \text{subject to} \quad & \frac{1}{2\beta h_k}\left(\mathrm{Tr}(\Sigma_k\Psi_k^{-1}) - n_{\tau_k}\right) + \varphi_k \le \bar{u}. \end{aligned}$$

Note that in Problem 2, Σ_k is known given $\mathcal{I}_{k-1}$. Problem 2 is difficult to solve since Ψ_k is involved in $\det(\Psi_k^{-1})$, $\mathrm{Tr}(\Psi_k)$ and $\mathrm{Tr}(\Sigma_k\Psi_k^{-1})$ — both the spectra of Ψ_k and of $\Sigma_k\Psi_k^{-1}$ should be taken into account. In [2], Problem 2 was approximated by replacing Ψ_k with ϵI such that $\Psi_k \preceq \epsilon I$. This leads to only suboptimal solutions. To overcome this drawbacks, we will next show that an optimal solution can be computed via solving a convex optimization problem. For that purpose, we first note that since Σ_k is Hermitian, it is unitary diagonalizable: there exist a unitary matrix V_k and a diagonal matrix Θ_k such that

$$\Sigma_k = V_k\Theta_k V_k^*. \tag{4}$$

Denote $\Theta_k \triangleq \mathrm{diag}\big(0,\dots,0,\theta_{k,1},\dots,\theta_{k,n_{\tau_k}}\big)$. Without loss of generality, let the scalars $\theta_{k,1},\dots,\theta_{k,n_{\tau_k}}$ be in nondecreasing order. Since Σ_k is known, V_k and Θ_k are also known matrices. For convenience, let $\Sigma_k^{1/2} \triangleq V_k\sqrt{\Theta_k}$ and define a new matrix

$$\Phi_k \triangleq \left(\Sigma_k^{1/2}\right)^* \Psi_k^{-1}\Sigma_k^{1/2}.$$

Similarly, $(\Psi_k^{-1})^{1/2}$ can be defined. By [2, Lemma 4.1], Φ_k is a block diagonal matrix satisfying

$$\Phi_k \triangleq \begin{bmatrix} 0 & 0 \\ 0 & \Phi_{k,2} \end{bmatrix} \geq \begin{bmatrix} 0 & 0 \\ 0 & I_{n_{\tau_k}\times n_{\tau_k}} \end{bmatrix}.$$

Similarly, there exist a unitary matrix U_k and a diagonal matrix Λ_k such that $\Phi_k = U_k\Lambda_k U_k^*$. Denote $\Lambda_k \triangleq \mathrm{diag}\big(0,\dots,0,\lambda_{k,1},\dots,\lambda_{k,n_{\tau_k}}\big)$, where $\lambda_{k,1},\dots,\lambda_{k,n_{\tau_k}}$ are in nondecreasing order. Then, by [2, Lemma 4.12],

$$\det(\Sigma_k)\det(\Psi_k^{-1}) = \det(\Phi_k) = \prod_{i=1}^{n_{\tau_k}}\lambda_{k,i} \tag{5}$$

and

$$\mathrm{Tr}(\Sigma_k\Psi_k^{-1}) = \mathrm{Tr}(\Phi_k) = \sum_{i=1}^{n_{\tau_k}}\lambda_{k,i}.$$

In addition, $\mathrm{Tr}(\Psi_k) = \mathrm{Tr}(\Sigma_k^{1/2}\Phi_k^{-1}(\Sigma_k^{1/2})^*) = \mathrm{Tr}(\Phi_k^{-1}\Theta_k)$. To derive our results, we require further observations on the structure of Φ_k to simplify Problem 2. This property can be interpreted by majorization, defined as follows:

Definition 1: Let $x \triangleq [x_i], y \triangleq [y_i] \in \mathbb{R}^n$ be given. The vector y is said to majorize x if $\sum_{i=1}^k y_i^\downarrow \geq \sum_{i=1}^k x_i^\downarrow$ for all $k = 1,\dots,n$ with equality for $k = n$, and denote it by $y \succ x$.

Lemma 4 ([4, Lemma 1]): Let H be a Hermitian matrix. the vector of eigenvalues of H majorizes the vector of diagonal entries of H.

Lemma 5 ([5, Rearrangement Inequality]): Let real numbers $x_1 \leq \cdots \leq x_n$ and $y_1 \leq \cdots \leq y_n$ be given. For any permutation $x_{\sigma(1)},\dots,x_{\sigma(n)}$ of $x_1,\dots,x_n$, we have $x_{\sigma(1)}y_1 + \cdots + x_{\sigma(n)}y_n \geq x_ny_1 + \cdots + x_1y_n$ and $x_{\sigma(1)}y_1 + \cdots + x_{\sigma(n)}y_n \leq x_1y_1 + \cdots + x_ny_n$.

For a given vector $y \in \mathbb{R}^n$, define a function: $\phi_y(x) = (x^\downarrow)'y^\uparrow$, where $x \in \mathbb{R}^n$. The function ϕ_y has the following property.

Definition 2: A real-valued function ϕ defined on a set $\mathcal{A} \subset \mathbb{R}^n$ is said to be Schur-cancave if $y \succ x$ on $\mathcal{A}$ implies $\phi(y) \leq \phi(x)$.

Lemma 6 ([4, Lemma 4]): For a given vector $y \in \mathbb{R}^n$, the function ϕ_y is Schur-concave on R^n.

Combining Lemmas 5 and 6, the next technical result follows.

Lemma 7: Let $x \triangleq [x_i], y \triangleq [y_i], z \triangleq [z_i] \in \mathbb{R}^n$ be given vectors. If $z \succ x$, then $x'y \geq (z^\downarrow)'y^\uparrow$.

Proof 1: By Lemma 6, we have $(x^\downarrow)'y^\uparrow = \phi_y(x) \geq \phi_y(z) = (z^\downarrow)'y^\uparrow$. Moreover, by Lemma 5, $x'y \geq (x^\downarrow)'y^\uparrow$, which completes the proof. □

We are in a position to present the following result.

Proposition 1: Suppose the pair $\big(\Psi_k^{\mathrm{opt}},\varphi_k^{\mathrm{opt}}\big)$ is a solution to Problem 2. Denote $\Phi_k^{\mathrm{opt}} \triangleq (\Sigma_k^{1/2})^*(\Psi_k^{\mathrm{opt}})^{-1}\Sigma_k^{1/2}$ and let the eigenvalues (in a nondecreasing order) of Φ_k^{opt} be $\{0,\dots,0,\lambda_{k,1}^{\mathrm{opt}},\dots,\lambda_{k,n_{\tau_k}}^{\mathrm{opt}}\}$. Therefore, the pair $\Big(\Psi_k^\dagger = \Sigma_k^{1/2}(\Lambda_k^{\mathrm{opt}})^{-1}(\Sigma_k^{1/2})^*,\varphi_k^{\mathrm{opt}}\Big)$ is also a solution, where $\Lambda_k^{\mathrm{opt}} \triangleq \mathrm{diag}\big(0,\dots,0,\lambda_{k,1}^{\mathrm{opt}},\dots,\lambda_{k,n_{\tau_k}}^{\mathrm{opt}}\big)$.

Proof 2: Suppose Φ_k^{opt} is not diagonal. There exists a unitary matrix U satisfying $\Phi_k^{\mathrm{opt}} = U\Lambda_k^{\mathrm{opt}}U^*$. Since Φ_k^{opt} is not diagonal, $U \neq I$ and $\lambda_{k,1}^{\mathrm{opt}},\dots,\lambda_{k,n_{\tau_k}}^{\mathrm{opt}}$ are not all equal. By Lemma 4, the vector of the nonzero diagonal entries of $(\Lambda_k^{\mathrm{opt}})^{-1}$ majorizes that of $(\Phi_k^{\mathrm{opt}})^{-1}$. Then $\mathrm{Tr}((\Phi_k^{\mathrm{opt}})^{-1}\Theta_k) \geq \mathrm{Tr}((\Lambda_k^{\mathrm{opt}})^{-1}\Theta_k)$ by Lemma 7. In addition, $\det((\Lambda_k^{\mathrm{opt}})^{-1}) = \det((\Phi_k^{\mathrm{opt}})^{-1})$ and $\mathrm{Tr}((\Lambda_k^{\mathrm{opt}})^{-1}) = \mathrm{Tr}((\Phi_k^{\mathrm{opt}})^{-1})$, which implies the optimality of the pair $\big(\Psi_k^\dagger,\varphi_k^{\mathrm{opt}}\big)$. □

Proposition 1 shows that there exists an optimal Q_k such that Σ_k and Ψ_k are simultaneously diagonalizable by a unitary similarity transformation. Thus, we can restrict our attention to a subset of the positive definite cone for searching Ψ_k. In this subset, Ψ_k is such that Φ_k is a diagonal matrix and $\Phi_k = \mathrm{diag}\big(0,\dots,0,\lambda_{k,1}^{\mathrm{opt}},\dots,\lambda_{k,n_{\tau_k}}^{\mathrm{opt}}\big)$.

Theorem 1: An optimal solution to Problem 2 is given by the following convex optimization problem:

Problem 3:

$$\underset{\substack{\{d_{k,i}\}_{i=1:n_{\tau_k}},\\ \varphi_k}}{\text{minimize}} \ \log\left(\sum_{i=1}^{n_{\tau_k}}\exp\Big(-\frac{1}{2}\sum_{j=1}^{n_{\tau_k}}d_{k,j} - \ln\theta_{k,i}d_{k,i} - \beta h_k\varphi_k\Big)\right),$$

$$\text{subject to } \frac{1}{2\beta h_k}\left(\sum_{i=1}^{n_{\tau_k}}\exp(d_{k,i}) - n_{\tau_k}\right) + \varphi_k \leq \bar{u},$$

$$\varphi_k,\ d_{k,i} \geq 0, \forall i = 1,\dots,n_{\tau_k},$$

where $\theta_{k,i}$'s are the nonzero eigenvalues of Θ_k. Then, $\Psi_k^{\mathrm{opt}} = \Sigma_k^{1/2}D_k(\Sigma_k^{1/2})^*$, where

$$D_k = \Big(\mathrm{diag}\big(0,\dots,0,\exp(d_{k,1})\dots,\exp(d_{k,n_{\tau_k}})\big)\Big)^{-1}.$$

Proof 3: Let $\lambda_{k,i} = \exp(d_{k,i})$. Problem 3 follows from Problem 2, equation (5), Proposition 1 and the fact that logarithm, as a monotonically increasing function, does not change extreme points of a function. In addition, since exponential functions and log-sum-exp functions are convex, Problem 3 is a convex optimization problem, which completes the proof. □

References

[1] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-Hill, 2007.
[2] J. Wu, Y. Li, D. E. Quevedo, V. Lau, and L. Shi, "Data-driven power control for state estimation: A Bayesian inference approach," *Automatica*, vol. 54, pp. 332 – 339, 2015.
[3] B. Anderson and J. Moore, *Optimal Filtering*. Prentice Hall, 1979.
[4] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint tx-rx beamforming design for multicarrier mimo channels: A unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, 2003.
[5] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge University Press, 1952.