

GUILLAUME SAGNOL, RADOSLAV HARMAN

Optimal Designs for Steady-state Kalman filters

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7
D-14195 Berlin-Dahlem

Telefon: 030-84185-0
Telefax: 030-84185-125

e-mail: bibliothek@zib.de
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Optimal Designs for Steady-state Kalman filters

Guillaume Sagnol, Radoslav Harman

Abstract

We consider a stationary discrete-time linear process that can be observed by a finite number of sensors. The experimental design for the observations consists of an allocation of available resources to these sensors. We formalize the problem of selecting a design that maximizes the information matrix of the steady-state of the Kalman filter, with respect to a standard optimality criterion, such as D - or A -optimality. This problem generalizes the optimal experimental design problem for a linear regression model with a finite design space and uncorrelated errors. Finally, we show that under natural assumptions, a steady-state optimal design can be computed by semidefinite programming.

1 Introduction

We consider a stationary discrete-time linear process with a state vector $\mathbf{x}_t \in \mathbb{R}^n$:

$$\mathbf{x}_t = F \mathbf{x}_{t-1} + L \boldsymbol{\nu}_t, \quad (t = 1, 2, \dots) \quad (1)$$

where F is an $n \times n$ transition matrix, L is an $n \times \ell$ noise selection matrix, and $\boldsymbol{\nu}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_\ell)$ is a process noise. In addition, we assume $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \Sigma_0)$. Uncorrelated observations $\mathbf{y}_t^{(1)}, \dots, \mathbf{y}_t^{(s)}$ of the process are available at each time step:

$$\forall i = 1, \dots, s, \quad \mathbf{y}_t^{(i)} = H_i \mathbf{x}_t + \mathbf{v}_t^{(i)}$$

where the i^{th} observation matrix H_i is $r_i \times n$ and the measurements errors satisfy $\mathbf{v}_t^{(i)} \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 \mathbf{I}_{r_i})$. We can group the measurements at time t , which gives a multidimensional observation

$$\mathbf{y}_t = H \mathbf{x}_t + \mathbf{v}_t \quad (2)$$

of size $r = \sum_{i=1}^s r_i$, with $H = [H_1^T, \dots, H_s^T]^T$, and $\mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, R)$ where R is the $r \times r$ block diagonal matrix whose i^{th} diagonal block is $\sigma_i^2 \mathbf{I}_{r_i}$. The random vectors $\{\mathbf{x}_0, \boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_t, \dots, \mathbf{v}_1, \dots, \mathbf{v}_t, \dots\}$ are assumed to be mutually independent.

In this article, we are concerned with the case where the variance σ_i^2 depends on the quantity w_i of resources dedicated to the i^{th} observation. More precisely, we assume that $\sigma_i^2 = \frac{1}{\mu_i(w_i)}$, where μ_i is a nondecreasing, concave and continuous function mapping \mathbb{R}_+ onto itself, and such that $\mu_i(0) = 0$. The interpretation for $w_i = 0$ is that $\sigma_i^2 = +\infty$, meaning that

$\mathbf{y}_t^{(i)}$ is unobserved if no resource is allocated to the i^{th} observation point. The vector $\mathbf{w} = [w_1, \dots, w_s] \in \mathbb{R}_+^s$ will be called a *measurement design*, or simply a *design* for the dynamic process (1)-(2).

The problem studied in this paper is the optimal allocation of resources to the s observation points, when the resources are limited and the design \mathbf{w} must be selected within a compact set $\mathcal{W} \subset \mathbb{R}_+^s$ prior to the beginning of the dynamic process.

The process described by Eq. (1)-(2) contains the natural ingredients to run a Kalman filter, cf. Eq. (4)-(8), which yields at each time t an unbiased estimator $\hat{\mathbf{x}}_t$ of \mathbf{x}_t that is linear with respect to the observations $\mathbf{y}_1, \dots, \mathbf{y}_t$, and with Loewner-minimum covariance matrix in the class of all linear unbiased estimators of \mathbf{x}_t ; see, e.g., [13, Section 5.2]. Under standard assumptions (see Section 2), the information matrix M_t , which is defined as the inverse of the variance-covariance matrix Σ_t of the error $(\hat{\mathbf{x}}_t - \mathbf{x}_t)$, converges to a constant matrix M_∞ . This limit depends only on the design \mathbf{w} (and not on the initial state \mathbf{x}_0 or the measurements $\mathbf{y}_1, \mathbf{y}_2, \dots$), and is the unique positive definite solution X of the discrete algebraic Riccati equation (written here in information form):

$$X = (FX^{-1}F^T + LL^T)^{-1} + M(\mathbf{w}), \quad (3)$$

where $M(\mathbf{w}) := \sum_{i=1}^s \mu_i(w_i) H_i^T H_i$. To stress this dependency, we denote by $M_\infty(\mathbf{w})$ the unique positive definite solution X of (3). A natural approach hence consists in choosing $\mathbf{w} \in \mathcal{W}$ so as to maximize an appropriate scalarization $\Phi(M_\infty(\mathbf{w}))$ of the steady-state information matrix. The main result of this paper (Theorem 3.5) shows that under natural conditions on $\Phi(\cdot)$, $\mu_i(\cdot)$ and \mathcal{W} , this problem can be solved using semidefinite programming (SDP).

The problem of maximizing $\Phi(M_\infty(\mathbf{w}))$ over \mathcal{W} is in fact a generalization of a classical problem which has been extensively studied by statisticians: in the *standard optimal experimental design problem*, the quality of a design \mathbf{w} is measured by a function of the form $\mathbf{w} \rightarrow \Phi(M(\mathbf{w}))$, where $M(\mathbf{w}) = \sum_{i=1}^s w_i H_i^T H_i$. This corresponds to the expression of $M_\infty(\mathbf{w})$ when no information can be gained from the observation of a dynamic process (so “ $LL^T \rightarrow +\infty$ ”), and $\mu_i(w_i) = w_i$ for all i . The approach presented in this paper thus extends the standard optimal design theory to deal with the situation where information can be gained from the knowledge of the system dynamics. We refer the reader to Pukelsheim [10] for a comprehensive monograph on the theory of optimal experimental designs.

Related Work The results presented in this paper answer a question raised in [14] by Singhal and Michailidis, who have considered a problem applicable in telecommunications, where $F = \mathbf{I}_n$ and each H_i has a single nonzero entry per row. The authors propose to use Second-Order Cone Programming to maximize the smallest element of the diagonal steady-state information matrix, i.e., they use the criterion of E -optimality. In contrast, the semidefinite programming approach of the present paper allows one to handle non-diagonal covariance matrices.

Steady-state sensor optimization problems have also been considered elsewhere: for example [5] use a gradient descent to minimize a bound of

the steady-state covariance matrix. Another related article is the sensor scheduling problem studied in [9], where the authors consider a continuous time model $\frac{d\mathbf{x}_t}{dt} = F\mathbf{x}_t + L\mathbf{v}_t$. In [9], the design weights w_i are interpreted as probabilities to activate the i^{th} sensor at time t , and are optimized by semidefinite programming with respect to a specific criterion, which is in fact weighted A -optimality.

Notation Throughout this article, we denote by \mathbb{S}_n ($\mathbb{S}_n^+, \mathbb{S}_n^{++}$) the set of $n \times n$ symmetric (positive semidefinite, positive definite) matrices. The symbol \preceq denotes the Löwner ordering ($A \preceq B \iff B - A \in \mathbb{S}_n^+$), and $A \prec B$ means that $B - A \in \mathbb{S}_n^{++}$.

2 The optimal design problem in a filtering context

Assume (temporarily) that $w_i > 0$ for all $i \in \{1, \dots, s\}$, so that $R < +\infty$ and the Kalman filter equations read (see e.g. [13]):

$$\hat{\mathbf{x}}_{t+1|t} = F\hat{\mathbf{x}}_t, \quad (4)$$

$$\hat{\Sigma}_{t+1|t} = F\Sigma_t F^T + LL^T, \quad (5)$$

$$K_t = \hat{\Sigma}_{t+1|t} H^T (H\hat{\Sigma}_{t+1|t} H^T + R)^{-1}, \quad (6)$$

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{x}}_{t+1|t} + K_t(\mathbf{y}_{t+1} - H\hat{\mathbf{x}}_{t+1|t}), \quad (7)$$

$$\Sigma_{t+1} = (\mathbf{I}_n - K_t H) \hat{\Sigma}_{t+1|t}, \quad (8)$$

where $\hat{\mathbf{x}}_{t+1|t}$ is the *a-priori* estimator of \mathbf{x}_{t+1} based on the observations up to the time t , $\hat{\Sigma}_{t+1|t}$ is the covariance matrix of $\hat{\mathbf{x}}_{t+1|t}$, the matrix K_t is the so-called optimal Kalman gain, $\hat{\mathbf{x}}_{t+1}$ is the *a-posteriori* estimator of \mathbf{x}_{t+1} based on the observations up to the time $t + 1$, and Σ_{t+1} is the covariance matrix of $\hat{\mathbf{x}}_{t+1}$.

Provided that (i) R is positive definite; (ii) the pair (F, L) is controllable [8, Section C3], i.e. $\text{rank}[L, FL, \dots, F^{n-1}L] = n$; and (iii) the pair (F, H) is detectable [8, Section C4], that is, $\text{rank}[F^T - \lambda\mathbf{I}_n, H^T] = n$ for all $\lambda \in \mathbb{C}$ such that $|\lambda| \geq 1$, it is well known that the sequence of covariance matrices of the a-priori estimator of the state $(\hat{\Sigma}_{t+1|t})_{t \in \mathbb{N}}$ converges to a constant matrix Σ_∞^- , that is the unique positive definite solution of the discrete algebraic Riccati equation (DARE), see [13, Section 7.3]. In this article, we work with information matrices rather than with covariance matrices, and so we shall consider an alternative Riccati equation in information form (3), which we call I-DARE. To derive it, note that the correction equation (8) of the Kalman filter is sometimes given under an alternative form, which can be obtained by using the Woodbury matrix identity: $\Sigma_{t+1} = \left(\hat{\Sigma}_{t+1|t}^{-1} + H^T R^{-1} H \right)^{-1}$. This gives a simple update formula for the *information matrix* $M_t := \Sigma_t^{-1}$ of the filter, which

implies the I-DARE, see Eq. (3):

$$M_{t+1} = \hat{\Sigma}_{t+1|t}^{-1} + H^T R^{-1} H = (FM_t^{-1}F^T + LL^T)^{-1} + \underbrace{\sum_{i=1}^s \mu_i(w_i) H_i^T H_i}_{M(\mathbf{w})}.$$

Now, let us remove the assumption that $w_i > 0$ for all $i \in \{1, \dots, s\}$: Recall that $w_i = 0$ means that the sequence $(\mathbf{y}_t^{(i)})$ is unobserved. Hence we define the reduced observation matrix $H_{\mathbf{w}} = [H_{i_1}^T, \dots, H_{i_q}^T]^T$, where $\{i_1, \dots, i_q\} := \{i \in \{1, \dots, s\} : w_i > 0\}$. Similarly, $R_{\mathbf{w}}$ is the block diagonal matrix whose k^{th} diagonal block is $\frac{1}{\mu_{i_k}(w_{i_k})} \mathbf{I}_{r_{i_k}}$. The equations of the Kalman filter are now obtained by substituting $H_{\mathbf{w}}$ for H and $R_{\mathbf{w}}$ for R in Equations (4)–(8). This leaves the I-DARE (3) unchanged, since $H_{\mathbf{w}}^T R_{\mathbf{w}}^{-1} H_{\mathbf{w}} = \sum_{i: w_i > 0} \mu_i(w_i) H_i^T H_i = \sum_{i=1}^s \mu_i(w_i) H_i^T H_i = M(\mathbf{w})$.

Now, for the rest of this article we assume that

- A1. The pair (F, L) is controllable.
- A2. The subset of *detectable designs*, $\mathcal{W}^+ := \{\mathbf{w} \in \mathcal{W} : (F, H_{\mathbf{w}}) \text{ is detectable}\}$ is nonempty.
- A3. The criterion $\Phi : \mathbb{S}_+^n \mapsto \mathbb{R}_+$ is isotonic (i.e., $A \succeq B \implies \Phi(A) \geq \Phi(B)$), continuous, concave, and $\Phi(M) = 0$ if and only if M is singular.

Assumption A3 is satisfied by most common criteria used in optimal design, such as $\Phi_D : M \rightarrow \det M^{\frac{1}{n}}$, $\Phi_E : M \rightarrow \lambda_{\min}(M)$ or $\Phi_A : M \rightarrow n / \text{trace } M^{-1}$, see [10]. An isotonic criterion Φ is said to be *strictly isotonic* if in addition it satisfies $A \neq B, A \succeq B \implies \Phi(A) > \Phi(B)$. For example, Φ_E is isotonic but not strictly isotonic.

Assumptions A1 and A2 ensure that Equation (3) has a unique positive definite solution, which we denote by $M_{\infty}(\mathbf{w})$, for all $\mathbf{w} \in \mathcal{W}^+$:

Lemma 2.1. *Equation (3) has a positive definite solution if and only if the design \mathbf{w} is detectable, i.e. $\mathbf{w} \in \mathcal{W}^+$. Moreover, this solution is unique.*

We omit the proof of this result for the sake of length.¹ The idea is to show that there is a one-to-one correspondence between the positive definite solutions of the standard DARE and its counterpart in information form I-DARE; then we can conclude by using known results on the DARE (see e.g. [13, Theorems 23 and 25]).

So the problem of maximizing $\Phi(M_{\infty}(\mathbf{w}))$ over \mathcal{W}^+ is well defined, and can be rewritten as follows:

$$\begin{aligned} & \sup_{\mathbf{w} \in \mathbb{R}^s, X \in \mathbb{S}_n} \Phi(X) & (9) \\ \text{s. t.} \quad & X = (FX^{-1}F^T + LL^T)^{-1} + \sum_{i=1}^s \mu_i(w_i) H_i^T H_i \\ & X \succ 0, \quad \mathbf{w} \in \mathcal{W}^+. \end{aligned}$$

¹proof at the discretion of the reviewers in Appendix

3 Semidefinite Programming formulation

We next give a series of propositions that basically show that the Riccati equation in (9) may be replaced by a linear matrix inequality (LMI). The proofs of these results will appear in a full-length version of this paper². In fact, Propositions 3.2 and 3.3 are similar to existing results concerning the inequality version of the standard DARE, see e.g. Appendix E in [8]. However, our LMI representation of the closure of the set $\{X \succ 0 : X \preceq (FX^{-1}F^T + LL^T)^{-1} + M(\mathbf{w})\}$, cf. Proposition 3.1, is completely new. Its proof is inspired by the LMI representation of the harmonic mean of two matrices, cf. § 4.1 in [2], and is presented at the end of this Section.

Let us first introduce the sets

$$\mathcal{X}(\mathbf{w}) := \left\{ X \succeq 0 : \exists U \in \mathbb{S}_n : \begin{array}{l} (i) : X = U + M(\mathbf{w}) \\ (ii) : \begin{pmatrix} X - F^T U F & F^T U L \\ L^T U F & \mathbf{I}_\ell - L^T U L \end{pmatrix} \succeq 0 \end{array} \right\}, \quad (10)$$

and

$$\mathcal{X}^+(\mathbf{w}) := \{X \succ 0 : f(X, \mathbf{w}) \succeq 0\},$$

where $f(X, \mathbf{w}) := (FX^{-1}F^T + LL^T)^{-1} + M(\mathbf{w}) - X$.

The first proposition of this series shows the relation between these two sets:

Proposition 3.1. *For all designs $\mathbf{w} \in \mathcal{W}$, we have $\mathcal{X}(\mathbf{w}) \cap \mathbb{S}_n^{++} = \mathcal{X}^+(\mathbf{w})$.*

Then, we shall see that $\mathcal{X}(\mathbf{w})$ is bounded, and hence $\mathcal{X}^+(\mathbf{w})$ is bounded as well:

Proposition 3.2. *For all designs $\mathbf{w} \in \mathcal{W}$, the set $\mathcal{X}(\mathbf{w})$ is bounded. Moreover, there exists a constant $\alpha \geq 0$ such that $X \in \mathcal{X}(\mathbf{w}) \implies \|X\|_2 \leq \alpha(1 + \sum_i \mu_i(w_i) \|H_i\|_2^2)$, where $\|M\|_2$ denotes the spectral norm of M .*

This proposition will be useful to show that $\mathcal{X}^+(\mathbf{w})$ has a maximal element:

Proposition 3.3. *Assume that $\mathcal{X}^+(\mathbf{w})$ is nonempty. Then, there is a matrix $X_{\mathbf{w}}^* \in \mathcal{X}^+(\mathbf{w})$ such that*

$$X \in \mathcal{X}^+(\mathbf{w}) \implies X \preceq X_{\mathbf{w}}^*.$$

Moreover, this maximal element necessarily satisfies $f(X_{\mathbf{w}}^*, \mathbf{w}) = 0$, so that \mathbf{w} is detectable and $X_{\mathbf{w}}^* = M_\infty(\mathbf{w})$.

In consequence, we can deduce equivalent statements for a design \mathbf{w} to be detectable:

Corollary 3.4. *The following statements are equivalent:*

- (i) *The design \mathbf{w} is detectable, i.e. $\mathbf{w} \in \mathcal{W}^+$;*
- (ii) *The I-DARE equation $f(X, \mathbf{w}) = 0$ has a positive definite solution $X \succ 0$;*
- (iii) *The LMI $f(X, \mathbf{w}) \succeq 0$ has a positive definite solution $X \succ 0$;*

²Proofs are presented in Appendix for the sake of the reviewing process

(iv) The set $\mathcal{X}^+(\mathbf{w})$ is nonempty;

(v) There is a pair $(X, U) \in \mathbb{S}_n^{++} \times \mathbb{S}_n$ satisfying the conditions (i) and (ii) of the definition (10).

Proof. The equivalence (i) \iff (ii) follows from Lemma 2.1 and (iii) \iff (iv) \iff (v) is clear from the definitions of $\mathcal{X}^+(\mathbf{w})$ and $\mathcal{X}(\mathbf{w})$ and Proposition 3.1. The implication (ii) \implies (iii) is trivial, and by Proposition 3.3 we have (iv) \implies (i). Hence the corollary is proved. \square

The main result of this article follows. It shows that Problem (9) can be reformulated by using linear matrix inequalities. As a consequence, a solution \mathbf{w} of the steady-state optimal design problem (9) can be computed by semidefinite programming (under natural assumptions on Φ, \mathcal{W} and the functions μ_i , see Remark 3.6):

Theorem 3.5. *Consider the following optimization problem:*

$$\max_{\substack{\mathbf{w} \in \mathbb{R}^s \\ X, U \in \mathbb{S}_n}} \Phi(X) \quad (11a)$$

$$\text{s. t.} \quad \begin{pmatrix} X - F^T U F & F^T U L \\ L^T U F & \mathbf{I}_\ell - L^T U L \end{pmatrix} \succeq 0 \quad (11b)$$

$$X = U + \sum_{i=1}^s \mu_i(w_i) H_i^T H_i \quad (11c)$$

$$X \succeq 0 \quad (11d)$$

$$\mathbf{w} \in \mathcal{W}. \quad (11e)$$

This problem has a solution, i.e. the problem is bounded and the maximum is reached for a triple (\mathbf{w}^, X^*, U^*) . Moreover, \mathbf{w}^* is a solution of the steady-state optimal design problem, $\max\{\Phi(M_\infty(\mathbf{w})) : \mathbf{w} \in \mathcal{W}^+\}$. If in addition Φ is strictly isotonic, then X^* coincides with the optimal steady-state information matrix $M_\infty(\mathbf{w}^+)$.*

Proof. We will prove this theorem in three steps:

1. We observe that the feasibility set of Problem (11) is compact, which guarantees the existence of an optimal solution (\mathbf{w}^*, X^*, U^*) by continuity of Φ . This is a direct consequence of the bound in Proposition 2, together with the compactness of \mathcal{W} and the continuity of the μ_i .
2. We show that this solution necessarily satisfies $X^* \succ 0$. Indeed, by Assumption A2 there exists a detectable design \mathbf{w} , so that $M_\infty(\mathbf{w}) \succ 0$ and we know from Assumption A3 that $\Phi(M_\infty(\mathbf{w})) > 0$. Hence the optimal value of Problem (11) must be positive, which implies that the optimal X^* cannot be singular (Assumption A3).
3. To conclude, observe that Problem (11) can be rewritten as

$$\max_{\mathbf{w} \in \mathcal{W}} \max_{X \in \mathcal{X}(\mathbf{w})} \Phi(X),$$

and by point 2, we can replace $X \in \mathcal{X}(\mathbf{w})$ by $X \in \mathcal{X}^+(\mathbf{w})$, see Proposition 3.1. Moreover by Corollary 3.4 the optimal design \mathbf{w}^* is

necessarily detectable (otherwise the maximization over X goes over the empty set and so it takes the value $-\infty$). Let $X^*(\mathbf{w})$ denote an optimal variable X of the inner problem, for a fixed $\mathbf{w} \in \mathcal{W}^+$. Since Φ preserves the Löwner ordering, the value $\Phi(X^*(\mathbf{w}))$ is necessarily equal to $\Phi(M_\infty(\mathbf{w}))$, because $M_\infty(\mathbf{w})$ is the maximal element of $\mathcal{X}^+(\mathbf{w})$, see Proposition 3.3. If moreover Φ is strictly isotonic, then the optimizer must be the maximal element: $X^*(\mathbf{w}) = M_\infty(\mathbf{w})$. This proves the theorem. \square

Remark 3.6. Assume that \mathcal{W}, Φ and the μ_i ($i \in \{1, \dots, s\}$) are *semidefinite-representable*: a precise definition can be found in [1], but basically it means that the constraint $\mathbf{w} \in \mathcal{W}$ can be replaced by an LMI, as well as constraints of the form $\Phi(M) \geq t$ and $\mu_i(w_i) \geq u_i$. (For example, it is known that the most common criteria Φ_A, Φ_D , and Φ_E are semidefinite representable [3], as well as all Kiefer's Φ_p -criteria for a value of $p \in \mathbb{Q}$ [12]; Concerning the scalar functions μ_i , every concave rational function is semidefinite representable [6].) Then, it is straightforward to reformulate Problem (11) as a semidefinite program (SDP). Note that interfaces such as CVX [4] or PICOS [11] allows one to easily pass Problem (11) to modern interior-point solvers, without further reformulations.

Proof of Proposition 3.1 Let $X \succ 0$. Let $[V_1^T, V_2^T]^T$ be a base of $\text{Ker}([F, L])$, i.e., $FV_1 + LV_2 = 0$. The matrix $[V_1^T, V_2^T]^T$ has full rank by rank-nullity theorem and controllability of (F, L) . So the matrix $(V_1^T V_1 + V_2^T V_2)$ is invertible.

The matrix $\Delta := \begin{pmatrix} FX^{-1} & -L \\ V_1^T & -V_2^T \end{pmatrix}$ is invertible. Indeed we can check by direct calculation that its inverse is $\begin{pmatrix} F^T T & A_1 \\ -L^T T & -A_2 \end{pmatrix}$, where

$$T = (FX^{-1}F^T + LL^T)^{-1},$$

$$A_1 = \left(V_1 - F^T T(FX^{-1}V_1 + LV_2) \right) (V_1^T V_1 + V_2^T V_2)^{-1},$$

$$A_2 = \left(V_2 - L^T T(FX^{-1}V_1 + LV_2) \right) (V_1^T V_1 + V_2^T V_2)^{-1}.$$

So, $\begin{pmatrix} X - F^T U F & F^T U L \\ L^T U F & I_\ell - L^T U L \end{pmatrix} \succeq 0$ if and only if

$$\Delta \begin{pmatrix} X - F^T U F & F^T U L \\ L^T U F & I_\ell - L^T U L \end{pmatrix} \Delta^T \succeq 0 \iff \Delta \begin{pmatrix} X & 0 \\ 0 & I_\ell \end{pmatrix} \Delta^T \succeq \Delta(F, -L)^T U(F, -L) \Delta^T.$$

We can simplify the last expression by using the relation $FV_1 + LV_2 = 0$. This yields a block diagonal LMI with following expressions on the two diagonal blocks:

$$FX^{-1}F^T + LL^T \succeq (FX^{-1}F^T + LL^T)U(FX^{-1}F^T + LL^T); \quad (12)$$

$$V_1^T X V_1 + V_2^T V_2 \succeq 0. \quad (13)$$

The LMI (13) is always satisfied, and LMI (12) reduces to $U \preceq (FX^{-1}F^T + LL^T)^{-1}$ (after pre- and post-multiplication by $(FX^{-1}F^T + LL^T)^{-1}$).

So, we have shown that $(X, U) \in \mathbb{S}_n^{++} \times \mathbb{S}_n$ satisfies the condition (ii) of Definition (10) if and only if $U \preceq (FX^{-1}F^T + LL^T)^{-1}$. The rest of the proof is easy. Let $(X, U) \in \mathbb{S}_n^{++} \times \mathbb{S}_n$ satisfy conditions (i) and (ii) of the definition of $\mathcal{X}(w)$. We have $X \preceq (FX^{-1}F^T + LL^T)^{-1} + M(w)$, that is, $f(X, w) \succeq 0$, and hence $X \in \mathcal{X}^+(w)$. Conversely, assume that $X \succ 0$, $f(X, w) \succeq 0$ and set $U' := (FX^{-1}F^T + LL^T)^{-1}$, $U := X - M(w)$. We know that (X, U') satisfies condition (ii) of (10), and since we have $f(X, w) \succeq 0 \iff U' \succeq U$, the pair (X, U) satisfies both conditions (i) and (ii), that is, $X \in \mathcal{X}(w) \cap \mathbb{S}_n^{++}$.

References

- [1] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2. Society For Industrial Mathematics, 1987.
- [2] R. Bhatia. *Positive definite matrices*. Princeton University Press, 2008.
- [3] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [4] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. <http://cvxr.com/cvx>, October 2010.
- [5] V. Gupta, T.H. Chung, B. Hassibi, and R.M. Murray. On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage. *Automatica*, 42(2):251–260, 2006.
- [6] D. Henrion. Semidefinite representation of convex hulls of rational varieties. *Acta applicandae mathematicae*, 115(3):319–327, 2011.
- [7] J.B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms: Fundamentals*. Springer, 1993.
- [8] T. Kailath, A.H. Sayed, and B. Hassibi. *Linear estimation*. 2000.
- [9] A.I. Mourikis and S.I. Roumeliotis. Optimal sensor scheduling for resource-constrained localization of mobile robot formations. *Robotics, IEEE Transactions on*, 22(5):917–931, 2006.
- [10] F. Pukelsheim. *Optimal Design of Experiments*. Wiley, 1993.
- [11] G. Sagnol. Picos, a python interface to conic optimization solvers. Technical Report 12-48, ZIB, 2012. <http://picos.zib.de>.
- [12] G. Sagnol. On the semidefinite representation of real functions applied to symmetric matrices. *Linear Algebra and its Applications*, 439(10):2829–2843, 2013.
- [13] D. Simon. *Optimal state estimation: Kalman, H_∞ , and nonlinear approaches*. Wiley, 2006.
- [14] H. Singhal and G. Michailidis. Optimal sampling in state space models with applications to network monitoring. In *SIGMETRICS'08*, Annapolis, Maryland, USA, 2008.

A Appendix: Proofs of intermediate results of Section 3

A.1 Proof of Lemma 2.1

Proof. We start to show that there is a one-to-one correspondence between the positive definite solutions of the the discrete algebraic Riccati Equation

$$P = FPF^T - FPH_w^T(H_wPH_w^T + R_w)^{-1}H_wPF^T + LL^T, \quad (14)$$

and of its counterpart in information form

$$M = (FM^{-1}F^T + LL^T)^{-1} + H_w^TR_w^{-1}H_w. \quad (15)$$

Let M be a positive definite solution of (15), and set $P := FM^{-1}F^T + LL^T$. Note that this matrix is positive definite, because it can be written as

$$P = \begin{pmatrix} F & L \end{pmatrix} \begin{pmatrix} M^{-1} & \\ & I_\ell \end{pmatrix} \begin{pmatrix} F & L \end{pmatrix}^T$$

and our controllability assumption (A1) implies that the matrix $[F, L]$ has full row rank. (Indeed, If $C = [L, FL, F^2L, \dots, F^{n-1}L]$ has rank n then CC^T has also rank n , but $CC^T = [F, L]B$, where B is some matrix. Thus, the $n \times (n + \ell)$ -matrix $[F, L]$ must have full rank n .) So we can write $M = P^{-1} + H_w^TR_w^{-1}H_w$. By the Woodbury matrix identity we have $M^{-1} = P - PH_w^T(H_wPH_w^T + R_w)^{-1}H_wP$, and it is now clear that P solves (14).

Conversely, if P is a positive solution of Eq (14), set $M := P^{-1} + H_w^TR_w^{-1}H_w$, which is also a positive definite matrix. As above, we can write $M^{-1} = P - PH_w^T(H_wPH_w^T + R_w)^{-1}H_wP$, and so we have $P = FM^{-1}F^T + LL^T$. Finally, we have $M := P^{-1} + H_w^TR_w^{-1}H_w = (FM^{-1}F^T + LL^T)^{-1} + H_w^TR_w^{-1}H_w$, i.e. M solves (15).

Moreover, the two maps that we have defined above are inverses of each other, between the set of positive definite solutions of the DARE and of the I-DARE. Indeed, the map $M \mapsto FM^{-1}F^T + LL^T$ can be rewritten as $M \mapsto (M - H_w^TR_w^{-1}H_w)^{-1}$ when M is a solution of the I-DARE, which is clearly the inverse mapping of $P \mapsto P^{-1} + H_w^TR_w^{-1}H_w$.

Then, we can apply known results about the former Riccati equation (14): under our assumption (A1) that (F, L) is controllable, there exists a positive semidefinite solution to Eq. (14) if and only if the pair (F, H_w) is detectable, and moreover this solution is unique and positive definite (see e.g. Theorems 23 and 25 in [13]). So the proof of the lemma is complete. \square

A.2 Proof of Proposition 3.2

Let $\mathbf{w} \in \mathcal{W}$, $X \in \mathcal{X}(\mathbf{w})$ and U be a matrix satisfying (i) and (ii) in the definition (10) of $\mathcal{X}(\mathbf{w})$. Define $M(\mathbf{w}) := \sum_{i=1}^s \mu_i(w_i)H_i^TH_i$. In particular,

$$\begin{pmatrix} U + M(\mathbf{w}) - F^TUF & F^TUL \\ L^TUF & I_\ell - L^TUL \end{pmatrix} \succeq 0.$$

Since (F, L) is controllable, there is a gain matrix N such that $F + LN$ has all his eigenvalues outside the unit disc of \mathbb{C} , see [8, Sections C3 and C4], and hence every eigenvalue of $\Gamma = (F + LN)^{-1}$ lies inside the open unit disc. Now, we obtain another positive semidefinite matrix by a congruence operation:

$$\begin{pmatrix} \mathbf{I}_n & -N^T \end{pmatrix} \begin{pmatrix} U + M(\mathbf{w}) - F^T U F & F^T U L \\ L^T U F & \mathbf{I}_\ell - L^T U L \end{pmatrix} \begin{pmatrix} \mathbf{I}_n \\ -N \end{pmatrix} \succeq 0.$$

This gives:

$$U + M(\mathbf{w}) + N^T N - (F + LN)^T U (F + LN) \succeq 0.$$

Pre and post-multiplying by Γ^T (resp. Γ), we obtain

$$U \preceq \Gamma^T U \Gamma + Q_{\mathbf{w}},$$

where $Q_{\mathbf{w}} := \Gamma^T (M(\mathbf{w}) + N^T N) \Gamma \succeq 0$. Denote by U_0 the unique solution of the discrete-time Lyapunov equation $U = \Gamma^T U \Gamma + Q_{\mathbf{w}}$, see [8, Section D1]. We have

$$U - U_0 \preceq \Gamma^T (U - U_0) \Gamma \preceq \dots \preceq (\Gamma^T)^k (U - U_0) \Gamma^k \preceq \dots,$$

and so $U \preceq U_0$ because Γ^k converges to 0 as $k \rightarrow \infty$.

Now, observe that U_0 has a closed-form expression that can be obtained by vectorizing the Lyapunov equation: $\text{vec}(U_0) = (\mathbf{I}_{n^2} - \Gamma^T \otimes \Gamma^T)^{-1} \text{vec}(Q_{\mathbf{w}})$, where \otimes denotes the Kronecker product. Using this relation, we obtain $\|U_0\|_F \leq \|(\mathbf{I}_{n^2} - \Gamma^T \otimes \Gamma^T)^{-1}\|_2 \cdot \|Q_{\mathbf{w}}\|_F$, where $\|M\|_F := \sqrt{\text{trace}(MM^T)}$ denotes the Frobenius norm of M and $\|M\|_2$ is its spectral norm. By using the definition of $Q_{\mathbf{w}}$, we can thus conclude that there exists a constant $\alpha' \geq 0$ such that $\|U\|_F \leq \alpha' (1 + \sum_i \mu_i(w_i) \|H_i\|_F^2)$. Finally, we obtain the bound of the proposition, $\|X\|_2 \leq \alpha (1 + \sum_i \mu_i(w_i) \|H_i\|_2^2)$ for some $\alpha \geq 0$, by using (i) in Eq. (10) and the fact that the Frobenius and spectral norms are equivalent.

A.3 Proof of Proposition 3.3

Let $\mathbf{w} \in \mathcal{W}$. We introduce the function g that maps \mathbb{S}_n^{++} onto \mathbb{S}_n , defined by $g(X) = f(X, \mathbf{w}) = (FX^{-1}F^T + LL^T)^{-1} + M(\mathbf{w}) - X$. Note that $X \in \mathcal{X}^+(\mathbf{w})$ if and only if $g(X) \succeq 0$.

The directional derivative of g in the direction of $\Delta \succeq 0$ can be found by using the formula $\frac{dA^{-1}}{dt} = -A^{-1} \frac{dA}{dt} A^{-1}$ and the chain rule:

$$\begin{aligned} Dg(M)[\Delta] &:= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (g(M + \epsilon\Delta) - g(M)) \\ &= (FM^{-1}F^T + LL^T)^{-1} FM^{-1} \Delta M^{-1} F^T (FM^{-1}F^T + LL^T)^{-1} - \Delta. \end{aligned}$$

Let X be a matrix in $\mathcal{X}^+(\mathbf{w})$, such that $g(X) \neq 0$. We are going to show that there exists a matrix X' in the neighborhood of X satisfying $X' \succeq X$, $X' \neq X$, $g(X') \succeq 0$. This is equivalent to the following first order property (see e.g., [7, Section VI.5.1]):

$$\exists \Delta \succeq 0 : \Delta \neq 0, \quad \forall \mathbf{u} \in \text{Ker}(g(X)), \quad \mathbf{u}^T Dg(X)[\Delta] \mathbf{u} \geq 0.$$

This condition is satisfied for $\Delta := g(X)$ indeed, because the first term of $Dg(X)[\Delta]$ is a positive semidefinite matrix, and the term $-\mathbf{u}^T \Delta \mathbf{u}$ vanishes for all $\mathbf{u} \in \text{Ker}(g(X))$. Hence, for all $X_0 \in \mathcal{X}^+(\mathbf{w})$ we can define a sequence of matrices $X_i \in \mathcal{X}^+(\mathbf{w})$ satisfying $X_0 \preceq X_1 \preceq \dots X_n \preceq \dots$ as follows: $X_{n+1} = \Psi(X_n)$, where

$$\Psi(X) = X + \left(\operatorname{argmax}_{t>0} \left\{ t : g(X + tg(X)) \succeq 0 \right\} \right) g(X).$$

The sequence (X_i) is bounded (by Propositions 3.1 and 3.2) and hence it converges. (Indeed, if it has two accumulation points X_1^* and X_2^* , then for all $\mathbf{u} \in \mathbb{R}^n$ the nondecreasing sequence $\mathbf{u}^T X_i \mathbf{u}$ converges to some value, which must be equal to both $\mathbf{u}^T X_1^* \mathbf{u}$ and $\mathbf{u}^T X_2^* \mathbf{u}$. Hence, $\mathbf{u}^T (X_1^* - X_2^*) \mathbf{u} = 0$ for all \mathbf{u} , which proves $X_1^* = X_2^*$.) Denote this limit by X^* . The above discussion shows that $g(X^*) = 0$, otherwise X^* cannot be a fixed point of Ψ . This means that X^* is the unique positive definite solution of the I-DARE (3). We have thus $X_0 \preceq X^* = M_\infty(\mathbf{w})$, and the proposition is proved.