2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

Optimal Sensor Design for Infinite-Time Kalman Filters

Shu-Xia Tang and Kirsten A. Morris

Abstract— **This paper is concerned with state estimation for systems governed by partial differential equations. Kalman filters are optimal state estimators in that they minimize the estimation error variance for given measurements. The focus of this paper is the achievement of additional minimization of the error variance by also optimizing over the sensor design. The optimal sensor design problem is thus incorporated into the estimation problem. Not only the sensor location but also other factors such as sensor shape and the effect of the sensors on system dynamics are included in the optimization criteria. The problem is first stated formally, and then it is shown to be well-posed and to possess an optimal solution. Applications to a one-dimensional diffusion equation and also to a two-dimensional wave equation are given. A computational framework for calculation of optimal shape is described.**

Index Terms— **distributed parameter systems, infinite-dimensional systems, optimal sensor design, optimal estimation, Kalman filter,**

I. INTRODUCTION

Estimator design arises in many applications. For estimation of distributed parameter systems, such as flexible structures and acoustic noise, there is choice of both the types and locations of sensors used in estimation. The development of smart materials [1] means that shape design is also possible. It is now understood from various examples that the performance of an estimator is affected by the number, locations and distribution of the sensors; see for example, [9], [11], [14]. Optimal sensor placement has been considered for decades, the first theoretical analysis is possibly [3]. See also [7] for optimal damping design in second order systems; [6]; and the review articles [8], [12], [21].

One family of approaches for sensor design optimization considers open-loop measures of performance. Often the observability of the system, quantified by a cost function related to the observability grammian, is maximized. The dual problem is actuator design optimization to maximize controllability, and there are several shortcomings with using the open-loop metrics in either problem. One drawback is that calculation relies on finite-dimensional approximations which in many cases are not convergent. First, a spillover-type phenomenon can occur where the best location for one approximation is very poor for a higher order approximation. See [18] for an example of spillover with the wave equation. The same issue is examined in [23] for the dual problem of maximizing controllability in a structure by actuator placement. This can be addressed by restricting the observation and optimizing over a finite-dimensional or other compact space; see for example [19]. Another major disadvantage with optimizing an open-loop criterion such as observability is that this approach does not in general leads to the best estimator performance. For instance, in the dual case of actuator location optimization, even if a finite-dimensional approximation of the partial differential equations (PDEs) is regarded as the true model, the locations calculated using the criterion of optimal controllability are not generally those with the best closed-loop control performance; see [23].

In recent years, results for concurrent optimal controller design and actuator location have been obtained. Conditions for well-posedness of the optimal linear-quadratic problem, along with conditions for using finite-dimensional approximations in computation are given in [15]. A numerical scheme for finding the global minimum of simultaneous optimal linear-quadratic (LQ) controller design and actuator location is described in [5]. Similar conditions have been obtained for H_2 control/actuator location [16] and also for the H_∞-problem [10]. These results were extended to general actuator design with quadratic performance in [17].

If a criterion such as minimizing the error covariance is used to design the estimator, using the same criterion in sensor design leads to consistent design optimization. The well-known Kalman filter for estimation minimizes the covariance between the estimated and true state. In [22] estimator design was considered on a finite-time interval with the sensors optimized over a set of possible locations. This result was generalized to infinite time in [24]. A computational framework for calculation of optimal sensor locations is given. The paper [24] also considers the issue of the number of sensors and their quality.

This paper considers concurrent optimal estimator and sensor design. Here, the approach in [22], [24] is extended to consider minimum error estimation where the form of the observation operator is not assumed to be fixed. That is, the shape as well as the location and number of the sensors is variable. Another generalization is that effect of the sensors on the system mass and dynamics is included in the problem formulation. The framework is illustrated with several examples.

II. PROBLEM FORMULATION

For the purpose of optimal state estimation the optimal sensor design problem in this paper includes not only optimization of the sensor location, but also optimization of other factors such as the sensor shape. Let Ω^{o} indicate the set of possible sensor designs and let $r \in \Omega^{o}$ a particular design. In the case of sensor location, Ω^{o} is the set of possible

K. A. Morris and S.-X. Tang are with the Department of Applied Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada s74tang@uwaterloo.ca; kmorris@uwaterloo.ca.

The financial support of AFOSR for this research is gratefully acknowledged.

locations. In the more general case, it is a set of functions that represent the allowable shapes of the sensors, as well as, indirectly, their number and locations. This leads to a family of observation systems indexed by the parameter $r \in \Omega^o$:

$$\dot{z}(t) = A(r)z(t) + Gw(t), \quad t \geq 0, \tag{1}$$
$$z(0) = z_0, \tag{2}$$
$$y(t) = C(r)z(t) + Hv(t), \quad t \geq 0, \tag{3}$$

where $z(t)$ is the system state variable. The process noise $w(t)$ and the measurement noise $v(t)$ are uncorrelated and are both assumed to be Gaussian white noise, with covariances $Q \geq 0$ and $R > 0$ respectively. It thus holds that

$$E[w(t)w(\tau)^T] = Q\delta(t-\tau), \tag{4}$$
$$E[v(t)v(\tau)^T] = R\delta(t-\tau), \tag{5}$$
$$E[w(t)v(\tau)^T] = 0, \tag{6}$$

where δ is the Dirac delta distribution. Moreover, the initial condition z_0 is assumed to be uncorrelated to both w and v. The initial state z_0 is a $\mathcal{Z}$-valued Gaussian random variable, with zero mean value and covariance P_0. This implies that the nuclear norm

$$||P_0||_1 = E\{||z_0||^2\} < \infty$$

[4, Definition 5.2]. The state space $\mathcal{Z}$ is a separable Hilbert space; the disturbance spaces $\mathcal{W}$ and $\mathcal{V}$ and the observation space $\mathcal{Y}$ are finite-dimensional Hilbert spaces.

For any parameter $r \in \Omega^o$,

(A1) **(i).** $A(r) : \mathcal{D}(A(r)) \subset \mathcal{Z} \to \mathcal{Z}$ is the infinitesimal generator of a strongly continuous (C_0) semigroup in $\mathcal{Z}$, denoted by $e^{tA(r)}$;
(ii). there exist constants $\omega_0 \in \mathbb{R}$ and $c_0 \geq 1$, independent of r, such that for all $r \in \Omega^o$ and $t \geq 0$, $\|e^{tA(r)}\| \leq c_0 e^{\omega_0 t}$;

(A2) $G \in \mathcal{L}(\mathcal{W}, \mathcal{Z}), C(r) \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$.

Note that the effect of the sensors on the dynamics of the system is included in the model. Not only does the output operator C depend on the sensor design r, but the system operator A and its domain may also depend on r.

Consider a general infinite-dimensional filter [4], [20]:

$$\dot{\hat{z}}(t) = A(r)\hat{z}(t) + K(t;r)(y(t) - C(r)\hat{z}(t)), \quad t \geq 0, \tag{7}$$
$$\hat{z}(0) = 0, \tag{8}$$

where $K(t;r) \in L^2(0,\infty;\mathcal{U})$ is the filter gain to be determined. For any finite $t \geq 0$, based on the noisy measured output $\{y(t); 0 \leq s \leq t\}$ in (3), the objective here is to choose the sensor design and filter $K(t,r)$ that optimizes a particular criterion. For a fixed observer design r, the Kalman filter minimizes the estimation error variance.

Theorem 1 *[4] Consider a fixed $r \in \Omega^o$. Let $[0, t_f]$ be a finite time interval and the operator $\Pi(t;r) \in Z)$ the unique, self-adjoint, non-negative solution the following differential Riccati equation (DRE):*

$$< \dot{\Pi}(t;r)h_1, h_2 >=< [A(r)\Pi(t;r) + \Pi(t;r)A^*(r) - \Pi(t;r)C^*(r)R^{-1}C(r)\Pi(t;r) + GQG^*]h_1, h_2 >_{\mathcal{Z}}, \quad \forall h_1, h_2 \in D(A^*), t \in (0, t_f), \tag{9}$$
$$\Pi(0;r) = \Pi_0(r). \tag{10}$$

Then, for the estimate $\hat{z}_*(t)$ obtained from (7)–(8) *with the filtering gain*

$$K(t;r) = \Pi(t;r)C(r)^*R^{-1}, \tag{11}$$

where $\Pi(t;r)$ is the solution to (9)–(10)*, it holds that the estimation error covariance*

$$E\{(z(t) - \hat{z}_*(t))(z(t) - \hat{z}_*(t))^T\} = \Pi(t;r), \tag{12}$$

where the true value $z(t)$ is the solution to (1)–(2) *with the initial datum z_0. Moreover, $\hat{z}_*(t)$ is the optimal estimate for $z(t)$ in the sense that for each $h \in \mathcal{Z}$*

$$E\{< z(t) - \hat{z}_*(t), h >_Z^2\} = \min_{\hat{z}} E\{< z(t) - \hat{z}(t), h >_Z^2\},$$

where the minimum is taken over all estimates $\hat{z}(t)$ of the form (7)–(8)*, , with $K(t;r)$ strongly continuous in time on $[0, t_f]$. Also,*

$$\|\Pi(t;r)\|_1 = E\{\|z(t) - \hat{z}_*(t)\|_{\mathcal{Z}}^2\} = \min_{\hat{z}} E\{\|z(t) - \hat{z}(t)\|_{\mathcal{Z}}^2\}, \tag{13}$$

where $\|\cdot\|_1$ indicates the nuclear norm.

As time approaches infinity, under certain conditions, the finite-time Kalman filter converges to a time-invariant filter, called the steady-state Kalman filter:

$$\dot{\hat{z}}(t) = A(r)\hat{z}(t) + K(r)(y(t) - C(r)\hat{z}(t)), \ t \geq 0, \tag{14}$$
$$\hat{z}(0) = \hat{z}_0. \tag{15}$$

Definition 1 *A C_0-semigroup $T(t)$ on a Hilbert space $\mathcal{X}$ is* exponentially stable *if there exist positive constants c and ω such that*

$$\|T(t)\|_{\mathcal{X}} \leq ce^{-\omega t}, \quad t \geq 0.$$

A family of C_0-semigroups $\{T_r(t)\}$, $r \in \Omega^o$ on $\mathcal{X}$ is uniformly exponentially stable *if there exist positive constants c and ω, independent of r, such that*

$$\|T_n(t)\|_{\mathcal{X}} \leq ce^{-\omega t}, \quad t \geq 0, \forall r \in \Omega^o.$$

Definition 2 *For a fixed $r \in \Omega^o$, the pair $\left(A(r), G(r)Q^{1/2}\right)$ is* exponentially stabilizable *if there exists an operator $K(r) \in \mathcal{L}(\mathcal{Z}, \mathcal{W})$ such that $A(r) - G(r)Q^{1/2}K(r)$ generates an exponentially stable semigroup; that is, for this operator $K(r)$, there exist positive constants c and ω such that*

$$\|e^{t(A(r)-G(r)Q^{1/2}K(r))}\|_{\mathcal{L}(\mathcal{Z})} \leq ce^{-\omega t}, \quad t \geq 0.$$

The family of operators $\left(A(r), G(r)Q^{1/2}\right)$ is uniformly exponentially stabilizable *with respect to $r \in \Omega^o$ if the constants c and ω are independent of r and if the operator $K(r)$ is uniformly bounded with respect to r.*

Definition 3 *For a fixed $r \in \Omega^o$, the pair $\big(A(r), C(r)\big)$ is* exponentially detectable *if there exists an operator $L(r) \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$ such that $A(r) - L(r)C(r)$ generates an exponentially stable C_0-semigroup; that is, for this operator $L(r)$, there exist positive constants c and ω such that*

$$\|e^{t(A(r)-L(r)C(r))}\|_{\mathcal{L}(\mathcal{Z})} \le ce^{-\omega t}, \quad t \ge 0.$$

The family of operators $\big(A(r), C(r)\big)$ is uniformly exponentially detectable *with respect to $r \in \Omega^o$ if the constants c and ω are independent of r and if the operator $L(r)$ is uniformly bounded with respect to r.*

Theorem 2 *[24, Theorem 2.6] Assume that for a fixed $r \in \Omega^o$, $(A(r), GQ^{1/2})$ is exponentially stabilizable and $(A(r), C(r))$ is exponentially detectable. If both the spaces $\mathcal{W}$ and $\mathcal{Y}$ are finite-dimensional, then there exists an optimal filter gain $K(r) \in L^2(\mathcal{U})$ and a corresponding optimal estimate $\hat{z}$ such that*

$$\lim_{t\to\infty} E\{\|z(t;r) - \hat{z}(t;r)\|^2\} = \min_{\check{z}} \lim_{t\to\infty} E\{\|z(t;r) - \check{z}(t;r)\|^2\}, \tag{16}$$

where $z(t;r)$ is the solution to (1)–(2) *with the initial datum z_0, and $\hat{z}(t;r)$ is the solution to* (14)–(15) *with the initial datum $\hat{z}_0$ and the optimal filter gain*

$$K(r) = \Pi_{ss}(r)C(r)^*R^{-1}, \tag{17}$$

where the self-adjoint, non-negative operator $\Pi_{ss}(r) \in \mathcal{L}(\mathcal{Z})$ is the unique non-negative solution the following algebraic Riccati equation (ARE):

$$< [A(r)\Pi_{ss}(r) + \Pi_{ss}(r)A^*(r) - \Pi_{ss}(r)C^*(r)R^{-1}C(r)\Pi_{ss}(r) + GQG^*]h_1, h_2 >_{\mathcal{Z}} = 0,\ \forall h_1, h_2 \in D(A^*(r)). \tag{18}$$

The operator $\Pi_{ss}(r)$ is nuclear and

$$\lim_{t\to\infty} \|\Pi(t;r) - \Pi_{ss}(r)\|_1 = 0. \tag{19}$$

Moreover, $A(r) - \Pi_{ss}(r)C(r)^*R^{-1}C(r)$ generates an exponentially stable C_0-semigroup, with the optimal estimation error variance

$$\|\Pi_{ss}(r)\|_1 = \lim_{t\to\infty} E\{\|z(t;r) - \hat{z}(t;r)\|^2\}. \tag{20}$$

For simplification, the subscript "ss" is deleted from $\Pi_{ss}(\cdot)$ in the sequel, and equation (18) is written

$$A(r)\Pi(r) + \Pi(r)A^*(r) - \Pi(r)C^*(r)R^{-1}C(r)\Pi(r) + GQG^* = 0, \tag{21}$$

with the understanding that it holds on $D(A^*(r))$.

III. Existence of an Optimal Sensor Design

The sensor design is now taken into consideration as part of minimizing the estimation error covariance. Let the set Ω^o of admissible designs be endowed with a topology, with respect to which the convergence of sequences in Ω^o and compactness are considered.

The following theorem is the main result of this paper. It provides sufficient conditions for the existence of an optimal sensor design. The proof can be found in [17].

Theorem 3 *Assume that **(A1)**, **(A2)** hold and that*

(A3) *the set Ω^o is sequentially compact;*

(A4) *the pairs of functions $\big(A(r), GQ^{1/2}\big)$ are uniformly exponentially stabilizable with respect to $r \in \Omega^o$;*

(A5) *the pairs of functions $\big(A(r), C(r)\big)$ are exponentially detectable for at least one $r \in \Omega^o$;*

(A6) *for any $r \in \Omega^o$ and for any sequence $\{r_n\} \subset \Omega^o$ that converges to $r \in \Omega^o$,*

$$\lim_{n\to\infty} \|C(r_n) - C(r)\|_{Z,\mathcal{Y})} = 0,$$

and moreover, for each $z \in \mathcal{Z}$,

(i) $\lim_{n\to\infty} \|e^{tA(r_n)}z - e^{tA(r)}z\|_{\mathcal{Z}} = 0,$

(ii) $\lim_{n\to\infty} \|e^{tA(r_n)^*}z - e^{tA(r)^*}z\|_{\mathcal{Z}} = 0$

uniformly in t on bounded intervals of time.

Then there exists an optimal sensor design $\hat{r} \in \Omega^o$ for the system (1)–(3) *such that*

$$\|\Pi(\hat{r})\|_1 = \min_{r\in\Omega^o} \|\Pi(r)\|_1, \tag{22}$$

where $\Pi(r)$ is the solution to (21) *for each $r \in \Omega^o$.*

Example: Diffusion equation

Consider the following one-dimensional diffusion equation

$$\begin{cases} z_t(t,x) = z_{xx}(t,x) + g(x)w(t), (t,x) \in (0,\infty)\times(0,1), \\ z(t,0) = z(t,1) = 0, t \in [0,\infty), \\ z(0,x) = z_0(x), x \in [0,1], \\ y(t) = \int_0^1 r(x)z(t,x)dx + v(t). \end{cases} \tag{23}$$

This equation can be used to describe the thermal diffusion along a rod of length 1 with zero temperature at the two ends. The state z denotes the temperature subject to a process noise $w(t)$ with the spatial distribution $g(x)$; and the output $y \in \mathbb{R}$ denotes the corresponding available measurement from one sensor subject to a sensor noise $v(t)$. The disturbances w and v are real-valued white Gaussian noises, with variances Q and R respectively. The initial condition z_0 and the noises $w(t), v(t)$ are assumed to be mutually uncorrelated. The objective is to find the optimal sensor shape described by $r(x)$ over the spatial interval $[0,1]$, which together with the corresponding infinite-time Kalman filter achieves minimum estimation error covariance.

Set

$$\Omega^o := \Big\{r(x) \in BV(0,1); |r(x)| \le \bar{r},\ TV(r) \le M\Big\},$$

where $BV(0,1)$ denotes the set of functions of bounded variations on $(0,1)$, $\bar{r}, M$ are positive constants and $TV(r)$ denotes the total variation of r. From Helly's selection theorem in Banach spaces [2, Theorem 1.126], the set Ω^o is sequentially compact in the following sense: for any sequence in Ω^o, there is a subsequence that converges pointwise for $x \in (0,1)$ to a function in Ω^o. Assumption **(A3)** is thus satisfied.

Let the state space $\mathcal{Z} = L^2(0,1)$, and define the operator $A : D(A) \to \mathcal{Z}$ as

$$Af = f'',$$

with the domain

$$\mathcal{D}(A) = \{f \in H^2(0,1); f(0) = f(1) = 0\} \subset \mathcal{Z}. \quad (24)$$

For any $r \in \Omega^o$, define the output operator $C(r) \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$ to be

$$C(r)f := < r, f >_{\mathcal{Z}}, \ \forall f \in \mathcal{Z}. \quad (25)$$

The model (23) can be then written into the following abstract form:

$$\dot{z}(t) = Az(t) + Gw(t), z(0) = z_0, t \geq 0, \quad (26)$$

$$y(t) = C(r)z(t) + v(t), \quad (27)$$

where

$$Gw = g(x)w \in \mathcal{Z}, \quad (28)$$

$$v(t) \in \mathcal{Y}, t \geq 0. \quad (29)$$

It is well-known that A generates an exponentially stable semigroup, and thus Assumptions **(A1)**, **(A4)** and **(A5)** are immediately satisfied. Moreover, by the definition of G in (28) and of $C(r)$ in (25) respectively, Assumptions **(A2)** and **(A6)** are satisfied. Hence, it follows from Theorem 3 that there exists an optimal shape for the sensor distribution, the cost functional being the estimation error variance with infinite-time Kalman filtering.

Example: Wave equation

Consider the two-dimensional wave equation

$$\begin{cases} z_{tt}(t,\mathbf{x}) = \Delta z(t,\mathbf{x}) + g(\mathbf{x})w(t), (t,\mathbf{x}) \in (0,\infty) \times \Xi, \\ z(t,\mathbf{x}) = 0, (t,\mathbf{x}) \in [0,\infty) \times \Gamma_0, \\ \frac{\partial z}{\partial \nu}(t,\mathbf{x}) = -a z_t(t,\mathbf{x}), (t,\mathbf{x}) \in [0,\infty) \times \Gamma_1, \\ z(0,\mathbf{x}) = z_0(\mathbf{x}), z_t(0,\mathbf{x}) = z_1(\mathbf{x}), \mathbf{x} \in \Xi, \\ y(t) = \int_\Xi r(\mathbf{x}) z(t,\mathbf{x}) d\mathbf{x} + v(t), \end{cases} \quad (30)$$

where $a > 0$ is a constant. The spatial variable $\mathbf{x} = (x_1, x_2)^T$, the spatial domain

$$\Xi = \{\mathbf{x} \in \mathbb{R}^2; 1 < x_1^2 + x_2^2 < 4\} \quad (31)$$

is an Euclidean annulus, and the boundary of Ξ

$$\partial \Xi = \Gamma_0 \cup \Gamma_1,$$

with

$$\Gamma_0 = \{\mathbf{x} \in \mathbb{R}^2; x_1^2 + x_2^2 = 1\},$$
$$\Gamma_1 = \{\mathbf{x} \in \mathbb{R}^2; x_1^2 + x_2^2 = 4\}.$$

The state z is subject to a process noise $w(t)$ with the spatial distribution $g(x)$; and the output $y \in \mathbb{R}$ denotes the corresponding available measurement from one sensor subject to a sensor noise $v(t)$. The disturbances w and v are real-valued white Gaussian noises, with variances Q and R respectively. The initial condition z_0, z_1 and the noises $w(t), v(t)$ are assumed to be mutually uncorrelated. The objective, as the previous example, is to find the optimal sensor shape described by $r(\mathbf{x})$ over the spatial region Ξ, that together with the corresponding infinite-time Kalman filter achieves minimum estimation error covariance.

Set

$$\Omega^o := \left\{ r(\mathbf{x}) \in L^\infty(\Xi); \underline{r} \leq r(\mathbf{x}) \leq \bar{r}, \int_0^1 r(\mathbf{x})\, \mathrm{d}\mathbf{x} \leq M \right\},$$

where $\underline{r}, \ \bar{r}, M$ are constants. The set Ω^o is sequentially compact in in the weak-star topology on $L^\infty(\Xi)$. Assumption **(A3)** is thus satisfied.

Define the state space $\mathcal{Z} = H^1(\Xi) \times L^2(\Xi)$, with the usual inner product

$$\langle (f_1, g_1)^T, (f_2, g_2)^T \rangle_{\mathcal{Z}} := \int_\Xi (\nabla f_1(\mathbf{x}) \nabla f_2(\mathbf{x}) + g_1(\mathbf{x}) g_2(\mathbf{x}))\, d\mathbf{x}. \quad (32)$$

Define the operator $A : D(A) \to \mathcal{Z}$ as

$$A(f,g)^T = (g, \Delta f)^T,$$

with the domain

$$\mathcal{D}(A) = \{(f,g)^T \in H^2(\Xi) \times H^1(\Xi); \left(\frac{\partial f}{\partial \nu} + ag \right) \bigg|_{\Gamma_1} = f|_{\Gamma_0} = 0\} \subset \mathcal{Z}. \quad (33)$$

For any $r \in \Omega^o$, denote the output operator $C(r) \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$

$$C(r)f := < r, f >_{\mathcal{Z}}, \ \forall f \in \mathcal{Z}. \quad (34)$$

The model (30) can be then written in the following abstract form:

$$\dot{\mathbf{z}}(t) = A\mathbf{z}(t) + Gw(t), \mathbf{z}(0) = (z_0, z_1)^T, t \geq 0, \quad (35)$$

$$y(t) = C(r)\mathbf{z}(t) + v(t), \quad (36)$$

where $G \in \mathcal{L}(\mathcal{W}, \mathcal{Z})$ and

$$Gw(t) = g(x)w(t) \in \mathcal{Z}, \quad (37)$$

$$v(t) \in \mathcal{Y}, t \geq 0. \quad (38)$$

The operator A generates an exponentially stable semigroup, see [13, Page 668], and thus Assumptions **(A1)**, **(A4)**, **(A5)** are immediately satisfied. Moreover, by the definition of G in (37) and of $C(r)$ in (34) respectively, Assumptions **(A2)** and **(A6)** are satisfied. Hence, Theorem 3 guarantees the existence of an optimal shape for the sensor distribution.

IV. Calculation of Optimal Sensor Design

For an infinite-dimensional system, an explicit solution $\Pi(r)$ to the ARE (21) is in general not available. In practice, approximation methods are used, which convert the original infinite-dimensional problem into finite-dimensional problems.

Consider finite-dimensional subspaces $\{\mathcal{Z}_n\}$ of $\mathcal{Z}$, where $n \in \mathbb{N}^+$ indicates the dimension, with $\{P_n\}$ the orthogonal projection of $\mathcal{Z}$ onto $\{\mathcal{Z}_n\}$. The space $\{\mathcal{Z}_n\}$ is equipped with the inner product inherited from $\mathcal{Z}$. The infinite-dimensional system (1) – (2) can be approximated by a finite-dimensional system using the Galerkin method. This leads to sequences of operators $A_n(r) \in \mathcal{L}(\mathcal{Z}_n, \mathcal{Z}_n)$, $G_n = P_n G \in \mathcal{L}(\mathcal{W}, \mathcal{Z}_n)$, $C_n(r) := C(r)|_{\mathcal{L}(\mathcal{Z}_n, \mathcal{Y})}$:

$$\dot{z}(t) = A_n(r)z(t) + G_n w(t), \quad t \geq 0, \tag{39}$$

$$z(0) = z_{0,n} = P_n z_0, \tag{40}$$

$$y_n(t) = C_n(r)z(t) + Hv(t), \quad t \geq 0 \tag{41}$$

that approximate (1) – (2) on $\mathcal{Z}_n$.

Since for any $n \in \mathbb{N}^+$, $\mathcal{Z}_n$ is also complete and therefore a Hilbert space, the results in Section III can be applied to the approximating system as well. Suppose that assumptions **(A1)**- **(A6)** are satisfied for the original problem on $\mathcal{Z}$. Then assumptions **(A1)**- **(A3)** are also satisfied by the approximation problems. Assume in addition that for each $n \in \mathbb{N}^+$,

- **(A9)** the pairs of functions $(A_n(r), G_n Q^{1/2})$ are uniformly exponentially stabilizable with respect to $r \in \Omega^o$;
- **(A10)** the pairs of functions $(A_n(r), C_n(r))$ are exponentially detectable for at least one $r \in \Omega^o$;
- **(A11)** for any $n \in \mathbb{N}^+$, for any $r \in \Omega^o$ and for any sequence $\{r_m\} \subset \Omega^o$ that converges to $r \in \Omega^o$,

  $$\lim_{n\to\infty} \|C_n(r_m) - C_n(r)\|_{\mathcal{L}(\mathcal{Z}_n, \mathcal{Y})} = 0,$$

  and moreover, for each $z \in \mathcal{Z}$,

  (i) $\lim_{n\to\infty} \|e^{tA(r_n)}z - e^{tA(r)}z\|_{\mathcal{Z}} = 0,$

  (ii) $\lim_{n\to\infty} \|e^{tA(r_n)^*}z - e^{tA(r)^*}z\|_{\mathcal{Z}} = 0$

  uniformly in t on bounded intervals.

It follows immediately from Theorem 3 that there exists an optimal sensor design $\hat{r}_n$ in the sense of minimizing the error variance. More precisely, the minimum cost for the approximating problems is

$$\|\Pi_n(\hat{r}_n)\|_1 = \min_{r \in \Omega^o} \|\Pi_n(r)\|_1, \tag{42}$$

where for each $r \in \Omega^o$, $\Pi_n(r)$ is the unique non-negative solution $X(r) \in \mathcal{L}(\mathcal{Z}_n, \mathcal{Z}_n)$ to the finite-dimensional ARE

$$A_n(r)X(r) + X(r)A_n^*(r) - X(r)C_n^*(r)R^{-1}C_n(r)X(r) + G_n(r)QG_n(r)^* = 0. \tag{43}$$

Under conditions **(A9)-(A11)** on the approximation scheme, for fixed r,

$$\lim_{n\to\infty} \|\Pi_n(r) - \Pi(r)\|_1 = 0.$$

(See [15][Thm. 3.8].) These conditions also imply convergence of the approximating minimium variance and sensor locations for the more restrictive problem of sensor location with no effect of the sensors on the system dynamics [24][Thm. 4.2]. It is anticipated that, under similar conditions on the approximation scheme, the optimal solutions $\Pi_n(\hat{r}_n)$ converge to $\Pi(\hat{r})$ where $\Pi(\hat{r})$ indicates the solution to the ARE (21) with $\hat{r}$ satisfying (22). That is, letting $\hat{r} \in \Omega^o$ indicate the the optimal sensor design for the $(A(r), G, C(r))$–problem there is a sequence of optimal sensor designs $\{\hat{r}_n\} \subset \Omega^o$ for the approximating $(A_n(r), G_n, C_n(r))$–problems such that

$$\lim_{n\to\infty} \|\Pi_n(\hat{r}_n)\|_1 = \|\Pi(\hat{r})\|_1. \tag{44}$$

V. Conclusions

The well-known Kalman filter minimizes the steady-state error variance of an estimator for a linear system under certain conditions on the noise. In this paper, minimum variance filter design for distributed parameter systems was combined with optimization of the sensor design. This is a generalization of the sensor location problem that includes sensor shape as well as location. The framework also allows for the effect of sensor mass and dynamics on the system dynamics. Provided that the admissible set is formulated so as to be sequentially compact in some topology, not necessarily the norm topology, an estimator that minimizes the error variance over possible filters and over possible sensor design exists. The result was illustrated with several examples.

It is anticipated that approximation algorithms [15] can be used in the calculation of optimal sensor design and the corresponding optimal cost. Convergence of the cost and the optimal sensor need to be established as the approximation order is increased. Current research is concerned with establishing sufficient conditions for approximations under which this convergence is guaranteed, and in the construction of a suitable algorithm.

References

[1] H. T. Banks, R. C. Smith, and Y. Wang, *Smart Material Structures: Modeling, Estimation, and Control*, New York: Wiley, 1996.

[2] V. Barbu and T. Precupanu, *Convexity and optimization in Banach spaces*, Springer Science & Business Media, 2012.

[3] A. Bensoussan, *Optimization of sensors' location in a distributed filtering problem*, Springer Berlin Heidelberg, 1972, pp. 62–84.

[4] R. F. Curtain and A. J. Pritchard, *Infinite dimensional linear systems theory*, vol. 8, Springer-Verlag Berlin, 1978.

[5] N. Darivandi, K. A. Morris, and A. Khajepour, *An algorithm for LQ-optimal actuator location*, Smart Materials and Structures, 22 (2013), p. 035001.

[6] F. Fahroo and M. A. Demetriou, *Optimal location of sensors and actuators for an active noise control problem*, Journal of Computational and Applied Mathematics, 114 (2000), pp. 137–158.

[7] F. Fahroo and K. Ito, *Variational formulation of optimal damping designs*, in Optimization methods in partial differential equations (South Hadley, MA, 1996), vol. 209 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1997, pp. 95–114.

[8] V. A. Gupta, M. B. Sharma, and N. Thakur, *Optimization criteria for optimal placement of piezoelectric sensors and actuators on a smart structure: A technical review*, J. Intell. Mat'l Sys. & Struct., 21 (2010), pp. 1227–1243.

[9] W. Hu, K. Morris, and Y. Zhang, *Sensor location in a controlled thermal fluid*, in Proc. of the 55th IEEE Conference on Decision and Control, 2016.

[10] D. Kasinathan and K. A. Morris, *H_∞-optimal actuator location*, IEEE Tran. on Automatic Control, 58 (2013), pp. 2522 – 2535.

[11] T. Khan, K. A. Morris, and M. Stastna, *Computation of the optimal sensor location for the estimation of a linear dispersive wave equation*, in American Control Conference, 2015.

[12] C. S. Kubrusly and H. Malebranche, *Sensors and controllers location in distributed systems-a survey*, Automatica, 21 (1985), pp. 117–128.

[13] I. Lasiecka and R. Triggiani, *Control Theory for Partial Differential Equations: Continuous and Approximation Theories, Volume II, Abstract Parabolic Systems over a Finite Time Horizon*, Cambridge University Press, 2000.

[14] K. A. Morris, *Noise reduction achievable by point control*, ASME Journal on Dynamic Systems, Measurement and Control, 120 (1998), pp. 216–223.

[15] K. A. Morris, *Linear-quadratic optimal actuator location*, IEEE Tran. on Automatic Control, 56 (2011), pp. 113–124.

[16] K. A. Morris, M. A. Demetriou, and S. D. Yang, *Using $\mathbb{H}_2$-control performance metrics for infinite-dimensional systems*, IEEE Tran. on Automatic Control, 60 (2015).
[17] K. A. Morris, S.-X. Tang, and A. Vest, *Optimal design with linear-quadratic performance measures*. preprint.
[18] Y. Privat, E. Trélat, and E. Zuazua, *Optimal observation of the one-dimensional wave equation*, Jour. of Fourier Analysis and Applications, 19 (2013), pp. 514–544.
[19] ———, *Optimal shape and location of sensors for parabolic equations with random initial data*, Archive for Rational Mechanics and Analysis, 216 (2015), pp. 921–981.
[20] D. Simon, *Optimal state estimation: Kalman, H_∞, and nonlinear approaches*, John Wiley & Sons, 2006.
[21] M. van de Wal and B. de Jager, *A review of methods for input/output selection*, Automatica, 37 (2001), pp. 487–510.
[22] X. Wu, B. Jacob, and H. Elbern, *Optimal control and observation locations for time-varying systems on a finite-time horizon*, SIAM J. Control and Optim., 54 (2015), pp. 291–316.
[23] S. D. Yang and K. A. Morris, *Comparison of actuator placement criteria for control of structures*, Journal of Sound and Vibration, 353 (2015), pp. 1–18.
[24] M. Zhang and K. Morris, *Sensor choice for minimum error variance estimation*, IEEE Transactions on Automatic Control, (2017).