2022 American Control Conference (ACC)
Atlanta, USA, June 8-10, 2022

Economic aspects of sensor selection optimization of finite and infinite dimensional dynamical systems

Michael A. Demetriou

Abstract— **The economic aspects as a new factor in the selection of sensors for improved filtering of dynamical systems are introduced. By using the price of a single sensor, reflected by high values of the associated covariance, an economic aspect of the sensor optimization for optimal filtering is introduced. Both the unit price and the total price of a network of inexpensive noisy sensors are used as an alternative to the performance of a single expensive and highly accurate sensor. Algorithms for the integrated sensor optimization for both finite and infinite dimensional systems are presented and examples are provided to demonstrate these effects.**

I. Introduction

The problem of sensor placement and selection in improved filtering of systems governed by partial differential equations (PDEs) has been explored as early as mid-70s with the paper by Curtain and Ichikawa [1]. and the paper by Amouroux, Babary, and Malandrakis [2]. Immediately after the initial works, many others considered deterministic or stochastic diffusion PDEs. The concern was for either the optimal sensor selection for filtering, or the optimal sensor and actuator selection for compensator design, [3], [4], [5].

Most of the optimal sensor selection works considered the optimal performance as expressed in terms of the location-parameterized solution to an operator Riccati equation. While the numerical aspects were of concern, the issue of the convergence of the optimal sensor (or actuator) location was not considered at that time. A somewhat related issue, that of the convergence of the matrix representations of the Riccati equations to their respective infinite dimensional operators was examined in the 80's, see [6] and references therein.

The missing part were the conditions that enable one to find the optimal sensor location of the finite dimensional representation of the distributed process and guarantee that it would converge to the optimal location of the infinite dimensional system. This was first considered in the context of optimal actuators in the optimal control (LQR/$\mathbb{H}^2$/$\mathbb{H}^\infty$) problem in the works by Morris [6], [7], [8], [9].

In the context of the sensor selection for optimal filtering, the work by Demetriou and Fahroo in [10] considered this problem for a class of infinite dimensional systems. At about the same time, Zhang and Morris, in [11], [12] considered a more general class of infinite dimensional systems and also brought forth the issue of the sensor quality, as expressed in terms of the sensing device covariance.

In this work, another level in the sensor optimization is added by incorporating the price of the devices. A pedestrian description would consider a single optimally placed expensive and highly accurate, in the sense of very low covariance, sensor and a network of optimally placed inexpensive and inaccurate sensors. If price is the deciding factor, would someone consider a compromised filter performance? If both price and performance are the deciding factors, would someone select a single accurate expensive sensor or a network of moderately priced and moderately accurate sensors?

The above issues will be considered in the context of sensor quality, number and location for the optimal filtering of both infinite and finite dimensional dynamical systems. As a performance measure of the optimal filter, the mean reconstruction error of the associated Kalman filter will be used and which is expressed in terms of the variance operator/matrix. This variance is subsequently optimized with respect to the price, location and number of candidate sensors in order to yield the optimal filter.

In Section II, the problem is formulated for a parabolic PDE in one spatial dimension and subsequently is extended in Section III to a class of infinite dimensional systems. The sensor accuracy, as reflected in the value of the sensor noise covariance is incorporated into the performance measure along with the sensor locations and the sensor numbers. Various optimization problems are stated involving the selection of the number, location and total price of sensors. A performance matching is incorporated whereby one matches the performance of many inexpensive sensors placed at optimal locations to the filter performance of a single expensive idealised sensor. The only criterion becomes that of total price. The same results are also presented for a class of linear finite dimensional dynamical systems, presented in Section IV. In this case, the sensor parametrization is made via the use of the unit vectors in the finite dimensional state space $\mathbb{R}^N$ and does not involve a sensor location, but rather a sensor selection. Numerical results for both a PDE and an N-dimensional system are included in Section V to provide insights on metrics for sensor selection in the filtering of dynamical systems. One of the revealing findings is that a single expensive and accurate sensor may not be the most economic choice for the filtering of dynamical systems. Using a network of inexpensive and noisy sensors, their filter performance is matched to the performance of a single expensive sensor and thus the selection criterion becomes that of the total price. As concluded in Section IV with the filter performances matched, *a network of noisy inexpensive filters is cheaper to use than a single expensive and accurate sensor*. Conclusions and extensions on the economic aspects for sensor selection are summarized in Section VI.

M. A. Demetriou is with Worcester Polytechnic Institute, Aerospace Engineering Dept., Worcester, MA 01609, USA, mdemetri@wpi.edu. The author acknowledges financial support from NSF-CMMI grant # 1825546.

II. PROBLEM FORMULATION

A representative infinite dimensional system is provided by the advection-diffusion PDE in one spatial dimension

$$\begin{aligned}\frac{\partial x(t,\xi)}{\partial t} &= \frac{\partial}{\partial \xi}\left(a\frac{\partial x(t,\xi)}{\partial \xi}\right)+b\frac{\partial x(t,\xi)}{\partial \xi}+cx(t,\xi) \\ &\quad +b_1(\xi)w(t)+b_2(\xi)u(t),\end{aligned} \qquad (1)$$

where $x(t,\xi)$ denotes the state at time $t\in\mathbb{R}^+$ and location $\xi\in\Omega=[0,L]$. The spatial functions $b_1(\xi), b_2(\xi)$ denote the spatial distributions of the disturbance (process noise) and actuator input, respectively. The temporal signals $w(t)$ and $u(t)$ denote the process noise and control input, respectively. The above equation is furnished with the appropriate boundary and initial conditions. For simplicity, it is assumed that Dirichlet boundary conditions are assumed $x(t,0)=0=x(t,L)$ and the initial condition is given by $x(0,\xi)=x_0(\xi)$.

The above PDE is supplemented with output measurements. Since two separate measurements associated with different sensing devices are considered, we present the "idealized" measurement as the output of an expensive sensor with minimal noise given by

$$y_0(t)=\int_0^L c_0(\xi;\xi_0)x(t,\xi)\,\mathrm{d}\xi+\nu_0(t). \qquad (2)$$

Similarly, we denote the measurements from the inexpensive noisy sensing devices by

$$y_i(t)=\int_0^L c_i(\xi;\xi_i)x(t,\xi)\,\mathrm{d}\xi+\nu_i(t), \quad i=1,\ldots,n. \qquad (3)$$

The function $\nu_0(t)$ denotes the noise corresponding to the expensive and idealized sensor placed at the location $\xi_0\in\Omega$ and having covariance N_0. Similarly, the functions $\nu_i(t)$, $i=1,\ldots,n$ denote the noise of the n inexpensive sensing devices placed at the locations ξ_i, $i=1,\ldots,n$ and having covariances given by N_i, $i=1,\ldots,n$. The measurement noise functions are assumed to be real-valued white noise with

$$\mathcal{E}\left[\nu_i(t)\nu_j(\tau)\right]=N\delta(t-\tau).$$

Note that since the n inexpensive sensors do not have the same performance as the single expensive sensor, then

$$N_i \gg N_0, \qquad i=1,\ldots,n.$$

The problem at hand can be summarized here: *given the process* (1) *and the idealized sensor* (2)*, find the optimal sensor location ξ_0 to provide the optimal state estimator performance. This performance now serves as the reference point for the selection of the inexpensive sensors* (3) *and the optimization problem becomes that of finding the minimum number n of sensing devices and their associated optimal locations ξ_i, $i=1,\ldots,n$ so that the state estimator associated with these n inexpensive sensing devices is "as close as" possible to the performance provided by the idealized expensive sensor. Differently put, find the smallest number of the inexpensive sensing devices and their associated optimal locations so that the resulting state estimator "matches" the filter performance provided by the single expensive sensor.*

With the filter performances being equal, one essentially replaces a single expensive sensing device by a number n of inexpensive sensors. The economic aspects of this optimization then become obvious.

To formulate the optimization problem for both the finite and infinite dimensional cases, we provide some additional information on the sensing devices and bring the system in (1) with (2) or with (3) in state space form, written as an evolution equation in a Hilbert space.

Assumption 1 (sensor price): The price of each of the inexpensive sensing devices associated with (3) is denoted by p_i, $i=1,\ldots,n$ and the price of the expensive sensor associated with (2) is denoted by p_0.

Following Assumption 2, the total price of the n sensors is

$$P_{\text{total}}=\sum_{i=1}^{n} p_i, \qquad (4)$$

and in the event that the n sensing devices are identical with $p_i=p_j$ for all $i,j=1,\ldots,n$ and denoted by p, then the total price of the n inexpensive sensing devices reduces to

$$P_{\text{total}}=np. \qquad (5)$$

Assumption 2 (sensor type): The n inexpensive sensors are identical and thus the total price is given by (5).

Note that when the single expensive sensor is replaced by the inexpensive noisy sensors the filter performance will be matched and additionally one would achieve $np \ll p_0$.

III. INFINITE DIMENSIONAL SYSTEMS

The PDE in (1) can be written as an evolution equation in the Hilbert space $H=L_2(\Omega)$. In addition to the state space, we consider the interpolating spaces $V=H_0^1(\Omega)$ and $V^*=H^{-1}(\Omega)$. The state operator $A\in\mathcal{L}(V,V^*)$ associated with (1) is given in weak form by

$$\langle A\phi,\psi\rangle=\int_0^L\left(\frac{\mathrm{d}}{\mathrm{d}\xi}\left(a\frac{\mathrm{d}\phi(\xi)}{\mathrm{d}\xi}\right)+b\frac{\mathrm{d}\phi(\xi)}{\mathrm{d}\xi}+c\phi(\xi)\right)\psi(\xi)\,\mathrm{d}\xi \qquad (6)$$

for all test functions $\phi,\psi\in V$. The input operators associated with the spatial functions $b_1(\xi), b_2(\xi)$ are given by

$$\langle B_i v(t),\psi\rangle=\int_0^L b_i(\xi)v(t)\psi(\xi)\,\mathrm{d}\xi, \quad i=1,2, \quad v=w,u, \qquad (7)$$

for all $\psi\in V$. Using (6), (7), the PDE in (1) is written as

$$\dot{x}(t)=Ax(t)+B_1w(t)+B_2u(t), \qquad (8)$$

with initial condition $x(0)\in\mathcal{D}(A)$. Similarly, the output signals associated with the single expensive sensor (2) and with the n inexpensive sensors (3) are given by

$$\begin{aligned}y_0(t) &= C_0x(t)+\nu_0(t) \\ &= \int_0^L c_0(\xi;\xi_0)x(t,\xi)\,\mathrm{d}\xi+\nu_0(t),\end{aligned} \qquad (9)$$

and by

$$\begin{aligned}\mathbf{y}(t) &= \begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t)\end{bmatrix}=\begin{bmatrix} C_1x(t)+\nu_1(t) \\ \vdots \\ C_nx(t)+\nu_n(t)\end{bmatrix} \\ &= \begin{bmatrix} \int_0^L c_1(\xi;\xi_1)x(t,\xi)\,\mathrm{d}\xi+\nu_1(t) \\ \vdots \\ \int_0^L c_n(\xi;\xi_n)x(t,\xi)\,\mathrm{d}\xi+\nu_n(t)\end{bmatrix}\end{aligned} \qquad (10)$$

Using a fixed location $\xi_0\in\Omega$ for the expensive sensor in (9) and using (8), one can design an associated Kalman filter to provide the optimal state estimate. This is given by

$$\dot{\widehat{x}}_0(t) = A\widehat{x}_0(t) + B_2(t)u(t) + K_0\left(y_0(t) - C_0\widehat{x}_0(t)\right), \quad (11)$$

with initial condition $\widehat{x}_0(0)$, where the filter operator gain $K_0 = \Sigma_0 C_0^* N_0^{-1}$ is given via the solution to the Operator Algebraic Riccati Equation

$$\begin{aligned}&\langle A\phi, \Sigma_0\psi\rangle + \langle \Sigma_0\phi, A^*\psi\rangle + \langle \phi, B_1B_1^*\psi\rangle \\ &\quad -\langle \phi, \Sigma_0 C_0^* N_0^{-1} C_0 \Sigma_0\psi\rangle = 0, \quad \phi,\psi \in \mathcal{D}(A^*).\end{aligned} \quad (12)$$

The abstract representation of (1) given in (8) via (6), (7), (9) can also be used to describe a large class of PDEs in two and three spatial domains. Thus, we consider evolution equations given by (8) representing distributed parameter systems. The state space H is a Hilbert space equipped with inner product and norm denoted by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$, respectively. The state operator A in (8) is the generator of a strongly continuous semigroup $T(t)$ on H, [13]. The control operator $B_2 \in \mathcal{L}(U,H)$ is a bounded linear operator from the control space U to H and the process noise operator $B_1 \in \mathcal{L}(W,H)$ is a bounded linear operator from the process noise space W to H. Indirectly assumed, the process and measurement uncertain signals w and v, respectively, are assumed to be square integrable to ensure well-posedness of (8). Finally, the output operator associated with the expensive sensor, i.e. device with "small" N_0, is $C_0 \in \mathcal{L}(Y,H)$ a bounded linear operator from the output space Y to the state space H.

To derive the baseline filter performance, the single expensive sensor must be optimally placed in the domain Ω. Thus, the output operator C_0 is parameterized by the candidate locations $\xi_0 \in \Theta$, where Θ is the parameter set and comprises the set of admissible sensor locations $\xi_0 \in \Omega$ that render the system approximately observable, [14]. Thus, we explicitly state this location dependence with (9) re-written as

$$y_0(t;\xi_0) = C_0(\xi_0)x(t) + \nu_0(t). \quad (13)$$

The output measurement due to the expensive sensor is explicitly dependent on the candidate locations $\xi_0 \in \Theta$. The parameter set is formally defined via

$$\Theta = \{\xi_0 \in \Omega : \left(C_0(\xi_0), A\right) \text{ is approximately observable}\} \quad (14)$$

A. Optimal location of single expensive sensor

The state estimator (11) is now expressed in terms of the location-parameterized output operator and thus

$$\begin{aligned}\dot{\widehat{x}}_0(t;\xi_0) &= A\widehat{x}_0(t;\xi_0) + B_2u(t) \\ &\quad + K_0(\xi_0)\left(y_0(t;\xi_0) - C_0(\xi_0)\widehat{x}_0(t;\xi_0)\right) \\ \widehat{x}(0;\xi_0) &\in \mathcal{D}(A), \quad \xi_0 \in \Theta.\end{aligned} \quad (15)$$

As denoted above, the operator $K_0(\xi_0) = \Sigma_0(\xi_0)C_0^*(\xi_0)N_0^{-1}$ is derived from the location-parameterized positive operator solution to the algebraic Riccati equation

$$\begin{aligned}&\langle A\phi, \Sigma_0(\xi_0)\psi\rangle_H + \langle \Sigma_0(\xi_0)\phi, A^*\psi\rangle_H + \langle \phi, B_1B_1^*\psi\rangle_H \\ &-\langle \phi, \Sigma_0(\xi_0)C_0^*(\xi_0)N_0^{-1}C_0(\xi_0)\Sigma_0(\xi_0)\psi\rangle_H = 0,\end{aligned} \quad (16)$$

for $\xi_0 \in \Theta$. To find the optimal estimator (16), one must perform a sensor location optimization. A suitable metric for location optimization, that is to find the "best" $\widehat{x}_0(t;\xi_0)$ over all $\xi_0 \in \Theta$, is the mean reconstruction error which is expressed in terms of the trace of the variance operator

$$\begin{aligned}J^{opt}(\xi_0) &= \mathcal{E}\left[\langle x(t) - \widehat{x}_0(t;\xi_0), x(t) - \widehat{x}_0(t;\xi_0)\rangle_H\right] \\ &= \text{trace } \left[\Sigma_0(\xi_0)\right], \quad \xi_0 \in \Theta.\end{aligned} \quad (17)$$

The optimal location of the single expensive sensor is then

$$\xi_0^{opt} = \arg \inf_{\xi_0 \in \Theta} \text{trace } \left[\Sigma_0(\xi_0)\right]. \quad (18)$$

The optimal performance associated with the optimal single expensive sensor is subsequently given by

$$\boxed{J^{opt}(\xi_0^{opt}) = \text{trace } \left[\Sigma_0(\xi_0^{opt})\right].} \quad (19)$$

Remark 1: Note that the optimal performance (19) is the one to be used to find the minimum number n and optimal locations ξ_i^{opt} of the inexpensive sensors such that the performance of the associated filter matches (19).

B. Optimal number and location of inexpensive sensors

To arrive at a location-parameterized filter, similar to (15), some simplifying assumptions must be made. The parameter space for the n sensing devices is assumed to be the same as for the single expensive device which means that each of the n inexpensive sensing devices can by itself result in an approximately observable pair. It should be noted that this type of observability is not achieved collectively by a number of sensing devices, but by each of the devices independently. We make the following assumptions.

Assumption 3: The sensing devices, namely the single expensive device and the n inexpensive devices, have identical spatial distributions, differing only on the measurement noise statistics. Thus, the corresponding output operators are

$$C_i(\xi_s) = C_0(\xi_s), \quad \forall \xi_s \in \Theta, \quad i = 1,\ldots,n.$$

With regards to the specific PDE in (1), and measurements (2), (3), the above assumption translates to

$$c_i(\xi;\xi_s) = c_0(\xi;\xi_s), \quad \forall \xi_s \in \Theta, \quad i = 1,\ldots,n.$$

The measurement vector due to the n inexpensive devices in (10), parameterized by the sensor locations is given by

$$\begin{aligned}\mathbf{y}(t;\theta) &= \begin{bmatrix} C_0(\xi_1)x(t) + \nu_1(t) \\ \vdots \\ C_0(\xi_n)x(t) + \nu_n(t)\end{bmatrix} \\ &= \mathbf{C}(\theta)x(t) + \boldsymbol{\nu}(t)\end{aligned} \quad (20)$$

where $\boldsymbol{\nu}(t) = \begin{bmatrix} \nu_1(t) & \cdots & \nu_n(t)\end{bmatrix}^T$, $\boldsymbol{\theta} = \{\xi_1,\ldots,\xi_n\} \in \prod^n \Theta = \boldsymbol{\Theta}$, $\mathbf{C}(\theta) = \text{diag}\{C_0(\xi_1),\ldots,C_0(\xi_n)\}$. The covariance of the sensor noise in (20) is given by $\mathbf{N} = \text{diag}\{N_1,\ldots,N_n\}$.

Following a similar procedure for the single expensive sensor, the state estimator associated with the n measurements (20), parameterized by the n locations $\boldsymbol{\theta} = \{\xi_1,\ldots,\xi_n\}$ is

$$\begin{aligned}\dot{\widehat{\mathbf{x}}}(t;\boldsymbol{\theta}) &= A\widehat{\mathbf{x}}(t;\boldsymbol{\theta}) + B_2u(t) \\ &\quad + \mathbf{K}(\boldsymbol{\theta})\left(\mathbf{y}(t;\boldsymbol{\theta}) - \mathbf{C}(\boldsymbol{\theta})\widehat{\mathbf{x}}(t;\boldsymbol{\theta})\right) \\ \widehat{\mathbf{x}}(0;\boldsymbol{\theta}) &\in \mathcal{D}(A), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}.\end{aligned} \quad (21)$$

where the filter gain $\mathbf{K}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{C}^*(\boldsymbol{\theta})\mathbf{N}^{-1}$ is obtained from the location-parameterized positive operator solution to the filter algebraic Riccati equation

$$\begin{aligned}&\langle A\phi, \boldsymbol{\Sigma}(\boldsymbol{\theta})\psi\rangle_H + \langle \boldsymbol{\Sigma}(\boldsymbol{\theta})\phi, A^*\psi\rangle_H + \langle \phi, B_1B_1^*\psi\rangle_H \\ &-\langle \phi, \boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{C}^*(\boldsymbol{\theta})\mathbf{N}^{-1}\mathbf{C}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta})\psi\rangle_H = 0, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}.\end{aligned} \quad (22)$$

For a fixed sensor location vector $\theta \in \boldsymbol{\Theta}$, the cost associated with the filter (21) is given by

$$\begin{aligned} \mathbf{J}^{opt}(\theta) &= \mathcal{E}\left[\langle x(t)-\widehat{\mathbf{x}}(t;\boldsymbol{\theta}), x(t)-\widehat{\mathbf{x}}(t;\boldsymbol{\theta})\rangle_H^2\right] \\ &= \text{trace}\ [\boldsymbol{\Sigma}(\boldsymbol{\theta})], \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}. \end{aligned} \tag{23}$$

Minimization of (23) over the admissible locations will provide the optimal location vector $\boldsymbol{\theta}^{opt}$ that minimizes trace$[\boldsymbol{\Sigma}(\boldsymbol{\theta})]$. However one is not simply interested in finding the optimal sensor locations. Instead, one is interested in finding the minimum number n of the optimal sensor locations that minimize trace $[\boldsymbol{\Sigma}(\boldsymbol{\theta})]$ that have the same performance as the single optimal expensive sensor (19) and minimize (5).

Sensor optimization O_1 (fixed price):

$$\begin{aligned} &\text{minimize} && n \\ &\text{subject to} && \inf_{\theta \in \boldsymbol{\Theta}} \text{trace}\,[\boldsymbol{\Sigma}(\boldsymbol{\theta})] = \inf_{\xi_0 \in \Theta} \text{trace}\,[\Sigma_0(\xi_0)], \\ &&& np < p_0. \end{aligned} \tag{24}$$

For a fixed number n of inexpensive sensing devices, one can find the optimal sensor location, given by

$$\boldsymbol{\theta}^{opt} = \arg \inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \text{trace}\,[\boldsymbol{\Sigma}(\boldsymbol{\theta})] \tag{25}$$

resulting in the optimal cost associated with n sensors

$$\mathbf{J}^{opt}(\boldsymbol{\theta}^{opt}) = \text{trace}\left[\boldsymbol{\Sigma}(\boldsymbol{\theta}^{opt})\right]. \tag{26}$$

In view of (25), (26), (24) is compactly written as

$$\boxed{\begin{aligned} &\text{minimize} && n \\ &\text{subject to} && \text{trace}\left[\boldsymbol{\Sigma}(\boldsymbol{\theta}^{opt})\right] = \text{trace}\left[\Sigma_0(\xi_0^{opt})\right], \\ &&& np < p_0. \end{aligned}} \tag{27}$$

Other related optimization problems can be considered.

Algorithm 1 Economic sensor optimization: fixed price

1: **initialize:** Find optimal location ξ_0^{opt} of single expensive sensor via (18)
2: **initialize:** Select integer $n = 2$
3: **loop**
4: define parameter space $\boldsymbol{\Theta}^n = \prod^n \Theta$
5: find optimal sensor location $\boldsymbol{\theta}^{opt,n}$ of inexpensive sensors using
$$\theta^{opt,n} = \arg \inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}^n} \text{trace}\ [\boldsymbol{\Sigma}(\boldsymbol{\theta})]$$
6: compute price of n inexpensive sensors using
$$P_{\text{total}}^n = np$$
7: **if** $\left|\text{trace}\,[\boldsymbol{\Sigma}(\boldsymbol{\theta}^{opt,n})] - \text{trace}\left[\Sigma_0(\xi_0^{opt})\right]\right| \geq \varepsilon$ or $P_{\text{total}}^n \geq p_0$ **then**
8: $n \leftarrow n+1$
9: **goto** 3
10: **else**
11: terminate
12: **end if**
13: **end loop**

For example, one may fix a priori both the number n and the sensor locations and optimize the filter (21) in order to match its performance (23) to (19) by selecting the sensor accuracy. The sensor accuracy is reflected in the value of the covariance, which itself is linked to the price p. The more reliable a sensing device is, equivalently the lower the N_i is, the higher the price. Assuming this relationship, a simple model relating the price to the accuracy (covariance) is

$$p \propto 1/N. \tag{28}$$

In this case the optimization becomes that of minimizing price or sensor covariance.

Sensor optimization O_2 (variable price):

$$\boxed{\begin{aligned} &\text{minimize} && p \\ &\text{subject to} && \inf_{p} \text{trace}\,[\boldsymbol{\Sigma}(\boldsymbol{\theta})] = \text{trace}\left[\Sigma_0(\xi_0^{opt})\right], \\ &&& np < p_0. \end{aligned}} \tag{29}$$

Other variants of the above arise, such as fixing n and optimizing both sensor locations and their price (equiv. covariance), but these are special cases of O_1 or O_2.

IV. Finite dimensional systems

Similar to the infinite dimensional case (8), a finite dimensional linear dynamical system is given by

$$\dot{x}(t) = Ax(t) + B_1 w(t) + B_2 u(t), \tag{30}$$

where the finite dimensional state $x \in \mathbb{R}^N$ and the matrices B_1, B_2 are of commensurate dimensions with the noise and control signals $w(t)$ and $u(t)$ being scalar signals.

While many sensor parametrization models appeared in the literature in the context of sensor selection, we consider a rather trivial parametrization; as such the i^{th} measurement has an output vector given by the unit vector $\mathbf{1}_i = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$ in $\mathbb{R}^N$

$$y_i(t) = \mathbf{1}_i x(t) + \nu_i(t), \quad i = 1, \ldots, n, \tag{31}$$

To appropriately parameterize the sensors by the sensor locations for the finite dimensional case, we use the notation

$$C_j(\theta_i) = \mathbf{1}_i, \quad \theta_i \in \Theta, \ i = 1, \ldots, N, \ j = 1, \ldots, n, \tag{32}$$

where C_j is the j^{th} device and θ_i is the location. The parameter space is similarly defined to be the space of all $1 \times N$ unit vectors $\mathbf{1}_i$ that render the pairs $(\mathbf{1}_i, A)$ detectable

$$\Theta = \{\mathbf{1}_i \in \mathbb{R}^{1\times N}, \mid (\mathbf{1}_i, A) \text{ detectabe}\}. \tag{33}$$

Please note that the parameter space has at most N elements.

To simplify further, it is assumed that all devices are identical, in the sense they have the same function on the state vector, differing on the accuracy of their readout, as quantified by the sensor covariance.

Assumption 4: All sensors, the single expensive sensor and the n inexpensive sensors, have the same output matrix,

$$C_j(\cdot) = C(\cdot), \quad j = 0, j = 1, \ldots, n.$$

The ideal expensive sensor has an output that is parameterized by the sensor location $\theta \in \Theta$ and given by

$$y_0(t;\theta) = C(\theta)x(t) + \nu_0(t), \tag{34}$$

where the sensor noise $\nu_0(t)$ has an associated covariance N_0. Please note that $\theta_i \in \Theta$ does not imply that $\theta_i = \mathbf{1}_i$; it simply means that the i^{th} sensing device $C(\theta_i)$ is an element of the parameter space Θ, or that $C(\theta_i)$ is equal to *any* of the unit vectors that yield detectability.

The n-dimensional output generated by the n inexpensive noisy sensing devices is given by

$$\mathbf{y}(t;\boldsymbol{\theta}) = \begin{bmatrix} C(\theta_1)x(t) + \nu_1(t) \\ \vdots \\ C(\theta_n)x(t) + \nu_n(t) \end{bmatrix} = \mathbf{C}(\boldsymbol{\theta}) + \boldsymbol{\nu}(t), \qquad (35)$$

where $\boldsymbol{\nu}(t) = \begin{bmatrix} \nu_1(t) & \dots & \nu_n(t) \end{bmatrix}^T$, $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_n\} \in \prod^n \Theta = \boldsymbol{\Theta}$, and $\mathbf{C}(\theta) = \text{diag}\{C(\theta_1), \dots, C(\theta_n)\}$. Similarly, the covariance in (33) is $\mathbf{N} = \text{diag}\{N_1, \dots, N_n\}$.

A. Optimal location of single expensive sensor

First, one selects the optimal sensor for the filter associated with the single expensive sensor. It is given by

$$\begin{aligned} \dot{\hat{x}}_0(t;\theta) &= A\hat{x}_0(t;\theta) + B_2 u(t) \\ &\quad + K_0(\theta)\left(y_0(t;\theta) - C(\theta)\hat{x}_0(t;\theta)\right), \quad \theta \in \Theta. \end{aligned} \qquad (36)$$

The N dimensional filter gain $K_0(\theta) = \Sigma_0(\theta)C^T(\theta)N_0^{-1}$ is derived from the now location-parameterized positive solution to the filter matrix algebraic Riccati equation

$$\begin{aligned} &A^T\Sigma_0(\theta) + \Sigma_0(\theta)A^T + B_1B_1^T \\ &-\Sigma_0(\theta)C^T(\theta)N_0^{-1}C(\theta)\Sigma_0(\theta) = 0, \ \ \theta \in \Theta. \end{aligned} \qquad (37)$$

The optimal sensor θ_0^{opt} for the single expensive sensor is

$$\theta_0^{opt} = \arg\inf_{\theta \in \Theta} \text{trace } [\Sigma_0(\theta)]. \qquad (38)$$

The optimal performance associated with the optimal single expensive sensor is given by

$$\boxed{J^{opt}(\theta_0^{opt}) = \text{trace } \left[\Sigma_0(\theta_0^{opt})\right].} \qquad (39)$$

B. Optimal number and location of inexpensive sensors

The state estimator associated with the n measurements (35), parameterized by the locations $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_n\}$ is

$$\begin{aligned} \dot{\hat{\mathbf{x}}}(t;\boldsymbol{\theta}) &= A\hat{\mathbf{x}}(t;\boldsymbol{\theta}) + B_2 u(t) \\ &\quad + \mathbf{K}(\boldsymbol{\theta})\left(\mathbf{y}(t;\boldsymbol{\theta}) - \mathbf{C}(\boldsymbol{\theta})\hat{\mathbf{x}}(t;\theta)\right), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}. \end{aligned} \qquad (40)$$

where the filter gain $\mathbf{K}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{C}^*(\boldsymbol{\theta})\mathbf{N}^{-1}$ is obtained from the location-parameterized positive solution to the filter matrix algebraic Riccati equation

$$\begin{aligned} &A\boldsymbol{\Sigma}(\boldsymbol{\theta}) + \boldsymbol{\Sigma}(\boldsymbol{\theta})A^T + B_1B_1^T \\ &-\boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{C}^T(\boldsymbol{\theta})\mathbf{N}^{-1}\mathbf{C}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}) = 0, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}. \end{aligned} \qquad (41)$$

For a fixed element $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the cost associated with (40) is

$$\begin{aligned} \mathbf{J}^{opt}(\boldsymbol{\theta}) &= \mathcal{E}\left[(x(t) - \hat{\mathbf{x}}(t;\boldsymbol{\theta}))(x(t) - \hat{\mathbf{x}}(t;\boldsymbol{\theta}))^T\right] \\ &= \text{trace } [\boldsymbol{\Sigma}(\boldsymbol{\theta})], \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}. \end{aligned} \qquad (42)$$

Similar to the infinite dimensional case, using (39), (42), the optimal number and location of the n inexpensive sensors is

$$\boxed{\begin{aligned} &\text{minimize} && n \\ &\text{subject to} && \text{trace } \left[\boldsymbol{\Sigma}(\boldsymbol{\theta}^{opt})\right] = \text{trace } \left[\Sigma_0(\theta_0^{opt})\right], \\ & && np < p_0. \end{aligned}} \qquad (43)$$

The associated algorithm is given by Algorithm 2 below.

V. Examples

Example 1 (advection-diffusion PDE): We consider (1) with

$$a(\xi) = 5 \times 10^{-3}\left(\frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} + 1 + 3\sin(3\pi\xi)\sin((\xi - 1)^2)\right)$$

Algorithm 2 Economic sensor optimization: fixed price

1: **initialize:** Find optimal location θ_0^{opt} of single expensive sensor via (38)
2: **initialize:** Select integer $n = 2$
3: **loop**
4: define parameter space $\boldsymbol{\Theta}^n = \prod^n \Theta$
5: find optimal sensor location $\boldsymbol{\theta}^{opt,n}$ of inexpensive sensors using
$$\theta^{opt,n} = \arg\inf_{\boldsymbol{\theta} \in \boldsymbol{\Theta}^n} \text{trace } [\boldsymbol{\Sigma}(\boldsymbol{\theta})]$$
6: compute price of n inexpensive sensors using
$$P^n_{\text{total}} = np$$
7: **if** $\left|\text{trace}\left[\boldsymbol{\Sigma}(\boldsymbol{\theta}^{opt,n})\right] - \text{trace}\left[\Sigma_0(\theta_0^{opt})\right]\right| \geq \varepsilon$ or $P^n_{\text{total}} \geq p_0$ **then**
8: $n \leftarrow n + 1$
9: **goto** 3
10: **else**
11: terminate
12: **end if**
13: **end loop**

---

n	N_i/N_0	$\mathbf{J}^{opt}(\boldsymbol{\theta})/J^{opt}(\theta_0^{opt})$	total price np
10	3.106	1	1.0366
100	31.04	1	0.1038
1000	310.3	1	0.0104

Table 1. Comparison with a priori selected sensor positions over a uniform grid, and with $J^{opt}(\theta_0^{opt}) = 0.1213$, $N_0 = 10^{-3}$, $\theta_0^{opt} = 0.76$.

with $\mu = 0.75L$, $\sigma = L/8$, $b = -10^{-2}$, $c = -3 \times 10^{-3}$. The initial condition is $x(0,\xi) = \sin(\pi(L-\xi)/L)\exp(-7(\xi - L)^2)$ and the spatial distribution of the noise is given by

$$b_1(\xi) = \sin(\pi(L-\xi)/L)\exp(-7(\xi - L).^2).$$

The spatial distribution of the single device in (2) and of the many inexpensive devices in (3) are identical and given by

$$\int_0^L c(\xi;\xi_i)\phi\, d\xi = \int_0^L \delta(\xi - \xi_i)\phi\, d\xi = \phi(\xi_i),$$

and this case the parameter set consists of all points that are not the zeros of the eigenfunctions of the spatial operator.

For the particular example, the expression in (28) is taken to be $p = k/N^2$ monetary units, where the constant k is selected for simplicity as $k = N_0^2$, and thus the total price in this case simplifies to

$$np = nk/N^2 = nN_0^2/N^2 = n/(N/N_0)^2$$

From Table 1, it is observed that it is cheaper to use $n = 1000$ sensors with a price of 1.041×10^{-5} monetary units/each than a single sensor placed at $\theta_0^{opt} = 0.76$ and with a price of 1 monetary units. Both filters have essentially the same performance but the one with a single sensor is 96.15 times more expensive. Another possibility is to use $n = 100$ inexpensive sensors with a total price of 0.1038 monetary units and which is 9.634 times cheaper than the single expensive sensor. Using $n = 10$ inexpensive sensors will not work, since their total price of 1.037 monetary units is above the price of a single expensive sensor. Figure 1 depicts the spatial distribution of $\hat{x}(t,\xi)$ for 4 different time instances.

case	norm
single sensor	0.0906
n sensors	0.0976

Table 2. $L_2(0,6;L_2(\Omega))$ norm of estimation error.

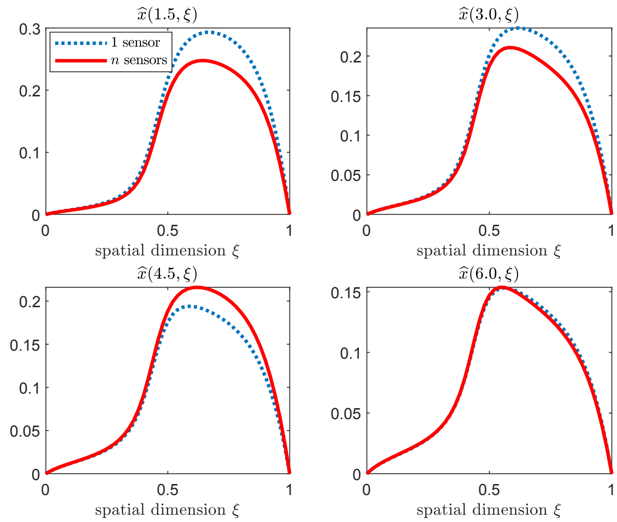

Fig. 1: Spatial distribution of $\widehat{x}(t,\xi)$ at different times.

Table 2 also summarizes the results of the estimation error norm. Both point to identical performance of the two filters, the price of the inexpensive sensors is significantly lower.

Example 2 (4th order system): Consider (30) with

$$A = \begin{bmatrix} -4 & 1 & 0 & 0 \\ 1 & -4 & 2 & 0 \\ 0 & 1 & -4 & 1 \\ 0 & 0 & 1 & -4 \end{bmatrix}, \quad Q = \mathbf{I}_{4\times 4}$$

and candidate sensors given by $C_i = e_i$, $i = 1,\ldots,4$. The single sensor is assumed to have $N_0 = 1\times 10^{-3}$. The case of using four sensors, thus having access to all states is only incorporated as a point of reference and does not represent a realistic situation. Another level of sensors is included, that of moderate price and moderate quality. It is assumed that if $n = 3$ inexpensive sensors are used, then $N_{i,3} = 30N_0$. If $n = 2$ sensors are used, they are moderately priced and moderately accurate with $N_{i,2} = 15N_0$. As the unrealistic case, of $n = 4$ sensors are used, they are extremely inexpensive and highly inaccurate. For reference, we have $N_{i,4} = 45.36N_0$.

Due to the low value of the state space dimension, we have performed an exhaustive search to obtain the filter performance and tabulate the results in Table 3. The best performance with 1.022% of that of a single sensor, is obtained for the three-sensor combination $\{C_1,C_2,C_3\}$ or $\{C_2,C_3,C_4\}$ with a total price 0.0033 . Close second is the two-sensor combination $\{C_2,C_3\}$ with a relative performance 1.0023% but with a higher total price of 0.0089. Despite the higher noise levels, the very noisy four-sensor combination yields the lowest possible cost of 0.0019, while being able to match the performance of the single expensive sensor.

VI. Conclusions

The effects of sensor price, assumed inversely proportional to noise covariance, were introduced as another level in the optimization for the optimal filtering of distributed and lumped parameter systems. When the total price of a sensor network was incorporated into the optimization problem, it revealed surprising results for their selection and placement. As observed in the examples, a network of inexpensive and noisy sensors may perform as well as a single accurate sensor with the added advantage of reduced cost.

The details of the convergence of the optimal sensor in terms of the approximation index, as presented in [10], [11] will be extended to include different spatial distributions and different noise covariance and price within a sensor network.

sensor selection	N_i/N_0	$\mathbf{J}^{opt}(\boldsymbol{\theta})/J^{opt}(\theta_0^{opt})$	total price
$\{C_1,C_2\}$	15	1.0355	0.0089
$\{C_1,C_3\}$	15	1.0181	0.0089
$\{C_1,C_4\}$	15	1.0927	0.0089
$\{C_2,C_3\}$	15	1.0023	0.0089
$\{C_2,C_4\}$	15	1.0105	0.0089
$\{C_3,C_4\}$	15	1.0458	0.0089
$\{C_1,C_2,C_3\}$	30	1.0022	0.0033
$\{C_1,C_2,C_4\}$	30	1.0170	0.0033
$\{C_1,C_3,C_4\}$	30	1.0242	0.0033
$\{C_2,C_3,C_4\}$	30	1.0022	0.0033
$\{C_1,C_2,C_3,C_4\}$	45.36	1.0000	0.0019

Table 3. Comparison with a priori selected sensor positions over a uniform grid, and with $J^{opt}(\theta_0^{opt}) = 0.4119$, $N_0 = 10^{-3}$, $\theta_0^{opt} = C_2$.

References

[1] R. F. Curtain and A. Ichikawa, "Optimal location of sensors for filtering for distributed systems," in *Distributed parameter systems: modelling and identification (Proc. IFIP Working Conf., Rome)*, 1976, pp. 236–255.

[2] M. Amouroux, J. P. Babary, and C. Malandrakis, "Optimal location of sensors for linear stochastic distributed parameter systems," in *Distributed parameter systems: modelling and identification (Proc. IFIP Working Conf., Rome)*, 1976, pp. 92–113.

[3] A. Ichikawa and E. P. Ryan, "Sensor and controller location problems for distributed parameter systems," *Automatica J. IFAC*, vol. 15, no. 3, pp. 347–352, 1979.

[4] C. G. Malandrakis, "Optimal sensor and controller allocation for a class of distributed parameter systems," *Internat. J. Systems Sci.*, vol. 10, no. 11, pp. 1283–1299, 1979.

[5] S. Omatu and J. H. Seinfeld, "Optimization of sensor and actuator locations in a distributed parameter system," *J. Franklin Inst.*, vol. 315, no. 5-6, pp. 407–421, 1983.

[6] K. Morris, "LQ-optimal actuator location and norm convergence of riccati operators," in *Proc. of the IEEE Conference on Decision and Control*, 9-11 Dec. 2008, pp. 209–214.

[7] K. Morris and M. A. Demetriou, "Using $\mathcal{H}_2$-control metrics for the optimal actuator location of infinite-dimensional systems," in *Proc. of the American Control Conf.*, 30 June-2 July 2010, pp. 4899–4904.

[8] D. Kasinathan and K. Morris, "Convergence of $\mathcal{H}_\infty$-optimal actuator locations," in *Proc. of the IEEE Conference on Decision and Control and European Control Conference*, 12-15 Dec. 2011, pp. 627–632.

[9] D. Kasinathan, K. Morris, and S. Yang, "Calculation of $\mathcal{H}_\infty$-optimal actuator location for distributed parameter systems," in *Proc. of the American Control Conference*, 2013, pp. 2086–2091.

[10] M. A. Demetriou and F. Fahroo, "Optimal sensor selection for optimal filtering of spatially distributed processes," in *Proc. of the 22nd International Symposium on Mathematical Theory of Networks and Systems*, University of Minnesota, MN, USA, July 12-15 2016.

[11] M. Zhang and K. Morris, "Effect of sensor noise on estimation of diffusion," *IFAC-PapersOnLine*, vol. 49, no. 8, pp. 60–65, 2016.

[12] ——, "Sensor choice for minimum error variance estimation," *IEEE Trans. Automat. Control*, vol. 63, no. 2, pp. 315–330, 2018.

[13] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*. NY: Springer-Verlag, 1983.

[14] R. F. Curtain and H. J. Zwart, *An Introduction to Infinite Dimensional Linear Systems Theory*. Berlin: Springer-Verlag, 1995.