THE INFINITE DIMENSIONAL OPTIMAL FILTERING PROBLEM WITH MOBILE AND STATIONARY SENSOR NETWORKS*

JOHN A. BURNS† AND CARLOS N. RAUTENBERG‡

Abstract. In this paper we introduce a framework to address filtering and smoothing with mobile sensor networks for distributed parameter systems. The main problem is formulated as the minimization of a functional involving the trace of the solution of a Riccati integral equation with constraints given by the trajectory of the sensor network. We prove existence and develop approximation of the solution to the Riccati equation in certain trace-class spaces. We also consider the corresponding optimization problem. Finally, we employ a Galerkin approximation scheme and implement a descent algorithm to compute optimal trajectories of the sensor network. Numerical examples are given for both stationary and moving sensor networks.

Key words. Optimal Sensor Networks, Riccati Integral Equations, Filtering

AMS subject classifications. 47B10, 45G15, 93B07

1. Introduction.

State estimation problems in a distributed parameter setting are a major source of engineering and applied science problems. In practice, almost all sources of information of a process (the temperature distribution inside a room, the concentration of a certain substance in the sea, etc.) are produced by devices capable of measuring a magnitude (temperature, concentration of a toxin, etc.) whose location is either a design variable or a known fact. We call *sensors* to these type of devices.

This research is motivated by applications to two distinct but related problem areas: **(1)** Determining optimal sensor/actuator locations for complex hybrid spatial systems to enhance tracking, estimation, information and effectiveness while limiting energy consumption. This area includes the sensor placement problem for energy efficient buildings. Here, the goal is to control room temperature and maximize solar energy production based on continuous changing dynamics determined by the weather, people inside the building and several sources of (stochastic) noise. The control/actuators include mechanical systems that position photovoltaic cells perpendicular to the incoming light beams and rotate glass panels on outside walls for capture (or deflection) of sunlight and heating and cooling systems inside the building. Due to the constant dynamic changes on weather conditions and occupancy of the building, feedback control based on sensor estimation is fundamental. Sensor tasks in this problem include appropriate location of incoming light, determination of weather conditions, temperature measurement and people location inside the building. **(2)** The second area is focused on optimal design and control of dynamic sensor fleets that include the dynamics of the mobile sensors. In this setting one must consider sensor dynamics as a constraint in the problems of optimal estimation and allow for information delays.

The first mathematically rigorous attempt to solve the optimal filtering problem for a wide class of linear distributed parameter systems was given by Bensoussan (see for example [7]). Also, this was the first attempt to provide a rigorous derivation of

*This research was supported in part by the Air Force Office of Scientific Research grant FA9550-10-1-0201 and by the DOE contract DE-EE0004261 under subcontract # 4345-VT-DOE-4261 from Penn State University.

† Interdisciplinary Center for Applied Mathematics, Virginia Tech, Blacksburg, VA (jaburns@vt.edu)

‡Department of Mathematics and Scientific Computing, Karl-Franzens-University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria. (carlos.rautenberg@uni-graz.at)

the Kalman-Bucy filter for linear distributed paramater systems (see Curtain's paper [16] for a detailed historical development of the optimal filtering problem).

In the case of finite dimensional systems, necessary conditions similar to those obtained later for the optimal filtering problem in the infinite dimensional setting were first developed by M. Athans ([2]) and these optimality conditions were obtained through Pontryagin's Maximum Principle.

A comprehensive treatment of the optimal sensor location problem was presented by Bensoussan (see [6]) where necessary conditions similar to the ones given previously by Athans (in [2]) were obtained for the infinite dimensional setting. Curtain obtained a general form of the Kalman-Bucy filter in infinite dimensions for general bounded generators of evolution operators (see [17]). Balakrishnan developed the Kalman-Bucy filter equations in another approach using integral equation theory.

The positioning of fixed sensors in order to achieve "maximal observability"of distributed parameters systems is fundamental for estimation and control. It should be noticed, however, that even in finite dimensional systems the term "maximal observability" is not always precisely defined. In the 1990's, Khapalov produced a series of papers (focused on questions of observability) on the design of optimal mobile sensors for a robust filtering problem and applied his results to parabolic and hyperbolic systems (see [33, 34, 35, 36, 37]).

The paper is organized as follows. The optimal filtering problem with PDE constraints which constitutes our main problem is given in §2. In section §3 we provide known results for the trace classes $\mathscr{I}_p(\mathscr{H})$ and prove necessary results associated with the Schatten classes $\mathscr{I}_p(X,\mathscr{H})$ which are fundamental for the description of the mappings $t\mapsto B(t)$ and $t\mapsto C(t)$. Results concerning Bochner integrability of the maps $t\mapsto BB^*(t)$ and $t\mapsto C^*C(t)$ are given in §4 together with results of compactness of the set of maps $t\mapsto C^*C(t)=C^*C(t,\bar{x}_i(t))$ with respect to the sensor trajectories $t\mapsto \bar{x}_i(t)$ (or locations in case of stationary sensor networks). Existence of optimal solutions to **Problem** $(\mathcal{P})$ is established for a wide class of output mappings in §5. Additionally, we also provide details on the existence and procedure to calculate the gradient of the solution of the Riccati equation w.r.t. map $t\mapsto C^*C(t)$. This fact makes it plausible to compute the gradient of the solution to the Riccati equation w.r.t. the controls of the trajectories $t\mapsto \bar{x}_i(t,\bar{x}_0^i,u_i(\cdot))$. Finally, we develop a computational scheme based on Galerkin type approximations and prove convergence of this algorithm in§6 and provide numerical examples to illustrate the theoretical results in §.

2. Problem Statement. Consider the convection-diffusion process in the n-dimensional unit cube $\Omega=(0,1)^n\subset\mathbb{R}^n$ defined by

$$\frac{\partial}{\partial t}T=(c^2\Delta+\mathbf{a}(x)\cdot\nabla)T+b(t,x)\eta(t), \tag{2.1}$$

where $\Delta=\sum_{i=1}^n\partial^2/\partial x_i^2$ is the $n-$dimensional Laplacian, $\mathbf{a}(x)\cdot\nabla=\sum_{i=1}^n a_i(x)\partial/\partial x_i$ is the convection operator and the functions $x\mapsto a_i(x)$ are smooth on $x\in\overline{\Omega}$. Here η is a real-valued Wiener process (a zero mean Gaussian process) and for each $t\in[0,t_f]$ the function $b(t,\cdot)$ belongs to $L^2(\Omega)$. We also assume boundary and initial conditions are of the form

$$T(t,x)\Big|_{\partial\Omega}=0, \qquad\qquad T(0,x)=T_0(x)+\xi,$$

where $T_0(\cdot)\in L^2(\Omega)$ and ξ is a $L^2(\Omega)$-valued gaussian random variable. Since the boundary $\partial\Omega$ is of Lipschitz class, the natural state space for the problem is $L^2(\Omega)$ and the domain of the differential operator $A = (c^2\Delta + \mathbf{a}(x)\cdot\nabla)$ is $\mathscr{D}(A) = H^2(\Omega)\cap H_0^1(\Omega)$ when $c > 0$.

Assume that one has p sensor-platforms (vehicles) moving in Ω, each with a sensor capable of measuring an average value of $T(t,x)$ within a fixed range of the location of the platform. Let $\bar{x}_i(t) \in \Omega$, $i = 1, 2, ..., p$ denote the position of the i^{th} sensor and time $t \in [0, t_f]$ and let

$$h_i(t) = \int_\Omega K(x, \bar{x}_i(t))T(t,x)\ \mathrm{d}x + \nu_i(t)\ . \tag{2.2}$$

denote the measured output which is the weighted average of the field $T(t,x)$ with weight $K(x, \bar{x}_i(t))$. Here, each $\nu = (\nu_1, \nu_2, \ldots, \nu_p)$ is a zero-mean *white* noise process and is uncorrelated with the process disturbance η. Observe that this definition also includes the one used by Khapalov (see [33], [34], [35], [36] and [37]) and offers a certain structure that allows for rigorous analysis when the dynamics of the vehicle network are included.

For a network of vehicle trajectories $\bar{x}_i(t) \in \Omega$, $i = 1, 2, ..., p$, we define the output map $C(t) : L_2(\Omega) \to \mathbb{R}^p$ by

$$C(t)\varphi = (C_1(t)\varphi,\ C_2(t)\varphi,\ C_3(t)\varphi, \ldots,\ C_p(t)\varphi)^T \in \mathbb{R}^p, \tag{2.3}$$

where

$$C_i(t)\varphi := \int_\Omega K(x, \bar{x}_i(t))\varphi(x)\ \mathrm{d}x. \tag{2.4}$$

The previous definitions can be used to formulate an abstract (infinite dimensional) model of the form

$$\dot{z}(t) = Az(t) + B(t)\eta(t) \in L^2(\Omega), \tag{2.5}$$

with $z(0) = z_0 + \xi$ and measured output

$$h(t) = C(t)z(t) + \nu(t). \tag{2.6}$$

Here, the state of the distributed parameter system is $z(t)(\cdot) = T(t,\cdot) \in L_2(\Omega)$. We will always assume that A is the infinitesimal generator of a C_0-semigroup of operators $S(t)$ over $L^2(\Omega)$. This is the standard abstract formulation of the convection-diffusion equations as a distributed parameter control system. This abstract model can be extended to include the case where η is a Wiener process with values in some separable Hilbert space X and $B(t) \in \mathscr{L}(X, L^2(\Omega))$ for each $t \in [0, t_f]$. This extension will be important for the work presented here.

One approach to optimal estimation is to observe that the variance equation for the optimal estimator is the (weak) solution to an infinite dimensional Riccati (partial) differential equation of the form

$$\dot{\Sigma}(t) = A\Sigma(t) + \Sigma(t)A^* + BR_2B^*(t) - \Sigma(t)(C^*R_1^{-1}C)(t)\Sigma(t)\ , \tag{2.7}$$

with initial condition $\Sigma(0) = \Sigma_0$. The operators $R_1(\cdot)$ and $R_2(\cdot)$ are the *incremental covariances* of the uncorrelated Wiener processes η and ν, respectively, and Σ_0 is the covariance operator of the $L^2(\Omega)$-valued Gaussian random variable ξ (see [16] and [6]). The expected value of $\|z(t) - \hat{z}(t)\|^2$ is the trace of the solution to the infinite dimensional Riccati equation at time t, i.e.,

$$\mathbb{E}\big\{\|z(t) - \hat{z}(t)\|^2\big\} = \mathrm{Tr}\,\Sigma(t),$$

where $\hat{z}(\cdot)$ is the stochastic $L^2(\Omega)$-valued process solution to the generalized Kalman-Bucy filter (see [6] and [7]). Therefore, for a sensor network defined by $t \mapsto C(t)$, the trace of the solution to the Riccati equation is a measure of error between the state and the state estimator. We use this measure to define the optimal sensor location problem.

In particular, we proceed as in [41] and consider the distributed parameter optimal control problem of finding $C_{opt}(t)$ to minimize

$$J(C(\cdot)) = \int_0^{t_f} \mathrm{Tr}\; Q(t)\; \Sigma(t)\; \mathrm{d}t \tag{2.8}$$

where $\Sigma(\cdot)$ is the mild solution of (2.7), $C(\cdot)$ is defined by (2.3)-(2.4), and for each $t \in [0, t_f]$, the operator $Q(t) : L_2(\Omega) \to L_2(\Omega)$ is a bounded linear operator. The (time-varying) map $Q(\cdot)$ allows one to weigh significant parts of the state estimate. For example, assume one has a control defined by a feedback operator $G(t) : Z \to \mathbb{R}^m$. If a re-constructed state (observer) is to be used in a feedback controller, then one might choose $Q(t) = G^*(t)G(t)$, in effect minimizing the error in the control produced by variance in the state estimate.

To complete the problem formulation for mobile sensors we include the sensor dynamics. Consider the case where $n = 3$. In particular, let $\bar{x}_i(t)$ denote the location of the i^{th} sensor platform at time t. The state of the sensor platform is defined by position and velocity so that $\bar{\theta}_i(t) = [\bar{x}_i(t), \dot{\bar{x}}_i(t)]^T$, where $\dot{\bar{x}}_i(t)$ represents the velocity of the sensor. The dynamics of the sensor are assumed to be governed by some controlled ordinary differential equations in $\mathbb{R}^6$ given by

$$\dot{\bar{\theta}}_i(t) = f_i(t, \bar{\theta}_i(t), u_i(t)), \tag{2.9}$$

$$\bar{\theta}_i(0) = \bar{\theta}_0^i = [\bar{x}_0^i, \dot{\bar{x}}_0^i]^T \tag{2.10}$$

where $u_i(\cdot)$ belongs to some admissible control set $\mathcal{U}$ and the initial conditions $\bar{\theta}_0^i$ belong to some compact set $\Theta_0 \subset \mathbb{R}^6$. For this paper we focus only on the position of the mobile sensor and observe that

$$\bar{x}_i(t) = [I \quad 0]\dot{\bar{\theta}}_i(t) = [I \quad 0] \begin{bmatrix} \bar{x}_i(t) \\ \dot{\bar{x}}_i(t) \end{bmatrix} =: M\dot{\bar{\theta}}_i(t) \tag{2.11}$$

is the output to the controlled system (2.9) - (2.10). It should be noted, however, that the theoretical results in this paper apply to the more general case. In particular, the position of the sensor is assumed to be given by $\bar{x}_i(t) = M\bar{\theta}_i(t)$ where $M \in \mathbb{R}^{3\times 6}$ is a constant matrix and $\bar{\theta}_i(t) \in \mathbb{R}^6$ for each $t \in [0, t_f]$. Observe that initial condition $\bar{x}_0^i$ for $\bar{x}_i(t)$ is given by $\bar{x}_0^i = M\bar{\theta}_0^i$ so that $X_0 = M\Theta_0 \subset \overline{\Omega}$ is the set of possible initial positions for the sensor location $\bar{x}_i(t)$. Finally, the moving sensor problem, for one sensor, is the following:

Problem $(\mathcal{P})$: Find $u_{opt}(\cdot) \in \mathcal{U}$ and $\bar{\theta}_0^{opt} \in \Theta_0$ that minimizes

$$J(\bar{\theta}^0, u(\cdot)) = \int_0^{t_f} \mathrm{Tr}\left(Q(t)\; \Sigma(t, \bar{x}(t))\right)\; \mathrm{d}t\;, \tag{2.12}$$

on the set $(\Theta_0, \mathcal{U})$, where $t \mapsto \Sigma(t, \bar{x}(t))$ is the solution to the constraint (2.7) and the output map $t \mapsto C(t)$ is of the form (2.4) and it is determined by the trajectory $t \mapsto \bar{x}(t) = M\bar{\theta}(t, \bar{\theta}_0, u)$.

The generalization of **Problem** $(\mathcal{P})$ for a finite number of sensors is straightforward and not stated here. Although, existence of solutions to this problem and approximations for the minimizers are also given in this paper.

There are several technical and computational challenges that must be addressed in order to solve **Problem** $(\mathcal{P})$ above. We cite the following issues:

(1) Since the variance equation is infinite dimensional, one must prove that the operator $\Sigma(t)$ is of trace class and integrable so that the cost functional (2.12) is well defined over the interval $[0, t_f]$. This can be a nontrival problem, but the results in [18], [20], [27], [28], [40], and [46] provide a background to develop the necessary structure.

(2) The approximation of the solution to the problem requires the introduction of numerical schemes. The basic theory and approximation schemes developed in [11], [13], [20], [27], [28], [32], [40], and [46] will be used as a starting points. We will consider the Galerkin approximation as the main tool for this problem.

3. Preliminaries and $\mathscr{I}_p(X, \mathscr{H})$ classes. Let X and $\mathscr{H}$ be separable complex Hilbert spaces. The space of bounded linear operators from X to $\mathscr{H}$ is denoted by $\mathscr{L}(X, \mathscr{H})$ and by $\mathscr{L}(\mathscr{H})$ if $X = \mathscr{H}$. If $A \in \mathscr{L}(X, \mathscr{H})$, then $\|A\|$ denotes the usual operator norm.

An operator $A \in \mathscr{L}(\mathscr{H})$, is said to be *non-negative* if $\langle Ax, x\rangle \geq 0$ for all $x \in \mathscr{H}$, *positive* if $\langle Ax, x\rangle > 0$ for all nonzero $x \in \mathscr{H}$ and *strictly positive* if there is a $c > 0$ such $\langle Ax, x\rangle \geq c\|x\|^2$ for all $x \in \mathscr{H}$. The notation $A \geq 0$, $A > 0$ and $A \gg 0$ is standard for non-negative, positive and strictly positive operators, respectively. If $A \geq 0$, and that both $\{\phi_n\}$ and $\{\psi_n\}$ are orthonormal bases of $\mathscr{H}$, then it follows that $\sum_n \langle \phi_n, A\phi_n\rangle = \sum_n \langle \psi_n, A\psi_n\rangle$ (we allow the case where both quantities are infinite). This observation motivates the definition of trace of an operator.

Definition 3.1. *If $A \geq 0$, then the trace of A is defined by* $\mathrm{Tr}\,(A) := \sum_{n=1}^{\infty} \langle \phi_n, A\phi_n\rangle$, *where $\{\phi_n\}_{n=1}^{\infty}$ is any orthonormal basis of $\mathscr{H}$.*

Each operator $A \in \mathscr{L}(\mathscr{H})$ admits a *polar decomposition* (see for example [44]) analogous to the decomposition $z = e^{i\mathrm{Arg}(z)}|z|$ when $z \in \mathbb{C}$. In particular, let $|A|$ be defined to be the unique non-negative operator such that $A = U|A|$, where U is the unique partial isometry such $\mathrm{Ker}\ U = \mathrm{Ker}\,|A|$. Since $|A| \geq 0$, then $|A|^p \geq 0$ for any $p \in \mathbb{N}$ and applying standard continuous functional calculus we can prove that $|A|^p \geq 0$ for any $1 \leq p < \infty$. Therefore, $\mathrm{Tr}\,(|A|^p)$ is well defined and leads to the following definition.

Definition 3.2. *Let $\mathscr{I}_p(\mathscr{H})$ for $1 \leq p < \infty$ (or simply $\mathscr{I}_p$ when the space $\mathscr{H}$ is understood) denote the set of all bounded operators over $\mathscr{H}$ such that* $\mathrm{Tr}\,(|A|^p) < \infty$. *If $A \in \mathscr{I}_p(\mathscr{H})$, then the $\mathscr{I}_p$-norm (or just the p-norm) of A is defined as* $\|A\|_p := \big(\mathrm{Tr}\,(|A|^p)\big)^{1/p} < \infty$.

The subspace of compact bounded linear operators acting on $\mathscr{H}$ is denoted by $\mathscr{I}_\infty(\mathscr{H})$ and, when $\mathscr{H}$ is understood, we simply use $\mathscr{I}_\infty$ for $\mathscr{I}_\infty(\mathscr{H})$. The norm in $\mathscr{I}_\infty$ is taken to be the operator norm, that is, $\|A\|_\infty := \|A\|$.

If $\mathscr{H}$ is a complex separable Hilbert space, then the linear space $\mathscr{I}_p$, endowed with the p-norm is a Banach space (see [45]). The classes $\mathscr{I}_1$ and $\mathscr{I}_2$ are called the space of Trace Class (or Nuclear) operators and the space of Hilbert-Schmidt operators, respectively. Actually, the space $\mathscr{I}_2$ is a Hilbert space under the inner product $\langle A, B\rangle_{\mathscr{I}_2} := \sum_{n=1}^{\infty} \langle A\phi_n, B\phi_n\rangle_{\mathscr{H}}$, where $A, B \in \mathscr{I}_2$ and $\{\phi_n\}_{n=1}^{\infty}$ is any orthonormal basis of $\mathscr{H}$. Note that $\langle A, A\rangle_{\mathscr{I}_2} = \sum_{n=1}^{\infty} \langle \phi_n, A^*A\phi_n\rangle_{\mathscr{H}}$. The operator $|A|$ is given by $|A| = \sqrt{A^*A}$, and the continuous functional calculus implies that

$|A|^2 = (\sqrt{A^*A})^2 = A^*A$. Consequently $\langle A, A\rangle_{\mathscr{I}_2} = \mathrm{Tr}\,(|A|^2) = \|A\|_2^2$.

It is well known (see for example [45]), that for $1 \le p \le \infty$, finite rank operators are dense (in the p-norm) in $\mathscr{I}_p$ and that $\mathscr{I}_p$ is a two-sided $*$-ideal in the ring $\mathscr{L}(\mathscr{H})$, i.e., $\mathscr{I}_p$ is a vector space and;

1) If $A \in \mathscr{I}_p$ and $B \in \mathscr{L}(\mathscr{H})$, then $AB \in \mathscr{I}_p$ and $BA \in \mathscr{I}_p$.

2) If $A \in \mathscr{I}_p$ then $A^* \in \mathscr{I}_p$.

It is also well known (see [29] for a proof) that if $1 \le p_1 < p_2 \le \infty$, and $A \in \mathscr{I}_{p_1}$ then $A \in \mathscr{I}_{p_2}$ and $\|A\|_{p_2} \le \|A\|_{p_1}$. Therefore, we have the continuous embedding: $\mathscr{I}_{p_1} \hookrightarrow \mathscr{I}_{p_2}$. As a result of this embedding, it follows by setting $p_2 = \infty$, that every operator in $\mathscr{I}_p$ is compact (See [22], [29] or [45]) and that $\|A\| \le \|A\|_p$ for all $1 \le p \le \infty$. We shall also need the following results (see [22], [29] and/or [45] for proof).

Lemma 3.3. *If $A \in \mathscr{I}_p$ with $1 \le p \le \infty$ and $B \in \mathscr{I}_q$ where $1/p + 1/q = 1$, then $AB, BA \in \mathscr{I}_1$ and*

$$\|AB\|_1 \le \|A\|_p \|B\|_q, \qquad \|BA\|_1 \le \|A\|_p \|B\|_q. \tag{3.1}$$

Moreover, $\|A\|_p = \|A^*\|_p$ and for any positive integer r we have $A^r \in \mathscr{I}_{p/r}$ and $\|A^r\|_{p/r} \le (\|A\|_p)^r$.

The trace is a continuous linear functional over $\mathscr{I}_1$ (see [22]). Consequently, if $A \in \mathscr{I}_1$, the value $\mathrm{Tr}\,(A) = \sum_{n=1}^{\infty} \langle \phi_n, A\phi_n\rangle$ does not depend on the choice of the orthonormal basis $\{\phi_n\}_{n=1}^{\infty}$. This result, combined with the previous Lemma, gives a simple characterization to the dual spaces of $\mathscr{I}_p$ (see [29]) given by the following proposition.

Proposition 3.4. *Let φ be a continuous linear functional over $\mathscr{I}_p$ with $1 < p \le \infty$, then there is an operator $A \in \mathscr{I}_q$ with $1/p + 1/q = 1$ such that $\varphi(X) = \mathrm{Tr}\,(AX)$, for all $X \in \mathscr{I}_p$, and $\|\varphi\|_{\mathscr{L}(X,\mathbb{C})} = \|A\|_q$. If φ is a bounded linear functional on $\mathscr{I}_1$, then there is a bounded linear operator $A \in \mathscr{L}(\mathscr{H})$ such that $\varphi(X) = \mathrm{Tr}\,(AX)$ for all $X \in \mathscr{I}_1$ and $\|\varphi\|_{\mathscr{L}(X,\mathbb{C})} = \|A\|$.*

The previous proposition implies that $(\mathscr{I}_p)^* \simeq \mathscr{I}_q$ when when $1 < p \le \infty$ and then $\mathscr{I}_p$ is reflexive when $1 < p < \infty$. Moreover $(\mathscr{I}_1)^* \simeq \mathscr{L}(\mathscr{H})$. If $A \in \mathscr{I}_\infty$, then it is well known (see [29]) that it has a norm convergent expansion given by $A(\cdot) = \sum_{n=1}^{\omega} s_n(A)\langle \phi_n, \cdot\rangle \psi_n$, with ω possibly infinite and $\{\phi_n\}_{n=1}^{\omega}$ and $\{\psi_n\}_{n=1}^{\omega}$ orthonormal sequences in $\mathscr{H}$. The elements of the sequence $\{s_n(A)\}_{n=1}^{\omega}$ are uniquely determined and called the *singular values of* A. In addition the singular values satisfy $s_n(A) \ge 0$ and $s_1(A) \ge s_2(A) \ge \cdots \ge 0$.

There are several equivalent ways to define the norm $\|A\|_p$ for an $A \in \mathscr{I}_p$. The following result uses the singular values of A and the results of the dual space of $\mathscr{I}_p$ to characterize $\|A\|_p$ (see [22] and [45]).

Proposition 3.5. *Let $A \in \mathscr{I}_p$ and $\{s_j(A)\}_{j=1}^{\omega}$ be its singular values and denote by $\mathscr{I}^0$ to the set of nonzero finite rank operators. Then, if $\frac{1}{p} + \frac{1}{q} = 1$, the norm $\|A\|_p$ satisfies*

$$\|A\|_p = \sup_{B \in \mathscr{I}^0} \frac{|\mathrm{Tr}\,(BA)\,|}{\|B\|_q} = \Big(\sum_{j=1}^{\omega} s_j^p(A)\Big)^{1/p}. \tag{3.2}$$

3.1. The Classes $\mathscr{I}_p(X, \mathscr{H})$. We are interested in studying input mappings of the form $B : I \to \mathscr{L}(X, \mathscr{H})$, for some real interval (commonly $I = [0, \tau]$ or $I = \mathbb{R}^+ = [0, \infty)$) where X and $\mathscr{H}$ are complex separable Hilbert spaces. Since X and $\mathscr{H}$ may not be the same, we need to define the spaces $\mathscr{I}_p$ for operators in

$\mathscr{L}(X,\mathscr{H})$. Some of the following results can be found on [52] but are included here for the sake of completeness.

Similarly as with the $\mathscr{L}(\mathscr{H})$ case, if $A \in \mathscr{L}(X,\mathscr{H})$ then it has a polar decomposition ([30]). That is, A can be written as $A = U|A|$ where $|A| \in \mathscr{L}(X)$ and $U \in \mathscr{L}(X,\mathscr{H})$ are the unique operators such that $|A| \geq 0$ and U is a partial isometry with $\text{Ker}\, U = \text{Ker}\, |A|$. Since $|A| \geq 0$, then for any $1 \leq p < \infty$ we observe that $|A|^p \geq 0$. This follows by the functional calculus, $\sigma(|A|) \subset [0,\infty)$ then $f(\lambda) = \lambda^p \geq 0$ defined as $f : \sigma(|A|) \to \mathbb{C}$ satisfies $f(\lambda) \geq 0$ and hence $f(|A|) = |A|^p \geq 0$. Therefore $\text{Tr}\,(|A|^p) = \sum_{n=1}^{\omega} \langle \phi_n, |A|^p \phi_n \rangle \geq 0$, with $\omega \leq \infty$, is independent (and could be finite or infinite) of the chosen orthonormal basis $\{\phi_n\}_{n=1}^{\omega}$ of X.

Definition 3.6. *Let X and $\mathscr{H}$ be separable complex Hilbert spaces. The set of all operators $A \in \mathscr{L}(X,\mathscr{H})$ such that* $\text{Tr}\,(|A|^p) < \infty$ *is denoted by $\mathscr{I}_p(X,\mathscr{H})$. That is, $A \in \mathscr{I}_p(X,\mathscr{H})$ if and only if $|A| \in \mathscr{I}_p(X)$.*

For $p = \infty$ we denote by $\mathscr{I}_\infty(X,\mathscr{H})$ the Banach space of compact operators in $\mathscr{L}(X,\mathscr{H})$ under the usual operator norm $\|A\|_{\mathscr{I}_\infty(X,\mathscr{H})} = \|A\|_{\mathscr{L}(X,\mathscr{H})} = \|A\|$. The sets $\mathscr{I}_p(X,\mathscr{H})$ are linear vector spaces. In fact, they are Banach Spaces when one uses $\text{Tr}\,(|A|^p) < \infty$ to define a norm. Although this result seems to be well known, we could not find a proof so we include the follow result for completeness.

Proposition 3.7. *The space $\mathscr{I}_p(X,\mathscr{H})$ with norm defined by*

$$\|A\|_{\mathscr{I}_p(X,\mathscr{H})} := \||A|\|_{\mathscr{I}_p(X)} = \left(\text{Tr}\,(|A|^p)\right)^{1/p} \tag{3.3}$$

is a Banach Space.

Proof. First we establish $\mathscr{I}_p(X,\mathscr{H})$ is a linear space. Let A_1 and A_2 be in $\mathscr{I}_p(X,\mathscr{H})$ and $A_i = U_i|A_i|$ be their polar decompositions. It follows that $|A_i| \in \mathscr{I}_p(X)$. Let $A_1 + A_2 = V|A_1 + A_2|$ be the polar decomposition of $A_1 + A_2$. Consequently, $|A_1 + A_2| = V^*U_1|A_1| + V^*U_2|A_2|$ and since $V^*U_1 \in \mathscr{L}(X)$ we have that $|A_1 + A_2| \in \mathscr{I}_p(X)$. This implies that $A_1 + A_2 \in \mathscr{I}_p(X,\mathscr{H})$. Also, if $\alpha \in \mathbb{C}$, then $|\alpha A_1| = |\alpha||A_1| \in \mathscr{I}_p(X)$, which implies that $\alpha A_1 \in \mathscr{I}_p(X,\mathscr{H})$ and this proves that $\mathscr{I}_p(X,\mathscr{H})$ is a linear space.

Next we establish that (3.3) defines a norm on $\mathscr{I}_p(X,\mathscr{H})$. By definition we have $\|A\|_{\mathscr{I}_p(X,\mathscr{H})} \geq 0$ and if $\|A\|_{\mathscr{I}_p(X,\mathscr{H})} = 0$, then $|A| = 0$ which yields $A = 0$. Also since $|\alpha A| = |\alpha||A|$ then it follows that $\|A\|_{\mathscr{I}_p(X,\mathscr{H})} = |\alpha|\||A|\|_{\mathscr{I}_p(X)} = |\alpha|\|A\|_{\mathscr{I}_p(X,\mathscr{H})}$. Finally, $|A_1+A_2| = V^*U_1|A_1|+V^*U_2|A_2|$ and each V, U_1 and U_2 is a partial isometry. This implies that $\|V^*U_i\| \leq 1$ for $i = 1,2$ and hence $\||A_1+A_2|\|_{\mathscr{I}_p(X)} \leq \||A_1|\|_{\mathscr{I}_p(X)} + \||A_2|\|_{\mathscr{I}_p(X)}$. We conclude that $\|A_1 + A_2\|_{\mathscr{I}_p(X,\mathscr{H})} \leq \|A_1\|_{\mathscr{I}_p(X,\mathscr{H})} + \|A_2\|_{\mathscr{I}_p(X,\mathscr{H})}$ and hence $(\text{Tr}\,(|A|^p))^{1/p}$ defines a norm.

In order to establish completeness, assume that $\{A_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathscr{I}_p(X,\mathscr{H})$. If $A_m - A_n = V_{mn}|A_m - A_n|$ is the polar decomposition of $A_m - A_n$, then we have

$$\begin{aligned}\|A_m - A_n\|_{\mathscr{L}(X,\mathscr{H})} &\leq \|V_{mn}\|_{\mathscr{L}(X,\mathscr{H})}\||A_m - A_n|\|_{\mathscr{L}(X)} \leq \||A_m - A_n|\|_{\mathscr{L}(X)} \\ &\leq \||A_m - A_n|\|_{\mathscr{I}_p(X)} = \|A_m - A_n\|_{\mathscr{I}_p(X,\mathscr{H})}.\end{aligned}$$

Thus, $\{A_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathscr{L}(X,\mathscr{H})$ and since this space is complete, it is convergent. Furthermore, since each A_n is compact, the limit is a compact operator, i.e., $\lim_{n\to\infty} A_n = A \in \mathscr{I}_\infty(X,\mathscr{H})$ in operator norm. This implies that $|A_n| \to |A|$ in operator norm as $n \to \infty$. The same holds for $|A_n|^p \to |A|^p$ for $p \in \mathbb{N}$ and we can use the continuous functional calculus to extend this to $1 \leq p < \infty$. All these claims follow from the continuity of $A \mapsto f(A)$ when $A \geq 0$ and f is continuous on the right hand complex semi-plane $\{z \in \mathbb{C} : \mathbf{Re}\, z \geq 0\}$ (see for example Halmos book [30]). First consider $f(\lambda) = \lambda^{1/2}$ and $B_n = A_n^* A_n$ which is self-adjoint and clearly $B_n \to B = A^*A$ if $A_n \to A$ and hence, $f(B_n) = |A_n| \to |A| = f(B)$. Secondly, consider the continuous function $g(\lambda) = \lambda^p$ on $\mathbb{C}$ for $p \geq 1$; since $|A_n| \geq 0$ and $|A| \geq 0$ and $|A_n| \to |A|$ in norm, $|A_n|^p \to |A|^p$ in norm.

Therefore, for any $\phi \in \mathscr{H}$, $\langle \phi, |A_n|^p \phi \rangle \to \langle \phi, |A|^p \phi \rangle$ as $n \to \infty$. Let $\{\phi_n\}_{n=1}^{\omega}$ be some orthonormal basis of $\mathscr{H}$. Since $\{A_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathscr{I}_p(X, \mathscr{H})$, then for any $N < \infty$,

$$\left(\sum_{k=1}^{N} \langle \phi_k, |A_n|^p \phi_k \rangle \right)^{1/p} \leq \| |A_n| \|_{\mathscr{I}_p(X)} = \|A_n\|_{\mathscr{I}_p(X,\mathscr{H})} \leq M < \infty,$$

where $M := \sup_{n \in \mathbb{N}} \|A_n\|_{\mathscr{I}_p(X,\mathscr{H})} < \infty$. Taking the limit as $n \to \infty$, we have

$$\left(\sum_{k=1}^{N} \langle \phi_k, |A|^p \phi_k \rangle \right)^{1/p} \leq M < \infty,$$

for any $N \in \mathbb{N}$. Therefore $\mathrm{Tr}\,(|A|^p) < \infty$, i.e., $A \in \mathscr{I}_p(X, \mathscr{H})$. ☐

In the case in which $X = \mathscr{H}$, we use $\mathscr{I}_p(\mathscr{H})$ or $\mathscr{I}_p$ as usual when the space is understood. We need to extend several previous results to the space $\mathscr{I}_p(X, \mathscr{H})$.

PROPOSITION 3.8. *Let $A \in \mathscr{I}_p(X, \mathscr{H})$ where X and $\mathscr{H}$ are separable complex Hilbert spaces and $1 \leq p \leq \infty$. Then $A \in \mathscr{I}_r(X, \mathscr{H})$ for all $p \leq r \leq \infty$. Also $A^* \in \mathscr{I}_p(\mathscr{H}, X)$ and $\|A\|_{\mathscr{I}_p(X,\mathscr{H})} = \|A^*\|_{\mathscr{I}_p(\mathscr{H},X)}$. If $A \in \mathscr{I}_{2p}(X, \mathscr{H})$ then $AA^* \in \mathscr{I}_p(\mathscr{H})$ and $\|AA^*\|_{\mathscr{I}_p(\mathscr{H})} \leq \|A\|_{\mathscr{I}_{2p}(X,\mathscr{H})} \|A^*\|_{\mathscr{I}_{2p}(\mathscr{H},X)}$.*

Proof. If $A \in \mathscr{I}_p(X, \mathscr{H})$, then $|A| \in \mathscr{I}_p(X)$. Hence $|A| \in \mathscr{I}_r(X)$ for any $p \leq r \leq \infty$ and this implies $A \in \mathscr{I}_r(X, \mathscr{H})$

Let $A = U|A| \in \mathscr{L}(X, \mathscr{H})$ be the polar decomposition of A, then the polar decomposition of the adjoint $A^* \in \mathscr{L}(\mathscr{H}, X)$ is given by $A^* = U^*|A^*|$ where $|A^*| = U|A|U^*$ (for a proof see [23]). Since $P_1 = U^*U \in \mathscr{L}(\mathscr{H})$ is a projection onto $(\mathrm{Ker}\, U)^{\perp} = (\mathrm{Ker}\, |A|)^{\perp} = \overline{(\mathrm{Range}|A|)}$ then $U^*U|A| = |A|$ and hence $|A^*|^p = U|A|^pU^*$ for $p \in \mathbb{N}$. By the continuous functional calculus $|A^*|^p = U|A|^pU^*$ holds for any $1 \leq p < \infty$.

Now, let $\{\phi_n\}_{n=1}^{\infty}$ be the orthonormal basis of $\mathscr{H}$ given by eigenvectors of $|A^*| \in \mathscr{L}(\mathscr{H})$. This is possible because $|A^*| \geq 0$ is compact and self-adjoint since A^* is compact. Then $\{\psi_n\}_{n=1}^{\infty} = \{U^*\phi_n\}_{n=1}^{\infty}$ is an orthonormal set (not necessarily a basis) in X. Since $\phi_n \in \mathrm{Range}(|A^*|)$ and $P_2 = UU^* \in \mathscr{L}(X)$ is a projection onto $\overline{\mathrm{Range}(|A^*|)}$, we observe that $\langle \psi_n, \psi_m \rangle_X = \langle UU^*\phi_n, \phi_m \rangle_{\mathscr{H}} = \langle \phi_n, \phi_m \rangle_{\mathscr{H}} = \delta_{n,m}$. Therefore,

$$\mathrm{Tr}\,(|A^*|^p) = \sum_n \langle \phi_n, U|A|^pU^*\phi_n \rangle_{\mathscr{H}} = \sum_n \langle \psi_n, |A|^p \psi_n \rangle_X \leq \mathrm{Tr}\,(|A|^p) < \infty.$$

Interchanging the roles of A and A^* in the proof, we obtain the reverse inequality $\mathrm{Tr}\,(|A|^p) \leq \mathrm{Tr}\,(|A^*|^p)$. This implies that $\|A\|_{\mathscr{I}_p(X,\mathscr{H})} = \|A^*\|_{\mathscr{I}_p(\mathscr{H},X)}$.

Now suppose that $A \in \mathscr{I}_{2p}(X, \mathscr{H})$. We observe that $|AA^*| = AA^* = U|A|^2U^* \geq 0$. Then $|AA^*|^p = U|A|^{2p}U^*$ for $p \in \mathbb{N}$ (since $U^*U|A| = |A|$ as we used before) and the continuous functional calculus extends this to any $1 \leq p < \infty$. Therefore

$$\mathrm{Tr}\,(|AA^*|^p) = \sum_n \left\langle \phi_n, U|A|^{2p}U^*\phi_n \right\rangle_{\mathscr{H}} = \sum \left\langle \psi_n, |A|^{2p}\psi_n \right\rangle_X \leq \mathrm{Tr}\,(|A|^{2p}) < \infty.$$

Finally, $\operatorname{Tr}(|A|^{2p}) = \||A|^p|A|^p\|_{\mathscr{I}_1(X)} \le \||A|^p\|_{\mathscr{I}_2(X)}\||A|^p\|_{\mathscr{I}_2(X)}$ and using the previous identies we have $\||A|^p\|_{\mathscr{I}_2(X)} = (\operatorname{Tr}(|A|^{2p}))^{1/2} = \|A\|^p_{\mathscr{I}_{2p}(X,\mathscr{H})} = \|A^*\|^p_{\mathscr{I}_{2p}(\mathscr{H},X)}$. Therefore $\operatorname{Tr}(|AA^*|^p) \le \|A\|^p_{\mathscr{I}_{2p}(X,\mathscr{H})}\|A^*\|^p_{\mathscr{I}_{2p}(\mathscr{H},X)}$ or

$$\|AA^*\|_{\mathscr{I}_p(\mathscr{H})} \le \|A\|_{\mathscr{I}_{2p}(X,\mathscr{H})}\|A^*\|_{\mathscr{I}_{2p}(\mathscr{H},X)}.$$

□

4. The Input and Output Maps. In this section, we provide conditions on the maps $t \mapsto B(t)$ and $t \mapsto C(t)$ so that the maps $t \mapsto BB^*(t)$ and $t \mapsto C^*C(t)$ are $\mathscr{I}_p$-valued, Bochner measurable and regular enough (in time) to guarantee existence of σ, a $\mathscr{I}_p$-valued solution to the Riccati equation. Additionally, we study compacity properties of the sets $\{t \mapsto C^*C(t, \bar{x}_i(t))\}$ determined by the trajectories of the sensors $\{t \mapsto \bar{x}_i(t) : \bar{x}_i(t) = M\bar{\theta}_i(t, \bar{\theta}^i_0, u_i(\cdot))\}$.

4.1. The Input Map $t \mapsto B(t)$. LEMMA 4.1. *Let X and $\mathscr{H}$ be separable complex Hilbert spaces, I be a real interval (bounded or unbounded) and let $B : I \to \mathscr{I}_{2p}(X, \mathscr{H})$ with $1 \le p \le \infty$. Then,*
i. *If $B(\cdot) \in L^2_{loc}(I; \mathscr{I}_{2p}(X, \mathscr{H}))$ then $BB^*(\cdot) \in L^1_{loc}(I; \mathscr{I}_p(\mathscr{H}))$.*
ii. *If $B(\cdot) \in \mathscr{C}(I; \mathscr{I}_{2p}(X, \mathscr{H}))$, then $BB^*(\cdot) \in \mathscr{C}(I; \mathscr{I}_p(\mathscr{H}))$.*

Proof. We first prove the measurability $BB^*(\cdot)$. Since $B(t) \in \mathscr{I}_{2p}(X, \mathscr{H})$ for each $t \in I$, then $B^*(t) \in \mathscr{I}_{2p}(\mathscr{H}, X)$ and $BB^*(t) := B(t)B^*(t) \in \mathscr{I}_p(\mathscr{H})$ by Proposition 3.8; then $BB^* : I \to \mathscr{I}_p(\mathscr{H})$. Since $B : I \to \mathscr{I}_{2p}(X, \mathscr{H})$ is measurable, there is a sequence of simple $\mathscr{I}_{2p}(X, \mathscr{H})$-valued functions $B_n(t) = \sum_{k=1}^n b_k(n)\chi_{E_k(n)}(t)$ for $t \in I$ with $b_k(n) \in \mathscr{I}_{2p}(X, \mathscr{H})$ and $E_k(n) \subset I$ measurable sets for $1 \le k \le n \in \mathbb{N}$ satisfying $\|(B - B_n)(t)\|_{\mathscr{I}_{2p}(X,\mathscr{H})} \to 0$ a.e. in $t \in I$ as $n \to \infty$. Note that $B_nB_n^*(t) := B_n(t)B_n^*(t)$ is a simple $\mathscr{I}_p(\mathscr{H})$-valued function and

$$\begin{aligned}\|(BB^* - B_nB_n^*)(t)\|_{\mathscr{I}_p(\mathscr{H})} &\le \|B(B^* - B_n^*)(t)\|_{\mathscr{I}_p(\mathscr{H})} + \|(B - B_n)B_n^*(t)\|_{\mathscr{I}_p(\mathscr{H})} \\ &= \|(B - B_n)B^*(t)\|_{\mathscr{I}_p(\mathscr{H})} + \|(B - B_n)B_n^*(t)\|_{\mathscr{I}_p(\mathscr{H})} \\ &\le \|(B - B_n)(t)\|_{\mathscr{I}_{2p}(X,\mathscr{H})}\Big(\|B^*(t)\|_{\mathscr{I}_{2p}(\mathscr{H},X)} + \|B_n^*(t)\|_{\mathscr{I}_{2p}(\mathscr{H},X)}\Big).\end{aligned}$$

Hence $\|(BB^* - B_nB_n^*)(t)\|_{\mathscr{I}_p(\mathscr{H})} \to 0$ a.e. in $t \in I$ as $n \to \infty$ and this implies that $BB^* : I \to \mathscr{I}_p(\mathscr{H})$ is measurable.

We now turn the attention to **i.**. Note that if $B(\cdot) \in L^2_{loc}(I; \mathscr{I}_{2p}(X, \mathscr{H}))$, then $\int_C \|B(t)\|^2_{\mathscr{I}_{2p}(X,\mathscr{H})}\,\mathrm{d}t < \infty$, for any compact interval $C \subset I$. From Proposition 3.8, we obtain $\|B(t)B^*(t)\|_{\mathscr{I}_p(\mathscr{H})} \le \|B(t)\|_{\mathscr{I}_{2p}(X,\mathscr{H})}\|B^*(t)\|_{\mathscr{I}_{2p}(\mathscr{H},X)} = \|B(t)\|^2_{\mathscr{I}_{2p}(X,\mathscr{H})}$, which implies that

$$\int_C \|BB^*(t)\|_{\mathscr{I}_p(\mathscr{H})}\,\mathrm{d}t \le \int_C \|B(t)\|^2_{\mathscr{I}_{2p}(X,\mathscr{H})}\,\mathrm{d}t < \infty.$$

Hence, $BB^*(\cdot) \in L^1_{loc}(I; \mathscr{I}_p(\mathscr{H}))$.

Let t and s belong to some compact interval $C \subset I$. It follows that

$$\begin{aligned}\|BB^*(t) - BB^*(s)\|_{\mathscr{I}_p(\mathscr{H})} &\le \|B(t)(B^*(t) - B^*(s))\|_{\mathscr{I}_p(\mathscr{H})} + \|(B(t) - B(s))B^*(s)\|_{\mathscr{I}_p(\mathscr{H})} \\ &\le \|(B(t) - B(s))B^*(t)\|_{\mathscr{I}_p(\mathscr{H})} + \|(B(t) - B(s))B^*(s)\|_{\mathscr{I}_p(\mathscr{H})} \\ &\le \|B(t) - B(s)\|_{\mathscr{I}_{2p}(X,\mathscr{H})}\Big(\|B^*(t)\|_{\mathscr{I}_{2p}(\mathscr{H},X)} + \|B^*(s)\|_{\mathscr{I}_{2p}(\mathscr{H},X)}\Big) \\ &\le M\|B(t) - B(s)\|_{\mathscr{I}_{2p}(X,\mathscr{H})},\end{aligned}$$

where $M = \sup_{t,s\in C}\left(\|B^*(t)\|_{\mathscr{I}_{2p}(\mathscr{H},X)} + \|B^*(s)\|_{\mathscr{I}_{2p}(\mathscr{H},X)}\right) < \infty$. Hence, if $B(\cdot) \in \mathscr{C}(I;\mathscr{I}_{2p}(X,\mathscr{H}))$, then $BB^*(\cdot) \in \mathscr{C}(I;\mathscr{I}_p(\mathscr{H}))$, hence **ii.** is proven. □

In many realistic control problems the number of inputs and sensed outputs are typically finite (and often small in numbers). In this case, the previous results become stronger when X is finite dimensional.

PROPOSITION 4.2. *Let X is a finite dimensional complex Hilbert space and $\mathscr{H}$ a separable complex Hilbert space. If $A \in \mathscr{L}(X,\mathscr{H})$, then $A \in \mathscr{I}_1(X,\mathscr{H})$.*

Proof. If $A = U|A|$ is the polar decomposition of A, then $|A| \in \mathscr{L}(X)$. Since X is finite dimensional, $|A| \in \mathscr{I}_1(X)$ and this implies that $A \in \mathscr{I}_1(X,\mathscr{H})$. □

As expected, for the same X and $\mathscr{H}$ as above, the same result holds for any operator $A \in \mathscr{L}(\mathscr{H},X)$. In this case $A \in \mathscr{I}_1(\mathscr{H},X)$, since A is of finite rank.

LEMMA 4.3. *Let X be a finite dimensional complex Hilbert space, $\mathscr{H}$ be a separable complex Hilbert space and I be a real interval (bounded or unbounded).*
i. *If $B(\cdot) \in L^2_{loc}(I;\mathscr{L}(X,\mathscr{H}))$, then $BB^*(\cdot) \in L^1_{loc}(I;\mathscr{I}_1(\mathscr{H}))$.*
ii. *If $B(\cdot) \in \mathscr{C}(I;\mathscr{L}(X,\mathscr{H}))$, then $BB^*(\cdot) \in \mathscr{C}(I;\mathscr{I}_1(\mathscr{H}))$.*

Proof. We first prove that for each $1 \le p < \infty$ there is an $c > 0$ such that $\|A\|_{\mathscr{I}_p(X,\mathscr{H})} \le c\|A\|_{\mathscr{L}(X,\mathscr{H})}$ for each $A \in \mathscr{L}(X,\mathscr{H})$. Since all norms are equivalent in finite dimensions and $\mathscr{L}(X)$ is finite dimensional there is a $c > 0$ such that $\||A|\|_{\mathscr{I}_p(X)} \le c\||A|\|_{\mathscr{L}(X)}$ since $|A| \in \mathscr{L}(X)$. Since $U^*A = |A|$ (where $A = U|A|$ is the polar decomposition of A), then $\||A|\|_{\mathscr{L}(X)} \le \|A\|_{\mathscr{L}(X,\mathscr{H})}$ this implies that $\||A|\|_{\mathscr{I}_p(X)} \le m\|A\|_{\mathscr{L}(X,H)}$. By definition $\|A\|_{\mathscr{I}_p(X,\mathscr{H})} = \||A|\|_{\mathscr{I}_p(X)}$, the claimed result follows.

Proposition 4.2 implies that $B(\cdot)$ is $\mathscr{I}_1(X,\mathscr{H})$-valued and hence also $\mathscr{I}_2(X,\mathscr{H})$-valued, since $\mathscr{I}_1(X,\mathscr{H}) \hookrightarrow \mathscr{I}_2(X,\mathscr{H})$. We now prove $B(\cdot)$is measurable as a $\mathscr{I}_2(X,\mathscr{H})$-valued function. Since $B : I \to \mathscr{L}(X,\mathscr{H})$ is Bochner measurable, then there is a sequence of simple functions $B_n : I \to \mathscr{L}(X,\mathscr{H})$ such that $\|B(t) - B_n(t)\|_{\mathscr{L}(X,\mathscr{H})} \to 0$ a.e. in $t \in I$ as $n \to \infty$. Since $\|B(t) - B_n(t)\|_{\mathscr{I}_2(X,\mathscr{H})} \le c\|B(t) - B_n(t)\|_{\mathscr{L}(X,\mathscr{H})}$ for some $c > 0$, we observe that $\|B(t) - B_n(t)\|_{\mathscr{I}_2(X,\mathscr{H})} \to 0$ a.e. in $t \in I$ as $n \to \infty$. This implies $B : I \to \mathscr{I}_2(X,\mathscr{H})$ is Bochner measurable.

If $B(\cdot) \in L^2_{loc}(I;\mathscr{L}(X,\mathscr{H}))$, since $\|B(t)\|_{\mathscr{I}_2(X,\mathscr{H})} \le c\|B(t)\|_{\mathscr{L}(X,\mathscr{H})}$ for each $t \in I$, then it follows that $B(\cdot) \in L^2_{loc}(I;\mathscr{I}_2(X,\mathscr{H}))$. Therefore, by Lemma 4.1 we observe that $BB^*(\cdot) \in L^1_{loc}(I;\mathscr{I}_1(\mathscr{H}))$ and **i.** is proven.

We now turn the attention to **ii.**. If $B(\cdot) \in \mathscr{C}(I;\mathscr{L}(X,\mathscr{H}))$, since $\|B(t)\|_{\mathscr{I}_2(X,H)} \le c\|B(t)\|_{\mathscr{L}(X,H)}$ for some $c > 0$ and for each $t \in I$, we observe that $B(\cdot) \in \mathscr{C}(I;\mathscr{I}_2(X,\mathscr{H}))$ and this implies by Lemma 4.1 that $BB^*(\cdot) \in \mathscr{C}(I;\mathscr{I}_1(\mathscr{H}))$.
□

4.2. The Output Map $t \mapsto C(t)$. For the case of p sensors, the operator $C(t)$ in (2.3) is given by

$$C(t)\varphi := (C_1(t)\varphi,\ C_2(t)\varphi, \dots,\ C_n(t)\varphi)^T$$

for $\varphi(\cdot) \in L^2(\Omega)$, where

$$C_i(t)\varphi := \int_\Omega K_i(t,x)\varphi(x)\,\mathrm{d}x,$$

for kernels $K_i : I \times \Omega \to \mathbb{C}$, for $i = 1, 2, \dots, m$. We assume that $K_i(t,\cdot) \in L^2(\Omega)$ for each $t \in I$, so that $C_i(t) \in \mathscr{L}(L^2(\Omega),\mathbb{C})$ and $C(t) \in \mathscr{L}(L^2(\Omega),\mathbb{C}^p)$ for each $t \in I$.

PROPOSITION 4.4. *If the map $t \mapsto K_i(t,\cdot)$ belongs to $L^\infty_{loc}(I;L^2(\Omega))$ for $i = 1, 2, \dots, p$ where I is a real (bounded or unbounded) interval, then it follows that $(C^*C)(\cdot) \in L^\infty_{loc}(I;\mathscr{I}_1(L^2(\Omega)))$.*

Proof. Suppose I is a bounded interval. Since $t \mapsto K_i(t,\cdot) \in L^\infty(I; L^2(\Omega))$, there is $0 < M < \infty$ such that $\operatorname{ess\,sup}_{t\in I} \|K_i(t,\cdot)\|_{L^2(\Omega)} < M$ for all i. Also, there are sequences $\left\{K_i^k\right\}_{k=1}^\infty$ of simple $L^2(\Omega)$-valued functions $K_i^k(t,x) = \sum_{j=1}^k P_i^j(k,x)\chi_{I_i^j(k)}(t)$ that converge point-wise a.e. to K_i with $P_i^j(k,\cdot) \in L^2(\Omega)$ for $1 \le i \le n$ and $1 \le j \le k$. Define the sequence $\left\{C_i^k(\cdot)\right\}_{k=1}^\infty$ by

$$(C_i^k(t)\varphi) = \int_\Omega K_i^k(t,x)\varphi(x)\,\mathrm{d}x = \sum_{j=1}^k \left(\int_\Omega P_i^j(k,x)\varphi(x)\,\mathrm{d}x\right)\chi_{I_i^j(k)}(t),$$

for $\varphi(\cdot) \in L^2(\Omega)$. For each i, $t \mapsto C_i^k(t)$ is a simple $\mathscr{L}(L^2(\Omega),\mathbb{C})$-valued function and satisfies

$$\begin{aligned}|\big(C_i(t) - C_i^k(t)\big)\varphi| &\le \int_\Omega |K_i(t,x) - K_i^k(t,x)||\varphi(x)|\,\mathrm{d}x \\ &\le \|K_i(t,\cdot) - K_i^k(t,\cdot)\|_{L^2(\Omega)}\|\varphi\|_{L^2(\Omega)}.\end{aligned}$$

Therefore, $\|C_i(t) - C_i^k(t)\|_{\mathscr{L}(L^2(\Omega),\mathbb{C})} \le \|K_i(t,\cdot) - K_i^k(t,\cdot)\|_{L^2(\Omega)}$ and hence each $t \mapsto C_i(t)$ is a Bochner measurable $\mathscr{L}(L^2(\Omega,\mathbb{C}))$-valued function with domain I. The same inequality shows that $C_i(\cdot) \in L^\infty(I, \mathscr{L}(L^2(\Omega);\mathbb{C}))$. This implies that $C(\cdot) \in L^\infty(I, \mathscr{L}(L^2(\Omega);\mathbb{C}^p))$ and $C^*(\cdot) \in L^\infty(I, \mathscr{L}(\mathbb{C}^p; L^2(\Omega)))$. In case I is unbounded, the above holds for each $I \cap [\alpha,\beta]$ and then in any case $C(\cdot) \in L^\infty_{loc}(I, \mathscr{L}(L^2(\Omega);\mathbb{C}^p))$ and $C^*(\cdot) \in L^\infty_{loc}(I, \mathscr{L}(\mathbb{C}^p; L^2(\Omega)))$.

Since $C^*(\cdot) \in L^\infty_{loc}(I, \mathscr{L}(\mathbb{C}^p; L^2(\Omega)))$ and $\mathbb{C}^p$ is finite dimensional we have that $C^*(\cdot) \in L^2_{loc}(I, \mathscr{L}(\mathbb{C}^n; L^2(\Omega)))$. If $[a,b] \subset I$, then

$$\int_{[a,b]} \|C^*(t)\|^2_{\mathscr{L}(\mathbb{C}^n;L^2(\Omega))}\,\mathrm{d}t \le (b-a)\Big(\operatorname*{ess\,sup}_{t\in[a,b]} \|C^*(t)\|_{\mathscr{L}(\mathbb{C}^n;L^2(\Omega))}\Big)^2.$$

Therefore, $C^*(\cdot) \in L^2_{loc}(I, \mathscr{L}(\mathbb{C}^p; L^2(\Omega)))$ and Lemma 4.3 implies that $C^*C(\cdot) \in L^\infty_{loc}(I; \mathscr{I}_1(L^2(\Omega)))$. □

4.2.1. The Stationary Network Case. Assume that one has p sensor-platforms *fixed* in $\overline{\Omega}$ compact, each with a sensor capable of measuring a weighted average value of the process of interest. We will denote the position of a sensor with a "bar" on top of the variable. For example, the position of the first sensor we will be denoted by "$\bar{x}_1$". We consider two important examples.

EXAMPLE 1. *Let $\bar{x}_i \in \overline{\Omega}$ and suppose the sensor measures an average value of each $\varphi(\cdot) \in L^2(\Omega)$ within a fixed radius $\delta > 0$ of the location $\bar{x}_i \in \overline{\Omega}$. In this case*

$$C_i(\bar{x}_i)\varphi = \int_\Omega \chi_\delta(x,\bar{x}_i)\varphi(x)\,\mathrm{d}x,$$

where $\chi_\delta(x,y) = 1$ if $\|x-y\|_{\mathbb{R}^n} < \delta$ and $\chi_\delta(x,y) = 0$ otherwise.

EXAMPLE 2. *Let $\bar{x}_i \in \overline{\Omega}$, $k > 0$ and $K(x) = e^{-k\|x-\bar{x}_i\|^2_{\mathbb{R}^n}}$, so that*

$$C_i(\bar{x}_i)\varphi = \int_\Omega e^{-k\|x-\bar{x}_i\|^2_{\mathbb{R}^n}}\varphi(x)\,\mathrm{d}x.$$

Here $C_i(\bar{x}_i)$ is a Gaussian-type kernel sensor.

The general form of the output operator when p sensors are placed in $\overline{\Omega}$ is given by

$$C(t,\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_p)\varphi = (C_1(t,\bar{x}_1)\varphi,\ C_2(t,\bar{x}_2)\varphi,\ldots,\ C_n(t,\bar{x}_p)\varphi)^T \tag{4.1}$$

for $\varphi(\cdot) \in L^2(\Omega)$, with

$$C_i(t, \bar{x}_i)\varphi = \int_\Omega K_i(t, x, \bar{x}_i)\varphi(x)\,\mathrm{d}x. \tag{4.2}$$

In applications, if the sensor is moved from the position $\bar{x}_i$ to a position $\bar{x}_i + \Delta\bar{x}_i$ with $\|\Delta\bar{x}_i\|_{\mathbb{R}^n} << 1$ we expect that the measurement at $\bar{x}_i$ is close to the one in $\bar{x}_i + \Delta\bar{x}_i$. This motivates the following definition.

Definition 4.5 (Continuity w.r.t. Location). *Let I be a real interval and $C : I \times \overline{\Omega}^p \to \mathscr{L}(L^2(\Omega); \mathbb{C}^p)$, be of the form (4.1). We say that C is* continuous with respect to location *if there is a continuous function $g : \mathbb{R}^+ \to \mathbb{R}^+$ such that $g(0) = 0$ and*

$$\|K_i(t, \cdot, y) - K_i(t, \cdot, z)\|_{L^2(\Omega)} \le g(\|y - z\|_{\mathbb{R}^n}), \qquad \forall t \in I \qquad \forall y, z \in \overline{\Omega},$$

and for $i = 1, 2, \ldots, p$. Here, $t \mapsto K_i(t, \cdot, x) \in L^2_{loc}(I; L^2(\Omega))$ for $i = 1, 2, \ldots, p$ and for each $x \in \overline{\Omega}$, K_i is the kernel of the integral representation for C_i in (4.2).

Since the solution of Riccati equation can be regarded a function of the mapping $t \mapsto C^*C(t, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p)$ we are interested in properties of the set $\{t \mapsto C^*C(t, \bar{\mathbf{x}}) : \text{ for } \bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p) \in X_0^p\}$, where X_0 is some compact set of interest that belongs to $\overline{\Omega}$. Since the stationary sensor network is just an example of the moving sensor network, the proof is provided in what follows on Lemma 4.6.

The condition (Definition 4.5) that states that there exist a continuous $g : \mathbb{R}^+ \to \mathbb{R}^+$ satisfying $g(0) = 0$ and $\|K_i(t, \cdot, y) - K_i(t, \cdot, y)\|_{L^2(\Omega)} \le g(\|y - z\|_{\mathbb{R}^n})$, for all $t \in I$ and all $y, z \in \overline{\Omega}$ is satisfied by the kernels we considered in the Examples 1 and 2 as we now prove.

Consider the sensor in Example 1 and let $y, z \in \overline{\Omega}$. Then

$$\int_{\mathbb{R}^n} |\chi_\delta(x, y) - \chi_\delta(x, z)|\,\mathrm{d}x = \int_{\mathbb{R}^n} |\chi_\delta(x, 0) - \chi_\delta(x, z - y)|\,\mathrm{d}x = 2m\left\{B_\delta(z - y) \setminus B_\delta(0)\right\},$$

where m refers to the Lebesgue measure in n dimensions and $B_\delta(x)$ is the closed $n-$dimensional ball of radius δ centered at x. The function $\Omega \ni w \mapsto m\left\{B_\delta(w) \setminus B_\delta(0)\right\}$ depends only on the norm of w, i.e., $G(\|w\|_{\mathbb{C}^n}) = 2m\left\{B_\delta(w) \setminus B_\delta(0)\right\}$. The function G is clearly monotonically increasing and satisfies $G(0) = m\emptyset = 0$. The continuity of $x \mapsto G(x)$ can be easily checked using the Lebesgue Dominated Convergence Theorem on the integral representation above. Since $|\chi_\delta(x, y) - \chi_\delta(x, z)|^2 = |\chi_\delta(x, y) - \chi_\delta(x, z)|$, then $\int_\Omega |\chi_\delta(x, y) - \chi_\delta(x, z)|\,\mathrm{d}x \le m(\Omega)(\int_{\mathbb{R}^n} |\chi_\delta(x, y) - \chi_\delta(x, z)|\,\mathrm{d}x)^{1/2}$ which yields

$$\int_\Omega |\chi_\delta(x, y) - \chi_\delta(x, z)|^2\,\mathrm{d}x \le m(\Omega)\big(G(\|y - z\|_{\mathbb{R}^n})\big)^{1/2}.$$

For the sensor in the Example 2 we consider $K(x, \bar{x}_i) = e^{-k\|x - \bar{x}_i\|^2_{\mathbb{R}^n}}$ for some $k > 0$. Since $t \mapsto e^{-kt} : [0, \infty) \to [0, 1]$ is Lipschitz continuous, it follows that

$$|K(x, y) - K(x, z)| \le C\Big|\|x - y\|^2_{\mathbb{R}^n} - \|x - z\|^2_{\mathbb{R}^n}\Big|.$$

The Law of Cosines yield

$$\|x - y\|^2_{\mathbb{R}^n} - \|x - z\|^2_{\mathbb{R}^n} = \|y - z\|^2_{\mathbb{R}^n} - 2\|y - z\|_{\mathbb{R}^n}\|x - z\|_{\mathbb{R}^n}\cos(\theta_z),$$

for some $\theta_z \in [-\pi, \pi]$ and since $\operatorname{diam}(\Omega) = \sup_{v, w \in \Omega} \|v - w\|_{\mathbb{R}^n} < \infty$, it follows that $|\|x - y\|^2_{\mathbb{R}^n} - \|x - z\|^2_{\mathbb{R}^n}| \le 3\operatorname{diam}(\Omega)\|y - z\|_{\mathbb{R}^n}$. Thus, we have

$$\int_\Omega |K(x, y) - K(x, z)|^2\,\mathrm{d}x \le \Big(9C^2\,(\operatorname{diam}(\Omega))^2 m(\Omega)\Big)\|y - z\|^2_{\mathbb{R}^n}.$$

4.2.2. The Moving Network Case. Here, we assume that $I = [0, \tau]$ for some fixed $\tau > 0$ and that that $\Omega \subset \mathbb{R}^3$ is open and bounded. Thus, $\overline{\Omega}$ is compact. Let $\bar{x}_i(t) \in \overline{\Omega}$ for each $t \in I$ be the position of the ith sensor at time t. In this case the general form of the output map for a moving sensor platform is given by

$$C_i(t, \bar{x}_i(t))\varphi = \int_\Omega K_i(t, x, \bar{x}_i(t))\varphi(x)\,\mathrm{d}x, \tag{4.3}$$

where $K_i : I \times \overline{\Omega} \times \overline{\Omega} \to \mathbb{R}$. For every continuous curve $t \mapsto \bar{x}_i(t)$ in $\overline{\Omega}$ we will assume that the map $t \mapsto K_i(t, \cdot, \bar{x}_i(t))$ is Bochner measurable as a $L^2(\Omega)$-valued map on the interval $I = [0, \tau]$ and also essentially bounded, i.e., $t \mapsto K_i(t, \cdot, \bar{x}_i(t))$ belongs to $L^\infty(I; L^2(\Omega))$.

As noted previously, in real problems, trajectories are determined by vehicles that are often governed by nonlinear controlled ordinary differential equations. We will assume that each of the vehicles satisfy the following hypotheses.

The sensor trajectories $t \mapsto \bar{x}(t)$ are given by $\bar{x}(t) = M\bar{\theta}(t)$ where $M \in \mathbb{R}^{3\times 6}$ is a constant matrix, $\bar{\theta}(t) \in \mathbb{R}^6$ for each $t \in [0, \tau]$. Also, the maps $t \mapsto \bar{\theta}(t)$ are outputs to a system of controlled ordinary differential equations

$$\begin{aligned}\dot{\bar{\theta}}(t) &= f(t, \bar{\theta}(t), u(t));\\ \bar{\theta}(0) &= \bar{\theta}_0;\end{aligned}$$

with $f \in C^1(\mathbb{R}^{1+6+q}; \mathbb{R}^6)$, $\bar{\theta}_0 \in \Theta_0$ with Θ_0 compact and $M\Theta_0 = X_0 \subset \overline{\Omega}$ and $u(\cdot) \in \mathcal{U}$, where

$$\mathcal{U} = \{u : u \text{ is measurable and } u(t) \in \Gamma \subset \mathbb{R}^q \text{ for all } t \in I = [0, \tau]\},$$

and Γ is compact. We make the following standard assumptions.

Ha) The response satisfies $\bar{\theta}(t, \bar{\theta}_0, u) \in \Theta_1$, with Θ_1 compact and $M\Theta_1 = \overline{\Omega}$, for all $u(\cdot) \in \mathcal{U}$, $\bar{\theta}_0 \in \Theta_0$ and all $t \in I = [0, \tau]$.
Hb) The set $V(\theta, t) = \{f(t, \theta, u) : u \in \Gamma\}$ is convex for each fixed (θ, t).

The previous hypotheses are the usual hypotheses required for the attainability set to be compact (see [39]) and to vary continuously with respect to $t \in [0, \tau]$. Note also that **Ha** implies that $\bar{\theta}(t, \bar{\theta}_0, u)$ is uniformly bounded in $([0, \tau], \Theta_0, \mathcal{U})$. This condition implies that there is an $m > 0$ such that $f(t, \bar{\theta}(t, \bar{\theta}_0, u), u) \le m$ for $t \in [0, \tau]$, $\bar{\theta}_0 \in \Theta_0$ and $u \in \mathcal{U}$ for any fixed $\tau > 0$.

Since we have p sensors, we denote $\bar{\mathbf{x}}(t, \bar{\boldsymbol{\theta}}_0, \mathbf{u}(\cdot))$ as the p-dimensional vector with $\bar{x}_i(t, \bar{\theta}_0^i, u_i(\cdot))$ for $i = 1, 2, \ldots, p$ as elements, where $\bar{\boldsymbol{\theta}}_0 = (\bar{\theta}_0^1, \bar{\theta}_0^2, \ldots, \bar{\theta}_0^p) \in \Theta_0^p$, and $\mathbf{u}(\cdot) = (u_1(\cdot), u_2(\cdot), \ldots, u_p(\cdot))$ such that $u_i(\cdot) \in \mathcal{U}$ for $i = 1, 2, \ldots, p$. We suppose that each $t \mapsto \bar{x}_i(t, \bar{\theta}_0^i, u_i(\cdot))$ is defined $\bar{x}_i(t, \bar{\theta}_0^i, u_i(\cdot)) = M\bar{\theta}_i(t, \bar{\theta}_0^i, u_i(\cdot))$ where $\bar{\theta}_i$ a solution to a controlled differential equation of the type described above and with initial conditions $\bar{\theta}_i(0) = \bar{\theta}_0^i$. Then, we have

$$(C(t, \bar{\mathbf{x}}(t, \bar{\boldsymbol{\theta}}_0, \mathbf{u}(\cdot)))\varphi) = \begin{pmatrix} (C_1(t, \bar{x}_1(t, \bar{\theta}_0^1, u_1(\cdot)))\varphi) \\ (C_2(t, \bar{x}_2(t, \bar{\theta}_0^2, u_2(\cdot)))\varphi) \\ \vdots \\ (C_p(t, \bar{x}_p(t, \bar{\theta}_0^p, u_p(\cdot)))\varphi) \end{pmatrix} \in \mathbb{C}^p, \tag{4.4}$$

with

$$C_i(t, \bar{x}_i(t, \bar{\theta}_0^i, u_i(\cdot)))\varphi = \int_\Omega K_i(t, x, \bar{x}_i(t, \bar{\theta}_0^i, u_i(\cdot)))\varphi(x)\,\mathrm{d}x. \tag{4.5}$$

We denote $\bar{\mathbf{x}}(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) = M\bar{\boldsymbol{\theta}}(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$, where $M\bar{\boldsymbol{\theta}}(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$ is shorthand for $\{M\bar{\theta}_i(t,\bar{\theta}_0^i,u_i(\cdot))\}_{i=1}^p$.

LEMMA 4.6. *Let $I=[0,\tau]$ and suppose that C is continuous with respect to location (see definition 4.5). Then, the set $\mathcal{F}$ defined by*

$$\mathcal{F}=\left\{C^*C(\cdot,M\bar{\boldsymbol{\theta}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}))\in L^\infty(I;\mathscr{I}_1(L^2(\Omega))):\ \text{for } \bar{\boldsymbol{\theta}}_0\in\Theta_0^p \text{ and } \mathbf{u}\in\mathcal{U}^p\right\},$$

is compact in $L^\infty([0,\tau];\mathscr{I}_1(L^2(\Omega)))$.

If the kernels K_i do not depend explicitly on t, so that $K_i(\cdot,y)\in L^2(\Omega)$ for each $y\in\overline{\Omega}$, then the set $\mathcal{F}$ defined as

$$\mathcal{F}=\left\{C^*C(M\bar{\boldsymbol{\theta}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}))\in\mathscr{C}(I;\mathscr{I}_1(L^2(\Omega))):\ \text{for } \bar{\boldsymbol{\theta}}_0\in\Theta_0^p \text{ and } \mathbf{u}\in\mathcal{U}^p\right\},$$

is compact in $\mathscr{C}([0,\tau];\mathscr{I}_1(L^2(\Omega)))$.

Proof. We present a proof for the case of one moving sensor. The generalization to multiple sensors is straightforward. For any sequence of $\bar{\theta}^k(\cdot,\bar{\theta}_0^k,u_k(\cdot))$ with $(\bar{\theta}_0^k,u_k(\cdot))\in(\Theta_0,\mathcal{U})$, there is a subsequence (that we will also call $\bar{\theta}^k(\cdot,\bar{\theta}_0^k,u_k(\cdot))$) such that

$$\sup_{t\in[0,\tau]}\|\bar{\theta}(t,\bar{\theta}_0,u(\cdot))-\bar{\theta}^k(t,\bar{\theta}_0^k,u_k(\cdot))\|_{\mathbb{R}^6}\to 0, \tag{4.6}$$

as $k\to\infty$, for some $\bar{\theta}_0\in\Theta_0$ and some $u(\cdot)\in\mathcal{U}$ (for a proof when see [39]).

Let $\bar{x}(t)=\bar{x}(t,\bar{\theta}_0,u(\cdot))$ and $\bar{x}^k(t)=\bar{x}^k(t,\bar{\theta}_0^k,u_k(\cdot))$ for $k\in\mathbb{N}$. Also, since C is continuous with respect to location on $\overline{\Omega}$, we have

$$\begin{aligned}|(C(t,\bar{x}(t))-C(t,\bar{x}^k(t)))\varphi| &\le \int_\Omega |K(t,x,\bar{x}(t))-K(t,x,\bar{x}^k(t))||\varphi(x)|\,\mathrm{d}x\\ &\le \|K(t,\cdot,\bar{x}(t))-K(t,\cdot,\bar{x}^k(t))\|_{L^2(\Omega)}\|\varphi(\cdot)\|_{L^2(\Omega)}\\ &\le g(\|\bar{x}(t)-\bar{x}^k(t)\|_{\mathbb{C}^p})\|\varphi(\cdot)\|_{L^2(\Omega)}.\end{aligned}$$

Equivalently,

$$|(C(t,\bar{x}(t))-C(t,\bar{x}^k(t)))\varphi|\le g(\|M(\bar{\theta}(t)-\bar{\theta}^k(t))\|_{\mathbb{R}^n})\|\varphi(\cdot)\|_{L^2(\Omega)}, \tag{4.7}$$

which implies that $\|(C(t,\bar{x}(t))-C(t,\bar{x}^k(t)))\|_{\mathscr{L}(L^2(\Omega),\mathbb{C})}\to 0$ for any $t\in I$ and that $\|(C(\cdot,\bar{x}(\cdot))-C(\cdot,\bar{x}^k(\cdot)))\|_{L^\infty(I;\mathscr{L}(L^2(\Omega),\mathbb{C}))}\to 0$ because g is continuous with $g(0)=0$ and by (4.6).

We also know that there is $c>0$ such that $\|C(t,\bar{x}(t))-C(t,\bar{x}^k(t))\|_{\mathscr{I}_1(L^2(\Omega),\mathbb{C})}\le c\|C(t,\bar{x}(t))-C(t,\bar{x}^k(t))\|_{\mathscr{L}(L^2(\Omega),\mathbb{C})}$ (see the proof of Lemma 4.3 and Proposition 3.8). Hence

$$\|C(\cdot,\bar{x}(\cdot))-C(\cdot,\bar{x}^k(\cdot))\|_{L^\infty(I;\mathscr{I}_1(L^2(\Omega),\mathbb{C}))}\to 0,$$

as $k\to\infty$ since $t\mapsto C(t,\bar{x})$ is also $\mathscr{I}_1(L^2(\Omega),\mathbb{C})$-valued and measurable

Clearly $(C^*C)(\cdot,\bar{x}(\cdot))\in\mathcal{F}$ and by the properties of Proposition 3.8, we have

$$\begin{aligned}&\|(C^*C)(t,\bar{x}(t))-(C^*C)(t,\bar{x}^k(t))\|_{\mathscr{I}_1(L^2(\Omega))}\le\\ &\qquad\|C^*(t,\bar{x}(t))\|_{\mathscr{I}_1(\mathbb{C},L^2(\Omega))}\|C(t,\bar{x}(t))-C(t,\bar{x}^k(t))\|_{\mathscr{I}_1(L^2(\Omega),\mathbb{C})}+\\ &\qquad\quad\|C(t,\bar{x}^k(t))\|_{\mathscr{I}_1(L^2(\Omega),\mathbb{C})}\|C^*(t,\bar{x}(t))-C^*(t,\bar{x}^k(t))\|_{\mathscr{I}_1(\mathbb{C},L^2(\Omega))}.\end{aligned}$$

Therefore,

$$\|(C^*C)(t,\bar{x}(t)) - (C^*C)(t,\bar{x}^k(t))\|_{\mathscr{I}_1(L^2(\Omega))} \le \left(\|C(t,\bar{x}(t))\|_{\mathscr{I}_1(L^2(\Omega),\mathbb{C})} + \|C(t,\bar{x}^k(t))\|_{\mathscr{I}_1(L^2(\Omega),\mathbb{C})}\right)\|C(t,\bar{x}(t)) - C(t,\bar{x}^k(t)\|_{\mathscr{I}_1(L^2(\Omega),\mathbb{C})}.$$

Taking the $\operatorname{ess\,sup}_{t\in I}$ we observe that the term in the parentheses is uniformly bounded in $n \in \mathbb{N}$ and hence we have

$$\|(C^*C)(\cdot,\bar{x}(\cdot)) - (C^*C)(\cdot,\bar{x}^k(\cdot))\|_{L^\infty(I;\mathscr{I}_1(L^2(\Omega)))} \to 0,$$

which proves the compactness of $\mathcal{F}$.

If the kernel K does not depend explicitly on t, then the inequality $\|K(\cdot,y) - K(\cdot,z)\|_{L^2(\Omega)} \le g(\|y-z\|_{\mathbb{R}^n})$ for any $y,z \in \overline{\Omega}$ implies that $t \mapsto C(\bar{x}(t))$ is continuous as we now prove. It follows that

$$\begin{aligned}|(C(\bar{x}(t)) - C(\bar{x}(s)))\phi(\cdot)| &\le \|K(\cdot,\bar{x}(t)) - K(\cdot,\bar{x}(s))\|_{L^2(\Omega)}\|\varphi(\cdot)\|_{L^2(\Omega)}\\ &\le g(\|M(\bar{\theta}(t)-\bar{\theta}(s))\|_{\mathbb{R}^n})\|\varphi(\cdot)\|_{L^2(\Omega)},\end{aligned}$$

which implies $\|C(\bar{x}(t)) - C(\bar{x}(s))\|_{\mathscr{L}(L^2(\Omega),\mathbb{C})} \le g(\|M(\bar{\theta}(t)-\bar{\theta}(s))\|_{\mathbb{R}^n})$. Since $\mathbb{C}$ is finite dimensional we have that

$$\|C(\bar{x}(t)) - C(\bar{x}(s))\|_{\mathscr{I}_1(L^2(\Omega),\mathbb{C})} \le cg(\|M(\bar{\theta}(t)-\bar{\theta}(s))\|_{\mathbb{R}^n})$$

for some constant $c > 0$ (see the proof of Lemma 4.3), independent of $t, s \in I$. Since $t \mapsto \bar{\theta}(t)$ is a continuous trajectory, this implies that $C(\bar{x}(\cdot)) \in \mathscr{C}(I;\mathscr{I}_1(L^2(\Omega),\mathbb{C}))$ and the same inequalities as before prove the compactness of $\mathcal{F}$ in this topology. ◻

5. Minimizers of Problem ($\mathcal{P}$) and Optimality Conditions. Several results and assumptions should be considered before proving existence for the minimization problem of interest. Recall, that we have assumed that A, the operator associated with the dynamics of the process in 2.5, is the infinitesimal generator of a C_0-semigroup $S(t)$ over $L^2(\Omega)$. For a general C_0-semigroup over a separable complex Hilbert space $\mathscr{H}$, we have the following result for solutions of the integral Riccati equation which we state without proof for the sake of brevity (for a proof see [43] or [15]).

THEOREM 5.1. *Let $\mathscr{H}$ be a separable complex Hilbert space, $I = [0,\tau]$ or $I = \mathbb{R}^+$, $S(t)$ be a C_0-semigroup on $\mathscr{H}$, and suppose that*
(i) *$\Sigma_0 \in \mathscr{I}_p$ and $\Sigma_0 \ge 0$;*
(ii) *$BB^*(\cdot) \in L^1_{loc}(I;\mathscr{I}_p)$, with $BB^*(t) \ge 0$ for $t \in I$;*
(iii) *$C^*C(\cdot) \in L^\infty_{loc}(I;\mathscr{L}(\mathscr{H}))$, with $C^*C(t) \ge 0$ for $t \in I$.*
Then, the equation

$$\Sigma(t) = S(t)\Sigma_0 S^*(t) + \int_0^t S(t-s)\big(BB^* - \Sigma(C^*C)\Sigma\big)(s)S^*(t-s)\,\mathrm{d}s, \tag{5.1}$$

where the integral is a Bochner integral, has a unique solution in the space $L^2_{loc}(I;\mathscr{I}_{2p})$, and even more the solution belongs to $\mathscr{C}(I;\mathscr{I}_p)$ and is point-wise self-adjoint and non-negative.

Suppose that $BB^*(\cdot) \in \mathscr{C}([0,\tau];\mathscr{I}_p)$ and that $C^*C(\cdot) \in \mathscr{C}([0,\tau];\mathscr{L}(\mathscr{H}))$, then there is a unique solution $\Sigma(\cdot)$ in $\mathscr{C}([0,\tau];\mathscr{I}_p)$ for (5.1). Since $\mathscr{H}$ is reflexive, $S^*(t)$ is a C_0-semigroup with generator A^* (see [42]). Let $x,y \in \mathscr{D}(A^*)$, and then $\Sigma(\cdot)$ satisfies

$$\langle\Sigma(t)x,y\rangle = \langle\Sigma_0 S^*(t)x, S^*(t)y\rangle + \int_0^t \big\langle\big(BB^* - \Sigma(C^*C)\Sigma\big)(s)S^*(t-s)x, S^*(t-s)y\big\rangle\,\mathrm{d}s.$$

Therefore, $t \mapsto \langle \Sigma(t)x, y\rangle$ is differentiable and a simple computation with the Leibniz integral rule (see [8] for a proof when BB^* and C^*C are constant mappings) shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \Sigma(t)x, y\rangle = \langle A^*y, \Sigma(t)x\rangle + \langle \Sigma(t)y, A^*x\rangle + \langle BB^*(t)x, y\rangle - \langle \Sigma(t)(C^*C)(t)\Sigma(t)x, y\rangle,$$

with $\langle \Sigma(0)x, y\rangle = \langle \Sigma_0 x, y\rangle$. Therefore, any solution in $\mathscr{C}([0,\tau];\mathscr{I}_p)$ of the integral Riccati equation is a weak solution of the differential equation

$$\dot{\Sigma}(t) = A\Sigma(t) + \Sigma(t)A^* + BB^*(t) - \Sigma(t)(C^*C)(t)\Sigma(t)\ , \tag{5.2}$$

with initial condition $\Sigma(0) = \Sigma_0$. Conversely, any weak solution to this equation can be proven to be a mild solution to the integral Riccati equation (5.1) (See [8] for a proof for constant mappings BB^* and C^*C. The extension for $BB^*(\cdot) \in \mathscr{C}([0,\tau];\mathscr{I}_p)$ and $C^*C(\cdot) \in \mathscr{C}([0,\tau];\mathscr{L}(\mathscr{H}))$ is straightforward). Since the unique solution of this latter equation in the space $\mathscr{C}([0,\tau];\mathscr{I}_p)$ is also a mild solution, these two are equivalent. Therefore, under the hypotheses of Theorem 5.1 and when $BB^*(\cdot) \in \mathscr{C}([0,\tau];\mathscr{I}_p)$ and $C^*C(\cdot) \in \mathscr{C}([0,\tau];\mathscr{L}(\mathscr{H}))$, any weak solution to (5.2) is $\mathscr{I}_p$-valued continuous solution of the integral Riccati equation (5.1).

We additionally require an approximation result for solutions (5.2) that is not only used to prove existence of solutions to **Problem** $(\mathcal{P})$ but also for the numerical scheme that will be implemented to approximate solutions. The proof of the following result can also be found in [43] or [15].

THEOREM 5.2. *Suppose that $S(t)$ is a C_0-semigroup of linear operators over $\mathscr{H}$, and that $\{S_n(t)\}$ is a sequence of uniformly continuous semigroups over the same Hilbert space $\mathscr{H}$ that satisfy, for each $x \in \mathscr{H}$,*

$$\|S(t)x - S_n(t)x\| \to 0 \qquad \text{and} \qquad \|S^*(t)x - S_n^*(t)x\| \to 0,$$

as $n \to \infty$, uniformly in compact intervals. Suppose also the following.

(i) *$\Sigma_0 \geq 0$ and the sequence $\{\Sigma_0^n\}_{n=1}^\infty$ are in $\mathscr{I}_p$, $\Sigma_0^n \geq 0$ for all $n \in \mathbb{N}$ and $\|\Sigma_0 - \Sigma_0^n\|_p \to 0$ as $n \to \infty$.*

(ii) *$BB^*(\cdot)$ and the sequence $\{D_n(\cdot)\}_{n=1}^\infty$ are in $L^1_{loc}(\mathbb{R}^+;\mathscr{I}_p)$, $BB^*(t) \geq 0$ and $D_n(t) \geq 0$ for all $t \in \mathbb{R}^+$ and all $n \in \mathbb{N}$ and satisfy*

$$\int_0^\tau \|BB^*(t) - D_n(s)\|_p \,\mathrm{d}s \to 0,$$

for any fixed $\tau > 0$ and as $n \to \infty$.

(iii) *$C^*C(\cdot)$ and the sequence $\{E_n(\cdot)\}_{n=1}^\infty$ are in $L^\infty(\mathbb{R}^+;\mathscr{L}(\mathscr{H}))$, $C^*C(t) \geq 0$ and $E_n(t) \geq 0$ for all $t \in \mathbb{R}^+$ and all $n \in \mathbb{N}$ and satisfy*

$$\operatorname*{ess\ sup}_{t\in[0,\tau]} \|C^*C(t) - E_n(t)\| \to 0,$$

for any fixed $\tau > 0$ and as $n \to \infty$.

Then, if $\Sigma(\cdot) \in \mathscr{C}([0,a],\mathscr{I}_p)$ is a solution of

$$\Sigma(t) = S(t)\Sigma_0 S^*(t) + \int_0^t S(t-s)\big(BB^* - \Sigma(C^*C)\Sigma\big)(s)S^*(t-s)\,\mathrm{d}s,$$

for some $a > 0$ and if $\Sigma_n(\cdot) \in \mathscr{C}(\mathbb{R}^+,\mathscr{I}_p)$ is the sequence of solutions of

$$\Sigma_n(t) = S_n(t)\Sigma_0^n S_n^*(t) + \int_0^t S_n(t-s)\big(D_n - \Sigma_n E_n \Sigma_n\big)(s)S_n^*(t-s)\,\mathrm{d}s,$$

we observe that

$$\sup_{t\in[0,a]} \|\Sigma(t) - \Sigma_n(t)\|_p \to 0,$$

as $n \to \infty$.

We are now in shape to prove that **Problem** $(\mathcal{P})$, described in the introduction of this work, has a solution. We present the proof for the mobile sensor network which also includes the solution for the stationary sensor networks. We assume throughout this section that $\mathscr{H} = L^2(\Omega)$.

THEOREM 5.3. *Consider the case of a moving sensor network of p sensors with trajectories given by* $\bar{\mathbf{x}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) = \left\{\bar{x}_i(\cdot,\bar{\theta}_0^i,u_i(\cdot))\right\}_{i=1}^p$, *where* $\bar{x}_i(t,\bar{\theta}_0^i,u_i(\cdot)) = M\bar{\theta}_i(t,\bar{\theta}_0^i,u_i(\cdot))$ *for all* $t\in[0,\tau]$ *and* $t\mapsto \bar{\theta}_i(t,\bar{\theta}_0^i,u_i(\cdot))$ *is a solution to a controlled ordinary differential equation of the type described in Section 4.2.2. Suppose that the initial set* Θ_0 *and the admissible control set* $\mathcal{U}$ *are also of the type described in Section 4.2.2.*

Suppose that the output map $(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))\mapsto C(t,\mathbf{x}(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)))$ *satisfies the conditions of Lemma 4.6. Also assume that* $BB^*(\cdot)\in L^1([0,\tau];\mathscr{I}_1)$ *with* $BB^*(t)\geq 0$ *for* $t\in[0,\tau]$, $S(t)$ *is a* C_0*-semigroup,* $0\leq\Sigma_0\in\mathscr{I}_1$ *and that* $(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))\mapsto\Sigma(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$ *is the unique solution in* $\mathscr{C}([0,\tau];\mathscr{I}_1)$ *to*

$$\Sigma(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) = S(t)\Sigma_0 S^*(t) + \int_0^t S(t-s)\big(BB^*(s) - \big(\Sigma(C^*C)\Sigma\big)(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))\big)S^*(t-s)\,\mathrm{d}s, \tag{5.3}$$

for each $(\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))\in\Theta_0^p\times\mathcal{U}^p$. *In addition, let* $Q(\cdot)\in L^\infty([0,\tau];\mathscr{L}(\mathscr{H}))$ *with* $Q(t)\geq 0$ *for* $t\in[0,\tau]$ *and* $J:\Theta_0^p\times\mathcal{U}^p\to\mathbb{R}$ *be defined as*

$$J(\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) = \int_0^\tau \mathrm{Tr}\,(Q(t)\Sigma(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)))\,\mathrm{d}t.$$

Then, there is $(\bar{\boldsymbol{\theta}}_0^{min},\mathbf{u}_{min}(\cdot))\in\Theta_0^p\times\mathcal{U}^p$ *such that*

$$\inf_{\bar{\boldsymbol{\theta}}_0\in X_0^p,\mathbf{u}(\cdot)\in\mathcal{U}^p} J(\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) = J(\bar{\boldsymbol{\theta}}_0^{min},\mathbf{u}_{min}(\cdot)),$$

i.e., ***Problem*** $(\mathcal{P})$ *has a solution.*

Proof. The set

$$\mathcal{F} = \left\{C^*C(\cdot,M\bar{\boldsymbol{\theta}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)))\in L^\infty([0,\tau];\mathscr{I}_1) : \bar{\boldsymbol{\theta}}_0\in\Theta_0^p \text{ and } \mathbf{u}(\cdot)\in\mathcal{U}^p\right\},$$

is compact in $L^\infty([0,\tau];\mathscr{I}_1)$ by Lemma 4.6, where $M\bar{\boldsymbol{\theta}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$ is shorthand for the set $\left\{M\bar{\theta}_i(\cdot,\bar{\theta}_0^i,u_i(\cdot))\right\}_{i=1}^p = \left\{\bar{x}_i(\cdot,\bar{\theta}_0^i,u_i(\cdot))\right\}_{i=1}^p$, hence $M\bar{\boldsymbol{\theta}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) = \bar{\mathbf{x}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$.

Theorem 5.2 implies that the map $t\mapsto\Sigma(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$ varies continuously in the $\sup_{t\in[0,\tau]}$ $\mathscr{I}_1$-norm with respect to $C^*C(\cdot,M\bar{\boldsymbol{\theta}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))\in\mathcal{F}$. If $\Sigma_1(\cdot),\Sigma_2(\cdot)\in\mathscr{C}([0,\tau];\mathscr{I}_1)$, then we have the inequality

$$\left|\int_0^\tau \mathrm{Tr}\,(Q(t)\Sigma_1(t))\,\mathrm{d}t - \int_0^\tau \mathrm{Tr}\,(Q(t)\Sigma_2(t))\,\mathrm{d}t\right| \leq \tau\|Q(\cdot)\|_{L^\infty([0,\tau];\mathscr{L}(\mathscr{H}))}\sup_{t\in[0,\tau]}\|(\Sigma_1-\Sigma_2)(t)\|_1.$$

Which implies that the map $\Sigma(\cdot) \mapsto \int_0^\tau \mathrm{Tr}\,(Q(t)\Sigma(t))\,\mathrm{d}t$ is uniformly continuous in $\mathscr{C}([0,\tau];\mathscr{I}_1)$. Since, by Theorem 5.1, $0 \leq \Sigma(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) \in \mathscr{I}_1$ we have that then $Q(t)\Sigma(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) \in \mathscr{I}_1$ and $\mathrm{Tr}\,(Q(t)\Sigma(t,\mathbf{x}))$ is well defined. It is also a non-negative since $Q(t)$ and $\Sigma(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$ are non-negative (see [29] for a proof). Therefore $J(\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))$ is well defined over $\Theta_0^p \times \mathcal{U}$ and by composition of continuous mappings, the map

$$C^*C(\cdot,\bar{\mathbf{x}}(\cdot,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot)) \mapsto \int_0^\tau \mathrm{Tr}\,(Q(t)\Sigma(t,\bar{\mathbf{x}}(t,\bar{\boldsymbol{\theta}}_0,\mathbf{u}(\cdot))))\,\mathrm{d}t$$

is continuous over the compact set $\mathcal{F}$. Then, the result follows. ☐

5.1. The Gradient of Σ w.r.t. the Map $t \mapsto C^*C(t)$. We are interested in using a gradient type algorithm over the trajectories of the sensors. For this matter, we will first prove that the solution of the integral Riccati equation, as a function of the mapping $t \mapsto C^*C(t)$, is Fréchet differentiable. We first need a lemma.

LEMMA 5.4. *Let $\mathscr{H}$ be a complex separable Hilbert space and $S(t)$ be a C_0-semigroup over $\mathscr{H}$. Suppose that $G(\cdot)$ and $\Sigma(\cdot)$ belong to $X = \mathscr{C}([0,\tau];\mathscr{I}_1)$. Then the equation $\Lambda = \hat{\gamma}(\Lambda)$ has a unique solution in $\mathscr{L}(X)$, where $\hat{\gamma} : \mathscr{L}(X) \to \mathscr{L}(X)$ and is defined by*

$$\big(\hat{\gamma}(\Lambda)K\big)(t) = -\int_0^t S(t-s)((\Lambda K)G\Sigma + \Sigma G(\Lambda K) + \Sigma K \Sigma)(s) S^*(t-s)\,\mathrm{d}s, \quad (5.4)$$

for all $K(\cdot) \in X$.

Proof. We will use the re-normalization technique first described in [9]. Define X_λ (with $\lambda > 0$) to be the set of all trace class continuous mappings $t \mapsto F(t)$ and domain $[0,\tau]$ with the norm $|||\cdot|||_{1,\lambda}$ defined as

$$|||F(\cdot)|||_{1,\lambda} = \sup_{t\in[0,\tau]} e^{-\lambda t}\|F(t)\|_1.$$

Note that X has norm given by $|||F(\cdot)|||_1 = \sup_{t\in[0,\tau]}\|F(t)\|_1$. It is obvious that X and X_λ coincide element-wise, and one can prove that they coincide topologically. Observe the equivalency of the norms of X_λ and X:

$$e^{-\lambda\tau}|||F(\cdot)|||_1 \leq |||F(\cdot)|||_{1,\lambda} \leq |||F(\cdot)|||_1.$$

Moreover, the spaces $\mathscr{L}(X_\lambda)$ and $\mathscr{L}(X)$ are topologically equivalent and

$$e^{-\lambda\tau}\|\Lambda\|_{\mathscr{L}(X)} \leq \|\Lambda\|_{\mathscr{L}(X_\lambda)} \leq e^{\lambda\tau}\|\Lambda\|_{\mathscr{L}(X)}.$$

If $\Lambda \in \mathscr{L}(X)$, then it follows immediately that $\hat{\gamma}(\Lambda)$ is also a linear operator acting on X. Also, let $\|S(t)\| \leq M_\tau$ for $t \in [0,\tau]$, and $m = \max\Big(|||G(\cdot)|||_1, |||\Sigma(\cdot)|||_1\Big)$. Then

$$|||\big(\hat{\gamma}(\Lambda)K\big)(\cdot)|||_1 \leq \tau M_\tau^2 m^2 |||K(\cdot)|||_1\Big(2\|\Lambda\|_{\mathscr{L}(X)} + 1\Big).$$

Hence $\hat{\gamma}(\Lambda) \in \mathscr{L}(X)$.

Let $\Lambda_1, \Lambda_2 \in \mathscr{L}(X)$ and $K(\cdot) \in X$. Then by the definition of the norm $|||\cdot|||_{1,\lambda}$ we observe

$$\begin{aligned}
\left\|\Big(\big(\hat\gamma(\Lambda_1)-\hat\gamma(\Lambda_2)\big)K\Big)(t)\right\|_1 &\le 2M_\tau^2 m^2 \int_0^t \|\big((\Lambda_1-\Lambda_2)K\big)(s)\|_1 \,\mathrm{d}s \\
&\le 2M_\tau^2 m^2 \int_0^t \|\big((\Lambda_1-\Lambda_2)K\big)(s)\|_1 e^{-\lambda s}e^{\lambda s} \,\mathrm{d}s \\
&\le 2M_\tau^2 m^2 |||\big((\Lambda_1-\Lambda_2)K\big)(\cdot)|||_{1,\lambda} \int_0^t e^{\lambda s}\,\mathrm{d}s \\
&\le \frac{2M_\tau^2 m^2}{\lambda} |||\big((\Lambda_1-\Lambda_2)K\big)(\cdot)|||_{1,\lambda}\ e^{\lambda t} \\
&\le \frac{2M_\tau^2 m^2}{\lambda} \|\Lambda_1-\Lambda_2\|_{\mathscr{L}(X_\lambda)} |||K(\cdot)|||_{1,\lambda}\ e^{\lambda t}
\end{aligned}$$

since $\int_0^t e^{\lambda s}\,\mathrm{d}s = \frac{e^{\lambda t}-1}{\lambda} < \frac{e^{\lambda t}}{\lambda}$ for $\lambda > 0$. Therefore

$$e^{-\lambda t}\left\|\Big(\big(\hat\gamma(\Lambda_1)-\hat\gamma(\Lambda_2)\big)K\Big)(t)\right\|_1 \le \frac{2M_\tau^2 m^2}{\lambda}\|\Lambda_1-\Lambda_2\|_{\mathscr{L}(X_\lambda)}|||K(\cdot)|||_{1,\lambda},$$

which implies by taking the sup over $t \in [0,\tau]$ that

$$\left|||\Big(\big(\hat\gamma(\Lambda_1)-\hat\gamma(\Lambda_2)\big)K\Big)(\cdot)\right|||_{1,\lambda} \le \frac{2M_\tau^2 m^2}{\lambda}\|\Lambda_1-\Lambda_2\|_{\mathscr{L}(X_\lambda)}|||K(\cdot)|||_{1,\lambda}.$$

Finally

$$\left\|\big(\hat\gamma(\Lambda_1)-\hat\gamma(\Lambda_2)\big)\right\|_{\mathscr{L}(X_\lambda)} \le \frac{2M_\tau^2 m^2}{\lambda}\|\Lambda_1-\Lambda_2\|_{\mathscr{L}(X_\lambda)}. \tag{5.5}$$

Suppose that $\lambda > 2M_\tau^2 m^2$. Then from Equation 5.5, we satisfy that the mapping $\hat\gamma : \mathscr{L}(X_\lambda) \to \mathscr{L}(X_\lambda)$ is a contraction, and by the contraction mapping principle there is a unique solution to $\Lambda = \hat\gamma(\Lambda)$ in $\mathscr{L}(X_\lambda)$. Since $\|\cdot\|_{\mathscr{L}(X)}$ and $\|\cdot\|_{\mathscr{L}(X_\lambda)}$ are equivalent norms, $\Lambda = \hat\gamma(\Lambda)$ has only one solution in in $\mathscr{L}(X)$. ☐

We now prove that the solution of the integral Riccati equation $\Sigma(\cdot)$ is Fréchet differentiable with respect to the mapping $t \mapsto C^*C(t)$.

Theorem 5.5. *Let $\mathscr{H}$ be a complex separable Hilbert space and $S(t)$ be a C_0-semigroup over $\mathscr{H}$. Suppose that $0 \le \Sigma_0 \in \mathscr{I}_1$ and $F(\cdot) \in L^1([0,\tau];\mathscr{I}_1)$. Let $\mathcal{D}$ be an open set of $X = \mathscr{C}([0,\tau];\mathscr{I}_1)$, then $\Sigma(\cdot, G) \in X$ the unique solution of*

$$\Sigma(t,G) = S(t)\Sigma_0 S^*(t) + \int_0^t S(t-s)\Big(F - \Sigma(G)G\Sigma(G)\Big)(s)S^*(t-s)\,\mathrm{d}s, \tag{5.6}$$

with $t \in [0,\tau]$ is Fréchet differentiable with respect to $G(\cdot) \in \mathcal{D}$. The Fréchet derivative of $\Sigma(\cdot;G)$ with respect to G is denoted by $\Lambda(G) \in \mathscr{L}(X)$ and is equal to the unique solution of

$$\begin{aligned}
\big(\Lambda(G)h\big)(t) = -\int_0^t S(t-s)\Big(\big(\Lambda(G)h\big)G\Sigma(G) + \Sigma(G)G\big(\Lambda(G)h\big) + \\
\Sigma(G)h\Sigma(G)\Big)(s)S^*(t-s)\,\mathrm{d}s,
\end{aligned} \tag{5.7}$$

for all $h(\cdot) \in X$ and all $t \in [0,\tau]$

Proof. If $F(\cdot) \in X$, we define $|||F(\cdot)|||_1 = \sup_{t\in[0,\tau]} \|F(t)\|_1$. By the continuity of the map $G \mapsto \Sigma(\cdot, G)$ obtained in Theorem 5.2, we know that $\big|\big|\big|\Sigma(\cdot, G+h) - \Sigma(\cdot, G)\big|\big|\big|_1 = O\big(|||h(\cdot)|||_1\big)$, and therefore

$$a(G,h) := \Big|\Big|\Big|\big(\Sigma(G+h)h\Sigma(G+h) - \Sigma(G)h\Sigma(G)\big)(\cdot)\Big|\Big|\Big|_1 = O\big(|||h(\cdot)|||_1^2\big).$$

Also since $\Lambda(G) \in \mathscr{L}(X)$, we satisfy that

$$b(G,h) := \Big|\Big|\Big|\Big(\big(\Sigma(G+h) - \Sigma(G)\big)G\big(\Lambda(G)h\big)\Big)(\cdot)\Big|\Big|\Big|_1 = O\big(|||h(\cdot)|||_1^2\big).$$

By direct calculation we observe that

$$\begin{aligned}
&\Sigma(t, G+h) - \Sigma(t,G) - (\Lambda(G)h)(t) = \\
&\quad -\int_0^t S(t-s)\Big[\Sigma(G+h)(G+h)\Sigma(G+h) - \Sigma(G)G\Sigma(G) + \\
&\quad - \big(\Lambda(G)h\big)G\Sigma(G) - \Sigma(G)G\big(\Lambda(G)h\big) - \Sigma(G)h\Sigma(G)\Big](s)S^*(t-s)\,\mathrm{d}s,
\end{aligned}$$

and that

$$\begin{aligned}
&\Sigma(t, G+h) - \Sigma(t,G) - (\Lambda(G)h)(t) = \\
&\quad -\int_0^t S(t-s)\Big[\Big(\Sigma(G+h) - \Sigma(G) - (\Lambda(G)h)\Big)G\Sigma(G) + \\
&\qquad\qquad \Sigma(G+h)G\Big(\Sigma(G+h) - \Sigma(G) - (\Lambda(G)h)\Big) + \\
&\qquad\qquad \Big(\Sigma(G+h) - \Sigma(G)\Big)G\big(\Lambda(G)h\big) + \\
&\qquad\qquad \Sigma(G+h)h\Sigma(G+h) - \Sigma(G)h\Sigma(G)\Big](s)S^*(t-s)\,\mathrm{d}s.
\end{aligned}$$

Therefore if $z(t,G,h) = \|\Sigma(t, G+h) - \Sigma(t,G) - (\Lambda(G)h)(t)\|_1$, then

$$z(t,G,h) = 2M_\tau^2\rho|||G(\cdot)|||_1 \int_0^t z(s,G,h)\,\mathrm{d}s + \tau M_\tau^2\Big(a(h,G) + b(h,G)\Big).$$

By the Grönwall's Lemma we observe that

$$|||\Sigma(\cdot, G+h) - \Sigma(\cdot,G) - (\Lambda(G)h)(\cdot)|||_1 \le \tau M_\tau^2\Big(a(h,G) + b(h,G)\Big)e^{2\tau M_\tau^2\rho|||G(\cdot)|||_1},$$

i.e.,

$$\frac{|||\Sigma(\cdot, G+h) - \Sigma(\cdot,G) - (\Lambda(G)h)(\cdot)|||_1}{|||h(\cdot)|||_1} = O(|||h(\cdot)|||_1),$$

since $a(h,G) + b(h,G) = O(|||h(\cdot)|||_1^2)$. This implies, as claimed, that $\Sigma'(G) = \Lambda(G)$, where the derivative is taken in the Fréchet sense. □

The previous result proves that the Fréchet derivative of Σ with respect to the mapping $t \mapsto C^*C(t)$ exists. We now show that in the case with moving sensors, we can take the derivative with respect to the controls.

We will denote a position at time t with input control u as $\bar{x}(t,u)$. Then we can regard $\bar{x}$ as a mapping from $L^2([0,\tau];\mathbb{R})$ to $\mathscr{C}([0,\tau];\overline{\Omega})$. Let

$$\bar{x}(t,u) = e^{At}\bar{x}_0 + \int_0^t e^{A(t-s)}bu(s)\,\mathrm{d}s.$$

Then the Fréchet derivative of $\bar{x}$ with respect to u satisfies $D_u\bar{x} \in \mathscr{L}\Big(L^2([0,\tau];\mathbb{R}), \mathscr{C}([0,\tau];\overline{\Omega})\Big)$ and it's given by

$$(D_u\bar{x}h)(t) = \int_0^t e^{A(t-s)} b h(s)\,\mathrm{d}s.$$

Also, let $G(t,\bar{x}(t)) = C^*C(t,\bar{x}(t))$ where the output map comes from one moving sensor. G is a mapping from $\mathscr{C}([0,\tau];\overline{\Omega})$ to $\mathscr{C}([0,\tau];\mathscr{I}_1(L^2(\Omega)))$. Then,

$$G(t,\bar{x}(t))\varphi(\cdot) = K(t,x,\bar{x}(t))\int_\Omega K(t,y,\bar{x}(t))\varphi(y)\,\mathrm{d}y,$$

and the regularity of $\bar{x}(t) \mapsto G(t,\bar{x}(t))$ is guaranteed by the smoothness of K, in which case, following the chain rule, we can compute $D_u\Sigma(\cdot, G(\cdot,\bar{x}(t,u)))$.

6. A Galerkin Approximation Scheme. In this section, we discuss the approximation of $\mathscr{I}_p$-valued solutions to the integral Riccati equation. We will make use of previous results by [19], [20] and [27]. Although, these references treat Hilbert-Schmidt operators or bounded operators, we will extend these results to $\mathscr{I}_p$ for $1 \le p \le \infty$.

Let, for each $n \in \mathbb{N}$, P_n be the projection from our initial Hilbert space $\mathscr{H}$ onto a finite dimensional Hilbert space $\mathcal{V}_n$ such that $\mathcal{V}_n \subset \mathscr{H}$ and $\mathcal{V}_n \subset \mathscr{D}(A)$, where the sequence $P_n^*P_n$ converges strongly to the identity and $[\mathcal{N}(P_n)]^\perp \subset \mathscr{D}(A)$ for each $n \in \mathbb{N}$. Since P_n are projections onto $\mathcal{V}_n \subset \mathscr{H}$, the norm $\|P_n\|$ is uniformly bounded.

Then $A_n = P_nAP_n^*$ is a bounded linear operator and the infinitesimal generator of the uniformly continuous semigroup $S_n(t) = e^{A_nt}$ on $\mathscr{H}$ where A is the infinitesimal generator of a semigroup $S(t)$ on $\mathscr{H}$. Consider the conditions:

H1) There are $M \ge 1$ and $\omega \in \mathbb{R}$ such that $\|S_n(t)\| \le Me^{\omega t}$.

H2) There is a dense subset $D \subset \mathscr{H}$ such that $D \subset \mathscr{D}(A)$ and if $x \in D$, then $A_nx \to Ax$ as $n \to \infty$ and there is a complex number λ_0, with $\mathbf{Re}\,\lambda_0 > \omega$ such that $\overline{(\lambda_0 - A)D} = \mathscr{H}$.

If **H1** and **H2** are satisfied, then

$$\|S(t)x - S_n(t)x\| \to 0, \tag{6.1}$$

for each $x \in \mathscr{H}$ as $n \to \infty$ and uniformly on compact intervals in t by an application of Trotter-Kato Theorem (see [5] and [42]).

A particular case is when P_n is an orthogonal projection, in which case $P_n^* = P_n$ and $P_n \to I$ strongly. In this case, we observe the following result.

PROPOSITION 6.1. *Let $\{P_n\}_{n=1}^\infty$ be a sequence of orthogonal projectors over a complex separable Hilbert space $\mathscr{H}$ that converge strongly to the identity, $0 \le \Sigma_0 \in \mathscr{I}_p$, $F(\cdot) \in L^1(I;\mathscr{I}_p)$ and $G(\cdot) \in \mathscr{C}(I;\mathscr{I}_1)$ where $I = [0,\tau]$ for some $\tau > 0$. Then*

i. *$P_n\Sigma_0P_n \in \mathscr{I}_p$ and $\|\Sigma_0 - P_n\Sigma_0P_n\|_p \to 0$ as $n \to \infty$.*

ii. *The map $t \mapsto P_nF(t)P_n$ belongs to $L^1(I;\mathscr{I}_p)$ and*

$$\int_I \|(F - P_nFP_n)(s)\|_p\,\mathrm{d}s \to 0,$$

as $n \to \infty$.

iii. *The map $t \mapsto P_nG(t)P_n$ belongs to $\mathscr{C}(I;\mathscr{I}_1)$ and*

$$\sup_{t\in I}\|(G - P_nGP_n)(t)\|_1 \to 0,$$

as $n \to \infty$.

Proof. **i.** Let $\Sigma_0 \geq 0$ be of rank one, so $\Sigma_0 x = \langle \varphi, x\rangle \varphi$ and define $\varphi_n = P_n\varphi$. Then $P_n\Sigma_0 P_n x = \langle \varphi_n, x\rangle \varphi_n$ since $P_n^* = P_n$, for all $x \in \mathscr{H}$. Then, $(\Sigma_0 - P_n\Sigma_0 P_n)x = \langle \varphi, x\rangle \varphi - \langle \varphi_n, x\rangle \varphi_n = \langle \varphi - \varphi_n, x\rangle \varphi + \langle \varphi_n, x\rangle (\varphi - \varphi_n)$, and then

$$\begin{aligned}|\mathrm{Tr}\,(B(\Sigma_0 - P_n\Sigma_0 P_n))\,| &\leq \sum_{k=1}^{\infty} |\,\langle \varphi - \varphi_n, \phi_k\rangle\,||\,\langle \phi_k, B\varphi\rangle\,| + |\,\langle \varphi_n, \phi_k\rangle\,||\,\langle \phi_k, B(\varphi - \varphi_n)\rangle\,| \\ &\leq \|\varphi - \varphi_n\|\|B\varphi\| + \|\varphi_n\|\|B(\varphi - \varphi_n)\| \\ &\leq \|B\|_p\Big(\|\varphi - \varphi_n\|\|\varphi\| + \|\varphi_n\|\|\varphi - \varphi_n\|\Big).\end{aligned}$$

Therefore, defining $\mathscr{I}^0$ to be the set of non-zero finite rank operators, we have

$$\|\Sigma_0 - P_n\Sigma_0 P_n\|_p = \sup_{B\in\mathscr{I}^0} \frac{|\mathrm{Tr}\,(B(\Sigma_0 - P_n\Sigma_0 P_n))\,|}{\|B\|_q} \leq \|\varphi - \varphi_n\|\|\varphi\| + \|\varphi_n\|\|\varphi - \varphi_n\|.$$

Then $\|\Sigma_0 - P_n\Sigma_0 P_n\|_p \to 0$ as $n \to \infty$, since $\varphi_n \to \varphi$. If Σ_0 is of finite rank, the same result follows easily. If $0 \leq \Sigma_0 \in \mathscr{I}_p$, there is a sequence of finite rank operators $\{\Sigma_0^m\}_{m=1}^{\infty}$ such that $\|\Sigma_0 - \Sigma_0^m\|_p \to 0$ as $m \to \infty$, then

$$\begin{aligned}\|\Sigma_0 - P_n\Sigma_0 P_n\|_p &\leq \|\Sigma_0 - \Sigma_0^m\|_p + \|\Sigma_0^m - P_n\Sigma_0^m P_n\|_p + \|P_n(\Sigma_0 - \Sigma_0^m)P_n\|_p \\ &\leq 2\|\Sigma_0 - \Sigma_0^m\|_p + \|\Sigma_0^m - P_n\Sigma_0^m P_n\|_p,\end{aligned}$$

taking then $\limsup_{n\to\infty} \|\Sigma_0 - P_n\Sigma_0 P_n\|_p \leq 2\|\Sigma_0 - \Sigma_0^m\|_p$ and hence $\|\Sigma_0 - P_n\Sigma_0 P_n\|_p \to 0$ as $n \to \infty$ since $\|\Sigma_0 - \Sigma_0^m\|_p \to 0$ as $m \to \infty$.

ii. Since P_n is not time dependent, it is elementary to check that $P_n F(\cdot) P_n \in L^1(I;\mathscr{I}_p)$. Let $F(t)$ be a step function $F(t) = \sum_{k=1}^{q} f_k \chi_{I_k}(t)$, then

$$\int_I \|(F - P_n F P_n)(t)\|_p \,\mathrm{d}t \leq \sum_{k=1}^{q} \|f_k - P_n f_k P_n\|_p m(I_k),$$

and from the previous result we observe that $\int_I \|(F - P_n F P_n)(t)\|_p \to 0$ as $n \to \infty$. Since, step functions are dense in $L^1(I;\mathscr{I}_p)$ the result will follows for any $F(\cdot) \in L^1(I;\mathscr{I}_p)$. Let $\{F_m(\cdot)\}_{m=1}^{\infty}$ be a sequence of step functions in $L^1(I;\mathscr{I}_p)$ such that $\|(F - F_m)(\cdot)\|_{L^1(I;\mathscr{I}_p)} \to 0$ as $m \to \infty$. Also,

$$(F - P_n F P_n)(t) = (F - F_m)(t) + P_n(F - F_m)(t)P_n + (F_m - P_n F_m P_n)(t),$$

and then

$$\|(F - P_n F P_n)(t)\|_p \leq 2\|(F - F_m)(t)\|_p + \|(F_m - P_n F_m P_n)(t)\|_p, \tag{6.2}$$

for $t \in I$, since $\|P_n\| \leq 1$. Hence,

$$\int_I \|(F - P_n F P_n)(t)\|_p \,\mathrm{d}t \leq 2\|(F - F_m)(\cdot)\|_{L^1(I;\mathscr{I}_p)} + \int_I \|(F_m - P_n F_m P_n)(t)\|_p \,\mathrm{d}t.$$

Therefore, $\overline{\lim}_{n\to\infty} \int_I \|(F - P_n F P_n)(t)\|_p \,\mathrm{d}t \leq 2\|(F - F_m)(\cdot)\|_{L^1(I;\mathscr{I}_p)}$, but $\|(F - F_m)(\cdot)\|_{L^1(I;\mathscr{I}_p)} \to 0$ as $m \to \infty$ from which the result follows.

iii. The continuity of the map $t \mapsto P_n G(t) P_n$ follows immediately. Since $I = [0,\tau]$ is compact, step functions are dense in $\mathscr{C}(I;\mathscr{I}_p)$. So if $G(\cdot)$ is a step function $G(t) = \sum_{k=1}^{q} g_k \chi_{I_k}(t)$, then

$$\sup_{t\in I} \|(G - P_n G P_n)(t)\|_p \leq \sum_{k=1}^{q} \|g_k - P_n g_k P_n\|_p,$$

and hence from **i**. it follows that $\sup_{t\in I}\|(G-P_nGP_n)(t)\|_p\to 0$ as $n\to\infty$. The density of step functions in $\mathscr{C}(I;\mathscr{I}_p)$ implies that there is a sequence of step functions $\{G_m(\cdot)\}_{m=1}^{\infty}$ in $\mathscr{C}(I;\mathscr{I}_p)$, such that $\sup_{t\in I}\|(G-G_m)(t)\|_p\to 0$ as $m\to\infty$. The inequality in (6.2), shows that

$$\sup_{t\in I}\|(G-P_nGP_n)(t)\|_p\le 2\sup_{t\in I}\|(G-G_m)(t)\|_p+\sup_{t\in I}\|(G_m-P_nG_mP_n)(t)\|_p.$$

Then, $\overline{\lim}_{n\to\infty}\sup_{t\in I}\|(G-P_nGP_n)(t)\|_p\le 2\sup_{t\in I}\|(G-G_m)(t)\|_p$. Since $\sup_{t\in I}\|(G-G_m)(t)\|_p\to 0$, the initial claim follows. □

6.1. The Convection-Diffusion Operator Case. Let $\Omega\subset\mathbb{R}^n$ be an open bounded domain with boundary $\partial\Omega$ of Lipschitz class (for example, the open unit cube in $\mathbb{R}^n$ has Lipschitz class boundary). Consider the differential operator of order 2

$$A(x,D)=-\epsilon^2\Delta-\sum_{|\alpha|\le 1}a_\alpha(x)D^\alpha, \tag{6.3}$$

with $\epsilon>0$ and where $\Delta=D^{(2,0,0,\dots,0)}+D^{(0,2,0,\dots,0)}+\cdots+D^{(0,0,\dots,0,2)}=\sum_{k=1}^n\frac{\partial^2}{\partial x_k^2}$ is the Laplacian operator on Ω and the functions $x\mapsto a_\alpha(x)$ are smooth complex values functions on $\overline{\Omega}$. Since $\epsilon^2>0$, $A(x,D)$ is strongly elliptic of order 2 (see [49] or [42]). We define A as $Ax=A(x,D)x$ for each $x\in H^2(\Omega)\cap H_0^1(\Omega)$ and denote $\mathscr{D}(-A)=H^2(\Omega)\cap H_0^1(\Omega)$ (Note that A in this section corresponds to $-A$ in the rest of the paper, this change is preferred since it simplifies some proofs). The operator $-A$ generates a C_0-semigroup $S(t)=e^{-At}$ over $L^2(\Omega)$ (see [49]) and the unique solution to

$$\begin{aligned}\frac{\partial u(t,x)}{\partial t}+A(x,D)u(t,x)&=0, &&\text{for } t>0 \text{ and } x\in\Omega\\ u(t,x)&=0, &&\text{for } t\ge 0 \text{ and } x\in\partial\Omega\\ u(0,x)&=u_0(x), &&\text{for } u_0(\cdot)\in L^2(\Omega),\end{aligned}$$

is given by $u(t,x)=(S(t)u_0)(x)$.

It is a well know fact that the Laplacian defined as $\Delta:H^2(\Omega)\cap H_0^1(\Omega)\to L^2(\Omega)$, has eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$ that can be arranged in decreasing order $0\ge\lambda_1\ge\lambda_2\ge\cdots$ such $\lambda_k\to-\infty$ as $k\to\infty$. Also, the eigenspaces are finite-dimensional and we can choose the eigenfunctions $\{\phi_k(\cdot)\}_{k=1}^{\infty}$ to be an orthonormal basis of $L^2(\Omega)$ and they are of class $C^\infty(\Omega)$.

Define then

$$\mathcal{V}_n=\operatorname{span}\{\phi_1,\phi_2,\dots,\phi_n\},$$

and let P_n be the orthogonal projector from $L^2(\Omega)$ to $\mathcal{V}_n$. Clearly, $\mathcal{V}_n\in\mathscr{D}(-A)$ and $P_n^*P_n=P_n^2=P_n\to I$ strongly as $n\to\infty$ since

$$\|(I-P_n)\psi\|^2=\sum_{k=n+1}^{\infty}|\langle\phi_k,\psi\rangle|^2\to 0,\qquad\text{as } n\to\infty,$$

and

$$\big(\mathcal{N}(P_n)\big)^{\perp}=\big(\overline{\operatorname{span}\{\phi_{n+1},\phi_{n+2},\dots\}}\big)^{\perp}=\operatorname{span}\{\phi_1,\phi_2,\dots,\phi_n\}=\mathcal{V}_n\subset\mathscr{D}(-A).$$

A well known result states that there is a $\hat{\lambda}_0 \geq 0$ (given by the Gårding's inequality) such that $-A_{\hat{\lambda}_0} = -(A + \hat{\lambda}_0 I)$ is the infinitesimal generator of a C_0-semigroup of contractions in $L^2(\Omega)$ (see [42]), i.e., $-A_{\hat{\lambda}_0} \in \mathbf{G}(1,0)$ (recall that by definition $A \in \mathbf{G}(M,\omega)$ means A is an infinitesimal generator of a C_0-semigroup $S(t)$ such that $\|S(t)\| \leq Me^{\omega t}$ for $t \geq 0$). Then $-A = -A_{\hat{\lambda}_0} + \hat{\lambda}_0 I$ and since $\hat{\lambda}_0 I$ is a bounded operator and $\|\hat{\lambda}_0 I\| = \hat{\lambda}_0$, $-A \in \mathbf{G}(1, \hat{\lambda}_0)$. Therefore, since $\mathcal{V}_n \in \mathscr{D}(A)$, $A_n = P_n A P_n$ satisfies $-A_n \in \mathbf{G}(1, \hat{\lambda}_0)$ (see [5] for a proof). This implies that the hypothesis **H1** is satisfied, since

$$\|S_n(t)\| \leq e^{\hat{\lambda}_0 t} \text{ for all } n \in \mathbb{N} \text{ and all } t \geq 0,$$

where $S_n(t)$ is the uniformly continuous semigroup generated by $-A_n$. Even more $\|S(t)\| \leq e^{\hat{\lambda}_0 t}$ for all $t \geq 0$ where $S(t)$ is the semigroup generated by $-A$.

We are now left to prove the Hypothesis **H2**. Let D be given by finite linear combinations of the $\{\phi_k\}_{k=1}^{\infty}$; that is

$$D = \operatorname{span}\{\phi_1, \phi_2, \ldots\}.$$

Since $\{\phi_k\}_{k=1}^{\infty}$ is an orthonormal basis of $L^2(\Omega)$, D is dense in $L^2(\Omega)$. Let $x \in D$, then $x = \sum_{k=1}^{N} \langle \phi_k, x\rangle \phi_k$ for some $N < \infty$. Now let $n \geq N$. Then

$$\|Ax - A_n x\| \leq \sum_{k=1}^{N} |\langle \phi_k, x\rangle| \|A\phi_k - A_n\phi_k\| \leq \sum_{k=1}^{N} |\langle \phi_k, x\rangle| \|A\phi_k - P_n A\phi_k\| \to 0$$

as $n \to \infty$ since $P_n \to I$ strongly as $n \to \infty$ and $N < \infty$. Since $-A \in \mathbf{G}(1, \hat{\lambda}_0)$, the last condition of **H2** states that there is a complex number λ_0 with $\mathbf{Re}\,\lambda_0 > \hat{\lambda}_0$ such that $\overline{(\lambda_0 + A)D} = L^2(\Omega)$. We follow almost word for word to Pazy's analysis (see [42]) on Parabolic Equations. We observe that $A(x, D)$ is strongly elliptic of order 2 with smooth coefficients $x \mapsto a_\alpha(x)$ on $\overline{\Omega}$. If we integrate by parts, we see that for every $\lambda \in \mathbb{C}$, $\langle(\lambda + A(x,D))u, v\rangle_0$ can be extended to a *continuous sesquilinear form* $(u,v) \mapsto B(u,v)$ on $H_0^1(\Omega) \times H_0^1(\Omega)$. If $\mathbf{Re}\,\lambda \geq \hat{\lambda}_0$, then it follows from Garding's inequality that this form is *coercive*. We can then apply the Lax-Milgram lemma to derive the existence of a unique solution $u(\cdot) \in H_0^1(\Omega)$ (it can actually be proven that $u(\cdot) \in H^2(\Omega)$) of the boundary value problem

$$(\lambda + A(x,D))u = f,$$

for every $f(\cdot) \in L^2(\Omega)$ and $\mathbf{Re}\,\lambda \geq \hat{\lambda}_0$. Hence, given any $f(\cdot) \in L^2(\Omega)$ there is a $u \in H_0^1(\Omega)$ such that $B(u,v) = \langle f, v\rangle_0$ for all $v(\cdot) \in H_0^1(\Omega)$. Since D is dense in $L^2(\Omega)$, there is a sequence $u_n \in D$ such that $u_n \to u$ in $H_0^1(\Omega)$ sense. Hence $B(u_n, v) \to \langle f, v\rangle$ as $n \to \infty$ for any $v(\cdot) \in H_0^1(\Omega)$, i.e.,

$$\overline{(\lambda + A)D} = L^2(\Omega),$$

for any $\lambda \in \mathbb{C}$ with $\mathbf{Re}\,\lambda \geq \hat{\lambda}_0$.

Since **H1** and **H2** are satisfied, we observe that

$$\|S(t)x - S_n(t)x\| \to 0$$

as $n \to \infty$ for each $x \in L^2(\Omega)$ and uniformly in compact intervals where $S(t)$ is the C_0-semigroup generated by $-A$ and $S_n(t)$ are the uniformly continuous semigroup generated by $P_n(-A)P_n$ for $n = 1, 2, \ldots$.

Since $A(x,D) = -\epsilon^2\Delta + \sum_{|\alpha|\le 1} a_\alpha(x)D^\alpha$, its formal adjoint $A^*(x,D)$ is defined (see [42]) by

$$A^*(x,D)u = -\epsilon^2\Delta u - \sum_{|\alpha|\le 1} D^\alpha\big(\overline{a_\alpha(x)}u\big),$$

and it is also strongly elliptic of order 2. Since the infinitesimal generator of our semigroup $-A$ is defined as $Ax = A(x,D)x$ for each $x \in H_0^1(\Omega)\cap H^2(\Omega)$, its adjoint can be proven to be $A^*x = A^*(x,D)x$ for each $x \in H_0^1(\Omega)\cap H^2(\Omega)$ (for a proof, see Pazy's book [42]). Therefore, exactly the same analysis that was carried out before can be applied to this case to imply that

$$\|S^*(t)x - S_n^*(t)x\| \to 0$$

as $n \to \infty$, for each $x \in L^2(\Omega)$ and uniformly in compact intervals where $S^*(t)$ is the C_0-semigroup generated by $-A^*$ and $S_n^*(t)$ are the uniformly continuous semigroup generated by $P_n(-A^*)P_n$ for $n = 1, 2, \ldots$.

A similar approach, for a much wider class of parabolic systems and for a general abstract approximation framework, was first developed by Banks and Kunisch in [4]. This approach satisfies the hypotheses, in most cases, of the finite element approach. Finally we can prove convergence of the approximation scheme.

Theorem 6.2. *Let X be a complex separable Hilbert space and let Y be a complex finite dimensional Hilbert space. Let $S(t)$ be the C_0-semigroup over $\mathscr{H} = L^2(\Omega)$ generated by the strongly elliptic operator $-A$ previously described and let $S_n(t)$ be the sequence generated by $-A_n = P_n(-A)P_n$. Suppose also that*
(i) $0 \le \Sigma_0 \in \mathscr{I}_p(\mathscr{H})$.
(ii) $B(\cdot) \in L^2([0,\tau];\mathscr{I}_{2p}(X,\mathscr{H}))$.
(iii) $C(\cdot) \in \mathscr{C}([0,\tau];\mathscr{L}(\mathscr{H},Y))$.
Then, $\Sigma(\cdot) \in \mathscr{C}([0,\tau];\mathscr{I}_p(\mathscr{H}))$, the unique solution of

$$\Sigma(t) = S^*(t)\Sigma_0 S(t) + \int_0^t S^*(t-s)(BB^* - \Sigma(s)(C^*C)(s)\Sigma(s))S(t-s)\,\mathrm{d}s,$$

and the sequence of solutions $\Sigma_n(\cdot) \in \mathscr{C}([0,\tau];\mathscr{I}_p(\mathscr{H}))$ of

$$\Sigma_n(t) = S_n^*(t)\big(P_n\Sigma_0P_n\big)S_n(t) + \int_0^t S_n^*(t-s)\Big(\big(P_nBB^*P_n\big) - \Sigma_n\big(P_nC^*CP_n\big)\Sigma_n\Big)(s)S_n(t-s)\,\mathrm{d}s,$$

satisfy

$$\sup_{t\in[0,\tau]} \|\Sigma(t) - \Sigma_n(t)\|_p \to 0 \tag{6.4}$$

as $n \to \infty$.

Proof. Hypothesis **(ii)** and **(iii)** imply that $BB^*(\cdot) \in L^1([0,\tau];\mathscr{I}_p(\mathscr{H}))$ and that $C^*C(\cdot) \in \mathscr{C}([0,\tau];\mathscr{I}_1(\mathscr{H}))$ by Lemma 4.1 and Lemma 4.3. The existence and uniqueness of the mappings $t \mapsto \Sigma(t)$ and $t \mapsto \Sigma_n(t)$ are given by the Theorem 5.1.

Proposition 6.1 implies that $\|\Sigma_0 - P_n\Sigma_0P_n\|_p \to 0$, $\|(BB^* - P_nBB^*P_n)\|_{L^1([0,\tau];\mathscr{I}_p)} \to 0$ and $\|(C^*C - P_nC^*CP_n)\|_{\mathscr{C}([0,\tau];\mathscr{I}_1)} \to 0$ as $n \to \infty$. These are the hypotheses required by Theorem 5.2 which implies the claimed result. □

Finally, we have to address conditions on the sequence $\{Q_n(\cdot)\}_{n=1}^{\infty}$ and $Q(\cdot)$ under which we can observe that $\int_0^\tau \mathrm{Tr}\,(Q_n\Sigma_n)\,(t)\,\mathrm{d}t \to \int_0^\tau \mathrm{Tr}\,(Q\Sigma)\,(t)\,\mathrm{d}t$.

COROLLARY 6.3. *Assume the hypotheses of Theorem 6.2 with $p = 1$ and suppose that the sequence $\{Q_n(\cdot)\}_{n=1}^{\infty}$ and $Q(\cdot)$ are in $L^\infty([0,\tau];\mathscr{L}(\mathscr{H}))$. Let $\{\Sigma_n(\cdot)\}_{n=1}^{\infty}$ and $\Sigma(\cdot)$ be the ones in the aforementioned Theorem. Therefore,*

$$\int_0^\tau \mathrm{Tr}\,(Q_n\Sigma_n)\,(t)\,\mathrm{d}t \to \int_0^\tau \mathrm{Tr}\,(Q\Sigma)\,(t)\,\mathrm{d}t,$$

if $\|(Q-Q_n)(\cdot)\|_{L^\infty([0,\tau];\mathscr{L}(\mathscr{H}))} \to 0$ as $n \to \infty$.

Proof. We observe the inequality

$$\|(Q\Sigma - Q_n\Sigma_n)(t)\|_1 \le \|Q(t)\|\|(\Sigma-\Sigma_n)(t)\|_1 + \|\Sigma_n(t)\|_1\|(Q-Q_n)(t)\|.$$

The initial hypotheses imply that $\sup_{t\in[0,\tau]}\|(\Sigma-\Sigma_n)(t)\|_1 \to 0$ and $\|(Q-Q_n)(t)\|_{L^\infty([0,\tau];\mathscr{L}(\mathscr{H}))}$ as $n \to \infty$. Also $\sup_{t\in[0,\tau]}\|\Sigma_n(t)\|_1 \le c$ for some $c > 0$ uniformly in $n \in \mathbb{N}$, therefore

$$\left|\int_0^\tau \mathrm{Tr}\,(Q\Sigma)\,(t)\,\mathrm{d}t - \int_0^\tau \mathrm{Tr}\,(Q_n\Sigma_n)\,(t)\,\mathrm{d}t\right| \le$$
$$\tau\Big(\|Q(\cdot)\|_{L^\infty([0,\tau];\mathscr{L}(\mathscr{H}))}\sup_{t\in[0,\tau]}\|(\Sigma-\Sigma_n)(t)\|_1 + c\|(Q-Q_n)(\cdot)\|_{L^\infty([0,\tau];\mathscr{L}(\mathscr{H}))}\Big),$$

and the claimed result follows. □

7. Numerical Implementation. We consider problems with the domain $\Omega = (0,1)\times(0,1)$ or $\Omega = (0,1)\times(0,1)\times(0,1)$ and with a convection-diffusion operator A of the form (6.3) and with constant $\mathbf{a} = \{a_\alpha\}_\alpha$. In 2D, the orthonormal set of eigenfunctions of the Laplacian Δ in the unit square are given by $\psi_{m,n}(x,y) = 2\sin(\pi m x)\sin(\pi n y)$. We order them using only one parameter first according to its associated eigenvalue $\lambda_{m,n} = -\pi^2(m^2+n^2)$. In the case of two functions sharing the same eigenvalue (e.g. $\psi_{1,3}$ and $\psi_{3,1}$), we put the one with the highest m first. Therefore, we define the sequence $\{\phi_n\}_{n=1}^{\infty}$ as $\psi_{1,1},\psi_{2,1},\psi_{1,2},\psi_{2,2},\ldots$. In 3D, the orthonormal set of eigenfunctions of the Laplacian Δ in the unit cube is given by $\psi_{l,m,n}(x,y,z) = 2^{3/2}\sin(\pi l x)\sin(\pi m y)\sin(\pi n z)$. We order them using only one parameter first according to its associated eigenvalue $\lambda_{l,m,n} = -\pi^2(l^2+m^2+n^2)$. In the case of two functions sharing the same eigenvalue (e.g. $\psi_{1,3,1}$ and $\psi_{3,1,1}$), we put first the one with the highest l. In the case, they share the same l, we order them according to the greater m and so on. Therefore, we define the sequence $\{\phi_n\}_{n=1}^{\infty}$ as $\psi_{1,1,1},\psi_{2,1,1},\psi_{1,2,1},\psi_{1,1,2},\psi_{2,2,1},\ldots$.

Let P_n be the orthogonal projector onto $\mathrm{span}\{\phi_1,\phi_2,\cdots,\phi_n\}$. Since $\langle\phi_i,\phi_j\rangle = \delta_{ij}$, then the matrix representation $[A_n] \in \mathbb{R}^{n\times n}$ of the approximation $A_n = P_nAP_n$ is given by

$$[A_n]_{ij} = \epsilon^2\,\langle\phi_i,\Delta\phi_j\rangle_{L^2(\Omega)} + \mathbf{a}\cdot\langle\phi_i,\nabla\phi_j\rangle_{L^2(\Omega)},$$

where $[A_n]_{ij}$ is the i row and j column element of $[A_n]$ and each one of these can be computed exactly.

The output map, for one sensor, is defined $C(t) : L^2(\Omega) \to \mathbb{R}$ is given by $C(t)\varphi = \int_\Omega c(t,x)\varphi(x)\,\mathrm{d}x$, where $c(t,x) = K(x-\bar{x}(t))$ and $\bar{x}(t)$ is the position of the sensor. at time t. The approximation $C_n(t)$ of $C(t)$, is going to be computed as

$$C_n(t)\phi = \int_\Omega c(t,x)(P_n\varphi)(x)\,\mathrm{d}x.$$

Therefore, its matrix representation is given by

$$[C_n](t) = \begin{pmatrix} \int_\Omega c(t,x)\phi_1(x)\,\mathrm{d}x & \int_\Omega c(t,x)\phi_2(x)\,\mathrm{d}x & \cdots & \int_\Omega c(t,x)\phi_n(x)\,\mathrm{d}x \end{pmatrix}$$

Since $C^*(t) : \mathbb{R} \to L^2(\Omega)$ is given by $C^*(t)a = ac(t,x)$, it is elementary to observe that $(C_n^* C_n)(t) = P_n(C^*C)(t)P_n$.

The input map is defined as $B : \mathbb{R} \to L^2(\Omega)$ and given by $Ba = b(x)a$. Then, its adjoint, B^*, satisfies $B^* : L^2(\Omega) \to \mathbb{R}$ and it is given by $B^*\varphi = \int_\Omega b(x)\varphi(x)\,\mathrm{d}x$. The matrix representation of the approximation $(BB^*)_n = P_n BB^* P_n$ is then given by

$$[(BB^*)_n] = \begin{pmatrix} f(1,1) & f(2,1) & \cdots & f(1,n) \\ f(2,1) & f(2,2) & \cdots & f(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ f(n,1) & f(n,2) & \cdots & f(n,n) \end{pmatrix},$$

where

$$f(i,j) = \Big(\int_\Omega b(x)\phi_i(x)\,\mathrm{d}x \Big) \Big(\int_\Omega b(x)\phi_j(x)\,\mathrm{d}x \Big).$$

The approximation to the (weak) solution of

$$\dot\Sigma = A\Sigma + \Sigma A^* + BB^* - \Sigma C^* C \Sigma,$$

is provided by solving the differential matrix Riccati equation

$$\frac{\mathrm{d}}{\mathrm{d}t}[\Sigma_n] = [A_n][\Sigma_n] + [\Sigma_n][A_n]^* + [(BB^*)_n] - [\Sigma_n][C_n]^*[C_n][\Sigma_n]. \tag{7.1}$$

Although one now has the option of selecting many possible quadrature rules to numerically integrate this equation, we used an implicit Euler method since we observed convergence using a relatively large time step. In particular,

$$\frac{[\Sigma_n^{k+1}] - [\Sigma_n^k]}{h} = [A_n][\Sigma_n^{k+1}] + [\Sigma_n^{k+1}][A_n]^* + [(BB^*)_n] - [\Sigma_n^{k+1}][C_n]^*[C_n][\Sigma_n^{k+1}],$$

where $h > 0$ is the time step and $\Sigma_n^k \simeq \Sigma_n(kh)$. Rearranging terms, we observe that

$$\Big(h[A_n] - \frac{1}{2}\Big)[\Sigma_n^{k+1}] + [\Sigma_n^{k+1}]\Big(h[A_n] - \frac{1}{2}\Big)^* - [\Sigma_n^{k+1}](\sqrt{h}[C_n])^*(\sqrt{h}[C_n])[\Sigma_n^{k+1}] + \\ + \left(h[(BB^*)_n] + [\Sigma_n^k] \right) = 0.$$

Thus, the approximation in each time step is reduced to the resolution of an algebraic Riccati equation with initial condition $\Sigma_n^0 = 0$.

In both problems we use the objective functionals $J(\Sigma) = \int_0^\tau \mathrm{Tr}\,(\Sigma(t))\,\mathrm{d}t$, where $\tau = 1$ in the stationary sensor problem and $\tau = 10$ in the moving sensor network. The approximation to the objective functional is accomplished by simple quadratures on $\int_0^\tau \mathrm{Tr}\,(\Sigma)(t)\,\mathrm{d}t \simeq \sum_k h(\mathrm{Tr}\,\big([\Sigma^k]\big)$.

7.1. 3D One Stationary Sensor. We consider $\epsilon^2 = 0.01$, the kernel of the sensor is of the form $K(x) = 10e^{-|x|^2_{\mathbb{R}^3}}$ and we consider different configurations of $\mathbf{a}$ and $x \mapsto b(x)$ that determines $B : \mathbb{R} \to L^2(\Omega)$ as $Ba = b(x)a$. The time step for the Euler's method is taken as $h = 0.1$ and the maximum number of eigenfunctions used is 33 although, the number in which computations stabilize is lower and different for each $\mathbf{a}$ and $x \mapsto b(x)$.

We consider $b(x) = 100$ and $a_x = a_y = a_z = 0$. Discrepancies of results between 11 and 33 eigenfunctions are negligible. The integrals involved in the matrix approximates $[(BB^*)_n]$ and $[C_n]$ are computed with relative tolerances of 10^{-6} and 10^{-3} respectively. The minimizer in this case is found exactly at the point $(0.5, 0.5, 0.5)$ and is given by $J(0.5, 0.5, 0.5) \simeq 12$. The value of the functional increases with the distance respect to the minimizer and hence the maximum value of the functional is attained in all vertices and it is approximately 20.

We use $b(x, y, z) = 1+20 \exp(-5|x-y_0|^2_{\mathbb{R}^2})$ with $y_0 = (0.2, 0.2, 0.2)$ and $a_x = a_y = a_z = 0$; that is, the intensity of the noise is higher on the point $y_0 = (0.2, 0.2, 0.2)$ and there is no convection. As in the previous case, discrepancies of results between 11 and 33 eigenfunctions are negligible. The integrals involved in the matrix approximates $[(BB^*)_n]$ and $[C_n]$ are computed with relative tolerances of 10^{-7} and 10^{-3} respectively. The minimizer in this case is found approximately in the point $(0.4, 0.4, 0.4)$ and has value $J(0.4, 0.4, 0.4) \simeq 1$ and the values of the functional increases with the distance with respect to the point $(0.4, 0.4, 0.4)$. The highest value attained by the functional is approximately 2. Note that the minimizer has been displaced from the center towards the location of the "noisiest" place in the cube. . Note that in this case we have increased the accuracy of the integrals of the approximation $[(BB^*)_n]$ from 10^{-6} to 10^{-7}, due to the rapid decay of the function b.

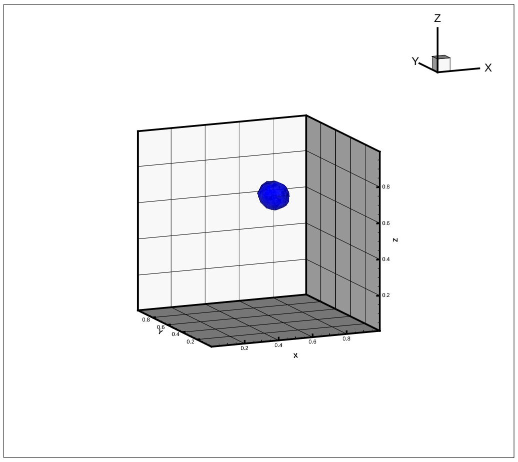
We use $b(x, y, z) = 100$ and $a_x = a_y = a_z = 20$. The number of modes used is 33, however the results using 28 up to 33 eigenfunctions show no significative difference. The integrals involved in the matrix approximates $[(BB^*)_n]$ and $[C_n]$ are computed with relative tolerances of 10^{-6} and 10^{-4} respectively. Several isosurfaces for this problem are shown on Figure 7.1. The minimizer in this case is found approximately in the point $(0.65, 0.65, 0.65)$ and has value $J(0.65, 0.65, 0.65) \simeq 42$. The highest value attained by the functional is approximately 177 and located at the origin. Note that the minimizer has been displaced from the center to a location upstream. In this case with a non-zero convective term, we require more eigenfunctions than in the previous ones (with a zero convective term) to observe convergence.

7.2. 2D Three Mobile Sensor Network. We use $\epsilon^2 = 0.01$ and the kernel of the sensor is of the form $K(x) = e^{-20|x|^2_{\mathbb{R}^2}}$ We will consider 3 sensors located initially at the points $(0.6, 0.4), (0.5, 0.5)$ and $(0.4, 0.6)$ and their trajectories are given by the integral equations $\bar{x}_i(t, u) = (x_i^0, y_i^0)^T + \int_0^t e^{\mathbf{A}(t-s)}\mathbf{b}u_i(s)\,\mathrm{d}s$ where $\mathbf{A} = \begin{pmatrix} -1 & 0.3 \\ 0 & -1 \end{pmatrix}$ and $\mathbf{b} = (1.5, -1)^T$. We assume that the initial conditions are fixed and are not a design variable.

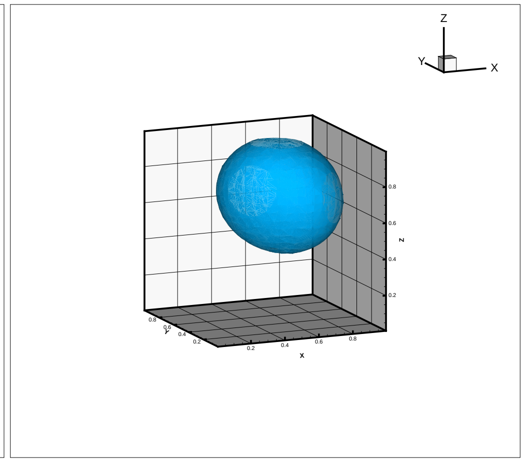
The functional to minimize is in this case is $J(u_1, u_2, u_3) = \int_0^{10} \mathrm{Tr}(\Sigma_{(u_1,u_2,u_3)}(t))\mathrm{d}t$, where $\Sigma_{(u_1,u_2,u_3)}$ refers to the solution of the Riccati equation where the output map is determined by the moving sensors with controls $(u_1, u_2, u_3) \in L^2([0, 1]) \times L^2([0, 1]) \times L^2([0, 1])$.

In order to approximate a local minimizer we will use a gradient descent method for this problem:

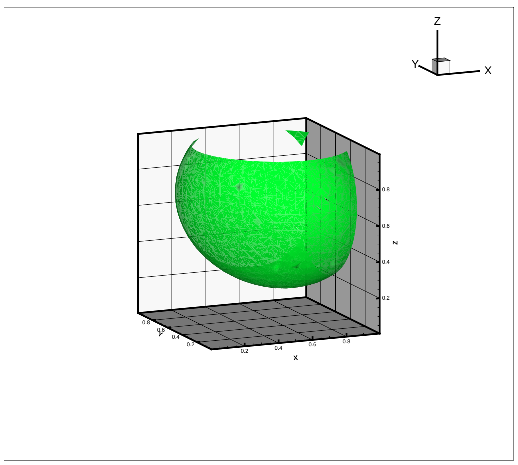
1. Start with the control with some choice $u^0(t) = (u_1^0(t), u_2^0(t), u_2^0(t))$.



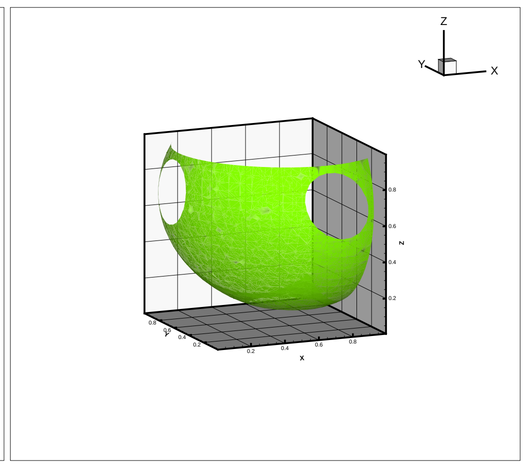
(a) Isosurface for $J(x) \simeq 43$



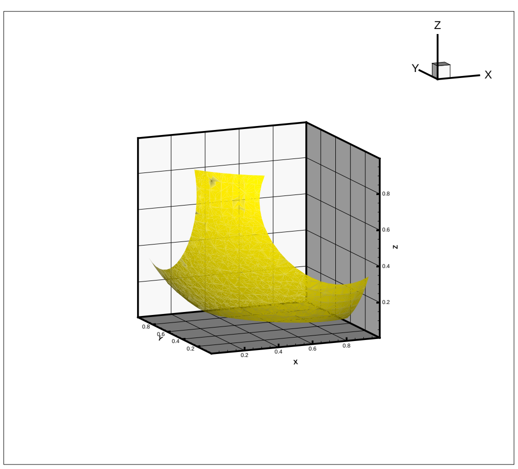
(b) Isosurface for $J(x) \simeq 50$



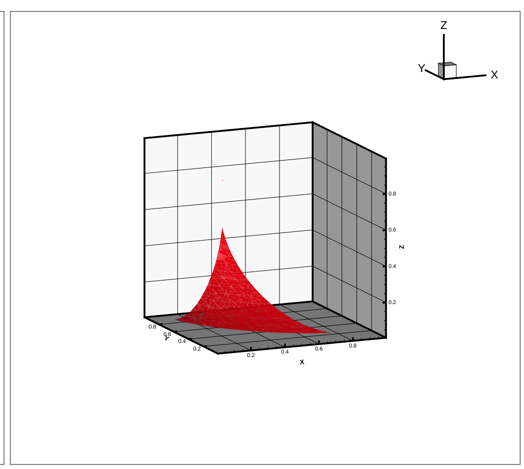
(c) Isosurface for $J(x) \simeq 61$



(d) Isosurface for $J(x) \simeq 74$



(e) Isosurface for $J(x) \simeq 87$



(f) Isosurface for $J(x) \simeq 118$

FIG. 7.1. *Isosurfaces of $J(x)$ for $b = 100$ and $a_x = a_y = a_z = 20$.*

2. Update the control as follows

$$u^{j+1}(t) = u^j(t) - \alpha_j J'(u^j)(t),$$

where $J'(u)$ is the gradient of J at u and α_n is chosen if possible as $\alpha_j = \arg\min_\alpha J(u^j - \alpha J'(u^j))$, and stop if $J(u^{j+1})$ is not decreased by at least 2% with respect to $J(u^j)$.

The termination condition for the algorithm does not involve any decrease condition on the gradient J'. This is because, in this case, there are no conditions that ensure that $J'(u^j) \to 0$ as $j \to \infty$. The computation of $\alpha_j = \arg\min_\alpha J(u^j - \alpha J'(u^j))$, is done using "brute force".

The approximation to the solution of the Riccati equation $t \mapsto \Sigma_n(t)$ is given by (7.1) The approximation to the derivative $D_{C^*C}\Sigma$ is computed using the sensitivity equation

$$\frac{\mathrm{d}}{\mathrm{d}t}[\Lambda_n] = [A_n][\Lambda_n] + [\Lambda_n][A_n]^* - [\Lambda_n][C_n(t)]^*[C_n(t)][\Sigma_n(t)] - [\Sigma_n(t)][\Sigma_n(t)] \\ - [\Sigma_n(t)][C_n(t)]^*[C_n(t)][\Lambda_n],$$

where $\Lambda_n(0) = 0$, $[A_n]$ and $t \mapsto [C_n(t)]$ are the matrix representations of the approximations to the operators A and the operator valued function $t \mapsto C(t)$, respectively.

The Fréchet derivative of $\bar{x}(t,u)$ with respect to u is given by $(D_u\bar{x}h)(t) = \int_0^t e^{\mathbf{A}(t-s)}\mathbf{b}h(s)\,\mathrm{d}s$, for each $h \in L^2([0,\tau])$ and $D_u\bar{x} \in \mathscr{L}(L^2([0,\tau]), \mathscr{C}([0,\tau];\overline{\Omega}))$. The map $C^*C : \mathscr{C}([0,\tau];\overline{\Omega}) \to \mathscr{C}([0,\tau];\mathscr{I}_1)$, is given by

$$(C^*C\varphi)(x) = K(x - \bar{x}(t)) \int_\Omega K(y - \bar{x}(t))\varphi(y)\,\mathrm{d}y.$$

Since, we use $K(x) = ae^{b\|x\|^2_{\mathbb{R}^n}}$ for some $a > 0$ and $b > 0$, then $D_{\bar{x}}K(x - \bar{x}(t))$ is well defined as a Fréchet derivative. Therefore, $D_{\bar{x}_i}C^*C \in \mathscr{L}(\mathscr{C}([0,\tau];\overline{\Omega}), \mathscr{C}([0,\tau];\mathscr{I}_1))$ is well defined as the Fréchet derivative of C^*C with respect to $\bar{x}$. Consequently, $D_uC^*C(\bar{x}(t,u)) \in \mathscr{L}(L^2([0,\tau]);\mathscr{C}([0,\tau];\mathscr{I}_1))$ is well defined. If we define $H(u) = D_uC^*C(\bar{x}(t,u))$, for $h(\cdot) \in L^2([0,\tau])$, the matrix form elements $[H_n(u)h]_{ij}$ of the approximation to $H(u)h$ are given by

$$\langle \phi_i, (H(u)h)\phi_j \rangle (t) = \\ = \int_0^t 2b\Bigg[\left(\int_\Omega K(x - \bar{x}(t))(x - \bar{x}(t))^T e^{\mathbf{A}t}\phi_i(x)\,\mathrm{d}x \right) \left(\int_\Omega K(y - \bar{x}(t))\phi_j(y)\,\mathrm{d}y \right) + \\ \left(\int_\Omega K(x - \bar{x}(t))\phi_i(x)\,\mathrm{d}x \right) \left(\int_\Omega K(y - \bar{x}(t))(y - \bar{x}(t))^T e^{\mathbf{A}t}\phi_j(y)\,\mathrm{d}y \right) \Bigg] e^{-\mathbf{A}s}\mathbf{b}h(s)\,\mathrm{d}s.$$

For the case of one sensor, $J(u) = \int_0^\tau \mathrm{Tr}\,(\Sigma(t))\,\mathrm{d}t$ has a Fréchet derivative $J'(u) \in \mathscr{L}(L^2([0,\tau]);\mathbb{R})$ and it is given by

$$J'(u)h = \int_0^\tau \mathrm{Tr}\,(\Lambda(t) \circ H(u)(t)h)\,\mathrm{d}t.$$

Hence, its approximation $(J'(u))_n$, is calculated as $(J'(u))_n h = \int_0^\tau \mathrm{Tr}(\Lambda_n(t) \circ H_n(u)(t)h)\mathrm{d}t$. and after a tedious algebraic manipulation, we obtain $(J'(u))_n h = \int_0^\tau \mathrm{Tr}(R_n(t))h(t)\mathrm{d}t$, for some $R_n(t)$. We identify $(J'(u))_n$ with $\mathrm{Tr}\,(R_n(t))$. The generalization for the case of three sensors is natural.

7.2.1. Case 1. We use $b(x, y) = 10$ and $a_x = a_y = 0$. The number of modes used is 16 and after 15 iterations the terminal condition is met. The initial and final controls are shown on Figure 7.2(a) and 7.2(c), respectively. The initial and final trajectories are shown in Figure 7.2(b) and Figure 7.2(d), respectively. Note that the initial position of the sensors is marked by a small circumference. Based on previous numerical results, in the case of one stationary sensor, the global minimizer is on the point $(0.5, 0.5)$. We should notice that the sensor with initial position in the center of the square, remains in this point for all t as we can observe on Figure 7.2(d). The other two trajectories, as we may expect, try to reach the center of the square.

7.2.2. Case 2. We use $b(x) = 10 + 10\exp(-5|x - (0.1, 0.9)|^2_{\mathbb{R}^2})$ so that the "nosiest" point on the domain is $(0.1, 0.9)$. The number of modes used is 16 and it takes 12 iterations until the termination criteria is met. The initial and final controls are shown on Figure 7.2(a) and 7.2(e), respectively. The initial and final trajectories are shown in Figure 7.2(b) and Figure 7.2(f), respectively. Note again that the initial position of the sensors is marked by a small circumference.

Based on previous numerical results, in the case of one stationary sensor, the global minimizer is on a the point in between $(0.5, 0.5)$ and $(0.1, 0.9)$ (the "noisiest" place in the square). We observe that trajectories tend to a region in between these two points (see Figure 7.2(f)).

REFERENCES

[1] W. Arendt, C. J. K. Batty, M. Hieber and F. Neubrander *Vector-valued Laplace Transforms and Cauchy Problems*, Birkhäuser Verlag, 2001.

[2] M. Athans, IFAC Congress Paris, June 1972 (1972).

[3] A. V. Balakrishnan, *Applied Functional Analysis*, Springer, New York, 1976.

[4] H. T. Banks and K. Kunisch *The Linear Regulator Problem for Parabolic Systems*, SIAM J. Controls and Optimization, Vol 22, No. 5, pp. 684-698, 1984.

[5] H. T. Banks and K. Kunisch *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, 1989.

[6] A. Bensoussan, *Optimization of Sensors' Location in a Distributed Filtering Problem*, Lecture Notes in Mathematics no. 294, Springer, New York, pp. 62-84.

[7] A. Bensoussan, *Filtrage Optimal des Systemes Lineaires*, Dunod, Paris, 1975.

[8] A. Bensoussan, G. Da Prato, M. Delfour and S. K. Mitter *Representation and Control of Infinite Dimensional Systems*, Vol. I, II, Birkhäuser, 1992.

[9] A. Bielecki *Une remarque sur la méthode de Banach-Cacciopoli-Tikhonov dans la theorie des équations différentielles ordinaires*, Bull. Acad. Polon. Sci, No. 4, Cl. III, pp. 261-264, 1956.

[10] J. A. Burns and B. B. King, A Note on the regularity of solutions of infinite dimensional Riccati equations, *Appl. Math. Lett.*, 7 (1994), 13–17.

[11] J. A. Burns and B. B. King, Optimal sensor location for robust control of distributed parameter systems, In *Proc. of the 33rd IEEE Control and Decision Conference*, Dec. 1994, 3967–3972.

[12] J. A. Burns, D. Rubio and B. B. King, Regularity of Feedback Operators for Boundary Control of Thermal Processes, In *Proc. First International Conference On Nonlinear Problems in Aviation and Aerospace*, S. Sivasundaram, Ed., Embry-Riddle Aeronautical Press, May, 1996, 67–73.

[13] J. A. Burns and B. B. King, A reduced basis approach to the design of low order feedback controllers for nonlinear continuous systems, *Journal of Vibration and Control*, 4 (1998), 297–323.

[14] J. A. Burns, E. M. Cliff, C. N. Rautenberg, *A distributed parameter control approach to optimal filtering and smoothing with mobile sensor networks*, pp.181-186, 2009 17th Mediterranean Conference on Control and Automation.

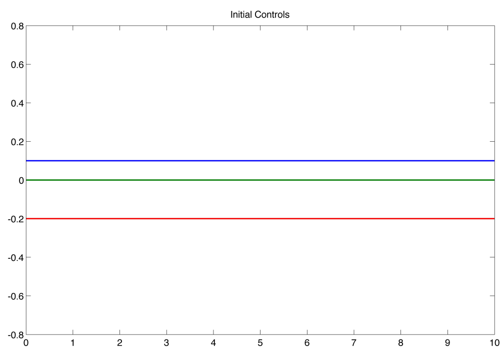
[15] J. A. Burns, C. N. Rautenberg, *Solutions and Approximations to the Riccati Integral Equation with Values in a Space of Compact Operators*, submitted.

[16] R. F. Curtain, *A Survey of Infinite-Dimensional Filtering*, SIAM Review, 17 (1975), 395-411.

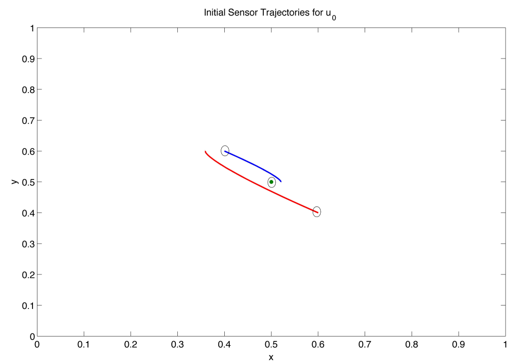
[17] R. F. Curtain, *Infinite-Dimensional Filtering*, SIAM J. Control , 13 (1975), pp. 89-104.

[18] R. F. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Springer, Berlin, 1995.
[19] G. Da Prato, *Linear Quadratic Control Theory for Infinite Dimensional Systems*, Lecture notes, Summer School on Mathematical Control Theory, Trieste, 2001.
[20] A. De Santis, A. Germani and L. Jetto, Approximation of the Algebraic Riccati Equation in the Hilbert Space of Hilbert-Schmidt Operators, *SIAM J. Control Opt.*, 31 (1993), 847–874.
[21] Sz. Dolecki, Observation for the One-dimensional Heat Equation, *Stadia Math.*, 48 (1973), 291–305.
[22] N. Dunford and J. T. Schwartz, *Linear Operators. Part II: Spectral Theory. Self Adjoint Operators in Hilbert Space*, Interscience Publishers, New York, 1963.
[23] D. E. Edmunds and W. D. Evans, *Spectral theory and differential operators*, Oxford, 2002.
[24] K-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer, 1991.
[25] A. El Jai and A. J. Pritchard, *Sensors and Actuators in the Analysis of Distributed Systems*, John Wiley, New York, 1988.
[26] A. Friedman, *Foundations of Modern Analysis*, Dover, New York, 1982.
[27] A. Germani, L. Jetto and M. Piccioni, Galerkin Approximation for Optimal Linear Filtering of Infinite-Dimensional Linear Systems, *SIAM J. Control and Opt.*, 26 (1988), 1287–1305.
[28] J. S. Gibson, The Riccati Integral Equations for Optimal Control Problems on Hilbert Spaces, *SIAM J. Control and Opt.*, 17 (1979), 537–565.
[29] I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, American Mathematical Society, 1969.
[30] P. R. Halmos, *A Hilbert Space Problem Book*, Springer, 1980.
[31] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, American Mathematical Society, 1957.
[32] K. Ito, Strong convergence and convergence rates of approximating solutions for algebraic Riccati equations in Hilbert spaces, In *Distributed Parameter Systems*, eds. F. Kappel, K. Kunisch, W. Schappacher, Springer-Verlag, (1987), 153–166.
[33] A.Yu. Khapalov, Optimal measurement trajectories for distributed parameter systems, *Systems Control Lett.*, 18 (1992), 467–477.
[34] A.Yu. Khapalov, Observability of hyperbolic systems with interior moving sensors, In *Lecture Notes in Control and Inform. Sci. 185*, R. F. Curtain, A. Bensoussan and J.-L. Lions, Eds., Springer-Verlag, New York, 1993, 489–499.
[35] A.Yu. Khapalov, L^∞-Exact Observability of the Heat Equation with Scanning Pointwise Sensor, *SIAM J. Control and Opt.*, 32 (1994), 1037–1051.
[36] A.Yu. Khapalov, Controllability of the wave equation with moving point control, *Appl. Math. Optim.*, 31 (1995), 155–175.
[37] A.Yu. Khapalov, Exact observability of the time-varying hyperbolic equation with finitely many moving internal observations, *SIAM J. Control and Opt.*, 33 (1995), 1256–1269.
[38] S. Kumar and J.H. Seinfeld, Optimal Location of Measurements for Distributed Parameter Estimation, *IEEE Trans. on Auto. Control* 23 (1978), 690-698.
[39] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, John Wiley, 1969.
[40] K. Morris, LQ-Optimal Actuator Location and Norm Convergence of Ricatti Operators, In *Proc. of the 47th IEEE Control and Decision Conference*, Dec.2008, 209–214.
[41] S.Omato, S. Koide, and T Soeda, Optimal Sensor Location Problem for a Linear Distributed Parameter System, *IEEE Trans. on Auto. Control* 23 (1978), 665-673.
[42] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, 1983.
[43] C. N. Rautenberg, *A Distributed Parameter Approach to Optimal Filtering and Estimation with Mobile Sensor Networks*, PhD dissertation, Virginia Polytechnic Institute and State University, 2010.
[44] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. Volume I: Functional Analysis*, Academic Press, 1980.
[45] J. R. Ringrose, *Compact Non-Self-Adjoint Operators*, Van Nostrand Reinhold Company, 1970.
[46] I. G. Rosen, On Hilbert-Schmidt Norm Convergence of Galerkin Approximation for Operator Riccati Equations, *International Series of Numerical Mathematics 91*, Birkhäuser, Basel, Switzerland, 1989, 335-349.
[47] Y. Sakawa, Controllability for partial differential equations of parabolic type, *SIAM J. Control*, 12 (1974), 389–400.
[48] Y. Sakawa, Observability and related problems for partial differential equations of parabolic type, *SIAM J. Control*, 13 (1975), 14–27.
[49] G. R. Sell and Y. You, *Dynamics of Evolutionary Equations*, Springer, 2002.

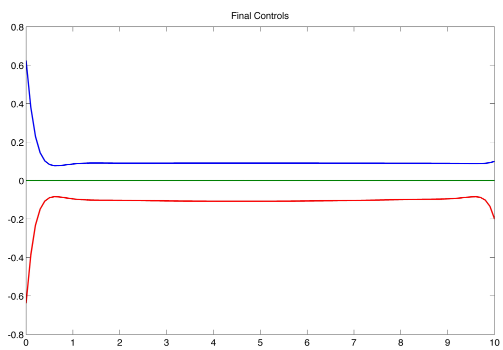
[50] B. Simon, *Trace Ideals and their Applications*, Cambridge University Press, 1979.
[51] J, Voigt, *On the Convex Compactness Property for the Strong Operator Topology*, Note Di Matematica, Vol. XII, pp. 259-269, (1992).
[52] J. Weidmann, *Linear Operators in Hilbert Space*, Springer-Verlag New York, 1980.



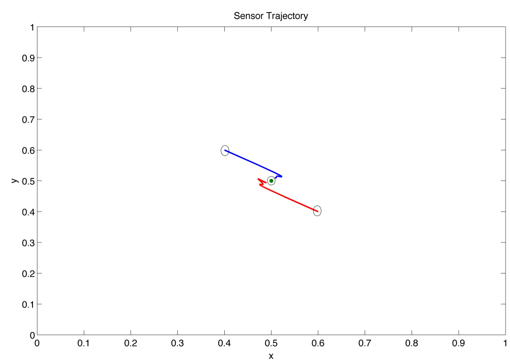
(a) Initial constant controls $(u_1^0(t), u_2^0(t), u_3^0(t))$



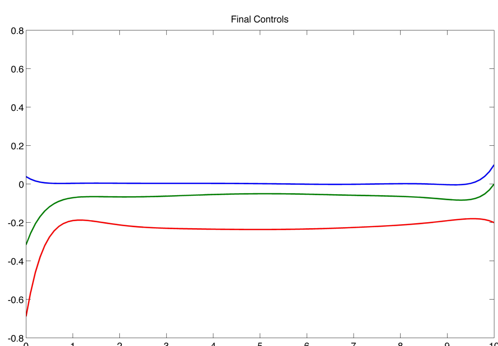
(b) Initial sensor trajectories.



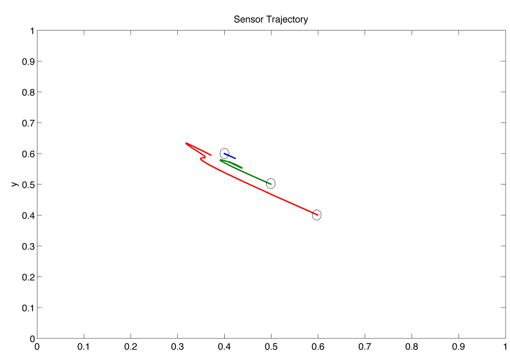
(c) Final controls for Case 1



(d) Final sensor trajectories for Case 1



(e) Final controls for Case 2



(f) Final sensor trajectories for Case 2

FIG. 7.2. *Sensor Controls and Trajectories*