

Gradient-based optimization methods for sensor & actuator placement in LTI systems

C. COLBURN, D. ZHANG, AND T. BEWLEY

University of California, San Diego, San Diego, USA
ccolburn, dazhang, bewley@ucsd.edu

(Received 14 May 2011)

This paper develops efficient techniques for calculating gradient information which may be used to optimize the placement of sensors & actuators of a given precision for the effective estimation and control of high-dimensional discretizations of infinite-dimensional linear time-invariant (LTI) systems. The necessary gradients are determined in this setting via adjoint analyses which quantify the effects of small variations of the observation and control operators. The approach can be modified appropriately to fit a variety of specific objectives within the Linear Quadratic Gaussian (LQG) estimation/control framework. Unlike other work in this area, we work directly with the covariance of the estimation error \mathbf{P} , rather than working with the Fischer information matrix \mathbf{M} , which is, in a sense, a best-case estimate of \mathbf{P}^{-1} that neglects the impact of the state disturbances on the evolution of the state estimation error. The method is tested by optimizing the placement of two sensors and two actuators in a 1D complex Ginzburg-Landau system.

1. Introduction

Sensor and actuator placement techniques for state estimation and control problems have broad applications in environmental studies, finance, and engineering. Significant applications include: actuator placement in vibration control of flexible structures (Hiramoto *et al.* 2000), sensor placement in environmental applications (Majumdar *et al.* 2002), explosion detection and contaminant plume tracking (Zhang *et al.* 2011), and estimation/control of chemical production/mixing procedures (Alonso *et al.* 2004). Although it is clear that the fidelity of the estimator in such problems is strongly dependent on the sensor locations chosen in addition to the sensor precisions used, there has been surprisingly little work on the development of rigorous, model-based, numerically-tractable algorithms for optimizing sensor placement in such high-dimensional systems.

In low-dimensional systems, this class of problems may be addressed effectively using the linear matrix inequality (LMI) formulation of Li *et al.* (2009). This approach does not actually address the placement of sensors and actuators of fixed precision, but rather assigns a cost associated with the precision of the sensors and actuators used (in pre-assigned locations), then optimizes these precisions in order to minimize this cost. By so doing, the problem is made convex. One can thus formulate and solve a problem that begins with a large number of sensors and actuators of undetermined precisions in candidate positions, then perform an LMI-based optimization of the precisions of the sensors and actuators used to minimize the cost. One can then, in an ad hoc fashion, eliminate those sensors and actuators with the smallest impact on the problem at hand, and reoptimize the precisions of the sensors and actuators that remain. The scaling of the complexity of this formulation with dimension of the system under consideration is poor; the present

gradient-based formulation might thus prove to be superior for the optimization of the placement of sensors & actuators of a given precision in high-dimensional discretizations of infinite-dimensional systems.

The majority of existing model-based sensor placement approaches considered in the literature are based on minimizing various measures of the Fisher information matrix, which is particularly convenient when considering problems of this sort (for an introduction, see, e.g., Cover & Thomas 2006, p. 392). In short, if a random variable x depends on an unknown parameter θ , then the Fisher information is a characterization of the ‘‘information’’ provided about θ via samples of x . In particular, if the probability density function (PDF) of the random variable x depends only weakly on θ , then the derivative of the conditional PDF $p(x; \theta)$ with respect to θ will be relatively small over all possible values of x , whereas if this dependence is strong, this derivative will, at least for some x , be large. Normalizing this derivative by the value of the conditional PDF itself, we define the *score* $v(x, \theta)$ as the sensitivity of the conditional PDF with respect to variation of θ for given values of x and θ ; that is,

$$v(x, \theta) = \frac{\frac{\partial}{\partial \theta} p(x; \theta)}{p(x; \theta)} = \frac{\partial}{\partial \theta} \ln p(x; \theta).$$

Note that the expected value of the score $v(x, \theta)$ over all possible values of x is zero:

$$\mathbb{E}\{v(x, \theta)\} = \int_{-\infty}^{\infty} v(x, \theta) p(x; \theta) dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} p(x; \theta) dx = \frac{\partial}{\partial \theta} \underbrace{\int_{-\infty}^{\infty} p(x; \theta) dx}_{=1} = 0.$$

It is thus the variance of the score $v(x, \theta)$ that is useful in characterizing the overall magnitude of the sensitivity of the conditional PDF with respect to variation of θ ; the *Fisher information* $M(\theta)$ is thus defined in this (scalar) case as the variance of the score:

$$M(\theta) = \mathbb{E}\{[v(x, \theta)]^2\} = \int_{-\infty}^{\infty} [v(x, \theta)]^2 p(x; \theta) dx.$$

Extension of this concept to vectors of random variables \mathbf{x} and vectors of unknown parameters $\boldsymbol{\theta}$ is straightforward, and leads immediately to the *Fisher information matrix* (FIM) $\mathbf{M}(\boldsymbol{\theta})$ via the appropriate outer product:

$$\mathbf{M}(\boldsymbol{\theta}) = \mathbb{E}\{[\mathbf{v}(\mathbf{x}, \boldsymbol{\theta})]^H [\mathbf{v}(\mathbf{x}, \boldsymbol{\theta})]\} = \int [\mathbf{v}(\mathbf{x}, \boldsymbol{\theta})]^H [\mathbf{v}(\mathbf{x}, \boldsymbol{\theta})] p(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x},$$

where $\mathbf{v}(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \ln p(\mathbf{x}; \boldsymbol{\theta})$ is considered to be a row vector and $[\cdot]^H$ denotes the conjugate transpose.

The inverse of the Fisher information matrix provides a lower bound of the estimation error covariance matrix \mathbf{P} of the Kalman filter via the Cramér-Rao inequality

$$\mathbf{P} \geq \mathbf{M}^{-1}; \quad (1.1)$$

derivation of this important bound is given in Goodwin & Payne 1977, Theorem 1.3.1.

As mentioned previously, scalar measures of the Fisher information matrix are typically considered when optimizing sensor locations. Three common such measures are

- A-optimality (trace):

$$J_A(\mathbf{M}) = \text{trace}(\mathbf{M}^{-1}), \quad (1.2a)$$

- D-optimality (determinant):

$$J_D(\mathbf{M}) = -\ln \det(\mathbf{M}), \quad (1.2b)$$

- E-optimality (max eigenvalue):

$$J_E(\mathbf{M}) = \lambda_{max}(\mathbf{M}^{-1}). \quad (1.2c)$$

Uciński (2005) reviews these cost functions and summarizes the impact of this choice on the overall optimization problem: “A D-optimum design minimizes the volume of the uncertainty ellipsoid for the estimates. An E-optimum design minimizes the length of the largest axis of the same ellipsoid. An A-optimum design suppresses the average variance of the estimates.”

Work on the sensor placement problem has focused heavily on the three cost functions listed above. For example, Faulds & King (2000) considered A-optimality measures of the FIM to analyze (but not optimize) a model-free method for placing sensors in the domain of the 2D heat equation using Centroidal Voronoi Tessellations (CVT). Similarly, Martínez & Bullo (2006) found methods for minimizing D-optimality measures of the FIM in target tracking problems. These two formulations are particularly attractive because they can be solved in a distributed framework (Cortes *et al.* 2004; Bullo & Cortes 2004; Kwok & Martinez 2010). However attractive these distributed formulations are, computational experiments in 2D Navier-Stokes systems (Zhang *et al.* 2011) indicate that centralized formulations which optimize sensor vehicle trajectories specifically targeting regions of high estimation uncertainty in a model predictive control setting generally provide superior estimator performance than CVT-based formulations. Porat & Nehorai (1996) propose a source-seeking estimation/tracking algorithm based on A-optimality measures of the FIM which seek to optimize sensor locations for estimating a contaminant source location via a relatively inefficient global search over feasible future measurement locations. Because this method scales poorly with problem size, they augment the algorithm by calculating gradients of the FIM at select locations within the feasible set for each sensor; the authors propose this optimization in a receding horizon setting, where measurement locations eventually converge to stationary points in the domain. In the robust setting, Flaherty *et al.* (2006) use E-optimality measures of the FIM to estimate parameter values in models of biological systems.

Although optimization of sensor locations via consideration of the FIM is convenient, it is somewhat unfortunate that Cramér-Rao only relates the FIM to a *lower* bound on the quantity of interest (that is, the covariance of the estimation error); though it is evident that the covariance of the estimation error and the inverse of FIM are somehow related, optimizing the sensor locations based on the FIM in fact provides no guarantees (upper bounds) on the resulting covariance of the estimation error. In fact, the covariance of the estimation error of the Kalman Filter only approaches the lower bound provided by the FIM in the limit that the state disturbances of the system model are made small. To illustrate this, Taylor (1979) puts the Kalman Filter for continuous time systems with discrete time measurements in context with the FIM and the Cramér-Rao bound. Assuming a continuous-time state transition matrix satisfying the differential equation

$$\frac{d\Phi(t, t_0)}{dt} = \mathbf{A}(t)\Phi(t, t_0), \quad (1.3a)$$

defining $\Phi_{k+1,k} = \Phi(t_{k+1}, t_k)$, and taking the initial condition $\Phi(t, t) = \mathbf{I}$, Taylor (1979) showed that the FIM for the discrete-time Kalman Filter can be written in the form

$$\mathbf{M}(t_k) = \Phi_{k,k-1}^{-H} \mathbf{M}(t_{k-1}) \Phi_{k,k-1}^{-1} + \mathbf{H}^H \mathbf{V}^{-1} \mathbf{H}. \quad (1.4)$$

Comparing this result with the propagation of the discrete-time information filter[†],

$$\mathbf{P}_{k|k}^{-1} = (\boldsymbol{\Phi}_{k,k-1} \mathbf{P}_{k-1|k-1} \boldsymbol{\Phi}_{k,k-1}^H + \mathbf{W})^{-1} + \mathbf{H}^H \mathbf{V}^{-1} \mathbf{H}, \quad (1.5)$$

where \mathbf{W} is the covariance of the state disturbances added to the state evolution equation, it is clear that, in the limit in which the state evolution is deterministic ($\mathbf{W} \rightarrow 0$), the covariance of the state estimation error approaches the lower bound predicted by the Crámer-Rao inequality.

Adjoint-based variational methods provide a powerful and broadly extensible framework for optimization problems of this sort, and scale well to high-dimensional discretizations of infinite-dimensional systems. Via successive linearization, they can also be used to optimize problems outside of the somewhat restrictive linear/quadratic setting. Note that adjoint-based optimizations have been applied broadly for shape design in aerodynamic systems (Jameson *et al.* 1998; Giles & Pierce 2000), adaptive grid refinement for error reduction in CFD simulations (Giles 1998), and a host of other practical applications. However, such methods have not yet been extended to optimize sensor distributions in fluid systems; the present work seeks to fill this void. Furthermore, only a few investigations have used adjoint methods to evaluate the sensitivity of solutions to Riccati equations. Specifically, De Farias *et al.* (2001, Appendix A) propose a strategy similar to that used here for extracting gradients while performing optimizations of LMIs, and Kenney & Hewer (1990) examined how solutions to Riccati equations change as a result of modeling errors in the actuation/measurement covariance matrix.

The remainder of this paper develops various adjoint-based methods for minimizing relevant scalar measures in the control and estimation problems: §2 presents this analysis for continuous-time systems, §3 performs the equivalent discrete-time analysis, and §4 presents an application of the continuous-time theory to the 1D complex Ginzburg-Landau equation.

2. Continuous-time Analyses

Consider a continuous-time Linear Time Invariant (LTI) system described by

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{B}(\mathbf{q}_a)\mathbf{u} + \mathbf{w}, \quad (2.1a)$$

$$\mathbf{y} = \mathbf{C}(\mathbf{q}_s)\mathbf{x} + \mathbf{v}, \quad (2.1b)$$

where[‡] $\mathbf{x}(t) \in \mathbb{C}^n$ is the state, $\mathbf{u}(t) \in \mathbb{C}^\ell$ is the control, $\mathbf{y}(t) \in \mathbb{C}^m$ is the measurement, $\mathbf{w}(t) \in \mathbb{C}^n$ is the state disturbance, $\mathbf{v}(t) \in \mathbb{C}^m$ is the measurement noise, $\mathbf{q}_a(t)$ parameterizes the actuator positions, and $\mathbf{q}_s(t)$ parameterizes the sensor positions. For simplicity below, we make the standard modeling assumptions that $\mathbf{w}(t)$ and $\mathbf{v}(t)$ are *uncorrelated, zero-mean, white continuous-time random processes* with spectral densities $\mathbf{W} \geq \mathbf{0}$ and $\mathbf{V} > \mathbf{0}$ respectively. The (bounded) functional dependence of the operators \mathbf{B} and \mathbf{C} on the actuator and sensor positions \mathbf{q}_a and \mathbf{q}_s is emphasized explicitly above, but is suppressed below for notational clarity. In the discussion below, we first treat the optimization of the sensor positions \mathbf{q}_s , then the optimization of the actuator positions \mathbf{q}_a ; the development of the gradient information necessary to optimize both actuator and sensor positions simultaneously follows similarly, and is discussed further in §2.3.

[†] The information filter is not to be confused with the FIM. The information filter is the propagation and update of the information matrix, which is defined as the inverse of the Kalman Filter covariance \mathbf{P} ; see, Anderson & Moore (1979).

[‡] For generality, our formulations are developed in §2 and §3 and tested in §4 on complex systems; it is trivial to restrict these formulations to the (more typical) setting of real systems.

2.1. Computing a gradient with respect to the sensor positions

Via standard (continuous-time) Kalman-Bucy filter theory, the best linear unbiased estimate $\hat{\mathbf{x}}(t)$ of the system (2.1) is given by

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = (\mathbf{A} - \mathbf{L}\mathbf{C})\hat{\mathbf{x}}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{L}(t)\mathbf{y}(t), \quad \mathbf{L}(t) = \mathbf{P}(t)\mathbf{C}^H\mathbf{V}^{-1}, \quad (2.2)$$

where the covariance $\mathbf{P}(t) = \mathbb{E}\{\tilde{\mathbf{x}}(t)\tilde{\mathbf{x}}^H(t)\} \geq 0$ of the estimation error $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t)$ evolves forward in time from given initial conditions $\mathbf{P}(0)$ according to the differential Riccati equation (DRE)

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{A}\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^H + \mathbf{W} - \mathbf{L}(t)\mathbf{V}\mathbf{L}^H(t). \quad (2.3)$$

This evolution equation for $\mathbf{P}(t)$ reveals that, as the estimator (2.2) evolves in time, the estimation error covariance $\mathbf{P}(t)$ is driven larger by the unmodelled state disturbances $\mathbf{w}(t)$ in the system (2.1), and is driven smaller by the feedback term $\mathbf{L}(t)\mathbf{y}(t)$ in the estimator (2.2). Given that the DRE (2.3) marches to a finite value at $t = T$ as $T \rightarrow \infty$ [that is, that the system (2.1) is *detectable*], the infinite-horizon solution of the DRE (2.3) may be computed directly by setting $d\mathbf{P}/dt = 0$, thus transforming the DRE (2.3) into the continuous-time algebraic Riccati equation (CARE)

$$\mathbf{0} = \mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^H + \mathbf{W} - \mathbf{L}\mathbf{V}\mathbf{L}^H, \quad \mathbf{L} = \mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}. \quad (2.4a)$$

Closed-form solutions to a CARE such as (2.4a) are generally unavailable, and thus iterative methods based on the Schur decomposition of a $2n \times 2n$ Hamiltonian matrix are typically used to solve them (Kailath 1980).

The matrices \mathbf{A} , \mathbf{W} , and \mathbf{V} in this LTI formulation are assumed to be given. The remaining matrix which affects \mathbf{P} in (2.4a) is \mathbf{C} , which is, in turn, a function of the sensor positions \mathbf{q}_s . Thus, an optimization problem may be posed to minimize some measure of \mathbf{P} in the infinite-horizon problem (2.4a) with respect to the (stationary) sensor locations \mathbf{q}_s . In particular, we will seek the optimal \mathbf{q}_s which minimizes the cost

$$J(\mathbf{q}_s) = \text{trace}(\mathbf{P}); \quad (2.4b)$$

alternative formulations based on different measures of \mathbf{P} are considered in §2.3.

The gradient-based optimization problem we develop here focuses on the selection of \mathbf{q}_s to minimize the cost $J(\mathbf{q}_s)$ in (2.4b), where J is related to \mathbf{q}_s via solution of the CARE (2.4a). We first select an initial \mathbf{q}_s essentially arbitrarily, subject only to the technical condition that (2.1) be detectable. Local gradients of the cost J with respect to the sensor positions \mathbf{q}_s are then iteratively optimized via a standard gradient-based minimization algorithm. The algebraically difficult step of this formulation is the efficient computation of the gradient $\nabla_{\mathbf{q}_s} J$.

A simple approach to computing the necessary gradient in this problem might be to apply a finite difference method or the (more accurate) complex-step derivative method to each component of each of the m sensor locations individually, solve a perturbation problem for each, then synthesize the results to assemble the gradient (see, e.g., Chen & Rowley 2010). An accurate and significantly more computationally efficient approach to compute the gradient is to instead perform a single adjoint computation, as discussed in detail below.

Starting from an initial set of sensor locations \mathbf{q}_s and the corresponding observation matrix \mathbf{C} , associated CARE solution \mathbf{P} , and cost J , consider the following Taylor series

expansion of (2.4b) about \mathbf{q}_s :

$$J(\mathbf{q}_s + \mathbf{q}'_s) = J(\mathbf{q}_s) + (\nabla_{\mathbf{q}_s} J)^H \mathbf{q}'_s + \dots = J(\mathbf{q}_s) + J'(\mathbf{q}_s, \mathbf{q}') + \dots \quad (2.5)$$

The constraint given in (2.4a) implies that a small perturbation of \mathbf{q}_s yields a small perturbation of J ; this expansion can also be written explicitly as a function of \mathbf{P}' :

$$J(\mathbf{q}_s + \mathbf{q}'_s) = \text{trace}(\mathbf{P} + \mathbf{P}') = \text{trace}(\mathbf{P}) + \text{trace}(\mathbf{P}') \quad \Rightarrow \quad J'(\mathbf{q}_s, \mathbf{q}') = \text{trace}(\mathbf{P}'). \quad (2.6)$$

The corresponding equations for the perturbation matrix \mathbf{P}' are

$$\mathbf{A}\mathbf{P}' + \mathbf{P}'\mathbf{A}^H - \mathbf{P}'\mathbf{C}^H\mathbf{L}^H - \mathbf{L}\mathbf{C}\mathbf{P}' = \mathbf{P}(\mathbf{C}')^H\mathbf{L}^H + \mathbf{L}\mathbf{C}'\mathbf{P} \quad (2.7a)$$

where

$$\mathbf{C}' = \left(\frac{d\mathbf{C}}{d\mathbf{q}_s} \right)^H \mathbf{q}'_s.$$

In general, the matrix \mathbf{C}' is a contraction of the rank-3 tensor $d\mathbf{C}/d\mathbf{q}_s$ with a vector of sensor perturbations \mathbf{q}'_s . For notational convenience, (2.7a) is written as two linear operations $\mathbf{U}(\mathbf{P}')$ and $\mathbf{V}(\mathbf{C}')$ such that

$$\mathbf{U}(\mathbf{P}') = \mathbf{V}(\mathbf{C}') \quad (2.7b)$$

where

$$\begin{aligned} \mathbf{U}(\mathbf{P}') &= \mathbf{A}\mathbf{P}' + \mathbf{P}'\mathbf{A}^H - \mathbf{P}'\mathbf{C}^H\mathbf{L}^H - \mathbf{L}\mathbf{C}\mathbf{P}', \\ \mathbf{V}(\mathbf{C}') &= \mathbf{P}(\mathbf{C}')^H\mathbf{L}^H + \mathbf{L}\mathbf{C}'\mathbf{P}. \end{aligned}$$

Comparing the right-hand-sides of (2.5) and (2.6) reveals that

$$(\nabla_{\mathbf{q}_s} J)^H \mathbf{q}'_s = \text{trace}(\mathbf{P}'). \quad (2.8)$$

The relationship between \mathbf{P}' and \mathbf{q}'_s can thus be used to compute the gradient. To proceed, define an appropriate matrix inner-product[†]

$$\langle \mathbf{X}, \mathbf{Z} \rangle = \Re[\text{trace}(\mathbf{X}^H \mathbf{Z})] \quad (2.9)$$

(where $\Re[\cdot]$ and $\Im[\cdot]$ denote the real and imaginary part, respectively) along with a matrix adjoint variable \mathbf{S} and an adjoint operator $\mathbf{U}^*(\cdot)$ defined such that

$$\begin{aligned} \langle \mathbf{S}, \mathbf{U}(\mathbf{P}') \rangle &= \langle \mathbf{U}^*(\mathbf{S}), \mathbf{P}' \rangle \quad (2.10) \\ \Rightarrow \quad \mathbf{U}^*(\mathbf{S}) &= \mathbf{A}^H \mathbf{S} + \mathbf{S} \mathbf{A} - \mathbf{S} \mathbf{L} \mathbf{C} - \mathbf{C}^H \mathbf{L}^H \mathbf{S} = (\mathbf{A} - \mathbf{L} \mathbf{C})^H \mathbf{S} + \mathbf{S} (\mathbf{A} - \mathbf{L} \mathbf{C}). \end{aligned}$$

Recognizing that the perturbation to the cost function (2.4b) may be expressed using (2.9), it follows from (2.7b) and (2.10) that, if $\mathbf{U}^*(\mathbf{S}) = \mathbf{I}$, then the first-order perturbation to the cost is exactly

$$\begin{aligned} J'(\mathbf{q}_s, \mathbf{q}') &= \text{trace}(\mathbf{P}') = \langle \mathbf{I}, \mathbf{P}' \rangle = \langle \mathbf{U}^*(\mathbf{S}), \mathbf{P}' \rangle = \langle \mathbf{S}, \mathbf{U}(\mathbf{P}') \rangle = \langle \mathbf{S}, \mathbf{V}(\mathbf{C}') \rangle \\ &= \Re \left[\text{trace} \left(2\mathbf{P} \mathbf{S} \mathbf{L} \frac{d\mathbf{C}}{d\mathbf{q}_s} \mathbf{q}'_s \right) \right] \\ &= \text{trace} \left(\Re \left[2\mathbf{P} \mathbf{S} \mathbf{L} \frac{d\mathbf{C}}{d\mathbf{q}_s} \right] \Re[\mathbf{q}'_s] \right) - \text{trace} \left(\Im \left[2\mathbf{P} \mathbf{S} \mathbf{L} \frac{d\mathbf{C}}{d\mathbf{q}_s} \right] \Im[\mathbf{q}'_s] \right). \quad (2.11) \end{aligned}$$

Thus, the gradient of the cost function (2.4b) with respect to the i 'th element of the

[†] Note that, if all matrices are real, (2.9) and (2.11) simplify, and the conjugate operation may be dropped from (2.12a).

sensor positions vector q_s^i can be extracted:

$$\nabla_{q_s^i} J = \text{trace} \left(\overline{2\mathbf{P}\mathbf{S}\mathbf{L} \frac{d\mathbf{C}}{dq_s^i}} \right), \quad (2.12a)$$

where the overbar denotes the complex conjugate, and where \mathbf{S} satisfies the associated continuous-time algebraic Lyapunov equation (CALE)

$$(\mathbf{A} - \mathbf{L}\mathbf{C})^H \mathbf{S} + \mathbf{S}(\mathbf{A} - \mathbf{L}\mathbf{C}) = \mathbf{I}. \quad (2.12b)$$

2.2. Computing a gradient with respect to the actuator positions

Following an analogous approach as that developed above for the sensor placement problem, we now consider the corresponding actuator placement problem. Standard continuous-time optimal control theory applied to the linear system (2.1) establishes that the cost function

$$J = \frac{1}{2} \int_0^T [\mathbf{x}^H(t) \mathbf{Q} \mathbf{x}(t) + \mathbf{u}^H(t) \mathbf{R} \mathbf{u}(t)] dt + \frac{1}{2} \mathbf{x}^H(T) \mathbf{Q}_T \mathbf{x}(T) \quad (2.13)$$

is minimized by the full-state feedback control policy

$$\mathbf{u}(t) = -\mathbf{K}(t) \mathbf{x}(t), \quad \mathbf{K}(t) = \mathbf{R}^{-1} \mathbf{B}^H \mathbf{Y}(t), \quad (2.14)$$

where the “cost-to-go” matrix $\mathbf{Y}(t) \geq 0$ evolves backward in time from the terminal condition $\mathbf{Y}(T) = \mathbf{Q}_T$ according to the DRE [cf. (2.3)]

$$-\frac{d\mathbf{Y}(t)}{dt} = \mathbf{A}^H \mathbf{Y}(t) + \mathbf{Y}(t) \mathbf{A} + \mathbf{Q} - \mathbf{K}^H(t) \mathbf{R} \mathbf{K}(t). \quad (2.15)$$

We identify $\mathbf{Y}(t)$ as a “cost-to-go” matrix because it can be shown that

$$\begin{aligned} J(\tau) &= \frac{1}{2} \int_\tau^T [\mathbf{x}^H(t) \mathbf{Q} \mathbf{x}(t) + \mathbf{u}^H(t) \mathbf{R} \mathbf{u}(t)] dt + \frac{1}{2} \mathbf{x}^H(T) \mathbf{Q}_T \mathbf{x}(T) \\ &= \frac{1}{2} \mathbf{x}^H(\tau) \mathbf{Y}(\tau) \mathbf{x}(\tau). \end{aligned} \quad (2.16)$$

Given that the DRE (2.15) marches to a finite value at $t = 0$ as $T \rightarrow \infty$ [that is, that the system (2.1) is *stabilizable*], the infinite-horizon solution of the DRE (2.15) may be computed directly by setting $d\mathbf{Y}/dt = 0$, thus transforming the DRE (2.15) into the CARE [cf. (2.4a)]

$$\mathbf{0} = \mathbf{A}^H \mathbf{Y} + \mathbf{Y} \mathbf{A} + \mathbf{Q} - \mathbf{K}^H \mathbf{R} \mathbf{K}, \quad \mathbf{K} = \mathbf{R}^{-1} \mathbf{B}^H \mathbf{Y}. \quad (2.17a)$$

The matrices \mathbf{A} , \mathbf{Q} , and \mathbf{R} in this LTI formulation are assumed to be given. The remaining matrix which affects \mathbf{Y} in (2.17a) is \mathbf{B} , which is, in turn, a function of the actuator positions \mathbf{q}_a . Note that (2.16) evaluated at $\tau = 0$ in the infinite-horizon limit $T \rightarrow \infty$ implies that the original cost metric in (2.13) is minimized when the eigenvalues of the symmetric matrix \mathbf{Y} are minimized. Towards this end, the following cost function may be proposed [cf. (2.4b)]

$$\min_{\mathbf{q}_a} J = \text{trace}(\mathbf{Y}). \quad (2.17b)$$

The actuator positions \mathbf{q}_a can now be iteratively optimized following an essentially identical analysis to that presented in §2.1. The resulting expression for the gradient [cf. (2.12a)] is

$$\nabla_{q_a^i} J = \text{trace} \left(\overline{2\mathbf{K}\mathbf{T}\mathbf{Y} \frac{d\mathbf{B}}{dq_a^i}} \right), \quad (2.18a)$$

where the matrix adjoint \mathbf{T} satisfies the associated CALE [cf. (2.12b)]

$$(\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{T} + \mathbf{T}(\mathbf{A} - \mathbf{B}\mathbf{K})^H = \mathbf{I}. \quad (2.18b)$$

2.3. Discussion

As mentioned previously, the gradient-based optimizations discussed in §2.1 and §2.2 first select initial sensor and actuator positions essentially arbitrarily, subject only to the technical conditions that (2.1) be detectable and stabilizable†. Local gradients of a relevant cost J with respect to the sensor and actuator positions are then successively calculated and used to efficiently (but locally) optimize the sensor and actuator locations via a standard minimization algorithm such as steepest descent or the nonquadratic conjugate gradient method. The algebraically difficult step is the efficient computation of the necessary gradients, which has been shown in both cases to arise in a straightforward fashion from the standard CARE for the estimation or control problem, together with an associated CALE to compute an adjoint matrix upon which the required gradient is based.

From the analyses performed in §2.1 and §2.2, it is clear that the RHS forcing in (2.12b) and (2.18b) is determined solely by the definition of the cost function. Alternative cost functions can also easily be considered, such as those appearing in (1.2a)-(1.2c) with the estimation error covariance \mathbf{P} , which the quantity of interest here, replacing the inverse of the FIM. Taking $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ as the eigenvalues of \mathbf{P} , a summary outlining the key results is sufficient to clarify‡:

- A-optimality (trace):

$$\begin{aligned} J_A &= \text{trace}(\mathbf{P}) = \lambda_1 + \dots + \lambda_n \\ \Rightarrow J'_A &= \text{trace}(\mathbf{P}') = \langle \mathbf{I}, \mathbf{P}' \rangle \\ \Rightarrow \mathbf{U}^*(\mathbf{S}) &= \mathbf{I}. \end{aligned} \tag{2.19a}$$

- D-optimality (determinant):

$$\begin{aligned} J_D &= -\ln \det(\mathbf{P}^{-1}) = \ln \det(\mathbf{P}) \\ \Rightarrow J'_D &= \text{trace}(\mathbf{P}^{-1} \mathbf{P}') = \langle \mathbf{P}^{-1}, \mathbf{P}' \rangle \\ \Rightarrow \mathbf{U}^*(\mathbf{S}) &= \mathbf{P}^{-1}. \end{aligned} \tag{2.19b}$$

- E-optimality (max eigenvalue, λ_1 , with corresponding eigenvector \mathbf{r}_1)

$$\begin{aligned} J_E &= \lambda_1(\mathbf{P}) \\ \Rightarrow J'_E &= \text{trace}(\mathbf{r}_1 \mathbf{r}_1^H \mathbf{P}') = \langle \mathbf{r}_1 \mathbf{r}_1^H, \mathbf{P}' \rangle \\ \Rightarrow \mathbf{U}^*(\mathbf{S}) &= \mathbf{r}_1 \mathbf{r}_1^H. \end{aligned} \tag{2.19c}$$

Another metric of interest is the square of the Frobenius norm of \mathbf{P} :

$$\begin{aligned} J_F &= \|\mathbf{P}\|_F^2 = \text{trace}(\mathbf{P}\mathbf{P}) = \sum_i \sum_j |p_{ij}|^2 = \lambda_1^2 + \dots + \lambda_n^2 \\ \Rightarrow J'_F &= \text{trace}(2\mathbf{P}\mathbf{P}') = \langle 2\mathbf{P}, \mathbf{P}' \rangle \\ \Rightarrow \mathbf{U}^*(\mathbf{S}) &= 2\mathbf{P}. \end{aligned} \tag{2.19d}$$

It has been shown that the adjoint method of computing the gradient in this class of problems is readily extensible to a broad range of different cost functions, with the only difference between the various cases being the RHS forcing of the associated adjoint.

† Lauga & Bewley (2003) showed that detectability and stabilizability are lost gradually in systems of this sort when the sensors and actuators are moved outside of the physical domain of interest (that is, where the significant dynamics of the open-loop PDE system take place), which can lead to numerical problems when using finite-precision arithmetic. It is thus advisable to choose reasonable initial placements of the sensors and actuators, well within the regions of significant dynamics of the open-loop system.

‡ Matrix identities from Petersen & Pedersen (2008) and Horn & Johnson (1990) are used in the analyses leading to (2.19b) and (2.19c), respectively.

The perhaps most notable extension to consider is the cost function associated with the \mathcal{H}_2 control problem. This problem is well known and has been studied extensively (see, e.g., Kwakernaak & Sivan (1972), Zhou & Doyle (1998), and Hassibi *et al.* (1999)). We proceed by appending the state equation with an additional output \mathbf{z} identifying the states of interest in the control problem

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{B}(\mathbf{q}_a)\mathbf{u} + \mathbf{B}_1\mathbf{w}, \quad (2.20a)$$

$$\mathbf{y} = \mathbf{C}(\mathbf{q}_s)\mathbf{x} + \mathbf{v}, \quad (2.20b)$$

$$\mathbf{z} = \mathbf{C}_1\mathbf{x}. \quad (2.20c)$$

Following the \mathcal{H}_2 approach, an estimate $\hat{\mathbf{x}}$ of the state \mathbf{x} is first developed, based on the measurements \mathbf{y} , as discussed in §2.1, then a full state feedback controller $\mathbf{u} = -\mathbf{K}\hat{\mathbf{x}}$ is developed, as discussed in §2.2. These two components are then connected by making the control feedback depend on the state estimate, $\mathbf{u} = -\mathbf{K}\hat{\mathbf{x}}$, rather than the state itself. Consolidating the disturbance vector $\mathbf{d} = [\mathbf{w}; \mathbf{v}]$ for the purpose of analysis, a new cost function can be written to characterize the \mathcal{H}_2 -norm of the closed-loop transfer function:

$$J_{\mathcal{H}_2} = \|T_{zd}\|_2^2 = \text{trace}(\mathbf{C}_1\mathbf{P}\mathbf{C}_1^H) + \text{trace}(\mathbf{V}^{-1}\mathbf{C}\mathbf{P}\mathbf{Y}\mathbf{P}\mathbf{C}^H), \quad (2.21a)$$

$$= \text{trace}(\mathbf{B}_1^H\mathbf{Y}\mathbf{B}_1) + \text{trace}(\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}\mathbf{P}\mathbf{Y}\mathbf{B}), \quad (2.21b)$$

where \mathbf{P} and \mathbf{Y} are the solutions to the CAREs (2.4a) and (2.17a), respectively. [As briefly mentioned above, this sensor/actuator placement problem was addressed by Chen & Rowley (2010), but will be reconsidered here since the gradient calculation is different.] By averaging (2.21a) and (2.21b) and then performing perturbation analysis on the averaged equations, the perturbation of $J_{\mathcal{H}_2}$ can be written

$$\begin{aligned} J'_{\mathcal{H}_2} &= \text{trace}(\mathbf{P}\mathbf{Y}\mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}') + \text{trace}(\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}\mathbf{P}\mathbf{Y}\mathbf{B}') \\ &\quad + \frac{1}{2}\text{trace}([\mathbf{C}_1^H\mathbf{C}_1 + \mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}\mathbf{P}\mathbf{Y} + \mathbf{Y}\mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}\mathbf{C} + \mathbf{Y}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}]\mathbf{P}') \\ &\quad + \frac{1}{2}\text{trace}([\mathbf{B}_1\mathbf{B}_1^H + \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}\mathbf{P} + \mathbf{P}\mathbf{Y}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H + \mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}\mathbf{P}]\mathbf{Y}') \\ &= \left\langle [\mathbf{P}\mathbf{Y}\mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}]^H, \mathbf{C}' \right\rangle + \left\langle [\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}\mathbf{P}\mathbf{Y}]^H, \mathbf{B}' \right\rangle \\ &\quad + \frac{1}{2}\left\langle [\mathbf{C}_1^H\mathbf{C}_1 + \mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}\mathbf{P}\mathbf{Y} + \mathbf{Y}\mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}\mathbf{C} + \mathbf{Y}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}]^H, \mathbf{P}' \right\rangle \\ &\quad + \frac{1}{2}\left\langle [\mathbf{B}_1\mathbf{B}_1^H + \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}\mathbf{P} + \mathbf{P}\mathbf{Y}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H + \mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}\mathbf{P}]^H, \mathbf{Y}' \right\rangle \end{aligned} \quad (2.22)$$

Thus, via a slight change of the RHS forcing of the adjoint CALEs (2.12b) and (2.18b),

$$\begin{aligned} (\mathbf{A} - \mathbf{L}\mathbf{C})^H\mathbf{S} + \mathbf{S}(\mathbf{A} - \mathbf{L}\mathbf{C}) &= \mathbf{C}_1^H\mathbf{C}_1 + \mathbf{Y}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y} \\ &\quad + \mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}\mathbf{P}\mathbf{Y} + \mathbf{Y}\mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}, \end{aligned} \quad (2.23)$$

$$\begin{aligned} (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{T} + \mathbf{T}(\mathbf{A} - \mathbf{B}\mathbf{K})^H &= \mathbf{B}_1\mathbf{B}_1^H + \mathbf{P}\mathbf{C}^H\mathbf{V}^{-1}\mathbf{C}\mathbf{P} \\ &\quad + \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H\mathbf{Y}\mathbf{P} + \mathbf{P}\mathbf{Y}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^H, \end{aligned} \quad (2.24)$$

the gradient can be reexpressed as a function of both the sensor and actuator positions

$$\nabla_{q_s^i} J = \text{trace} \left(\overline{\mathbf{P}(\mathbf{Y} + \mathbf{S})\mathbf{L} \frac{d\mathbf{C}}{dq_s^i}} \right), \quad (2.25)$$

$$\nabla_{q_a^i} J = \text{trace} \left(\overline{\mathbf{K}(\mathbf{P} + \mathbf{T})\mathbf{Y} \frac{d\mathbf{B}}{dq_a^i}} \right). \quad (2.26)$$

3. Discrete-Time Analyses

We now present the discrete-time analogs of the derivations given in §2 in the continuous-time case, as there are some subtle differences. Consider a discrete-time linear system described by

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{G}(\mathbf{q}_a)\mathbf{u}_k + \mathbf{w}_k, \quad (3.1a)$$

$$\mathbf{y}_k = \mathbf{H}(\mathbf{q}_s)\mathbf{x}_k + \mathbf{v}_k, \quad (3.1b)$$

where $\mathbf{x}_k \in \mathbb{C}^n$, $\mathbf{u}_k \in \mathbb{C}^\ell$, $\mathbf{y}_k \in \mathbb{C}^m$, $\mathbf{w}_k \in \mathbb{C}^n$, $\mathbf{v}_k \in \mathbb{C}^m$ are the discrete-time equivalents of the corresponding quantities in (2.1). Similarly, \mathbf{q}_a and \mathbf{q}_s parametrize the locations of sensors and actuators at each timestep k . We again make the standard modeling assumptions that \mathbf{w}_k and \mathbf{v}_k are *uncorrelated, zero-mean, white continuous-time random processes* with covariance $\mathbf{W} \geq \mathbf{0}$ and $\mathbf{V} > \mathbf{0}$ respectively. The (bounded) functional dependence of the operators \mathbf{G} and \mathbf{H} on the actuator and sensor positions \mathbf{q}_a and \mathbf{q}_s is again emphasized explicitly above, but is suppressed below for notational clarity.

3.1. Computing a gradient with respect to the sensor positions

Via standard (discrete-time) Kalman filter theory, the best linear unbiased estimate of the system (3.1) is given by a two-step update [cf. (2.2)]

$$\text{time update:} \quad \hat{\mathbf{x}}_{k+1|k} = \mathbf{F}\hat{\mathbf{x}}_{k|k} + \mathbf{G}\mathbf{u}_k, \quad (3.2a)$$

$$\text{measurement update:} \quad \hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{L}_{k+1}(\mathbf{y}_{k+1} - \mathbf{H}\hat{\mathbf{x}}_{k+1|k}), \quad (3.2b)$$

where

$$\mathbf{L}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{H}^H(\mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^H + \mathbf{V})^{-1}. \quad (3.2c)$$

The discrete-time estimation error covariance obeys a similar two-step update known as the Riccati difference equation (RDE) [cf. (2.3)]

$$\text{time update:} \quad \mathbf{P}_{k+1|k} = \mathbf{F}\mathbf{P}_{k|k}\mathbf{F}^H + \mathbf{W}, \quad (3.3a)$$

$$\text{measurement update:} \quad \mathbf{P}_{k+1|k+1} = (\mathbf{I} - \mathbf{L}_{k+1}\mathbf{H})\mathbf{P}_{k+1|k}. \quad (3.3b)$$

Using this standard notation for the discrete-time estimation setting, the notation $\hat{\mathbf{x}}_{k|j}$ denotes the maximum likelihood estimate of \mathbf{x} at time t_k given all measurements up to and including time t_j , and $\mathbf{P}_{k|j}$ denotes the covariance corresponding to this estimate. In particular, $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ are often called the *prior estimate* and *prior covariance*, whereas $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ are often called the *posterior estimate* and *posterior covariance*.

As in §2.1, we now consider the minimization of the trace of the infinite-horizon covariance matrix. Because the discrete-time Kalman filter is characterized as a two-step process, there are two possible choices to make as to whether the covariance should be minimized before or after the measurement update, as shown below. Both formulations are presented; which is more appropriate to use in practice is application dependant.

The *infinite-horizon prior covariance* of (3.3) is computed by substituting (3.3b) into (3.3a) and applying (3.2c), then defining $\mathbf{P}_- = \mathbf{P}_{k+1|k} = \mathbf{P}_{k|k-1}$, thus transforming the

RDE (3.3) into the prior form of the discrete-time algebraic Riccati equation (DARE) [cf. (2.4a)]

$$\mathbf{P}_- = \mathbf{F}\mathbf{P}_-\mathbf{F}^H + \mathbf{W} - \mathbf{F}\mathbf{P}_-\mathbf{H}^H(\mathbf{H}\mathbf{P}_-\mathbf{H}^H + \mathbf{V})^{-1}\mathbf{H}\mathbf{P}_-\mathbf{F}^H \quad (3.4a)$$

Alternatively, the *infinite-horizon posterior covariance* of (3.3) is computed by substituting (3.3a) into (3.3b) and (3.2c) and combining, then defining $\mathbf{P}_+ = \mathbf{P}_{k+1|k+1} = \mathbf{P}_{k|k}$ and applying the Matrix-Inversion Lemma

$$(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} = \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1},$$

thus transforming the RDE (3.3) into the posterior form of the DARE

$$\mathbf{P}_+^{-1} = (\mathbf{F}\mathbf{P}_+\mathbf{F}^H + \mathbf{W})^{-1} + \mathbf{H}^H\mathbf{V}^{-1}\mathbf{H}. \quad (3.4b)$$

The matrices \mathbf{F} , \mathbf{W} , and \mathbf{V} in this LTI formulation are assumed to be given. The remaining matrix which affects \mathbf{P}_- in (3.4a), and \mathbf{P}_+ in (3.4b), is \mathbf{H} , which is a function of the sensor positions \mathbf{q}_s . Thus, an optimization problem may be posed to minimize some measure of \mathbf{P}_- , or \mathbf{P}_+ , with respect to the (stationary) sensor locations \mathbf{q}_s .

3.1.1. Prior covariance optimization

We first seek the optimal \mathbf{q}_s which minimizes the cost [cf. (2.4b)]

$$J_- = \text{trace}(\mathbf{P}_-), \quad (3.5)$$

subject to (3.4a). The associated first-order perturbations are

$$\begin{aligned} J'_- &= \text{trace}(\mathbf{P}'_-) = \langle \mathbf{I}, \mathbf{P}'_- \rangle, \\ \mathbf{P}'_- &= \mathbf{F}(\mathbf{I} - \mathbf{L}\mathbf{H})\mathbf{P}'_-(\mathbf{I} - \mathbf{L}\mathbf{H})^H\mathbf{F}^H \\ &\quad + \mathbf{F}\mathbf{L}(\mathbf{H}'\mathbf{P}_-\mathbf{H}^H + \mathbf{H}\mathbf{P}'_-(\mathbf{H}')^H)\mathbf{L}^H\mathbf{F}^H - \mathbf{F}\mathbf{P}'_-(\mathbf{H}')^H\mathbf{L}^H\mathbf{F}^H - \mathbf{F}\mathbf{L}\mathbf{H}'\mathbf{P}_-\mathbf{F}^H \\ \mathbf{H}' &= \left(\frac{d\mathbf{H}}{dq_s^i} \right)^H \mathbf{q}_s'. \end{aligned}$$

The above relations are derived in a manner analogous to the continuous-time case. Note that the perturbation of $(\mathbf{H}\mathbf{P}_-\mathbf{H}^H + \mathbf{V})^{-1}$ is determined leveraging the identity $(\boldsymbol{\Phi}^{-1})' = -\boldsymbol{\Phi}^{-1}\boldsymbol{\Phi}'\boldsymbol{\Phi}^{-1}$ (see Petersen & Pedersen (2008)), thus leading to

$$((\mathbf{H}\mathbf{P}_-\mathbf{H}^H + \mathbf{V})^{-1})' = (\mathbf{H}\mathbf{P}_-\mathbf{H}^H + \mathbf{V})^{-1}(\mathbf{H}'\mathbf{P}_-\mathbf{H}^H + \mathbf{H}\mathbf{P}'_-(\mathbf{H}')^H + \mathbf{H}\mathbf{P}_-(\mathbf{H}')^H)(\mathbf{H}\mathbf{P}_-\mathbf{H}^H + \mathbf{V})^{-1}.$$

Defining an adjoint matrix \mathbf{S}_- and the inner product (2.9) and performing the necessary rearrangements in a manner analogous to the continuous-time case, one ultimately arrives at the gradient of J_- ,

$$\nabla_{q_s^i} J_- = \text{trace} \left(\overline{2\mathbf{P}_-(\mathbf{I} - \mathbf{L}\mathbf{H})^H\mathbf{F}^H\mathbf{S}_-\mathbf{F}\mathbf{L}\frac{d\mathbf{H}}{dq_s^i}} \right), \quad (3.6a)$$

where \mathbf{S}_- satisfies the associated discrete-time algebraic Lyapunov equation (DALE)

$$(\mathbf{I} - \mathbf{L}\mathbf{H})^H\mathbf{F}^H\mathbf{S}_-\mathbf{F}(\mathbf{I} - \mathbf{L}\mathbf{H}) - \mathbf{S}_- = \mathbf{I}. \quad (3.6b)$$

3.1.2. Posterior covariance optimization

We now seek the optimal \mathbf{q}_s which minimizes the cost [cf. (2.4b)]

$$J_+ = \text{trace}(\mathbf{P}_+), \quad (3.7)$$

subject to (3.4b). The associated first-order perturbations in this case are

$$-\mathbf{P}_+^{-1}\mathbf{P}'_+\mathbf{P}_+^{-1} = -(\mathbf{F}\mathbf{P}_+\mathbf{F}^H + \mathbf{W})^{-1}\mathbf{F}\mathbf{P}'_+\mathbf{F}^H(\mathbf{F}\mathbf{P}_+\mathbf{F}^H + \mathbf{W})^{-1} + (\mathbf{H}^H)' \mathbf{V}^{-1}\mathbf{H} + \mathbf{H}^H\mathbf{V}^{-1}\mathbf{H}'.$$

Using a similar procedure as before, one ultimately arrives at the gradient of J_+ ,

$$\nabla_{q_s^i} J_+ = \text{trace} \left(\overline{2\mathbf{S}_+ \mathbf{H}^H \mathbf{V}^{-1} \frac{d\mathbf{H}}{dq_s^i}} \right), \quad (3.8a)$$

where \mathbf{S}_+ satisfies the DALE

$$\mathbf{P}_+ \mathbf{F}^H (\mathbf{F} \mathbf{P}_+ \mathbf{F}^H + \mathbf{W})^{-1} \mathbf{S}_+ (\mathbf{F} \mathbf{P}_+ \mathbf{F}^H + \mathbf{W})^{-1} \mathbf{F} \mathbf{P}_+ - \mathbf{S}_+ = \mathbf{P}_+^2, \quad (3.8b)$$

The gradients in (3.6a) and (3.8a) are slightly different, because the cost functions they minimize are different. Noting (3.3a), it is evident that

$$J_- = \text{trace}(\mathbf{P}_-) = \text{trace}(\mathbf{F} \mathbf{P}_+ \mathbf{F}^H) + \text{trace}(\mathbf{W}),$$

whereas $J_+ = \text{trace}(\mathbf{P}_+)$. The gradients and optimal solutions of these two formulations thus coincide only if $\mathbf{F}^H \mathbf{F} = \mathbf{I}$.

3.2. Computing a gradient with respect to the actuator positions

In a final analysis analogous to those of the previous sections, we now consider the corresponding discrete-time actuator placement problem. Standard discrete-time optimal control theory applied to the linear system (3.1) establishes that the cost function

$$J = \frac{1}{2} \sum_{k=1}^{N-1} [\mathbf{x}_k^H \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^H \mathbf{R} \mathbf{u}_k] + \frac{1}{2} \mathbf{x}_N^H \mathbf{Q}_N \mathbf{x}_N. \quad (3.9)$$

is minimized by the full-state feedback control policy

$$\mathbf{u}_k = -\mathbf{K}_k \mathbf{x}_k, \quad \mathbf{K}_k = (\mathbf{R} + \mathbf{G}^H \mathbf{Y}_{k+1} \mathbf{G})^{-1} \mathbf{G}^H \mathbf{Y}_{k+1} \mathbf{F}, \quad (3.10)$$

where the matrix $\mathbf{Y}_k \geq 0$ evolves backward in time from the terminal condition $\mathbf{Y}_N = \mathbf{Q}_N$ according to the RDE [cf. (2.3)]

$$\mathbf{Y}_k = \mathbf{F}^H \mathbf{Y}_{k+1} \mathbf{F} - \mathbf{F}^H \mathbf{Y}_{k+1} \mathbf{G} (\mathbf{R} + \mathbf{G}^H \mathbf{Y}_{k+1} \mathbf{G})^{-1} \mathbf{G}^H \mathbf{Y}_{k+1} \mathbf{F}^H + \mathbf{Q}. \quad (3.11)$$

Similar to the continuous-time case, we identify \mathbf{Y}_k as the ‘‘cost-to-go’’ matrix because it can be shown [cf. (2.16)] that

$$J(\kappa) = \frac{1}{2} \sum_{k=\kappa}^{N-1} [\mathbf{x}_k^H \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^H \mathbf{R} \mathbf{u}_k] + \frac{1}{2} \mathbf{x}_N^H \mathbf{Q}_N \mathbf{x}_N \quad (3.12)$$

$$= \frac{1}{2} \mathbf{x}_\kappa^H \mathbf{Y}_\kappa \mathbf{x}_\kappa. \quad (3.13)$$

Given that the RDE (3.11) marches to a finite value at $k = 0$ as $N \rightarrow \infty$ [that is, that the system (3.1) is *stabilizable*], the infinite-horizon solution may be computed directly by setting $\mathbf{Y} = \mathbf{Y}_k = \mathbf{Y}_{k+1}$, thus transforming the RDE (3.11) into a DARE [cf. (2.4a)]

$$\mathbf{Y} = \mathbf{F}^H \mathbf{Y} \mathbf{F} - \mathbf{F}^H \mathbf{Y} \mathbf{G} (\mathbf{R} + \mathbf{G}^H \mathbf{Y} \mathbf{G})^{-1} \mathbf{G}^H \mathbf{Y} \mathbf{F} + \mathbf{Q}. \quad (3.14a)$$

As in §2.2, for $\kappa = 0$ in the infinite-horizon limit $N \rightarrow \infty$, the cost matrix (3.9) is minimized when the eigenvalues of \mathbf{Y} are minimized. With this in mind, the following cost function may be considered [cf. (2.4b)]

$$\min_{\mathbf{q}_a} J = \text{trace}(\mathbf{Y}). \quad (3.14b)$$

The steps for gradient calculation via perturbation analysis are essentially identical to

those presented previously. The resulting expression for the gradient [cf. 2.12a] is

$$\nabla_{q_a^i} J = \text{trace} \left(\overline{2\mathbf{K}\mathbf{T}(\mathbf{F} - \mathbf{G}\mathbf{K})^H \mathbf{Y} \frac{d\mathbf{G}}{dq_a^i}} \right) \quad (3.15a)$$

where the matrix adjoint \mathbf{T} must satisfy the associated DALE [cf. 2.12b]

$$(\mathbf{F} - \mathbf{G}\mathbf{K})\mathbf{T}(\mathbf{F} - \mathbf{G}\mathbf{K})^H - \mathbf{T} = \mathbf{I}. \quad (3.15b)$$

4. Application to the complex Ginzburg-Landau equation

The 1D complex Ginzburg-Landau (CGL) system (Chomaz *et al.* 1987; Roussopoulos & Monkewitz 1996) shares some interesting dynamic features of 3D Navier-Stokes (NS) systems. Notable similarities include transient energy growth (due to non-normality of the system eigenvectors) and extensively-studied stability characteristics (including well-identified thresholds between stability, convective instability, and global instability). For this reason, and its relative computational simplicity, CGL systems are a useful 1D PDE testbed for estimation and control strategies being developed for ultimate application in 3D NS systems.

The linear CGL equation for a flow perturbation variable $\phi(\xi, t)$ may be written

$$\frac{\partial \phi}{\partial t} = \left(-U \frac{\partial}{\partial \xi} + \mu(\xi) + \gamma \frac{\partial^2}{\partial \xi^2} \right) \phi \quad (4.1)$$

where U , $\mu(\xi)$, γ are complex coefficients which parameterize the advection, amplification, and diffusion properties of the flow, respectively, and ξ denotes the streamwise coordinate of the system. This flexible parameterization has been tuned to match a variety of physical phenomena; for example, Roussopoulos & Monkewitz (1996) tuned the parameters to model vortex shedding behind a circular cylinder. A recent review of the CGL model by Bagheri *et al.* (2009) surveys several such studies of this system.

In the results presented below, the parameters were selected to coincide with the convectively unstable case mentioned by Bagheri *et al.* (2009), with $U = 2 + 0.2i$, $\mu(\xi) = 0.38 - 0.01\xi^2/2$, $\gamma = 1 - i$. The resulting variable-coefficient PDE has a locally unstable domain with $\mu(\xi) > 0$ for all $\xi \in (-8.72, 8.72)$.

Bagheri *et al.* (2009) also provide a convenient codebase for discretization and simulation of the CGL equation using a collocation approach based on a Hermite polynomial expansion. Following this approach, the state $\phi(\xi, t)$ in (4.1) is considered as a linear combination of n orthogonal polynomials[†] defined on $\xi \in (-\infty, \infty)$,

$$\phi(\xi, t) \triangleq \sum_{j=1}^n \hat{\phi}_j(t) H_j(\xi) \quad \text{where} \quad H_j(\xi) = (-1)^j e^{\xi^2} \frac{d^j(e^{-\xi^2})}{d\xi^j}.$$

The perturbation variable $\phi(\xi, t)$ may now be discretized on a set of n collocation points ξ_j , for $j = 1, \dots, n$, and assembled as a state vector \mathbf{x} , where the n collocation points are selected as roots of $H_n(\xi)$. With this discretization, the transformation given above, and the relationships between the derivatives of the Hermite polynomials $H_n(\xi)$, it is straightforward to write the discretized system (4.1) in collocation form with the appropriate

[†] Note that the Hermite polynomials are orthogonal on $\xi \in (-\infty, \infty)$ using the weighting function $w(\xi) = e^{-\xi^2}$.

forcing and measurement variables added:

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \bar{\mathbf{B}}\bar{\mathbf{w}}, \quad (4.2)$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{v}, \quad (4.3)$$

where the matrix \mathbf{A} approximates the spatially-varying linear operator on the RHS of (4.1), \mathbf{B} models the effect of the control inputs \mathbf{u} on the system near the actuators located at $\xi = q_a^j$ for $j = 1, \dots, n_a$, and \mathbf{C} models the measurements \mathbf{y} of the system taken from the sensors located at $\xi = q_s^j$ for $j = 1, \dots, n_s$. The random vectors $\bar{\mathbf{w}}$ and \mathbf{v} are independent and normally distributed with covariances $\bar{\mathbf{W}}$ and \mathbf{V} respectively. The matrix $\bar{\mathbf{B}}$ models the effect of the disturbance inputs $\bar{\mathbf{w}}$ applied to the system near $\xi = q_d^j$ for $j = 1, \dots, n_d$; note that it is straightforward to cast this system in the standard continuous-times state-space form given in (2.1) by considering a new disturbance vector \mathbf{w} with covariance $\mathbf{W} = \bar{\mathbf{B}}\bar{\mathbf{W}}\bar{\mathbf{B}}^H$.

In the results presented below, we actuate the system with $n_a = 1$ or 2 actuators (at locations q_a^j that we will optimize), we sense the system with $n_s = 1$ or 2 sensors (at locations q_s^j that we will optimize), and we disrupt the system with $n_d = 1$ disturbance (at $q_d^1 = -11.0$); the corresponding matrices are all chosen to approximate narrow Gaussians in space:

$$\begin{aligned} [\mathbf{B}]_{i,j} &= \exp(-(q_a^j - \xi_i)^2/2\sigma^2), \\ [\mathbf{C}]_{i,j} &= m_i \exp(-(q_s^j - \xi_i)^2/2\sigma^2), \\ [\bar{\mathbf{B}}]_{i,1} &= \exp(-(q_d^1 - \xi_i)^2/2\sigma^2), \end{aligned}$$

where the width of the Gaussians used in the simulations reported below is $\sigma^2 = 1/2$, and where a trapezoidal integration weighting factor m_i is used in the definition of \mathbf{C} ,

$$m_i = \begin{cases} (\xi_2 - \xi_1)/2 & i = 1, \\ (\xi_{i+1} - \xi_{i-1})/2 & 1 < i < n, \\ \xi_n - \xi_{n-1} & i = n, \end{cases}$$

so that the sum of the elements on any row of \mathbf{C} approximates the integral of the corresponding Gaussian, independent of the sensor locations q_s^j . By selecting a parameterization of this sort, the input and output operators represent sensors and actuators of a given sensitivity, it is only their *locations* that change when q_a^j and q_s^j are modified.

Before analyzing the influence of measurements and control on the statistics of the estimation error and the statistics of the disturbed system, it is important to consider first the statistics of the disturbed CGL system itself. Figure 1 thus depicts the modulus of the covariance of the state itself, as given by the solution to the infinite-horizon Lyapunov equation [that is, (2.3) with $\mathbf{L} = 0$], when the CGL system (4.1) is forced with the disturbances \mathbf{w} , but no measurements are used for state estimation. As expected, it is seen in these statistics that disruptions of the state tend to grow as they convect through the locally unstable region of the domain and then decay after that; thus, the peak in these statistics is on the diagonal near $\xi = 8.72$.

4.1. Optimal sensor placement in the estimation problem

Finding the optimal placement of a single sensor $q_s^1 \in (-\infty, \infty)$ is a relatively straightforward task that may be achieved with a simple line search. The problem becomes more interesting when considering the simultaneous placement of two or more sensors with $q_s^i \in (-\infty, \infty)$, as the dimension of the optimization space is increased and thus a

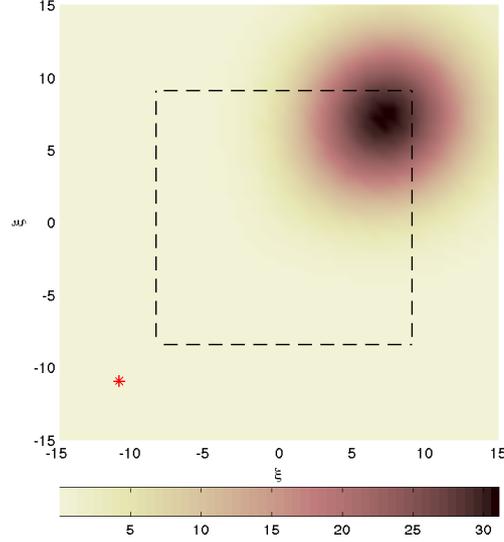


FIGURE 1. The modulus of the covariance of the state itself in the disturbed CGL system [that is, $E\{\mathbf{x}(t)\mathbf{x}^H(t)\}$]; see also Bagheri *et al.* (2009), Figure 16]. Note that this coincides with the covariance $\mathbf{P} = E\{\tilde{\mathbf{x}}(t)\tilde{\mathbf{x}}^H(t)\}$ of the state estimation error when no measurement information is used. The red star indicates the location of the disturbance forcing, $q_w^1 = -11.0$, and the dashed box indicates the region of local instability, $\mu(\xi) > 0$.

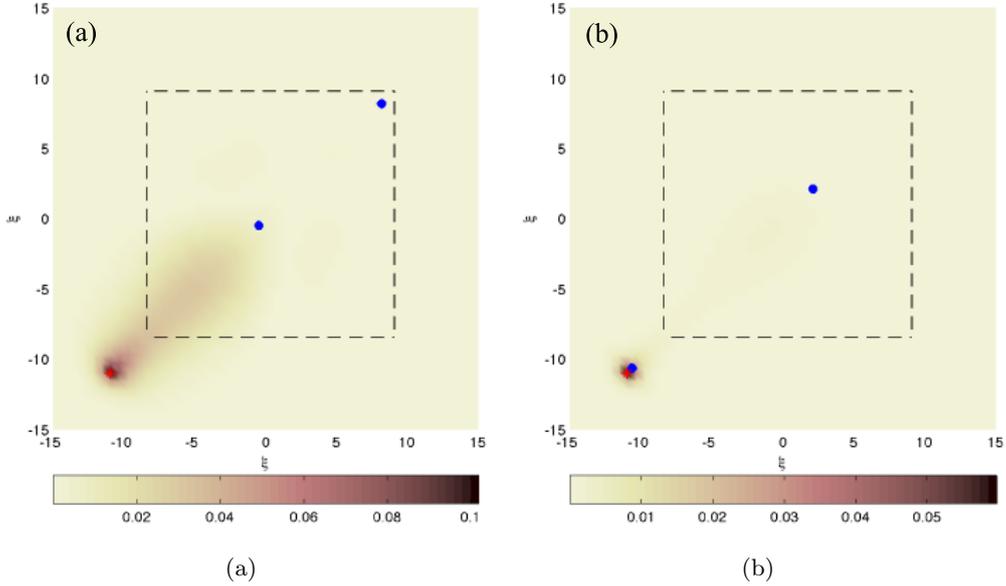


FIGURE 2. The modulus of the covariance of the state estimation error, $\mathbf{P} = E\{\tilde{\mathbf{x}}(t)\tilde{\mathbf{x}}^H(t)\}$, for two different placements of a pair of sensors (blue dots). Figure 2(a) uses a heuristic sequential method of placing the sensors (see text), thereby reducing the covariance depicted in Fig. 1 by nearly 4 orders of magnitude. Figure 2(b) uses a gradient-based method of optimizing the placement of both sensors simultaneously, as described in §2, thereby further reducing the covariance depicted in Fig. 2(a) by another order of magnitude.

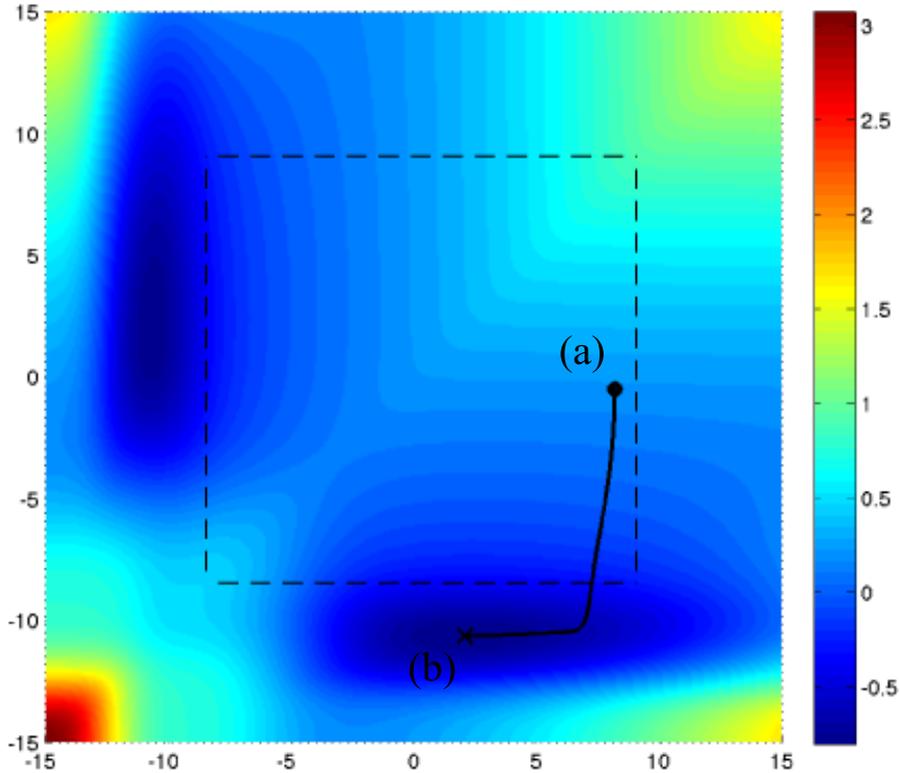


FIGURE 3. A \log_{10} plot of the optimization surface for the two-sensor placement problem in the CGL system with the optimization path superposed, with the axes representing the positions of the two sensors. The optimization was initialized as depicted in Fig. 2(a), and converged to the solution depicted in Fig. 2(b). Since the two sensors are identical, the plot is symmetric across the diagonal.

gradient-based optimization approach is motivated. We thus consider now the optimization of the placement of two sensors in this system.

Figure 2 depicts the modulus of the covariance of the state estimation error for two different configurations of a pair of sensors. The sensor configuration in Fig. 2(a) was chosen heuristically, first placing one sensor at the location of maximum covariance in Figure 1, then placing the other sensor at the location of maximum state estimation error in the estimator that results. The sensor configuration in Fig. 2(b), on the other hand, was optimized using the algorithm described in §2.1. Figure 3 depicts of the full optimization surface as a function of the locations of the two sensors, indicating the path taken during the optimization process from the initial configuration at $\{q_s^1, q_s^2\} = \{12.60, -0.46\}$ [see Fig. 2(a)] to the optimized configuration at $\{q_s^1, q_s^2\} = \{2.10, -10.65\}$ [see Fig. 2(b)]. Note at each step that the path taken is downhill (normal to the isocontours), which is consistent with the fact that a steepest descent method was used in the optimization. Also, Figure 3 is symmetric about the diagonal $q_s^2 = q_s^1$, as the sensors in this case are identical; had sensors of different precision been used (i.e., $\mathbf{V} = \text{diag}\{[v_1, v_2]\}$ with $v_1 \neq v_2$), the symmetry in Figure 3 would be broken, and the gradient-based optimization algorithm would converge to a local minimum.

4.2. Optimal actuator placement in the full information control problem

The two-actuator placement problem is analogous to the two-sensor placement problem discussed in the previous section. Figure 4 depicts the full optimization surface as a function of the configuration of the two actuators, indicating the path taken during the optimization process to minimize the cost (2.4b). The initial configuration of the actuators in this case was taken simply as the optimized sensor configuration found in the previous section. As in the estimation problem, at each step the path taken is downhill (normal to the isocontours). Also, Figure 4 is symmetric about the diagonal $q_a^2 = q_a^1$, as the actuators in this case are identical. In this full-information setting, the optimized solution at $\{q_a^1, q_a^2\} = \{-4.66, 2.36\}$ places both actuators inside the locally unstable region, thus effectively leveraging the positive local amplification term of the CGL; this is in contrast with the optimized sensor configuration presented previously, in which the upstream sensor is actually placed outside the unstable domain, relatively close to where the disturbance is introduced into the system.

Further understanding of this result is given by Fig. 5, which depicts the diagonal of the “cost-to-go” matrix \mathbf{Y} of the controlled CGL system. It is reasonable that the area of the three lobes of the optimized configuration are approximately equal, indicating essentially that contributions to the cost function are, effectively, evenly distributed over the physical domain when the actuator positions are properly optimized.

4.3. \mathcal{H}_2 optimal actuator/sensor placement

Though many control-oriented studies of the CGL system have appeared in the literature, only recently has the question of optimizing sensor and actuator placements in such problems been considered. In particular, Chen & Rowley (2010) found actuator/sensor configurations that minimize the H_2 norm in such problems by simultaneously optimizing both sensor and actuator positions. As intuition suggests, optimizing the sensor and actuators positions separately leads to reasonable but not optimal performance in the full H_2 problem.

As established in §2.3, the gradient-based procedure outlined in §2.1 and §2.2 may easily be extended to optimize sensor and actuator locations simultaneously in the full H_2 setting; results are depicted in Fig. 6. The optimization surface is depicted in Fig. 6(a) for the one-actuator, one-sensor \mathcal{H}_2 problem, and superposed is the path taken by the full optimization algorithm. Note again at each step that the path taken is downhill (normal to the isocontours), thus indicating the correctness of the gradient computation. The initial configuration, $\{q_s, q_a\} = \{-2.47, -1.94\}$, was generated by solving separately the optimal sensor placement problem for one sensor and the (full-information) optimal actuator placement problem for one actuator. The gradient-based method discussed in §2.3 was then used to find the optimized configuration $\{q_s, q_a\} = \{-3.08, -4.66\}$, as depicted in Fig. 6(b).

A similar procedure was performed for the two-actuator, two-sensor case, as depicted in Fig. 6(c), where the initial configuration (based on solving the actuator and sensor placement problems separately) and the optimized configuration $\{q_s^1, q_s^2, q_a^1, q_a^2\} = \{1.09, -10.57, -10.32, 0.49\}$ are compared side-by-side. Though this problem is four-dimensional and thus difficult to visualize, optimizations from the various random initial conditions tested all appear to converge to the same optimized configuration in this case.

The present gradient-based formulation appears to extend naturally to high-dimensional discretizations of various 2D and 3D Navier-Stokes systems, which is left for future studies.

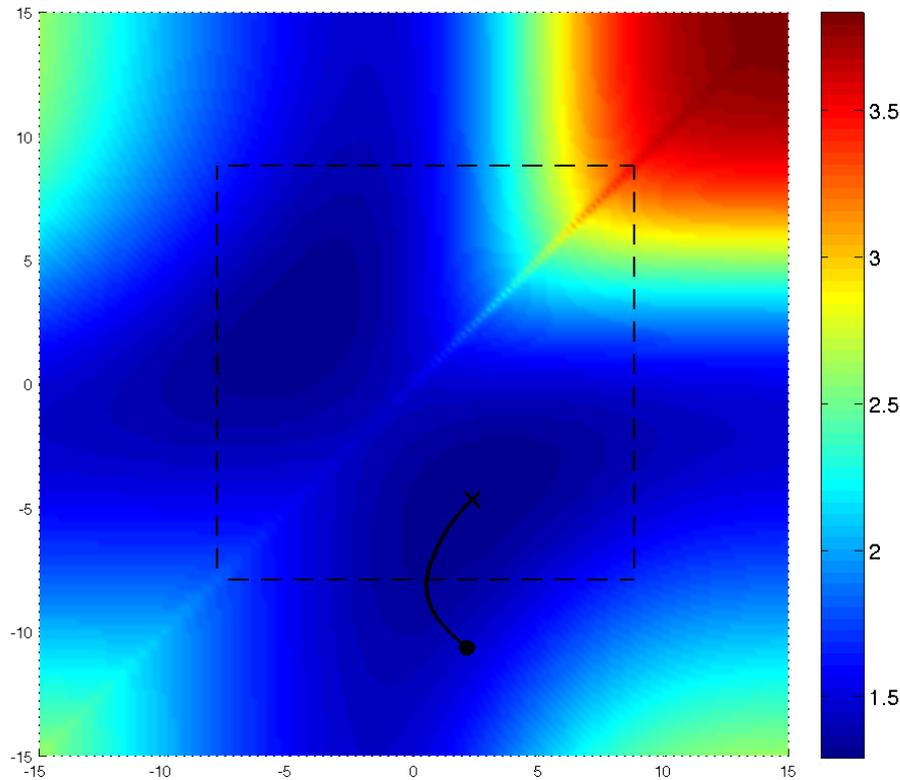


FIGURE 4. A \log_{10} plot of the optimization surface for the full-information two-actuator placement problem in the CGL system with the optimization path superposed, with the axes representing the positions of the two actuators.

REFERENCES

- ALONSO, A.A., KEVREKIDIS, I.G., BANGA, J.R. & FROUZAKIS, C.E. 2004 Optimal sensor location and reduced order observer design for distributed process systems. *Computers & Chemical Engineering* **28** (1-2), 27–35.
- ANDERSON, B.D.O. & MOORE, J.B. 1979 *Optimal filtering*. Prentice-Hall Englewood Cliffs, NJ.
- BAGHERI, S., HENNINGSON, D.S., HÖPFNER, J. & SCHMID, P.J. 2009 Input-output analysis and control design applied to a linear model of spatially developing flows. *Applied Mechanics Reviews* **62**, 020803.
- BULLO, F. & CORTES, J. 2004 Adaptive and distributed coordination algorithms for mobile sensing networks. *Cooperative Control* pp. 431–434.
- CHEN, K. & ROWLEY, C. 2010 Optimal actuator and sensor placement in the linearized complex Ginzburg-Landau system. *Bulletin of the American Physical Society* **55**.
- CHOMAZ, J.M., HUERRE, P. & REDEKOPP, L.G. 1987 Models of hydrodynamic resonances in separated shear flows. In *Proc. 6th Symposium on Turbulent Shear Flows*, pp. 321–326.
- CORTES, J., MARTINEZ, S., KARATAS, T. & BULLO, F. 2004 Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation* **20** (2), 243–255.
- COVER, THOMAS M. & THOMAS, JOY A. 2006 *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience.
- DE FARIAS, DP, DE OLIVEIRA, MC & GEROMEL, JC 2001 Mixed H_2/H_∞ Control of Flexible Structures. *Mathematical Problems in Engineering* **6** (6), 557–598.

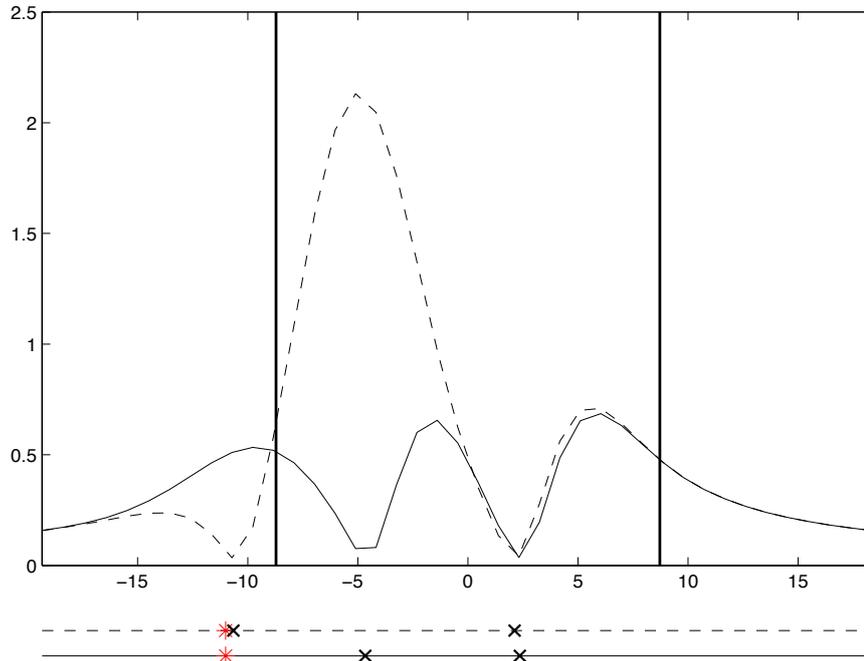
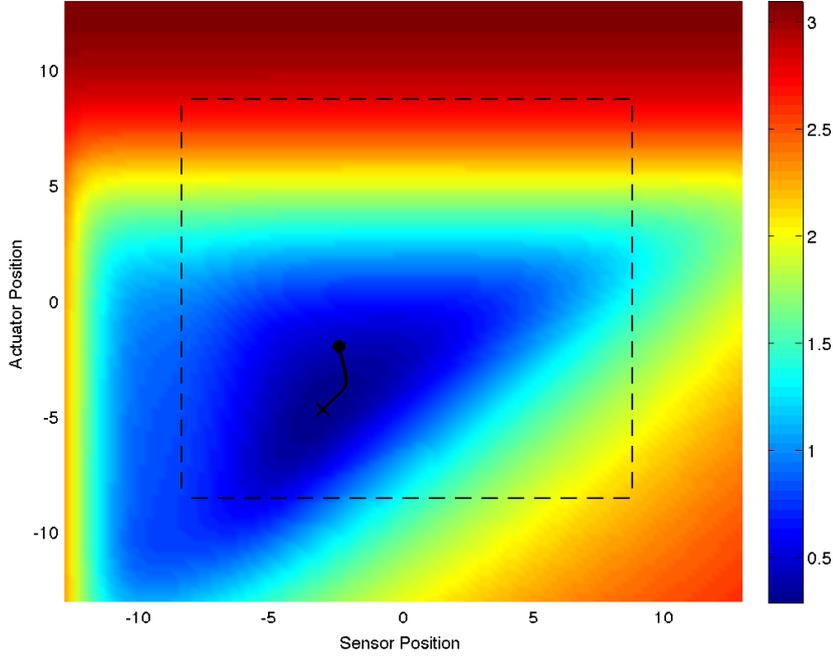


FIGURE 5. The diagonal of \mathbf{Y} (the “cost-to-go” matrix) as a function of streamwise coordinate ξ in the CGL system. The integral of each line is the total cost (the trace of the covariance matrix). The two lines represent solutions for different actuator configurations: the dashed line corresponds to the initial configuration of the optimization, $\{q_a^1, q_a^2\} = \{2.10, -10.65\}$, and the solid line corresponds to the solution of the optimization, $\{q_a^1, q_a^2\} = \{2.36, -4.66\}$. For clarity, below the plot, the disturbance location is denoted by $*$ and the actuator positions are denoted by \times in the initial (dashed) and final (solid) configurations.

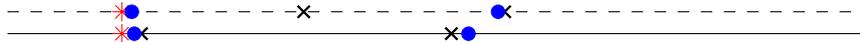
- FAULDS, A.L. & KING, B.B. 2000 Sensor location in feedback control of partial differential equation systems. In *Proceedings of the 2000 IEEE International Conference on Control Applications, 2000*, pp. 536–541. IEEE.
- FLAHERTY, P., JORDAN, M., ARKIN, A. *et al.* 2006 Robust design of biological experiments. *Advances in Neural Information Processing Systems* **18**, 363–370.
- GILES, M.B. 1998 On adjoint equations for error analysis and optimal grid adaptation in CFD. *Frontiers of Computational Fluid Dynamics* pp. 155–169.
- GILES, M.B. & PIERCE, N.A. 2000 An introduction to the adjoint approach to design. *Flow, Turbulence and Combustion* **65** (3), 393–415.
- GOODWIN, G.C. & PAYNE, R.L. 1977 *Dynamic system identification: experiment design and data analysis*. Academic Press.
- HASSIBI, B., SAYED, A.H. & KAILATH, T. 1999 *Indefinite-Quadratic Estimation and Control: A Unified Approach to H_2 and $H[\infty]$ Theories*. Society for Industrial Mathematics.
- HIRAMOTO, K., DOKI, H. & OBINATA, G. 2000 Optimal sensor/actuator placement for active vibration control using explicit solution of algebraic Riccati equation. *Journal of Sound and Vibration* **229** (5), 1057–1075.
- HORN, R.A. & JOHNSON, C.R. 1990 *Matrix Analysis*. Cambridge University Press.
- JAMESON, A., MARTINELLI, L. & PIERCE, N.A. 1998 Optimum aerodynamic design using the Navier–Stokes equations. *Theoretical and Computational Fluid Dynamics* **10** (1), 213–237.
- KAILATH, T. 1980 *Linear systems*. Prentice-Hall Englewood Cliffs, NJ.
- KENNEY, C. & HEWER, G. 1990 The sensitivity of the algebraic and differential Riccati equations. *SIAM Journal on Control and Optimization* **28**, 50.
- KWAKERNAAK, H. & SIVAN, R. 1972 *Linear Optimal Control Systems*. Wiley-Interscience New York.



(a) The optimization surface for the one-sensor, one-actuator \mathcal{H}_2 problem. Also shown is the path taken during the optimization process. The initial condition, \bullet , was generated by performing independent optimizations of the sensor and actuator placements. The combined gradient-based optimization formulation converges from this (suboptimal) initial guess to a (significantly improved) optimal solution, \times , of the combined problem.



(b) Initial condition (dashed) and optimal solution (solid) for the one-sensor, one-actuator \mathcal{H}_2 optimization problem depicted in Fig. 6(a). Note that, through the optimization, the sensor and actuator actually swap their relative positions



(c) As in Fig. 6(b), but for the two-sensor, two-actuator \mathcal{H}_2 problem.

FIGURE 6. Optimized sensor/actuator placements for the combined \mathcal{H}_2 estimation/control problem. The disturbance (indicated with $*$) is located at $q_w = -11.0$ in all instances. Actuators and sensors are denoted with the symbols \times , and \bullet , respectively.

- KWOK, A. & MARTINEZ, S. 2010 Unicycle coverage control via hybrid modeling. *IEEE Transactions on Automatic Control* **55** (2), 528–532.
- LAUGA, ERIC & BEWLEY, THOMAS R 2003 The decay of stabilizability with Reynolds number in a linear model of spatially developing flows. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **459** (2036), 2077–2095.
- LI, FAMING, OLIVEIRA, MAURICIO C. & SKELTON, ROBERT E. 2009 Designing Instrumentation for Control. In *Model-Based Control: Bridging Rigorous Theory and Advanced Technology*, pp. 71–88. Springer US.
- MAJUMDAR, S.J., BISHOP, C.H., ETHERTON, B.J. & TOTH, Z. 2002 Adaptive Sampling with

- the Ensemble Transform Kalman Filter. Part II: Field Program Implementation. *Monthly Weather Review* **130** (5), 1356–1369.
- MARTÍNEZ, S. & BULLO, F. 2006 Optimal sensor placement and motion coordination for target tracking. *Automatica* **42** (4), 661–668.
- PETERSEN, K. B. & PEDERSEN, M. S. 2008 The matrix cookbook. Technical University of Denmark.
- PORAT, B. & NEHORAI, A. 1996 Localizing vapor-emitting sources by moving sensors. *IEEE Transactions on Signal Processing* **44** (4), 1018–1021.
- ROUSSOPOULOS, K. & MONKEWITZ, P.A. 1996 Nonlinear modelling of vortex shedding control in cylinder wakes. *Physica D: Nonlinear Phenomena* **97** (1-3), 264–273.
- TAYLOR, J. 1979 The Cramér-Rao Estimation Error Lower Bound Computation for Deterministic Nonlinear Systems. *IEEE Transactions on Automatic Control* **24** (2), 343–344.
- UCIŃSKI, D. 2005 *Optimal measurement methods for distributed parameter system identification*. CRC.
- ZHANG, D., COLBURN, C.H. & BEWLEY, T.R. 2011 Estimation and Adaptive Observation of Environmental Plumes. In *American Control Conference (ACC), 2011*. IEEE, accepted for publication.
- ZHOU, K. & DOYLE, J.C. 1998 *Essentials of Robust Control*. Prentice Hall New Jersey.