2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Kalman Filtering with Optimal Sensor Motion Planning

Islam I. Hussein
Mechanical Engineering
Worcester Polytechnic Institute
Worcester, MA 01609 USA
ihussein@wpi.edu

Abstract— **In this paper the Kalman filter equations are re-derived for systems with mobile sensors. The motion of the sensor impacts the measurement noise statistical properties through the measurement covariance matrix and, consequently, the estimation error covariance matrix. Moreover, since the sensors are mobile, their motion has an associated control cost. Minimizing the control effort expended during the measurement process is of critical importance in many applications, in particular where fuel and energy resources scarce. With this realization, the Kalman filter equations are re-derived in this paper to include sensor dynamics and their impact on the estimation error. While the classical Kalman filter seeks to minimize a measure of the estimation error, in this paper the goal is to minimize both the estimation error as well as control energy expended during the estimation process. Necessary optimality conditions are derived, and are further simplified for the case where the process to be estimated is composed of a set of decoupled and uncorrelated processes. A mathematical continuation approach is proposed to solve the resulting nonlinear two-point value problem. A numerical example is provided to illustrate the work of this paper.**

I. Introduction

The problem of estimating a noisy signal using a noisy sensor is a classic problem in estimation theory. The optimal solution to the problem of estimating a linear time varying process is the Kalman filter [1]. With the accelerated progress made in sensor and autonomous vehicle technologies that can host these sensors, as well as the emergence of applications that need to employ such technologies, a new feature in the estimation problem arises. This feature is the introduction of a new degree of freedom that needs to be optimized, in particular if it influences the estimation process. This new degree of freedom is the sensor motion itself, which is now seen as a variable that influences the quality of the estimate. Moreover, since sensor motion requires energy expenditure and since energy resources may be scarce or very valuable in many (in particular, long time range) applications, the amount of control energy expended during the estimation process becomes another quantity that we may need to minimize along with the estimation error. Hence, one now has to ask: How should the sensors be best moved in order to maximize the estimation quality and minimize the control energy expended in the maneuver?

The question of mobility in multi- and single sensor estimation problems, though not including optimality, has been addressed by many researchers in recent years. Most of the literature has been focused on extending the Kalman filter to mobile sensor networks from different perspectives, including distributed Kalman filtering *with* consensus filters [2], *for* solving consensus problems [3], [4] and sensor fusion [5], Kalman filtering for coverage control and simultaneous localization and mapping (SLAM) [6]–[8], where we note that not all the cited articles involve mobile sensors. This list of papers is by no means comprehensive, and we refer the reader to these papers and references therein on Kalman filtering for fixed and mobile sensor networks.

The earliest paper on decentralized Kalman filtering in a *combined sensing and control* setting is the paper by Speyer [9]. In this paper a decentralized control problem involving multiple sensor/actuator nodes is formulated, where the goal is to share the information of each sensor, processed with a local Kalman estimator, with all the other nodes so that the controllers can be computed using the best estimate of the state of the system given the information from all the sensors. The controls are determined so that the expected value of a quadratic performance index is minimized. Hence, the paper [9] and the present work share the common theme of obtaining controls to maximize a performance metric, except that in the present work decentralization is not considered and where the control is not used to influence the process to be estimated, but, instead, to control the location of the sensors to maximize a given performance metric.

The question of *optimal* filtering with *mobile* sensors has not been widely addressed in the past. In [10], by allowing sensors to continuously and autonomously correlate the outcomes of their previous observations and use them to plan for next observations, more effective sensory activities can be achieved. The authors propose a mathematical formulation which ties together, the state of uncertainty of the sensor and the parameters that control its sensing activities in an optimal control setting. In [11], the authors consider the nonlinear filtering problem of a diffusion process where several noisy vector observations with possibly different dimensions are available. At each time the problem considered is the optimal selection of a schedule of available sensors to optimally estimate a function of the state at the final time.

In this paper we formulate a cost functional that is a weighted sum of (1) the error covariance matrix at the terminal time (a soft constraint on the terminal error covariance matrix), (2) the control energy expended throughout the maneuver, and (3) the error covariance matrix throughout the maneuver. The maneuver is required to be complete over

978-1-4244-2079-7/08/$25.00 ©2008 AACC.

a predetermined time interval given a set of initial sensor positions and velocities. The sensor vehicle dynamics are assumed linear. Terminal sensor positions and velocities can be kept free or can be set before hand, though in this paper we keep the terminal sensor state free. There are several ways to derive the optimality conditions for the classical Kalman filter. However, the one that is most in line with the optimal control approach of this paper is the paper [12], where the author derives the Kalman filter equations using Pontryagin's maximum principle [13]. We follow the same approach (as opposed to, say, a variational approach) in this paper, where the optimality conditions involve solving a two-point boundary value problem, which is a difficulty when compared to the classical Kalman filter, which involves solving an initial value problem. Given this difficulty, we proposed a method based on mathematical continuation [14] to solve the two-point boundary value problem. This solution approach may not converge to the *global* optimal solution, but will converge to local optima (which may in fact be suboptimal). The present paper derives necessary optimality conditions and the question of sufficient conditions and/or uniqueness of the solutions is the subject of future research.

The paper is organized as follows. In Section II, the classical Kalman filter is reviewed and the optimal control problem of interest in this paper is introduced. In Section III, the necessary optimality conditions are derived and in Section IV we propose a method to compute trajectories that satisfy the optimality conditions. In Section V, the optimality conditions for the simple case where the processes to be estimated are decoupled and uncorrelated is derived in a simple form and numerical simulation is provided. We conclude the paper with current and future research.

II. Problem Formulation

A. The Constraining Equations

Consider a linear process described by the following set of linear time varying equations:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{w}(t),\ \mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t), \quad (1)$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the state of the process, $\mathbf{w}(t) \in \mathbb{R}^r$ is a vector-valued white-noise Gaussian process with zero mean $E[\mathbf{w}(t)] = 0,\ \forall t$ and covariance given by $E[\mathbf{w}(t)\mathbf{w}'(t)] = \delta(t-\tau)\mathbf{Q}(t)$, where $\delta(\cdot)$ is the Dirac delta function and $\mathbf{Q}(t)$ being a symmetric positive definite matrix for all t, and $\mathbf{y}(t) \in \mathbb{R}^m$ is the output to be measured by a set of *moving* sensors. The matrix $\mathbf{A}(t)$ is $n \times n$, $\mathbf{B}(t)$ is $n \times r$, and $\mathbf{C}(t)$ is $m \times n$.

It is assumed that the initial state $\mathbf{x}(t_0)$ is a vector-valued Gaussian random variable independent of the process noise $\mathbf{w}(t)$, with a known mean $E[\mathbf{x}(t_0)] = \bar{\mathbf{x}}_0$ and covariance matrix $E[\mathbf{x}(t_0)\mathbf{x}'(t_0)] = \mathbf{\Sigma}_0$. Hence, $\mathbf{x}(t)$ and, consequently, $\mathbf{y}(t)$ are Gaussian random processes.

The sensor measurement is described by the equation

$$\mathbf{z}(t) = \mathbf{y}(t) + \boldsymbol{\nu}(t) = \mathbf{C}(t)\mathbf{x}(t) + \boldsymbol{\nu}(t), \quad (2)$$

where $\boldsymbol{\nu}(t) \in \mathbb{R}^m$ is a Gaussian white-noise process with zero mean, $E[\boldsymbol{\nu}(t)] = 0$, and covariance given by $E[\boldsymbol{\nu}(t)\boldsymbol{\nu}'(t)] = \delta(t-\tau)\mathbf{R}(t)$, $\mathbf{R}(t)$ being a symmetric positive definite matrix for all t. In this paper, we will assume that $\mathbf{R}(t)$ is in fact a function of the relative distance between the sensors and the process location. We will return to this issue, which is at the crux of this paper, in the next section.

In this paper, the Linear time-varying filter is given by

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\hat{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{z}(t), \quad (3)$$

where $\mathbf{F}(t)$ is a $n \times n$ matrix[1] and $\mathbf{G}(t)$ is a $n \times m$ matrix. The estimation error is then given by

$$\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t) \quad (4)$$

and it satisfies the error dynamics given by

$$\begin{aligned}\dot{\mathbf{e}}(t) &= (\mathbf{A}(t) - \mathbf{F}(t) - \mathbf{G}(t)\mathbf{C}(t))\,\mathbf{x}(t) \\ &\quad + \mathbf{F}(t)\mathbf{e}(t) + \mathbf{B}(t)\mathbf{w}(t) - \mathbf{G}(t)\boldsymbol{\nu}(t), \quad (5)\end{aligned}$$

To guarantee an unbiased estimate $\hat{\mathbf{x}}(t)$, $\mathbf{F}(t)$ and $\mathbf{G}(t)$ are constrained to satisfy [12]

$$\mathbf{F}(t) = \mathbf{A}(t) - \mathbf{G}(t)\mathbf{C}(t). \quad (6)$$

Defining the error covariance matrix $\mathbf{\Sigma}(t) := E[\mathbf{e}(t)\mathbf{e}'(t)]$, one can verify that the error covariance matrix satisfies the equation [12]

$$\begin{aligned}\dot{\mathbf{\Sigma}}(t) &= (\mathbf{A}(t) - \mathbf{G}(t)\mathbf{C}(t))\,\mathbf{\Sigma}(t) \\ &\quad + \mathbf{\Sigma}(t)\,(\mathbf{A}(t) - \mathbf{G}(t)\mathbf{C}(t))' \\ &\quad + \mathbf{B}(t)\mathbf{Q}(t)\mathbf{B}'(t) + \mathbf{G}(t)\mathbf{R}(t)\mathbf{G}'(t),\ \mathbf{\Sigma}(t_0) = \mathbf{\Sigma}_0, \quad (7)\end{aligned}$$

where we have eliminated $\mathbf{F}(t)$ as given in equation(6). Note that in deriving the error covariance matrix equation the fact that $\mathbf{R}$ may be dependent on the sensor positions played no role. *Equation (7) represents the first constraining equation for the optimal control problem developed in the next section.*

We now introduce the dynamics for each mobile sensor. Each sensor is assumed to satisfy dynamics of the form

$$\dot{\mathbf{x}}_{s,i}(t) = \mathbf{A}_{s,i}\mathbf{x}_{s,i} + \mathbf{B}_{s,i}\mathbf{u}_i,\ \mathbf{x}_{s,i}(t_0) = \mathbf{x}^0_{s,i}, \quad (8)$$

$i = 1, \ldots, N$, where N is the number of sensors, $\mathbf{x}_{s,i} \in \mathbb{R}^4$ is the state of sensor i, and $\mathbf{u}_i \in \mathbb{R}^2$ is the control force acting on sensor i. $\mathbf{x}_{s,i} = (\mathbf{p}_i(t), \mathbf{v}_i(t))$ is thus a vector of length[2] 4, where $\mathbf{p}_i(t) \in \mathbb{R}^2$ denotes the position of sensor i and $\mathbf{v}_i(t) \in \mathbb{R}^2$ denotes the velocity of sensor i. Aggregating the sensor dynamics into a single vector $\mathbf{x}_s(t) = (\mathbf{x}_{s,1}(t), \mathbf{x}_{s,2}(t), \ldots, \mathbf{x}_{s,N}) \in \mathbb{R}^{4N}$, the control vectors $\mathbf{u}_i$ into the single vector $\mathbf{u}(t) = (\mathbf{u}_1, \ldots, \mathbf{u}_N(t)) \in \mathbb{R}^{2N}$, the matrices $\mathbf{A}_{s,i}$ into the single matrix $\mathbf{A}_s = \mathrm{diag}\,(\mathbf{A}_{s,1}, \ldots, \mathbf{A}_{s,N})$, and the matrices $\mathbf{B}_{s,i}$ into the single matrix $\mathbf{B}_s = \mathrm{diag}\,(\mathbf{B}_{s,1}, \ldots, \mathbf{B}_{s,N})$, the sensors' dynamics can be written in the compact form

$$\dot{\mathbf{x}}_s(t) = \mathbf{A}_s\mathbf{x}_s(t) + \mathbf{B}_s\mathbf{u}_s(t),\ \mathbf{x}_s(t_0) = \mathbf{x}^0_s. \quad (9)$$

More general dynamics may also be easily considered. Above, $\mathrm{diag}\,(\mathbf{A}, \mathbf{B}, \ldots, \mathbf{Z})$ denotes a block diagonal matrix whose block diagonal components are $\mathbf{A}, \mathbf{B}, \ldots, \mathbf{Z}$. *Equation (9) represents the second constraining equation for the optimal control problem developed in the next section.*

B. The Optimal Control Problem

If the sensors are fixed the only free variable is $\mathbf{G}(t)$. Moreover, if the goal is to minimize the cost

$$J_{\mathrm{kf}} = \mathrm{Tr}\,[\mathbf{M}(t_f)\mathbf{\Sigma}(t_f)]\,, \quad (10)$$

[1]Here we assume that the filter has the same dimension as the process, though it is standard to generalize to higher dimensional filters.

[2]The dimension of $\mathbf{p}_i(t)$ and $\mathbf{v}_i(t)$ may increase to higher dimensions if three-dimensional and rigid body motions, or three-dimensional vehicle models are of interest.

where Tr denotes the trace of a square matrix and $\mathbf{M}(t_f)$ is a positive definite matrix, subject to the covariance matrix equation (7), then the optimal choice for $\mathbf{G}(t)$ is given by

$$\mathbf{G}(t) = \mathbf{\Sigma}(t)\mathbf{C}'(t)\mathbf{R}^{-1}(t). \tag{11}$$

Substituting this and equation (6) into equation (3) one obtains the classical Kalman filter equations [12].

If the sensors are mobile, additional degrees of freedom become available. These are namely the trajectories of the sensors in the system. Since the locations of the sensors influence the quality of the estimated signal, via the observation covariance matrix, one can then ask "what is the optimal trajectory to achieve the maximum estimation quality?" If energy efficiency is not a concern, the answer to this question is simply to move the agents to the "sweet spots" of the sensors, where optimal estimation is achieved based on the sensor characteristics [15], [16]. If we also have the additional requirement that, in addition to minimizing the estimation error, the trajectory minimizes the control energy expended during the maneuver. Control effort efficiency is a desirable property, in particular, when energy is scarce or when battery recharging or re-fueling is not readily available. Including energy efficiency in the cost may have the consequence that sensor sweet spots are not necessarily the optimal sensor locations since getting to these locations may require a prohibitive amount of control energy. Hence, the goal is to find a trajectory that minimizes a weighted sum of estimation error as well as control energy. This discussion motivates the following optimal control problem.

Problem II.1. Given an initial time t_0 and a final time t_f, a set of N mobile sensors that satisfy the equations of motion (9), and the error covariance matrix dynamics (7), find the optimal values for $\mathbf{G}(t)$ and $\mathbf{u}_s(t)$ that minimize

$$J = \mathrm{Tr}\left[\mathbf{M}(t_f)\mathbf{\Sigma}(t_f)\right] + \int_{t_0}^{t_f} \frac{\alpha}{2}\left\|\mathbf{u}_s\right\|^2 + \mathrm{Tr}\left[\mathbf{K}\mathbf{\Sigma}(t)\right] \mathrm{d}t, \tag{12}$$

where $\alpha \geq 0$ is a weighting parameter and $\mathbf{K}$ is a *symmetric positive semi-definite* $n \times n$ weighting matrix. No terminal conditions on the sensor location or velocity are imposed.

Two important remarks are in order. In the case when the sensors are immobile, one may be tempted, in an effort to recover the classical Kalman filter equations, to just set $\alpha = 0$ and $\mathbf{K} = 0$. As will be mathematically seen from the optimality conditions, sensor immobility is in fact reflected in the problem formulation by setting $\alpha = \infty$. By setting $\alpha = \infty$, any nonzero control vector $\mathbf{u}_s$ immediately results in an infinite cost, which is obviously non-optimal since setting $\mathbf{u}_s = 0$ will result in a finite cost. With $\mathbf{u}_s = 0$ and $\mathbf{K} = 0$, the resulting finite cost is exactly that which results from the classical Kalman filter with immobile sensors. In the next section, we derive the necessary optimality conditions, where it should become obvious that $\alpha = \infty$ corresponds to the immobile sensors case. Setting $\alpha = 0$ corresponds to the case where control effort is of no concern and an infinite source of energy is available to us. $\alpha = 0$ results in a singular optimal control problem. In fact, in the next section we will restrict α to being strictly positive. In the immobile sensor case, the only free variable, as in the classical Kalman filter, is just $\mathbf{G}(t)$. In the mobile sensor case, both $\mathbf{G}(t)$ and $\mathbf{u}(t)$ are free variables.

The second remark is in regards to terminal conditions on sensor locations and velocities, which are kept free in problem II.1. One may also impose terminal conditions on the sensor locations and velocities. An important application is a scenario where an autonomous aerial vehicle equipped with a radar detection sensor flies through a way-point at a particular speed. This defines the initial aircraft position and velocity. The goal is to pass through a second way-point at a particular speed. This defines the terminal condition on the sensor position and velocity. The optimal control problem II.1 seeks to find the trajectory that satisfies the sensor boundary conditions, its dynamics, the filter constraining equations, and that optimizes between a weighted sum of the fuel cost and estimation quality. The above optimal control problem and the ensuing derivation of the optimality conditions can easily be modified (in fact, only transversality conditions will be modified below) to deal with terminal sensor position and velocity conditions. The reason we restrict attention to the free terminal sensor condition case is that there is a straightforward mathematical solution procedure to solve the optimal control problem that will be described in Section IV.

III. Necessary Optimality Conditions

As in [12], the optimality conditions are derived using the Pontryagin maximum principle [13]. We first form the pre-Hamiltonian, which is given by

$$\begin{aligned}\hat{H} &= \mathrm{Tr}\left[\dot{\mathbf{\Sigma}}(t)\mathbf{\Lambda}'(t)\right] + \boldsymbol{\lambda}(t)\cdot\dot{\mathbf{x}}_s(t) \\ &\quad -\frac{\alpha}{2}\left\|\mathbf{u}_s(t)\right\|^2 - \mathrm{Tr}\left[\mathbf{K}\mathbf{\Sigma}(t)\right],\end{aligned} \tag{13}$$

where $\mathbf{\Lambda}(t)$ is an $n \times n$ generalized momentum matrix dual to $\mathbf{\Sigma}(t)$ and $\boldsymbol{\lambda}(t) \in \mathbb{R}^n$ is a generalized momentum vector dual to $\mathbf{x}_s(t)$. After substituting equations (7) and (9) into equation (13), the Hamiltonian is obtained by finding the critical value of $\hat{H}$ with respect to $\mathbf{G}(t)$ and $\mathbf{u}_s(t)$, equating to zero to obtain expressions for $\mathbf{G}(t)$ and $\mathbf{u}_s(t)$ in terms of the multipliers $\mathbf{\Lambda}(t)$ and $\boldsymbol{\lambda}(t)$, and finally substituting back into $\hat{H}$. Firstly, the critical value of $\hat{H}$ with respect to the free variables $\mathbf{G}(t)$ and $\mathbf{u}_s(t)$ gives

$$\partial\hat{H}/\partial\mathbf{G}(t) = 0,\ \partial\hat{H}/\partial\mathbf{u}(t) = 0,$$

which result in

$$\begin{aligned}0 &= -\mathbf{\Lambda}(t)\mathbf{\Sigma}(t)\mathbf{C}'(t) - \mathbf{\Lambda}'(t)\mathbf{\Sigma}(t)\mathbf{C}'(t) \\ &\quad + \left(\mathbf{\Lambda}'(t) + \mathbf{\Lambda}(t)\right)\mathbf{G}(t)\mathbf{R}(\mathbf{x}_s(t)) \\ 0 &= -\alpha\mathbf{u}(t) + \mathbf{B}_s'\boldsymbol{\lambda}(t),\end{aligned} \tag{14}$$

where we used the identities [17]

$$\partial\mathrm{Tr}\left[\mathbf{A}\mathbf{X}'\right]/\partial\mathbf{X} = \mathbf{A},\ \partial\mathrm{Tr}\left[\mathbf{A}\mathbf{X}'\mathbf{B}\right]/\partial\mathbf{X} = \mathbf{B}\mathbf{A}. \tag{15}$$

Remark. Note that the partial derivative of a function with respect to a matrix is known as a matrix gradient. The result is again a matrix. Matrix gradients can also be of vector and matrix maps of matrices. The result is a higher order tensor. We will come across such gradients shortly. We refer the reader to [17], [18] for more on derivatives of vector and matrix maps with respect to a matrix. •

The adjoint variables satisfy the following equations:

$$\dot{\mathbf{\Lambda}}(t) = -\partial\hat{H}/\partial\mathbf{\Sigma},\ \dot{\boldsymbol{\lambda}}(t) = -\partial\hat{H}/\partial\mathbf{x}_s,$$

which result in

$$\dot{\mathbf{\Lambda}}(t) = -\left(\mathbf{A}(t) - \mathbf{G}(t)\mathbf{C}(t)\right)' \mathbf{\Lambda}(t) - \mathbf{\Lambda}(t)\left(\mathbf{A}(t) - \mathbf{G}(t)\mathbf{C}(t)\right) - \mathbf{K} \quad (16)$$

$$\dot{\boldsymbol{\lambda}}(t) = -\mathbf{A}_s'\boldsymbol{\lambda}(t) - \left(\mathbf{I}_{4N} \otimes \frac{\partial f(\mathbf{R})}{\partial\left(\mathrm{rs}\left(\mathbf{R}\right)\right)}\right)\left(\frac{\partial\left(\mathrm{rs}\left(\mathbf{R}\right)\right)'}{\partial \mathbf{x}_s}\right),$$

where we used the following identities [17]:

$$\partial \mathrm{Tr}\left[\mathbf{XA}\right]/\partial \mathbf{X} = \mathbf{A}',\ \partial \mathrm{Tr}\left[\mathbf{AXB}'\right]/\partial \mathbf{X} = \mathbf{A}'\mathbf{B}$$
$$\partial \mathrm{Tr}\left[\mathbf{AXBX}'\right]/\partial \mathbf{X} = \mathbf{A}'\mathbf{XB}' + \mathbf{AXB} \quad (17)$$

and the fact that $\mathbf{\Sigma}$ is symmetric. In the above $\otimes$ is the Kronecker product, and $\mathrm{rs}\left(\mathbf{A}\right)$, where $\mathbf{A}$ is $n \times m$, is a row vector obtained by taking the rows of $\mathbf{A}$ and stacking them horizontally to obtain a $1 \times nm$ row matrix. In obtaining the expression for $\dot{\boldsymbol{\lambda}}$, one had to apply a *matrix* chain rule. There are several conventions in the literature, where we adopted the one found in [18]. The function $f(\mathbf{R})$ is given by

$$f(\mathbf{R}) = \mathrm{Tr}\left[\mathbf{G}(t)\mathbf{R}\mathbf{G}'\mathbf{\Lambda}'\right]. \quad (18)$$

Transversality conditions give

$$\mathbf{\Lambda}(t_f) = \frac{\partial}{\mathbf{\Sigma}(t_f)}\mathrm{Tr}\left[\mathbf{M}(t_f)\mathbf{\Sigma}(t_f)\right] = \mathbf{M}(t_f),\ \boldsymbol{\lambda}(t_f) = 0. \quad (19)$$

The right hand side of the second equation is zero since there are no soft or hard terminal conditions on $\mathbf{x}_s$. In the case where terminal soft or hard constraints are imposed on the sensor locations and velocities, the right hand side of the second equation will not be zero.

Note that the equation for $\dot{\mathbf{\Lambda}}$ in (16) is a linear matrix differential equation in $\mathbf{\Lambda}$. Note also the $\mathbf{\Lambda}(t_f)$ from the transversality conditions is symmetric and positive definite since so is $\mathbf{M}(t_f)$. These two facts imply that $\mathbf{\Lambda}(t)$ is *symmetric* and *positive definite*. Using these two properties of $\mathbf{\Lambda}$, one can simplify equations (14) in the form:

$$\mathbf{G}^*(t) = \mathbf{\Sigma}(t)\mathbf{C}'(t)\mathbf{R}^{-1}(\mathbf{x}_s(t)),\ \mathbf{u}^*(t) = \frac{1}{\alpha}\mathbf{B}_s'(t)\boldsymbol{\lambda}(t). \quad (20)$$

Remark. Note that:

1. The Kalman gain $\mathbf{G}(t)$ in the present case is exactly equal to that in the classical Kalman filter [12].
2. If $\alpha = \infty$, the control force is zero, which corresponds to the immobile sensor case. In the case where $\alpha = 0$, we have a singularity that results in infinite control forces, which is not a feasible solution. Hence, from hereon *we restrict α to being a strictly positive real number*. •

Substituting the optimal expressions for $\mathbf{G}$ and $\mathbf{u}_s$ from (20) in the pre-Hamiltonian, one obtains the Hamiltonian function H, which is a function of the states $\mathbf{x}_s(t)$ and $\mathbf{\Sigma}(t)$, and the co-states $\boldsymbol{\lambda}(t)$ and $\mathbf{\Lambda}(t)$. The equations governing the optimal values of $\mathbf{\Lambda}^*$ and $\boldsymbol{\lambda}^*$ are then given by

$$\begin{aligned}
\dot{\mathbf{\Lambda}}^*(t) &= -\frac{\partial H}{\partial \mathbf{\Sigma}} = -\left(\mathbf{A}(t) - \mathbf{G}^*(t)\mathbf{C}(t)\right)'\mathbf{\Lambda}^*(t) \\
&\quad -\mathbf{\Lambda}^*(t)\left(\mathbf{A}(t) - \mathbf{G}^*(t)\mathbf{C}(t)\right) + \mathbf{K} \\
\dot{\boldsymbol{\lambda}}^*(t) &= -\frac{\partial H}{\partial \mathbf{\Sigma}} = -\mathbf{A}_s'\boldsymbol{\lambda}^*(t) \quad (21) \\
&\quad + \left(\mathbf{I}_{4N} \otimes \frac{\partial f_2(\mathbf{R})}{\partial\left(\mathrm{rs}\left(\mathbf{R}\right)\right)}\right)^* \left(\frac{\partial\left(\mathrm{rs}\left(\mathbf{R}\right)\right)'}{\partial \mathbf{x}_s}\right)^* \\
\mathbf{\Lambda}^*(t_f) &= \mathbf{M}(t_f) \\
\boldsymbol{\lambda}^*(t_f) &= 0,
\end{aligned}$$

where $(\cdot)^*$ means that all the variables inside the parentheses are evaluated along the optimal trajectory, and where

$$f_2(\mathbf{R}) = \mathrm{Tr}\left[\mathbf{\Sigma}(t)\mathbf{C}'(t)\mathbf{R}^{-1}(\mathbf{x}_s(t))\mathbf{C}(t)\mathbf{\Sigma}(t)\mathbf{\Lambda}'(t)\right]. \quad (22)$$

Equations (21) are the necessary optimality conditions. We have the following result.

Theorem III.1. The equations (21) along with the equations of motion (7) and (9) evaluated along the optimal trajectory are the necessary optimality conditions for the problem II.1.

IV. A Numerical Procedure for Solving the Necessary Optimality Conditions

Note that the set of optimality conditions is comprised of four first order matrix and vector differential equations (21), (7) and (9) with two point boundary conditions. This is a significant difficulty in the present result. While the Kalman filter results in solving an initial value problem that can be solved in realtime, the optimal control problem presented here requires the solution of a two-point boundary value problem. However, this is a standard difficulty in optimal control approaches and many techniques such as receding horizon control (see for example [19] and references therein) can be used to obtain (suboptimal) solutions for the optimal control problem in realtime.

For off-line determination of the optimal solution, one can use, for example, Matlab®'s bvp4c.m function, the two-point boundary value problem solver [20], [21]. This uses a simple shooting method that requires an initial guess for the time parameterized states: $\mathbf{\Sigma}(t)$, $\mathbf{x}_s(t)$, $\mathbf{\Lambda}(t)$, and $\boldsymbol{\lambda}(t)$. Since an initial guess is hard to obtain for the nonlinear, possibly time-varying optimality conditions, one can make use of the standard Kalman filter solution, and use a continuation method (homotopy) to solve the problem (a general discussion of the method applied to two point boundary value problems can be found in Chapter 7 in [14].) This is done as follows.

Consider the following cost functional:

$$J_\epsilon = \mathrm{Tr}\left[\mathbf{M}(t_f)\mathbf{\Sigma}(t_f)\right] + \frac{1}{\epsilon}\left(\int_{t_0}^{t_f} \frac{\alpha}{2}\left\|\mathbf{u}_s\right\|^2 + \mathrm{Tr}\left[\mathbf{K}\mathbf{\Sigma}(t)\right]\mathrm{d}t\right), \quad (23)$$

where $0 \leq \epsilon \leq 1$. Note that if $\epsilon = 0$, then we are placing an infinite wight on the integral part of the cost and the solution to problem II.1 is the classical Kalman filter solution, which is easily computable since it only involves an initial value problem as explained in Section II. Here is where we highlight the need for having free terminal time sensor locations and velocities. If the terminal sensor vehicle conditions were not free but constrained to be different from the initial conditions, then clearly the Kalman filter is not a solution to the problem with $\epsilon = 0$ since the classical Kalman filter necessarily requires an immobile sensor, which violates terminal sensor vehicle conditions. Hence, the main reason why we keep the terminal vehicle conditions free is to be able to use the classical Kalman filter as a starting solution for the modified optimal control problem when $\epsilon = 0$. If we need to impose terminal conditions on sensor vehicle position and velocity, then an alternate solution method has to be pursued. This is the focus of current research.

For a nonzero value of ϵ, the necessary optimality conditions are identical to those given in the previous section except that the optimal value of $\mathbf{u}_s$ is given by

$$\mathbf{u}_s^\epsilon\Big|^* = \frac{\epsilon}{\alpha}\mathbf{B}_s'(t)\boldsymbol{\lambda}(t). \tag{24}$$

In a continuation method, one uses the solution to the problem with $\epsilon = \epsilon_0 = 0$ (that is, the classical Kalman filter which is easily computable) as the initial guess for $\epsilon_1 = \delta$, where δ is a sufficiently small parameter. Assuming that the problem with $\epsilon = \epsilon_1 = \delta$ has been successfully solved, one then uses this solution as the initial guess for the next step with, say, $\epsilon = \epsilon_2 = 2\delta$. This is repeated until $\epsilon = \epsilon_j = j\delta$ is sufficiently close to $\epsilon = 1$. At this point, we are able to solve the original two point boundary value problem (21), (7) and (9), since $\epsilon = 1$ corresponds to the original problem we seek to solve. In the next section, we will apply the mathematical continuation approach to solve the optimal control problem.

V. Linear Time Invariant Case with Decoupled, Uncorrelated Processes

In this section, we simplify the above equations for the situation where the state to be estimated $\mathbf{x}$ is composed of a set of n decoupled and uncorrelated processes, each satisfying a linear time-invariant process. This case arises naturally in many applications and is also of much interest. In this case, the matrix $\mathbf{A} = \operatorname{diag}(A_1, \ldots, A_n)$ and $\mathbf{B} = \operatorname{diag}(B_1, \ldots, B_n)$ is a diagonal $n \times n$ matrix (here we have set $r = n$ for simplicity). We will also assume that the noise corrupting the process is uncorrelated and, hence, that the covariance matrix $\mathbf{Q}$ is a diagonal matrix of the form: $\mathbf{Q} = \operatorname{diag}(Q_1, \ldots, Q_n)$. The matrix $\mathbf{C}$ is chosen to be the $Nn \times n$ row block matrix composed by concatenating m identity matrices of dimension $n \times n$ row-wise (i.e., each sensor has access to all states to be measured. Hence, $m = Nn$ since the number of outputs is equal to the number of sensors times the number of states.

Assuming that the sensor measurements are uncorrelated (which is a natural assumption to make), the covariance matrix $\mathbf{R}(\mathbf{x}_s)$ is a $Nn \times Nn$ block diagonal matrix of the form $\mathbf{R}(\mathbf{x}_s) = \operatorname{diag}(\mathbf{R}^1(\mathbf{p}_1), \ldots, \mathbf{R}^N(\mathbf{p}_N)$, where each $\mathbf{R}^i(\mathbf{p}_i)$ $(i = 1, \ldots, N)$ is a $n \times n$ diagonal matrix: $\mathbf{R}^i(\mathbf{p}_i) = \operatorname{diag}\left(R_1^i(\mathbf{p}_i), \ldots, R_n^i(\mathbf{p}_i)\right)$, and where we recall that $\mathbf{p}_i$ is the position of sensor i. In this example, we will assume that each $R_j^i(\mathbf{p}_i)$ is a function of the distance between between the sensor i and the location of the j^{th} state of the process to be estimated. Here we assume that the state $\mathbf{x}$ has a fixed location and, hence, the measurement covariance matrix is a function of the position of the sensor location alone as is the case with the earlier developments in this paper. The case where the process to be estimated is the position of, say, a target (as in target tracking), the problem is a bit more complex since the covariance matrix $\mathbf{R}$ itself is an estimate as opposed to being a deterministic quantity that is based on some sensor model. This situation is the subject of future work.

Based on the above the covariance matrix $\boldsymbol{\Sigma}$ is a diagonal matrix of the form $\boldsymbol{\Sigma} = \operatorname{diag}(\Sigma_1, \ldots, \Sigma_n)$. As described in Section II, the state of the sensors $\mathbf{x}_s = ((\mathbf{p}_1, \mathbf{v}_1), \cdots (\mathbf{p}_N, \mathbf{v}_N))$. We will assume that all sensor vehicles are identical. The matrix $\mathbf{A}_s$ is a block diagonal matrix of the form $\mathbf{A}_s = \operatorname{diag}(\mathbf{A}_{s,1}, \ldots, \mathbf{A}_{s,N})$ where

$$\mathbf{A}_{s,i} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

for all $i = 1, \ldots, N$. The matrix $\mathbf{B}_s$ is a block diagonal matrix of the form $\mathbf{B}_s = \operatorname{diag}(\mathbf{B}_{s,1}, \ldots, \mathbf{B}_{s,N})$, where

$$\mathbf{B}_{s,i} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}'.$$

The covector $\boldsymbol{\Lambda}$ will then be diagonal of the form $\boldsymbol{\Lambda} = \operatorname{diag}(\Lambda_1, \ldots, \Lambda_n)$. The covector $\boldsymbol{\lambda}$ is a vector of length $4N$. We will partition $\boldsymbol{\lambda}$ in the following manner. Let $\boldsymbol{\lambda}_i^1$ denote the covector associated with the position vector of sensor i and $\boldsymbol{\lambda}_i^2$ denote the covector associated with the velocity vector of sensor i. Hence, we write $\boldsymbol{\lambda} = \left((\boldsymbol{\lambda}_1^1, \boldsymbol{\lambda}_1^2), (\boldsymbol{\lambda}_2^1, \boldsymbol{\lambda}_2^2), \ldots, (\boldsymbol{\lambda}_N^1, \boldsymbol{\lambda}_N^2)\right)$. Finally, we will assume that both $\mathbf{M}(t_f)$ and $\mathbf{K}$ are diagonal $n \times n$ matrices, where $\mathbf{M} = \operatorname{diag}(M_1, \ldots, M_n)$ and $\mathbf{K} = \operatorname{diag}(K_1, \ldots, K_n)$. One can show that

$$\dot{\Sigma}_i == 2\Sigma_i a_i - \Sigma_i^2 \sum_{j=1}^{N} \left(r_i^j\right)^{-1} + Q_i. \tag{25}$$

The sensor vehicle equations of motion are given by

$$\dot{\mathbf{p}}_i = \mathbf{v}_i,\ \dot{\mathbf{v}}_i = \frac{1}{\alpha}\boldsymbol{\lambda}_i^2,\ i = 1, \ldots, N. \tag{26}$$

One can verify that $\boldsymbol{\Lambda}$ is diagonal matrix, where the i^{th} diagonal of $\boldsymbol{\Lambda}$ is given by

$$\dot{\Lambda}_i = -2\Lambda_i \left(a_i - \Sigma_i \sum_{j=1}^{N} \left(r_i^j\right)^{-1}\right) + K_i. \tag{27}$$

Our final goal is to obtain an expression for $\dot{\boldsymbol{\lambda}}$. One can show that

$$\dot{\boldsymbol{\lambda}}_i^1 = -\sum_{j=1}^{n} \Lambda_j \Sigma_j^2 \frac{1}{\left(r_j^i\right)^2} \frac{\partial r_j^i}{\partial \mathbf{p}_i},\ \dot{\boldsymbol{\lambda}}_i^2 = -\boldsymbol{\lambda}_i^1. \tag{28}$$

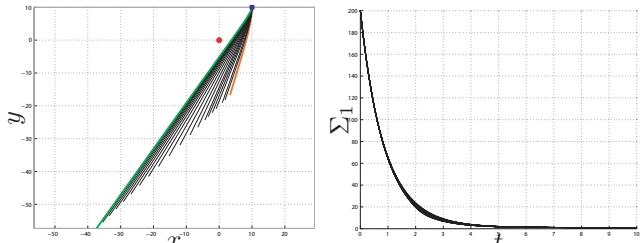
The above equations can then be solved using a numerical procedure as described in Section IV.

In the following simulation, we assume that the measurement noise covariance matrix is quadratic in the distance between the location of the process to be estimated and the sensor location. Hence, each r_j^i is of the form:

$$r_j^i = \beta^i \left\|\mathbf{p}_i - \mathbf{s}_j\right\|^2 + \gamma^i,$$

where $\mathbf{s}_j$ is the position vector that describes the location of process j, and β^i and γ^i are positive scalar parameters that characterize the sensor performance. More realistic models can be used, though we will use the simplified model given above.

For the sake of simplicity, consider a one dimensional process ($n = 1$) and a single sensor ($N = 1$) with $a_1 = -0.5$, $B_1 = 1$, $Q_1 = 1$, $\mathbf{s}_1 = (0, 0)$ (red circle in Figure 1), and r_1^1 having the above form with $\beta^1 = 5$ and $\gamma^1 = 3$. The initial position for the sensor is given by $\mathbf{p}_1(0) = (10, 10)$ and initial velocity $\mathbf{v}_1(0) = (0, -2)$. The initial value for $\boldsymbol{\Sigma} = \Sigma_1$ was set to be $\Sigma_1(0) = 200$. The control weighting parameter is set to $\alpha = 0.004$, the terminal covariance weight was set to $M_1 = 20$ and the weighting parameter $K_1 = 0$. The final time t_f was set to 10.



(a) Motion in the plane (b) $\Sigma_1(t)$

Fig. 1. (a) Motion in the plane withe the various trajectories correspond to a varying value of ϵ from $\epsilon = 0.05$ (red trajectory) to $\epsilon = 1$ (green trajectory), and (b) error covariance variable $\Sigma_1(t)$ for various values of ϵ.

The results are shown in Figure 1, where we have applied the mathematical continuation approach of Section IV. As can be seen as ϵ is varied, more weight is placed on the integral component of the cost function. Hence, for $\epsilon = 0$ we get the classical Kalman filter solution with a fixed sensor location (blue square) that corresponds to its original location. As ϵ is increased (in increments of $\delta = 0.05$) the trajectory is adjust until the solution converges to a trajectory that satisfies the original optimal control problem with $\epsilon = 1$. In the figure, one can also see the variation in Σ_1 as epsilon is incremented from 0 to 1. Due to the scale of the variable Σ_1 it is hard to distinguish between the final solution with $\epsilon = 1$ and other solutions. However, a close-up view of the figure will show that the final solution is not necessarily the one with minimum covariance value at each point in time since the final trajectory may compromise signal quality at the cost of improved control energy efficiency. In a future archival publication, we will also consider the more interesting case with more than a single sensor and a single process, including the possibility that the signals be coupled and correlated. More realistic sensor models will also be considered.

VI. Current and Future Work

Currently, the author is seeking to determine whether the above necessary conditions are sufficient or not. Under some conditions (specifically, convexity in terms of the state and control variables) on the process and sensor dynamics and cost functional, the necessary conditions are also sufficient (see, for example, the Corollary on page 214 and Theorem 10 on page 216 in [22]). Another crucial question is the case where $\mathbf{R}$ is a function of both the state to be estimated as well as the position and/or velocity of the sensor vehicle dynamics. In this case $\mathbf{R}$ is itself a function of estimated variables which renders the problem more complicated. This situation is of particular interest in optimal target tracking problems. Other questions of interest include decentralizing the above result, where information sharing and estimation is distributed to reduce communication costs at the expense of estimation quality.

Acknowledgements

The author gratefully acknowledges financial support from the WPI Faculty Development Grant.

References

[1] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Transactions ASME, Journal of Basic Engineering, Series D*, vol. 83, pp. 95–108, March 1961.

[2] R. Olfati-Saber, "Distributed kalman filter with embedded consensus filters," in *Proceedings of the IEEE Conference on Decision and Control*, Seville, Spain, 2005.

[3] W. Ren, R. W. Beard, and D. B. Kingston, "Multi-agent kalman consensus with relative uncertainty," in *Proceedings of the American Control Conference*, Portland, OR, 2005, pp. 1865–1870.

[4] M. Alighanbari and J. How, "An unbiased Kalman consensus algorithm," in *Proceedings of the American Control Conference*, Minneapolis, MN, 2006, pp. 3519–3524.

[5] R. Olfati-Saber, "Distributed kalman filtering and sensor fusion in sensor networks," *2005 Workshop on Network Embedded Sensing and Control*, October 2005.

[6] I. I. Hussein, "A Kalman filter-based control strategy for dynamic coverage control," in *Proceedings of the American Control Conference*, New York, NY, 2007, pp. 3271–3276.

[7] H. F. Durrant-Whyte, M. W. M. G. Dissanayake, and P. Gibbons, *Robotics Research: The Ninth International Symposium*. Springer Verlag, 2000, ch. Towards simultaneous localization and map building (SLAM) in large unstructured environments, pp. 162–168.

[8] B. Grocholsky, "Information-theoretic control of multiple sensor platforms," Ph.D. dissertation, The University of Sydney, 2002.

[9] J. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-gaussian control problem," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 266–269, 1979.

[10] O. A. Basir and H. C. Shen, "Vision-based intelligent vehicles: State of the art and perspectives," *IEEE International Conference on Systems, Man and Cybernetics*, vol. 2, pp. 999–1004, October 1995.

[11] J. S. Baras and A. Bensoussan, "Optimal sensor scheduling in nonlinear filtering of diffusion processes," *SIAM Journal of Control and Optimization*, vol. 27, pp. 786–813, July 1989.

[12] M. Athans and E. Tse, "A direct derivation of the optimal linear filter using the maximum principle," *IEEE Transactions on Automatic Control*, vol. AG-12, pp. 690–698, 1967.

[13] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*. New York, NY: The MacMillan Company, 1964, translated by D. E. Brown.

[14] V. I. Shalashilin and E. B. Kuznetsov, *Parametric Continuation and Optimal Parametrization in Applied Mathematics and Mechanics*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003.

[15] T. H. Chung, V. Gupta, J. W. Burdick, and R. M. Murray, "On a decentralized active sensing strategy using mobile sensor platforms in a network," in *Proceedings of the IEEE Conference on Decision and Control*, Paradise Island, Bahamas, 2004, pp. 1914–1919.

[16] Y. Mostofi, T. Chung, R. Murray, and J. Burdick, "Communication and sensing trade offs in decentralized mobile sensor networks: A cross-layer design approach," in *International Conference on Information Processing in Sensor Networks*, April 2005. [Online]. Available: http://www.cds.caltech.edu/∼yasi/papers/IPSN.pdf

[17] M. Athans and F. C. Schweppe, "Gradient matrices and matrix calculations," MIT Lincoln Laboratories, Technical Note 1965-53, November 1965.

[18] W. J. Vetter, "Derivative operations on matrices," *IEEE Transactions on Automatic Control*, pp. 241–244, April 1970.

[19] D. Q. Mayne and H. Michalska, "Receding horizon control of nonlinear systems," *IEEE Transactions on Automatic Control*, no. 7, pp. 814–824, July 1990.

[20] MATLAB®, Web, April 2006. [Online]. Available: http://www.mathworks.com/products/matlab/

[21] L. F. Shampine, M. Reichelt, and J. Kierzenka, "Solving boundary value problems for ordinary differential equations in matlab with bvp4c." [Online]. Available: http://www.mathworks.com/support/solutions/files/s8314/bvp-paper.pdf

[22] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*. New York, NY: John Wiley & Sons, Inc., 1967.