2015

A Study on False Information Injection Attack on Dynamic State Estimation in Multi-Sensor Systems

Jingyang Lu
luj2@vcu.edu

A Study on False Information Injection Attack on Dynamic State Estimation in Multi-Sensor Systems

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science at Virginia Commonwealth University

by

Jingyang Lu

Adviser: Ruixin Niu
Department of Electrical and Computer Engineering

Virginia Commonwealth University
Richmond, Virginia
May 8, 2015

CONTENTS

Abstract

A STUDY ON FALSE INFORMATION INJECTION ATTACK ON DYNAMIC STATE ESTIMATION IN MULTI-SENSOR SYSTEMS

By Jingyang Lu, Master

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2015.

Major Director: Ruixin Niu

In this thesis, the impact of false information injection is investigated for linear dynamic systems with multiple sensors. It is assumed that the Kalman filter system is unaware of the existence of false information and the adversary is trying to maximize the negative effect of the false information on the Kalman filter's estimation performance. First, a brief introduction to the Kalman filter is shown in the thesis. We mathematically characterize the false information attack under different conditions. For the adversary, many closed-form results for the optimal attack strategies that maximize the Kalman filter's estimation error are theoretically derived. It is shown that by choosing the optimal correlation coefficients among the bias noises and allocating power optimally among sensors, the adversary could significantly increase the Kalman filter's estimation errors. To be concrete, a target tracking system is used as an example in the thesis. From the adversary's point of view, the best attack strategies are obtained under different scenarios, including a single-sensor system with both position and velocity measurements, and a multi-sensor system with position and velocity measurements. Under a constraint on the total

power of the injected bias noises, the optimal solutions are solved from two perspectives: trace and determinant of the mean squared error matrix. Numerical results are also provided in order to illustrate the negative effect which the proposed attack strategies could inflict on the Kalman filter.

II. Introduction

System state estimation in the presence of an adversary that injects false information into sensor readings is an important problem with wide application areas, such as target tracking with compromised sensors, secure monitoring of dynamic electric power systems, and radar detection and tracking in the presence of jammers. This topic has attracted considerable attention and interest recently [1]–[9]. In [1], the problem of how to take advantage of the power system configuration to introduce arbitrary bias to the system was investigated. In [2], the authors showed the impact of malicious attacks on real-time electricity market and how the attackers can make profit by manipulating certain values of the measurements. They also provided certain strategies to find the optimal single attack vector. The relationship between the attackers and the control center was discussed in [3], where both the adversary's attack strategies and the control center's attack detection algorithms have been proposed. False data attacks on the electricity market have also been investigated in [4] and [5]. In [6], the data frame attack was formulated as a quadratically constrained quadratic program (QCQP). The data frame attack aiming to mislead the power system control center was studied in [7], where it was shown that the system could be made unobservable by controlling only half of a critical set of measurements. Subspace method was presented in [8] showing how to learn the system operating subspace from the measurement and launch the attack accordingly either by hiding the false information in the subspace or misleading the system to remove the data not being attacked. In [9], the relation between a target and a MIMO radar was characterized as a two-person zero-sum game. However, in the aforementioned publications, only the problem of *static* system state estimation has been considered.

In this thesis, for a linear *dynamic* system, we analyze the impact of the injected false information on the Kalman filter's state estimation performance over time, which has not got much attention in the literature. Some related publications exist on sensor management [10]–[13], where the authors showed how to manage the sensors to minimize the mean squared estimation error or its lower bound so that a more accurate state estimate can be obtained. This problem is clearly opposite to the problem we study in the thesis, where the goal for the adversary is to maximize the mean squared state estimation error, and to confuse the Kalman filter. In [14], the problem of sensor bias estimation and compensation for target tracking has been addressed.

Interested readers are referred to [14] and the references therein for details. In [15], impact of the injected biases on a Kalman filter's estimation performance has been studied showing that if the false information is injected at a single time, its impact converges to zero as time goes on; if the false information is injected into the system continuously, the estimation error tends to reach a steady state. In this thesis, we derive the best strategies for the adversary to attack the Kalman filter system from the perspective of the trace of the mean squared error (MSE) matrix, and obtain some closed-form results. We also derive the optimal attack strategy for the adversary, which maximizes the impact of the false information from the determinant perspective. By adopting the objective function as the determinant of the MSE matrix, we change the problem significantly. As shown later in the thesis, the optimal attack strategy that maximizes the determinant of the MSE matrix is a function of the Kalman filter's state estimation covariance and hence "adaptive" to the Kalman filter; whereas that maximizing the trace of the MSE matrix is not a function of the Kalman filter's state estimation covariance.

The rest of thesis is organized as follows. Chapter III gives a brief introduction to the Kalman filter system. The false information attack problem in a general discrete-time linear dynamic system is formulated in the Section IV. Chapter V mathematically characterizes the impact of deterministic or random false information on the Kalman filter's system. Chapter VI and VII analyze how to get the best strategy to attack the Kalman filter's system by maximizing the trace and determinant of the MSE matrix from the perspective of the adversary. Under the constraint on the adversary's total sensor bias noise power, different strategies are derived to maximize the Kalman filter's mean squared state estimation error for different scenarios. Chapter VIII provides the simulation results and Chapter IX concludes the thesis.

III. KALMAN FILTER SYSTEM

A. Linear Dynamic State Estimation

The discrete-time linear dynamic system [16] can be described as below,

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{u}_k + \mathbf{v}_k \tag{1}$$

where $\mathbf{F}_k$ is the system state transition matrix, $\mathbf{x}_k$ is the system state vector at time k, $\mathbf{u}_k$ is a known input vector, $\mathbf{G}_k$ is the input gain matrix, and $\mathbf{v}_k$ is a zero-mean white Gaussian process noise with covariance matrix $E[\mathbf{v}_k \mathbf{v}_k^T] = \mathbf{Q}_k$. The measurement equation is

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k \tag{2}$$

where $\mathbf{w}_k$ is the sequence of zero-mean white Gaussian measurement noise, and

$$E[\mathbf{w}_k \mathbf{w}_k^T] = \mathbf{R}_k \tag{3}$$

The matrices $\mathbf{F}_k$, $\mathbf{G}_k$, $\mathbf{H}_k$, $\mathbf{Q}_k$, and $\mathbf{R}_k$ are assumed to be known with proper dimensions and possibly time varying. The initial state $\mathbf{x}_0$ in general is unknown and modeled as Gaussian distributed with known mean and covariance. The two noise sequences and the initial state are mutually independent. Sometimes, $\mathbf{v}_k$ is taken as $\mathbf{\Gamma}_k \mathbf{v}_k$ with $\mathbf{v}_k$ being an n_v-dimensional vector and $\mathbf{\Gamma}_k$ a known $n_x \times n_v$ matrix. Then the covariance matrix of the noise in the state equation can be written as

$$E\left[\left(\mathbf{\Gamma}_k \mathbf{v}_k\right)\left(\mathbf{\Gamma}_k \mathbf{v}_k\right)^T\right] = \mathbf{\Gamma}_k \mathbf{Q}_k \mathbf{\Gamma}_k^T \tag{4}$$

The linearity of (1) and (2) ensures the preservation of the Gaussian property of the state and measurements. The estimate of the system state $\mathbf{x}_i$ based on the observations up to time k can be written as,

$$\hat{\mathbf{x}}_{i|k} = E\left[\mathbf{x}_i | \mathbf{Z}^k\right] \tag{5}$$

where

$$\mathbf{Z}^k = \{\mathbf{z}_i : i \leq k\} \tag{6}$$

If $i = k$, the conditional mean is called the estimate of the system; if $i < k$, the conditional mean is called the smoothed value of the state; if $i > k$, the conditional mean is called predicted value of the state. The estimation error is defined as

$$\tilde{\mathbf{x}}_{i|k} = \mathbf{x}_i - \hat{\mathbf{x}}_{i|k} \tag{7}$$

The conditional covariance matrix of $\mathbf{x}_i$ given the data $\mathbf{Z}^k$ or the covariance associated with the estimate is

$$\mathbf{P}_{i|k} = E\left[\left(\mathbf{x}_i - \hat{\mathbf{x}}_{i|k}\right)\left(\mathbf{x}_i - \hat{\mathbf{x}}_{i|k}\right)^T |\mathbf{Z}^k\right] \tag{8}$$

B. The Recursive Estimation Algorithm

In terms of observation $\mathbf{z}$ according to the minimum mean squared error (MMSE) criterion, the estimate of $\mathbf{x}$ with prior information $\mathbf{x} \sim N(\bar{\mathbf{x}}, \mathbf{P}_{xx})$ is

$$\hat{\mathbf{x}} = E\left[\mathbf{x}|\mathbf{z}\right] = \bar{\mathbf{x}} + \mathbf{P}_{xz}\mathbf{P}_{zz}^{-1}(\mathbf{z} - \bar{\mathbf{z}}) \tag{9}$$

and the corresponding mean squared error (MSE) is

$$\mathbf{P}_{xx|z} = E\left[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T\right] = \mathbf{P}_{xx} - \mathbf{P}_{xz}\mathbf{P}_{zz}^{-1}\mathbf{P}_{zx} \tag{10}$$

Given the initial estimate $\hat{\mathbf{x}}_{0|0}$ of $\mathbf{x}_0$ and the associated initial covariance $\mathbf{P}_{0|0}$, the cycle of the dynamic estimation will consider mapping the estimate

$$\hat{\mathbf{x}}_{k|k} = E\left[\mathbf{x}_k|\mathbf{Z}^k\right] \tag{11}$$

which is the conditional mean of the state at the time k, and the covariance matrix

$$\mathbf{P}_{k|k} = E\left[[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}][\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]^T|\mathbf{Z}^k\right] \tag{12}$$

into the corresponding variables at the next stage, that is to say, $\hat{\mathbf{x}}_{k+1|k+1}$ and $\mathbf{P}_{k+1|k+1}$. Since the process noise is white and Gaussian, the predicted state $\hat{\mathbf{x}}_{k+1|k}$ is

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k} &= E\left[\mathbf{x}_{k+1}|\mathbf{Z}^k\right] = E\left[\mathbf{F}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{u}_k + \mathbf{v}_k|\mathbf{Z}^k\right] \\ &= \mathbf{F}_k\hat{\mathbf{x}}_{k|k} + \mathbf{G}_k\mathbf{u}_k \end{aligned} \tag{13}$$

The state prediction error, namely the difference between the system state and state prediction is

$$\tilde{\mathbf{x}}_{k+1|k} = \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k\tilde{\mathbf{x}}_{k|k} + \mathbf{v}_k \tag{14}$$

Using the equation above, we can get the state prediction covariance as

$$\begin{aligned} \mathbf{P}_{k+1|k} &= E\left[\tilde{\mathbf{x}}_{k+1|k}\tilde{\mathbf{x}}_{k+1|k}^T|\mathbf{Z}^k\right] \\ &= \mathbf{F}_k E\left[\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T|\mathbf{Z}^k\right]\mathbf{F}_k^T + E\left[\mathbf{v}_k\mathbf{v}_k^T\right] \\ &= \mathbf{F}_k\mathbf{P}_{k|k}\mathbf{F}_k^T + \mathbf{Q}_k \end{aligned} \tag{15}$$

The predicted measurement is the expectation of the measurement conditioned on $\mathbf{Z}^k$,

$$\begin{aligned} \mathbf{z}_{k+1|k} &= E\left[\mathbf{z}_{k+1}|\mathbf{Z}^k\right] \\ &= E\left[\mathbf{H}_{k+1}\mathbf{x}_{k+1} + \mathbf{w}_{k+1}|\mathbf{Z}^k\right] \\ &= \mathbf{H}_{k+1}\hat{\mathbf{x}}_{k+1|k} \end{aligned} \tag{16}$$

The measurement prediction error is

$$\tilde{\mathbf{z}}_{k+1|k} = \mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k} = \mathbf{H}_{k+1}\tilde{\mathbf{x}}_{k+1|k} + \mathbf{w}_{k+1} \tag{17}$$

Thus the measurement prediction covariance, which is defined as $\mathbf{S}_{k+1}$, is

$$\mathbf{S}_{k+1} = \mathbf{H}_{k+1}\mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T + \mathbf{R}_{k+1} \tag{18}$$

The covariance between the state and measurement is

$$\begin{aligned} E\left[\tilde{\mathbf{x}}_{k+1|k}\tilde{\mathbf{z}}_{k+1|k}^T|\mathbf{Z}^k\right] &= E\left[\tilde{\mathbf{x}}_{k+1|k}\left[\mathbf{H}_{k+1}\tilde{\mathbf{x}}_{k+1|k} + \mathbf{w}_{k+1}\right]^T|\mathbf{Z}^k\right] \\ &= \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T \end{aligned} \tag{19}$$

The filter gain can be calculated as

$$\mathbf{W}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T\mathbf{S}_{k+1}^{-1} \tag{20}$$

Thus the updated state estimate can be written as

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{W}_{k+1}\tau_{k+1} \tag{21}$$

where

$$\tau_{k+1} = \mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k} = \tilde{\mathbf{z}}_{k+1|k} \tag{22}$$

which is called innovation or measurement residual. Finally, the updated covariance of the state at time $k+1$ is,

$$\begin{aligned} \mathbf{P}_{k+1|k+1} &= \mathbf{P}_{k+1|k} - \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T S_{k+1}^{-1}\mathbf{H}_{k+1}\mathbf{P}_{k+1|k} \\ &= \mathbf{P}_{k+1|k} - \mathbf{W}_{k+1}\mathbf{S}_{k+1}\mathbf{W}_{k+1}^T \end{aligned} \tag{23}$$

An alternative form for the covariance update can be provided as

$$\mathbf{P}_{k+1|k+1}^{-1} = \mathbf{P}_{k+1|k}^{-1} + \mathbf{H}_{k+1}^T R_{k+1}^{-1}\mathbf{H}_{k+1} \tag{24}$$

C. Statistical Test for Filter Consistency

Under the linear-Gaussian assumption, the conditional probability density function of the state $\mathbf{x}_k$ at the time k is

$$p(\mathbf{x}_k|\mathbf{Z}^k) = \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_{k|k}) \tag{25}$$

Based on (25), we can get the first two moments,

$$\begin{aligned} &E\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right] = E\left[\tilde{\mathbf{x}}_{k|k}\right] = 0 \\ &E\left[\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right]\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right]^T\right] = E\left[\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T\right] = \mathbf{P}_{k|k} \end{aligned} \tag{26}$$

Define the normalized estimation error squared as

$$\epsilon_k = \tilde{\mathbf{x}}_{k|k}^T \mathbf{P}_{k|k}^{-1} \tilde{\mathbf{x}}_{k|k} \tag{27}$$

Under hypothesis H_0 that the filter is consistent and linear Gaussian assumption, ϵ_k is Chi-square distributed with n_x degrees of freedom, where n_x is the dimension of the system state $\mathbf{x}$, and

$$E\left[\epsilon_k\right] = n_x \tag{28}$$

Based on the Monte Carlo simulations with N independent samples $\epsilon_k^i, i = 1, ..., N$, the sample average of ϵ_k can be obtained,

$$\bar{\epsilon}_k = \frac{1}{N}\sum_{i=1}^{N} \epsilon_k^i \tag{29}$$

It can be shown that $N\bar{\epsilon}_k$ follows a Chi-square distribution with Nn_x degrees of freedom. The hypothesis of H_0 is accepted if

$$\bar{\epsilon}_k \in [r_1, r_2] \tag{30}$$

where the acceptance interval is determined such that

$$P\{\bar{\epsilon}_k \in [r_1, r_2]|H_0\} = 1 - \alpha \tag{31}$$

and α is the power of the test.

IV. System Model

Let us assume that M sensors are used by the linear system. The measurement at time k collected by sensor i is

$$\mathbf{z}_{k,i} = \mathbf{H}_{k,i}\mathbf{x}_{k,i} + \mathbf{w}_{k,i} \tag{32}$$

with $\mathbf{H}_{k,i}$ being the measurement matrix, and $\mathbf{w}_{k,i}$ a zero-mean white Gaussian measurement noise with covariance matrix $E[\mathbf{w}_{k,i}\mathbf{w}_{k,i}^T] = \mathbf{R}_{k,i}$, for $i = 1, \cdots, M$. We further assume that the measurement noises are independent across sensors. The matrices $\mathbf{F}_k$, $\mathbf{G}_k$, $\mathbf{H}_{k,i}$, $\mathbf{Q}_{k,i}$, and $\mathbf{R}_{k,i}$ are assumed to be known with proper dimensions. For such a linear and Gaussian dynamic system, the Kalman filter is the optimal state estimator. In this thesis, we assume that a bias $\mathbf{b}_{k,i}$ is injected by the adversary into the measurement of the ith sensor at time k intentionally. Therefore, the measurement equation (32) becomes

$$\mathbf{z}'_{k,i} = \mathbf{H}_{k,i}\mathbf{x}_k + \mathbf{w}_{k,i} + \mathbf{b}_{k,i} = \mathbf{z}_{k,i} + \mathbf{b}_{k,i} \tag{33}$$

where $\mathbf{z}'_{k,i}$ is the corrupted measurement, $\mathbf{b}_{k,i}$ is either an unknown constant or a random variable independent of $\{\mathbf{v}_{k,i}\}$ and $\{\mathbf{w}_{k,i}\}$.

For compactness, let us denote the system sensor observation as $\mathbf{z}_k = [\mathbf{z}_{k1}^T, \cdots, \mathbf{z}_{kM}^T]^T$, which contains the observations from all the M sensors. Similarly, let us denote the system bias vector as $\mathbf{b}_k = [\mathbf{b}_{k1}^T, \cdots, \mathbf{b}_{kM}^T]^T$ which includes the biases at all the M sensors. Correspondingly, the measurement matrix becomes

$$\mathbf{H}_k = [\mathbf{H}_{k1}^T, \cdots, \mathbf{H}_{kM}^T]^T \tag{34}$$

With these notations, it is easy to convert (32) and (33) into the following equations respectively.

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k \tag{35}$$

and

$$\mathbf{z}'_k = \mathbf{z}_k + \mathbf{b}_k \tag{36}$$

Further, we have the measurement error covariance matrix corresponding to $\mathbf{w}_k$ is

$$\mathbf{R}_k = \begin{bmatrix} \mathbf{R}_{k,1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{R}_{k,M} \end{bmatrix} \tag{37}$$

which is obtained by using the assumption that measurement noises are independent across sensors.

V. Impact of False Information Injection

In this thesis, let us assume that the adversary attacks the system by injecting false information into the sensors while the Kalman filter is unaware of such attacks. We start with the case where biases ($\mathbf{b}_k$) are continuously injected into the system starting from a certain time K. Note that single injection is just a special case of continuous injection when $\mathbf{b}_k$ are set to be nonzero at time K and zero otherwise.

In the continuous injection case, the Kalman filter' extra state estimation error, which is caused by the continuous bias injection alone, is derived in [17] and provided as follows.

Proposition 1. *The Kalman filter's state estimation error at time $K+N$ is*

$$\begin{aligned}\hat{\mathbf{x}}'_{K+N|K+N} - \mathbf{x}_{K+N} &= \hat{\mathbf{x}}_{K+N|K+N} - \mathbf{x}_{K+N} \\ &+ \sum_{m=0}^{N} \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \mathbf{b}_{K+N-m}\end{aligned} \tag{38}$$

where $\hat{\mathbf{x}}'_{K+N|K+N}$ is the Kalman filter's state estimate in the presence of the bias sequence $\{\mathbf{b}_k\}$, $\hat{\mathbf{x}}_{K+N|K+N}$ is the Kalman filter's state estimate in the absence of the bias,

$$\mathbf{B}_K \triangleq (\mathbf{I} - \mathbf{W}_K \mathbf{H}_K)\, \mathbf{F}_{K-1}, \tag{39}$$

$\mathbf{I}$ is the identity matrix, and $\mathbf{W}_K$ is the Kalman filter gain [16] at time K. As a result, the extra state estimation error at time $K+N$ due to the continuous bias $\mathbf{b}_k$ injected at and after time K is

$$\sum_{m=0}^{N} \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \mathbf{b}_{K+N-m}, \tag{40}$$

If $\{\mathbf{b}_k\}$ is a zero-mean, random, and independent sequence, the extra mean squared error (EMSE) at a particular time instant $K+N$ due to the bias alone is provided in the following proposition. Using the results from Proposition 1, the proof of Proposition 2 is provided as well.

Proposition 2. *When the bias sequence $\{\mathbf{b}_k\}$ is zero mean, random, and independent over time, the $EMSE$ at time $K+N$ due to the biases injected at and after time K, denoted as $\mathbf{A}_{K+N}$, is*

$$\mathbf{A}_{K+N} = \sum_{m=0}^{N} \mathbf{D}_m \mathbf{\Sigma}_{K+N-m} \mathbf{D}_m^T \tag{41}$$

where

$$\mathbf{D}_m = \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \tag{42}$$

$\prod_{i=0}^{-1} \mathbf{B}_{K+N-i} = \mathbf{I}$ *is an identity matrix, and* $\mathbf{\Sigma}_{K+N-m}$ *is the covariance matrix of* $\mathbf{b}_{K+N-m}$.

Proof Sketches: Let us denote $\tilde{\mathbf{x}}_{K+N|K+N} = \hat{\mathbf{x}}_{K+N|K+N} - \mathbf{x}_{K+N}$ as the Kalman filter's state estimation error in the absence of any false information, and

$$\mathbf{a}_m = \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \mathbf{b}_{k+N-m} \tag{43}$$

From (38), we can get

$$\begin{aligned}
&\mathbf{A}_{K+N} \\
&= E\left[\left(\tilde{\mathbf{x}}_{K+N|K+N} + \sum_{m=0}^{N} \mathbf{a}_m \right) \left(\tilde{\mathbf{x}}_{K+N|K+N} + \sum_{n=0}^{N} \mathbf{a}_n \right)^T \right] \\
&- E\left(\tilde{\mathbf{x}}_{K+N|K+N} \tilde{\mathbf{x}}_{K+N|K+N}^T \right) \\
&= E\left(\tilde{\mathbf{x}}_{K+N|K+N} \sum_{n=0}^{N} \mathbf{a}_n^T \right) + E\left(\sum_{m=0}^{N} \mathbf{a}_m \tilde{\mathbf{x}}_{K+N|K+N}^T \right) \\
&+ E\left(\sum_{m=0}^{N} \sum_{n=0}^{N} \mathbf{a}_m \mathbf{a}_n^T \right) \\
&= E\left(\sum_{m=0}^{N} \sum_{n=0}^{N} \mathbf{a}_m \mathbf{a}_n^T \right)
\end{aligned}$$

where the last line is due to the fact that $\mathbf{a}_m$ and $\mathbf{a}_n$ have zero mean, are independent from each other when $m \neq n$, and are independent from $\tilde{\mathbf{x}}_{K+N|K+N}$. Using this fact again, we further have

$$\begin{aligned}
E\left(\sum_{m=0}^{N} \sum_{n=0}^{N} \mathbf{a}_m \mathbf{a}_n^T \right) &= E\left(\sum_{m=0}^{N} \mathbf{a}_m \mathbf{a}_m^T \right) \\
&= \sum_{m=0}^{N} \mathbf{D}_m \mathbf{\Sigma}_{K+N-m} \mathbf{D}_m^T
\end{aligned} \tag{44}$$

where $\mathbf{D}_m$ has been defined in Proposition 2.

VI. The Optimal Attack Strategy

A. Problem Formulation for a General Linear System

In this thesis, we investigate the optimal attack strategy that an adversary can adopt to maximize the system estimator's estimation error. This problem can be formulated as a constrained optimization problem. Without loss of generality, let us consider that the attacker is interested in maximizing the system state estimation error at time K right after a single false bias is injected at time K. In this case, we are interested in designing the injected random bias' covariance matrix such that

$$\begin{aligned} &\max_{\mathbf{\Sigma}_K} \mathrm{Tr}\left[\mathbf{P}_{K|K} + \mathbf{A}_K(\mathbf{\Sigma}_K)\right] \\ &s.t. \quad \mathrm{Tr}(\mathbf{\Sigma}_K) = a^2 \end{aligned} \tag{45}$$

where a is a constant, $\mathrm{Tr}(\cdot)$ is the matrix trace operator, and $\mathbf{P}_{K|K}$ is the Kalman filter's state estimation error covariance matrix at time K in the absence of any false information. Note that it is meaningful to have a constraint on the trace of $\mathbf{\Sigma}_K$, since it can be deemed as the power of injected sensor bias $\mathbf{b}_K$, and a smaller power for $\mathbf{b}_K$ reduces the probability that the adversary is detected by the system estimator using an innovation based detector. Note that the optimization problem is equivalent to the one that maximizes $\mathrm{Tr}\left(\mathbf{A}_K(\mathbf{\Sigma}_K)\right)$, since $\mathbf{P}_{K|K}$ is not a function of $\mathbf{\Sigma}_K$, and trace is a linear operator. If one is more interested in the determinant of the mean squared estimation error matrix, a similar optimization problem can be easily formulated as follows.

$$\begin{aligned} &\max_{\mathbf{\Sigma}_K} \left|\mathbf{P}_{K|K} + \mathbf{A}_K(\mathbf{\Sigma}_K)\right| \\ &s.t. \quad \mathrm{Tr}(\mathbf{\Sigma}_K) = a^2 \end{aligned} \tag{46}$$

B. Equivalent Measurement in Multi-Sensor Systems

To simplify the mathematical analysis, it is helpful to derive the equivalent sensor measurement, which is a linear combination of the observations from all the sensors, and is a sufficient statistic containing all the information about the systems state. The equivalent sensor measurement vector and its corresponding covariance matrix should have much smaller dimensionality than the original measurement vector and its covariance, making the mathematical manipulation

and derivation later in the thesis much simpler. In a information filter recursion [16], which is equivalent to the Kalman filter recursion, we have

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{z}_k \tag{47}$$

where $\hat{\mathbf{y}}_{k|k} = \mathbf{P}_{k|k}^{-1} \mathbf{x}_{k|k}$ and $\hat{\mathbf{y}}_{k|k-1} = \mathbf{P}_{k|k-1}^{-1} \mathbf{x}_{k|k-1}$. It is clear that $\hat{\mathbf{y}}_{k|k-1}$ represents the prior knowledge about the system state based on past sensor data, and the second term in (47) represents the new information from the new sensor data $\mathbf{z}_k$, which can be expanded by using (34) and (37) as follows.

$$\begin{aligned} &\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{z}_k \\ &= [\mathbf{H}_{k1}^T, \cdots, \mathbf{H}_{kM}^T] \begin{bmatrix} \mathbf{R}_{k1}^{-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{R}_{kM}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{k1} \\ \vdots \\ \mathbf{z}_{kM} \end{bmatrix} \\ &= \sum_{i=1}^{M} \mathbf{H}_{ki}^T \mathbf{R}_{ki}^{-1} \mathbf{z}_{ki} \end{aligned} \tag{48}$$

In the following derivations, we skip the time index k for simplicity. Our purpose is to find an equivalent measurement $\mathbf{z}_e$ such that

$$\mathbf{z}_e = \mathbf{H}_e \mathbf{x} + \mathbf{w}_e \tag{49}$$

where $\mathbf{w}_e \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_e)$, and

$$\mathbf{H}_e^T \mathbf{R}_e^{-1} \mathbf{z}_e = \sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{z}_i \tag{50}$$

Let us consider two cases. First, suppose all the $\mathbf{H}_i$s are the same ($\mathbf{H}_i = \mathbf{H}$) , then it is natural to set $\mathbf{H}_e = \mathbf{H}$. Note that a sufficient condition for (50) to be true is

$$\mathbf{z}_e = \mathbf{R}_e \sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{z}_i \tag{51}$$

Taking the covariance on the both sides of (51), we get

$$\begin{aligned} \mathbf{R}_e &= \mathbf{R}_e \mathrm{cov} \left(\sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{z}_i \right) \mathbf{R}_e^T \\ &= \mathbf{R}_e \left[\sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{R}_i (\mathbf{R}_i^{-1})^T \right] \mathbf{R}_e^T \end{aligned} \tag{52}$$

This implies that

$$\mathbf{R}_e = \left(\sum_{i=1}^{M} \mathbf{R}_i^{-1} \right)^{-1} \tag{53}$$

In the second case, let us assume that the system state $\mathbf{x}$ is observable based on the observations from all the sensors, meaning that the Fisher information matrix $\sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i$ is invertible. In this case, by setting $\mathbf{H}_e = \mathbf{I}$, using (50), and following a similar procedure as in the first case, we have

$$\mathbf{z}_e = \mathbf{R}_e \sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{z}_i \tag{54}$$

and

$$\mathbf{R}_e = \left(\sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i \right)^{-1} \tag{55}$$

VII. A TARGET TRACKING EXAMPLE

In this thesis, we give a concrete target tracking example. We assume that the target moves in a 1-dimensional space according to a discrete white noise acceleration model [16], which can still be described by the plant and measurement equations given in (1) and (32). In such a system, the state is defined as $\mathbf{x_k} = [\xi_k \;\; \dot{\xi}_k]^T$, where ξ_k and $\dot{\xi}_k$ denote the target's position and velocity at time k respectively. The input $\mathbf{u}_k$ is a zero sequence. The state transition matrix is

$$\mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{56}$$

where T is the time between measurements. The process noise is $\mathbf{v}_k = \mathbf{\Gamma} v_k$, where v_k is a zero mean white acceleration noise, with variance σ_v^2, and the vector gain multiplying the scalar process noise is given by $\mathbf{\Gamma}^T = [T^2/2 \quad T]$. The covariance matrix of the process noise is therefore $\mathbf{Q} = \sigma_v^2 \mathbf{\Gamma}\mathbf{\Gamma}^T$.

In this thesis, we investigate the attack strategies for two scenarios. In the first scenario, only position measurements are available to the sensors, whereas in the second scenario, the sensors measure both position and velocity of the target.

A. Attack Strategy Analysis from Trace Perspective

1) Attack Strategy for Multiple Position Sensors: In this case, it is assumed that at each sensor, only the position measurement is available, so that $\mathbf{H}_i = [1 \;\; 0]$. At each sensor, the measurement noise process is zero-mean, white, and with variance, $\sigma_{w_i}^2$. In order to simplify the problem, we think of $\mathbf{z}_{e_k}$ as the equivalent measurement, which is a linear combination of the measurements from all the sensors. Using the results we derived in Section VI-B for the first case, namely (51) and (53), the measurement equation (33) becomes

$$z'_k = z_{ek} + b_{ek} \tag{57}$$

where

$$z_{ek} = \sum_{m=0}^{M} c_i z_{ki} \tag{58}$$

$$b_{ek} = \sum_{m=0}^{M} c_i b_{ki} \tag{59}$$

and

$$c_i = \frac{1/\sigma_{w_i}^2}{\sum_{j=1}^{M}\left(1/\sigma_{w_j}^2\right)} \tag{60}$$

which is the corresponding coefficient/weight for the ith sensor. In this target tracking problem, let us first consider the strategy that maximizes the trace of the Kalamn filter mean squared estimation error matrix, which is the solution of (45) in Section VI-A. In this case,

$$\mathbf{\Sigma}_K = \begin{bmatrix} \sigma_{b_1}^2 & \rho_{12}\sigma_{b_1}\sigma_{b_2} & \cdots & \rho_{1M}\sigma_{b_1}\sigma_{b_M} \\ \rho_{12}\sigma_{b_1}\sigma_{b_2} & \sigma_{b_2}^2 & \cdots & \rho_{2M}\sigma_{b_2}\sigma_{b_M} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1M}\sigma_{b_1}\sigma_{b_M} & \rho_{2M}\sigma_{b_2}\sigma_{b_M} & \cdots & \sigma_{b_M}^2 \end{bmatrix} \tag{61}$$

where $\sigma_{b_i}^2$ is the variance of the random bias injected at the ith sensor (b_i), and ρ_{ij} is the correlation coefficient between b_i and b_j. Therefore, (45) is equivalent to

$$\begin{aligned} &\max \mathrm{Tr}\left[\mathbf{A}_K\right] \\ &s.t. \quad \sum_{i=1}^{M} \sigma_{b_i}^2 = a^2 \\ &\quad -1 \le \rho_{ij} \le 1, \ \text{ for } \ 1 \le i,j \le M \end{aligned} \tag{62}$$

To simplify this problem, we first use the equivalent measurement approach to convert the multi-sensor problem to a single sensor problem. Namely, in Proposition 2 by replacing

$$\mathbf{H}_k = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix}$$

with $\mathbf{H}_e = [1\ 0]$, and replacing $\mathbf{\Sigma}_K$ with

$$\begin{aligned} \Sigma_{e_K} &= E[b_{e_K}^2] \\ &= E\left[\left(\sum_{i=1}^{M} c_i b_i\right)^2\right] \\ &= \sum_{i=1}^{M} c_i^2 \sigma_{b_i}^2 + \sum_{i}\sum_{j\neq i} 2\rho_{ij} c_i c_j \sigma_{b_i}\sigma_{b_j} \end{aligned} \tag{63}$$

we can easily show that $\mathbf{A}_K = \mathbf{D}_0 \Sigma_{e_K} \mathbf{D}_0^T$. Since Σ_{e_K} is a scalar and $\mathbf{D}_0$ is not a function of $\mathbf{\Sigma}_K$, maximizing the trace of $\mathbf{A}_K$ is equivalent to maximizing Σ_{e_K}.

First, let us consider the case where the random biases at different sensors are independent meaning that $\rho_{i,j} = 0$ for $1 \leq i, j \leq M$. The optimal strategy for the adversary in this case is clearly to put all the bias power to the sensor with the largest coefficient c_i:

Proposition 3. *For a system with M sensors, if the adversary injects independent random noises, the best strategy is to allocate all the power to the sensor with the smallest noise variance.*

Next, let us consider the more general case where the random biases are dependent. By inspecting (63), it is clear that to maximize Σ_{e_K}, we need to set all the ρ_{ij}s to 1. As a result, (63) becomes

$$\Sigma_{e_K} = \left(\sum_{i=1}^{M} c_i \sigma_{b_i} \right)^2 \tag{64}$$

Now, the optimization problem in (62) has been converted to the following problem:

$$\begin{aligned} &\max \left(\sum_{i=1}^{M} c_i \sigma_{b_i} \right)^2 \\ &s.t. \quad \sum_{i=1}^{M} \sigma_{b_i}^2 = a^2 \end{aligned} \tag{65}$$

The above problem can be solved by using standard constrained optimization techniques [18] based on gradient and Hessian, which are rather involved. Here we solve the problem using a much simpler geometric solution, which has been shown to give the same solution as that by the standard optimization techniques. We start with the simplest case with two sensors, in which we need to solve the following optimization problem.

$$\begin{aligned} \max \quad & c_1 \sigma_{b_1} + c_2 \sigma_{b_2} \\ \text{s.t.} \quad & \sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2 \end{aligned} \tag{66}$$

We can get the optimal solution by analyzing the problem geometrically with the norm vector $(c_1, c_2)^T$ of the objective function as shown in Fig. 1. The constraint of the problem is represented by the circle with a radius of a. We move the line l_1 with the slope $-\frac{c_1}{c_2}$ to get the largest intercept between l_1 and σ_2 axis under the constraint that there is an intersection between the circle and the line l_1. The corresponding optimal solution is found when l_1 becomes a tangent line to the

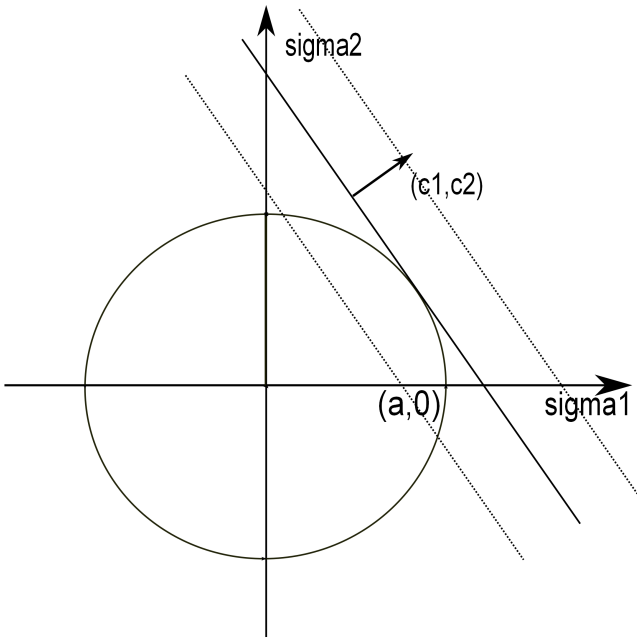

Fig. 1. Geometric solution for systems with two sensors.

circle, which is

$$\begin{aligned} \sigma_1 &= \frac{c_1 a}{\sqrt{c_1^2 + c_2^2}} \\ \sigma_2 &= \frac{c_2 a}{\sqrt{c_1^2 + c_2^2}} \end{aligned} \tag{67}$$

For a system with arbitrary number of sensors, we can repeat the same procedure to find the optimal solution by using hyper planes and hyperspheres. In general, the optimal attack strategy can be found and summarized as follows.

Theorem 1. *For a system with M sensors, the optimal strategy for the adversary is to inject statistically correlated random noises with a pairwise correlation coefficient of 1. The random bias power is allocated such that*

$$\sigma_{b_i} = \frac{c_i a}{\sqrt{\sum_{j=1}^{M} c_j^2}}, \quad \text{for } i = 1, \cdots, M. \tag{68}$$

2) Attack Strategy for a Single Position and Velocity Sensor: In this case, let us assume that the sensors collect both position and velocity measurements of the target. Therefore, the measurement matrix for the ith sensor is $\mathbf{H}_i = \mathbf{I}_2$, where $\mathbf{I}_2$ is a 2×2 identity matrix. At the

ith sensor, the adversary injects the bias noise vector $\mathbf{b}_{k_i}$ to the sensor measurement $\mathbf{z}_{k_i}$, where $\mathbf{b}_{k_i} = [b_{p_i} \;\; b_{v_i}]^T$ consists of biases in position and velocity measurements. Let us assume that the system bias vector $\mathbf{b}_k = [\mathbf{b}_{k1}^T, \cdots, \mathbf{b}_{kM}^T]^T$ is zero-mean and has a $2M \times 2M$ covariance matrix $\mathbf{\Sigma}_K$. Further, the $(i,\ j)$th 2×2 submatrix for $\mathbf{\Sigma}_K$ is defined as

$$\mathbf{\Sigma}_K(i,j) = \begin{bmatrix} \rho_{b_{p_i},b_{p_j}}\sigma_{b_{p_i}}\sigma_{b_{p_j}} & \rho_{b_{p_i},b_{vj}}\sigma_{b_{p_i}}\sigma_{b_{vj}} \\ \rho_{b_{vi},b_{p_j}}\sigma_{b_{vi}}\sigma_{b_{p_j}} & \rho_{b_{vi},b_{vj}}\sigma_{b_{vi}}\sigma_{b_{vj}} \end{bmatrix} \tag{69}$$

for $1 \leq i,\ j \leq M$. $\sigma_{b_{p_i}}$ and $\sigma_{b_{vi}}$ are the position and velocity bias noise standard deviations at the ith sensor respectively. The ρs are defined as the proper correlation coefficients between components of the bias vector, and $\rho_{b_{p_i},b_{p_i}} = \rho_{b_{vi},b_{vi}} = 1$, for $1 \leq i \leq M$. Since the position bias b_p and velocity bias b_v have different units, we need an appropriate constraint for bias noise power. Here we assume that the total noise power is defined as

$$\sum_{i=1}^{M} \sigma_{b_{p_i}}^2 + T^2\sigma_{b_{v_i}}^2 \tag{70}$$

Note that this is a meaningful power definition, since the two terms in the above equation has the same unit. Recall that according to the target tracking system plant equation and ignoring the system process noise, we have $\xi_{k+1} = \xi_k + T\dot{\xi}_k$. Therefore, the power defined in (70) can be interpreted as the summation of the extra mean squared errors for the position estimate caused by independent bias injections. We can see that the best attack strategy derived under a constraint on power defined in (70) can be easily adjusted and extended for other power definitions, as long as in the new definition, the second term is proportional to $T^2\sigma_{b_{v_i}}^2$.

As we can use the equivalent sensor to represent the multiple sensors, we focus on the single-sensor case first. If we are interested in the case of $N = 0$, maximizing the trace of $\mathbf{A}_K$ is equivalent to maximize the $\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T$. We assume that the adversary knows the system models and the prior information $\mathbf{P}_{0|0}$ at time zero, so that he/she can calculate the offline Kalman filter gain matrix $\mathbf{W}_k$ recursively. Therefore, the best strategy the adversary can adopt to attack the system is the solution to the following optimization problem:

$$\begin{aligned} &\max_{\mathbf{\Sigma}_K} \mathrm{Tr}\left[\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right] \\ &s.t. \quad \sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = a^2 \\ &\qquad -1 \leq \rho_{b_p,b_v} \leq 1 \\ &\qquad \sigma_{b_p}, \sigma_{b_v} > 0 \end{aligned} \tag{71}$$

where

$$\mathbf{\Sigma}_K = \begin{bmatrix} \sigma_{b_p}^2 & \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} \\ \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} & \sigma_{b_v}^2 \end{bmatrix} \tag{72}$$

and

$$\mathbf{W}_K = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \tag{73}$$

It is easy to show that

$$\begin{aligned} \mathrm{Tr}\left[\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right] &= \mathrm{Tr}\left[\mathbf{W}_K^T\mathbf{W}_K\mathbf{\Sigma}_K\right] \\ &= (w_{11}^2 + w_{21}^2)\sigma_{b_p}^2 + (w_{12}^2 + w_{22}^2)\sigma_{b_v}^2 \\ &+ 2(w_{11}w_{12} + w_{21}w_{22})\rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} \end{aligned} \tag{74}$$

According to the sign of $(w_{11}w_{12} + w_{21}w_{22})$, we can set the value of the ρ_{b_p,b_v} to maximize the objective function. For example, if $(w_{11}w_{12} + w_{21}w_{22})$ is positive, we set $\rho_{b_p,b_v} = 1$ and the optimization problem becomes

$$\begin{aligned} &\max(w_{11}\sigma_{b_p} + w_{12}\sigma_{b_v})^2 + (w_{21}\sigma_{b_p} + w_{22}\sigma_{b_v})^2 \\ &s.t. \quad \sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = a^2 \\ &\quad \sigma_{b_p}, \sigma_{b_v} \geq 0 \end{aligned} \tag{75}$$

We have solved the optimization problem in (75), and summarize the results in the following theorem.

Theorem 2. *For a system with one sensor observing position and velocity of the target, the optimal strategy for the adversary is to inject random noise that has dependent position and velocity components. If* $w_{11}w_{12} + w_{21}w_{22} > 0$*, the correlation coefficient* ρ_{b_p,b_v} *should be set as*

1, *and the random bias power is allocated such that*

$$\begin{aligned}
&\sigma_{b_p} = a\sin(\theta^*) \\
&\sigma_{b_v} = \frac{a}{T}\cos(\theta^*) \\
&\theta^* = \frac{\pi}{4} - \frac{\phi}{2} \\
&\phi = \arctan\left[\frac{\beta_2 - \beta_1 T^2}{2T(\alpha_1+\alpha_2)}\right] \\
&w_{11}^2 + w_{21}^2 = \beta_1 \\
&w_{12}^2 + w_{22}^2 = \beta_2 \\
&w_{11}w_{12} = \alpha_1 \\
&w_{21}w_{22} = \alpha_2
\end{aligned} \tag{76}$$

When $w_{11}w_{12}+w_{21}w_{22} < 0$*, we should set* $\rho_{b_p,b_v} = -1$ *and set* $\alpha_1 = -w_{11}w_{12}$ *and* $\alpha_2 = -w_{21}w_{22}$*. The rest of the equations in formula (76) remain the same.*

Proof Sketches: Let us first denote

$$\begin{aligned}
w_{11}^2 + w_{21}^2 &= \beta_1 \\
w_{12}^2 + w_{22}^2 &= \beta_2 \\
w_{11}w_{12} &= \alpha_1 \\
w_{21}w_{22} &= \alpha_2
\end{aligned} \tag{77}$$

The constraint in (71) can be written as

$$\frac{\sigma_{b_p}^2}{T^2} + \sigma_{b_v}^2 = \frac{a^2}{T^2} = a_1^2 \tag{78}$$

Now we set $\sigma_{b_p} = a_1 T\sin(\theta)$ and $\sigma_{b_v} = a_1\cos(\theta)$. Plugging σ_{b_p} and σ_{b_v} into the objective function, we have the following equivalent optimization problem

$$\begin{aligned}
&\max_{\theta} a_1^2\left[\frac{\beta_1 T_1^2 + \beta_2}{2} + A\sin(2\theta+\phi)\right] \\
&s.t. \quad 0 \le \theta \le \frac{\pi}{2}
\end{aligned} \tag{79}$$

where

$$A = \sqrt{\frac{1}{4}\left(\beta_2 - \beta_1 T^2\right)^2 + T^2(\alpha_1 + \alpha_2)^2} \tag{80}$$

$$\tan(\phi) = \frac{\beta_2 - \beta_1 T^2}{2T(\alpha_1 + \alpha_2)} \tag{81}$$

Clearly, the optimal solution is

$$\theta^* = \frac{\pi}{4} - \frac{\phi}{2} \tag{82}$$

3) Attack Strategy for Multiple Position and Velocity Sensors: In this case, let us consider the case of $M = 2$, where the measurement matrix is $\mathbf{H} = [\mathbf{I}_2 \ \mathbf{I}_2]^T$. The measurement covariance matrix for the ith sensor is assumed to be

$$\mathbf{R}_i = \begin{bmatrix} \sigma_{p_i}^2 & 0 \\ 0 & \sigma_{v_i}^2 \end{bmatrix} \tag{83}$$

Now, according to (55), we have

$$\begin{aligned} \mathbf{R}_e &= [\mathbf{R}_1^{-1} + \mathbf{R}_2^{-1}]^{-1} \\ &= \begin{bmatrix} \left(\sigma_{p_1}^{-2} + \sigma_{p_2}^{-2}\right)^{-1} & 0 \\ 0 & \left(\sigma_{v_1}^{-2} + \sigma_{v_2}^{-2}\right)^{-1} \end{bmatrix} \end{aligned} \tag{84}$$

According to (54), we define

$$\begin{aligned} \mathbf{C}_i &= \mathbf{R}_e \mathbf{H}_i^T \mathbf{R}_i^{-1} \\ &= \begin{bmatrix} \frac{\sigma_{p_i}^{-2}}{\sigma_{p_1}^{-2} + \sigma_{p_2}^{-2}} & 0 \\ 0 & \frac{\sigma_{v_i}^{-2}}{\sigma_{v_1}^{-2} + \sigma_{v_2}^{-2}} \end{bmatrix} \end{aligned} \tag{85}$$

as the weighting matrix for the ith sensor's observation $\mathbf{z}_i$. Further, we define

$$\begin{aligned} c_{p_i} &= \mathbf{C}_i(1,1) \\ c_{v_i} &= \mathbf{C}_i(2,2) \end{aligned} \tag{86}$$

both of which are positive numbers. The equivalent noise injection is therefore

$$\mathbf{b}_{eK} = \sum_{i=1}^{2} \mathbf{C}_i \mathbf{b}_{K_i} \tag{87}$$

So the covariance matrix of the equivalent bias vector is

$$\mathbf{\Sigma}_{eK} = \sum_{i=1}^{2} \sum_{i=j}^{2} \mathbf{C}_i \mathbf{\Sigma}_K(i,j) \mathbf{C}_j^T \tag{88}$$

where $\mathbf{\Sigma}_K(i,j)$ has been defined in (69). It can be shown that

$$\mathbf{\Sigma}_{eK} = \begin{bmatrix} s_1 & s_2 \\ s_2 & s_3 \end{bmatrix} \tag{89}$$

where

$$\begin{aligned} s_1 &= c_{p_1}^2\sigma_{b_{p_1}}^2 + c_{p_2}^2\sigma_{b_{p_2}}^2 + 2\rho_{b_{p_1},b_{p_2}}c_{p_1}c_{p_2}\sigma_{b_{p_1}}\sigma_{b_{p_2}} \\ s_3 &= c_{v_1}^2\sigma_{b_{v_1}}^2 + c_{v_2}^2\sigma_{b_{v_2}}^2 + 2\rho_{b_{v_1},b_{v_2}}c_{v_1}c_{v_2}\sigma_{b_{v_1}}\sigma_{b_{v_2}} \end{aligned} \tag{90}$$

$$\begin{aligned} s_2 = {} & c_{p_1}c_{v_1}\rho_{b_{p_1},b_{v_1}}\sigma_{b_{p_1}}\sigma_{b_{v_1}} + c_{p_1}c_{v_2}\rho_{b_{p_1},b_{v_2}}\sigma_{b_{p_1}}\sigma_{b_{v_2}} \\ & + c_{p_2}c_{v_1}\rho_{b_{p_2},b_{v_1}}\sigma_{b_{p_2}}\sigma_{b_{v_1}} + c_{p_2}c_{v_2}\rho_{b_{p_2},b_{v_2}}\sigma_{b_{p_2}}\sigma_{b_{v_2}} \end{aligned} \tag{91}$$

The optimization problem can be written as follows.

$$\begin{aligned} &\max_{\mathbf{\Sigma}_{eK}} \mathrm{Tr}\left[\mathbf{W}_{eK}\mathbf{\Sigma}_{eK}\mathbf{W}_{eK}^T\right] \\ &s.t. \quad \sigma_{b_{p_1}}^2 + \sigma_{b_{p_2}}^2 + T^2\sigma_{b_{v_1}}^2 + T^2\sigma_{b_{v_2}}^2 = a^2, \\ &\quad -1 \le \rho_{p_i,v_j} \le 1, \\ &\quad -1 \le \rho_{v_i,v_j} \le 1, \\ &\quad -1 \le \rho_{p_i,p_j} \le 1, \\ &\quad \sigma_{p_i}, \sigma_{v_i} \ge 0, \quad \forall i,j \in \{1,2\} \end{aligned} \tag{92}$$

where

$$\mathbf{W}_{eK} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \tag{93}$$

is the Kalman filter gain calculated using the equivalent measurement covariance matrix $\mathbf{R}_e$ and equivalent measurement matrix $\mathbf{H}_e$. It is easy to show that

$$\begin{aligned} &\mathrm{Tr}\left[\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right] = \mathrm{Tr}\left[\mathbf{W}_K^T\mathbf{W}_K\mathbf{\Sigma}_K\right] \\ &= (w_{11}^2 + w_{21}^2)^2 s_1 + (w_{12}^2 + w_{22}^2)^2 s_3 \\ &+2(w_{11}w_{12} + w_{21}w_{22})s_2 \end{aligned} \tag{94}$$

Clearly, all the ρs that appear in s_1 and s_3 should be set as 1 to maximize the objective function. The optimal values for ρs in s_2 depend on the Kalman filter gain $\mathbf{W}_{eK}$. More specifically, when

$w_{11}w_{12} + w_{21}w_{22} > 0$, all the ρs that appear in s_2 should be set to 1; otherwise, they should be set as -1. Let us first suppose that $w_{11}w_{12} + w_{21}w_{22} > 0$ is true, then we have

$$\begin{aligned}
&\mathrm{Tr}\left[\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right] = (w_{11}^2 + w_{21}^2)^2(c_{p_1}\sigma_{p_1} + c_{p_2}\sigma_{p_2})^2 \\
&+ (w_{12}^2 + w_{22}^2)^2(c_{v_1}\sigma_{v_1} + c_{v_2}\sigma_{v_2})^2 \\
&+ 2(w_{11}w_{12} + w_{21}w_{22})(c_{p_1}c_{v_1}\sigma_{p_1}\sigma_{v_1} + c_{p_1}c_{v_2}\sigma_{p_1}\sigma_{v_2} \\
&+ c_{p_2}c_{v_1}\sigma_{p_2}\sigma_{v_1} + c_{p_2}c_{v_2}\sigma_{p_2}\sigma_{v_2})
\end{aligned} \tag{95}$$

So far, we have converted the objective function in (92), which involves 10 variables to one that involves only 4 variables. Considering that the power constraint reduces one degree of freedom, we only need to solve an optimization problem in a 3-dimensional space.

4) Strategy for a Single Sensor with Multiple Time Attacks: Based on Proposition 2, we get the extra mean squared error matrix,

$$\mathbf{A}_{K+N} = \sum_{m=0}^{N} \mathbf{D}_m\mathbf{\Sigma}_{K+N-m}\mathbf{D}_m^T$$

Supposing that the adversary attacks the system continuously from time K to $K+N$, the weighted extra mean squared error matrices at different time are as shown below,

$$\begin{aligned}
&\mathbf{A}_{K+0}^{'} = \alpha_0(\mathbf{D}_0\mathbf{\Sigma}_K\mathbf{D}_0^T) \\
&\mathbf{A}_{K+1}^{'} = \alpha_1(\mathbf{D}_0\mathbf{\Sigma}_{K+1}\mathbf{D}_0^T + \mathbf{D}_1\mathbf{\Sigma}_K\mathbf{D}_1^T) \\
&... \\
&\mathbf{A}_{K+N}^{'} = \alpha_N(\mathbf{D}_0\mathbf{\Sigma}_{K+N}\mathbf{D}_0^T + ... + \mathbf{D}_N\mathbf{\Sigma}_K\mathbf{D}_N^T)
\end{aligned} \tag{96}$$

where $\alpha_i, i \in N$ is the weight of extra mean squared error matrix at time i, and $\sum_{i=0}^{N}\alpha_i = 1$. The objective function in the multi-time attack problem is the trace of weighted sum of the extra mean squared error matrices at different time:

$$\mathrm{Tr}\left(\sum_{i=0}^{N}\alpha_i\mathbf{A}_{K+i}\right) = \mathrm{Tr}\left(\sum_{i=0}^{N}\mathbf{A}_{K+i}^{'}\right) \tag{97}$$

Maximizing the term above is equivalent to maximize the trace of the weighted sum of the mean squared error matrices of the state estimates over time, because once the system reaches its steady state, $\mathbf{P}_{K+i|K+i}$ will be a constant, and the weighted sum of $\mathbf{P}_{K+i|K+i}$ will remain the

same. First we study the case with position sensors only, where all the terms in (96) are scalars. Using lower case d, σ_p to denote $\mathbf{D}, \Sigma$, we can formulate the optimization problem below,

$$\begin{aligned}
\max \quad & \sum_{i=0}^{N} \alpha_i \mathbf{A}_{K+i} = \sum_{i=0}^{N} \mathbf{A}^{'}_{K+i} \\
& = \sigma^2_{p_{K+0}}(\alpha_0 d_0^2 + \alpha_1 d_1^2 + ... + \alpha_N d_N^2) \\
& + \sigma^2_{p_{K+1}}(\alpha_1 d_0^2 + \alpha_2 d_1^2 + ... + \alpha_N d_{N-1}^2) \\
& + \sigma^2_{p_{K+2}}(\alpha_2 d_0^2 + \alpha_3 d_1^2 + ... + \alpha_N d_{N-2}^2) \\
& + ... \\
& + \sigma^2_{p_{K+N}}(\alpha_N d_0^2) \\
s.t. \quad & \sum_{i=K}^{K+N} \sigma^2_{p_i} \leq a^2 \\
& \sum_{i=0}^{N} \alpha_i = 1
\end{aligned} \tag{98}$$

The adversary can allocate the power based on the coefficient of the variance variables at different time. For example, if the weights $\alpha_i' s$ are all the same, the best strategy is to allocate all the power to the sensor at the first beginning (at time K) because the coefficient for $\sigma^2_{p_{K+0}}$ is the largest.

Second, let us consider the case with position and velocity sensors. The optimization problem

can be characterized as follows,

$$\begin{aligned}
\max \quad & \mathrm{Tr}\left[\sum_{m=0}^{N}\alpha_i \mathbf{A}_{K+i}\right] = \mathrm{Tr}\left[\sum_{m=0}^{N}\mathbf{A}^{'}_{K+i}\right] \\
& = \mathrm{Tr}\left[\mathbf{\Sigma}_{K+0}(\alpha_0\mathbf{D}_0^T\mathbf{D}_0 + ... + \alpha_N\mathbf{D}_N^T\mathbf{D}_N)\right] \\
& + \mathrm{Tr}\left[\mathbf{\Sigma}_{K+1}(\alpha_1\mathbf{D}_0^T\mathbf{D}_0 + ... + \alpha_N\mathbf{D}_{N-1}^T\mathbf{D}_{N-1})\right] \\
& + \mathrm{Tr}\left[\mathbf{\Sigma}_{K+2}(\alpha_2\mathbf{D}_0^T\mathbf{D}_0 + ... + \alpha_N\mathbf{D}_{N-2}^T\mathbf{D}_{N-2})\right] \\
& + ... \\
& + \mathrm{Tr}\left[\mathbf{\Sigma}_{K+N}(\alpha_N\mathbf{D}_0^T\mathbf{D}_0)\right] \\
s.t. \quad & \sum_{i=K}^{K+N}\sigma_{p_i}^2 + T^2\sigma_{v_i}^2 \leq a^2 \\
& \sum_{i=0}^{N}\alpha_i = 1
\end{aligned} \tag{99}$$

Since $\mathbf{\Sigma}_i$ and $\mathbf{D}_j^T\mathbf{D}_j$ are positive semidefinite matrices, $\mathrm{Tr}\left[\Sigma_i(\mathbf{D}_j^T\mathbf{D}_j)\right] \geq 0$. Trace is a monotonically increasing function of the positive semidefinite matrix, that is to say, if $\mathbf{A}$ and $\mathbf{B}$ are both positive semidefinite matrices, and $\mathbf{A} > \mathbf{B}$, then $Tr(\mathbf{A}) > Tr(\mathbf{B})$. So the best strategy for the adversary to attack the system is to allocate all the power to $\mathbf{\Sigma}_i$ with the largest positive semidefinite matrix of $\sum_{i=s}^{N}\alpha_i\mathbf{D}_{i-s}^T\mathbf{D}_{i-s}$. If the adversary's goal is to maximize the average effect of the false information, the best strategy is to put all the power at time K.

B. Attack Strategy Analysis from Determinant Perspective

1) Attack Strategy for Multiple Position Sensors: We are also interested in the effect of bias information on the Kalman filter's mean squared estimation error from the determinant perspective. By using the equivalent measurement approach as in Section VII-A1, we have

$$\begin{aligned}
|\mathbf{P}_{K|K} + \mathbf{A}_K| &= |\mathbf{P}_{K|K} + \Sigma_{eK}\mathbf{D}_0\mathbf{D}_0^T| \\
&= |\mathbf{P}_{K|K}||\mathbf{I} + \Sigma_{eK}\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T|
\end{aligned} \tag{100}$$

where $\mathbf{D}_0$ can be obtained using (42) and Σ_{eK} is defined in (63). As $\mathbf{P}_{K|K}$ is constant and positive definite, $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$ is positive semidefinite meaning that all the eigenvalues of the $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$ are non-negative. First, let us denote $\mathbf{C}$ as a square matrix whose columns are the

eigenvectors of $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$. Then through eigendecomposition, (100) can be written concisely as,

$$\begin{aligned} &|\mathbf{P}_{K|K}||\mathbf{C}\mathbf{I}\mathbf{C}^{-1} + \Sigma_{eK}\mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}| \\ &= |\mathbf{P}_{K|K}||\mathbf{I} + \Sigma_{eK}\mathbf{\Lambda}| \end{aligned} \tag{101}$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues of the $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$. So we just need to maximize Σ_{eK} in order to maximize the determinant of $\mathbf{P}_{K|K} + \mathbf{A}_K$. This is equivalent to maximizing the trace of $\mathbf{P}_{K|K} + \mathbf{A}_K$ as discussed in Section VII-A1.

2) Attack Strategy for a Single Position and Velocity Sensor: We assume that the adversary knows the system model and the prior information $\mathbf{P}_{0|0}$ at time zero, so that he/she can calculate the offline Kalman filter gain matrix $\mathbf{W}_k$ recursively. The best attack strategy is the solution to the following optimization problem.

$$\begin{aligned} \max_{\mathbf{\Sigma}_K} \ & \left|\mathbf{P}_{K|K} + \mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right| \\ s.t. \ \ & \sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = a^2 \\ & -1 \le \rho_{b_p,b_v} \le 1 \\ & \sigma_{b_p}, \sigma_{b_v} > 0 \end{aligned} \tag{102}$$

where $\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T = \mathbf{A}_K$, and

$$\mathbf{\Sigma}_K = \begin{bmatrix} \sigma_{b_p}^2 & \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} \\ \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} & \sigma_{b_v}^2 \end{bmatrix} \tag{103}$$

Using the properties of the determinant, we get the formula as follows.

$$\begin{aligned} &|\mathbf{P}_{K|K} + \mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T| \\ &= |\mathbf{P}_{K|K}||\mathbf{I}_n + \mathbf{\Sigma}_K\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K| \end{aligned} \tag{104}$$

Since $\mathbf{P}_{K|K}$ is independent of $\mathbf{\Sigma}_K$, the optimization problem can be further written as:

$$\begin{aligned} \max_{\mathbf{\Sigma}_K} \ & \left|\mathbf{I}_n + \mathbf{\Sigma}_K\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\right| \\ s.t. \ \ & \sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = a^2 \\ & -1 \le \rho_{b_p,b_v} \le 1 \\ & \sigma_{b_p}, \sigma_{b_v} > 0 \end{aligned} \tag{105}$$

By defining

$$\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K = \begin{bmatrix} m_1 & m_2 \\ m_2 & m_3 \end{bmatrix} \tag{106}$$

and after simplifying (105), the objective function becomes

$$\begin{aligned} &\left| \mathbf{I}_n + \mathbf{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \right| \\ &= 1 + (1 - \rho_{b_p,b_v}^2)\sigma_{b_p}^2 \sigma_{b_v}^2 (m_1 m_3 - m_2^2) \\ &+ \sigma_{b_p}^2 m_1 + \sigma_{b_v}^2 m_3 + 2\rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} m_2 \end{aligned} \tag{107}$$

The solution to the optimization problem will be the best strategy to attack the system. In order to get closed-form solutions, we denote $\mathbf{\Sigma}_K = \mathbf{R}^T\mathbf{R}$ where $\mathbf{\Sigma}_K$ is invertible, and

$$\begin{aligned} & \left| \mathbf{I}_n + \mathbf{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \right| \\ = \quad & \left| \mathbf{I}_n + \mathbf{R}^T \mathbf{R} \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \right| \\ = \quad & \left| \mathbf{I}_n + \mathbf{R} \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T \right| \end{aligned} \tag{108}$$

To facilitate the derivations, we introduce two lemmas from [19], which are provided below.

Lemma 1. *Suppose* $\mathbf{A}$ *and* $\mathbf{B}$ *are* $n \times n$ *positive semidefinite matrices with eigendecomposition* $\mathbf{A} = \mathbf{D}_A \mathbf{\Sigma}_A \mathbf{D}_A^T$ *and* $\mathbf{B} = \mathbf{D}_B \mathbf{\Sigma}_B \mathbf{D}_B^T$*, the eigenvalues of* $\mathbf{A}$ *and* $\mathbf{B}$ *satisfy that* $\alpha_1 \geq \alpha \geq \cdots \alpha_n$ *and* $\beta_1 \geq \beta \geq \cdots \beta_n$*, then*

$$\Pi_{i=1}^n (\alpha_i + \beta_i) \leq \det(\mathbf{A} + \mathbf{B}) \leq \Pi_{i=1}^n (\alpha_i + \beta_{n+1-i}) \tag{109}$$

where the upper bound is achieved if and only if $\mathbf{D}_A = \mathbf{D}_B \mathbf{P}$*, and the lower bound is achieved if and only if* $\mathbf{D}_A = \mathbf{D}_B$*, where* $\mathbf{P}$ *is defined below,*

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & \cdots & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \tag{110}$$

The optimal solution achieving the upper bound is the best attack strategy with the most effect on the Kalman filter system, and that achieving the lower bound is the strategy with the least effect on the Kalman filter system.

Lemma 2. *Given an $n \times n$ matrix $\mathbf{V}_1$ and an $n \times n$ positive semidefinite matrix $\mathbf{Q}_1$ with $\mathbf{V}_1\mathbf{Q}_1\mathbf{V}_1^T$ being a diagonal matrix with its diagonal elements in increasing order, it is always possible to find another $n \times n$ matrix $\bar{\mathbf{V}}_1$ such that $\bar{\mathbf{V}}_1\mathbf{Q}_1\bar{\mathbf{V}}_1^T = \beta\mathbf{V}_1\mathbf{Q}_1\mathbf{V}_1^T$ with $Tr(\mathbf{V}_1\mathbf{V}_1^T) = Tr(\bar{\mathbf{V}}_1\bar{\mathbf{V}}_1^T)$ where $\beta \geq 1$. $\bar{\mathbf{V}}_1$ can be written as $\mathbf{\Sigma}_Q\mathbf{D}_1^T$, where $\mathbf{D}_1$ is the unitary matrix whose columns are the eigenvectors corresponding to the eigenvalues of $\mathbf{Q}_1$ in increasing order, and $\mathbf{\Sigma}_Q$ is a diagonal matrix.*

By combining the two lemmas together, we can get the final optimal solution to the optimization problem above. It is obvious that $\mathbf{I}_n$ and $\mathbf{R}\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}^T$ are both positive semidefinite matrices. The eigendecomposition of the two items above can be written as follows,

$$\begin{aligned} \mathbf{I}_n &= \mathbf{D}_I\mathbf{\Sigma}_1\mathbf{D}_I^T \\ \mathbf{R}\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}^T &= \mathbf{D}_2\mathbf{\Sigma}_2\mathbf{D}_2^T \end{aligned} \tag{111}$$

with identity matrix $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2 = diag([\sigma_{2,1}, \cdots, \sigma_{2,n}])$. Based on Lemma 1, we have

$$\left|\mathbf{I}_n + \mathbf{R}\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}^T\right| \leq \Pi_{i=1}^n(\sigma_{1,i} + 1) \tag{112}$$

where $\mathbf{D}_2\mathbf{P} = \mathbf{D}_1$, and

$$\begin{aligned} &|\mathbf{I}_n + \mathbf{R}\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}^T| \\ &= |\mathbf{D}_1^T||\mathbf{I}_n + \mathbf{R}\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}^T||\mathbf{D}_1| \\ &= |\mathbf{I}_n + \mathbf{D}_1^T\mathbf{R}\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}^T\mathbf{D}_1| \end{aligned} \tag{113}$$

Set $\mathbf{R}_1 = \mathbf{D}_1^T\mathbf{R}$ and $\mathbf{\Sigma}_3 = \mathbf{P}\mathbf{\Sigma}_2\mathbf{P}^T$ with the eigenvalues in increasing order and we know that $Tr(\mathbf{R}\mathbf{R}^T) = Tr(\mathbf{R}_1\mathbf{R}_1^T)$. So the optimization problem can be written as below,

$$\begin{aligned} \max \quad & |\mathbf{I}_n + \mathbf{R}_1\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}_1^T| \\ s.t. \quad & Tr(\mathbf{R}_1\mathbf{R}_1^T) \leq a^2 \\ & \mathbf{R}_1\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K\mathbf{R}_1^T = \mathbf{\Sigma}_3 \end{aligned} \tag{114}$$

Let us set $\mathbf{W}_K^T\mathbf{P}_{K|K}^{-1}\mathbf{W}_K = \tilde{\mathbf{Q}}$. Based on Lemma 2, $\mathbf{R}_1\tilde{\mathbf{Q}}\mathbf{R}_1^T = \mathbf{\Sigma}_3$, we can surely find a matrix $\bar{\mathbf{R}}$ such that $\bar{\mathbf{R}}_1\tilde{\mathbf{Q}}\bar{\mathbf{R}}_1^T = \beta\mathbf{R}_1\tilde{\mathbf{Q}}\mathbf{R}_1^T$, for $\beta \geq 1$. Note that the determinant is a monotonic increasing function of the positive semidefinite matrix. So

$$|\mathbf{I}_n + \mathbf{R}_1\tilde{\mathbf{Q}}\mathbf{R}_1^T| \leq |\mathbf{I}_n + \bar{\mathbf{R}}_1\tilde{\mathbf{Q}}\bar{\mathbf{R}}_1^T| \tag{115}$$

So the optimal solution $\bar{\mathbf{R}}$ should be in the form of $\bar{\mathbf{V}}$. The eigendecompostion of $\tilde{\mathbf{Q}}$ is as follows,

$$\tilde{\mathbf{Q}} = \mathbf{V}_Q \mathbf{\Sigma}_Q \mathbf{V}_Q^T \tag{116}$$

where $\mathbf{\Sigma}_Q = diag([\sigma_{q,1}, \sigma_{q,2}, \cdots, \sigma_{q,n}])$ with its diagonal elements in increasing order. $\mathbf{V}_Q$ is a unitary matrix whose column vectors correspond to the eigenvalues of $\tilde{\mathbf{Q}}$. The problem can be written as

$$\begin{aligned} \max_{\sigma_{b,i}^2} \quad & \sum_{i=1}^{n} log(\sigma_{b,i}^2 \sigma_{q,i} + 1) \\ s.t. \quad & \sum_{i=1}^{n} (\sigma_{b,i}^2) \leq a^2 \end{aligned} \tag{117}$$

The objective function above is a concave and increasing function. The optimal solution is achieved through Lagrangian multipliers yielding the water-filling strategy,

$$\sigma_{b,i}^2 = \left(\frac{1}{\lambda} - \frac{1}{\sigma_{q,i}} \right)^+ \tag{118}$$

where the value of λ can be obtained by solving

$$\sum_{i=1}^{n} \left(\frac{1}{\lambda} - \frac{1}{\sigma_{q,i}} \right)^+ = a^2 \tag{119}$$

The solution is

$$\mathbf{R}^{opt} = \mathbf{D}_1 [\mathbf{\Sigma}_b^{1/2}]^T \mathbf{V}_Q^T \tag{120}$$

Finally, the optimal solution of (105) is,

$$\mathbf{\Sigma}_K = \mathbf{V}_Q \mathbf{\Sigma}_b \mathbf{V}_Q^T \tag{121}$$

Theorem 3. *For a system with one sensor that provides position and velocity measurements, the optimal strategy for the adversary to attack the system in order to maximize the determinant of mean squared error matrix is*

$$\mathbf{\Sigma}_K = \mathbf{V}_Q \mathbf{\Sigma}_b \mathbf{V}_Q^T \tag{122}$$

where $\mathbf{\Sigma}_b = diag([\sigma_{b,1}^2, \sigma_{b,2}^2, \cdots, \sigma_{b,n}^2])$, $\sigma_{b,i}^2 = \left(\frac{1}{\lambda} - \frac{1}{\sigma_{q,i}} \right)^+$ *and* $\sum_{i=1}^{n} \left(\frac{1}{\lambda} - \frac{1}{\sigma_{q,i}} \right)^+ = a^2$, $\mathbf{V}_Q$ *is a unitary matrix whose column vectors corresponds to the eigenvalues* $\sigma_{q,i}$*'s of* $\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K$, $\sigma_{q,i}, i \in \{1, n\}$ *is in increasing order.*

3) Attack Strategy for Multiple Position and Velocity Sensors: For a system with multiple sensors, the best strategy to allocate the bias noise power and set the correlation coefficients among the bias noises at different sensors is also investigated. Let us denote the number of sensors as M, and the measurement matrix as $\mathbf{H} = [\mathbf{I}_2, \cdots, \mathbf{I}_2]^T$. The measurement covariance matrix for the ith sensor is assumed to be

$$\mathbf{R}_i = \begin{bmatrix} \sigma_{p_i}^2 & 0 \\ 0 & \sigma_{v_i}^2 \end{bmatrix} \tag{123}$$

Now, according to (53), we have

$$\begin{aligned} \mathbf{R}_e &= \left(\sum_{i=1}^{M} \mathbf{R}_i^{-1} \right)^{-1} \\ &= \begin{bmatrix} \left(\sum_{i=1}^{M} \sigma_{p_i}^{-2} \right)^{-1} & 0 \\ 0 & \left(\sum_{i=1}^{M} \sigma_{v_i}^{-2} \right)^{-1} \end{bmatrix} \end{aligned} \tag{124}$$

According to (51), we define

$$\begin{aligned} \mathbf{C}_i &= \mathbf{R}_e \mathbf{R}_i^{-1} \\ &= \begin{bmatrix} \frac{\sigma_{p_i}^{-2}}{\sum_{j=1}^{M} \sigma_{p_j}^{-2}} & 0 \\ 0 & \frac{\sigma_{v_i}^{-2}}{\sum_{j=1}^{M} \sigma_{v_j}^{-2}} \end{bmatrix} \end{aligned} \tag{125}$$

as the weighting matrix for the ith sensor's observation $\mathbf{z}_i$.

The equivalent injected bias noise is therefore

$$\mathbf{b}_{eK} = \sum_{i=1}^{M} \mathbf{C}_i \mathbf{b}_{K_i} \tag{126}$$

and the covariance matrix of the equivalent bias vector is

$$\mathbf{\Sigma}_{eK} = \sum_{i=1}^{M} \sum_{j=1}^{M} \mathbf{C}_i E\left[\mathbf{b}_i \mathbf{b}_j^T \right] \mathbf{C}_j^T \tag{127}$$

Now the optimization problem can be formulated as follows.

$$\max_{\mathbf{\Sigma}_{eK}} \left| \mathbf{P}_{K|K} + \mathbf{W}_{eK} \mathbf{\Sigma}_{eK} \mathbf{W}_{eK}^T \right| \tag{128}$$

$$\begin{aligned} s.t. \quad & \sum_{i=1}^{N} \sigma_{b_{p_i}}^2 + T^2 \sum_{j=1}^{N} \sigma_{b_{v_i}}^2 = a^2, \\ & -1 \leq \rho_{b_{p_i}, b_{v_j}} \leq 1, \\ & -1 \leq \rho_{b_{v_i}, b_{v_j}} \leq 1, \\ & -1 \leq \rho_{b_{p_i}, b_{p_j}} \leq 1, \\ & \sigma_{b_{p_i}}, \sigma_{b_{v_i}} \geq 0, \quad \forall i, j \in \{1, M\} \end{aligned}$$

where $\mathbf{W}_{eK}$ is the Kalman filter gain calculated using $\mathbf{H}_e$ and $\mathbf{R}_e$. The optimal solution of (128) can be obtained numerically as shown later in Chapter VIII of the thesis.

VIII. Numerical Results

Numerical results are presented in this chapter to demonstrate the effectiveness of the derived attack strategies.

A. Systems with Position Sensors

The parameters used in the target tracking example are provided below. The system sampling interval is $T = 1$. The adversary injects bias information to two sensors with $\sigma_{w_1}^2 = 3$ and $\sigma_{w_2}^2 = 4$, respectively. The variance of the system process noise is $\sigma_v^2 = 0.25$. The biases b_is are zero-mean Gaussian random variables with variances $\sigma_{b_i}^2$s. For the power constraint we discussed earlier, we set the sum of $\sigma_{b_i}^2$ to be 3000.

The effect of the bias injection on the Kalman filter is measured by the normalized MSE. More specifically, we use the sum of the normalized MSE over N_m Monte-Carlo runs

$$q_k = \sum_{j=1}^{N_m} \left[\hat{\mathbf{x}}_{k|k}'^j - \mathbf{x}_k^j\right]^T \mathbf{P}_{k|k}^{-1} \left[\hat{\mathbf{x}}_{k|k}'^j - \mathbf{x}_k^j\right] \tag{129}$$

where at time k, $\mathbf{P}_{k|k}$ is the nominal state covariance matrix calculated by the Kalman filter, $\hat{\mathbf{x}}_{k|k}'^j$ is the state estimate, and $\mathbf{x}_k^j$ is the true state, during the jth Monte-Carlo run. First, if the random biases injected to different sensors are independent, we should allocate all the bias power to the sensor with the smallest measurement noise variance. This is clearly true as demonstrated in Fig. 2, where allocating all the power to sensor 1 causes the maximum mean squared estimation error. In Fig. 3, three dependent-noise attack strategies are compared, including the optimal one according to (67), allocating the power equally among the sensors, and allocating all the power to the sensor with the smallest measurement error variance. It is clear that the optimal solution has the largest impact on the estimation performance, and it outperforms the best independent-noise attack strategy significantly.

B. Systems with Position and Velocity Sensors

We now consider the case where the adversary attacks the Kalman filtering system with a vector sensor observation containing both position and velocity measurements. We first consider a single-sensor system, and the sensor has a position measurement variance of 3 and a velocity measurement variance of 4. We set the sum of $\sigma_{b_{p_1}}^2$ and $T^2\sigma_{b_{v_1}}^2$ to be 3000. In this particular

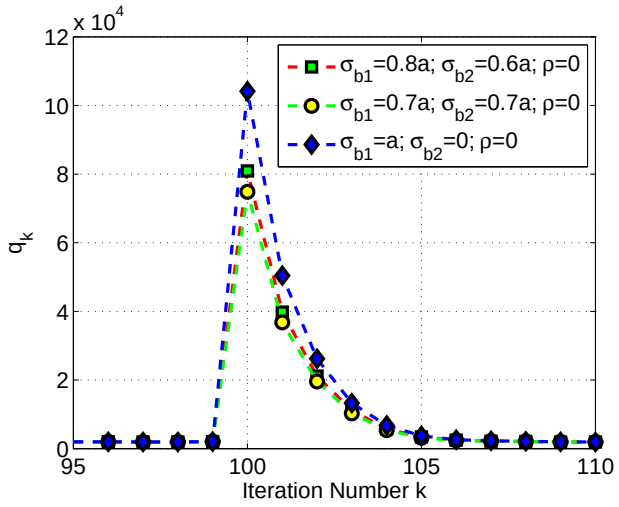

Fig. 2. The normalized MSE when independent biases are used. $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2$ for each case.

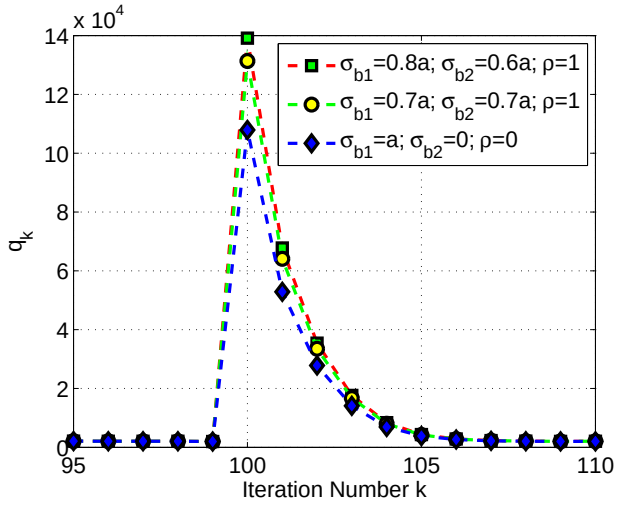

Fig. 3. The normalized MSE for dependent biases. $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2$ for each case.

case, $w_{11}w_{12} + w_{21}w_{22} > 0$, so the optimal choice is $\rho_{b_p,b_v} = 1$. Based on Theorem 2, the best strategy is to set $\sigma_{b_p} = 52.3$ and $\sigma_{b_v} = 16.2$. It is clear from Fig. 4 that the strategy provided in Theorem 2 maximizes the MSE of the Kalman filter system by injecting vector bias information.

Next we consider a system with two sensors. The first sensor is the same as the one described above, and the second one is with position measurement variance 4 and velocity measurement

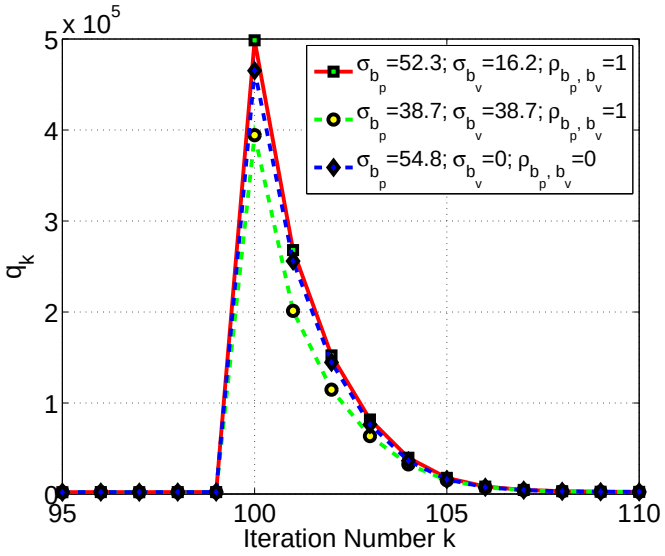

Fig. 4. The normalized MSE for a system with a single sensor. $\sigma_{p_1}^2 + T^2\sigma_{v_1}^2 = a^2$ for each case.

variance 5. In this particular case, again we have $w_{11}w_{12} + w_{21}w_{22} > 0$, so all the ρs in s_1, s_2, and s_3 should be set as 1. We first use a systematic grid search to find an approximate globally optimal solution and then we use the FMINCON function in Matlab, a local search algorithm, to refine this approximate globally optimal solution. The optimal solution we have obtained is $\sigma_{b_{p_1}}^2 = 1826, \sigma_{b_{p_2}}^2 = 1023, \sigma_{b_{v_1}}^2 = 81, \sigma_{b_{v_2}}^2 = 68$. For comparison purpose, we also implement an attack strategy that allocate power equally among the observation components and among the two sensors, which is $\sigma_{b_{p_1}}^2 = \sigma_{b_{p_2}}^2 = \sigma_{b_{v_1}}^2 = \sigma_{b_{v_2}}^2 = 750$. The simulation result is shown in Fig. 5. As we can see, the optimal attack strategy has a much greater impact than the one that allocates power equally. Based on the optimal solution, we can find that allocating more power to the measurement with lower variance will have a greater effect on the Kalman filter system.

C. Determinant Case

Numerical results are presented in this section to illustrate the effectiveness of the proposed attack strategies. Assuming that the injected bias noise $\mathbf{b}_k$ is zero-mean and Gaussian distributed, we can show that the posterior probability density function (PDF) of the target state conditioned on the past observations and the current corrupted observation is

$$p(\mathbf{x}_K|\mathbf{z}_{1:K-1}, \mathbf{z}'_K) = \mathcal{N}(\hat{\mathbf{x}}_{K|K}, \mathbf{P}_{K|K} + \mathbf{A}_K) \tag{130}$$

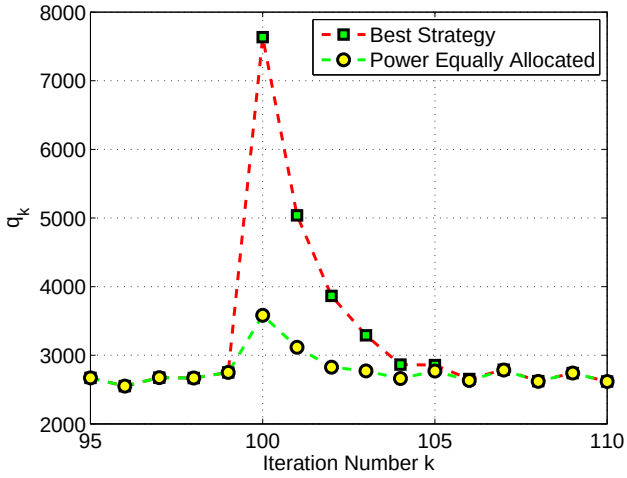

Fig. 5. The normalized MSE for a system with two sensors. $\sigma_{p_1}^2 + \sigma_{p_2}^2 + T^2\sigma_{v_1}^2 + T^2\sigma_{v_2}^2 = a^2$ for each case.

where $\hat{\mathbf{x}}_{K|K}$ is the updated state estimate calculated by the Kalman filter, which is unaware of the presence of the injected false information. Then the target state $\mathbf{x}_K$ will be in the following confidence region (or error ellipse)

$$\left\{\mathbf{x} : (\mathbf{x} - \hat{\mathbf{x}}_{K|K})^T (\mathbf{P}_{K|K} + \mathbf{A}_K)^{-1} (\mathbf{x} - \hat{\mathbf{x}}_{K|K}) \leq \gamma\right\} \tag{131}$$

with a probability that is determined by the threshold γ [20]. The volume of the confidence region defined by (131) corresponding to the threshold γ is

$$V(K) = c_{n_x} |\gamma(\mathbf{P}_{K|K} + \mathbf{A}_K)|^{1/2} \tag{132}$$

where n_x is the dimension of the target state $\mathbf{x}$,

$$c_n = \frac{\pi^{n/2}}{\Gamma(n/2+1)} \tag{133}$$

and $\Gamma(\cdot)$ is the gamma function. First, let us consider a single-sensor case, where the sensor has a position measurement with noise variance of 3, which is independent of the velocity measurement with noise variance of 4. We set the bias noise power constraint as $\sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = 3000$. Based on Theorem 3, we have $\sigma_{b_p}^2 = 1500, \sigma_{b_v}^2 = 1500$, and $\rho_{b_{p,v}} = 0.063$, which is the same as the solution obtained numerically. In Fig. 6, error ellipses for different attack strategies are plotted. For all the different attack strategies, we set $\rho_{b_{p,v}} = 0.063$. As we can see, under normal condition without false information injection, the error ellipse has the smallest area, while the optimal

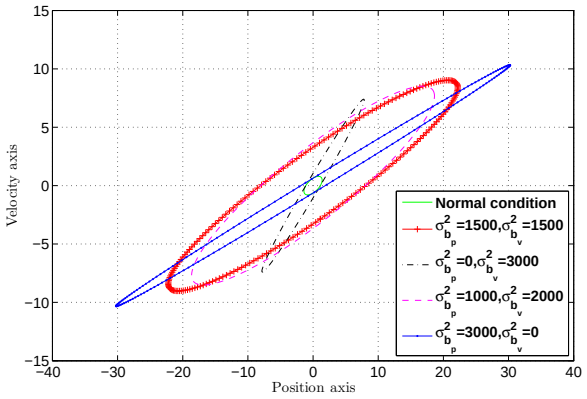

Fig. 6. Error ellipses for different power allocation strategies

attack strategy leads to an error ellipse with the largest area. In Figs. 7 and 8, the volume (area) of the error ellipse is provided as a function of $\rho_{b_{p,v}}$ and the ratio $\kappa = \frac{\sigma_{b_p}}{\sigma_{b_v} T}$. We can see that when the $\kappa = \frac{\sigma_{b_p}}{\sigma_{b_v} T} = 1$, the area of the ellipse is maximized. Also from Figs. 7 and 8, it is clear that the area of ellipse increases as the absolute value of ρ decreases. In Fig. 9, the trend of the error ellipses as the ρ changes from -1 to $+1$ is illustrated.

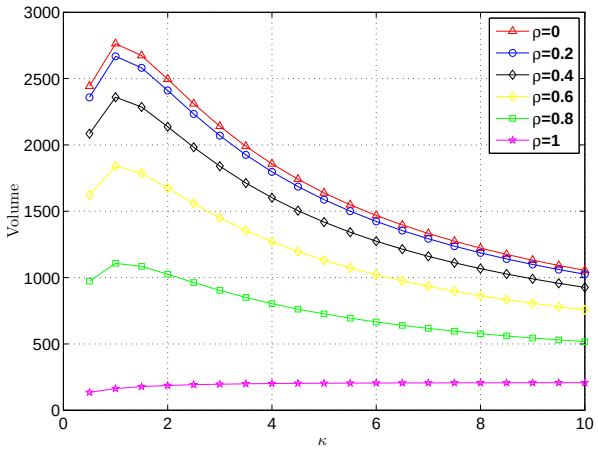

Fig. 7. Error ellipses volume

In this particular case, since $\sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = 3000$, $\mathbf{\Sigma}_K$ is large and in (102) the second term $(\mathbf{W}_K \mathbf{\Sigma}_K \mathbf{W}_K^T)$ dominates. Therefore, in (107) the identity matrix in the objective function is

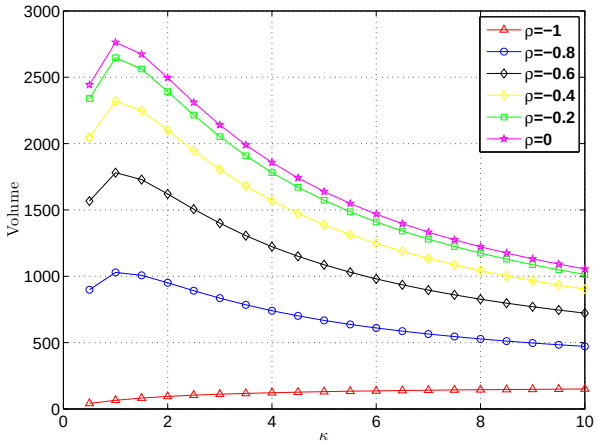

Fig. 8. Error ellipses volume

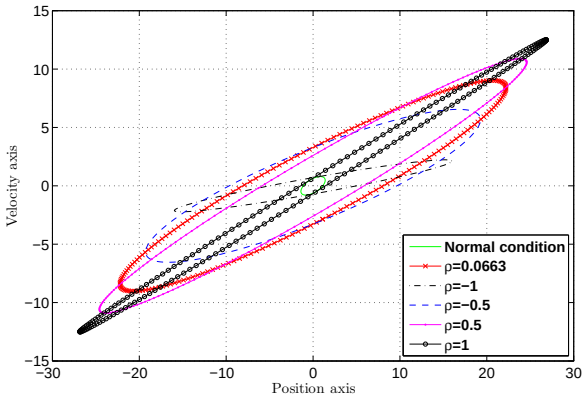

Fig. 9. Error ellipses for different ρs

relatively small comparing to the second term, and approximately we have

$$\begin{aligned} &\left|\mathbf{I}_n + \mathbf{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K\right| \\ &\approx |\mathbf{\Sigma}_K| \left|\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K\right| \end{aligned} \tag{134}$$

The second term in the second line of the above equation is a constant. Hence, in order to get the maximum determinant, we should set $\sigma_{b_p}^2 = \sigma_{b_v}^2 T^2$ and $\rho_{b_p,b_v} = 0$. This is almost the same solution as we have obtained using the optimal water-filling strategy. Next we consider a system with two sensors. The first sensor is the same as the one described above, and the

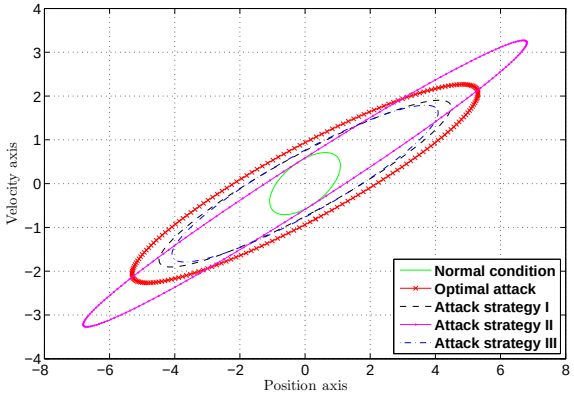

Fig. 10. Error ellipses for different power allocation strategies

second one is with position measurement variance 4 and velocity measurement variance 5. To solve the optimization problem formulated in (128), we first use a systematic grid search to find an approximate globally optimal solution and then we use the FMINCON function in Matlab, a local search algorithm, to refine this approximate globally optimal solution. The optimal solution we have obtained is $\sigma^2_{b_{p_1}} = 1100,\ \sigma^2_{b_{p_2}} = 600,\ \sigma^2_{b_{v_1}} = 750,\ \sigma^2_{b_{v_2}} = 550,$ $\rho_{b_{p_1,p_2}} = 0.99,\ \rho_{b_{p_1,v_1}} = -0.83,\ \rho_{b_{p_1,v_2}} = 0.75,\ \rho_{b_{v_1,p_2}} = 0.89,\ \rho_{b_{p_2,v_2}} = -0.23,\ \rho_{b_{v_1,v_2}} = 0.95.$ For comparison purpose, we introduce three sub-optimal attack strategies: Strategy I with all the ρs being 0s, and $\sigma^2_{b_{p_1}} = 1100, \sigma^2_{b_{p_2}} = 600, \sigma^2_{b_{v_1}} = 750, \sigma^2_{b_{v_2}} = 550$; Strategy II with all the ρs being 1s, and $\sigma^2_{b_{p_1}} = 1100, \sigma^2_{b_{p_2}} = 600, \sigma^2_{b_{v_1}} = 750, \sigma^2_{b_{v_2}} = 550$; and Strategy III with the ρs being the same as those for the optimal strategy, and $\sigma^2_{b_{p_1}} = \sigma^2_{b_{p_2}} = \sigma^2_{b_{v_1}} = \sigma^2_{b_{v_2}} = 750$. The numerical results are shown in Fig. 10. As we can see, the optimal attack strategy has a greater impact than those sub-optimal attack strategies, resulting in the largest error ellipse.

IX. Conclusion

In this thesis, the impact of false information injection on the Kalman filter's state estimation was studied. We derived the EMSE due to the injected random biases for a Kalman filter in a linear dynamic system. This allows us to find how to allocate the bias power among multiple sensors in order to maximize the effect of the false information on the Kalman filter from two perspectives: trace and determinant of the MSE matrix. By using the equivalent sensor to denote the multiple sensors in the Kalman filter system, the analysis of optimization problem is simplified a lot. A concrete example of multi-sensor target tracking system has been provided. In this example, we investigated both the case where the sensors provide position measurements and the case where they collect both position and velocity measurements. For the case where the sensors provide only position measurements, we have found that the dependent false information will incur more Kalman filter system state estimation error than the independent one does. From the trace and determinant perspectives, many closed-form results have been provided for the optimal attack strategies. In the future, we will use game theory and hypothesis testing techniques to characterize the model in order to have a better understanding of the false information attacks and Kalman filter's defense against such attacks.

REFERENCES

[1] Y. Liu, M.K. Reiter, and P. Ning, "False data injection attacks agianst state estimation in electric power grids," in *Proc. the 16th ACM Conference on Computer and Communications Security*, Chicago, IL, November 2009.

[2] L. Jia, R.J. Thomas, and L. Tong, "Malicious data attack on real-time electricity market," in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, Prague, Czech Republic, May 2011, pp. 5952–5955.

[3] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious Data Attack on Smart Grid State Estimation: Attack Strategies and Countermeasures," in *Proc. First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Gaithersburg, MD, Oct. 2010, pp. 220–225.

[4] L. Jia, R. J. Thomas, and L. Tong, "On the nonlinearity effects on malicious data attack on power system," in *Proc. Power and Energy Society General Meeting*, San Diego, CA, July 2012, pp. 1–8.

[5] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *Proc. Global Communications Conference*, San Diego, CA, Dec. 2012, pp. 3153–3158.

[6] J. Kim, L. Tong, and R. J. Thomas, "Data framing attack on state estimation," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1637–1653, July 2013.

[7] J.Kim, L. Tong, and R.J. Thomas, "Data framing attack on state estimation with unknown network parameters," in *Proc. of the Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov 2013, pp. 1388 – 1392.

[8] J.Kim, L. Tong, and R.J. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," *IEEE Trans. on Signal Processing*, vol. 63, no. 5, pp. 1102 – 1114, Dec 2014.

[9] X. Song, P. Willett, S. Zhou, and P. B. Luh, "The mimo radar and jammer games," *IEEE Trans. on Signal Processing*, vol. 60, no. 2, pp. 687–699, February 2012.

[10] C. Yang, L. Kaplan, and E. Blasch, "Performance measures of covariance and information matrices in resource management for target state estimation," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2594–2612, July 2012.

[11] C. Yang, L. Kaplan, E. Blasch, and M. Bakich, "Optimal placement of heterogeneous sensors for targets with gaussian priors," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1637–1653, July 2013.

[12] E. Masazade, R. Niu, P.K. Varshney, and M. Keskinoz, "Energy aware iterative source localization for wireless sensor networks," *IEEE Trans. on Signal Processing*, vol. 58, no. 9, pp. 4824–4835, September 2010.

[13] E. Masazade, R. Niu, and P.K. Varshney, "Dynamic bit allocation for object tracking in wireless sensor networks," *IEEE Trans. on Signal Processing*, vol. 60, no. 10, pp. 5048–5063, October 2012.

[14] X. Lin and Y. Bar-Shalom, "Multisensor target tracking performance with bias compensation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 3, pp. 1139–1149, July 2006.

[15] R. Niu and L. Huie, "System State Estimation in the Presence of False Information Injection," in *Statistical Signal Processing Workshop (SSP)*, Ann Arbor, MI, Aug. 2012, pp. 385–388.

[16] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley, New York, 2001.

[17] R. Niu, "Dynamic System State Estimation in the Presence of Continuous False Information Injection," Tech. Rep., Extension Grant from Visiting Faculty Research Program, Air Force Research Laboratory Information Directorate, March 2012.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, Cambridge, U.K., 2004.

[19] B. Tang, J. Tang, and Y. Peng, "Mimo radar waveform design in colored noise based on information theory," *IEEE Trans. on Signal Processing*, vol. 58, no. 9, pp. 4684 – 4697, September 2010.

[20] Y. Bar-Shalom, P.K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*, YBS Publishing, Storrs, CT, 2011.