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

Jingyang Lu and Ruixin Niu
Department of Electrical and Computer Engineering
Virginia Commonwealth University
Richmond, VA 23284, U.S.A.
{luj2, rniu}@vcu.edu

***Abstract*— In this paper, the impact of false information injection is investigated for linear dynamic systems with multiple sensors. It is assumed that the 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. 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 multi-sensor target tracking system with either position sensors or position and velocity sensors has been used as an example to illustrate the theoretical results.**

I. Introduction

System state estimation in the presence of adversary that injects false information into sensor readings is an important problem with wide application areas, such as target tracking with compromised sensors, and secure monitoring of dynamic electric power systems. This topic has attracted considerable attention and interest recently [1]–[4]. In [1], the authors showed the impact of malicious attacks on real-time electricity market and that 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 close relationship between these attackers and the control center was discussed in [2], 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 [3] and [4]. However, in the aforementioned publications, only the problem of *static* system state estimation has been considered.

In this paper, 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 received little attention in literature. Some related publications exist, where the problem of sensor bias estimation and compensation for target tracking has been addressed. Interested readers are referred to [5] and references therein for details. In our previous work [6], we have investigated the impact of the injected biases on a Kalman filter's estimation performance. We have shown that if the false data are injected at a single time, the impact of the false information converges to zero as time goes on; if the false data are injected into the system continuously, the estimation error tends to reach a steady state. Based on our previous work in [6], in this paper our goal is to find the optimal attack strategy for the adversary, which maximizes the effect of the false information injection on Kalman filter's estimation performance. We mathematically derive the Kalman filter's mean squared state estimation error and maximize it under a constraint on the sensor biases' total power. To be concrete, we provide a multi-sensor target tracking example, where both the case with position sensors and that with both position and velocity sensors are investigated. For various scenarios, closed-form solutions are derived for the optimal attack strategies. Our results show that coordinated attacks (dependent attacks) always outperform independent attacks, and allocating bias noise power optimally will significantly increase the the Kalman filter's estimation errors.

II. System Model

The discrete-time linear dynamic system 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 $E[\mathbf{v}_k \mathbf{v}_k^T] = \mathbf{Q}_k$. Let us assume that M sensors are used by the linear system. The measurement at time k collected by sensor i is

$$\mathbf{z}_{k,i} = \mathbf{H}_{k,i} \mathbf{x}_{k,i} + \mathbf{w}_{k,i} \tag{2}$$

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 paper, 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 (2) becomes

$$\mathbf{z}'_{k,i} = \mathbf{H}_{k,i} \mathbf{x}_k + \mathbf{w}_{k,i} + \mathbf{b}_{k,i} = \mathbf{z}_{k,i} + \mathbf{b}_{k,i} \tag{3}$$

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

$$\mathbf{H}_k = [\mathbf{H}_{k1}^T, \cdots, \mathbf{H}_{kM}^T]^T \tag{4}$$

With these notations, it is easy to convert (2) and (3) into the following equations respectively.

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k \tag{5}$$

and

$$\mathbf{z}'_k = \mathbf{z}_k + \mathbf{b}_k \tag{6}$$

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

$$\mathbf{R}_k = \begin{bmatrix} \mathbf{R}_{k,1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{R}_{k,M} \end{bmatrix} \tag{7}$$

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

III. Impact of False Information Injection

In this paper, 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 [7] and provided as follows.

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

$$\begin{aligned} &\hat{\mathbf{x}}'_{K+N|K+N} - \mathbf{x}_{K+N} = \hat{\mathbf{x}}_{K+N|K+N} - \mathbf{x}_{K+N} \\ &+ \sum_{m=0}^{N} \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \mathbf{b}_{K+N-m} \end{aligned} \tag{8}$$

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,

$$\mathbf{B}_K \triangleq (\mathbf{I} - \mathbf{W}_K \mathbf{H}_K) \mathbf{F}_{K-1}, \tag{9}$$

$\mathbf{I}$ is the identity matrix, and $\mathbf{W}_K$ is the Kalman filter gain [8] 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

$$\sum_{m=0}^{N} \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \mathbf{b}_{K+N-m}, \tag{10}$$

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*

$$\mathbf{A}_{K+N} = \sum_{m=0}^{N} \mathbf{D}_m \mathbf{\Sigma}_{K+N-m} \mathbf{D}_m^T \tag{11}$$

where

$$\mathbf{D}_m = \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \tag{12}$$

$\prod_{i=0}^{-1} \mathbf{B}_{K+N-i} = \mathbf{I}$ 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

$$\mathbf{a}_m = \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i} \right) \mathbf{W}_{K+N-m} \mathbf{b}_{k+N-m} \tag{13}$$

From (8), 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

$$\begin{aligned} E\left(\sum_{m=0}^{N} \sum_{n=0}^{N} \mathbf{a}_m \mathbf{a}_n^T \right) &= E\left(\sum_{m=0}^{N} \mathbf{a}_m \mathbf{a}_m^T \right) \\ &= \sum_{m=0}^{N} \mathbf{D}_m \mathbf{\Sigma}_{K+N-m} \mathbf{D}_m^T \end{aligned} \tag{14}$$

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

IV. The Optimal Attack Strategy

1) Problem Formulation for a General Linear System : In this paper, 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

$$\begin{aligned} &\max_{\mathbf{\Sigma}_K} \mathrm{Tr}\left[\mathbf{P}_{K|K} + \mathbf{A}_K(\mathbf{\Sigma}_K)\right] \\ &s.t. \quad \mathrm{Tr}(\mathbf{\Sigma}_K) = a^2 \end{aligned} \tag{15}$$

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 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.

$$\begin{aligned} &\max_{\mathbf{\Sigma}_K} \left|\mathbf{P}_{K|K} + \mathbf{A}_K(\mathbf{\Sigma}_K)\right| \\ &s.t. \quad \mathrm{Tr}(\mathbf{\Sigma}_K) = a^2 \end{aligned} \tag{16}$$

2) 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 paper much simpler. In a information filter recursion [8], which is equivalent to the Kalman filter recursion, we have

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{z}_k \tag{17}$$

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 (17) represents the new information from the new sensor data $\mathbf{z}_k$, which can be expanded by using (4) and (7) as follows.

$$\begin{aligned} &\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{z}_k \\ &= [\mathbf{H}_{k1}^T, \cdots, \mathbf{H}_{kM}^T] \begin{bmatrix} \mathbf{R}_{k1}^{-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{R}_{kM}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{k1} \\ \vdots \\ \mathbf{z}_{kM} \end{bmatrix} \\ &= \sum_{i=1}^{M} \mathbf{H}_{ki}^T \mathbf{R}_{ki}^{-1} \mathbf{z}_{ki} \end{aligned} \tag{18}$$

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

$$\mathbf{z}_e = \mathbf{H}_e \mathbf{x} + \mathbf{w}_e \tag{19}$$

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

$$\mathbf{H}_e^T \mathbf{R}_e^{-1} \mathbf{z}_e = \sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{z}_i \tag{20}$$

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 (20) to be true is

$$\mathbf{z}_e = \mathbf{R}_e \sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{z}_i \tag{21}$$

Taking the covariance on the both sides of (21), we get

$$\begin{aligned} \mathbf{R}_e &= \mathbf{R}_e \mathrm{cov}\left(\sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{z}_i\right) \mathbf{R}_e^T \\ &= \mathbf{R}_e \left[\sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{R}_i (\mathbf{R}_i^{-1})^T\right] \mathbf{R}_e^T \end{aligned} \tag{22}$$

This implies that

$$\mathbf{R}_e = \left(\sum_{i=1}^{M} \mathbf{R}_i^{-1}\right)^{-1} \tag{23}$$

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 (20), and following a similar procedure as in the first case, we have

$$\mathbf{z}_e = \mathbf{R}_e \sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{z}_i \tag{24}$$

and

$$\mathbf{R}_e = \left(\sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i\right)^{-1} \tag{25}$$

V. A Target Tracking Example

In this paper, 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 [8], which can still be described by the plant and measurement equations given in (1) and (2). 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

$$\mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{26}$$

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 $\mathbf{\Gamma}^T = \begin{bmatrix} T^2/2 & T \end{bmatrix}$. The covariance matrix of the process noise is therefore $\mathbf{Q} = \sigma_v^2 \mathbf{\Gamma}\mathbf{\Gamma}^T$.

In this paper, 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 Strategies for 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, $\sigma^2_{w_i}$.

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 IV-2 for the first case, namely (21) and (23), the measurement equation (3) becomes

$$z'_k = z_{ek} + b_{ek} \tag{27}$$

where

$$z_{ek} = \sum_{m=0}^{M} c_i z_{ki} \tag{28}$$

$$b_{ek} = \sum_{m=0}^{M} c_i b_{ki} \tag{29}$$

and

$$c_i = \frac{1/\sigma^2_{w_i}}{\sum_{j=1}^{M}\left(1/\sigma^2_{w_j}\right)} \tag{30}$$

which is the corresponding coefficient/weight for the ith sensor.

1) Maximizing the Trace of $\mathbf{A}_K$: In this target tracking problem, let us first consider the strategy that maximizes the trace of the Kalamn filter estimation error, which is the solution of (15) in Section IV-1. In this case,

$$\mathbf{\Sigma}_K = \begin{bmatrix} \sigma^2_{b_1} & \rho_{12}\sigma_{b_1}\sigma_{b_2} & \cdots & \rho_{1M}\sigma_{b_1}\sigma_{b_M} \\ \rho_{12}\sigma_{b_1}\sigma_{b_2} & \sigma^2_{b_2} & \cdots & \rho_{2M}\sigma_{b_2}\sigma_{b_M} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1M}\sigma_{b_1}\sigma_{b_M} & \rho_{2M}\sigma_{b_2}\sigma_{b_M} & \cdots & \sigma^2_{b_M} \end{bmatrix} \tag{31}$$

where $\sigma^2_{b_i}$ 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, (15) is equivalent to

$$\begin{aligned} &\max \mathrm{Tr}\left[\mathbf{A}_K\right] \\ &s.t. \quad \sum_{i=1}^{M} \sigma^2_{b_i} = a^2 \\ &\quad -1 \le \rho_{ij} \le 1, \ \text{ for } \ 1 \le i,j \le M \end{aligned} \tag{32}$$

To simplify this problem, we first use the equivalent measurement 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

$$\begin{aligned} \Sigma_{e_K} &= E[b^2_{e_K}] \\ &= E\left[\left(\sum_{i=1}^{M} c_i b_i\right)^2\right] \\ &= \sum_{i=1}^{M} c_i^2 \sigma^2_{b_i} + \sum_i \sum_{j\neq i} 2\rho_{ij} c_i c_j \sigma_{b_i}\sigma_{b_j} \end{aligned} \tag{33}$$

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 $1 \le i,j \le M$. 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 smallest noise variance.*

Next, let us consider the more general case where the random biases are dependent. By inspecting (33), it is clear that to maximize Σ_{e_K}, we need to set all the ρ_{ij}s to 1. As a result, (33) becomes

$$\Sigma_{e_K} = \left(\sum_{i=1}^{M} c_i \sigma_{b_i}\right)^2 \tag{34}$$

Now, the optimization problem in (32) has been converted to the following problem:

$$\begin{aligned} &\max \left(\sum_{i=1}^{M} c_i \sigma_{b_i}\right)^2 \\ &s.t. \quad \sum_{i=1}^{M} \sigma^2_{b_i} = a^2 \end{aligned} \tag{35}$$

The above problem can be solved by using standard constrained optimization techniques [9] 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.

$$\begin{aligned} \max \quad & c_1\sigma_{b_1} + c_2\sigma_{b_2} \\ \text{s.t.} \quad & \sigma^2_{b_1} + \sigma^2_{b_2} = a^2 \end{aligned} \tag{36}$$

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 the 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 $-\dfrac{c_1}{c_2}$ 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

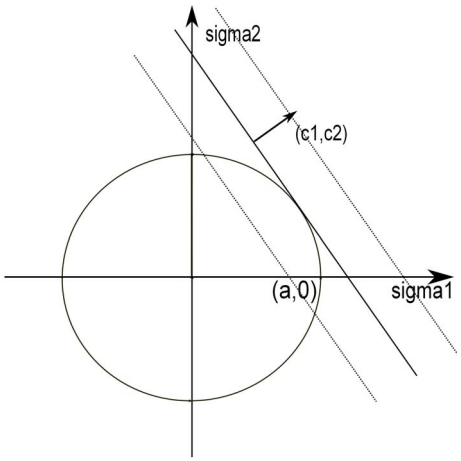

Fig. 1. Geometric solution for systems with two sensors.

solution is found when l_1 becomes a tangent line to the circle, which is

$$\begin{aligned}\sigma_1 &= \frac{c_1 a}{\sqrt{c_1^2+c_2^2}} \\ \sigma_2 &= \frac{c_2 a}{\sqrt{c_1^2+c_2^2}}\end{aligned} \tag{37}$$

For a system with arbitrary number of sensors, we can repeat the same procedure to find the optimal solution by using hyperplanes 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 dependent random noises with a pairwise correlation coefficient of* 1*. The random bias power is allocated such that*

$$\sigma_{b_i} = \frac{c_i a}{\sqrt{\sum_{j=1}^{M} c_j^2}}, \quad \text{for } i = 1, \cdots, M. \tag{38}$$

2) Maximizing the Determinant of $\mathbf{P}_{K|K} + \mathbf{A}_K$: 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 V-A1, we have

$$\begin{aligned}&|\mathbf{P}_{K|K} + \mathbf{A}_K| = |\mathbf{P}_{K|K} + \Sigma_{eK}\mathbf{D}_0\mathbf{D}_0^T| \\ &= |\mathbf{P}_{K|K}||\mathbf{I} + \Sigma_{eK}\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T|\end{aligned} \tag{39}$$

where $\mathbf{D}_0$ can be obtained using (12) and Σ_{eK} is defined in (33). 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, (39) can be written concisely as,

$$\begin{aligned}&|\mathbf{P}_{K|K}||\mathbf{C}\mathbf{I}\mathbf{C}^{-1} + \Sigma_{eK}\mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}| \\ &= |\mathbf{P}_{K|K}||\mathbf{I} + \Sigma_{eK}\mathbf{\Lambda}|\end{aligned} \tag{40}$$

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 V-A1.

B. Attack Strategies for Position and Velocity Sensors

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 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

$$\mathbf{\Sigma}_K(i,j) = \begin{bmatrix} \rho_{b_{p_i},b_{p_j}}\sigma_{b_{p_i}}\sigma_{b_{p_j}} & \rho_{b_{p_i},b_{vj}}\sigma_{b_{p_i}}\sigma_{b_{vj}} \\ \rho_{b_{vi},b_{p_j}}\sigma_{b_{vi}}\sigma_{b_{p_j}} & \rho_{b_{vi},b_{vj}}\sigma_{b_{vi}}\sigma_{b_{vj}} \end{bmatrix} \tag{41}$$

for $1 \le i,\ j \le M$. $\sigma_{b_{p_i}}$ and $\sigma_{b_{v_i}}$ 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 $1 \le i \le M$. 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

$$\sum_{i=1}^{M} \sigma_{b_{p_i}}^2 + T^2\sigma_{b_{v_i}}^2 \tag{42}$$

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 (42) 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 (42) can be easily adjusted and extended for other power definitions, as long as in the new definition, the second term is proportional to $T^2\sigma_{b_{v_i}}^2$.

1) Systems with a Single Sensor: 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:

$$\begin{aligned}&\max_{\mathbf{\Sigma}_K} \operatorname{Tr}\left[\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right] \\ &s.t. \quad \sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = a^2 \\ &\qquad -1 \le \rho_{b_p,b_v} \le 1 \\ &\qquad \sigma_{b_p}, \sigma_{b_v} > 0\end{aligned} \tag{43}$$

where

$$\mathbf{\Sigma}_K = \begin{bmatrix} \sigma_{b_p}^2 & \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} \\ \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} & \sigma_{b_v}^2 \end{bmatrix} \tag{44}$$

and

$$\mathbf{W}_K = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \tag{45}$$

It is easy to show that

$$\begin{aligned} \mathrm{Tr}\left[\mathbf{W}_K \mathbf{\Sigma}_K \mathbf{W}_K^T\right] &= \mathrm{Tr}\left[\mathbf{W}_K^T \mathbf{W}_K \mathbf{\Sigma}_K\right] \\ &= (w_{11}^2 + w_{21}^2)\sigma_{b_p}^2 + (w_{12}^2 + w_{22}^2)\sigma_{b_v}^2 \\ &+ 2(w_{11}w_{12} + w_{21}w_{22})\rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} \end{aligned} \tag{46}$$

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

$$\begin{aligned} &\max (w_{11}\sigma_{b_p} + w_{12}\sigma_{b_v})^2 + (w_{21}\sigma_{b_p} + w_{22}\sigma_{b_v})^2 \\ &s.t. \quad \sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = a^2 \\ &\quad \sigma_{b_p}, \sigma_{b_v} \geq 0 \end{aligned} \tag{47}$$

To solve this constrained optimization problem, let us first denote

$$\begin{aligned} w_{11}^2 + w_{21}^2 &= \beta_1 \\ w_{12}^2 + w_{22}^2 &= \beta_2 \\ w_{11}w_{12} &= \alpha_1 \\ w_{21}w_{22} &= \alpha_2 \end{aligned} \tag{48}$$

The constraint in (43) can be written as

$$\frac{\sigma_{b_p}^2}{T^2} + \sigma_{b_v}^2 = \frac{a^2}{T^2} = a_1^2 \tag{49}$$

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

$$\begin{aligned} &\max_\theta a_1^2 \left[\frac{\beta_1 T_1^2 + \beta_2}{2} + A\sin(2\theta + \phi)\right] \\ &s.t. \quad 0 \leq \theta \leq \frac{\pi}{2} \end{aligned} \tag{50}$$

where

$$A = \sqrt{\frac{1}{4}\left(\beta_2 - \beta_1 T^2\right)^2 + T^2(\alpha_1 + \alpha_2)^2} \tag{51}$$

$$\tan(\phi) = \frac{\beta_2 - \beta_1 T^2}{2T(\alpha_1 + \alpha_2)} \tag{52}$$

Clearly, the optimal solution is

$$\theta^* = \frac{\pi}{4} - \frac{\phi}{2} \tag{53}$$

We summarize this result 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*

$$\begin{aligned} \sigma_{b_p} &= a\sin(\theta^*) \\ \sigma_{b_v} &= \frac{a}{T}\cos(\theta^*) \\ \theta^* &= \frac{\pi}{4} - \frac{\phi}{2} \\ \phi &= \arctan\left[\frac{\beta_2 - \beta_1 T^2}{2T(\alpha_1 + \alpha_2)}\right] \\ w_{11}^2 + w_{21}^2 &= \beta_1 \\ w_{12}^2 + w_{22}^2 &= \beta_2 \\ w_{11}w_{12} &= \alpha_1 \\ w_{21}w_{22} &= \alpha_2 \end{aligned} \tag{54}$$

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 (54) remains the same.

2) System with Two Senors: In this case, $M = 2$, and 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

$$\mathbf{R}_i = \begin{bmatrix} \sigma_{p_i}^2 & 0 \\ 0 & \sigma_{v_i}^2 \end{bmatrix} \tag{55}$$

Now, according to (25), we have

$$\begin{aligned} \mathbf{R}_e &= [\mathbf{R}_1^{-1} + \mathbf{R}_2^{-1}]^{-1} \\ &= \begin{bmatrix} \left(\sigma_{p_1}^{-2} + \sigma_{p_2}^{-2}\right)^{-1} & 0 \\ 0 & \left(\sigma_{v_1}^{-2} + \sigma_{v_2}^{-2}\right)^{-1} \end{bmatrix} \end{aligned} \tag{56}$$

According to (24), we define

$$\begin{aligned} \mathbf{C}_i &= \mathbf{R}_e \mathbf{H}_i^T \mathbf{R}_i^{-1} \\ &= \begin{bmatrix} \frac{\sigma_{p_i}^{-2}}{\sigma_{p_1}^{-2} + \sigma_{p_2}^{-2}} & 0 \\ 0 & \frac{\sigma_{v_i}^{-2}}{\sigma_{v_1}^{-2} + \sigma_{v_2}^{-2}} \end{bmatrix} \end{aligned} \tag{57}$$

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

$$\begin{aligned} c_{p_i} &= \mathbf{C}_i(1,1) \\ c_{v_i} &= \mathbf{C}_i(2,2) \end{aligned} \tag{58}$$

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

$$\mathbf{b}_{eK} = \sum_{i=1}^{2} \mathbf{C}_i \mathbf{b}_{K_i} \tag{59}$$

So the covariance matrix of the equivalent bias vector is

$$\mathbf{\Sigma}_{eK} = \sum_{i=1}^{2}\sum_{i=j}^{2} \mathbf{C}_i \mathbf{\Sigma}_K(i,j) \mathbf{C}_j^T \tag{60}$$

where $\mathbf{\Sigma}_K(i,j)$ has been defined in (41). It can be shown that

$$\mathbf{\Sigma}_{eK} = \begin{bmatrix} s_1 & s_2 \\ s_2 & s_3 \end{bmatrix} \tag{61}$$

Where

$$\begin{aligned} s_1 &= c_{p_1}^2\sigma_{b_{p_1}}^2 + c_{p_2}^2\sigma_{b_{p_2}}^2 + 2\rho_{b_{p_1},b_{p_2}}c_{p_1}c_{p_2}\sigma_{b_{p_1}}\sigma_{b_{p_2}} \\ s_3 &= c_{v_1}^2\sigma_{b_{v_1}}^2 + c_{v_2}^2\sigma_{b_{v_2}}^2 + 2\rho_{b_{v_1},b_{v_2}}c_{v_1}c_{v_2}\sigma_{b_{v_1}}\sigma_{b_{v_2}} \end{aligned} \tag{62}$$

$$s_2 = c_{p_1}c_{v_1}\rho_{b_{p_1},b_{v_1}}\sigma_{b_{p_1}}\sigma_{b_{v_1}} + c_{p_1}c_{v_2}\rho_{b_{p_1},b_{v_2}}\sigma_{b_{p_1}}\sigma_{b_{v_2}} + c_{p_2}c_{v_1}\rho_{b_{p_2},b_{v_1}}\sigma_{b_{p_2}}\sigma_{b_{v_1}} + c_{p_2}c_{v_2}\rho_{b_{p_2},b_{v_2}}\sigma_{b_{p_2}}\sigma_{b_{v_2}} \quad (63)$$

The optimization problem can be written as follows.

$$\begin{aligned} \max_{\mathbf{\Sigma}_{eK}} \ & \mathrm{Tr}\left[\mathbf{W}_{eK}\mathbf{\Sigma}_{eK}\mathbf{W}_{eK}^T\right] \quad (64)\\ s.t. \ & \sigma_{b_{p_1}}^2 + \sigma_{b_{p_2}}^2 + T^2\sigma_{b_{v_1}}^2 + T^2\sigma_{b_{v_2}}^2 = a^2,\\ & -1 \le \rho_{p_i,v_j} \le 1,\\ & -1 \le \rho_{v_i,v_j} \le 1,\\ & -1 \le \rho_{p_i,p_j} \le 1,\\ & \sigma_{p_i}, \sigma_{v_i} \ge 0, \quad \forall i,j \in \{1,2\} \end{aligned}$$

where

$$\mathbf{W}_{eK} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \quad (65)$$

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

$$\begin{aligned} &\mathrm{Tr}\left[\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right] = \mathrm{Tr}\left[\mathbf{W}_K^T\mathbf{W}_K\mathbf{\Sigma}_K\right] \quad (66)\\ &= (w_{11}^2 + w_{21}^2)^2 s_1 + (w_{12}^2 + w_{22}^2)^2 s_3\\ &+2(w_{11}w_{12} + w_{21}w_{22})s_2 \end{aligned}$$

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

$$\begin{aligned} &\mathrm{Tr}\left[\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right] = (w_{11}^2 + w_{21}^2)^2(c_{p_1}\sigma_{p_1} + c_{p_2}\sigma_{p_2})^2 \quad (67)\\ &+ (w_{12}^2 + w_{22}^2)^2(c_{v_1}\sigma_{v_1} + c_{v_2}\sigma_{v_2})^2\\ &+ 2(w_{11}w_{12} + w_{21}w_{22})(c_{p_1}c_{v_1}\sigma_{p_1}\sigma_{v_1} + c_{p_1}c_{v_2}\sigma_{p_1}\sigma_{v_2}\\ &+ c_{p_2}c_{v_1}\sigma_{p_2}\sigma_{v_1} + c_{p_2}c_{v_2}\sigma_{p_2}\sigma_{v_2}) \end{aligned}$$

So far, we have converted the objective function in (64), 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.

VI. Numerical Results

Numerical results are presented in this section to illustrate the theoretical results.

A. System 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 a Chi-squared test. More specifically, we use the sum of the normalized MSE over N_m Monte-Carlo runs

$$q_k = \sum_{j=1}^{N_m} \left[\hat{\mathbf{x}}_{k|k}^{\prime j} - \mathbf{x}_k^j\right]^T \mathbf{P}_{k|k}^{-1} \left[\hat{\mathbf{x}}_{k|k}^{\prime j} - \mathbf{x}_k^j\right] \quad (68)$$

where at time k, $\mathbf{P}_{k|k}$ is the nominal state covariance matrix calculated by the Kalman filter, $\hat{\mathbf{x}}_{k|k}^{\prime j}$ 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.

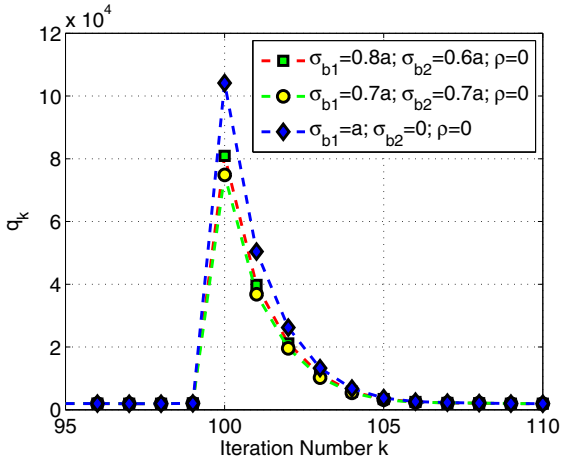

Fig. 2. The normalized MSE when independent biases are used. $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2$ for each case.

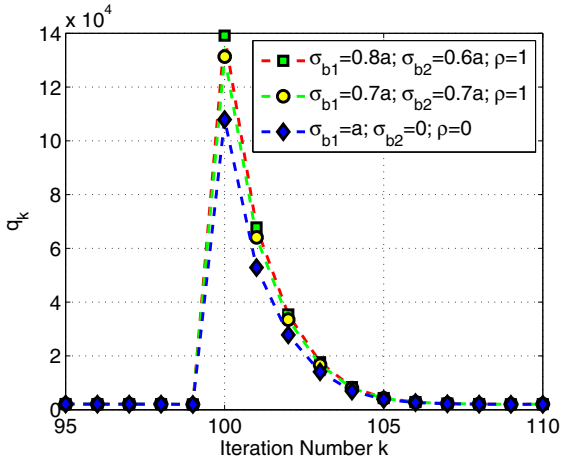

Fig. 3. The normalized MSE for dependent biases. $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2$ for each case.

In Fig. 3, three dependent-noise attack strategies are compared, including the optimal one according to (37), allocating the power equally among the sensors, and allocating all the power to the sensor with 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 $\sigma^2_{b_{p_1}}$ and $T^2\sigma^2_{b_{v_1}}$ to be 3000. In this particular 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.

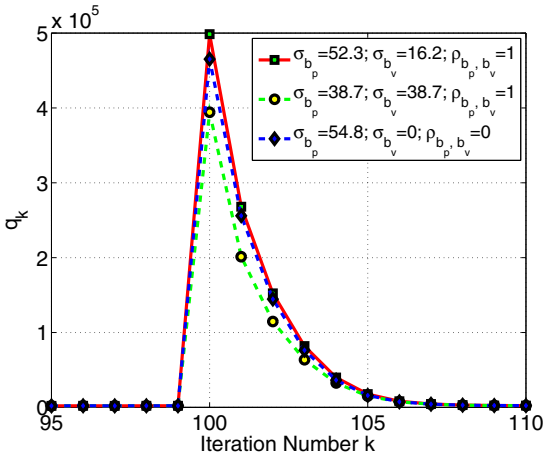

Fig. 4. The normalized MSE for a system with a single sensors. $\sigma^2_{p_1} + T^2\sigma^2_{v_1} = a^2$ for each case.

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 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 $\sigma^2_{b_{p_1}} = 1826, \sigma^2_{b_{p_2}} = 1023, \sigma^2_{b_{v_1}} = 81, \sigma^2_{b_{v_2}} = 68$. For comparison purpose, we also implement an attack strategy that allocate power equally among the observation components and among the two sensors, which is $\sigma^2_{b_{p_1}} = \sigma^2_{b_{p_2}} = \sigma^2_{b_{v_1}} = \sigma^2_{b_{v_2}} = 750$.

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.

VII. Conclusions

In this paper, 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. 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. In both cases, we have theoretically proved that dependent biases can incur more system estimation error than the independent ones. Further,

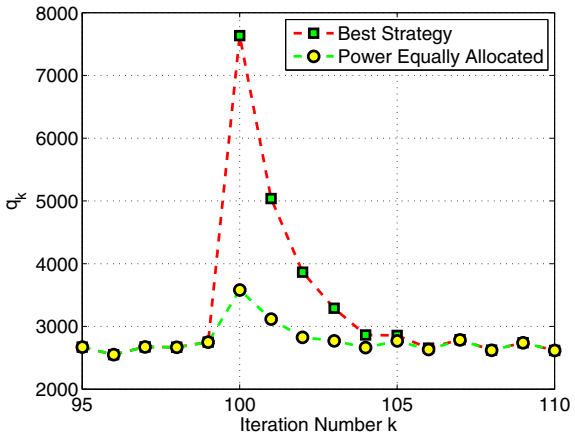

Fig. 5. The normalized MSE for a system with two sensors. $\sigma^2_{p_1} + \sigma^2_{p_2} + T^2\sigma^2_{v_1} + T^2\sigma^2_{v_2} = a^2$ for each case.

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 defense against such attacks.

References

[1] 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.

[2] 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.

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

[4] 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.

[5] 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.

[6] 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.

[7] 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.

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

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