Bayesian Surprise in Linear Gaussian Dynamic Systems: Revisiting State Estimation

1st Yeganeh Zamiri-Jafarian
Department of Electrical and Computer Engineering
University of Toronto
Toronto, Canada
yeganeh.zamiri@mail.utoronto.ca

2nd Konstantinos N. Plataniotis
Department of Electrical and Computer Engineering
University of Toronto
Toronto, Canada
kostas@ece.utoronto.ca

***Abstract*—This article proposes a Bayesian surprise minimization scheme to perform adaptive estimation for a family of linear Gaussian dynamic models. It is shown that the redefined Bayesian surprise in linear Gaussian dynamic systems is a function of the Kalman filter parameters and plays a key role in the state-estimation process. The proposed representation of the Kalman filter illustrates that the information from the Bayesian surprise and the innovation process contributes to the estimation of the state vector and its covariance matrix. This unique approach yields a new set of linear estimation algorithms, where filtering is purely performed with respect to the Bayesian surprise. Simulation results confirm that the information in Bayesian surprise can be sufficient to achieve optimal estimation. In addition, an alternative approach is proposed to test filter consistency based on Bayesian surprise.**

***Index Terms*—Linear Gaussian Dynamic Systems, Bayesian Surprise Minimization, Kalman Filter, Bayesian Surprise, Innovation Process, Mahalanobis Distance.**

I. Introduction

Uncertain events that occur far from expectation, play a significant role in guiding human behavior [1]. The element of surprise, underlying uncertain events, triggers attention and motivation, which as a result influences learning and decision making [2], [3]. Therefore, surprising events are central to understanding and gaining information from the environment through perception and action [4].

Surprise presents useful information on how one's perception is changed and updated over time when encountering uncertain events [5]. The Bayesian surprise, in particular, measures the Kullback-Leibler (KL) information between the prior belief and its update when a new observation is made [6]. Numerous research has considered the Bayesian surprise or some form of it, to acquire information from data for different models and applications [7]–[11]. In [7], the Bayesian surprise measures human attention by analyzing eye movement to improve computer vision applications. A similar concept of the Bayesian surprise is applied in associative learning [8], where surprise is used as an error-correction learning rule for the Rescorla-Wagner model [12]. A confidence-corrected expression for the surprise is presented in [9], where the Bayesian surprise is modified to consider the effect of one's confidence when encountering unexpected events. The principle of free energy suggests that intelligent agents adjust their models to make better predictions by reducing the Bayesian surprise [10]. This concept enhanced the analysis of brain imaging data by unraveling the patterns of cortical activity and connectivity among different regions [11].

Despite significant achievements in specific scenarios [10], [11], a holistic approach to exploring Bayesian surprise for a broad family of system models is not yet accomplished. A structural and unified framework is necessary to examine various aspects of the Bayesian surprise for a wide range of applications (e.g. navigation and guidance systems, radar tracking, etc.). One practical application is the design of advanced vehicular radar systems responsible for determining target motion in autonomous vehicles [13]. Since the environmental conditions of such systems are unexpectedly changing, the exact estimation of a target's motion becomes a challenging problem [14]. To address this issue, new adaptive filtering techniques are required that are capable of applying information from distinct sources. Fortunately, the Bayesian surprise provides a suitable platform to examine how different types of information contribute to adaptive estimation. In this regard, this paper comprehensively exploits the Bayesian surprise in the context of linear Gaussian state-space models for designing novel adaptive filtering algorithms.

The purpose of this research is to demonstrate the effectiveness of Bayesian surprise in state-estimation problems for a class of linear Gaussian dynamic models. The traditional Kalman filter [15] is chosen for performing filtering since it's the optimal estimator in the mean square error sense. The paper also suggests that a surprise minimization mechanism is implemented in the Kalman algorithm. It is proposed that the Bayesian surprise is reduced during the state-estimation process.

Furthermore, the article revisits the derivation of the prediction-estimation steps of the traditional Kalman filter by considering the Bayesian surprise. In the context of linear Gaussian dynamic system, the Bayesian surprise is expressed in terms of the Kalman filter parameters. The main accomplishment is the alternative representation of the Kalman algorithm, where the estimated state vector and its covariance are updated over time based on the information of Bayesian surprise and the information of the innovation process. Also, iterative forms for the Bayesian surprise covariance matrix are examined in both estimation and prediction modes.

Multiple experiments are designed to evaluate the contribution of the Bayesian surprise in the state-estimation process. In this regard, two additional algorithms originated from our proposed representation of Kalman filter are derived. While one algorithm only adopts the information of the Bayesian surprise in estimating the state of the system, the other algorithm functions entirely on the information in the innovation process. The performance of these two algorithms is compared to the conventional Kalman filter algorithm for a simple system model [16]. Results also examine whether a Bayesian surprise minimization mechanism is adopted in the Kalman filter. The statistical properties of the Bayesian surprise and its close connection to the Mahalanobis distance are investigated. To this end, a revised consistency test is proposed that checks filter operation with respect to Bayesian surprise.

The remainder of this paper is organized as follows. Section II describes the state-estimation problem in linear Gaussian dynamic systems. Section III re-defines the Bayesian surprise in the context of linear Gaussian dynamic systems. A new representation of the Kalman filter in terms of Bayesian surprise and innovation process is demonstrated in Section IV. Section V presents the numerical results of the proposed approach and finally Section VI concludes this paper.

II. LINEAR GAUSSIAN DYNAMIC SYSTEMS

This research adopts the linear Gaussian state-space model, where the evolution of the state follows first-order Markov-chain, and the noise elements are assumed additive zero-mean white Gaussian processes [17]. The state equation and the measurement equation at discrete time k, are defined as

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k \tag{1}$$

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

where in (1) $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{F}_k \in \mathbb{R}^{n\times n}$ and $\mathbf{w}_k \in \mathbb{R}^n$ are the state vector, the transition matrix and the state noise, respectively. In (2), $\mathbf{z}_k \in \mathbb{R}^m$, $\mathbf{H}_k \in \mathbb{R}^{m\times n}$ and $\mathbf{v}_k \in \mathbb{R}^m$ are respectively, the measurement vector, the measurement matrix and the measurement noise. The noise elements are zero mean Gaussian processes, assigned as $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}_{n\times 1}, \mathbf{Q}_k)$ and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}_{m\times 1}, \mathbf{R}_k)$, where $\mathbf{Q}_k \in \mathbb{R}^{n\times n}$ is the state noise covariance and $\mathbf{R}_k \in \mathbb{R}^{m\times m}$ is the measurement noise covariance. The model presumes $\mathbb{E}[\mathbf{w}_k \mathbf{w}_{k+j}^T] = \mathbf{0}_{n\times n}$ and $\mathbb{E}[\mathbf{v}_k \mathbf{v}_{k+j}^T] = \mathbf{0}_{m\times m}$, for $\forall k, j$ and $j \neq 0$. In addition, the initial state is set to follow Gaussian distribution, denoted as $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}(0|0), \mathbf{P}(0|0))$, where $\hat{\mathbf{x}}(0|0)$ and $\mathbf{P}(0|0)$ are respectively the estimated state mean and estimated state covariance at time index $k = 0$. Also, $\mathbf{x}_0$, $\mathbf{w}_k$ and $\mathbf{v}_k$ are assumed mutually uncorrelated, where $\mathbb{E}[\mathbf{x}_0 \mathbf{w}_k^T] = \mathbf{0}_{n\times n}$, $\mathbb{E}[\mathbf{x}_0 \mathbf{v}_k^T] = \mathbf{0}_{n\times m}$, $\mathbb{E}[\mathbf{w}_k \mathbf{v}_j^T] = \mathbf{0}_{n\times m}$, for $\forall k, j$.

For the linear Gaussian state-space model given in (1) and (2), the estimation problem is finding the estimated state mean, $\hat{\mathbf{x}}(k|k) = \mathbb{E}[\mathbf{x}_k|\mathbf{Z}_k] \in \mathbb{R}^n$, and the estimated state covariance, $\mathbf{P}(k|k) = \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}(k|k))(\mathbf{x}_k - \hat{\mathbf{x}}(k|k))^T|\mathbf{Z}_k] \in \mathbb{R}^{n\times n}$, given all measurements up to time k, denoted as $\mathbf{Z}_k = \{\mathbf{z}_i, i \leq k\}$. Since $\hat{\mathbf{x}}(k|k)$ and $\mathbf{P}(k|k)$ embody the entire information about

Algorithm 1 Traditional Kalman Filter [17]

Measurement update:
$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}_k \tilde{\mathbf{z}}(k|k-1)$
$\mathbf{P}(k|k) = \mathbf{P}(k|k-1) - \mathbf{K}_k \mathbf{S}(k|k-1) \mathbf{K}_k^T$
$\mathbf{K}_k = \mathbf{P}(k|k-1) \mathbf{H}_k^T \mathbf{S}(k|k-1)^{-1}$
Time update:
$\hat{\mathbf{x}}(k+1|k) = \mathbf{F}_k \hat{\mathbf{x}}(k|k)$
$\mathbf{P}(k+1|k) = \mathbf{Q}_k + \mathbf{F}_k \mathbf{P}(k|k) \mathbf{F}_k^T$

the state $\mathbf{x}_k$ at time k. Under the assumption that the parameters of the system (i.e. $\mathbf{F}_k, \mathbf{H}_k, \mathbf{Q}_k, \mathbf{R}_k, \hat{\mathbf{x}}(0|0), \mathbf{P}(0|0)$) are known, the Kalman filter [15] is the optimal estimator in the mean square error sense to estimate $\hat{\mathbf{x}}(k|k)$ and $\mathbf{P}(k|k)$. Algorithm 1 presents the traditional Kalman filter algorithm for state prediction and estimation [17]. As shown in Algorithm 1, the Kalman filter reduces the state covariance and gains new information from the measurements over time.

While state-estimation in Kalman filtering is a well-established subject, this paper approaches the problem from an alternative aspect, by adopting the Bayesian surprise. To that end, the following section re-introduces the Bayesian surprise in the context of linear Gaussian state-space models.

III. BAYESIAN SURPRISE IN LINEAR GAUSSIAN DYNAMIC SYSTEMS

The Bayesian surprise measures the impact of a new observation by calculating the Kullback-Leibler (KL) distance between a prior distribution and a posterior distribution [6]. For the models (1) and (2), let's consider $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ as the prior distribution (i.e. before observing $\mathbf{z}_k$) and $p(\mathbf{x}_k|\mathbf{Z}_k)$ as the posterior distribution (i.e. after observing $\mathbf{z}_k$). In this regard, the Bayesian surprise at time index k, denoted as $\mathcal{S}_k^B(\mathbf{z}_k)$, is determined as

$$\begin{aligned}\mathcal{S}_k^B(\mathbf{z}_k) &= D_{KL}(p(\mathbf{x}_k|\mathbf{Z}_{k-1}), p(\mathbf{x}_k|\mathbf{Z}_k)) \\ &= \int_{\mathbf{x}_k \in \mathbb{R}^n} p(\mathbf{x}_k|\mathbf{Z}_{k-1}) \ln \frac{p(\mathbf{x}_k|\mathbf{Z}_{k-1})}{p(\mathbf{x}_k|\mathbf{Z}_k)} \mathbf{d}\mathbf{x}_k \\ &= \frac{1}{2}[\ln \frac{|\mathbf{P}(k|k)|}{|\mathbf{P}(k|k-1)|} + \text{tr}\{\mathbf{P}(k|k)^{-1}\mathbf{P}(k|k-1)\} \\ &\qquad - n + ||\hat{\mathbf{x}}(k|k) - \hat{\mathbf{x}}(k|k-1)||^2_{\mathbf{P}(k|k)^{-1}}]\end{aligned} \tag{3}$$

where $\hat{\mathbf{x}}(k|k-1) \in \mathbb{R}^n$ and $\mathbf{P}(k|k-1) \in \mathbb{R}^{n\times n}$ are respectively the predicted state mean and predicted state covariance. The final line in (3) is due to $p(\mathbf{x}_k|\mathbf{Z}_{k-1}) = \mathcal{N}(\hat{\mathbf{x}}(k|k-1), \mathbf{P}(k|k-1))$ and $p(\mathbf{x}_k|\mathbf{Z}_k) = \mathcal{N}(\hat{\mathbf{x}}(k|k), \mathbf{P}(k|k))$, where tr$\{.\}$ calculates trace of a matrix. Note that, $|.|$ and $||.||$ are the determinant operator and the norm operator, respectively.

In part A.1 of the Appendix section, it is shown that the Bayesian surprise depends on the quadratic form $||\hat{\mathbf{x}}(k|k) - \hat{\mathbf{x}}(k|k-1)||^2_{\mathbf{P}(k|k)^{-1}}$, and thus can be approximated as follows

$$\mathcal{S}_k^B(\mathbf{z}_k) \approx \frac{1}{2}||\hat{\mathbf{x}}(k|k) - \hat{\mathbf{x}}(k|k-1)||^2_{\mathbf{P}(k|k)^{-1}} \tag{4}$$

By substituting the estimated state mean from the Kalman algorithm, the expression in (4) can be further simplified to

$$\mathcal{S}_k^B(\mathbf{z}_k) \approx \frac{1}{2}||\tilde{\mathbf{z}}(k|k-1)||^2_{\mathbf{K}_k^T\mathbf{P}(k|k)^{-1}\mathbf{K}_k} \tag{5}$$

where $\tilde{\mathbf{z}}(k|k-1) = \mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}(k|k-1) \in \mathbb{R}^m$ corresponds to the innovation vector and $\mathbf{K}_k \in \mathbb{R}^{n\times m}$ is the Kalman gain. For convenience, here we introduce a new parameter, denoted as $\mathbf{P}_{\mathcal{S}^B}(k|k-1) \in \mathbb{R}^{m\times m}$, which represents the predicted Bayesian surprise covariance. According to (5), the inverse of the predicted Bayesian surprise covariance, also referred to as the predicted Bayesian surprise information matrix (i.e. $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$), is defined as follows

$$\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1} = \mathbf{K}_k^T\mathbf{P}(k|k)^{-1}\mathbf{K}_k \tag{6}$$

The above expression indicates that $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ is a function of the Fischer information matrix, $\mathbf{P}(k|k)^{-1}$, that is projected on to the Kalman gain. For different expressions of the Kalman gain, the following forms for $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ are achieved,

$$\begin{aligned}\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1} &\overset{(1)}{=} \mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{K}_k \\ &\overset{(2)}{=} \mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1} \\ &\overset{(3)}{=} \mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{P}(k|k-1)\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1} \\ &\overset{(4)}{=} \mathbf{R}_k^{-1} - \mathbf{S}(k|k-1)^{-1}\end{aligned} \tag{7}$$

where the first and second expressions are due to substituting $\mathbf{K}_k = \mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1}$. The third expression is the result of using $\mathbf{K}_k = \mathbf{P}(k|k-1)\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1}$; where $\mathbf{S}(k|k-1) = \mathbb{E}[\tilde{\mathbf{z}}(k|k-1)\tilde{\mathbf{z}}(k|k-1)^T|\mathbf{Z}_{k-1}] = \mathbf{R}_k + \mathbf{H}_k\mathbf{P}(k|k-1)\mathbf{H}_k^T$ is the innovation covariance. The final expression in (7) is derived from replacing $\mathbf{H}_k\mathbf{P}(k|k-1)\mathbf{H}_k^T$ with $\mathbf{R}_k - \mathbf{S}(k|k-1)$ in the third definition. In this paper, we refer to $\mathbf{S}(k|k-1)^{-1} \in \mathbb{R}^{m\times m}$ as innovation information matrix. Note that, the first expression in (7), shows that the predicted Bayesian surprise information matrix is closely connected to the filter parameters. Therefore, Bayesian surprise simply depends on the filter components.

IV. Kalman Filtering with Bayesian Surprise

This section revisits the state-estimation problem and illustrates the role of Bayesian surprise in Kalman filtering. Since the Kalman filter is a special case of the Bayesian filter [18], it is anticipated that the elements of Bayesian surprise appear in the Kalman algorithm.

In this regard, we first consider the estimated state mean, $\hat{\mathbf{x}}(k|k)$, which is given in Algorithm 1. To derive $\hat{\mathbf{x}}(k|k)$ in terms of the predicted Bayesian surprise information matrix, we refer to the fourth expression in (7). By substituting $\mathbf{R}_k^{-1} = \mathbf{S}(k|k-1)^{-1} + \mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ and $\mathbf{K}_k = \mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1}$ in $\hat{\mathbf{x}}(k|k)$, the following is achieved

$$\begin{aligned}\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) &+ \mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1) \\ &+ \mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1)\end{aligned} \tag{8}$$

which by applying $\mathbf{P}(k|k)\mathbf{H}_k^T = \mathbf{K}_k\mathbf{R}_k$, leads to

$$\begin{aligned}\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) &+ \mathbf{K}_k\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1) \\ &+ \mathbf{K}_k\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1)\end{aligned} \tag{9}$$

The second term in (9), denoted as $\mathbf{K}_k\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1)$, corresponds to the amount of information that the innovation process contributes to the estimation of $\hat{\mathbf{x}}(k|k)$. Meanwhile, $\mathbf{K}_k\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1)$ indicates the amount of information within the Bayesian surprise that plays a part in estimating the state mean. The terms $\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}$ and $\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ show that how the uncertainty within the measurements are balanced by the amount of information within innovation sequence and information gained from the Bayesian surprise, respectively. From (7), it is clear that $\mathbf{R}_k\mathbf{S}(k|k-1)^{-1} + \mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1} = \mathbf{I}_{m\times m}$; where $\mathbf{I}_{m\times m}$ is the identity matrix. This results in the original estimated state mean of the Kalman filter depicted in Algorithm 1.

In addition, an expression for the estimated state covariance in terms of the predicted Bayesian surprise information matrix could be derived. According to Algorithm 1, using $\mathbf{S}(k|k-1) = \mathbf{R}_k + \mathbf{H}_k\mathbf{P}(k|k-1)\mathbf{H}_k^T$ in $\mathbf{P}(k|k)$, results the following

$$\begin{aligned}\mathbf{P}(k|k) = \mathbf{P}(k|k-1) &- \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\ &- \mathbf{K}_k\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\mathbf{S}(k|k-1)\mathbf{K}_k^T\end{aligned} \tag{10}$$

where $\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\mathbf{S}(k|k-1)$ is substituted instead of $\mathbf{H}_k\mathbf{P}(k|k-1)\mathbf{H}_k^T$ (see expression three in (7)). A simple re-arrangement of (10) provides a suitable representation for $\mathbf{P}(k|k)$,

$$\begin{aligned}\mathbf{P}(k|k) = \mathbf{P}(k|k-1) &- \mathbf{K}_k\{\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T \\ &- \mathbf{K}_k\{\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T\end{aligned} \tag{11}$$

where $\mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T = \mathbf{K}_k\{\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T$ includes the contribution of the innovation process to estimating $\mathbf{P}(k|k)$. On the other hand, the term $\mathbf{K}_k\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\mathbf{S}(k|k-1)\mathbf{K}_k^T$ illustrates the contribution of the Bayesian surprise to the estimation of state covariance. Similar to (9), the terms $\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}$ and $\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ also appear in the estimated state covariance expression.

The expression in (9) and (11) present the measurement update steps of the Kalman filter based on Bayesian surprise and innovation process. Since $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ and $\mathbf{S}(k|k-1)^{-1}$ are associated with the parameters of the measurement equation (given in (2)); they don't directly affect the calculation of $\hat{\mathbf{x}}(k|k-1)$ and $\mathbf{P}(k|k-1)$. Hence, the time update equations for $\hat{\mathbf{x}}(k|k-1)$ and $\mathbf{P}(k|k-1)$ remain unchanged (see Algorithm 1).

The calculation of the predicted Bayesian surprise information matrix (or covariance) is required to update $\hat{\mathbf{x}}(k|k)$ and $\mathbf{P}(k|k)$ at each time cycle. No doubt, a recursive relation for $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ (or $\mathbf{P}_{\mathcal{S}^B}(k|k-1)$) provides an elegant approach to express the Kalman algorithm in terms of Bayesian

Algorithm 2 Kalman Filter in terms of Bayesian surprise and innovation process

Measurement update:

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}_k\{\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\}\tilde{\mathbf{z}}(k|k-1) + \mathbf{K}_k\{\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\}\tilde{\mathbf{z}}(k|k-1)$$

$$\mathbf{P}(k|k) = \mathbf{P}(k|k-1) - \mathbf{K}_k\{\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T - \mathbf{K}_k\{\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T$$

$$\mathbf{K}_k = \mathbf{P}(k|k-1)\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1}$$

$$\mathbf{P}_{\mathcal{S}^B}(k|k) = \mathbf{P}_{\mathcal{S}^B}(k|k-1) + \mathbf{R}_k$$

$$\mathbf{S}(k|k) = \mathbf{S}(k|k-1) - \mathbf{H}_k\mathbf{K}_k\{\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T\mathbf{H}_k^T - \mathbf{H}_k\mathbf{K}_k\{\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T\mathbf{H}_k^T$$

Time update:

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{F}_k\hat{\mathbf{x}}(k|k)$$

$$\mathbf{P}(k+1|k) = \mathbf{Q}_k + \mathbf{F}_k\mathbf{P}(k|k)\mathbf{F}_k^T$$

$$\mathbf{P}_{\mathcal{S}^B}(k+1|k) = \mathbf{S}(k+1|k)\{\mathbf{H}_{k+1}[\mathbf{Q}_k + \mathbf{F}_k \{\mathbf{H}_k^T\mathbf{S}(k|k)^{-1}\mathbf{P}_{\mathcal{S}^B}(k|k)\mathbf{R}_k^{-1}\mathbf{H}_k\}^{-1}\mathbf{F}_k^T]\mathbf{H}_{k+1}^T\}^{-1}\mathbf{R}_{k+1}$$

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

surprise. Since the Kalman filter computes uncertainty rather than information, this work demonstrates recursive relations for $\mathbf{P}_{\mathcal{S}^B}(k|k-1)$ and $\mathbf{P}_{\mathcal{S}^B}(k|k)$.

Let's first consider the estimated Bayesian surprise covariance, denoted as $\mathbf{P}_{\mathcal{S}^B}(k|k)$. To proceed, a definition for $\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1}$ is required. Inspired by the third expression of $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ in (7), $\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1}$ is defined as follows,

$$\begin{aligned}\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1} &\overset{(1)}{=} \mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{S}(k|k)^{-1} \\ &\overset{(2)}{=} \mathbf{R}_k^{-1} - \mathbf{S}(k|k)^{-1}\end{aligned} \tag{12}$$

where $\mathbf{S}(k|k) = \mathbf{R}_k + \mathbf{H}_k\mathbf{P}(k|k)\mathbf{H}_k^T$ is the measurement post-fit residual covariance [19]. Equivalent to the fourth expression in (7), a similar approach is applied to achieve the final term in (12). By replacing $\mathbf{S}(k|k)^{-1}$ with its definition and applying the matrix inversion lemma to (12); the following alternative form for $\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1}$ is obtained

$$\begin{aligned}\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1} &= \mathbf{R}_k^{-1}\mathbf{H}_k(\mathbf{P}(k|k)^{-1} + \mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k)^{-1}\mathbf{H}_k^T\mathbf{R}_k^{-1} \\ &= \mathbf{R}_k^{-1}\mathbf{H}_k(\mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1} - \mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{S}(k|k)^{-1}\mathbf{H}_k\mathbf{K}_k) \\ &= \mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1} - \mathbf{P}_{\mathcal{S}^B}(k|k)^{-1}\mathbf{H}_k\mathbf{K}_k\end{aligned} \tag{13}$$

where $\mathbf{K}_k = \mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1}$, (7) and (12) are used. From (13), $\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1}$ is expressed in terms of $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$, given as

$$\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1} = \mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}(\mathbf{I}_{m\times m} + \mathbf{H}_k\mathbf{K}_k)^{-1} \tag{14}$$

which eventually leads to

$$\begin{aligned}\mathbf{P}_{\mathcal{S}^B}(k|k) &\overset{(1)}{=} (\mathbf{I}_{m\times m} + \mathbf{H}_k\mathbf{K}_k)\mathbf{P}_{\mathcal{S}^B}(k|k-1) \\ &\overset{(2)}{=} \mathbf{P}_{\mathcal{S}^B}(k|k-1) + \mathbf{R}_k\end{aligned} \tag{15}$$

The second definition in (15) is the result of $\mathbf{R}_k = \mathbf{H}_k\mathbf{K}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)$ (see (7)). For the time update expression, let's first consider $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$. The iterative relation for $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ is determined by replacing $\mathbf{P}(k|k-1)$ in the third expression of (7),

$$\begin{aligned}\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1} &= \mathbf{R}_k^{-1}\mathbf{H}_k[\mathbf{Q}_{k-1} + \mathbf{F}_{k-1}\mathbf{P}(k-1|k-1)\mathbf{F}_{k-1}^T]\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1} \\ &= \mathbf{R}_k^{-1}\mathbf{H}_k[\mathbf{Q}_{k-1} + \mathbf{F}_{k-1} \\ &\quad \{\mathbf{H}_{k-1}^T\mathbf{S}(k-1|k-1)^{-1}\mathbf{P}_{\mathcal{S}^B}(k-1|k-1)\mathbf{R}_{k-1}^{-1}\mathbf{H}_{k-1}\}^{-1} \\ &\quad \mathbf{F}_{k-1}^T]\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1}\end{aligned} \tag{16}$$

where $\mathbf{P}(k-1|k-1) = \{\mathbf{H}_{k-1}^T\mathbf{S}(k-1|k-1)^{-1}\mathbf{P}_{\mathcal{S}^B}(k-1|k-1)\mathbf{R}_{k-1}^{-1}\mathbf{H}_{k-1}\}^{-1}$ is computed from (12) for time step $k-1$. The inverse of (16) yields the recursive form of $\mathbf{P}_{\mathcal{S}^B}(k|k-1)$,

$$\begin{aligned}\mathbf{P}_{\mathcal{S}^B}(k|k-1) = \mathbf{S}(k|k-1)\{&\mathbf{H}_k[\mathbf{Q}_{k-1} + \mathbf{F}_{k-1} \\ &\{\mathbf{H}_{k-1}^T\mathbf{S}(k-1|k-1)^{-1}\mathbf{P}_{\mathcal{S}^B}(k-1|k-1)\mathbf{R}_{k-1}^{-1}\mathbf{H}_{k-1}\}^{-1} \\ &\mathbf{F}_{k-1}^T]\mathbf{H}_k^T\}^{-1}\mathbf{R}_k\end{aligned} \tag{17}$$

Note that, both (15) and (17) present significant insight on the performance of the filter. According to (15) and (17), the uncertainty associated with the Bayesian surprise is increased over time until the system reaches a steady state. While the Bayesian surprise covariance is increased, $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ (or $\mathbf{P}_{\mathcal{S}^B}(k|k)^{-1}$) is reduced over time. In other words, as the filter extracts the innovative information from the data (i.e. increase in $\mathbf{S}(k|k-1)^{-1}$), there remains no uncertainty or surprising element within the measurements; which from (5), the Bayesian surprise is minimized. Therefore, a Bayesian surprise minimization mechanism is employed during the state-estimation process of the Kalman filter. The Bayesian surprise minimization approach demonstrated in this work is aligned with the principle of free energy [10]. Similar to biological agents, the Kalman filter reaches a steady state by reducing the Bayesian surprise.

Deriving a recursive form for $\mathbf{S}(k|k)$ may ease the computation of the predicted Bayesian surprise covariance at each cycle. Based on (11), $\mathbf{S}(k|k)$ is obtained as

$$\begin{aligned}\mathbf{S}(k|k) = \mathbf{S}(k|k-1) &- \mathbf{H}_k\mathbf{K}_k\{\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T\mathbf{H}_k^T \\ &- \mathbf{H}_k\mathbf{K}_k\{\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\}\mathbf{S}(k|k-1)\mathbf{K}_k^T\mathbf{H}_k^T\end{aligned} \tag{18}$$

where the expressions for the innovation covariance is considered. Note that, unlike $\mathbf{S}(k|k)$, a recursive relation for $\mathbf{S}(k|k-1)$ doesn't exist. This is because $\mathbf{S}(k|k-1)$ corresponds to novel part of measurements that are independent over time. Algorithm 2 presents the Kalman filter algorithm in terms of Bayesian surprise and innovation process.

V. Simulation and Numerical Results

In this paper, two numerical results are carried out to evaluate the performance of the Kalman filter based on Bayesian surprise. The first experiment examines and compares the state-estimation performance of the proposed filter by calculating the root mean square relative error (RMSRE). In the meantime, the second experiment is designed to analyze the

TABLE I
MODEL PARAMETERS USED IN SIMULATION [16]

$\mathbf{F}$	$\mathbf{H}$	$\mathbf{Q}$	$\mathbf{R}$	$\hat{\mathbf{x}}(0\|0)$	$\mathbf{P}(0\|0)$
$\begin{bmatrix} 0.82 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	1	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$

potential of Bayesian surprise for checking filter consistency. Numerical results are achieved through multiple Monte Carlo simulations for a simple system model given in Table I [16].

Experiment 1

In this experiment, the impact of the Bayesian surprise on the performance of the Kalman filter is evaluated. To measure the effectiveness of Bayesian surprise, we consider three versions of the Kalman filter. The first one is the traditional Kalman filter given in Algorithm 1. Meanwhile, the second filter only considers the impact of Bayesian surprise, and the third filter only functions on the innovation process. The measurement update algorithms, based on only Bayesian surprise and only innovation process, are illustrated in Algorithm 3 and Algorithm 4, respectively. Note that, the time update steps remain unchanged and are the same as Algorithm 2. The performance of each Kalman filter is evaluated by computing their RMSRE at time index k, defined as follows

$$RMSRE_k = \frac{\sqrt{\sum_{j=1}^{N_{mc}} ||\hat{\mathbf{x}}^j(k|k) - \mathbf{x}_k^j||^2}}{\sqrt{\sum_{j=1}^{N_{mc}} ||\mathbf{x}_k^j||^2}} \tag{19}$$

where $\mathbf{x}_k^j$ and $\hat{\mathbf{x}}^j(k|k)$ are the state vector and the estimated state vector from the j-th Monte Carlo simulation at time step k. It is noteworthy to mention that RMSRE is unit-less, which represents the relative error of the estimation.

Algorithm 3 Kalman Filter in terms of the Bayesian surprise

Measurement update:

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}_k\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1)$$
$$\mathbf{P}(k|k) = \mathbf{P}(k|k-1) - \mathbf{K}_k\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\mathbf{S}(k|k-1)\mathbf{K}_k^T$$
$$\mathbf{K}_k = \mathbf{P}(k|k-1)\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1}$$
$$\mathbf{P}_{\mathcal{S}^B}(k|k) = \mathbf{P}_{\mathcal{S}^B}(k|k-1) + \mathbf{R}_k$$
$$\mathbf{S}(k|k) = \mathbf{S}(k|k-1) - \mathbf{H}_k\mathbf{K}_k\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\mathbf{S}(k|k-1)\mathbf{K}_k^T\mathbf{H}_k^T$$

Algorithm 4 Kalman Filter in terms of innovation process

Measurement update:

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}_k\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1)$$
$$\mathbf{P}(k|k) = \mathbf{P}(k|k-1) - \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T$$
$$\mathbf{K}_k = \mathbf{P}(k|k-1)\mathbf{H}_k^T\mathbf{S}(k|k-1)^{-1}$$
$$\mathbf{P}_{\mathcal{S}^B}(k|k) = \mathbf{P}_{\mathcal{S}^B}(k|k-1) + \mathbf{R}_k$$
$$\mathbf{S}(k|k) = \mathbf{S}(k|k-1) - \mathbf{H}_k\mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T\mathbf{H}_k^T$$

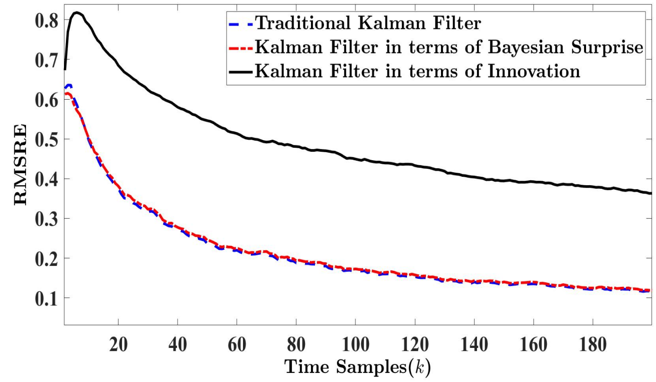

Fig. 1. Comparison between the RMSRE performance of three Kalman filters.

The RMSRE of the three Kalman filters over $N_{mc} = 2000$ rounds of Monte Carlo simulations are depicted in Fig. 1. Fig. 1 shows that by omitting the information in the innovation that is projected by the measurement noise ($\mathbf{R}_k\mathbf{S}(k|k-1)^{-1}$), the performance of the system is analogous to the traditional Kalman filter. In the meantime, by ignoring the information in the Bayesian surprise that is projected on the measurement noise ($\mathbf{R}_k\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$), the performance of the filter drops significantly compared to the original Kalman filter. Let's investigate whether the results achieved in Fig. 1 is anticipated or not. This requires observing the behaviour of the predicted Bayesian surprise information, $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$, and the innovation information, $\mathbf{S}(k|k-1)^{-1}$. The reasoning is traced back to (9) and (11), where these two elements play a part in the estimation of state mean and state covariance, respectively. Here we have considered the sample-based innovation information, denoted as $\bar{\mathbf{S}}(k|k-1)^{-1}$, and the sample-based Bayesian surprise information, denoted as $\bar{\mathbf{P}}_{\mathcal{S}^B}(k|k-1)^{-1}$. The sample-based versions are calculated from the innovation vector; such that, $\bar{\mathbf{S}}(k|k-1) = \frac{1}{k}\sum_{j=1}^{k}\tilde{\mathbf{z}}(j|j-1)\tilde{\mathbf{z}}(j|j-1)^T$ and $\bar{\mathbf{P}}_{\mathcal{S}^B}(k|k-1) = \mathbf{R}_k^{-1} - \bar{\mathbf{S}}(k|k-1)^{-1}$. Fig. 2 illustrates $\bar{\mathbf{S}}(k|k-1)^{-1}$ and $\bar{\mathbf{P}}_{\mathcal{S}^B}(k|k-1)^{-1}$ averaged over 2000 run of Monte Carlo simulations. According to Fig. 2, the information measured by the Bayesian surprise is on average more reliable

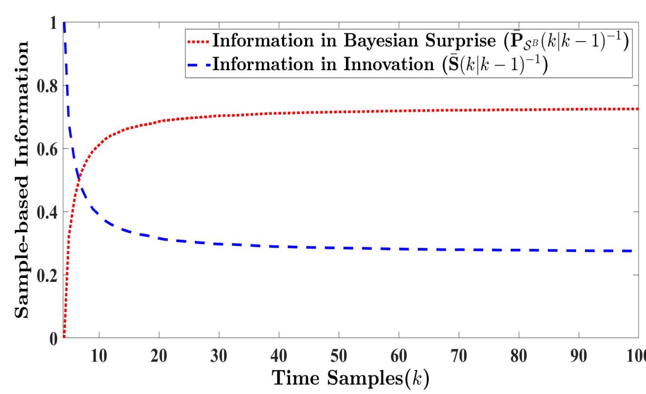

Fig. 2. The sample-based values of the predicted Bayesian surprise information and the innovation information averaged over 2000 trials.

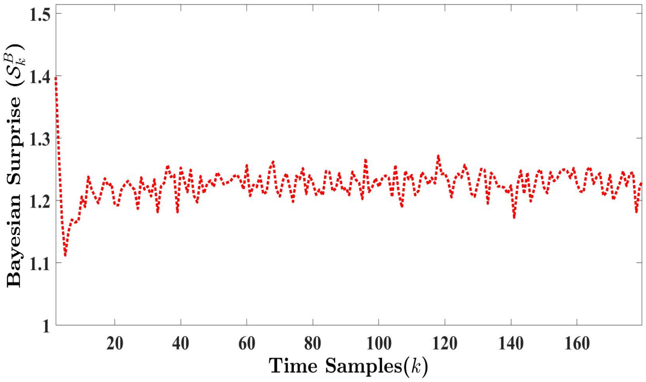

Fig. 3. The Bayesian surprise averaged over 2000 trials.

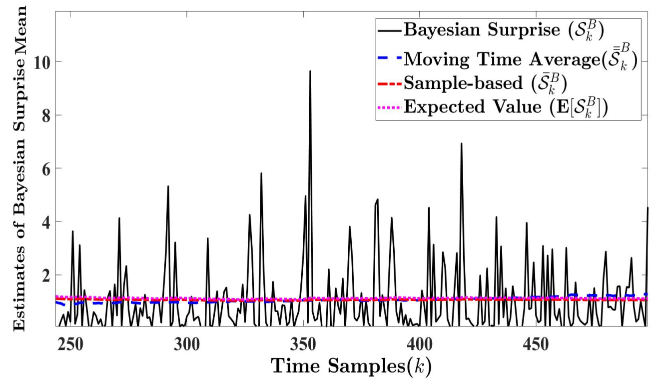

Fig. 4. The comparison of sample-based Bayesian surprise, moving time average and the expected value of Bayesian surprise.

compared to the information in innovation. Note that, in general the contribution of Bayesian surprise and innovation process vary depending on the system model and the application. $\mathbf{H}_k$ and $\mathbf{Q}_k$ considerably impacts the calculation of the predicted Bayesian surprise information; such that it is increased by factor $\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^T$ (see (16)). The opposite effect is applied in computing $\mathbf{S}(k|k-1)^{-1}$. Therefore, the result achieved in Fig. 1 becomes inevitable for the model in Table I.

In addition, the minimization of Bayesian surprise during state-estimation is also examined. Fig. 3 illustrates that the Bayesian surprise, $\mathcal{S}_k^B \approx \frac{1}{2}||\tilde{\mathbf{z}}(k|k-1)||^2_{\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}}$, is reduced and a minimum is reached in steady state. This outcome is traced back to the proposed iterative form for $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ (see (16)), where over time information in Bayesian surprise is decreased by the Kalman filter. The result in Fig. 3 is averaged over 2000 trials.

Experiment 2

This experiment is designed to investigate filter consistency in terms of Bayesian surprise. Rather than computing the state error (e.g. RMSRE), alternative techniques can also be applied to test if the filter is working correctly or not. The most common approach for consistency verification is conducting the chi-square testing of the squared Mahalanobis distance [19]. Let's briefly address the chi-square testing in adaptive filters. For a consistent filter, the innovation vector, $\tilde{\mathbf{z}}(k|k-1)$, follows zero mean Gaussian distribution with covariance matrix, $\mathbf{S}(k|k-1)$ (i.e. $\tilde{\mathbf{z}}(k|k-1) \sim \mathcal{N}(\mathbf{0}_{m\times 1}, \mathbf{S}(k|k-1))$). If the filter is working correctly, then the squared Mahalanobis distance, defined as $\mathbf{d}_k^2 = ||\tilde{\mathbf{z}}(k|k-1)||^2_{\mathbf{S}(k|k-1)^{-1}}$, follows a chi-squared distribution with m degrees of freedom (i.e. $\mathbf{d}_k^2 \sim \mathcal{X}_m^2$). Since the following relation exists between the squared Mahalanobis distance and the Bayesian surprise (see (5), (6), (7)),

$$\mathcal{S}_k^B \approx \frac{1}{2}||\tilde{\mathbf{z}}(k|k-1)||^2_{\mathbf{R}_k^{-1}} - \frac{1}{2}\mathbf{d}_k^2 \tag{20}$$

this work proposes the Bayesian surprise as a possible metric to verify filter consistency. To derive the consistency test, the initial step requires the distribution of the Bayesian surprise. In part A.2 of the Appendix section, we have shown that the Bayesian surprise follows a Gamma distribution, with shape κ_{sum} and scale $\frac{\theta_{sum}}{2}$ (i.e. $\mathcal{S}_k^B \sim \Gamma(\kappa_{sum}, \frac{\theta_{sum}}{2})$). The expressions for κ_{sum} and θ_{sum} are derived in the Appendix section (see 35). For the filter to pass the consistency test based on the Bayesian surprise, the following relationship should hold,

$$\mathbb{E}[\mathcal{S}_k^B] = \frac{\kappa_{sum}\theta_{sum}}{2} = \frac{1}{2}\sum_{l=1}^{m}\sigma_l^2 \tag{21}$$

where σ_l^2 is computed from (34). To estimate $\mathbb{E}[\mathcal{S}_k^B]$, the sample-based Bayesian surprise, defined as $\bar{\mathcal{S}}_k^B = \frac{1}{k}\sum_{j=1}^{k}\mathcal{S}_k^B$, or the moving time average of Bayesian surprise, defined as $\bar{\bar{\mathcal{S}}}_k^B = \frac{1}{N_m}\sum_{j=1}^{N_m}\mathcal{S}_k^B$, may apply. Here, $\bar{\mathcal{S}}_k^B$ and $\bar{\bar{\mathcal{S}}}_k^B$ are computed for a single run of the filter, when the moving average window is set to $N_m = 200$. Since the innovations are ergodic for the system model in Table I, both the sample-based Bayesian surprise and the moving average yield to the same value. Fig. 4 depicts $\bar{\mathcal{S}}_k^B$ and $\bar{\bar{\mathcal{S}}}_k^B$, while comparing them with the expected value of the Bayesian surprise derived in (21). As shown, over time the sample-based Bayesian surprise and moving time average align closely with $\mathbb{E}[\mathcal{S}_k^B]$. This shows that the filter is consistent and works correctly by applying Bayesian surprise. Thus, Bayesian surprise is a suitable measure to perform filter consistency test. Note that, to compute $\mathbb{E}[\mathcal{S}_k^B]$, we have used the sample-based innovation covariance in the definition of σ_l^2.

VI. Conclusion

This article demonstrated the state-estimation problem of linear Gaussian dynamic systems by adopting the Bayesian surprise. The re-interpretation of the conventional Kalman filter indicated that the estimated state vector and its covariance matrix are updated over time based on information in the Bayesian surprise and information in the innovation process. This work demonstrated that optimal estimation in the Kalman filter is achieved through Bayesian surprise minimization. For a simple system model, simulation results illustrated that the contribution of Bayesian surprise outperforms the innovation process. Also, an alternative approach for testing filter consistency in terms of Bayesian surprise was determined.

VII. Appendix

A.1: Derivation of Bayesian Surprise

This section presents an approximate expression for the Bayesian surprise. To find an elegant expression for the Bayesian surprise given in (3), let's simplify $\text{tr}\{\mathbf{P}(k|k)^{-1}\mathbf{P}(k|k-1)\}$ and $\ln\frac{|\mathbf{P}(k|k)|}{|\mathbf{P}(k|k-1)|}$, by using the Kalman (or information) filter equations. By considering $\mathbf{P}(k|k)^{-1} = \mathbf{P}(k|k-1)^{-1} + \mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k$, the following is achieved

$$\text{tr}\{\mathbf{P}(k|k)^{-1}\mathbf{P}(k|k-1)\} = n + \text{tr}\{\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{P}(k|k-1)\} \tag{22}$$

where $\text{tr}\{\mathbf{I}_{n\times n}\} = n$. Meanwhile, a closed term for $\ln\frac{|\mathbf{P}(k|k)|}{|\mathbf{P}(k|k-1)|}$ is derived by adopting $\mathbf{P}(k|k) = (\mathbf{I}_{n\times n} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}(k|k-1)$ from Kalman equations; which results the following

$$\ln\frac{|\mathbf{P}(k|k)|}{|\mathbf{P}(k|k-1)|} = \ln(|\mathbf{I}_{m\times m} - \mathbf{H}_k\mathbf{K}_k|) \tag{23}$$

where the determinant properties used in (23) includes, $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ (i.e. for any square matrix $\mathbf{A}$, $\mathbf{B}$) and $|\mathbf{I}_{n\times n} - \mathbf{AB}| = |\mathbf{I}_{m\times m} - \mathbf{BA}|$ (i.e. for $\mathbf{A} \in \mathbb{R}^{n\times m}$,$\mathbf{B} \in \mathbb{R}^{m\times n}$. Also, for any positive definite matrix $\mathbf{C}$, lower and upper bounds for the log determinant could be defined, $\text{tr}\{\mathbf{I} - \mathbf{C}^{-1}\} \leq \ln|\mathbf{C}| \leq \text{tr}\{\mathbf{C} - \mathbf{I}\}$. Since, $\mathbf{I}_{m\times m}$ and $\mathbf{H}_k\mathbf{K}_k$ are positive definite matrices, $\mathbf{I}_{m\times m} - \mathbf{H}_k\mathbf{K}_k$ is also a positive definite matrix. Therefore, the lower bound, denoted as LB, and the upper bound, denoted as UB, of $\ln(|\mathbf{I}_{m\times m} - \mathbf{H}_k\mathbf{K}_k|)$ are respectively obtained as

$$\begin{aligned} LB &= \text{tr}\{-\mathbf{H}_k(\mathbf{I}_{n\times n} - \mathbf{K}_k\mathbf{H}_k)^{-1}\mathbf{K}_k\} \\ &= -\text{tr}\{(\mathbf{I}_{n\times n} - \mathbf{K}_k\mathbf{H}_k)^{-1}\mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\} \end{aligned} \tag{24}$$

$$UB = \text{tr}\{-\mathbf{H}_k\mathbf{K}_k\} = -\text{tr}\{\mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\} \tag{25}$$

where the trace properties used above includes, $\text{tr}\{c\mathbf{A}\} = c\text{tr}\{\mathbf{A}\}$ and $\text{tr}\{\mathbf{ABC}\} = \text{tr}\{\mathbf{BCA}\} = \text{tr}\{\mathbf{CAB}\}$ (i.e. for any matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$). Applying the matrix inversion lemma and setting $\mathbf{K}_k = \mathbf{P}(k|k)\mathbf{H}_k^T\mathbf{R}_k^{-1}$, leads to the final expressions in (24) and (25). By adopting (22), (24) and (25) in (3), the bounds for the Bayesian surprise become,

$$LB_{\mathcal{S}^B} = \frac{1}{2}||\hat{\mathbf{x}}(k|k) - \hat{\mathbf{x}}(k|k-1)||^2_{\mathbf{P}(k|k)^{-1}} \tag{26}$$

$$\begin{aligned} UB_{\mathcal{S}^B} &= \frac{1}{2}||\hat{\mathbf{x}}(k|k) - \hat{\mathbf{x}}(k|k-1)||^2_{\mathbf{P}(k|k)^{-1}} \\ &+ \frac{1}{2}\text{tr}\{\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{K}_k\mathbf{H}_k\mathbf{P}(k|k-1)\} \end{aligned} \tag{27}$$

where the Kalman equation regarding the estimated state covariance is used to achieve (26) and (27). Note that, the Bayesian surprise varies based on the parameters of the state-space model. However, in steady state condition it is clear that Bayesian surprise only depends on $\frac{1}{2}||\hat{\mathbf{x}}(k|k) - \hat{\mathbf{x}}(k|k-1)||^2_{\mathbf{P}(k|k)^{-1}}$. Hence, the following approximation can be made,

$$\mathcal{S}_k^B(\mathbf{z}_k) \approx \frac{1}{2}||\hat{\mathbf{x}}(k|k) - \hat{\mathbf{x}}(k|k-1)||^2_{\mathbf{P}(k|k)^{-1}} \tag{28}$$

which concludes (4).

A.2: Distribution of Bayesian Surprise

The distribution of the Bayesian surprise is obtained by adopting the concept of eigenvalue and eigenvectors. To this end, let's express $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$, $\mathbf{S}(k|k-1)^{-1}$ and $\mathbf{R}_k^{-1}$ in terms of there eigenvalue and eigenvectors as follows,

$$\begin{aligned} \mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1} &= \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T = \sum_{l=1}^{m}\lambda_l^{-1}\mathbf{u}_l\mathbf{u}_l^T \\ \mathbf{S}(k|k-1)^{-1} &= \mathbf{V}\mathbf{D}^{-1}\mathbf{V}^T = \sum_{l=1}^{m}d_l^{-1}\mathbf{v}_l\mathbf{v}_l^T \\ \mathbf{R}_k^{-1} &= \mathbf{C}\mathbf{B}^{-1}\mathbf{C}^T = \sum_{l=1}^{m}b_l^{-1}\mathbf{c}_l\mathbf{c}_l^T \end{aligned} \tag{29}$$

where the columns of the $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{C}$ are the corresponding eigenvectors $\{\mathbf{u}\}_{i=1}^m$, $\{\mathbf{v}\}_{i=1}^m$, $\{\mathbf{c}\}_{i=1}^m$ of the inverse covariances $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$, $\mathbf{S}(k|k-1)^{-1}$ and $\mathbf{R}_k^{-1}$, respectively. Also $\mathbf{\Lambda}^{-1}$, $\mathbf{D}^{-1}$ and $\mathbf{B}^{-1}$ are diagonal matrices containing the eigenvalues $\{\lambda_i^{-1}\}_{i=1}^m$, $\{d_i^{-1}\}_{i=1}^m$, and $\{b_i^{-1}\}_{i=1}^m$ of the corresponding covariances. By substituting $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ from (29) in the Bayesian surprise, the following is achieved,

$$\begin{aligned} \mathcal{S}_k^B &= \frac{1}{2}\tilde{\mathbf{z}}(k|k-1)^T\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}\tilde{\mathbf{z}}(k|k-1) \\ &= \frac{1}{2}\tilde{\mathbf{z}}(k|k-1)^T(\sum_{l=1}^{m}\lambda_l^{-1}\mathbf{u}_l\mathbf{u}_l^T)\tilde{\mathbf{z}}(k|k-1) \\ &= \frac{1}{2}\sum_{l=1}^{m}Y_l^2 \end{aligned} \tag{30}$$

where $Y_l = \lambda_l^{-\frac{1}{2}}\mathbf{u}_l^T\tilde{\mathbf{z}}(k|k-1)$ is a new random variable. Since, $\tilde{\mathbf{z}}(k|k-1) \sim \mathcal{N}(\mathbf{0}_{m\times m}, \mathbf{S}(k|k-1))$; then according to (30), Y_l is a univariate gaussian distribution that follows $Y_l \sim \mathcal{N}(0, \sigma_l^2)$. According to (30), σ_l^2 is determined as,

$$\sigma_l^2 = (\lambda_l^{-\frac{1}{2}}\mathbf{u}_l^T)\mathbf{S}(k|k-1)(\lambda_l^{-\frac{1}{2}}\mathbf{u}_l) = \lambda_l^{-1}\mathbf{u}_l^T\mathbf{S}(k|k-1)\mathbf{u}_l \tag{31}$$

To further simplify σ_l^2, let's find an expression for $\lambda_l^{-1}\mathbf{u}_l^T$. Since $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1} = \mathbf{R}_k^{-1} - \mathbf{S}(k|k-1)^{-1}$, the following relationship holds (from (29)),

$$\mathbf{\Lambda}^{-1}\mathbf{U}^T = \mathbf{U}^T(\mathbf{C}\mathbf{B}^{-1}\mathbf{C}^T - \mathbf{V}\mathbf{D}^{-1}\mathbf{V}^T) \tag{32}$$

where $\mathbf{U}^{-1} = \mathbf{U}^T$. Note that, $\mathbf{P}_{\mathcal{S}^B}(k|k-1)^{-1}$ is presumed symmetric and thus, its eigenvectors are orthogonal. With some manipulations, for the l-th column, the following expression is derived,

$$\lambda_l^{-1}\mathbf{u}_l^T = \mathbf{u}_l^T[\sum_{l=1}^{m}b_l^{-1}\mathbf{c}_l\mathbf{c}_l^T - \sum_{l=1}^{m}d_l^{-1}\mathbf{v}_l\mathbf{v}_l^T] \tag{33}$$

by substituting (33) in (31), σ_l^2 is simplified to

$$\begin{aligned}\sigma_l^2 &= \mathbf{u}_l^T[\sum_{l=1}^{m} b_l^{-1}\mathbf{c}_l\mathbf{c}_l^T - \sum_{l=1}^{m} d_l^{-1}\mathbf{v}_l\mathbf{v}_l^T](\sum_{l=1}^{m} d_l\mathbf{v}_l\mathbf{v}_l^T)\mathbf{u}_l \\ &= \mathbf{u}_l^T[\mathbf{R}_k^{-1}\mathbf{S}(k|k-1) - \mathbf{I}_{m\times m}]\mathbf{u}_l\end{aligned} \quad (34)$$

where $\mathbf{S}(k|k-1) = \sum_{l=1}^{m} d_l\mathbf{v}_l\mathbf{v}_l^T$ is used. Note that, (34) provides a closed form expression to compute σ_l^2. Due to $Y_l \sim \mathcal{N}(0, \sigma_l^2)$, then Y_l^2, follows a chi-squared distribution with l degrees of freedom that is scaled by, σ_l^2; $Y_l^2 \sim \sigma_l^2\mathcal{X}_1^2$. Since chi-squared distribution is a special case of Gamma distribution, an alternate representation for Y_l^2 is the Gamma distribution, denoted as $Y_l^2 \sim \Gamma(\frac{1}{2}, 2\sigma_l^2)$, with shape $\kappa_l = \frac{1}{2}$ and scale $\theta_l = 2\sigma_l^2$. Note that, the summation of $\sum_{l=1}^{m} Y_l^2 = Y_1^2 + Y_2^2 + ... + Y_m^2$ is approximated by a Gamma distribution, where the shape, denoted as κ_{sum}, and the scale, denoted as θ_{sum}, are defined as [20]

$$\begin{aligned}\kappa_{sum} &= \frac{(\sum_{l=1}^{m}\theta_l\kappa_l)^2}{\sum_{l=1}^{m}\theta_l^2\kappa_l} = \frac{(\sum_{l=1}^{m}\sigma_l^2)^2}{\sum_{l=1}^{m}2(\sigma_l^2)^2} \\ \theta_{sum} &= \frac{\sum_{l=1}^{m}\theta_l\kappa_l}{\kappa_{sum}} = \frac{\sum_{l=1}^{m}2(\sigma_l^2)^2}{\sum_{l=1}^{m}\sigma_l^2}\end{aligned} \quad (35)$$

where according to (30) the distribution of the Bayesian surprise becomes, $\mathcal{S}_k^B \sim \Gamma(\kappa_{sum}, \frac{\theta_{sum}}{2})$.

References

[1] M. Browning and C. J. Harmer, "Expectancy and surprise predict neural and behavioral measures of attention to threatening stimuli," *Neuroimage*, vol. 59, no. 2, pp. 1942–1948, 2012.

[2] L. Itti and P. Baldi, "Bayesian surprise attracts human attention," *Vision research*, vol. 49, no. 10, pp. 1295–1306, 2009.

[3] C. Ranganath and G. Rainer, "Cognitive neuroscience: Neural mechanisms for detecting and remembering novel events," *Nature Reviews Neuroscience*, vol. 4, no. 3, p. 193, 2003.

[4] J. M. Fuster, *Cortex and mind: Unifying cognition.* Oxford university press, 2003.

[5] A. Barto, M. Mirolli, and G. Baldassarre, "Novelty or surprise?," *Frontiers in psychology*, vol. 4, p. 907, 2013.

[6] P. Baldi, "A computational theory of surprise," in *Information, Coding and Mathematics*, pp. 1–25, Springer, 2002.

[7] P. Baldi and L. Itti, "Of bits and wows: a bayesian theory of surprise with applications to attention," *Neural Networks*, vol. 23, no. 5, pp. 649–666, 2010.

[8] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[9] M. Faraji, K. Preuschoff, and W. Gerstner, "Balancing new against old information: the role of puzzlement surprise in learning," *Neural computation*, vol. 30, no. 1, pp. 34–83, 2018.

[10] K. Friston, "The free-energy principle: a unified brain theory?," *Nature reviews neuroscience*, vol. 11, no. 2, p. 127, 2010.

[11] A. Razi and K. J. Friston, "The connected brain: causality, models, and intrinsic dynamics," *IEEE Signal Processing Magazine*, vol. 33, no. 3, pp. 14–35, 2016.

[12] R. A. Rescorla, A. R. Wagner, *et al.*, "A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement," *Classical conditioning II: Current research and theory*, vol. 2, pp. 64–99, 1972.

[13] T. Luettel, M. Himmelsbach, and H.-J. Wuensche, "Autonomous ground vehicles—concepts and a path to the future," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1831–1839, 2012.

[14] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: A review of signal processing techniques," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 22–35, 2017.

[15] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[16] D. G. Lainiotis, "Adaptive dynamic modeling, risk assessment, failure detection and correction," in *Advances in Structural Reliability, Proceedings of the Advanced Seminar on Structural Reliability, JRC-Ispra*, 1984.

[17] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[18] S. Haykin, *Cognitive dynamic systems: perception-action cycle, radar and radio*. Cambridge University Press, 2012.

[19] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[20] B. L. Welch, "The generalization of student's' problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1/2, pp. 28–35, 1947.