The Role of Kalman Gain and Noise Covariance Selection on the Convergence in State Estimation

Lamia Alyami
Centre for Environmental Mathematics,
Faculty of Environment, Science and Economy,
University of Exeter, Penryn Campus,
Cornwall TR10 9FE, United Kingdom.
Email: la424@exeter.ac.uk, lmlasloom@nu.edu.sa

Saptarshi Das, *Member, IEEE*
Centre for Environmental Mathematics,
Faculty of Environment, Science and Economy,
University of Exeter, Penryn Campus,
Cornwall TR10 9FE, United Kingdom.
Email: S.Das3@exeter.ac.uk, saptarshi.das@ieee.org

***Abstract*—In state estimation problems, the Kalman filter (KF) algorithm considers the noise in the measurements and the systems facilitating convergence to the true state. This paper presents the Bayesian derivation of the discrete-time KF algorithm for a simple example known as the random walk model. However, if the KF coefficients are not well-tuned, it can significantly impact the estimation accuracy and may lead to algorithmic inconsistency. The Kalman gain is a quantitative measure which plays a crucial role in achieving the optimum convergence and stability. In this study, we evaluate the importance of the Kalman gain in the KF algorithm across several choices of the error covariance within the context of the random walk model. Furthermore, we demonstrate that the optimal Kalman gain is determined by minimizing the mean squared error (MSE), producing an unbiased and efficient estimate. This adaptive adjustment enables the KF to tune parameters easily. The theoretical and numerical investigations were carried out using the random walk plus noise model.**

***Index Terms*—Kalman filter, Kalman gain, convergence, random walk model, state estimation, noise covariance**

I. Introduction

Accurate state estimation is fundamental in estimation theory and control systems for a wide range of applications. The Kalman filter (KF) [1] is a standard mathematical tool that provides an optimal and unbiased solution for linear and Gaussian systems by minimizing mean squared error (MMSE) between the state variables and its estimates. The KF algorithm works recursively by incorporating the noisy measurements in the dynamic system model to estimate the hidden variables or partially observed systems along with their uncertainty [2]. The KF algorithm works in two phases. The first phase is the prediction step which is the estimation of the current state along with their uncertainties. The second phase is the correction step which works once the noisy measurement has arrived. Then the previous step is updated and improved by scaling it using a weighted average value called the Kalman gain matrix K_t. This leads to a new state estimate that places between the predicted and the measured state with a reduced uncertainty degree. The Kalman gain matrix K_t is a crucial component of the Kalman filter algorithm, which is responsible for stabilizing and converging the direction of the algorithm see e.g. [3]–[5]. There is little literature addressing the significant role of the Kalman gain. In [6], the Kalman gain matrix was employed to mitigate the influence of sample size on the estimation process. In [7], the authors investigated the characteristic of the Kalman gain within a type of Kalman filter known as the cubature Kalman filter. In [8], the Kalman gain has been used to estimate the ellipsoidal states. The comparisons of the Kalman gain regularization and covariance regularization methods were reported in [9] which found that Kalman gain regularization is more accurate for predicting the real data. However, it is important to note that there are several modifications of the conventional Kalman filter algorithm, such as the unscented Kalman filter, extended Kalman filter, ensemble Kalman filter, and skew Kalman filter. The Kalman gain may have a different formula for each type, but they are equivalent in functionality and serve the same purpose. We will discuss in this paper the derivation of the KF algorithm from the Bayesian perspective and then the significance of the Kalman gain K_t selection in a simple Gaussian dynamic linear model. Numerous studies have addressed Kalman filter tuning, e.g. [10]–[14]. In this study, we aim to address the Kalman filter tuning problem via the Kalman gain performance since it remained an unsolved problem for a long time. This challenge has remained without a definitive solution, particularly in the context of high-dimensional systems. The default method is by trial and error approach which continues to be popular in the KF setting in an interactive way.

II. Kalman Filter from a Bayesian Perspective

The KF algorithm has been derived from different perspectives such as the Bayesian filtering approach [15], [16], maximum likelihood [17], Newton method [18] and least squares estimation [2]. The most popular derivations of the KF are Bayesian filtering and the least squares estimation, providing optimal solutions of the KF which are recognized in the theoretical literature. However, a fully detailed derivation of the KF algorithm using a purely Bayesian approach is missing as explained in [19]. The reason behind this is that the Bayesian filtering cannot be performed analytically in the non-Gaussian case and least squares estimation is preferable

979-8-3503-0551-7/23/$31.00 ©2023 IEEE

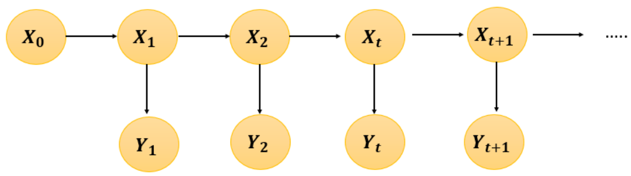

Fig. 1: Structure for the state space model shows the dependency between the states x_t and the observations Y_t.

in those cases. Despite that the Bayesian approach is still utilized in complex tasks e.g. an interesting application that used the Bayesian approach is searching for the disappearance of flight MH370 as mentioned in [20]. Therefore, in this paper, we will review the Bayesian approach in the KF algorithm, within the field of state estimation for a dynamic linear-Gaussian model where this approach is a powerful tool for dealing with uncertainty, combining the available information and facilitating updated estimations.

The state space models formally consist of two components, the observed process $Y_t = (y_1, \ldots, y_t)$ and the unobserved states $X_t = (x_1, \ldots, x_t)$ satisfying the the two main assumptions:

Assumption 1. $\{X_t, t = 0, 1, 2, \ldots\}$ is a Markov chain i.e. X_t does not depend on the entire past sequences but only on the previous value X_{t-1}.

Assumption 2. The Y_t's are independent observations and conditional on $(X_t,\ t = 0, 1, 2, \ldots)$. It follows that $Y_t|X_t$ has a joint conditional density $\prod_{t=1}^{n} \pi(Y_t|X_t)$, $n \geq 1$. Figure. 1 represents graphically the assumptions of the state space models.

The assumptions (1)and (2) for satisfying the state space model assign the initial density $\pi(X_0)$, the state transition density $\pi(X_t|X_{t-1})$ and the conditional density $\pi(Y_t|X_{t-1})$. We consider the discrete-time state-space model with two equations; the system equation and the observation equation.

The system equation is:

$$X_{t+1} = FX_t + W_t, \tag{1}$$

where, X_t is the state of a variable at time t, denoted as $X_t \in \mathbb{R}^n$, F is a state transition matrix of size $n \times n$ that has the Markovian property which describes the state movements from time $(t-1)$ to t and W_t is the process noise assumed to be Gaussian with zero mean and covariance matrix Q which can be expressed as $W_t \sim \mathcal{N}(0, Q)$.

The observation equation is:

$$Y_t = HX_t + V_t, \tag{2}$$

where $Y_t \in \mathbb{R}^m$ is the measurement vector of the observed data at time t, H is the observation matrix of size $m \times n$ which maps between the measurements and the state and the V_t is the observation noise which is assumed to be Gaussian with zero mean and covariance R. This matrix can be expressed as $V_t \sim \mathcal{N}(0, R)$, and the noise sequences $\{W_t\}$ and $\{V_t\}$ are assumed to be uncorrelated and independent. To begin deriving the KF algorithm, we can describe it as a recursive Bayesian filtering technique based on the Bayes' rule, restricted to linear Gaussian estimation problems only. The main purpose of the Bayesian approach is to obtain the conditional posterior probability of state $\pi(X_t|Y_t)$. To achieve the Bayesian filtering, two stages are required: prediction and update. This approach is successfully applied in the filtering literature, e.g. [21], [22]. The prediction and the correction steps are carried out through the Bayes rule. The one-step-ahead predictive density $\pi(X_t|Y_{t-1})$ can be computed as:

$$\pi(X_t|Y_{t-1}) = \int \pi(X_{t-1}|Y_{t-1})\pi(X_t|X_{t-1})dX_{t-1}. \tag{3}$$

The state transition density $\pi(X_t|X_{t-1})$ is the probability density function (PDF) of the state at time t given the previous state at time $(t-1)$ which satisfies the Markovian property mentioned in the assumption (1). The above equation is called Chapman-Kolmogorov equation [23] and is used to estimate the prior distribution and in some literature *(old posterior)* and the $\pi(X_t|X_{t-1})$ can be estimated from the system equation (1).

The updated PDF of the current state is derived using the Bayes rule:

$$\begin{aligned}\pi(X_t|Y_t) &= \frac{\pi(Y_t|X_t)\pi(X_t|Y_{t-1})}{\pi(Y_t|Y_{t-1})},\\ \pi(X_t|Y_t) &= \frac{\pi(Y_t|X_t)\pi(X_t|Y_{t-1})}{C},\\ \pi(X_t|Y_t) &\propto \pi(Y_t|X_t)\pi(X_t|Y_{t-1}),\end{aligned} \tag{4}$$

where, $C = \pi(Y_1, Y_2, \ldots Y_n)$ is the normalizing constant, and is obtained by:

$$\pi\left(Y_t \mid Y_{t-1}\right) = \int \pi\left(Y_t \mid X_t\right)\pi\left(X_t \mid Y_{t-1}\right)dX_t. \tag{5}$$

The measurement (likelihood) density $\pi\left(Y_t \mid X_t\right)$ in the equation (5) can be computed from the measurement equation (2) and it follows the assumption (2) which states they are conditionally independent given the current state as $Y_t \| Y_{1:t-1} \mid X_t$. According to the equations (4), the last expression summarizes the Bayesian state estimation principle that moves from $(t-1)$ to t where the posterior $\pi(X_t|Y_t)$ contains all the information about the current state X_t by combining the prior distribution and the likelihood density,

Posterior $\propto$ Observed likelihood $\times$ Prior.

In the Gaussian context, it is likely to convert the Bayesian estimation to the point estimation and express the posterior PDF by the mean and covariance which can be computed recursively. Then the general solution of the linear Gaussian problems lead to the KF algorithm. Since the assumptions of the prior probability density $\pi(X_t|Y_{t-1})$ and the likelihood density $\pi(Y_t|Y_{t-1})$ result in Gaussian distribution then the sub-vector, conditional/marginal densities are Gaussian and then the posterior probability density $\pi(X_t|Y_t)$ is also Gaussian. Then we can derive the KF algorithm in the Bayesian scheme as a sequential estimator by computing the means $\mathbb{E}(.)$ and the covariances $\mathrm{Cov}(.)$ of the states over time t for the quantities in the equation (6) as shown in Figure **??**.

$$\underbrace{\pi\left(X_{t-1} \mid Y_{t-1}\right)}_{\text{posterior known at time } (t-1)} \longrightarrow \underbrace{\pi\left(X_t \mid Y_{t-1}\right)}_{\text{prediction at time } t} \longrightarrow \underbrace{\pi\left(X_t \mid Y_t\right)}_{\text{correction at time } t}. \tag{6}$$

Now, we will introduce the conventional KF algorithm which has the following steps:

Theorem 1 (Kalman filter algorithm, [24])**.** *Let $\{H_t, F_t, Q_t, R_t\}$ be a Gaussian dynamic linear model in discrete time satisfying the assumptions (1)and (2). If we have $(X_{t-1}|D_{t-1}) \sim N_p(\hat{X}_{t-1}, \hat{P}_{t-1})$ where $t \geq 1$ and denote the collection of the available information of $Y_1, ..., Y_t$ as D_t. Then,*

1) *The one-step-ahead state predictive density of X_t given Y_{t-1} is a Gaussian with parameters*

$$\pi(X_t|D_{t-1}) \sim N_p(\hat{X}_{t-1}, \hat{P}_{t-1}),$$

where, $\hat{X}_{t-1} = F\hat{X}_{t-1}$ and $\hat{P}_{t-1} = F\hat{P}_{t-1}F^T + Q_t$, for $t \geq 1$.

2) *One-step forecast density of y_t given D_{t-1} is Gaussian with parameters*

$$\pi(y_t|D_{t-1}) \sim N_m(\hat{y}_t, \epsilon_t),$$

where, $\hat{y}_t = H_t\hat{X}_{t-1}$ and $\epsilon_t = H_t\hat{P}_{t-1}H_t^T + R_t$.

3) *The filtering density of X_t given D_t is a Gaussian with parameters*

$$\pi(X_t|D_t) \sim N_p(\hat{X}_t, \hat{P}_t),$$

where $\hat{X}_{t|t} = \hat{X}_{t-1} + \hat{P}_{t-1}F_t^T\epsilon_t e_t$,
$\hat{P}_t = \hat{P}_{t-1} - \hat{P}_{t-1}H_t^T\epsilon_t^{-1}H_t\hat{P}_{t-1}$ with $e_t = y_t - \hat{y}_t$.
where e_t is called the measurement innovation.

Proof. 1) From the system equation (1), taking the expectation value on both sides:

$$\begin{aligned}\mathbb{E}(X_t|D_{t-1}) &= \mathbb{E}(F_tX_{t-1} + V_t|D_{t-1}) \\ &= \mathbb{E}(F_tX_{t-1}|D_{t-1}) + \mathbb{E}(\underbrace{V_t|Y_{t-1}}_{0}) \\ &= \underbrace{F_t\hat{X}_{t-1}}_{\text{prior mean } \hat{X}_{t-1}}.\end{aligned} \tag{7}$$

$\mathbb{E}(\xi_t|Y_{t-1}) = 0$ for all t as the process noise term does not correlate with the measurement noise. Now, taking the covariance value for both sides:

$$\begin{aligned}\mathrm{Cov}(X_t|D_{t-1}) &= \mathrm{Cov}(F_tX_{t-1} + V_t|D_{t-1}) \\ &= \mathrm{Cov}(F_tX_{t-1}|D_{t-1}) + \mathrm{Cov}(V_t|D_{t-1}) \\ &= \underbrace{F_t\hat{P}_{t-1}F_t^T + Q_t}_{\text{prior covariance } \hat{p}_{t-1}}.\end{aligned} \tag{8}$$

At this stage, we compute the KF prediction of the state at time t conditional in the time $(t-1)$, and in the next steps, we compute the KF update equations.

2) From the observation equation (2), taking the expectation value on both sides:

$$\begin{aligned}\mathbb{E}(Y_t|D_{t-1}) &= \mathbb{E}(H_tX_t + W_t|D_{t-1}) \\ &= \underbrace{H_t\hat{X}_{t|t-1}}_{\text{mean forecasting}}.\end{aligned} \tag{9}$$

Now taking the covariance value for the observation equation on both sides as:

$$\begin{aligned}\mathrm{Cov}(Y_t|D_{t-1}) &= \mathrm{Cov}(H_tX_t + W_t|D_{t-1}) \\ &= \underbrace{H_t\hat{P}_{t|t-1}H_t^T + R_t}_{\text{variance forecasting}}.\end{aligned} \tag{10}$$

3) Using Bayes formula to compute $\pi(X_t|D_t)$ by taking $\pi(X_t|D_{t-1})$ as the prior distribution and $\pi(y_t|D_{t-1})$ as the likelihood function, we have:

$$\pi(X_t|D_t) \propto \pi(X_t|D_{t-1})\pi(Y_t|D_{t-1}). \tag{11}$$

Now substitute the means and variances from equations (7), (8),(9), (10) into equation (12) which is the multivariate Gaussian PDF as:

$$\begin{aligned}\pi(X_t|D_t) \propto\, & e^{\left\{\frac{-1}{2}(X_t - F_t\hat{X}_{t-1})^T(F_t\hat{P}_{t-1}F_t^T + Q)^{-1}(X_t - F_t\hat{X}_{t-1})\right\}} \\ & e^{\left\{\frac{-1}{2}(Y_t - H_t\hat{X}_{t|t-1})^T(H_t\hat{p}_{t|t-1}H_t^T + R_t)^{-1}(Y_t - H\hat{X}_{t|t-1})\right\}}.\end{aligned} \tag{12}$$

After algebraic manipulation, the posterior distribution of $\pi(X_t|D_t)$ will be obtained in the Gaussian distribution. The updated posterior mean $\hat{X}_t$ can be written as:

$$\begin{aligned}\hat{X}_t = \hat{X}_{t|t-1} + & \underbrace{\hat{P}_{t|t-1}H_t^T(H_t\hat{P}_{t|t-1}H_t^T + R_t)^{-1}}_{K_t} \\ & \underbrace{(Y_t - H_t\hat{X}_{t|t-1})}_{e_t},\end{aligned} \tag{13}$$

where, the term $\hat{P}_{t|t-1}H_t^T(H_t\hat{P}_{t|t-1}H_t^T + R_t)^{-1}$ is the weight of the correction term by the gain matrix called the Kalman gain matrix K_t and it is $0 \leq K_t \leq 1$, which adjusts the estimation process between the predicted state and the measurements based on their uncertainties in the correction step to minimize the estimation error. The term e_t presents the measurement errors which are also called innovations and for a sequence of $\{e_t\}$, $t \geq 1$ this error has an expected value equal to zero and the error sequences e_t are uncorrelated independent across time steps [24].

The updated posterior covariance $\hat{p}_t$ is expressed as:

$$\hat{P}_t = \hat{P}_{t-1} + \hat{P}_{t-1}H_t^T\{H_t\hat{P}_{t-1}H_t^T + R_t\}^{-1}H_t\hat{P}_{t-1}. \tag{14}$$

Summarizing the KF equations as:

- The posterior mean is

$$\hat{X}_t = \hat{X}_{t-1} + K_te_t, \tag{15}$$

which gives the KF estimate of the state.

- The Kalman gain is

$$K_t = \hat{P}_{t|t-1} H_t^T (H_t \hat{P}_{t|t-1} H_t^T + R_t)^{-1}. \quad (16)$$

- The posterior covariance is

$$\hat{P}_t = \hat{P}_{t|t-1} + K_t H_t \hat{P}_{t|t-1}, \quad (17)$$

which gives the covariance of the KF state estimate.

□

III. Example of Gaussian Dynamic Linear Model as the Univariate Random Walk Plus Noise Model

This section presents an example of a dynamic linear model called the random walk plus noise model. Then we will describe the recursive steps for computing the Kalman filter algorithm where the assumptions of linearity and Gaussianity are held. The random walk plus noise model [25], [26], also called the local level model as presented in [24] and the steady model as called in [27] is a simple polynomial model used to represent a time series that has random steps influenced by the assumed noise. This model is usually used in the case where the rate of change in the process is stable or the system has gradual drift with consideration of noisy measurements. The random walk model is applied in many applications in natural sciences such as anthropological research in [28], to estimate the true level of the Nile river in [24], to model the stock prices in [29]. The random walk plus noise model is appropriate for a short time interval where the process has a clear trend where changes such as persistent growth or decline are either absent or limited. Consequently, the model may not be suitable for a long-term behaviour estimate or high data variation where the main objective is to capture the current level/state within a short time frame. This model can be expressed and analysed in the univariate case, in a simpler way.

The observation equation is

$$y_t = x_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad (18)$$

where, y_t is the univariate observation, x_t is the random change in the level at time t and the ε_t is a measurement random error at time t Gaussian distributed with zero mean and covariance σ_ε^2. The corresponding system equation is

$$x_t = x_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \quad (19)$$

where, η_t is a random system disturbance at time t, Gaussian distributed with zero mean and covariance σ_η^2. The error components $\{\varepsilon_t\}$ and $\{\eta_t\}$ for all t, are both independent within them and between them. The dynamic linear representation of the random walk plus noise has $m = n = 1$, $F_t = 1$, $H_t = 1$, $R_t = \sigma_\varepsilon^2$ and $Q_t = \sigma_\eta^2$. This is the simple form of the model and may contains extra terms such as the trend, expenditure or seasonality according to the application of interest.

IV. Kalman Filter Derivations for the Random Walk Plus Noise Model

Now we can illustrate the KF theory for the random walk noise model in the above example using the Bayesian approach.

- Initialize the state x_0 with mean $\hat{x}_0$ and covariance $\hat{p}_0$.
- Define the posterior distribution for the random walk noise model as $\pi(x_{t-1}|D_{t-1})$ with expected value equal to $\hat{x}_{t-1}$ and covariance equal to $\hat{p}_{t-1}$. Then:

$$\pi(x_{t-1}|D_{t-1}) \sim \mathcal{N}(\hat{x}_{t-1}, \hat{p}_{t-1}). \quad (20)$$

- Applying Bayes rule to estimate the posterior estimate of the state x_t we get:

$$\pi(x_t|D_t) \propto \pi(x_t|D_{t-1}) \pi(y_t|x_t). \quad (21)$$

To find the $\pi(y_t|x_t)$ and $\pi(x_t|D_{t-1})$ recall the system and measurement equations of the random walk noise model in equation (18) and equation (19). First, compute the conditional mean and the covariance for the posterior density $\pi(x_{t-1}|D_{t-1})$ from the system equation (19) as:

$$\begin{aligned} \mathbb{E}(x_t|D_{t-1}) &= \mathbb{E}(x_{t-1} + \eta_t|D_{t-1}) \\ &= \mathbb{E}(x_{t-1}|D_{t-1}) + \underbrace{\mathbb{E}(\eta_t|D_{t-1})}_{0} \\ &= \hat{x}_{t-1}. \\ \mathbb{V}(x_t|D_{t-1}) &= \mathbb{V}(x_{t-1} + \eta_t|D_{t-1}) \\ &= \mathbb{V}(x_t|D_{t-1}) + \mathbb{V}(\eta_t|D_{t-1}) \\ &= \hat{p}_{t-1} + \sigma_\eta^2, \\ &= r_t. \end{aligned} \quad (22)$$

Then the Gaussian probability distribution of the prior density $\pi(x_t|D_{t-1})$ is

$$\begin{aligned} &\pi(x_t|D_{t-1}) \sim \mathcal{N}(\hat{x}_{t-1}, r_t), \\ &\pi(x_t|D_{t-1}) \propto e^{-\frac{1}{2r_t}(x_t - \hat{x}_{t-1})^2}. \end{aligned} \quad (23)$$

Now find the probability distribution of the predictive density for the observations $\pi(y_t|x_t)$ from the observation equation (18) as:

$$\begin{aligned} \mathbb{E}(y_t|x_t) &= \mathbb{E}(x_t + \epsilon_t|x_t), \\ &= \mathbb{E}(x_t|x_t) + \underbrace{\mathbb{E}(\epsilon_t|x_t)}_{0}, \\ &= x_t, \\ \mathbb{V}(y_t|x_t) &= \mathbb{V}(x_t + \epsilon_t|x_t), \\ &= \mathbb{V}\underbrace{(x_t|x_t)}_{0} + \mathbb{V}(\epsilon_t|x_t), \\ &= \sigma_\epsilon^2. \end{aligned} \quad (24)$$

Then the Gaussian probability distribution of the prior density $\pi(x_t|D_{t-1})$ is

$$\begin{aligned} &\pi(y_t|x_t) \sim \mathcal{N}(x_t, \sigma_\epsilon^2) \\ &\pi(y_t|x_t) \propto e^{-\frac{1}{2\sigma_\epsilon^2}(y_t - x_t)^2}. \end{aligned} \quad (25)$$

Now substituting $\pi(x_t|D_{t-1})$ and $\pi(y_t|x_t)$ in the equation (21) we get:

$$\pi\left(x_t \mid D_t\right) \propto \exp\left\{-\frac{1}{2}\frac{\left(y_t-x_t\right)^2}{\sigma_\varepsilon^2}+\frac{\left(x_t-\hat{x}_{t-1}\right)^2}{r_t}\right\}. \tag{26}$$

Taking the natural logarithm and multiplying both sides of the equation (26) by -2, we get:

$$-2\ln\pi\left(x_t \mid D_t\right) \propto \frac{\left(y_t-x_t\right)^2}{\sigma_\varepsilon^2}+\frac{\left(x_t-\hat{x}_{t-1}\right)^2}{r_t}, \tag{27}$$

$$-2\ln\pi\left(x_t \mid D_t\right) \propto \frac{y_t^2}{\sigma_\varepsilon^2}-2\frac{y_t x_t}{\sigma_\varepsilon^2}+\frac{x_t^2}{\sigma_\varepsilon^2}+\frac{x_t^2}{r_t}-2\frac{\hat{x}_{t-1}x_t}{r_t}+\frac{x_{t-1}^2}{r_t} \tag{28}$$

$$= \text{constant}+x_t^2\left(\frac{1}{\sigma_\varepsilon^2}+\frac{1}{r_t}\right)-2\left(\frac{\hat{x}_{t-1}}{r_t}+\frac{y_t}{\sigma_\varepsilon^2}\right)x_t. \tag{29}$$

Now put,

$$A_1=\frac{1}{\sigma_\varepsilon^2}+\frac{1}{r_t} \quad \text{and} \quad A_2=\frac{\hat{x}_{t-1}}{r_t}+\frac{y_t}{\sigma_\varepsilon^2}. \tag{30}$$

$$\begin{aligned} A_1 &= \frac{\sigma_\varepsilon^2+r_t}{\sigma_\varepsilon^2 r_t} \\ &= \hat{p}_t^{-1}, \end{aligned} \tag{31}$$

and

$$\begin{aligned} A_2 &= \frac{\hat{x}_{t-1}}{r_t}+\frac{y_t}{\sigma_\varepsilon^2}+\frac{\hat{x}_{t-1}}{\sigma_\varepsilon^2}-\frac{\hat{x}_{t-1}}{\sigma_\varepsilon^2}, \\ &= \hat{x}_{t-1}\left(\frac{\sigma_\varepsilon^2+r_t}{\sigma_\varepsilon^2 r_t}\right)+\frac{1}{\sigma_\varepsilon^2}\left(y_t-\hat{x}_{t-1}\right). \end{aligned} \tag{32}$$

$$\begin{aligned} &= \frac{\sigma_\varepsilon^2+r_t}{\sigma_\varepsilon^2 r_t}\left(\hat{x}_{t-1}+\left(\frac{r_t}{(\sigma_\varepsilon^2+r_t)}\right)\left(y_t-\hat{x}_{t-1}\right)\right), \\ &= \hat{p}_t^{-1}\left(\hat{x}_{t-1}+\frac{r_t}{(\sigma_\varepsilon^2+r_t)}\left(y_t-\hat{x}_{t-1}\right)\right), \\ &= \hat{p}_t^{-1}\hat{x}_t, \end{aligned} \tag{33}$$

where, $\hat{x}_t$ is the mean posterior which is given by:

$$\hat{x}_t=\hat{x}_{t-1}+\frac{r_t}{(\sigma_\varepsilon^2+r_t)}\left(y_t-\hat{x}_{t-1}\right). \tag{34}$$

Now, substitute the A_1 and A_2 values in the Eq. 29

$$-2\ln\pi\left(x_t \mid D_t\right)=\text{constant}+p_t^{-1}x_t^2-2\hat{p}_t^{-1}\hat{x}_t x_t. \tag{35}$$

where the constant term does not contain the parameter of interest.
After subtracting and adding the term $p_t^{-1}\hat{x}_t^2$ into equation (35), we will have

$$-2\ln\pi\left(x_t \mid D_t\right) \propto \hat{p}_t^{-1}(x_t-\hat{x}_t)^2, \tag{36}$$

Thus,

$$\pi(x_t|D_t) \propto e^{-\frac{1}{2\hat{p}_t}(x_t-\hat{x}_t)}. \tag{37}$$

We can express the KF equations for the random walk plus noise model as the following:
The updated mean is $\hat{x}_t$ is

$$\hat{x}_t=\hat{x}_{t-1}+k_t(y_t-\hat{x}_{t-1}). \tag{38}$$

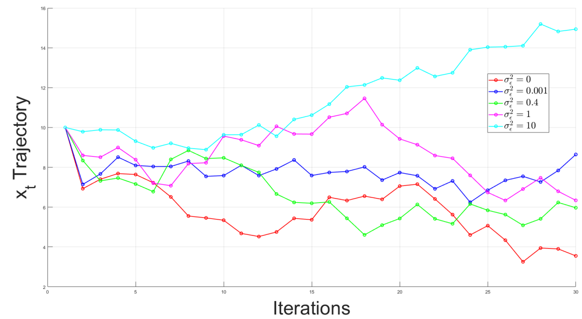

Fig. 2: The random walk model approach with different values of the noise measurements covariance σ_ϵ^2.

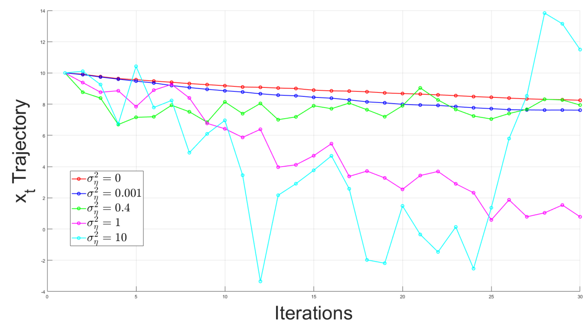

Fig. 3: The random walk model approach with different values of the covariance noise system σ_η^2.

The Kalman gain k_t is

$$k_t=\frac{r_t}{\sigma_\varepsilon^2+r_t}. \tag{39}$$

The updated covariance $\hat{p}_t$ is

$$\hat{p}_t=\sigma_\varepsilon^2 k_t. \tag{40}$$

We have seen how we obtained the updated posterior PDF in a Gaussian distribution after successive iterations through the tractable Bayesian approach, which makes the KF algorithm an optimal solution for the linear Gaussian problems. It is worth mentioning that the choice between Bayesian and Frequentist approaches in the KF derivations depends on the availability of prior knowledge and the nature of the phenomena e.g. linear, nonlinear, Gaussian, and non-Gaussian. The Bayesian approach is often flexible in linear case and automatically capture the uncertainty which is helpful for a wider range of situations. But it requires the specification of prior distributions which is unlikely in the frequentist methods.

V. Simulation Results of Kalman Filtering with Random Walk Plus Noise Model

In this section, we conduct a simulation over 30 iterations that utilises arbitrary values to practically investigate the effect of noise in the filtering and convergence processes. We examine the noise in the system and the observations, along with the behaviour of the Kalman gain k_t and its implications in the random walk plus noise model. The Kalman gains stabilise the estimation process and mitigate the effect of noise and disturbances in either the measurements or the system. Thus tuning and understanding of these parameters is essential to achieve good performance of the filter algorithm as also mentioned in [5].

A. Understanding the Impact of Varying the Observation Noise Variance on the Model

We examine the impact of different levels of uncertainty in the observation noise on the accuracy and stability of the system which could be stock prices, weather forecasting, or any other measurable quantity. From Figure 2 even in the absence of covariance $\sigma_\epsilon^2 = 0$ or small value $\sigma_\epsilon^2 = 0.001$ the system responds to the changes and quantifies the trend/random components with repeats quite regularly in its behaviour over time. However, increasing the variance σ_ϵ^2 does not have a significant impact on the model as we see the $\sigma_\epsilon^2 = 1, 10$ which means the small value of σ_ϵ^2 that added it can be adequate.

B. Understanding the Impact of Varying the System Noise Covariance on the Model

In this case, we assume the system is perfect with no error. The model shows more sensitivity to the change of the system variance σ_η^2. From Figure 3 with the absence of the variance and the small value the system remains constant over time and the model will not be able to capture the patterns in these values which may be the invalid case. In contrast, with σ_η^2 =10, the model has a complex pattern which adds more flexibility to capture a complex time series. Then, this suggests the influence of the σ_η^2 is important and should be adjusted to have appropriate predictive outcomes.

C. Understanding the Impact of Varying the Kalman Gain on the Model

The Kalman gain k_t fundamentally can be used to stabilize the filter, to improve the initialization and determine the level of confidence on the predicted state estimate versus the available measurements since it is a ratio of the covariance matrices of the measurement error and the state estimate error. The Kalman gain k_t ranges between 0 and 1 and then gradually converges to a constant value in for convergence. Otherwise, if the Kalman gain is not stable, then it indicates improper initialization of error covariances leading to poor performance of the filter. A Kalman gain value closer to 0 indicates that more weight is given to the predicted state estimate, the uncertainty in the measurements is very high and the updating process is primarily driven by the model. Then, the estimates do not change from one step to the next step meaning $x_t = x_{t-1}$. A Kalman gain value closer to 1 indicates that more weight is given to the observed measurement, and the update is primarily driven by the new measurement information. The simulation evolves recursively by applying the KF equations with fixed initial mean and covariance as $x_0 = 10$, $p_0 = 0.02$.

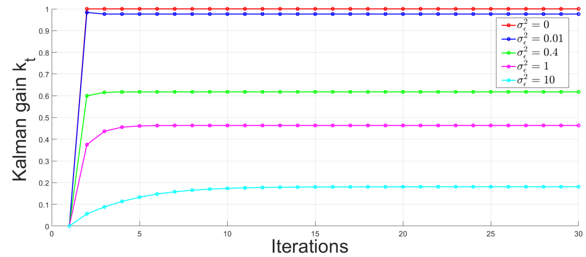

Fig. 4: Kalman gain convergence with different values of σ_ϵ^2 and fixed value $\sigma_\eta^2 = 0.4$.

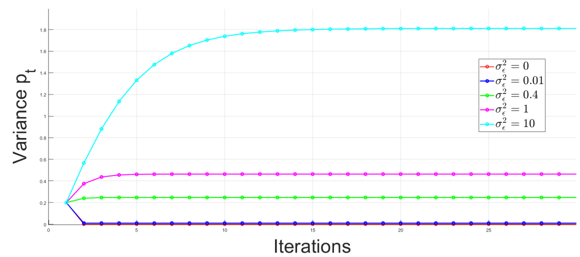

Fig. 5: Variance p_t trajectory that associated with variability values of of σ_ϵ^2.

We conducted a simulation of the KF to investigate the Kalman gain trajectory with different values of covariance σ_ϵ^2 as shown in Figure 4. A Kalman gain value for this case presented in Figure 4 is approximately close to 1 with σ_ϵ^2 = 0 and 0.01 which quickly stabilises at $t = 1$ leading to the steady state with reduced covariance estimate p_t in Figure 5 with a value between 0 and 0.000618. This means that the filter is trying to make a balance between the predicted state and the measurements which is a sign that the convergence to the true state and k_t gives more weight to the current measurement. By dynamically adapting to these changing conditions, the Kalman gain enhances the filter's ability to accurately estimate the true state. In Figure 6, we test the same range of σ_η^2 values that were used in model 3 where we fixed the $\sigma_\epsilon^2 = 0.40$. We obtained a small Kalman gain for $\sigma_\eta^2 = 0, 0.01$ over time which takes the value of $k_t = 0.020408$ with variance p_t value approximately to 0.00816 as shown in Figure 7. Moreover, the Kalman gain in approach to model 1 with σ_η^2 =10 with the largest value of variance p_t approximately equal to 0.39 as shown in Figure 7. To conclude this simulation we highlighted the behaviour of the Kalman gain and observed the adaptability of the Kalman gain which is a key feature that allows the KF to achieve accurate state estimation and suggests where the filter can be adjusted. It's important to note that, through the visualization, the Kalman gain is influenced by different factors such as the uncertainties in the process and measurement noise and if we increase the noises will have a non-stationary variance but the Kalman gain still might be able to succeed in stabilizing the variance over time.

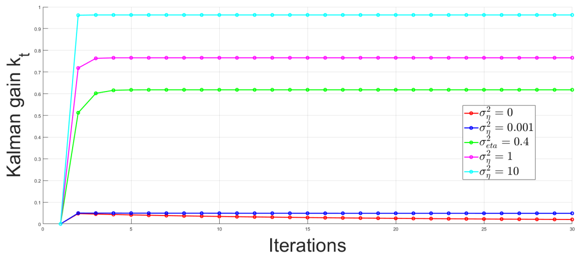

Fig. 6: Kalman gain convergence with different values of σ_η^2 and fixed value $\sigma_\epsilon^2 = 0.40$.

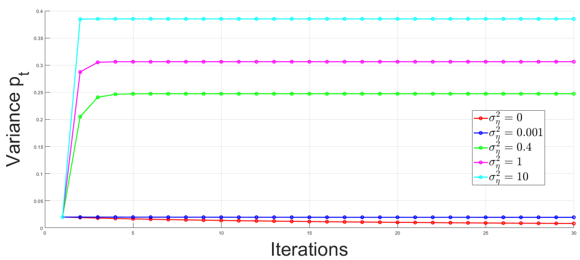

Fig. 7: Variance p_t trajectory that associated with variability values of of σ_η^2.

VI. Conclusion

This paper presents the random walk plus noise model in the univariate case with derived KF equations. The performances of the model in the absence and presence of the uncertainties are addressed and visualized. We show the random walk plus noise model is capable of capturing irregularities and jumps, but it needs further effort to adjust the variance values. Understanding the Kalman Gain's role in the Kalman filter helps in solving complex problems. Furthermore, the Kalman gain's performance in different scenarios have been visualized and aligned with the trajectories of the variance to minimize the mean square error which is a useful indicator of how good the filtering algorithm works. All The interpretations of the underlying system are synthetic, in the future, we will use the model to estimate real-time applications in epidemiology, robotics, navigation, engineering/natural systems [30], [31].

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[2] D. Simon, *Optimal state estimation: Kalman, H_∞, and nonlinear approaches*. John Wiley & Sons, 2006.

[3] S. So and K. K. Paliwal, "Suppressing the influence of additive noise on the kalman gain for low residual noise speech enhancement," *Speech Communication*, vol. 53, no. 3, pp. 355–378, 2011.

[4] R. V. Soares, C. Maschio, and D. J. Schiozer, "Applying a localization technique to kalman gain and assessing the influence on the variability of models in history matching," *Journal of Petroleum Science and Engineering*, vol. 169, pp. 110–125, 2018.

[5] A. Yadav, N. Naik, M. Ananthasayanam, A. Gaur, and Y. Singh, "A constant gain kalman filter approach to target tracking in wireless sensor networks," in *2012 IEEE 7th International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 2012, pp. 1–7.

[6] Y. Zhang and D. S. Oliver, "Improving the ensemble estimate of the kalman gain by bootstrap sampling," *Mathematical Geosciences*, vol. 42, pp. 327–345, 2010.

[7] S. Wang, J. Feng, and K. T. Chi, "Analysis of the characteristic of the kalman gain for 1-d chaotic maps in cubature kalman filter," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 229–232, 2013.

[8] B. Noack, F. Pfaff, and U. D. Hanebeck, "Optimal kalman gains for combined stochastic and set-membership state estimation," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 4035–4040.

[9] Y. Zhang and D. S. Oliver, "Evaluation and error analysis: Kalman gain regularization versus covariance regularization," *Computational Geosciences*, vol. 15, pp. 489–508, 2011.

[10] B. Boulkroune, K. Geebelen, J. Wan, and E. van Nunen, "Auto-tuning extended kalman filters to improve state estimation," in *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2023, pp. 1–6.

[11] A. S. Tummala and P. Ramanarao, "Tuning of extended kalman filter for power systems using two lbest particle swarm optimization," *International Journal Of Control Theory and Applications*, vol. 10, no. 5, pp. 197–206, 2017.

[12] L. Zhang, D. Sidoti, A. Bienkowski, K. R. Pattipati, Y. Bar-Shalom, and D. L. Kleinman, "On the identification of noise covariances and adaptive kalman filtering: A new look at a 50 year-old problem," *IEEE Access*, vol. 8, pp. 59 362–59 388, 2020.

[13] J. Duník, O. Straka, O. Kost, and J. Havlík, "Noise covariance matrices in state-space models: A survey and comparison of estimation methods—part i," *International Journal of Adaptive Control and Signal Processing*, vol. 31, no. 11, pp. 1505–1543, 2017.

[14] L. Alyami, D. K. Panda, and S. Das, "Bayesian noise modelling for state estimation of the spread of covid-19 in saudi arabia with extended kalman filters," *Sensors*, vol. 23, no. 10, p. 4734, 2023.

[15] Y. Ho and R. Lee, "A bayesian approach to problems in stochastic estimation and control," *IEEE Transactions on Automatic Control*, vol. 9, no. 4, pp. 333–339, 1964.

[16] V. Peterka, "Bayesian approach to system identification," in *Trends and Progress in System identification*. Elsevier, 1981, pp. 239–304.

[17] Y.-X. Lin, "An alternative derivation of the kalman filter using the quasi-likelihood method," *Journal of Statistical Planning and Inference*, vol. 137, no. 5, pp. 1627–1633, 2007.

[18] J. Humpherys and J. West, "Kalman filtering with newton's method [lecture notes]," *IEEE Control Systems Magazine*, vol. 30, no. 6, pp. 101–106, 2010.

[19] R. Gurajala, P. B. Choppala, J. S. Meka, and P. D. Teal, "Derivation of the kalman filter in a bayesian filtering perspective," in *2021 2nd International Conference on Range Technology (ICORT)*. IEEE, 2021, pp. 1–5.

[20] S. Davey, N. Gordon, I. Holland, M. Rutten, and J. Williams, *Bayesian Methods in the Search for MH370*. Springer Nature, 2016.

[21] J. Spragins, "A note on the iterative application of bayes' rule," *IEEE Transactions on Information Theory*, vol. 11, no. 4, pp. 544–549, 1965.

[22] R. S. Bucy, "Bayes theorem and digital realizations for non-linear filters," *Journal of the Astronautical Sciences*, vol. 17, p. 80, 1969.

[23] A. Jazwinski, "Stochastic processes and filtering theory. vol. 64 of math," *Science and Engineering. Academic Press. New York*, 1970.

[24] G. Petris, S. Petrone, and P. Campagnoli, "Dynamic linear models," in *Dynamic Linear Models with R*. Springer, 2009, pp. 31–84.

[25] R. N. Mantegna and H. E. Stanley, *Introduction to econophysics: correlations and complexity in finance*. Cambridge University Press, 1999.

[26] M. West and J. Harrison, *Bayesian forecasting and dynamic models*. Springer Science & Business Media, 2006.

[27] P. J. Harrison and C. F. Stevens, "Bayesian forecasting," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 38, no. 3, pp. 205–228, 1976.

[28] C. Perreault, "Time-averaging slows down rates of change in the archaeological record," *Journal of Archaeological Method and Theory*, vol. 25, pp. 953–964, 2018.

[29] P. Chen, "A random walk or color chaos on the stock market? time-frequency analysis of s&p indexes," *Studies in Nonlinear Dynamics & Econometrics*, vol. 1, no. 2, 1996.

[30] M. Zeinali and M. Shafiee, "A new kalman filter based 2d ar model parameter estimation method," *IETE Journal of Research*, vol. 63, no. 2, pp. 151–159, 2017.

[31] U. Iqbal, J. Georgy, M. J. Korenberg, and A. Noureldin, "Nonlinear modeling of azimuth error for 2d car navigation using parallel cascade identification augmented with kalman filtering," *International Journal of Navigation and Observation*, vol. 2010, 2010.