2022 American Control Conference (ACC)
Atlanta, USA, June 8-10, 2022

Optimal Moving Average Estimation of Noisy Random Walks using Allan Variance-informed Window Length

Hossein Haeri[1], Behrad Soleimani[2], and Kshitij Jerath[3]

Abstract— **Moving averages are widely used to estimate time-varying parameters, especially when the underlying dynamic model is unknown or uncertain. However, the selection of the optimal window length over which to evaluate the moving averages remains an unresolved issue in the field. In this paper, we demonstrate the use of Allan variance to identify the *characteristic timescales* of a noisy random walk from historical measurements. Further, we provide a closed-form, analytical result to show that the Allan variance-informed averaging window length is indeed the optimal averaging window length in the context of moving average estimation of noisy random walks. We complement the analytical proof with numerical results that support the solution, which is also reflected in the authors' related works. This systematic methodology for selecting the optimal averaging window length using Allan variance is expected to widely benefit practitioners in a diverse array of fields that utilize the moving average estimation technique for noisy random walk signals.**

I. Introduction

Despite the vast investigation in advanced data filtering techniques, Simple Moving Average Estimations (SMAE) are still among the widely used filtering solutions in many fields of research such as trend forecasting [1] communication systems [2], econometrics [3], circuit design [4], cardiac signal analysis [5] and, general data monitoring [6]. Its widespread adoption is primarily driven by its ease of use (as a linear low-pass filter with a constant computational complexity), and often by the inability to use model-based techniques such as the Kalman filtering which need prior knowledge of the signal and noise characteristics as well as a reliable, high-fidelity dynamical model in the specific domain [7]. Moreover, simple moving average filters are easy to interpret which makes them an excellent candidate for visualizing noisy data.

While the concepts of moving averages and MA estimation have been around since the early twentieth century, the challenge of systematically selecting an appropriate timescale or window length over which to average measurements persists to this day. While considering longer windows deliver smoother results, it introduces a significant amount of lag between the input data and the estimate. Conversely, choosing a short window, despite delivering a low-lag signal, may not be effective enough for noise-cancellation purposes. We previously addressed this trade-off by introducing the *characteristic timescale* as the window length that includes the most relevant measurements or information for a given estimation task, where the concept of data relevancy is determined by its positive/negative effect on the expected estimation error at the very recent timestep. We also showed this timescale can be estimated by finding the timescale that minimizes the Allan Variance (AVAR) [8].

In this work, we focus on estimating noisy random walk signals which are typical building blocks for stochastic drift models across a wide spectrum of scientific disciplines. For example, financial markets use random walks to model changes in stock prices according to the random walk theory [9]. Random walk models are also vastly used in modeling diffusion processes [10] [11] that in turn has many applications in biology [12] geology [13], and hydrology [14]. In this paper, we first derive a closed-form solution for the optimal window length that should be used to filter noisy random walk signals using simple moving average estimation (SMAE). We then propose the use of Allan variance to estimate this window length without having any prior knowledge about the signal or noise properties. We also prove the timescale suggested by AVAR is indeed an unbiased estimate of the optimal timescale (i.e. averaging window length).

II. Simple Moving Average Estimation of Random Walks

As a first step towards discussing the optimal moving average estimation, we formulate the simple moving average estimation problem. In this context, the goal is to estimate the value of a time-varying parameter based on historical measurements or an incoming data stream. First, the time-varying dynamics of the target parameter may be modeled as a random walk stochastic process $\{X_k\}$. For practical purposes of this approach, we may assume that the incremental evolution of the random walk is obtained via numerical integration of Gaussian white noise process $\{W_k\}$ with zero mean and variance σ_w^2. The dynamics of the time-varying parameter are thus modeled as:

$$x_k = x_{k-1} + w_k \tag{1}$$

where w_k are realized from $\{W_k\}$, i.e. $w_k \sim \mathcal{N}(0, \sigma_w^2)$. Second, the measurements of the time-varying parameter are themselves assumed to be corrupted by Gaussian white noise $\mathcal{N}(0, \sigma_v^2)$, i.e. with mean zero and variance σ_v^2, and is

[1]Hossein Haeri is with Department of Mechanical Engineering University of Massachusetts Lowell, 1 University Ave, Lowell, MA 01854 hossein_haeri@uml.edu

[2]Behrad Soleimani is with Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20740 behrad@umd.edu

[3]Kshitij Jerath is with Department of Mechanical Engineering University of Massachusetts Lowell, 1 University Ave, Lowell, MA 01854 kshitij_jerath@uml.edu

978-1-6654-5196-3/$31.00 ©2022 AACC

independent of the process $\{W_k\}$. Thus, the measurements of the random walk signal are given by:

$$y_k = x_k + v_k \tag{2}$$

where v_k are realized from a white noise process $\{V_k\}$ with the variance σ_v^2.

The Simple Moving Average Estimation (SMAE) attempts to recover the value of the parameter x_n, given historical measurements from the time series data $\boldsymbol{y} = \{y_1, y_2, ..., y_n\}$. We may note that it is not a requirement that the data be varying over the time dimension, the same approach is equally applicable for estimating noisy random walk signals that vary across space or any other valid underlying attribute of interest [13]. The simple moving average estimate of the time-varying parameter is given by:

$$\hat{x}_n^{(m)} = \frac{1}{m}\sum_{i=1}^{m} y_{n-m+i} \tag{3}$$

where $\hat{x}_n^{(m)}$ is the estimate of parameter x obtained at time step n using the past m measurements, i.e. $m \in \mathbb{Z}^+ : m \leq n$ is the window length (or timescale) over which the historical measurements are averaged. We assume that information is available sequentially (ordered in time, space, or other valid attributes) and that the parameter x is estimated for subsequent time steps as well when new measurements arrive. In the work presented below, we focus only on causal moving average estimators which do not rely on future measurements to estimate the current value of the parameter. In principle, however, the approach presented below can easily be used with limited modification for the purpose of non-causal filtering as well.

III. Optimal Simple Moving Average Estimation

The estimation approach of SMAE is straightforward and has been used in the past for real-time data-driven estimation, especially in the absence of a viable dynamical model. However, there is a lack of clarity on how to choose the optimal averaging window length in a manner that minimizes the error in the estimate. Here we refer to the optimal averaging window length as the *characteristic timescale*) and denote it by m_c. Specifically, we may determine the characteristic timescale by minimizing the mean-square estimation error:

$$m_c = \underset{m \in Z^+}{\operatorname{argmin}}\ \mathbb{E}\left[(\hat{x}_n^{(m)} - x_n)^2\right] \tag{4}$$

We drop the $\mathbb{Z}^+$ notation with the understanding that m is an integer-valued positive number with $m \leq n$. The analytical expression for m_c is obtained as shown in Thm. 1.

Theorem 1. *If the drift signal x_k is realized from a random walk process $\{X_k\}$ (shown in (1)) and the measurement signal is realized according to $\{X_k + V_k\}$ (shown in (2)), then the characteristic timescale m_c is given by:*

$$m_c = \left\lfloor \sqrt{3\frac{\sigma_v^2}{\sigma_w^2} + \frac{1}{2}} \right\rceil \tag{5}$$

where $\lfloor . \rceil$ denotes rounding to the adjacent integer with the smaller estimation error in (11).

Proof. We first substitute $\hat{x}_n^{(m)}$ from (3) into (4). For simplicity, we will omit the superscript (m) in $\hat{x}_n^{(m)}$ from here onwards. Thus, the minimization problem to find the characteristic timescale (optimal averaging window length) may now be written as:

$$m_c = \underset{m}{\operatorname{argmin}}\ \mathbb{E}\left[\left\{\left(\frac{1}{m}\sum_{i=1}^{m} y_{n-m+i}\right) - x_n\right\}^2\right] \tag{6}$$

Using the random walk model in (1) and assuming $x_1 = 0$ (without loss of generality), we can rewrite x_n as the sum of the individual w_i terms until time step n, i.e. $x_n = \sum_{i=1}^{n} w_i$. Together with the measurement equation given by (2), we can expand the summation term on the right-hand side of (6) and write the following expression for the expectation term:

$$\mathbb{E}\left[\left(\frac{1}{m}\sum_{i=1}^{m} x_{n-m+i} + \frac{1}{m}\sum_{i=1}^{m} v_{n-m+i} - \sum_{i=1}^{n} w_i\right)^2\right]$$

We may reasonably assume that the v_k terms realized from the white noise process $\{V_k\}$ are independently and identically distributed, and are also independent of the stochastic processes $\{W_k\}$ (and consequently $\{X_k\}$). Thus, by considering that $\mathbb{E}[((1/m)\sum_{i=1}^{m} v_{n-m+i})^2] = (1/m)\,\sigma_v^2$, the expression for the expectation term can be simplified to:

$$\mathbb{E}\left[\left(\frac{1}{m}\sum_{i=1}^{m} x_{n-m+i} - \sum_{i=1}^{n} w_i\right)^2\right] + \frac{\sigma_v^2}{m} \tag{7}$$

Furthermore, accounting for the fact that the w_i terms ($i \in [1, n-m]$) occur in all subsequent x_i terms ($i \in [n-m+1, n]$), the mean of the most recent m parameter values can be re-written as follows:

$$\frac{1}{m}\sum_{i=1}^{m} x_{n-m+i} = \sum_{i=1}^{n-m} w_i + \sum_{i=n-m+1}^{n} \frac{n-i+1}{m} w_i \tag{8}$$

The expression on the right-hand side of (8) can be substituted into the expression in (7) to yield:

$$\mathbb{E}\left[\left(\sum_{1}^{n-m} w_i + \sum_{i=n-m+1}^{n} \frac{n-i+1}{m} w_i - \sum_{i=1}^{n} w_i\right)^2\right] + \frac{\sigma_v^2}{m} \tag{9}$$

which, upon simplification, leads to the following:

$$m_c = \underset{m}{\operatorname{argmin}}\ \mathbb{E}\left[\left(\sum_{i=1}^{m} \frac{m-i}{m} w_i\right)^2\right] + \frac{\sigma_v^2}{m} \tag{10}$$

Since w_i are independently and identically distributed, $\mathbb{E}[w_i w_j] = 0$ for $i \neq j$. Consequently, using the expression for sum of squares of first $(m-1)$ natural numbers, the above term can be calculated as:

$$\mathbb{E}\left[(\hat{x}_n - x_n)^2\right] = \frac{2m^2 - 3m + 1}{6m}\ \sigma_w^2 + \frac{1}{m}\sigma_v^2 \tag{11}$$

To minimize (11), we let $\frac{\partial}{\partial m}\,\mathbb{E}[(\hat{x}_k - x_k)^2] = 0$. Although this is a discrete optimization problem to be solved for positive integer values of m, equation (11) is convex for $m > 0$. As a result, we can practically solve this for continuous values of m and choose either the left or the right integer that yields the smaller error. Differentiating the right-hand side of (11) with respect to m, we obtain:

$$\frac{(2m^2-1)}{6m^2}\sigma_w^2 - \frac{1}{m^2}\sigma_v^2 = 0 \tag{12}$$

which upon additional algebraic manipulation leads to the expression for the optimal characteristic timescale:

$$m_c = \underset{m\in Z^+}{\text{argmin}}\ \mathbb{E}\left[(\hat{x}_n^{(m)} - x_n)^2\right] = \left\lfloor \sqrt{3\frac{\sigma_v^2}{\sigma_w^2}+\frac{1}{2}} \right\rceil \tag{13}$$

□

Thus, averaging at the characteristic timescale m_c, i.e., the optimal averaging window length, minimizes the mean-square error between the estimated and actual values of the time-varying parameter x_n at the current time instant n.

From equation (13) as well as Figure 1, it is evident that the magnitude of the measurement noise (σ_v^2) and the magnitude of the white noise used to generate the random walk (σ_w^2) both have linear relationships with the size of the optimal averaging window, but with opposite effects. Specifically, if σ_v^2 increases compared to σ_w^2, the incoming data stream looks increasingly like white noise, with measurements failing to accurately reflect the underlying random walk process. In the extreme, this corresponds to $\sigma_v^2/\sigma_w^2 \to \infty$, indicating that the optimal averaging window length $m_c \to \infty$ as well. Thus, if the measurement noise dominates the random walk signal, the moving average estimator is more accurate if averaging is performed over long windows, i.e. the characteristic timescale is large.

On the other hand, if σ_v^2 decreases compared to σ_w^2, the incoming data stream increasingly takes on the characteristics of a random walk, with minimal impact of the measurement noise. This corresponds to $\sigma_v^2/\sigma_w^2 \to 0$, indicating that the optimal averaging window length $m_c \to 1$. Thus, for a 'pure' random walk with no measurement noise, the optimal window size is 1, i.e. the current measurement is trivially the optimal estimate of the current state of the stochastic process.

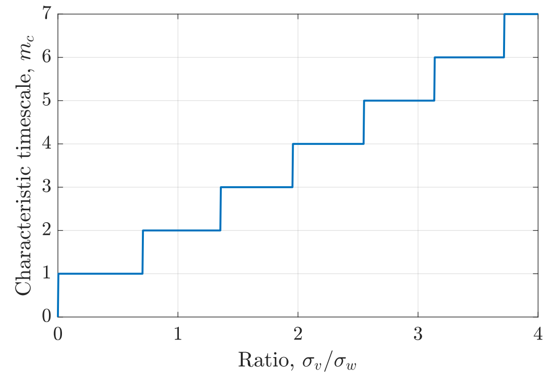

Fig. 1. The characteristic timescale m_c indicates the optimal window length to use for moving average estimation. The value of m_c increases as the ratio σ_v/σ_w increases. For very noisy measurements, longer window lengths provide optimal moving average estimation.

IV. Allan Variance-based Optimal Moving Average Estimation

As is evident from the discussion in Section III, to find the optimal window length m_c we need to know the ratio σ_v/σ_w. However, this information is not always readily available, which leaves the only possible recourse of using (a) heuristic estimates, or (b) knowledge from domain experts. Moreover, it is challenging to obtain this ratio from the raw measurements themselves since the effects of the stochastic processes $\{W_k\}$ and $\{V_k\}$ are confounded in the output measurements $\{X_k + V_k\}$. In this section, we present analytical results for a new method to systematically estimate the characteristic timescale m_c (i.e. the optimal averaging window length) by finding the timescale that minimizes the Allan Variance of the signal [8].

A. Allan Variance

Allan Variance (AVAR) was originally developed to address the issue of frequency stability and time synchronization between atomic clocks [15] but it has quickly become a useful tool for modeling and de-noising inertial sensors [16][17]. Allan Variance measures the signal bias at a certain timescale. We recently showed the timescale minimizing AVAR can be effectively used as an estimate of the characteristic timescale in moving average estimators [8] which will be discussed in more detail in the next subsection. A more detailed analysis on noise modeling and characterization using Allan Variance can be found in [18].

The non-overlapping Allan Variance σ_A^2, also known as two-sample variance, for a time series $\boldsymbol{y} = \{y_1, y_2, ..., y_n\}$ is defined as a function of timescale m as follows:

$$\sigma_A^2(m) = \frac{1}{2}\mathbb{E}\left[(\bar{y}_k - \bar{y}_{k-m})^2\right] \tag{14}$$

where $\bar{y}_k$ is the average of measurements in the window of size m calculated at the time step k. We omit the symbol m when writing the window averages $\bar{y}_k$ to make the notation simpler:

$$\bar{y}_k = \frac{1}{m}\sum_{i=1}^{m} y_{k-i+1} \tag{15}$$

The expectation operator in (14) may be evaluated, so that the Allan Variance can be numerically estimated as:

$$\hat{\sigma}_A^2(m) = \frac{1}{2(n-2m)}\sum_{k=2m+1}^{n}(\bar{y}_k - \bar{y}_{k-m})^2 \tag{16}$$

where n is the number of measurements or data points in the entire time series $\boldsymbol{y}$ and m denotes the timescale at which AVAR is being evaluated. Alternatively, as shown in [19], the expressions in (14) can be also written as:

$$\hat{\sigma}_A^2(m) = \frac{1}{2m^2}\mathbb{E}\left[(z_{k+2m} - 2z_{k+m} + z_k)^2\right] \tag{17}$$

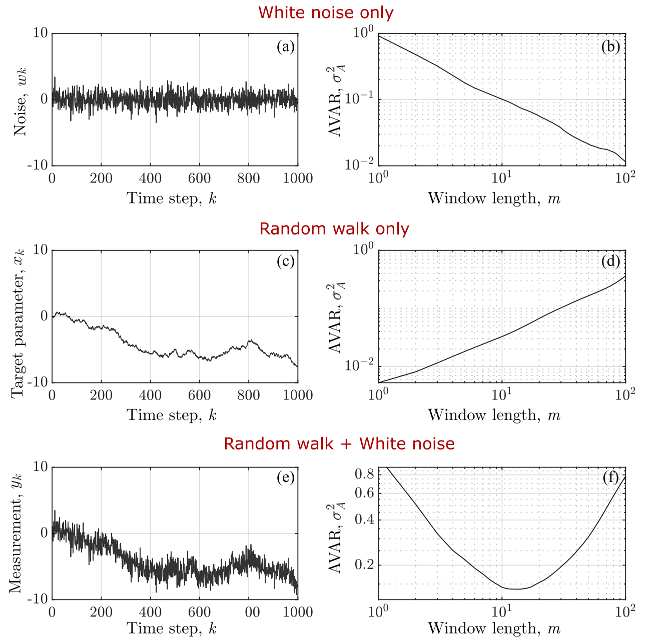

Fig. 2. **(a-b)** A Gaussian white noise realization and its corresponding AVAR. **(c-d)** A random walk realization and its corresponding AVAR. **(e-f)** Measurement signal composed from the previous signals and its AVAR that is minimized at a timescale of $m = 12$. The AVAR values are estimated using equation (18).

where $z_k = \sum_{i=1}^{k} y_i$ is the accumulated signal value at the time step k. Thus, the value of the Allan variance in (14) can be estimated with the computational complexity of $\mathcal{O}(n)$ as follows:

$$\hat{\sigma}_A^2(m) = \frac{1}{2m^2(n-2m)} \sum_{k=1}^{n-2m} (z_{k+2m} - 2z_{k+m} + z_k)^2 \tag{18}$$

where n denotes the length of the signal being analyzed, and m is the window length in units of time steps.

Figure 2 shows three signals and their corresponding Allan variance as a function of window length. Fig. 2(a-b) shows a white noise signal (which is uncorrelated in time) and its corresponding AVAR as a function of window length. Visualizing this in a log-log scale, we observe that the Allan variance of the white noise decreases for increasing window lengths, indicating that measurement bias reduces as we include more data points (longer windows) in the averaging process (see [18] for more details). Fig. 2(c-d) show a random walk signal and its AVAR that is increasing as a function of window length. Fig. 2(e-f) shows a measurement signal that is composed of the random walk and the white noise time series data in the other plots. In this example, the Allan variance curve has a minimum at $m = 12$, indicating the minimum measurement bias occurs at this particular timescale.

B. Optimality of Moving Average Estimator using AVAR-informed timescale

In this subsection, we prove that the Allan Variance is also minimized at the *same* characteristic timescale as derived in (5). A significant implication of this result is that we can estimate the characteristic timescale m_c (i.e. the optimal moving average window length) by directly estimating AVAR, *without* requiring any explicit information about σ_v or σ_w. Consequently, AVAR enables a practical way to perform optimal moving average estimation in noisy random walks, without any knowledge of the underlying noise characteristics.

Theorem 2. *If the drift signal x_k is realized from a random walk process as shown in (1), and the measurement signal is realized according to (2), then the optimal averaging window length m_c minimizes the Allan Variance as defined in (14).*

Proof. We begin by expanding the expression for evaluating the non-overlapping Allan variance in (14) as follows:

$$\sigma_A^2(m) = \frac{1}{2}\mathbb{E}\left[\left(\frac{1}{m}\sum_{i=1}^{m} y_{k-i+1} - \frac{1}{m}\sum_{i=1}^{m} y_{k-m-i+1}\right)^2\right]$$

Replacing y_i with $x_i + v_i$, and decomposing the summations into three independent terms, we obtain:

$$\sigma_A^2(m) = \frac{1}{2m^2}\mathbb{E}\left[\left(\sum_{i=1}^{m} x_{k-i+1} - x_{k-m-i+1} + \sum_{i=1}^{m} v_{k-i+1} - \sum_{i=1}^{m} v_{k-m-i+1}\right)^2\right] \tag{19}$$

Since summation terms of the random walk process and measurement noise are assumed to be independent of each other and white noise is uncorrelated, the cross terms in the expectation term are zero (i.e. $\mathbb{E}[x_i v_j] = 0$ for all i, j, and $\mathbb{E}[v_i v_j] = 0$ for all $i \neq j$). Consequently, noise terms v_i can be extracted out from the expectation operator as follows:

$$\sigma_A^2(m) = \frac{1}{2m^2}\mathbb{E}\left[\left(\sum_{i=1}^{m} x_{k-i+1} - x_{k-m-i+1}\right)^2\right] + \frac{\sigma_v^2}{m}$$

Further, using the fact that $x_b - x_a = \sum_{i=a+1}^{b} w_i$ for any $a < b \in \mathbb{Z}^+$, we can simplify the above expression for the Allan variance to yield:

$$\sigma_A^2(m) = \frac{1}{2m^2}\mathbb{E}\left[\left(\sum_{i=1}^{m}\sum_{j=k-m-i+2}^{k-i+1} w_j\right)^2\right] + \frac{\sigma_v^2}{m} \tag{20}$$

which can be rewritten in expanded form as:

$$\sigma_A^2(m) = \frac{1}{2m^2}\mathbb{E}\Big[\Big\{w_{k-2m+2} + 2w_{k-2m+3} + ... + (m-1)w_{k-m} + mw_{k-m+1} + (m-1)w_{k-m+2} + ... + 2w_{k-1} + w_k\Big\}^2\Big] + \frac{\sigma_v^2}{m} \tag{21}$$

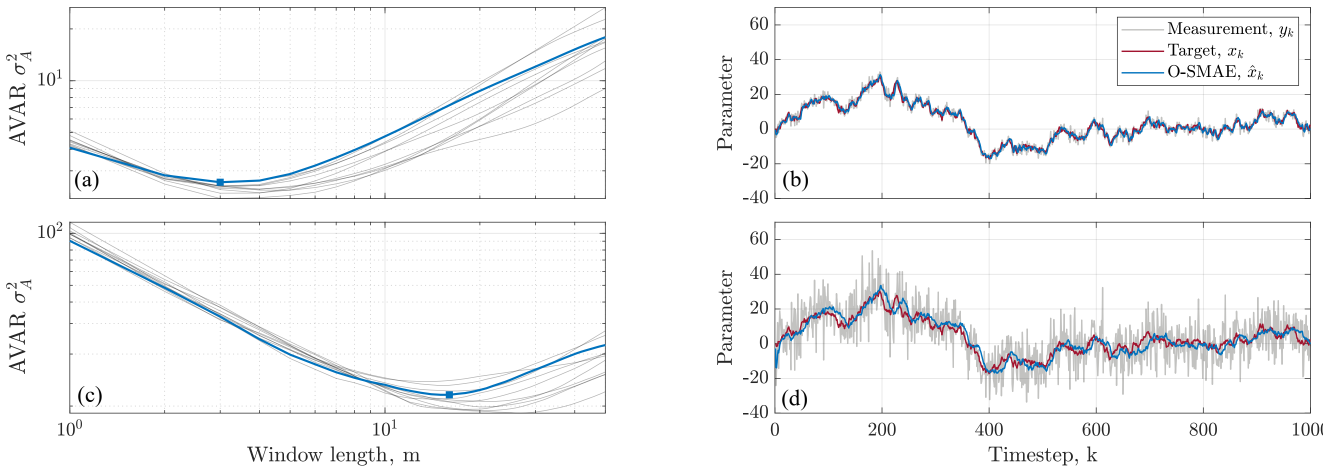

Fig. 3. Two examples of Optimal Simple Moving Average Estimation (O-SMAE) of noisy random walks using Allan Variance-informed window lengths. Value of $\sigma_w = 1$ for both cases. For the first example (a-b), $\sigma_v = 2$, whereas in the second one (c-d), $\sigma_v = 10$. The optimal averaging window length is estimated by finding the timescale that minimizes AVAR (blue square). For a better perspective, the gray lines in (a) and (c) show the corresponding AVAR curves of several other random realizations of noisy random walks.

Again, we know that $\mathbb{E}[w_i w_j] = 0$ for $i \neq j$, since the random walk is generated by uncorrelated Gaussian white noise. After some arithmetic operations (including summing the squares of first m natural numbers), the expression for Allan variance in (21) simplifies to:

$$\sigma_A^2(m) = \frac{2m^2+1}{6m}\sigma_w^2 + \frac{1}{m}\sigma_v^2 \quad (22)$$

To minimize the Allan variance, we differentiate the right-hand side of (22), i.e. letting $\frac{\partial}{\partial m}\sigma_A^2(m) = 0$ results in:

$$\frac{(2m^2-1)}{6m^2}\sigma_w^2 - \frac{1}{m^2}\sigma_v^2 = 0 \quad (23)$$

which is the same equation as (12) and yields the optimal timescale m_c in (5), reproduced here for clarity:

$$m_c = \underset{m \in Z^+}{\text{argmin}}\ \mathbb{E}\left[(\hat{x}_n^{(m)} - x_n)^2\right] = \left\lfloor \sqrt{3\frac{\sigma_v^2}{\sigma_w^2} + \frac{1}{2}} \right\rceil \quad (24)$$

□

As an immediate result of the above theorem, and in the absence of any information about the underlying noise characteristics, we propose to generate an estimate of the characteristic timescale, i.e. $\hat{m}_c$, as follows:

$$\hat{m}_c = \text{argmin}\ \hat{\sigma}_A^2(m) \quad (25)$$

where $\hat{\sigma}_A^2(m)$ can be evaluated using (18).

A consequence of this result is that we can optimally perform moving average estimation by systematically selecting an Allan variance-informed window length, without the need to rely on heuristics. It may also be noted that AVAR can be estimated via other estimators proposed in the literature. While the AVAR estimator in (18) yields a precise estimate, there are other AVAR estimators that are computationally faster such as the one proposed in [20], which estimates the AVAR recursively using a hierarchical structure.

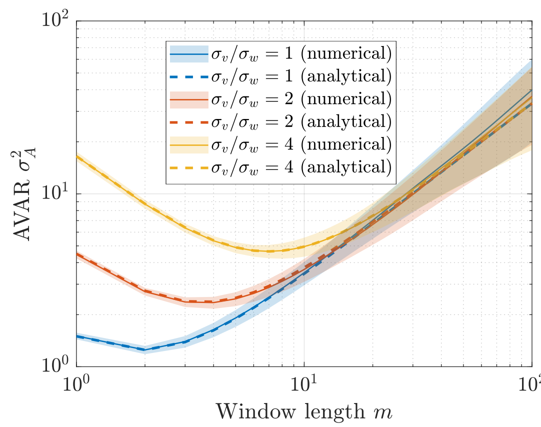

Fig. 4. Numerical estimation of the Allan Variance (solid lines) using (18) and analytically calculated Allan Variance (dashed lines) using (22). Note that AVAR curves are minimized at the exact optimal window length m_c illustrated on Fig. 1.

V. Results and Concluding Remarks

In this section, we provide numerical support to demonstrate that the proposed SMAE method that utilizes the Allan variance-informed characteristic timescale does indeed provide optimal estimates. Fig. 3 shows an optimal simple moving average estimation of a noisy random walk with two different noise properties. As shown in Fig. 3 (a-c), the AVAR (estimated using (18)) is minimized at different timescales for each example. In Fig. 3(a-b), there is only a small amount of noise added to the random walk, hence measurements are tightly correlated and a short averaging window yields better results. On the other hand, Fig. 3(c-d) indicates the presence of a significant amount of white noise which makes measurements less correlated in time, so averaging over longer window length gives better estimates.

Additionally, as shown in Fig. 4, AVAR estimation yields an unbiased estimation of the characteristic timescale. The numerical (solid lines) AVAR values have been evaluated according to (18) and have been averaged over 100 Monte Carlo simulations. The shaded bounds represent the standard deviation associated with the Monte Carlo simulations. The dashed lines represent the actual AVAR values assuming the variances σ_v^2 and σ_w^2 are known which is represented by (22). In all the cases, the value of the σ_w is set to 1 while σ_v varies from 1 to 4.

Finally, we analyze the difference between the characteristic timescale obtained numerically from Allan variance *estimates*, i.e. $\hat{m}_c$ in (25), and the optimal averaging window length obtained from analytical solution associated with the mean-square error, i.e. m_c in (13). The results show that while there may be a small error in the size of the optimal window length determined via Allan variance, the error $|\hat{m}_c - m_c| \to 0$ as the number of available measurements in the data stream increases. Even with a moderate number of historical measurements, the difference between the optimal and the Allan-variance informed window length is negligibly small. Another point of interest is that the variance of the error $|\hat{m}_c - m_c|$ (shown with the lightly shaded region) also tends to zero for a moderately large number of historical measurements.

A consequence of these results is that we can now optimally perform moving average estimation of noisy random walks by systematically selecting an Allan variance-informed window length, without the need to rely on heuristics or any prior knowledge of signal and noise characteristics. The proposed method can be deployed for a vectorial formulation of the state space as long as the white noise processes are uncorrelated across all dimensions. As a result, several application areas can benefit from this approach. The closed-form solution and knowledge that the Allan variance-informed characteristic timescale is indeed the optimal averaging window length lends additional confidence to practitioners for using this method in real-world applications.

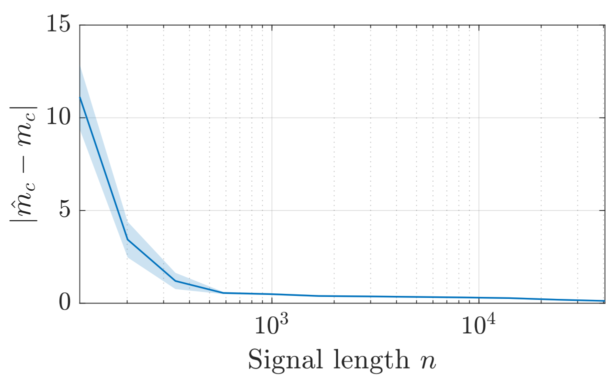

Fig. 5. Characteristic timescale estimation error, i.e., the difference between the timescale minimizing AVAR in (18), $\hat{m}_c$, and the one minimizing the estimation error in (11), m_c, versus length of the measurement signal n. For this particular result we set $\sigma_w = 1$ and $\sigma_v = 2$. Also at each signal length, the results have been averaged over 100 Monte Carlo realizations.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1932138. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] Y. Shynkevich, T. M. McGinnity, S. Coleman, Y. Li, and A. Belatreche, "Forecasting stock price directional movements using technical indicators: investigating window size effects on one-step-ahead forecasting," in *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*. IEEE, 2014, pp. 341–348.

[2] P. Luo, M. Zhang, Y. Liu, D. Han, and Q. Li, "A moving average filter based method of performance improvement for ultraviolet communication system," in *2012 8th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP)*. IEEE, 2012, pp. 1–4.

[3] "Is smarter better? A comparison of adaptive, and simple moving average trading strategies," *Research in International Business and Finance*, vol. 19, no. 3, pp. 399–411, 2005.

[4] J. Forbes, M. Ordonez, and M. Anun, "Improving the dynamic response of power factor correctors using simple digital filters: Moving average filter comparative evaluation," in *2013 IEEE Energy Conversion Congress and Exposition*. IEEE, 2013, pp. 4814–4819.

[5] H. Chen and S.-W. Chen, "A moving average based filtering system with its application to real-time qrs detection," in *Computers in Cardiology, 2003*. IEEE, 2003, pp. 585–588.

[6] Y. Chen, D. Li, Y. Li, X. Ma, and J. Wei, "Use moving average filter to reduce noises in wearable ppg during continuous monitoring," in *eHealth 360°*. Springer, 2017, pp. 193–203.

[7] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.

[8] H. Haeri, C. E. Beal, and K. Jerath, "Near-optimal moving average estimation at characteristic timescales: An Allan variance approach," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1531–1536, 2020.

[9] J. C. Van Horne and G. G. Parker, "The random-walk theory: an empirical test," *Financial Analysts Journal*, vol. 23, no. 6, pp. 87–92, 1967.

[10] O. C. Ibe, *Elements of random walk and diffusion processes*. John Wiley & Sons, 2013.

[11] N. Suciu, C. Vamoş, I. Turcu, C. Pop, and L. Ciortea, "Global random walk modelling of transport in complex systems," *Computing and Visualization in Science*, vol. 12, no. 2, pp. 77–85, 2009.

[12] E. A. Codling, M. J. Plank, and S. Benhamou, "Random walk models in biology," *Journal of the Royal Society Interface*, vol. 5, no. 25, pp. 813–834, 2008.

[13] B. Berkowitz, A. Cortis, M. Dentz, and H. Scher, "Modeling non-fickian transport in geological formations as a continuous time random walk," *Reviews of Geophysics*, vol. 44, no. 2, 2006.

[14] A. Elfeki and J. Bahrawi, "Application of the random walk theory for simulation of flood hazards: Jeddah flood 25 november 2009," *International Journal of Emergency Management*, vol. 13, no. 2, pp. 169–182, 2017.

[15] D. W. Allan, "Statistics of atomic frequency standards," *Proceedings of the IEEE*, vol. 54, no. 2, pp. 221–230, 1966.

[16] K. Jerath and S. N. Brennan, "GPS-Free Terrain-Based Vehicle Tracking Performance as a Function of Inertial Sensor Characteristics," in *Proceedings of ASME 2011 Dynamic Systems and Control Conference*, 2011, pp. 367–374.

[17] M. Kok, J. D. Hol, and T. B. Schön, "Using inertial sensors for position and orientation estimation," *arXiv preprint arXiv:1704.06053*, 2017.

[18] K. Jerath, S. Brennan, and C. Lagoa, "Bridging the gap between sensor noise modeling and sensor characterization," *Measurement*, vol. 116, pp. 350 – 366, 2018.

[19] N. El-Sheimy, H. Hou, and X. Niu, "Analysis and modeling of inertial sensors using allan variance," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 1, pp. 140–149, 2007.

[20] S. Prasad Maddipatla, H. Haeri, K. Jerath, and S. Brennan, "Fast Allan Variance (FAVAR) and Dynamic Fast Allan Variance (D-FAVAR) Algorithms for both Regularly and Irregularly Sampled Data," *To appear in Modeling, Estimation and Control Conference*, 2021.