Optimal Averaging Time for Improving Observer Accuracy of Stochastic Dynamical Systems

Pedro Balaguer*

Jaume I University, Castellón

Asier Ibeas

Universitat Autònoma de Barcelona

Abstract

In the problem of remote estimation by a centralized observer, improvements to the accuracy of observer estimates come at a cost of higher communication bandwidth and energy consumption. In this article we improve observer estimation accuracy by reducing the measurement variance on the sensor node before its transmission to the centralized observer node. The main contribution is to show that measurement variance is a trade-off between dynamical system variance and sensor variance. As a result there is an optimal averaging time that minimizes measurement variance, providing more accurate measurement to the observer. The optimal averaging time is computable by solving a univariate optimization problem.

Keywords: Smart Sensors, Wireless Sensor Networks, Observer, Stochastic Processes, Optimal Averaging Time, Kalman Filter

1. Introduction

Technological developments in sensor networks (SN), the Internet of Things (IoT), and networked control systems (NCS) have led to a multitude of sensors that send information for supervision, estimation, and control [1, 2, 3]. In

*Corresponding author
Email address: pbalague@uji.es (Pedro Balaguer)

this context, centralized remote observers receive measurements through the communication network from sensor nodes and provide a state estimate of the dynamical system.

In centralized remote observers, there is a trade-off between the estimate accuracy and network resource consumption. In fact, an increase in estimate accuracy is normally accompanied by an increase in communication bandwidth and the consumption of node energy. This trade-off is tackled in the bibliography by event triggered estimation [4], and by sensor scheduling [5]. In event triggered estimation, the sensor nodes decide when to transmit on the basis of triggering rules [6]. In sensor scheduling, the approach is to temporally schedule the sensor nodes to minimize expected error covariance performance [7]. In both approaches, the goal is to orchestrate the transmission of measurements by reducing the communication burden while minimizing estimate accuracy degradation. However, no improved measurement is pursued at the sensor node before the measurement is transmitted.

The existing approaches for measurement improvement are based on digital filtering [8, 9]. Digital filters are designed for noisy sensors based on the minimization of some criterion [10], and are technologically implemented by means of oversampling and averaging [11], output filtering [12], or by using integrating analog to digital converters (ADC) [13]. Note however that the optimal processing of the signal not only depends on the sensor noise characteristics but also on the relationship between stochastic dynamical system noise (i.e. process noise) and sensor noise.

The discretization of the measurement equation in stochastic systems leads to measurement averaging [14], because in this case the discretized sensor noise $v_k := \frac{1}{\Delta t}\int_0^{\Delta t} v(\tau)d\tau$ and the continuous sensor noise $v(t)$ have the same spectral densities [15]. The averaging procedure makes the measurement variance depend not only on sensor variance but also on process variance. Furthermore, both variances have effects that compete with the averaging time Δt, because while the sensor variance is reduced with Δt, the process variance is increased with Δt. As a result, there is expected to be an optimal sampling time Δt^* such that the variance of the averaged measurement is minimized.

Bibliographical results for measurement averaging, also called integrated measurements, are scarce [16], and focus on state estimation and fusion of fast-rate measurements with slow-rate averaged measurements [17]. In these studies, the existence of averaged measurements is a technological consequence when laboratory measurements of samples taken over time are available to be used for state estimation. However, the bibliographical contributions do not provide any guidelines for selecting the optimal averaging time. The contribution of the present study to the field of averaging measurements consists of computing the optimal averaging/integrating time that provides the measurement with minimum variance.

In summary, the aim of this study is to reduce remote observer variance by sending the measurement with minimum variance obtained by optimal averaging on the sensor node, and by modifying the observer equations to account for the departure from instant sampling towards averaged sampling. The contributions of this article are:

1. The continuous-time stochastic dynamical model and the output model are discretized with distinct time periods. Although the current literature uses the same discretization period for the dynamical equation and the output equation [18], the use of different discretization periods allows for separation between the measurement averaging on the sensor node and the measurement communication to the observer node.
2. The measurement variance as a function of the averaging time is derived, and it is shown that it captures a trade-off between process noise and sensor noise. The shorter the averaging time the lesser the influence of process noise on the measurement, but the sensor noise cannot be filtered. On the contrary, the longer the averaging time the lesser the influence of the sensor noise, but the process noise increases the measurement variance. We show that there is an optimal averaging time that minimizes measurement variance that can be computed by solving a univariate minimization

problem.

3. The standard Kalman filter equations are modified for use with averaged measurements. This is necessary to account for the use of averaged measurements on the Kalman filter equations instead of instant measurements. We show the improvement in observer accuracy.

2. Problem Statement

The dynamical system is a stochastic multivariable continuous-time linear time-invariant (LTI) system modelled by the following stochastic differential equation (SDE) [19, 20] and output equation

$$\dot{x}(t) = A_c x(t) + B_c u(t) + G_c w(t) \tag{1}$$

$$x(t_0) = x_0$$

$$y(t) = Cx(t) + v(t) \tag{2}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^{n_u}$, the input vector, $y(t) \in \mathbb{R}^{n_y}$, the output vector, $w(t) \sim N(0, Q_c) \in \mathbb{R}^{n_w}$ and $v(t) \sim N(0, R_c) \in \mathbb{R}^{n_y}$ are normally distributed white noise processes that are uncorrelated with each other and with the initial state $x_0 \sim N(\bar{x}_0, P_0)$, with $Q_c\delta(t-\tau) \in \mathbb{R}^{n_w \times n_w}$, $R_c\delta(t-\tau) \in \mathbb{R}^{n_y \times n_y}$, and $P_0 \in \mathbb{R}^{n \times n}$ covariance matrices [21], and $\delta(t-\tau)$ the Dirac delta distribution. $A_c \in \mathbb{R}^{n \times n}$, $B_c \in \mathbb{R}^{n \times n_u}$, $G_c \in \mathbb{R}^{n \times n_w}$, and $C \in \mathbb{R}^{n_y \times n}$ are the system matrices.

Consider the problem of estimating the state $x(t)$ of the stochastic dynamical system by a remote observer node within a distributed network of sensor nodes, as represented in Fig. 1. The goal is to provide the optimal estimator $\hat{x}(t)$, that is, the estimator with minimum variance. The Kalman filter [14] provides the minimum variance estimator for Gaussian, zero-mean, uncorrelated, and white noises [14].

The observer node requires the discretized dynamic Eq. (1) for non-periodic updates on the received measurements at each time instant t_k, resulting in a variable sampling time $\mu_k := t_{k+1} - t_k$. The sensor node requires the discretized

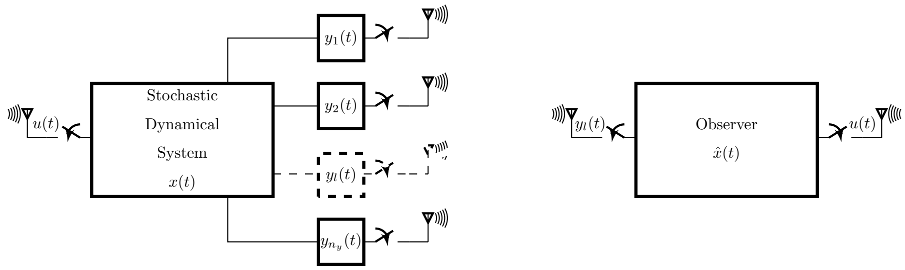

Figure 1: Remote estimation problem in sensor networks. The set of n_y sensor nodes send the measurements $y_l, \quad l \in \{1, 2, \ldots, n_y\}$ to the observer node. Likewise, input $u(t)$ is sent to the stochastic dynamical system and to the observer.

measurement Eq. (2) with discretization time $\delta_l := t_{k+1} - t_k^{y_l}$. Note that we consider a constant averaging time δ_l that is independent from the sampling interval k. The reason is that as the stochastic processes are stationary, the optimal averaging time δ_l^* does not depend on any particular time interval. In Fig. 2 we can see both discretization times, μ_k and δ_l.

The discretization of the measurement equation results in a measurement y_{δ_l} that is the average of the continuous time output [15], that is, $y_{\delta_l} := \frac{1}{\delta_l} \int_{t_k^{y_l}}^{t_{k+1}} y_l(\tau) d\tau$. The variance of the averaged measurement for sensor l depends on the averaging time δ_l. There is an optimal value for the averaging time δ_l^* that provides the measurement with minimum variance, optimizing the trade-off between sensor noise variance, reduced with larger δ_l, and process noise variance, reduced with shorter δ_l. One contribution of the article is to compute, for each sensor node $l \in \{1, 2, \ldots, n_y\}$, its optimal averaging time δ_l^* that provides the measurement with minimum variance to the observer node.

3. Stochastic Process Discretization

The discretization of the stochastic system provided in the bibliography [14] uses the same discretization time period for the dynamical model (1) and the output model (2). Although this approach might be adequate for wired or centralized systems, the dynamical model and the output model may be discretized

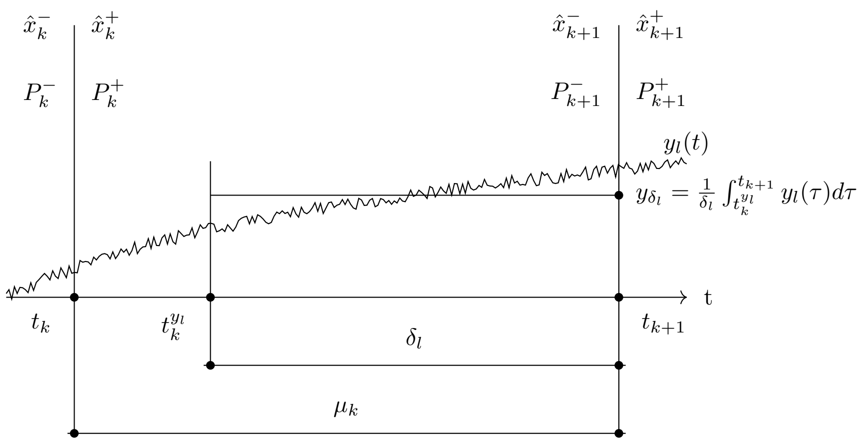

Figure 2: The observer node receives a measurement at time t_k. The measurement update is computed giving estimate $\hat{x}_k^+$ and covariance P_k^+. Before a new measurement is received at time t_{k+1} the time update gives estimate $\hat{x}_{k+1}^-$ and convariance P_{k+1}^-. The process repeats with measurement at time t_{k+1} providing new estimate $\hat{x}_{k+1}^+$ and covariance P_{k+1}^+. The time between received measurements at observer node is $\mu_k := t_{k+1} - t_k$. The sensor node l sends a measure to the observer node at time instant t_{k+1}. In this article, instead of sending the output value sampled at time instant t_{k+1}, that is $y_l(t_{k+1})$, we consider sending an averaged value $y_{\delta_l} := \frac{1}{\delta_l}\int_{t_k^{y_l}}^{t_{k+1}} y_l(\tau)d\tau$, with $\delta_l := t_{k+1} - t_k^{y_l}$. We show that there is an optimal value δ_l^* that provides the measurement y_{δ_l} with minimum variance, hence improving the observer estimation accuracy.

using distinct time intervals because they attend to different goals. The dynamical model discretization time is related to the data communication between sensor nodes and observer node. The output model discretization results in an averaged measure that is controlled by the sensor node. In what follows we discretize the continuous-time process using distinct discretization periods for the model and the output equations.

3.1. Dynamical Model Discretization

The solution of the SDE (1) for an arbitrary time interval $\mu_k := t_{k+1} - t_k$ is

$$x(t_{k+1}) \quad = \quad e^{A_c \mu_k} x(t_k) + \underbrace{\int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau)} B_c u(\tau) d\tau}_{\text{deterministic integral}} + \underbrace{\int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau)} G_c w(\tau) d\tau}_{\text{stochastic integral}} \tag{3}$$

with $x(t_0) = x_0$.

The first integral on the right hand side of (3) is a deterministic integral, whereas the second integral is of a stochastic nature. The discretization of (1), when the control input $u(t)$ is a zero-order hold (ZOH), (i.e. $u(t) = u(t_k) \quad \forall t \in [t_k, t_{k+1}), \quad k = 0, 1, \ldots, N-1$), is given by

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{4}$$

with $x_k := x(t_k)$, $u_k := u(t_k)$, and

$$\begin{aligned}
A &:= e^{A_c \mu_k} && (5)\\
B &:= \int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau)} B_c d\tau \xrightarrow{\lambda = t_{k+1}-\tau} \int_0^{\mu_k} e^{A_c \lambda} B_c d\lambda && (6)\\
w_k &:= \int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau)} G_c w(\tau) d\tau \xrightarrow{\lambda = t_{k+1}-\tau} \int_0^{\mu_k} e^{A_c \lambda} G_c w(t_{k+1}-\lambda) d\tau && (7)
\end{aligned}$$

We remark that system matrices A and B, and stochastic signal w_k, depend on time increment μ_k and are independent from the particular time instant t_k and t_{k+1}. The change of variable $\lambda = t_{k+1} - \tau$ on the integrals is performed to show their dependence on the sampling time μ_k.

Now, the stochastic properties of w_k are derived. First, the expectation of w_k is 0, because as stated in the problem statement $\mathbb{E}[w(t)] = 0$ and the integral is a linear operator. The covariance $Q_s := \mathbb{E}[w_k w_k^T]$ is given as follows

$$
\begin{aligned}
Q_s &= \mathbb{E}\left[\int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau_1)} G_c w(\tau_1) d\tau_1 \int_{t_k}^{t_{k+1}} w(\tau_2)^T G_c^T e^{A_c^T(t_{k+1}-\tau_2)} d\tau_2\right] \\
&= \mathbb{E}\left[\int\int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau_1)} G_c w(\tau_1) w(\tau_2)^T G_c^T e^{A_c^T(t_{k+1}-\tau_2)} d\tau_1 d\tau_2\right] \\
&= \int\int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau_1)} G_c \mathbb{E}[w(\tau_1) w(\tau_2)^T] G_c^T e^{A_c^T(t_{k+1}-\tau_2)} d\tau_1 d\tau_2 \\
&= \int\int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau_1)} G_c Q_c \delta(\tau_2-\tau_1) G_c^T e^{A_c^T(t_{k+1}-\tau_2)} d\tau_1 d\tau_2 \\
&= \int_{t_k}^{t_{k+1}} e^{A_c(t_{k+1}-\tau)} G_c Q_c G_c^T e^{A_c^T(t_{k+1}-\tau)} d\tau \\
&\xrightarrow{\lambda = t_{k+1}-\tau} \int_0^{\mu_k} e^{A_c\lambda} G_c Q_c G_c^T e^{A_c^T\lambda} d\lambda
\end{aligned}
\tag{8}
$$

The computation of Q_s does not generally have a closed form and must be computed numerically [14]. The most common approach is the one proposed by Van Loan [22]. However, we must keep the dependence of covariance Q_s with time interval μ_k for optimization purposes, hence the computation of Q_s is based on the exponential matrix expansion of $e^{A_c\mu_k}$, as presented in the Appendix 9.

Example 1 (From [23], Example 2.6 page 87). Consider the continuous-time dynamical system of a mass with a force given by

$$
\dot{x}(t) = \underbrace{\begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}}_{A_c} x(t) + \underbrace{\frac{1}{m}\begin{vmatrix} 0 \\ 1 \end{vmatrix}}_{B_c} u(t) + \underbrace{\begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}}_{G_c} w(t), \quad Q_c = q_a \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}
\tag{9}
$$

with state $x(t) = [x_1(t) \quad x_2(t)]^T$, being $x_1(t)$ the mass position and $x_2(t)$ the mass velocity, $u(t)$ is the input force, and $w(t) = [w_1 \quad w_2]^T$ the process noise. The continuous-time covariance matrix Q_c shows that there is a random disturbing force whereas the velocity is not affected by any random disturbance. Matrix A_c is nilpotent because $A_c^n = 0$ for $n \geq 2$. As a result, the calculation equations provided in the Appendix 9 are exact for $N = 2$, giving the following discrete-time dynamical model

$$x_{k+1} = \underbrace{\begin{vmatrix} 1 & \mu_k \\ 0 & 1 \end{vmatrix}}_{A} x_k + \underbrace{\frac{1}{m}\begin{vmatrix} \frac{\mu_k^2}{2} \\ \mu_k \end{vmatrix}}_{B} u_k + w_k, \quad Q_s = q_a \begin{vmatrix} \frac{\mu_k^3}{3} & \frac{\mu_k^2}{2} \\ \frac{\mu_k^2}{2} & \mu_k \end{vmatrix} \tag{10}$$

3.2. Output Model Discretization

Recall the vector output Eq. (2)

$$y(t) \quad = \quad Cx(t) + v(t) \tag{11}$$

with $y(t) = [y_1(t) \dots y_l(t) \dots y_{n_y}(t)]^T$, $x(t) = [x_1(t) \dots x_m(t) \dots x_n(t)]^T$, $v(t) = [v_1(t) \dots v_l(t) \dots v_{n_y}(t)]^T$, and $C \in \mathbb{R}^{n_y \times n}$. Consider the discretization of the scalar output $y_l(t) \in \mathbb{R}$ with $l \in \{1, 2, \dots, n_y\}$. Discretization of the measurement process with interval δ_l results in measurement averaging, because in this case the discretized sensor noise v_{δ_l} and the continuous sensor noise $v_l(t)$ have the same spectral densities [14, 15]. As a result the discretized measurement is

$$\underbrace{\frac{1}{\delta_l}\int_{t_k^{y_l}}^{t_{k+1}} y_l(\tau)d\tau}_{=:y_{\delta_l}} \quad = \quad c_l \underbrace{\frac{1}{\delta_l}\int_{t_k^{y_l}}^{t_{k+1}} x(\tau)d\tau}_{=:x_{\delta_l}} + \underbrace{\frac{1}{\delta_l}\int_{t_k^{y_l}}^{t_{k+1}} v(\tau)d\tau}_{=:v_{\delta_l}} \tag{12}$$

with c_l the lst row of output matrix C. We can write the discretized output in compact form as

$$y_{\delta_l} = c_l x_{\delta_l} + v_{\delta_l} \tag{13}$$

Note that in case $\delta_l \to 0$, then $y_{\delta_l} = y(t_{k+1})$ and we recover the instant sampled signal as a particular case.

We now derive the statistics of the random variable v_{δ_l}. The expectation of v_{δ_l} is 0 because $\mathbb{E}[v_l(\tau)] = 0$, and the expectation is a linear operator. The covariance $r := \mathbb{E}[v_{\delta_l} v_{\delta_l}^T]$ is

$$
\begin{aligned}
\mathbb{E}[v_{\delta_l} v_{\delta_l}^T] &= \mathbb{E}\left[\frac{1}{\delta_l}\int_{t_k^{y_l}}^{t_{k+1}} v_l(\tau_1)d\tau_1 \frac{1}{\delta_l}\int_{t_k^{y_l}}^{t_{k+1}} v_l(\tau_2)^T d\tau_2\right] \\
&= \mathbb{E}\left[\frac{1}{\delta_l^2}\int\int_{t_k^{y_l}}^{t_{k+1}} v_l(\tau_1)v_l(\tau_2)^T d\tau_1 d\tau_2\right] \\
&= \frac{1}{\delta_l^2}\int\int_{t_k^{y_l}}^{t_{k+1}} \mathbb{E}[v_l(\tau_1)v_l(\tau_2)^T]d\tau_1 d\tau_2 \\
&= \frac{1}{\delta_l^2}\int\int_{t_k^{y_l}}^{t_{k+1}} r_{cl}\delta_l(\tau_2-\tau_1)\tau_1 d\tau_2 \\
&= \frac{1}{\delta_l^2}\int_{t_k^{y_l}}^{t_{k+1}} r_{cl} d\tau \\
&= \frac{r_{cl}}{\delta_l} \qquad (14)
\end{aligned}
$$

with r_{cl} the continuous-time density variance of $v_l(t)$. Note that the covariance of the discretized noise is inversely proportional to the discretization time δ_l.

In what follows we relate the measurement process as performed by electronic circuits. The objective is to link the mathematical meaning of integration time of sensor l, that is δ_l, to the physical meaning on the signals processed. The typical measurement process is to convert an analogue signal, provided by the transducer and the signal conditioning circuitry, to a digital value through an analogue to digital converter (ADC). The ADC converter performs a digital conversion on the basis of a stable analogue signal by means of the sampler and hold (S&H) circuit, with T_{SH} being the time for which switch S_{SH} is closed, as can be seen in Fig. 3. Note that the analogue signal converted is the measured signal $y(t)$ filtered by a first order low pass filter composed of the follower operational amplifier output impedance and the capacitor C_{SH}. The time constant of this filter is very small. Hence we have the conversion of an instant output value $y(t_{k+1})$.

The implementation of the averaged measurement Eq. (13) requires integration of the measure that is to be converted, as shown in the schematic of Fig. 4. Thus the analog signal is integrated for time δ_l. This is controlled by switch S_I on the first operational amplifier. The sampler and hold works in the same

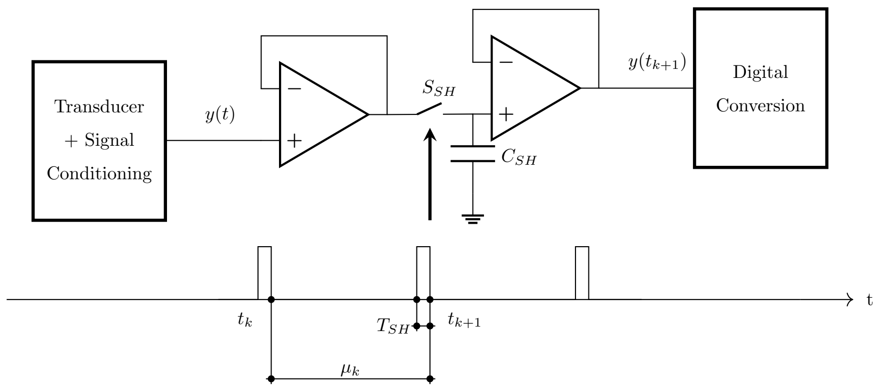

Figure 3: Analogue signal and ADC converter with sampler and hold. The train of pulses on the switch set the sampler and hold times.

way as before, but now holds the value of the integrated measure. Varying the integration time we are able to control the sensor variance. In whatever case, we claim that the implementation of the averaged measurement should be performed by the simple circuitry presented in Fig. 4. Another implementation of the same idea can be performed digitally by oversampling and averaging [11], or using integrating ADC converters [13].

Example 2. Following on from the previous example, consider the continuous-time output equation

$$y(t) = \underbrace{\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}}_{C=\begin{vmatrix} c_1 \\ c_2 \end{vmatrix}} x(t) + v(t), \quad R_c = \begin{vmatrix} r_{c1} & 0 \\ 0 & r_{c2} \end{vmatrix} \tag{15}$$

where $y(t) = [y_1(t) \;\; y_2(t)]^T$ is the output with $y_1(t)$ the measurement of the mass position and $y_2(t)$ the measurement of the mass velocity, $v(t) = [v_1(t) \;\; v_2(t)]^T$ are the position and velocity noise respectively, with diagonal covariance matrix given by R_c. The discretized output equations, each with its integration time $\delta_i \;\; i \in \{1, 2\}$, are

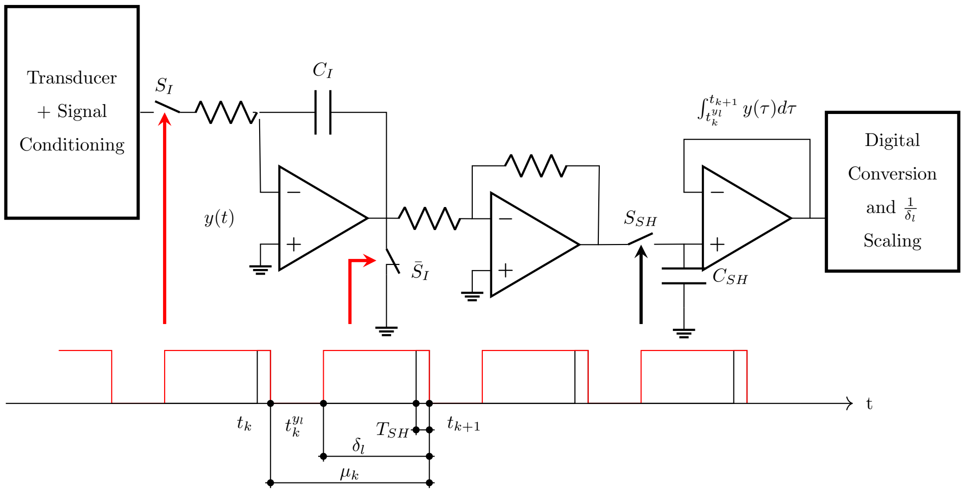

Figure 4: Electronic implementation of the averaged measurement equation. Switch S_I determines the integration time. Its complementary switch $\bar{S}_I$ is used for discharging capacitor C_I for resetting purposes. By varying the integration time δ_l we are able to modify the sensor noise variance r_{cl}/δ_l.

$$y_{\delta_1} = c_1 x_{\delta_1} + v_{\delta_1}, \quad r_1 = \frac{r_{c1}}{\delta_1} \tag{16}$$

$$y_{\delta_2} = c_2 x_{\delta_2} + v_{\delta_2}, \quad r_2 = \frac{r_{c2}}{\delta_2} \tag{17}$$

where c_1 and c_2 are the first and second row of matrix C respectively.

4. Measurement Variance

In this section, we derive the expression of the measurement variance for sensor l, that is $\mathbb{E}[(y_{\delta_l} - \mathbb{E}[y_{\delta_l}])(y_{\delta_l} - \mathbb{E}[y_{\delta_l}])^T]$, with respect to the averaging time δ_l. The measurement variance, not to be confused with the previously derived sensor variance r_l, depends not only on the sensor noise but also on the process noise. As a result, the output variance is a trade-off between sensor and process noise.

Theorem 1. *The measurement variance for sensor l is given by*

$$\mathbb{E}[(y_{\delta_l}-\mathbb{E}[y_{\delta_l}])(y_{\delta_l}-\mathbb{E}[y_{\delta_l}])^T] = c_l\left(\frac{1}{\delta_l^2}\int_0^{\delta_l}\left(\int e^{A_c\lambda}d\lambda\right)G_cQ_cG_c^T\left(\int e^{A_c^T\lambda}d\lambda\right)d\lambda\right)c_l^T+\frac{r_{cl}}{\delta_l} \quad (18)$$

The first term in Eq. (18) is the effect of the process noise whereas the second term is the effect of the sensor noise, both as a function of the averaging time δ_l.

PROOF. The variance of the averaged output y_{δ_l} is by definition $\mathbb{E}[(y_{\delta_l}-\mathbb{E}[y_{\delta_l}])(y_{\delta_l}-\mathbb{E}[y_{\delta_l}])^T]$ and by substitution of the output Eq. (13) results

$$\begin{aligned}\mathbb{E}[(y_{\delta_l}-\mathbb{E}[y_{\delta_l}])(y_{\delta_l}-\mathbb{E}[y_{\delta_l}])^T] &= \mathbb{E}[(c_lx_{\delta_l}+v_{\delta_l}-\mathbb{E}[c_lx_{\delta_l}+v_{\delta_l}])(c_lx_{\delta_l}+v_{\delta_l}-\mathbb{E}[c_lx_{\delta_l}+v_{\delta_l}])^T]\\ &= \mathbb{E}[(c_lx_{\delta_l}+v_{\delta_l}-c_l\mathbb{E}[x_{\delta_l}])(c_lx_{\delta_l}+v_{\delta_l}-c_l\mathbb{E}[x_{\delta_l}])^T]\\ &= \mathbb{E}[(c_l(x_{\delta_l}-\mathbb{E}[x_{\delta_l}])(x_{\delta_l}-\mathbb{E}[x_{\delta_l}])^Tc_l^T]+\mathbb{E}[v_{\delta_l}v_{\delta_l}^T] \quad (19)\end{aligned}$$

The second equality follows because $\mathbb{E}[v_{\delta_l}]=0$. The third equality follows because the expectation of the cross products $\mathbb{E}[c_l(x_{\delta_l}-\mathbb{E}[x_{\delta_l}])v_{\delta_l}^T]$ and $\mathbb{E}[v_{\delta_l}(c_l(x_{\delta_l}-\mathbb{E}[x_{\delta_l}]))^T]$ are also zero, as the expectation of sensor noise is zero and sensor and process noise are uncorrelated.

Define the integral of the state space as $x_I(t):=\int x(t)dt$, then the averaged state space x_{δ_l} is easily related to the integrated state space by means of δ_l, because $x_{\delta_l}=\frac{1}{\delta_l}x_I$. As a result

$$\begin{aligned}\mathbb{E}[((x_{\delta_l}-\mathbb{E}[x_{\delta_l}])(x_{\delta_l}-\mathbb{E}[x_{\delta_l}])^T] &= \mathbb{E}\left[\left(\frac{1}{\delta_l}x_I-\mathbb{E}\left[\frac{1}{\delta_l}x_I\right]\right)\left(\frac{1}{\delta_l}x_I-\mathbb{E}\left[\frac{1}{\delta_l}x_I\right]\right)^T\right]\\ &= \frac{1}{\delta_l^2}\mathbb{E}[(x_I-\mathbb{E}[x_I])(x_I-\mathbb{E}[x_I])^T] \quad (20)\end{aligned}$$

The variance of x_I can be obtained by extending the state space of the dynamic model (1) with the integral $x_I(t)$ and computing its variance in an analogous manner as shown in Section 3.1. The differential equation to add is $\dot{x}_I(t)=x(t)$. As a result we have the extended dynamical system

$$\begin{vmatrix} \dot{x}(t) \\ \dot{x}_I(t) \end{vmatrix} = \underbrace{\begin{vmatrix} A_c & 0 \\ I & 0 \end{vmatrix}}_{A_e} \begin{vmatrix} x(t) \\ x_I(t) \end{vmatrix} + \underbrace{\begin{vmatrix} B_c \\ 0 \end{vmatrix}}_{B_e} u_k + \underbrace{\begin{vmatrix} G_c \\ 0 \end{vmatrix}}_{G_e} w_k \tag{21}$$

As shown in Section 3.1, the system variance is defined as

$$\begin{aligned} Q &= \int_0^{\delta_l} e^{A_e \lambda} G_e Q_c G_e^T e^{A_e^T \lambda} d\lambda \\ &= \begin{vmatrix} \overbrace{\mathbb{E}[(x-\mathbb{E}[x])(x-\mathbb{E}[x])^T]}^{Q_s} & \mathbb{E}[(x-\mathbb{E}[x])(x_I-\mathbb{E}[x_I])^T] \\ \mathbb{E}[(x_I-\mathbb{E}[x_I])(x-\mathbb{E}[x])^T] & \underbrace{\mathbb{E}[((x_I-\mathbb{E}[x_I])(x_I-\mathbb{E}[x_I])^T]}_{Q_I} \end{vmatrix} \end{aligned} \tag{22}$$

The calculus of Eq. (22) requires first the computation of $e^{A_e \lambda}$, which is done by expanding the matrix exponential in series as follows:

$$\begin{aligned} e^{A_e \lambda} &= I + A_e \lambda + A_e^2 \frac{\lambda^2}{2!} + A_e^3 \frac{\lambda^3}{3!} + \ldots \\ &= \begin{vmatrix} I & 0 \\ 0 & I \end{vmatrix} + \begin{vmatrix} A_c & 0 \\ I & 0 \end{vmatrix} \lambda + \begin{vmatrix} A_c^2 & 0 \\ A_c & 0 \end{vmatrix} \frac{\lambda^2}{2!} + \begin{vmatrix} A_c^3 & 0 \\ A_c^2 & 0 \end{vmatrix} \frac{\lambda^3}{3!} + \ldots \\ &= \begin{vmatrix} I + A_c \lambda + A_c^2 \frac{\lambda^2}{2!} + A_c^3 \frac{\lambda^3}{3!} + \ldots & 0 \\ \lambda + A_c \frac{\lambda^2}{2!} + A_c^2 \frac{\lambda^3}{3!} + \ldots & I \end{vmatrix} \\ &= \begin{vmatrix} e^{A_c \lambda} & 0 \\ \int e^{A_c \lambda} d\lambda & I \end{vmatrix} \end{aligned} \tag{23}$$

By substitution of the matrix exponential $e^{A_e \lambda}$ as given by Eq. (23) into Eq. (22) and multiplying matrices, we have the resulting expression for computing the covariance matrix Q

$$
\begin{aligned}
Q &= \int_0^{\delta_l} e^{A_e\lambda} G_e Q_c G_e^T e^{A_e^T\lambda} d\lambda \\
&= \left| \begin{array}{cc} \overbrace{\int_0^{\delta_l} e^{A_c\lambda} G_c Q_c G_c^T e^{A_c^T\lambda} d\lambda}^{Q_s} & \int_0^{\delta_l} e^{A_c\lambda} G_c Q_c G_c^T \left(\int e^{A_c^T\lambda} d\lambda\right) d\lambda \\ \int_0^{\delta_l} \left(\int e^{A_c\lambda} d\lambda\right) G_c Q_c G_c^T e^{A_c^T\lambda} d\lambda & \underbrace{\int_0^{\delta_l} \left(\int e^{A_c\lambda} d\lambda\right) G_c Q_c G_c^T \left(\int e^{A_c^T\lambda} d\lambda d\lambda\right)}_{Q_I} \end{array} \right|
\end{aligned}
$$

Thus the integrated state's covariance matrix Q_I is

$$
\mathbb{E}[((x_I - \mathbb{E}[x_I])(x_I - \mathbb{E}[x_I])^T] = \int_0^{\delta_l} \left(\int e^{A_c\lambda} d\lambda\right) G_c Q_c G_c^T \left(\int e^{A_c^T\lambda} d\lambda\right) d\lambda \quad (24)
$$

The final result follows by substitution of Eqs. (24) and (20) in the first term of Eq. (19), and by substituting Eq. (14) in the second term of Eq. (19). □

In summary, in this section we have found the expression of the measurement variance as a function of δ_l. In the next section we show how the optimal δ_l^* can be found to minimize the output variance.

5. Minimization of the Measurement Variance

In this section, we provide one of the main contributions of the article, namely to show how to compute the optimal averaging time that minimizes the variance of the measured output, due to the trade-off between sensor noise and process noise. With the optimal averaging time, we can program the sensors averaging time to provide, in expectation, the measurement with minimum variance.

Given the expression of the measured output variance, the objective is to minimize it by solving the following problem

$$
\min_{\delta_l \in [T_{SH}, \mu_k]} c_l \left(\frac{1}{\delta_l^2} \int_0^{\delta_l} \left(\int e^{A_c\lambda} d\lambda\right) G_c Q_c G_c^T \left(\int e^{A_c^T\lambda} d\lambda\right) d\lambda \right) c_l^T + \frac{r_{cl}}{\delta_l} \quad (25)
$$

Note that this is a constrained optimization problem, because the averaging time δ_l must be greater than the sampler and hold time T_{SH}, and smaller than the time between the communication of measurements μ_k. The first term of the cost function (25), despite its apparent complexity, has a series expansion as shown in Appendix 9. For the sake of readability define $M := G_c Q_c G_c^T$, hence we have

$$\int_0^{\delta_l} \left(\int e^{A_c \lambda} d\lambda \right) M \left(\int e^{A_c^T \lambda} d\lambda \right) d\lambda = \sum_{n=0}^{\infty} \left(\sum_{j=0}^{n} \frac{A_c^{n-j} M (A_c^j)^T}{(n+1-j)!(j+1)!} \right) \frac{\delta_l^{n+3}}{n+3} \quad (26)$$

As a result, the cost function in Eq. (25) can be approximated using a finite number N of terms of the series as

$$\underbrace{c_l \left(\sum_{n=0}^{N} \left(\sum_{j=0}^{n} \frac{A_c^{n-j} M (A_c^j)^T}{(n+1-j)!(j+1)!} \right) \frac{\delta_l^{n+1}}{n+3} \right) c_l^T}_{\text{scalar}} + \frac{r_{cl}}{\delta_l} \quad (27)$$

Cost function (27) is a univariate function of δ_l that is amenable to be minimized by appropriate existing methods. Once solved it provides a sub-optimal value of the averaging time. However, the series being convergent, the accuracy of the solution can be improved by increasing the number of terms N.

One simple approach to optimize cost function (27) is the use of search methods (e.g. Golden search). Another approach is to use standard calculus to obtain the necessary condition for a minimum as follows

$$\begin{aligned} \frac{d}{d\delta_l} \left(c_l \left(\sum_{n=0}^{N} \left(\sum_{j=0}^{n} \frac{A_c^{n-j} M (A_c^j)^T}{(n+1-j)!(j+1)!} \right) \frac{\delta_l^{n+1}}{n+3} \right) c_l^T + \frac{r_{cl}}{\delta_l} \right) &= 0 \\ c_l \left(\sum_{n=0}^{N} \left(\sum_{j=0}^{n} \frac{A_c^{n-j} M (A_c^j)^T}{(n+1-j)!(j+1)!} \right) \frac{n+1}{n+3} \delta_l^{n+2} \right) c_l^T - r_{cl} &= 0 \quad (28) \end{aligned}$$

As a result the possible candidates for optimal averaging time are given as the roots of Eq. (28). Eq. (28) is in fact a univariate polynomial in δ_l, so optimal averaging time candidates are the roots of a univariate polynomial, which is a standard problem.

Example 3. We follow with example 1. Lets compute the first term of the cost function (27). Recall matrix A_c is nilpotent, as a result the computation of the first term of Eq. (27) is exact for $N = 2$. First lets compute the following series expansion

$$\sum_{n=0}^{N=2} \left(\sum_{j=0}^{n} \frac{A_c^{n-j} M (A_c^j)^T}{(n+1-j)!(j+1)!} \right) \frac{\delta_l^{n+1}}{n+3} = M \frac{\delta_l}{3} + \left(\frac{A_c M}{2!1!} + \frac{M A_c^T}{1!2!} \right) \frac{\delta_l^2}{4} + \frac{A_c M A_c^T}{2!2!} \frac{\delta_l^3}{5}$$

$$= \begin{vmatrix} 0 & 0 \\ 0 & q_a \end{vmatrix} \frac{\delta_l}{3} + \left(\frac{1}{2} \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} \begin{vmatrix} 0 & 0 \\ 0 & q_a \end{vmatrix} + \frac{1}{2} \begin{vmatrix} 0 & 0 \\ 0 & q_a \end{vmatrix} \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} \right) \frac{\delta_l^2}{4} +$$

$$+ \frac{1}{4} \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} \begin{vmatrix} 0 & 0 \\ 0 & q_a \end{vmatrix} \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} \frac{\delta_l^3}{5}$$

$$= \begin{vmatrix} 0 & 0 \\ 0 & q_a \end{vmatrix} \frac{\delta_l}{3} + \frac{1}{2} \begin{vmatrix} 0 & q_a \\ q_a & 0 \end{vmatrix} \frac{\delta_l^2}{4} + \frac{1}{4} \begin{vmatrix} q_a & 0 \\ 0 & 0 \end{vmatrix} \frac{\delta_l^3}{5}$$

$$= \begin{vmatrix} \frac{q_a}{20} \delta_l^3 & \frac{q_a}{8} \delta_l^2 \\ \frac{q_a}{8} \delta_l^2 & \frac{q_a}{3} \delta_l \end{vmatrix}$$

The minimum variance problem of the position measurement has $c_1 = |1 \quad 0|$, with r_{c1} the continuous time variance of the position sensor. The optimization problem, which is solvable by standard calculus, and the optimal integration time for the position sensor δ_1^* are

$$\min_{\delta_1} \frac{q_a}{20} \delta_1^3 + \frac{r_{c1}}{\delta_1} \quad \rightarrow \quad \delta_1^* = \sqrt[4]{\frac{20}{3} \frac{r_{c1}}{q_a}} \tag{29}$$

Likewise, for the velocity output we have $c_2 = |0 \quad 1|$, with r_{c2} the continuous time variance of the velocity sensor. The optimization problem and the optimal integration time for the velocity sensor δ_2^* are

$$\min_{\delta_2} \frac{q_a}{3} \delta_2 + \frac{r_{c2}}{\delta_2} \quad \rightarrow \quad \delta_2^* = \sqrt[2]{3 \frac{r_{c2}}{q_a}} \tag{30}$$

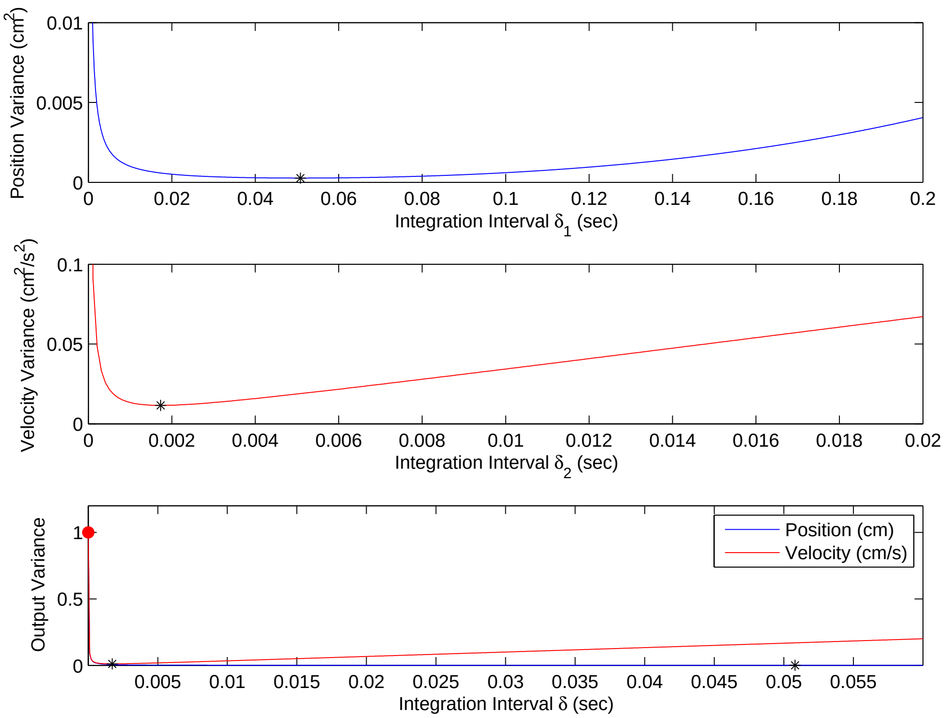

Figure 5: (**a**) Position variance and optimal integration time. (**b**) Velocity variance and optimal integration time. (**c**) Comparisson of position and velocity optimal integration times and achieved output variances. The red dot is the output variance when the averaging time is 10^{-5} seconds.

Consider the following numerical values $q_a = 10 \ \ (cm/sec^2)^2$, $r_{c1} = 10^{-5} \ \ cm^2$, and $r_{c2} = 10^{-5} \ \ (cm/sec)^2$. The cost function together with the optimal variance and integration time are shown in Fig. 5. The optimal integration time for the position variance is $\delta_1^* = 0.0508$ seconds. On the contrary, for the velocity measurement, the optimal integration time is $\delta_2^* = 0.0017$ seconds, a smaller order of magnitude. Although both sensors have the same variance, their optimal averaging time differs because the state variances are different. As a result it is not possible to determine a priori the optimal averaging time based only on the sensor characteristics, and process characteristics must also be considered. Note that the optimal solution depends on the ratio between sensor variance and process variance. If the process variance is much greater than the sensor

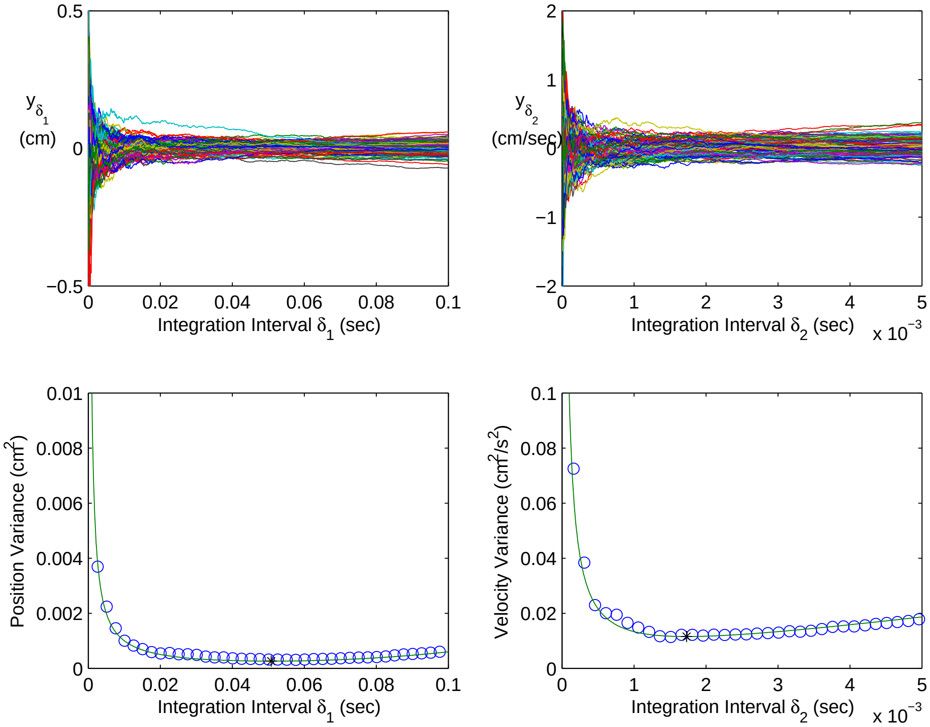

Figure 6: (**a**) Averaged position y_{δ_1} and velocity y_{δ_2} outputs as a function of integration time δ_1 and δ_2 respectively. It can be seen that for optimal averaging values $\delta_1^* = 0.0508$ seconds and $\delta_2^* = 0.0017$ seconds the realizations show the minimum variance. (**b**) The previous intuition is confirmed by computing the numerical variance shown by the blue dots. It can be seen that there is a close match with the theoretical output variance function plotted in green.

variance, the optimal integration time tends to zero, which means instant sampling. On the contrary, if the sensor variance is much greater than the process variance, the optimal integration time tends to infinity, because noise can be better averaged without being affected by process variance.

In Fig. 6 we plot the averaged position y_{δ_1} and averaged velocity y_{δ_2}, as a function of the integration time δ_1 and δ_2 respectively, for 100 realizations. For averaging times smaller than 0.0508 seconds the averaged position y_{δ_1} has higher variance due to sensor noise. For averaging times greater than 0.0508 seconds the averaged position has higher variance due to process noise. Below, we graph the theoretical variance as a function of δ_1 together with the experimental variance computed from the 100 realizations. The match confirms the accordance with the theory. The same discussion can be performed for velocity variance but with optimal averaging time given by 0.0017 seconds.

6. Observer Equations

6.1. From instant error to averaged error

In this section we derive the equations of the Kalman filter with averaged measurements and show how its estimates reduce their variance. First we modify the standard Kalman filter equations to account for the departure from instant sampling values to averaged values. Recall that the standard Kalman Filter estimation equation with instant measurement is

$$\hat{x}_k^+ = \hat{x}_k^- + K_k \underbrace{(y_k - C\hat{x}_k^-)}_{\text{Instant Error}} \tag{31}$$

However with the use of optimally averaged measurements Eq. (31) must be modified as follows

$$\hat{x}_k^+ = \hat{x}_k^- + K_k \underbrace{(y_\delta - C\hat{x}_\delta)}_{\text{Averaged Error}} \tag{32}$$

Thus the error is the difference between averaged measurements y_δ and averaged model estimations $C\hat{x}_\delta$, as shown in Eq. (32). The averaged model estimation $C\hat{x}_\delta$ is defined as

$$C\hat{x}_\delta := \begin{vmatrix} c_1\hat{x}_{\delta_1} \\ c_2\hat{x}_{\delta_2} \\ \vdots \\ c_{n_y}\hat{x}_{\delta_{n_y}} \end{vmatrix} \in \mathbb{R}^{n_y} \tag{33}$$

with $c_l, \quad l \in \{1, 2, \ldots, n_y\}$ the lst row of the output matrix C, and $\hat{x}_{\delta_l}, \quad l \in \{1, 2, \ldots, n_y\}$ the averaged model state with time δ_l, that is given by

$$\hat{x}_{\delta_l} = \frac{1}{\delta_l} \int_{t_k^{y_l}}^{t_{k+1}} \hat{x}(t) dt \tag{34}$$

The computation of the term $\hat{x}_\delta$ requires the integration of the model state evolution $\hat{x}(t)$ given by

$$\hat{x}(t) = e^{A_c(t - t_k^{y_l})} x(t_k^{y_l}) + \int_{t_k^{y_l}}^{t} e^{A_c(t-\tau)} B_c u(\tau) d\tau \tag{35}$$

By substitution of Eq. (35) in Eq. (34) we are able to compute the averaged state with time δ_l as

$$\hat{x}_{\delta_l} = \frac{1}{\delta_l} \int_{t_k^{y_l}}^{t_{k+1}} e^{A_c(t - t_k^{y_l})} dt \hat{x}(t_k^{y_l}) + \int_{t_k^{y_l}}^{t_{k+1}} \int_{t_k^{y_l}}^{t} e^{A_c(t-\tau)} d\tau dt B_c u(\tau) \tag{36}$$

Note, however, that before computing Eq. (36) we still need to compute the state at time $t_k^{y_l}$, that is $\hat{x}(t_k^{y_l})$. This is the time when the sensor starts to average the measurements (See Fig. 2). This is easy because we know the a posteriori estimate at time instant t_k, $\hat{x}_k^+$, which was already calculated in previous iteration by the observer. $\hat{x}(t_k^{y_l})$ is given by

$$\hat{x}(t_k^{y_l}) = e^{A_c(t_k^{y_l} - t_k)} \hat{x}_k^+ + \int_{t_k}^{t_k^{y_l}} e^{A_c(t_k^{y_l} - \tau)} B_c u(\tau) d\tau \tag{37}$$

6.2. Observer Implementation

The averaged Kalman filter measurements are given by the following equations, which are computed for each time step $k = 1, 2, \ldots$, with initial observer variance matrix P_0:

$$P_k^- = A_k P_{k-1}^+ A_k + Q_{k-1} \tag{38}$$

$$K_k = P_k^- C^T (C P_k^- C^T + R_{\delta*})^{-1} \tag{39}$$

$$\hat{x}_k^- = A_{k-1} \hat{x}_{k-1}^+ + B_{k-1} u_{k-1} \tag{40}$$

$$\hat{x}_k^+ = \hat{x}_k^- + K_k (y_\delta - C \hat{x}_\delta) \tag{41}$$

$$P_k^+ = (I - K_k C) P_k^- (I - K_k C)^T + K_k R_{\delta*} K_k^T \tag{42}$$

1. The dynamical system matrices A_k, B_k are given by Eq. (5) and Eq. (6), respectively. The output matrix C is given in Eq. (2).
2. The system covariance matrix Q_{k-1} is given by Eq. (8) and the optimal measurement covariance matrix R_{δ^*} is a diagonal matrix with $l \times l$ entries given by Eq. (18) particularized for each sensor l with optimal averaging time δ_l^*.
3. The averaged model estimations $C\hat{x}_\delta$ are given by computing Eq. (36) for all sensors $l \in \{1, 2, \ldots, n_y\}$ and computing Eq. (33).

In summary, in this section we have provided the new observer equations with optimal averaged measurements. We have referred the observer matrices to the theoretically derived ones in the text. In the Appendix 9 we provide computationally amenable forms to implement the observer numerically.

Example 4. Following the previous example, we consider observer implementation when only velocity is measured. We start by computing $\hat{x}(t_k^{y_l})$ as given by Eq. (37). We make use of the expressions (52) and (53) in the Appendix 9.

$$\hat{x}(t_k^{y_l}) = \begin{vmatrix} 1 & \mu_k - \delta_2 \\ 0 & 1 \end{vmatrix} \hat{x}_k^+ + \begin{vmatrix} \frac{(\mu_k - \delta_2)^2}{2} \\ \mu_k - \delta_2 \end{vmatrix} u_k \tag{43}$$

We now apply Eq. (36) to compute $\hat{x}_{\delta_2}$. Using again (53) results in

$$\hat{x}_{\delta_2} = \begin{vmatrix} 1 & \frac{\delta_2}{2} \\ 0 & 1 \end{vmatrix} \hat{x}(t_k^{y_l}) + \begin{vmatrix} \frac{\delta_2^2}{3!} \\ \frac{\delta_2}{2} \end{vmatrix} u_k \tag{44}$$

Finally we have to compute $c_2\hat{x}_{\delta_2}$, with $c_2 = |0 \;\; 1|$, which results in

$$c_2\hat{x}_{\delta_2} = \begin{vmatrix} 0 & 1 \end{vmatrix} \hat{x}_k^+ + \left(\mu_k - \frac{\delta_2}{2}\right) u_k \tag{45}$$

We implement the modified Kalman filter with a constant measurement communication period of $\mu_k = 2.8 \times 10^{-4}$ seconds. The observer initial state is the zero vector and the initial covariance matrix is $10^6 I_{2\times 2}$. The input is zero up to time 4.48×10^{-2} seconds, a positive ramp up to time 6.44×10^{-2} seconds, and a final negative ramp. Three simulations are performed with sensor averaging times $\delta_2 = \{10^{-5}, 1.7 \times 10^{-4}, 2.8 \times 10^{-4}\}$ seconds. Recall that 1.7×10^{-4} seconds is the optimal averaging time that minimizes the measurement output variance of the measured velocity. In Fig. 7 we can see the state estimates of position and velocity for each averaging time. The left plot represents the position and the right plot the velocity. In blue we have the true state evolution and the red dots show the estimation provided by the observer. The upper line shows the Kalman filter performance for measurements averaged for 10^{-5} seconds. The measurement variance is very high due to sensor noise and, although the Kalman filter improves the state estimation, it is not accurate. The middle line shows the Kalman filter performance for optimal averaging of 1.7×10^{-4} seconds. The observer accuracy is dramatically improved. Finally, the bottom line shows the Kalman filter performance for measurements integrated in 2.8×10^{-4} seconds, a longer time than the optimal integral time. The observer accuracy degrades slightly, although the averaging time has doubled.

7. Application Examples

7.1. Two Tanks [17]

In [17] a laboratory experiment consisting of two interconnected tanks is presented to evaluate the performance of the Integrated Measurement Kalman Filter (IMKF), designed for state estimation and fusion of fast rate measurements with averaged slow rate measurements. In the two-tank system, water is pumped into the left tank which is connected to the middle tank and each tank is equipped with a differential pressure cell (DP-cell) level sensor. The linearized continuous-time model is given by

$$\dot{x}(t) = \underbrace{\begin{vmatrix} -0.0101 & 0.0101 \\ 0.0101 & -0.0147 \end{vmatrix}}_{A_c} x(t) + \underbrace{\begin{vmatrix} 0.0041 \\ 0 \end{vmatrix}}_{B_c} q_i(t) + w(t) \tag{46}$$

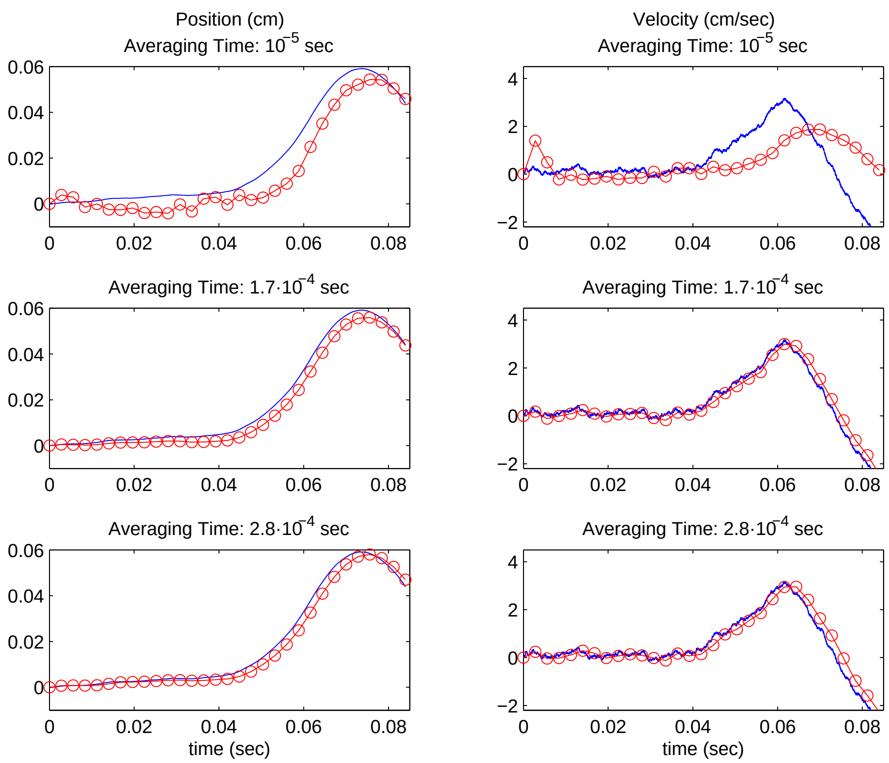

Figure 7: The left plots represent the position and the right plots the velocity. In blue we have the true state evolution and the red dots show the estimation provided by the observer. (**a**) The first line plots show the simulation for $\delta_2 = 10^{-5}$ seconds. (**b**) The second line plots show the Kalman filter performance for $\delta_2 = 1.7 \times 10^{-4}$ seconds, which is the optimal averaging value. (**c**) The third line plots shows the Kalman filter performance for $\delta_2 = 2.8 \times 10^{-4}$ seconds.

Only the height of the middle tank is measured, so the output equation is

$$y(t) = \underbrace{\begin{vmatrix} 0 & 1 \end{vmatrix}}_{C} x(t) + v(t) \tag{47}$$

The process noise and sensor noise are obtained experimentally as

$$Q_c = 10^{-5} \begin{vmatrix} 1.305 & -0.061 \\ -0.061 & 1.198 \end{vmatrix} (cm^2), \quad R_c = 2.420 \times 10^{-6} (cm^2), \tag{48}$$

Furthermore in [17] the middle tank DP-cell is artificially contaminated with some extra Gaussian random noise with variance $R_m = 5 \times 10^{-3} \ \ cm^2$. This is important for the discussion that follows.

Fig. 8 shows the middle tank level variance as a function of averaging time. The upper plot considers only the DP-cell sensor noise with variance $2.420 \times 10^{-6} \ \ cm^2$. The optimal averaging time that provides the minimum variance measurement is $\delta^* = 1.2182$ seconds. In [17] the measurement averaging time proposed is 10 seconds. Note that this provides a measurement variance that is three times longer than the optimal one (see the red dot in Fig. 8). It can be concluded that the 10 second averaging time does not provide a more accurate measurement than the 1 second averaging time.

The lower plot is the middle tank level variance with the DP-cell contaminated with extra Gaussian noise of variance $5 \times 10^{-3} \ \ cm^2$. The optimal averaging time that provides the minimum variance measurement is in this case $\delta^* = 43.5088$ seconds. In [17] the proposed 10 second averaging time provides a measurement variance that is again larger than the optimal one (see the red dot in Fig. 8). However, in this case the measurement variance is similar to the optimal one due to the flatness of the cost function around the minimum. The contribution of this article allows computation of the optimal averaging time so as to produce the minimum variance measurement, to be used for instance in the IMKF.

Finally we assess the impact of measurement accuracy in the modified Kalman filter derived in Section 6. The observer initial state is the zero vector and the

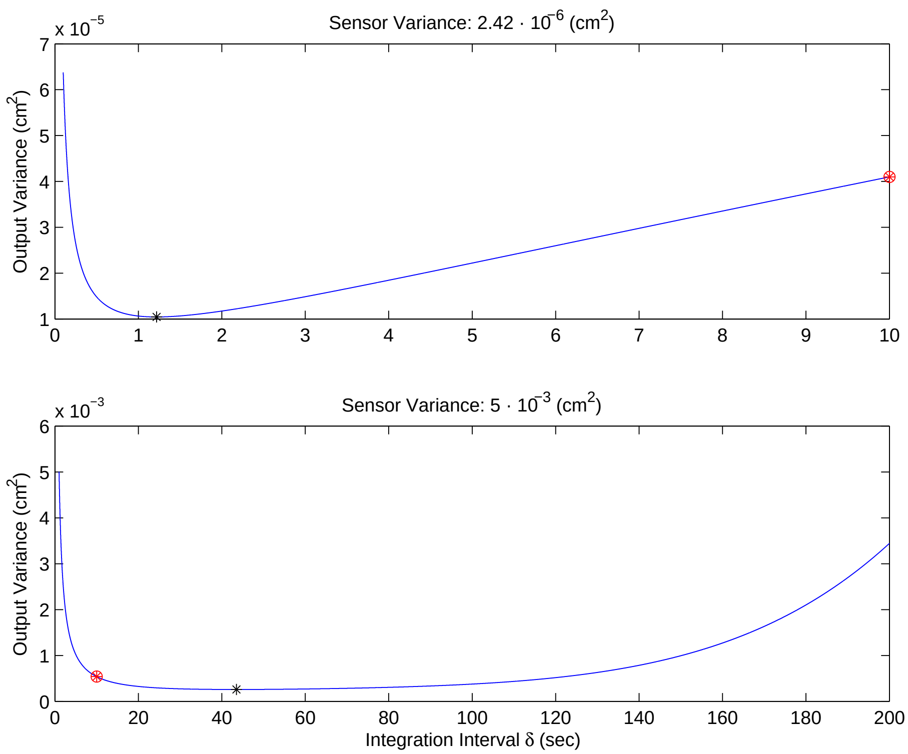

Figure 8: Middle tank level variance as a function of averaging time. (**a**) Upper plot is for sensor noise with variance $2.420 \times 10^{-6}\ \ cm^2$. The optimal averaging time is 1.2182 seconds. (**b**) Lower plot is for sensor noise with variance $5 \times 10^{-3}\ \ cm^2$. The optimal averaging time is 43.5088 seconds

initial covariance matrix is $10^6 I_{2\times 2}$. We apply the modified Kalman filter with three sampling rates of 1 second, 10 seconds, and 40 seconds. The averaging time in each case is equal to the sampling rate. The results can be seen in the first, second, and third rows of Fig. 9, respectively. Note that in this case the data transmission rate is not constant. The best observer performance is obtained with the averaging time of 40 seconds, approximately the optimal one. The observer performance with averaging time of 1 and 10 seconds is less accurate, despite the data transmission rate being higher than in the 40 second case. This example shows the impact of optimal averaging time at the observer node prior to communication to the remote observer.

7.2. Aerial Navigation

We consider the problem of estimating the position of an airplane with a continuous-time model given by

$$\dot{x}(t) = \underbrace{\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}}_{B_c} u(t) + w(t) \tag{49}$$

with $x(t) := [P_x \;\; P_y \;\; P_z]$ the position in the three axes, $u(t) := [V_x \;\; V_y \;\; V_z]$ the input velocity in the three axes, and $w(t)$ a random disturbance. We consider that the airplane is flying at a constant velocity in the x axis, hence $V_x = 10 \;\; m/s$ and $V_y = V_z = 0 \;\; m/s$, with initial altitude of 500 meters. We measure the position on the three axes, hence the output equation is

$$y(t) = \underbrace{\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}}_{C} x(t) + v(t) \tag{50}$$

with $y(t) = [y_x \;\; y_y \;\; y_z]$ the position measures in the three axes, and $v(t)$ the

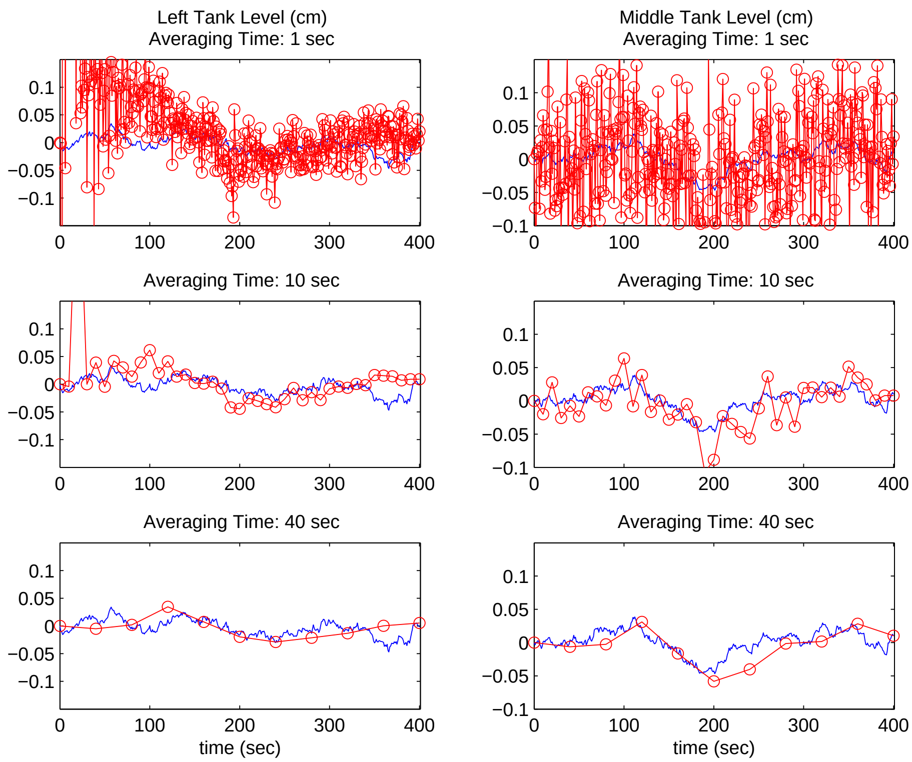

Figure 9: Tank level variations around the set point. The left plots represent the left tank and the right plots the middle tank. In blue we have the true state evolution and the red dots show the estimation provided by the observer. (**a**) Modified Kalman filter performance with averaging time and sampling rate equal to 1 second. (**b**) Modified Kalman filter performance with averaging time and sampling rate equal to 10 seconds. (**c**) Modified Kalman filter performance with averaging time and sampling rate equal to 40 seconds.

sensor noise. The process noise and the sensor noise covariance matrices are given as

$$Q_c = \begin{vmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{vmatrix} (m^2/s^2), \quad R_c = \begin{vmatrix} 0.04 & 0 & 0 \\ 0 & 0.04 & 0 \\ 0 & 0 & 1 \end{vmatrix} (m^2), \tag{51}$$

The sensor noise in the x and y axis measurement has equal variance of value $0.04 \ \ m^2$ because in both cases the position is given as the integration of accelerometer measurements. On the contrary, the position of the axis z (i.e. the altitude) is measured by a radar with larger variance of value $1 \ \ m^2$. However, the radar provides an absolute measure in contrast to the accelerometers that are subject to error accumulation over time.

In Fig. 10 we plot the measurement variance of the accelerometers and the radar as a function of averaging time. The optimal averaging time for the accelerometers is 0.49 seconds, whereas for the radar it is 2.45 seconds. Fig. 11 shows the position estimation in the three axes. The position estimation in the x and y with optimally averaged measurements is shown in red. In black we show the estimation with averaging time of 2.45 seconds, which is larger than the optimal averaging time of 0.49 seconds. We can see that the estimate's accuracy degrades and that there is a lag in the estimates, due to the increase in the averaging time. The position estimation in the z axis, measured with radar, with optimally averaging time is shown in red. In black we show the estimation with averaging time of 0.49 seconds, which is shorter than the optimal averaging time of 2.45 seconds. We can see that the estimate's accuracy degrades because of the sensor noise, due to the decrease in the averaging time. This example shows the impact on appropriate selection of the averaging time on the variance of the Kalman filter estimates.

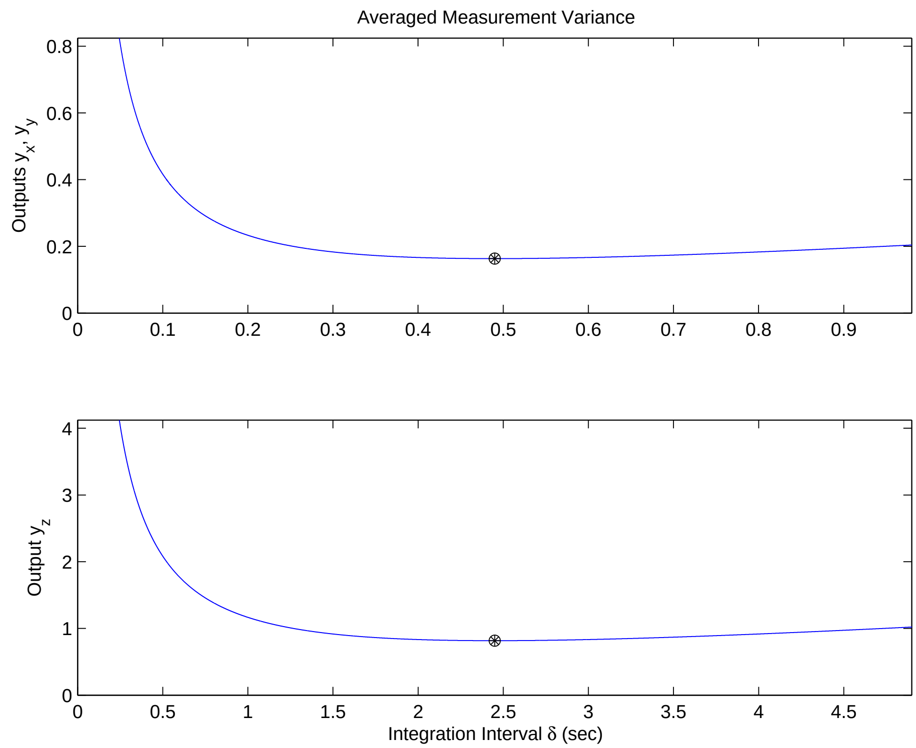

Figure 10: Averaged measurement variance as a function of the averaging time δ for measurement of position y_x, y_y, and y_z. (**a**) Upper plot is for measurement of positions y_x and y_y given by integration of accelerometers with sensor variance 0.04 m^2. The optimal averaging time is 0.49 seconds. (**b**) Lower plot is for measurement of position y_z given by radar with variance 1 m^2. The optimal averaging time is 2.45 seconds.

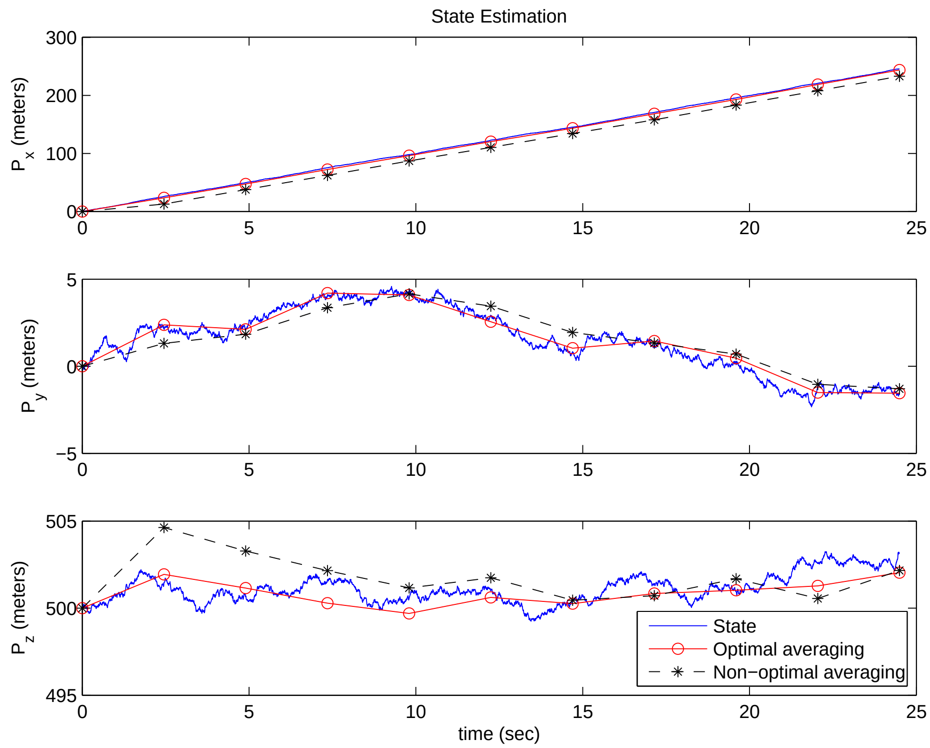

Figure 11: Airplane position estimation. In blue we have the true state evolution, the red dots show the estimation provided by the observer with optimal averaging time, and the black dots show the estimation provided by the observer with non-optimal averaging time. (**a**) In red, estimation with optimal averaging time equal to 0.49 seconds. In black, estimation with non-optimal averaging time equal to 2.45 seconds. (**b**) In red, estimation with optimal averaging time equal to 0.49 seconds. In black, estimation with non-optimal averaging time equal to 2.45 seconds. (**c**) In red, estimation with optimal averaging time equal to 2.45 seconds. In black, estimation with non-optimal averaging time equal to 0.49 seconds.

8. Conclusions

In this article, we have shown that for stochastic dynamical systems there is an optimal averaging time that provides the measurement with minimum variance. The existence of the optimal averaging time is an outcome of the trade-off between process noise and sensor noise, which have competing effects on the variance of the averaged measurement. This optimal averaging time is obtained by solving a univariate optimization problem. However, for a particular realization, the optimal averaging time may differ from the optimal in expected averaging time. This opens the possibility for programming smart sensors with algorithms that provide minimum variance measures for any realization. Finally, the a posteriori estimation equations of the classical Kalman filter have been adapted to take into account averaged measurements. The new filter might be designed to tackle the trade-off between sampling time and accuracy, in such a way that the optimal averaging time could be obtained by minimizing the estimate variance instead of the measurement variance. This is an open research area under investigation.

Acknowledgements

We would like to thank Jorge Galindo Pastor and Fernando Casas Pérez from the Mathematics Research Institute (IMAC) at University Jaume I of Castellón for its invitation to attend the Mini-course on Stochastic Differential Equations lectured by David Cohen from Umeå University.

9. Appendix

The article results require computation of integrals of exponential matrices that, in general, have no closed form. Although there are well established numerical methods to calculate these integrals, we still need to obtain the calculations as a function of a generic time interval for optimization purposes. To accomplish this goal we make use of the series expansion of the matrix exponential [24]

$$e^{A_c\lambda} = I + A_c\lambda + A_c^2\frac{\lambda^2}{2!} + A_c^3\frac{\lambda^3}{3!} + \dots \tag{52}$$

As a result, by substituting the term $e^{A_c\lambda}$ with its series expansion we are able to approximate the following integrals that appear throughout the article.

The series expansion is absolutely convergent [25], thus the integral of the exponential matrix is

$$\int_0^\lambda e^{A_c\lambda}d\lambda = \sum_{n=0}^{\infty}\frac{A_c^n}{(n+1)!}\lambda^{n+1} \tag{53}$$

This is used in the discretization of the dynamical model in (6) and in the computation of the state and the averaged state in Eqs. (35) and (36).

The following integral is required to compute the covariance of the state space in Eq. (8)

$$\int_0^\lambda e^{A_c\lambda}Me^{A_c^T\lambda}d\lambda = \sum_{n=0}^{\infty}\left(\sum_{j=0}^{n}\frac{A_c^{n-j}M(A_c^j)^T}{(n-j)!j!}\right)\frac{\lambda^{n+1}}{n+1} \tag{54}$$

Finally, the next integral computes the covariance matrix of the integral of the state space, which appears in the minimization problem of the output variance in Eq. (25)

$$\int_0^\lambda\left(\int e^{A_c\lambda}d\lambda\right)M\left(\int e^{A_c^T\lambda}d\lambda\right)d\lambda = \sum_{n=0}^{\infty}\left(\sum_{j=0}^{n}\frac{A_c^{n-j}M(A_c^j)^T}{(n+1-j)!(j+1)!}\right)\frac{\lambda^{n+3}}{n+3} \tag{55}$$

By truncating the expansion of the previous convergent series by taking its first $N+1$ terms, the desired accuracy approximation can be found. In Table 1 we summarize the computation formulae for the truncated series up to a certain index N. Note that the computation is a univariate matrix polynomial in λ.

Table 1: Summary of the proposed computation formulae

Definition	**Computation**
$e^{A_c \lambda}$	$\sum_{n=0}^{N} \frac{A_c^n}{n!} \lambda^n$
$\int_0^{\lambda} e^{A_c \lambda} B_c d\lambda$	$\sum_{n=0}^{N} \frac{A_c^n B_c}{(n+1)!} \lambda^{n+1}$
$\int_0^{\lambda} e^{A_c \lambda} M e^{A_c^T \lambda} d\lambda$	$\sum_{n=0}^{N} \left(\sum_{j=0}^{n} \frac{A_c^{n-j} M (A_c^j)^T}{(n-j)! j!} \right) \frac{\lambda^{n+1}}{n+1}$
$\int_0^{\lambda} \left(\int e^{A_c \lambda} d\lambda \right) M \left(\int e^{A_c^T \lambda} d\lambda \right) d\lambda$	$\sum_{n=0}^{N} \left(\sum_{j=0}^{n} \frac{A_c^{n-j} M (A_c^j)^T}{(n+1-j)!(j+1)!} \right) \frac{\lambda^{n+3}}{n+3}$

References

[1] J. Zheng, A. Jamalipour, Wireless sensor networks : a networking perspective, Wiley-IEEE Press, 2009.

[2] X. Ge, F. Yang, Q.-L. Han, Distributed networked control systems: A brief overview, Information Sciences 380 (2017) 117–131. doi:10.1016/J.INS.2015.07.047.

[3] T. Tran, Q. Ha, Dependable control systems with Internet of Things, ISA Transactions 59 (2015) 303–313. doi:10.1016/J.ISATRA.2015.08.008.

[4] J. Wu, Q.-S. Jia, K. H. Johansson, L. Shi, Event-Based Sensor Data Scheduling: Trade-Off Between Communication Rate and Estimation Quality, IEEE Transactions on Automatic Control 58 (4) (2013) 1041–1046. doi:10.1109/TAC.2012.2215253.

[5] A. S. Leong, S. Dey, D. E. Quevedo, Sensor Scheduling in Variance Based Event Triggered Estimation With Packet Drops, IEEE Transactions on Automatic Control 62 (4) (2017) 1880–1895. doi:10.1109/TAC.2016.2602499.

[6] D. Shi, L. Shi, T. Chen, Event- Based State Estimation, Springer, 2016.

[7] L. Shi, P. Cheng, J. Chen, Optimal Periodic Sensor Scheduling With Limited Resources, IEEE Transactions on Automatic Control 56 (9) (2011) 2190–2195. doi:10.1109/TAC.2011.2152210.

[8] S. C. Dutta Roy, A State-of-the-Art Survey on Linear Phase Digital Filter Design, in: Topics in Signal Processing, Springer Singapore, Singapore, 2020, pp. 193–200. doi:10.1007/978-981-13-9532-1_18.
URL http://link.springer.com/10.1007/978-981-13-9532-1{_}18

[9] N. Singh, A. Potnis, A review of different optimization algorithms for a linear phase FIR filter, in: International Conference on Recent Innovations in Signal Processing and Embedded Systems, RISE 2017, Vol. 2018-January, Institute of Electrical and Electronics Engineers Inc., 2018, pp. 44–48. doi:10.1109/RISE.2017.8378122.

[10] R. V. Ravi, K. Subramaniam, T. V. Roshini, S. P. B. Muthusamy, G. K. D. Prasanna Venkatesan, Optimization algorithms, an effective tool for the design of digital filters; a review, Journal of Ambient Intelligence and Humanized Computing (sep 2019). doi:10.1007/s12652-019-01431-x.

[11] S. Labs, AN118 Improving ADC resolution by oversampling and averaging, Tech. rep. (2013).

[12] T. Hägglund, Signal Filtering in PID Control*, Tech. rep. (2012). doi:10.3182/20120328-3-IT-3014.00002.

[13] W. Kester, J. Bryant, MT-027 Tutorial ADC Architectures VIII: Integrating ADCs, Tech. rep.

[14] F. L. Lewis, L. Xie, D. Popa, F. L. Lewis, Optimal and robust estimation : with an introduction to stochastic control theory, CRC Press, 2007, 2008.

[15] J. L. Crassidis, J. L. Junkins, Optimal Estimation of Dynamic Systems, Second Edition, Chapman and Hall/CRC, 2011, 2000.

[16] Y. Guo, B. Huang, State estimation incorporating infrequent, delayed and integral measurements, Automatica 58 (2015) 32–38. doi:10.1016/J.AUTOMATICA.2015.05.001.

[17] A. Fatehi, B. Huang, State Estimation and Fusion in the Presence of Integrated Measurement, IEEE Transactions on Instrumentation and Measurement 66 (9) (2017) 2490–2499. doi:10.1109/TIM.2017.2701143.

[18] D. Simon, Optimal state estimation : Kalman, H_∞ and nonlinear approaches, Wiley-Interscience, 2006.

[19] B. K. Øksendal, Stochastic differential equations : an introduction with applications, Springer, 2003.

[20] L. C. Evans, An introduction to stochastic differential equations, Mathematical Association of America.

[21] A. Gelb, J. F. Kasper Jr., R. A. Nash Jr., C. F. Price, A. A. Sutherland Jr., Applied Optimal Estimation, no. 16, 2001. doi:10.1109/PROC.1976.10175.

[22] C. Van Loan, Computing integrals involving the matrix exponential, IEEE Transactions on Automatic Control 23 (3) (1978) 395–404. doi:10.1109/TAC.1978.1101743.

[23] D. P. Frank L. Lewis, Lihua Xie, Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory, Second Edition.

[24] R. A. Horn, C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.

[25] C. Moler, C. V. Loan, Nineteen Dubious Ways to Compute the Exponential of a Matrix , Twenty-Five Years, SIAM Review 45 (1) (2003) 801–836.