Multimodal Kalman Filtering

Anthony Bourrier, Pierre-Olivier Amblard, Olivier Michel, Christian Jutten

MULTIMODAL KALMAN FILTERING

Anthony Bourrier *Pierre-Olivier Amblard* *Olivier Michel* *Christian Jutten*

Gipsa-Lab, 11 Rue des Mathématiques, 38400 St-Martin-d'Hères, France

ABSTRACT

A difficult aspect of multimodal estimation is the possible discrepancy between the sampling rates and/or the noise levels of the considered data. Many algorithms cope with these dissimilarities empirically. In this paper, we propose a conceptual analysis of multimodality where we try to find the "optimal" way of combining modalities. More specifically, we consider a simple Kalman filtering framework where several noisy sensors with different sampling frequences and noise variances regularly observe a hidden state. We experimentally underline some relationships between the sampling grids and the asymptotic variance of the maximum a posteriori (MAP) estimator. However, the explicit study of the asymptotic variance seems intractable even in the simplest cases. We describe a promising idea to circumvent this difficulty: exploiting a stochastic measurement model for which one can more easily study the average asymptotic behavior.

1. INTRODUCTION

The term "multimodality" generally refers to the observation of a latent phenomenon through different acquisition media [1]. Making the most of these different measurements is still a challenge, and may lead to better estimates of the latent phenomenon than the estimation from a sole set of measurements. However, a recurrent problem in multimodal estimation is the discrepancy between data, which may not have the same nature, dimension, sampling rate, noise level or time delay. Examples of applications for which such discrepancies occur include neuronal activity acquisition devices [2] or remote sensing [3].

In this work, we propose to consider the problem of multimodal estimation from a more theoretical point of view, basically considering the question: What is the best possible estimation one can obtain from multimodal measurements? To tackle this problem, we consider a simple multimodal model where one observes a continuous latent variable through different noisy sensors, each having its own measurement noise variance and its own sampling frequency. Taking a Kalman filter [4] based approach to estimate the hidden state, the most natural way to evaluate the quality of estimation consists in studying the asymptotic mean variance of the estimator.

The paper is organized as follows: in Section 2, we present the multimodal Kalman estimation model we consider. In Section 3, we experimentally underline the sampling layout which minimizes the asymptotic variance of the estimator in the case of two modalities. Since the explicit layout is too tedious to compute even in simple cases, we present in Section 4 a promising approach to provide bounds on the variance of such a multimodal estimator by replacing the deterministic model by a stochastic model.

This work has been partly supported by the European project ERC-2012-AdG-320684-CHESS.

Earlier works. Multimodal Kalman filtering has mainly been considered in the case of sensor networks, where several sensors measuring the same hidden state are connected. In this setup, most work focus on estimation with constraints such as decentralized estimation [5], unreliable communication channels subject to packet losses [6] or time delayed observations [7]. We focus here on a centralized multimodal estimation with no constraints, showing that even this simple case is not fully understood. Kalman filtering with an stochastic observation model has been extensively studied [8, 9, 10] and applied to multimodal estimation for sensor networks [11]. We believe that such studies have the potential to convey answers on the optimal sampling layout even in the centralized unconstrained case we consider.

2. MODEL DESCRIPTION

Consider a real Brownian motion θ_t, satisfying, for $t > s$, $\theta_t - \theta_s \sim \mathcal{N}(0, (t-s)\sigma^2)$. Suppose n sensors can make measurements of the form

$$X_t^i = \theta_t + N_t^i \tag{1}$$

at time t, where $1 \leq i \leq n$ corresponds to the index of the sensor and N_t^i is a centered white Gaussian noise of variance v_i. We suppose N_t^i and N_t^j are independent for $i \neq j$ and that sensor i performs regular measurements with period T_i. Without loss of generality, we can suppose that $\sigma = 1$ (if not, replacing each T_i by T_i/σ boils down to the same model with time being dilated by a factor $1/\sigma$).

The Kalman filter framework can apply and the maximum likelihood estimate $\hat{\theta}_t$ of θ_t can be computed at any time t, supposing we have an unbiased estimate $\hat{\theta}_0$ of θ_0 with variance V_0. In this case, denoting $V_t = \text{Var}(\hat{\theta}_t)$ and s the time were the latest measurement was performed (by any sensor), we have at time t [12]:

- if there was no measurements between s and t,

$$\hat{\theta}_t = \hat{\theta}_s \quad \text{and} \quad V_t = V_s + (t-s)\sigma^2 = V_s + (t-s); \tag{2}$$

- if a measurement is performed at time t by the i^{th} sensor,

$$\hat{\theta}_t = \hat{\theta}_s + \frac{V_s + (t-s)}{v_i + V_s + (t-s)}(X_t^i - \hat{\theta}_s) \tag{3}$$

$$\text{and} \quad V_t = \frac{v_i(V_s + (t-s))}{v_i + V_s + (t-s)}. \tag{4}$$

Our main goal is to compare the behavior of V_t in a multimodal case, quantifying the gain in the expected mean square error (MSE) when one exploits several types of measurements. To this end, we first establish the behavior of V_t under unimodal regular sampling.

2.1. Unimodal estimation

Let us suppose $X_t^1 = X_t = \theta_t + N_t$, the only observation of θ_t with measurements taken at times $T\mathbb{N} = \{0, T, 2T, \dots\}$ and note $v_1 = v$ the noise variance. We can explicit the asymptotic mean value of V_t.

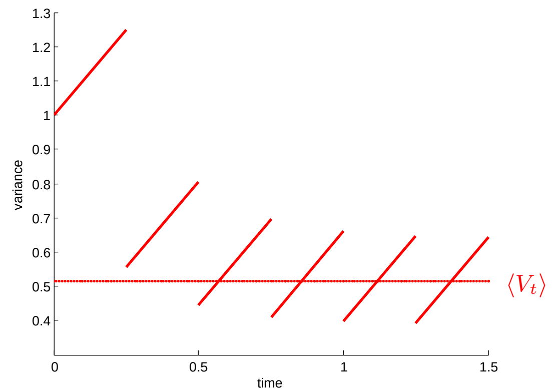

Fig. 1. Representation of V_t with respect to t for a unimodal estimation of a Brownian motion. The parameters used are $T = 0.25$ and $v = 1$, and the Kalman filter is initialized with an estimated value of variance v at $t = 0$. The discontinuities correspond to the measurements of the sensor (update equation (4)), the lines correspond to the linear increase of V_t in the absence of measurement (update equation (2)). The dotted line represents the asymptotic mean value of V_t.

Property 1. *Suppose X_t observes a Brownian motion of variance 1 at times $T\mathbb{N}$, with a white Gaussian measurement noise N_t of variance v. Then*

$$\langle V_t \rangle = \lim_{k\to\infty} \frac{1}{T}\int_{kT}^{(k+1)T} V_t dt = \frac{T}{2}\sqrt{1+4\frac{v}{T}}. \qquad (5)$$

A representation of V_t in a simple case is given in Figure 1.

2.2. Multimodal estimation

Let us now suppose $n > 1$. In this case, the general sampling grid is not entirely determined by the sampling periods T_i but also depends on the "shifts" of the sampling grid of each sensor relatively to the origin (all sensors do not necessarily begin to sample at time 0). We therefore consider that sensor i samples at times $h_i + T_i\mathbb{N} = \{h_i, h_i + T_i, h_i + 2T_i, \ldots\}$. For $1 \le i \le n$ and $t \ge 0$, denote

$$f_{i,t} : x \mapsto \frac{v_i(x+t)}{v_i + x + t}. \qquad (6)$$

If a measurement is performed at time t by the i^{th} sensor, the variance is updated as $V_t = f_{i,t-s}(V_s)$, where s is the time when the previous measurement was performed. An illustration of the sampling grid and the corresponding updates of the variances is illustrated in Figure 2 in the case of two modalities.

Let us establish that if the ratios between the T_i's are all rational, then V_t has a periodic asymptotic behavior. Let us first consider the case of two modalities with $pT_1 = qT_2$ $(p, q \in \mathbb{N})$ and $h_1 = h_2 = 0$. The measurements up until time pT_1 are performed at times kT_1, $k \in [\![0, p-1]\!]$ for modality 1 and at times kT_2, $k \in [\![0, q-1]\!]$ for modality 2. At time $pT_1 = qT_2$, a measurement is performed by both modalities (as for $t = 0$) and the layout of the measurements for times $[pT_1, 2pT_1[$ is the same as for times $[0, pT_1[$. This periodicity in the layout of the measurements also holds for any offset h, that is the layout of the measurements is the same in all intervals of the form $[h + kpT_1, h + (k+1)pT_1[$, $k \in \mathbb{N}$. When h_1 and/or h_2 are nonzero, this periodicity is subsequently conserved.

In the case of n modalities, it is sufficient to suppose that $p_iT_1 = q_iT_i$ with all p_i's and q_i's integers, in which case all T_i/T_j are rational. The layout of the measurements is the same in all intervals of the form $[h + kpT_1, h + (k+1)pT_1[$, $k \in \mathbb{N}$, where p is the least common multiple of the p_i's. Therefore, if a measurement is performed at time $s \in [0, pT_1[$, the variance at time $t_k = t + kpT_1$ can be expressed as $V_{t_k} = f_s^k(s)$, where f_s is a composition of functions of the form $f_{i,t}$, corresponding to the change in the variance depending on the layout of the measurements between times t_k and t_{k+1}.

Since all $f_{i,t}$ are homographic functions with positive coefficients, any function f_s is also a homographic function with positive coefficients. The study of such a function shows that it always has a unique positive fixed point v_s. Since f_s is bounded and increasing, the sequence $(V_{t_k})_{k\ge0}$ converges to v_s. Finally, the asymptotic behavior of V_t is periodic with period pT_1.

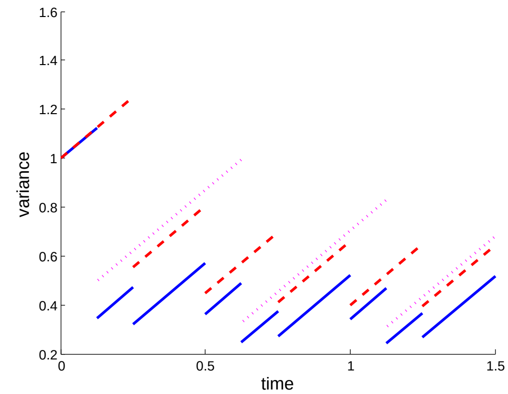

Fig. 3. Bimodal estimation of a Brownian motion. Plot of the variance when using 2 modalities (blue solid line) and using only modality 1 (dashed line) or 2 (dotted line). The parameters used are $T_1 = 0.25 = T_2/2$, $v_1 = 1 = 2v_2$, $h_1 = 0$ and $h_2 = T_1/2$. The discontinuities are located at the measurement times of any sensor. Note that when taking into account both sets of measurements, the variance is below the variance when taking only one measurement.

3. WHAT IS THE BEST SAMPLING GRID?

In this section, we consider the case where the measurements are performed with two sensors. In this case, the behavior of V_t is represented in Figure 3 (see explanations in the caption). Figure 4 shows that the shifts h_i may have an importance in the value of $\langle V_t \rangle$, which is an indicator of the average asymptotic quality of the estimator. In the example of Figure 4, the minimal asymptotic variance is achieved when $h_2 \approx T_1/2$, meaning that to reduce the variance of the estimator through time, one should spread measurements over time instead of making several measurements at the same time. Also note that the shift that minimizes the inferior limit of the variance (which corresponds to the asymptotic variance after the "best" measurement) does not necessarily minimize the asymptotic mean variance.

By analogy with the unimodal case of Section 2.1, one can aim at obtaining a closed form expression for $\langle V_t \rangle$ in the multimodal case. To illustrate the difficulty of the asymptotic study, let us consider the

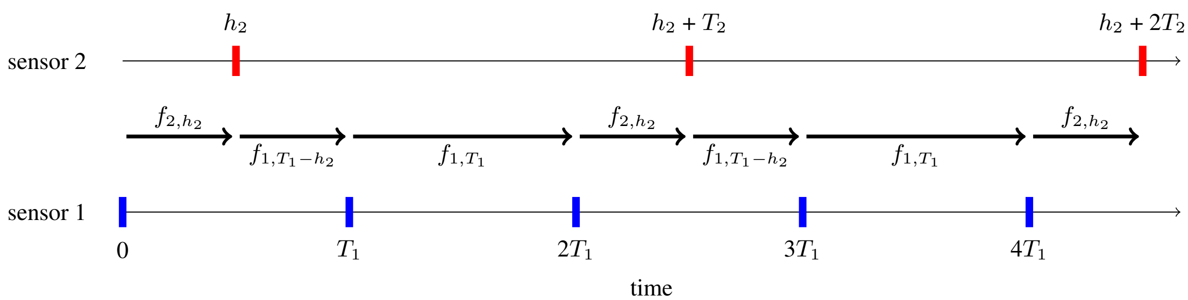

Fig. 2. Updates on V_t in the case of two modalities, with $T_2 = 2T_1 = 1$, $h_1 = 0$ and $h_2 = 1/2$. Between each successive timestamp is represented the function applied to the variance to get the new value. Since T_1/T_2 is rational, the sampling grid exhibits a periodic pattern (see the end of Section 2.2).

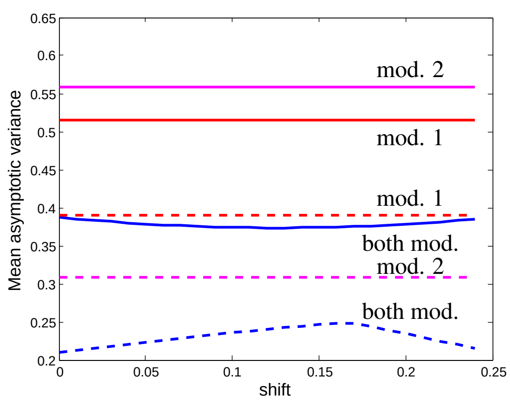

Fig. 4. Influence of the shift h_2 of modality 2 with respect to modality 1 on the value of $\langle V_t \rangle$. Here, $T_1 = T_2 = 1$. Solid lines represent the value of $\langle V_t \rangle$ when taking into account the two modalities or either one, dashed lines represent the values of $\liminf V_t$.

simple bimodal framework where $T_1 = T_2 = T$ and $h_1 = 0$. Let us denote $h = h_2$. In this case, there are two different update steps for the variance : $f_{2,h}$ and $f_{1,T-h}$. Asymptotically, V_{nT} converges to the fixed point of $f_{1,T-h} \circ f_{2,h}$ while V_{nT+h} converges to the fixed point of $f_{2,h} \circ f_{1,T-h}$. One can compute explicit expressions for these fixed points because these functions are homographies. The unique positive fixed point of the homography $x \mapsto \frac{ax+b}{cx+d}$ is expressed as

$$\frac{1}{2c}\left[\sqrt{(d-a)^2 + 4bc} + a - d\right] \tag{7}$$

This allows us to explicitly compute the fixed points λ_1 and λ_2 of $f_{1,T-h} \circ f_{2,h}$ and $f_{2,h} \circ f_{1,T-h}$. The asymptotic mean of V_t is then given by

$$\langle V_t \rangle = \frac{1}{T}\left[h\left(\lambda_1 + \frac{h}{2}\right) + (T-h)\left(\lambda_2 + \frac{T-h}{2}\right)\right]. \tag{8}$$

Finding the best shift h boils down to minimizing expression (8). However, the expressions of λ_1 and λ_2 are heavy and this minimization is tedious even in this simple case. The situation is even worse when more modalities are considered or when the sampling periods are not equal since the expression of $\langle V_t \rangle$ involves fixed points of compositions of many functions $f_{i,t}$. Figure 5 represents the quantity (8) with respect to h for several values of T, v_1 and v_2. Two conclusions can be drawn from Figure 5:

- Figure 5 *Left*: As the (common) sampling period T increases, so is the importance of the shift in the value of $\langle V_t \rangle$. Therefore, the fewer the measurements the more crucial it is to spread them well over time.
- Figure 5 *Right*: As the ratio v_2/v_1 increases, the importance of the shift in the value of $\langle V_t \rangle$ decreases. When a sensor is much less reliable than the other, the placement of the sampling grid is less important than when the two sensors have similar precision.

However the general study of the covariance of a multimodal Kalman estimator, even in elementary cases, seems too tedious to be performed. The next section presents a promising approach to provide a study of multimodal Kalman filtering on a slightly different model.

4. OUTLOOK: STOCHASTIC MODELING

In [8], the authors introduce a multidimensional stochastic measurement model where time is discretized and at a time t, a (unique) sensor has probability $0 \le p \le 1$ to perform a measurement while no measurement is performed with probability $1-p$, these measurement triggers being independent of one another. More precisely, the model is:

$$\theta_{t+1} = A\theta_t + N_t^\theta \tag{9}$$
$$X_t = C\theta_t + N_t, \tag{10}$$

where θ_t is now a multidimensional vector and N_t^θ and N_t being centered Gaussian noises of respective covariance matrices Q and R. Denote V_t the covariance matrix of the MAP estimator of θ_t knowing all observations up to time t. At time t, the update equation for V_t depends on whether a measurement has been performed or not at time t. Therefore V_t becomes a random variable which depends on the times and locations of the previous measurements. The authors study the asymptotic behavior of $\mathbb{E}[V_t]$, proving in particular that $\mathbb{E}[V_t]$ is upper bounded at each time by a matrix U_t which is obtained by a recursive relationship with a unique function. Thus, even though the behavior of V_t is random and can involve many different updates layouts, the average behavior can be upper bounded by a recursive sequence with a unique update equation.

Subsequent works using such a stochastic formulation have been performed to study sensor networks with unreliable communication

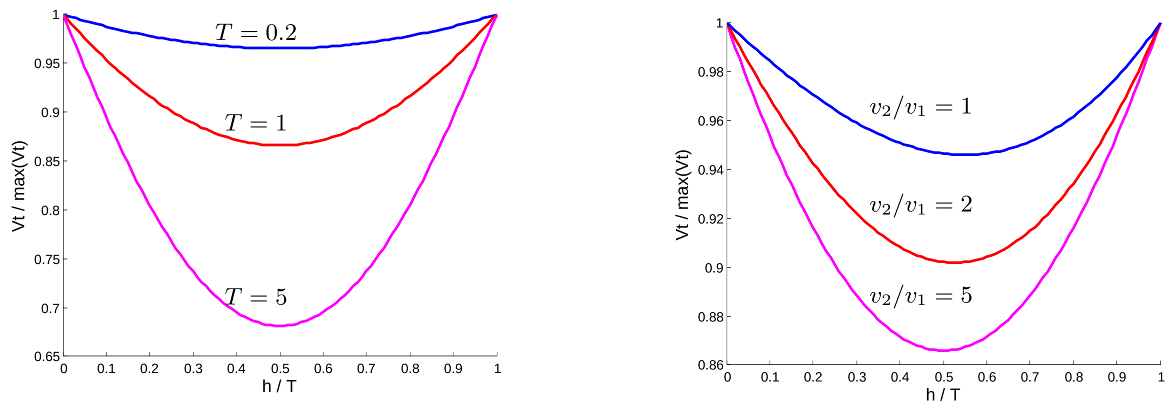

Fig. 5. Representation of the variation of $\langle V_t \rangle$ with respect to its maximal value with respect to the shift of the second modality. *Left:* $v_1 = v_2 = 1$, representation for various values of T. *Right:* $T_1 = T_2 = 1$, representation for various values of v_2/v_1.

channels [13, 6, 11]. However, we believe that this formulation also has the potential to be a useful tool to study the asymptotic behavior of a centralized deterministic multimodal Kalman filter. Indeed, as we have seen, the difficulty in this case is the potentially huge number of variance update functions to consider. Instead, one can consider replacing the deterministic, regular measurements considered in Section 2.2 by stochastic measurements where the probability that a sensor performs a measurement is proportional to its sampling frequency.

Let p_i stand for the probability that a measurement is performed on sensor i, $i = 1 \dots n$. It is assumed that either no measurement occurs with probability noted p_{n+1} or a single measurement occurs on sensor i with probability p_i, $i = 1 \dots n$ (normalization for these n sensors framework reads $\sum_{i=1}^{n+1} p_i = 1$). The model of [8] can be extended to a multimodal framework by modifying Equation 10 into n equations:

$$X_t^i = C_i \theta_t^i + N_t^i, \quad i = 1, \dots, n, \tag{11}$$

where N_t^i are independent centered Gaussian noises of covariance matrices R_i. In this setup, the (random) update equation for V_t is:

$$V_{t+1} = f(V_t) - \sum_{i=1}^{n} \gamma_{t+1}^i f(V_t) C_i^T \left[C_i f(V_t) C_i^T + R_i \right]^{-1} C_i f(V_t), \tag{12}$$

where $f(X) = AXA^T + Q$ and $(\gamma_{t+1}^1, \dots, \gamma_{t+1}^n)$ equals e_i (the i^{th} vector of the canonical basis) if a measurement is performed by the i^{th} sensor (which occurs with probability p_i) or equals $(0, \dots, 0)$ if no measurement is performed at time $t+1$ (which occurs with probability $p = \sum_{i=1}^{n} p_i$). The bounds of [8] can be generalized:

Property 2. *Let V_0 be a symmetric positive (SP) matrix and (V_t) the random process defined from V_0 by equation* (12). *Then for all t, we get*

$$\mathbb{E}[V_t] \preceq U_t \text{ [1]} \tag{13}$$

with $U_0 = V_0$, $U_{t+1} = g(AU_tA^T + Q)$ and

$$g(X) = X - \sum_{i=1}^{n} p_i X C_i^T (C_i X C_i^T + R_i)^{-1} C_i X. \tag{14}$$

This result is valid even if the time discretization step is not 1 but any $\epsilon > 0$. Getting back to the formulation of Section 2, we can consider for ϵ small enough that $p_i = \epsilon / T_i$ so that the average waiting time between two consecutive measurements by sensor i is T_i. We further have $Q = \epsilon$, $A = C_i = 1$ and $R_i = v_i$. Property 2 allows us to get an upper bound on $\mathbb{E}[V_t]$ which only depends on a single recursive equation.

Although this upper bound gives a simpler way to study the asymptotic variance of a stochastic multimodal Kalman filter, the model is not exactly analog to the deterministic model. An interesting outlook is to find upper bounds on $\mathbb{E}[V_t]$ when the update function h in the update equation $V_{t+1} = h(V_t)$ is randomly chosen among several analog functions, such as the $f_{i,t}$ functions of Equation 6. This would allow us to study a model closer to the initial deterministic regular sampling of Section 2.

5. CONCLUSION

In this paper, we proposed to raise the question of the "optimal" estimation of a latent variable in the presence of observations having different sampling rates and/or noise levels. Despite the fact that this type of multimodal estimation is involved in many applications, there is no pristine answer on how to design the best sampling pattern even in the simplest cases such as the estimation of a Brownian motion by two sensors having the same sampling period. The experimental results we observed seem to underline the fact that spreading the measurements over time instead of synchronizing them can lead to substantial differences in the average MSE of the MAP estimator and its limit inferior. Depending on whether one aims at favoring an estimator which is more reliable over time or at specific timestamps, the sampling layout will not be the same. However, deriving explicit rules for a given problem seems too tedious because the involved expressions are untractable.

Relying on a stochastic measurement model seems to be an elegant and efficient approach to the conceptual study of multimodal Kalman estimation. Indeed, replacing the various deterministic updates by random updates can lead, with a proper upper bound on the MSE, to a more explicit asymptotic control of the variance of the estimator.

[1] $A \preceq B$ means that $B - A$ is a positive matrix.

References

[1] Dana Lahat, Tülay Adalı, and Christian Jutten, "Multimodal Data Fusion: An Overview of Methods, Challenges and Prospects," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1449–1477, Aug. 2015.

[2] Jing Sui, Tlay Adali, Qingbao Yu, Jiayu Chen, and Vince D. Calhoun, "A review of multivariate methods for multimodal fusion of brain imaging data," *Journal of Neuroscience Methods*, vol. 204, no. 1, pp. 68 – 81, 2012.

[3] Mauro Dalla Mura, S Prasad, Fabio Pacifici, Paolo Gamba, and Jocelyn Chanussot, "Challenges and opportunities of multimodality and Data Fusion in Remote Sensing," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, Lisbonne, Portugal, Sept. 2014, pp. 106–110.

[4] Peter S. Maybeck, *Stochastics Models, Estimation, and Control*, Academic Press, 1979.

[5] Shu-Li Sun and Zi-Li Deng, "Multi-sensor optimal information fusion kalman filter," *Automatica*, vol. 40, no. 6, pp. 1017 – 1023, 2004.

[6] Li Yu Wen-An Zhang, Gang Feng, "Multi-rate distributed fusion estimation for sensor networks with packet losses," *Automatica*, vol. 48, no. 9, pp. 2016 – 2028, 2012.

[7] C.L. Robinson and P.R. Kumar, "Sending the most recent observation is not optimal in networked control: Linear temporal coding and towards the design of a control specific transport protocol," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2005.

[8] Bruno Sinopoli, Luca Schenato, Massimo Franceschetti, Kameshwar Poolla, Michael I. Jordan, and Shankar S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1453–1464, 2004.

[9] Soummya Kar, Bruno Sinopoli, and José M. F. Moura, "Kalman filtering with intermittent observations: Weak convergence to a stationary distribution," *IEEE Trans. Automat. Contr.*, vol. 57, no. 2, pp. 405–420, 2012.

[10] Yilin Mo and Bruno Sinopoli, "Kalman filtering with intermittent observations: Tail distribution and critical value," *IEEE Trans. Automat. Contr.*, vol. 57, no. 3, pp. 677–689, 2012.

[11] Sean Weerakkody, Yilin Mo, Bruno Sinopoli, Duo Han, and Ling Shi, "Multi-sensor scheduling for state estimation with event-based, stochastic triggers," *CoRR*, vol. abs/1502.03068, 2015.

[12] Andrew H. Jazwinski, *Stochastic processes and filtering theory*, Mathematics in science and engineering. Academic Press, New York, 1970, UKM.

[13] Orhan C. Imer, Serdar Yksel, and Tamer Baar, "Optimal control of lti systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429 – 1439, 2006.