Event-Based Sensor Data Scheduling: Trade-Off Between Communication Rate and Estimation Quality

Junfeng Wu, Qing-Shan Jia, Karl Henrik Johansson, and Ling Shi

***Abstract*—We consider sensor data scheduling for remote state estimation. Due to constrained communication energy and bandwidth, a sensor needs to decide whether it should send the measurement to a remote estimator for further processing. We propose an event-based sensor data scheduler for linear systems and derive the corresponding minimum squared error estimator. By selecting an appropriate event-triggering threshold, we illustrate how to achieve a desired balance between the sensor-to-estimator communication rate and the estimation quality. Simulation examples are provided to demonstrate the theory.**

***Index Terms*—Estimation performance, event-based scheduling, Kalman filter, sensor scheduling.**

I. Introduction

Networked control systems have received much attention in the last decade and are found in a wide spectrum of applications, e.g., in civil structure maintenance, environmental monitoring, battlefield surveillance. In many of these applications, sensor nodes are battery-powered. Replacing old batteries that are running out of energy are costly operations and may not even be possible. At the same time, the communication network may be shared by many nodes, and consequently the communication bandwidth might be scarce and uncertain. Thus it is practically important to minimize the sensor-to-estimator communication rate. A too low rate may, however, lead to poor estimation quality. It is of significant interest to reduce the sensor-to-estimator communication rate while guarantee a certain level of desired estimation quality.

Related research on remote estimation under communication constraint and sensor scheduling in various forms have appeared in recent years. The problem of sensor scheduling can be traced back to the 1970s. Athans [1] first formulated a class of optimization problems dealing with selecting one measurement provided by one out of many sensors. Gupta *et al.* [2] proposed a stochastic sensor scheduling scheme among multiple sensors for one process and provided the optimal probability distribution over the sensors to be selected. In control of modern networked systems, actions are often desired to be taken only after certain events occur. These events may contain useful information about the system [3], and using an appropriate event-based scheduler, the performance of the estimator can be improved. Imer and Basar [4] considered optimal estimation with limited measurements where the stochastic process was a scalar linear system. They showed that the optimal observer policy has a solution in an event-triggered form. Cogill *et al.* [5] considered a sensor data scheduling problem and used a feedback policy to choose the transmission times which provides a trade-off between the communication rate and the estimation error. Ambrosino *et al.* [6] considered the channel capacity constraint. In recent work by Li *et al.* [7], an event-triggered approach was used to trigger the data transmission from a sensor to a remote observer in order to minimize the mean squared estimation error at the observer subject to a constraint on transmission frequency. Closely related works are also given by Riberio *et al.* [8] and Msechu *et al.* [9] where quantized Kalman filter were considered. The main distinctions between our work and [8], [9] include the different communication models (packed-based versus finite-bit channels) and different estimation procedures. While we design an event-based scheduler to optimize the tradeoff between the sensor-to-estimator communication rate and the remote estimation quality, the work of [8], [9] focused on designing encoder-decoder pairs to improve the estimation quality over a bit-limited channel.

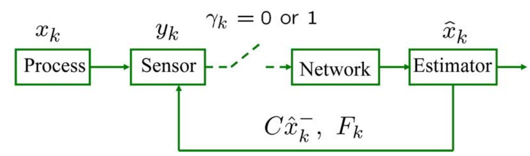

Fig. 1. Event-based scheduling for remote state estimation.

This paper focuses on the design of sensor data scheduler and the corresponding networked state estimator illustrated by the architecture in Fig. 1. We propose an event-based sensor data scheduler and derive the corresponding minimum mean-squared error (MMSE) estimator. By adopting an approximation technique from nonlinear filtering, we derive a simple form of an accurate MMSE estimator, from which an illustrative relationship between the sensor-to-estimator communication rate and the remote estimation quality can be obtained.

The remainder of this paper is organized as follows. In Section II, we provide the mathematical problem formulation. In Section III, we derive the exact MMSE estimator and an approximate MMSE estimator for an event-based sensor data scheduler. Via simulation examples in Section IV, we demonstrate how a desired trade-off between the sensor communication rate and the estimation quality can be achieved. It is also shown that the approximate MMSE estimator produces accurate results. In Section V, some concluding remarks are given.

Notation: $\mathbb{S}_+^n$ is the set of $n \times n$ positive semi-definite matrices. When $X \in \mathbb{S}_+^n$, we simply write $X \geq 0$; Similarly, $X \geq Y$ means $X - Y \geq 0$. $f_{\mathrm{x}}(x)$ represents the probability density function (pdf) of the random variable (r.v.) x, and $f_{\mathrm{x}|\mathrm{y}}(x|y)$ denotes the pdf of a r.v. x conditional on the variable y. $\mathcal{N}(\mu, \Sigma)$ denotes Gaussian distribution with mean μ and covariance matrix Σ. $\mathbb{E}[\cdot]$ denotes the mathematical expectation and $\Pr(\cdot)$ denotes the probability of a random event. $\mathrm{Tr}\{\cdot\}$ denotes the trace of a matrix and $\|\cdot\|_\infty$ denotes the Hölder infinity-norm of a vector.

Manuscript received August 26, 2011; revised January 30, 2012, August 06, 2012, and August 09, 2012; accepted August 20, 2012. Date of publication August 24, 2012; date of current version March 20, 2013. The work of J. Wu and L. Shi was supported in part by a HK RGC CRF grant HKUST11/CRF/10. The work of Q.-S. Jia iwas supported in part by the National Science Foundation of China under Grants 60704008, 60736027, 61174072, and 90924001, by the Specialized Research Fund for the Doctoral Program of Higher Education (No. 20070003110), by the National 111 International Collaboration Project (No. B06002), and by the TNLIST Cross Disciplinary Research Funding. The work of K. H. Johansson was supported by the Knut and Alice Wallenberg Foundation and the Swedish Research Council. Recommended by Associate Editor L. Schenato.

J. Wu and L. Shi are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: jfwu@ust.hk; eesling@ust.hk).

Q.-S. Jia is with the Center for Intelligent and Networked Systems, Department of Automation, TNLIST, Tsinghua University, Beijing, China (e-mail: jiaqs@tsinghua.edu.cn).

K. H. Johansson is with the ACCESS Linnaeus Center, School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden (e-mail: kallej@ee.kth.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2012.2215253

II. Problem Setup

A. System Model

Consider the following linear system:

$$\mathrm{x}_{k+1} = A\mathrm{x}_k + w_k \quad (1)$$
$$\mathrm{y}_k = C\mathrm{x}_k + v_k \quad (2)$$

where $\mathrm{x}_k \in \mathbb{R}^n$ is the state vector, $\mathrm{y}_k \in \mathbb{R}^m$ is the sensor measurement, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are mutually uncorrelated white Gaussian noises with covariances $Q \geq 0$ and $R > 0$, respectively. The initial state x_0 is zero-mean Gaussian with covariance matrix $\mathbb{E}[\mathrm{x}_0\mathrm{x}_0'] = \Pi_0 \geq 0$, and is uncorrelated with w_k and v_k for all $k \geq 0$. (A, C) and $(A, \sqrt{Q})$ are observable and controllable, respectively. After y_k (the measured value of y_k) is taken, the sensor decides whether it will send y_k to a remote estimator for further processing. Let $\gamma_k = 1$ or 0 be the decision variable whether y_k shall be sent or not. Define $\mathrm{I}_k \triangleq \{\gamma_0 y_0, \ldots, \gamma_k y_k\}$ with $\mathrm{I}_{-1} \triangleq \emptyset$,

$$\hat{x}_k^- \triangleq \mathbb{E}[\mathrm{x}_k|\mathrm{I}_{k-1}], \mathrm{e}_k^- \triangleq \mathrm{x}_k - \hat{x}_k^-, P_k^- \triangleq \mathbb{E}\left[\mathrm{e}_k^- \mathrm{e}_k^{-\prime}|\mathrm{I}_{k-1}\right] \quad (3)$$

and

$$\hat{x}_k \triangleq \mathbb{E}[\mathrm{x}_k|\mathrm{I}_k], \quad \mathrm{e}_k \triangleq \mathrm{x}_k - \hat{x}_k, \quad P_k \triangleq \mathbb{E}\left[\mathrm{e}_k\mathrm{e}_k'|\mathrm{I}_k\right]. \quad (4)$$

The estimates $\hat{x}_k^-$ and $\hat{x}_k$ are called the *a priori* and *a posteriori* MMSE estimate, respectively. Further define the measurement innovation z_k as

$$\mathrm{z}_k \triangleq \mathrm{y}_k - \mathbb{E}[\mathrm{y}_k|\mathrm{I}_{k-1}]. \quad (5)$$

Define the functions h, $\tilde{g}_\lambda$, g_λ and $g_\lambda^k \colon \mathbb{S}_+^n \to \mathbb{S}_+^n$ as follows:

$$\tilde{g}_\lambda(X) \triangleq X - \lambda XC'[CXC' + R]^{-1}CX$$
$$h(X) \triangleq AXA' + Q \quad g_\lambda(X) \triangleq \tilde{g}_\lambda \circ h(X)$$

where $\circ$ denotes the function composition. In the sequel, if $\lambda = 1$, g_1 and $\tilde{g}_1$ will be written as g and $\tilde{g}$ for brevity. We can write the update equation for P_k in a compact form as

$$P_k = \begin{cases} g(P_{k-1}), & \text{if } \gamma_k = 1 \\ h(P_{k-1}), & \text{if } \gamma_k = 0. \end{cases}$$

Notice that $h(P_{k-1}) \geq g(P_{k-1})$ as $\tilde{g}(X) \leq X$ for any $X \geq 0$. This has an intuitive explanation: the measurement y_k (or alternatively, the innovation z_k) always reduces the estimation error covariance.

B. Event-Based Sensor Scheduler

We consider in this paper applications where feedback is available from the estimator to the sensor, see Fig. 1.[1]

Consider the following two cases for the Kalman filter when $z_k = 0$:

1) $\gamma_k = 1 : \hat{x}_k = A\hat{x}_{k-1}$ and $P_k = g(P_{k-1})$;
2) $\gamma_k = 0 : \hat{x}_k = A\hat{x}_{k-1}$ and $P_k = h(P_{k-1})$.

[1]Examples of such applications can be found in remote estimation based on the IEEE 802.15.4/ZigBee protocol: sensor devices can be scheduled to communicate to the so-called Personal Area Network coordinator which also serves as a remote estimator. The coordinator broadcasts information to all devices at the beginning of each periodic superframe and can then incorporate the required feedback information.

The estimate $\hat{x}_k$ for the two cases are the same, but the error covariances are different. Therefore if the sensor finds that z_k is zero and does not send y_k to the estimator, and at the same time, the estimator is aware of this information, then even without receiving y_k, the estimator knows that $\hat{x}_k = A\hat{x}_{k-1}$ has error covariance $g(P_{k-1})$, which is smaller than $h(P_{k-1})$.

Since $CP_k^-C' + R > 0$, there exists a unitary matrix $U_k \in \mathbb{R}^{m\times m}$ such that

$$U_k'\left(CP_k^-C' + R\right)U_k = \Lambda_k$$

where $\Lambda_k = \mathrm{diag}(\lambda_k^1, \ldots, \lambda_k^m) \in \mathbb{R}^{m\times m}$ and $\lambda_k^1, \ldots, \lambda_k^m \in \mathbb{R}$ are the eigenvalues of $CP_k^-C' + R$. Define $F_k \in \mathbb{R}^{m\times m}$ as

$$F_k \triangleq U_k\Lambda_k^{-\frac{1}{2}}. \quad (6)$$

Evidently, $F_k'F_k = (CP_k^-C' + R)^{-1}$. The matrix F_k is computed by the remote estimator and is sent back to the sensor along with $C\hat{x}_k^-$ at each time, see Fig. 1. Define ϵ_k as

$$\epsilon_k \triangleq F_k'\mathrm{z}_k. \quad (7)$$

This transformation is called the Mahalanobis transformation. The coordinates of z_k are decorrelated, so ϵ_k has m-variable standard Gaussian distribution, which contains a set of independent principal components of z_k.

We consider the following event-based sensor data scheduler:

$$\gamma_k = \begin{cases} 0, & \text{if } \|\epsilon_k\|_\infty \leq \delta \\ 1, & \text{otherwise} \end{cases} \quad (8)$$

where $\delta \geq 0$ is a fixed threshold. Under this scheduler, if $\gamma_k = 0$, the estimator can infer that $\|\epsilon_k\|_\infty \leq \delta$. It is this additional information that helps reduce the estimation error at the remote estimator. With a slight abuse of notation, we redefine the information I_k received by the remote estimator till k as

$$\mathrm{I}_k \triangleq \{\gamma_0 y_0, \ldots, \gamma_k y_k\} \cup \{\gamma_0, \ldots, \gamma_k\}.$$

Define the average sensor communication rate as

$$\gamma \triangleq \limsup_{T\to+\infty} \frac{1}{T+1}\sum_{k=0}^{T}\mathbb{E}[\gamma_k]. \quad (9)$$

Notice that both the average rate γ and the estimation error covariance matrix P_k depend on the threshold δ. For example, if $\delta = 0$, then $\Pr(\|\epsilon_k\|_\infty \leq 0) = 0$ and the sensor sends y_k at each k (almost surely). Consequently, $\gamma = 1$ and $P_k = g(P_{k-1})$. On the other hand, if $\delta = +\infty$, then the sensor keeps y_k for all k, thus making $\gamma = 0$. As the event $\|\epsilon_k\|_\infty \leq +\infty$ provides no extra information on the innovation, the estimator is in this case equivalent to an open-loop predictor. Therefore, $P_k = h(P_{k-1})$. In the latter case, when A is unstable, P_k diverges as $k \to +\infty$. Apparently there is a tradeoff between the communication rate and the estimation quality.

We now state the main problems considered in this paper.

1) Under the event-based sensor scheduler (8), what is the MMSE estimator?
2) How to choose the threshold δ in (8) to achieve a desirable trade-off between the communication rate and the estimation quality?

We will provide answers to these two problems in the remainder of the paper.

III. Event-Based State Estimation

In this section, we derive the MMSE estimator under the event-based sensor data scheduler (8), first the exact estimator and then an accurate approximation.

A. The Exact MMSE Estimator

The MMSE estimate is uniquely specified as the conditional mean given all available information [10]. In this subsection, we provide an exact MMSE estimator corresponding to the event-based scheduler (8) using the following two-step updating procedure.

1) Time Update: The *a priori estimate* $\hat{x}_k^-$, which is the conditional mean of x_k given the information set I_{k-1}, is derived as

$$\hat{x}_k^- = \mathbb{E}[\mathrm{x}_k|\mathrm{I}_{k-1}] = \int_{\mathbb{R}^m} x f_{\mathrm{x}_k}(x|\mathrm{I}_{k-1})\mathrm{d}x \tag{10}$$

and the corresponding estimation error covariance P_k^- is given by

$$P_k^- = \int_{\mathbb{R}^m} \left(x - \hat{x}_k^-\right)\left(x - \hat{x}_k^-\right)' f_{\mathrm{x}_k}(x|\mathrm{I}_{k-1})\mathrm{d}x.$$

2) Measurement Update: The *a posteriori estimate* $\hat{x}_k$, which is the conditional mean of x_k given I_k, is derived as follows. Depending on whether $\gamma_k = 0$ or 1, we have the following two cases:

1) $\gamma_k = 0$. The sensor does not send y_k to the remote estimator, but the estimator is aware of that $\|\epsilon_k\|_\infty \le \delta$. Consequently, $\hat{x}_k$ is given by

$$\hat{x}_k = \mathbb{E}[\mathrm{x}_k|\hat{\mathrm{I}}_k] = \int_{\mathbb{R}^m} x f_{\mathrm{x}_k}\left(x|\hat{\mathrm{I}}_k\right)\mathrm{d}x \tag{11}$$

where we denote $\hat{\mathrm{I}}_k = \mathrm{I}_{k-1} \cup \{\gamma_k = 0\}$. Define the set $\Omega \subset \mathbb{R}^m$ as

$$\Omega \triangleq \left\{\epsilon_k \in \mathbb{R}^m : \|\epsilon_k\|_\infty \le \delta\right\} \tag{12}$$

then one can compute $f_{\mathrm{x}_k}(x|\hat{\mathrm{I}}_k)$ using Bayes' rule as

$$f_{\mathrm{x}_k}(x|\hat{\mathrm{I}}_k) = \frac{f_{\mathrm{x}_k}(x|\mathrm{I}_{k-1})\int_\Omega f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1}, \mathrm{x}_k)\mathrm{d}\epsilon}{\int_\Omega f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon} \tag{13}$$

where $f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1}, \mathrm{x}_k) = \mathcal{N}(F_k'\mathrm{x}_k - F_k'C\hat{x}_k^-, F_k'RF_k)$ and

$$\mathrm{z}_k = C\mathrm{x}_k + v_k - C\hat{x}_k^- = C\mathrm{e}_k^- + v_k. \tag{14}$$

The *a posteriori* error covariance P_k is given by

$$P_k = \int_{\mathbb{R}^m} (x - \hat{x}_k)(x - \hat{x}_k)' f_{\mathrm{x}_k}(x|\hat{\mathrm{I}}_k)\mathrm{d}x.$$

2) $\gamma_k = 1$. The sensor sends y_k to the remote estimator. Denote the measured value of the innovation z_k as z. Then I_k becomes $\mathrm{I}_k = \mathrm{I}_{k-1} \cup \{\mathrm{z}_k = z\}$. The remote estimator updates $\hat{x}_k$ as in (11), but the conditional pdf $f_{\mathrm{x}_k}(x|\mathrm{I}_k)$ is now calculated using Bayes' rule as

$$f_{\mathrm{x}_k}(x|\mathrm{I}_k) = \frac{f_{\mathrm{x}_k}(x|\mathrm{I}_{k-1}) f_{z_k}(z|\mathrm{I}_{k-1}, \mathrm{x}_k)}{f_{z_k}(z|\mathrm{I}_{k-1})} \tag{15}$$

where, from (14), one easily sees that

$$f_{z_k}(z|\mathrm{I}_{k-1}, \mathrm{x}_k) = \mathcal{N}\left(C\mathrm{x}_k - C\hat{x}_k^-, R\right).$$

The *a posteriori* estimation error covariance P_k is given by

$$P_k = \int_{\mathbb{R}^m} (x - \hat{x}_k)(x - \hat{x}_k)' f_{\mathrm{x}_k}(x|\mathrm{I}_{k-1}, \mathrm{z}_k = z)\mathrm{d}x.$$

Remark 3.1: Although the above two steps produce the MMSE estimate $\hat{x}_k$ corresponding to the event-based scheduler (8), each updating step requires numerical integration. The amount of computation involved make this estimator intractable in general, which motivates us to consider an approximate MMSE estimator. As we will demonstrate, by using a standard technique in nonlinear filtering, we can derive an approximate MMSE estimator in a simple recursive form.

B. Approximate MMSE Estimator

A commonly used approximation technique in nonlinear filtering is to assume that the conditional distribution of x_k given I_{k-1} is Gaussian, i.e.,

$$f_{\mathrm{x}_k}(x|\mathrm{I}_{k-1}) = \mathcal{N}\left(\hat{x}_k^-, P_k^-\right). \tag{16}$$

This assumption reduces the estimation problem from the tracking of a general pdf, which is usually computationally intractable, to the tracking of its mean and covariance matrix. The approximation is widely used in the literature, e.g., [8], [9], [11]. Unless specifically mentioned, our analysis in the rest of this paper is based on this assumption. The approximation leads to a very simple form of the estimator, as shown by the following result.

Theorem 3.2: Consider the remote state estimation in Fig. 1 with the event-based sensor scheduler (8). Under the assumption (16), the MMSE estimator is given recursively as follows:

1) *Time update*:

$$\begin{cases} \hat{x}_k^- = A\hat{x}_{k-1} \\ P_k^- = h(P_{k-1}). \end{cases} \tag{17}$$

2) *Measurement update*:

$$\begin{cases} \hat{x}_k = \hat{x}_k^- + \gamma_k L_k z_k \\ P_k = \gamma_k \tilde{g}\left(P_k^-\right) + (1 - \gamma_k)\tilde{g}_{\beta(\delta)}\left(P_k^-\right) \end{cases}$$

where

$$\beta(\delta) = \frac{2}{\sqrt{2\pi}}\delta e^{-\frac{\delta^2}{2}}\left[1 - 2Q(\delta)\right]^{-1} \tag{18}$$

and $Q(\cdot)$ is the standard Q-function defined by

$$Q(\delta) \triangleq \int_\delta^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\mathrm{d}x. \tag{19}$$

Before we present the proof, we state a few preliminary results.

From (14) and (16), z_k is zero-mean Gaussian conditioned on I_{k-1}. Furthermore, z_k is jointly Gaussian with x_k conditioned on I_{k-1}. From (14)

$$\mathbb{E}\left[\mathrm{z}_k\mathrm{z}_k'|\mathrm{I}_{k-1}\right] = C\mathbb{E}\left[\mathrm{e}_k^-\mathrm{e}_k^{-\prime}|\mathrm{I}_{k-1}\right]C' + R = CP_k^-C' + R \tag{20}$$

and

$$\mathbb{E}\left[\mathrm{e}_k^-\mathrm{z}_k'|\mathrm{I}_{k-1}\right] = \mathbb{E}\left[\mathrm{e}_k^-\mathrm{e}_k^{-\prime}|\mathrm{I}_{k-1}\right]C' = P_k^-C'. \tag{21}$$

Now let us take a look at ϵ_k defined in (7). From (20)

$$\mathbb{E}\left[\epsilon_k\epsilon_k'|\mathrm{I}_{k-1}\right] = F_k'\mathbb{E}\left[\mathrm{z}_k\mathrm{z}_k'|\mathrm{I}_{k-1}\right]F_k = I_m.$$

Thus, given I_{k-1}, ϵ_k is a zero-mean Gaussian multivariate random variable with unit variance. Denote ϵ_k^i as the ith element of ϵ_k. Then ϵ_k^i and ϵ_k^j are mutually independent if $i \neq j$. Notice that $\gamma_k = 0$ implies that the event $\|\epsilon_k\|_\infty \leq \delta$ happens. We then have the following result.

Lemma 3.3: $F_k'\mathbb{E}[\mathrm{z}_k\mathrm{z}_k'|\hat{\mathrm{I}}_k]F_k = \mathbb{E}[\epsilon_k\epsilon_k'|\hat{\mathrm{I}}_k] = [1-\beta(\delta)]I_m$.

Proof: Straightforward calculation yields the first equality. Given I_{k-1}, due to the independence of ϵ_k^i and ϵ_k^j for $i \neq j$ along with Lemma A.1 in the Appendix, we have

$$\mathbb{E}\left[\left(\epsilon_k^i\right)^2|\hat{\mathrm{I}}_k\right] = \mathbb{E}\left[\left(\epsilon_k^i\right)^2|\mathrm{I}_{k-1}, \|\epsilon_k\|_\infty \leq \delta\right] = \mathbb{E}\left[\left(\epsilon_k^i\right)^2|\mathrm{I}_{k-1}, \left|\epsilon_k^i\right| \leq \delta\right] = 1-\beta(\delta)$$

and

$$\mathbb{E}\left[\epsilon_k^i\epsilon_k^j|\hat{\mathrm{I}}_k\right] = \mathbb{E}\left[\epsilon_k^i\epsilon_k^j|\mathrm{I}_{k-1}, \left|\epsilon_k^i\right| \leq \delta, \left|\epsilon_k^j\right| \leq \delta\right] = 0.$$

Thus

$$\mathbb{E}\left[\epsilon_k\epsilon_k'|\hat{\mathrm{I}}_k\right] = [1-\beta(\delta)]\,I_m.$$

∎

The following lemma is used in deriving the main result.

Lemma 3.4: The following equalities hold:

$$\mathbb{E}\left[\mathrm{e}_k^-\mathrm{z}_k'|\hat{\mathrm{I}}_k\right] = L_k\mathbb{E}\left[\mathrm{z}_k\mathrm{z}_k'|\hat{\mathrm{I}}_k\right] \tag{22}$$

$$\mathbb{E}\left[\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\mathrm{z}_k'|\hat{\mathrm{I}}_k\right] = 0 \tag{23}$$

$$\mathbb{E}\left[\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)'|\mathrm{I}_{k-1}, \mathrm{z}_k = z\right] = \tilde{g}\left(P_k^-\right) \tag{24}$$

$$\mathbb{E}\left[\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)'|\hat{\mathrm{I}}_k\right] = \tilde{g}\left(P_k^-\right) \tag{25}$$

where $L_k = P_k^-C'[CP_k^-C'+R]^{-1}$.

Proof: We first prove (22). From Lemma A.2, (20), and (21)

$$\mathbb{E}\left[\mathrm{x}_k|\mathrm{I}_{k-1}, \mathrm{z}_k = z\right] = \hat{x}_k^- + P_k^-C'\left[CP_k^-C'+R\right]^{-1}z. \tag{26}$$

Since given I_{k-1}, ϵ_k is Gaussian with zero mean and unit covariance, we can define $p_\delta \triangleq \Pr(\|\epsilon_k\|_\infty \leq \delta|\mathrm{I}_{k-1})$. Using the conditional pdf

$$f_{\epsilon_k}(\epsilon|\hat{\mathrm{I}}_k) = \begin{cases} \frac{f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})}{p_\delta}, & \text{if } \|\epsilon_k\| \leq \delta \\ 0, & \text{otherwise} \end{cases} \tag{27}$$

we obtain

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}_k^-\mathrm{z}_k'|\hat{\mathrm{I}}_k\right] \\
&= \frac{1}{p_\delta}\int_\Omega \mathbb{E}\left[\mathrm{e}_k^-|\mathrm{I}_{k-1}, \mathrm{z}_k = F_k'^{-1}\epsilon\right]\epsilon'F_k^{-1}f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon \\
&= \frac{1}{p_\delta}\int_\Omega \mathbb{E}\left[\mathrm{x}_k - \hat{x}_k^-|\mathrm{I}_{k-1}, \mathrm{z}_k = F_k'^{-1}\epsilon\right]\epsilon'F_k^{-1}f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon \\
&= \frac{1}{p_\delta}\int_\Omega \left(\mathbb{E}\left[\mathrm{x}_k|\mathrm{I}_{k-1}, \mathrm{z}_k = F_k'^{-1}\epsilon\right] - \hat{x}_k^-\right)\epsilon'F_k^{-1}f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon \\
&= P_k^-C'\left(CP_k^-C'+R\right)^{-1}F_k'^{-1}\frac{1}{p_\delta}\int_\Omega \epsilon\epsilon'f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon F_k^{-1} \\
&= L_kF_k'^{-1}\mathbb{E}\left[\epsilon_k\epsilon_k'|\hat{\mathrm{I}}_k\right]F_k^{-1} \\
&= L_k\mathbb{E}\left[\mathrm{z}_k\mathrm{z}_k'|\hat{\mathrm{I}}_k\right]
\end{aligned}$$

where the last equality is from Lemma 3.3. From (22), we have

$$\mathbb{E}\left[\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\mathrm{z}_k'|\hat{\mathrm{I}}_k\right] = \mathbb{E}\left[\mathrm{e}_k^-\mathrm{z}_k'|\hat{\mathrm{I}}_k\right] - L_k\mathbb{E}\left[\mathrm{z}_k\mathrm{z}_k'|\hat{\mathrm{I}}_k\right] = 0$$

which shows (23). To prove (24), using Lemma A.2, we have

$$\begin{aligned}
&\mathbb{E}\Big[\left(\mathrm{x}_k - \mathbb{E}\left[\mathrm{x}_k|\mathrm{I}_{k-1}, \mathrm{z}_k = z\right]\right)\left(\mathrm{x}_k - \mathbb{E}\left[\mathrm{x}_k|\mathrm{I}_{k-1}, \mathrm{z}_k = z\right]\right)' \\
&\quad \times |\mathrm{I}_{k-1}, \mathrm{z}_k = z\Big] \\
&= P_k^- - P_k^-C'\left(CP_k^-C'+R\right)^{-1}CP_k^- \\
&= \tilde{g}\left(P_k^-\right).
\end{aligned} \tag{28}$$

Notice that (26) leads to

$$\mathrm{x}_k - \mathbb{E}[\mathrm{x}_k|\mathrm{I}_{k-1}, \mathrm{z}_k = z] = \mathrm{x}_k - \hat{x}_k^- - L_kz = \mathrm{e}_k^- - L_kz$$

which together with (28) shows (24). Now from (24), one obtains

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)'|\hat{\mathrm{I}}_k\right] \\
&= \int_\Omega \mathbb{E}\left[\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)'|\mathrm{I}_{k-1}, \mathrm{z}_k = F_k'^{-1}\epsilon\right] \\
&\quad \times \frac{f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})}{p_\delta}\mathrm{d}\epsilon \\
&= \frac{1}{p_\delta}\tilde{g}\left(P_k^-\right)\int_\Omega f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon \\
&= \tilde{g}\left(P_k^-\right)
\end{aligned}$$

where to get the second last equality, we note that from (27) we have $\int_\Omega f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon = p_\delta$ ∎

Proof to Theorem 3.2: The proof of the time update is simple, shown as follows:

$$\begin{aligned}
\hat{x}_k^- &= A\mathbb{E}\left[\mathrm{x}_{k-1}|\mathrm{I}_{k-1}\right] = A\hat{x}_{k-1} \\
P_k^- &= \mathbb{E}\left[(A\mathrm{e}_{k-1}+w_{k-1})(A\mathrm{e}_{k-1}+w_{k-1})'|\mathrm{I}_{k-1}\right] \\
&= AP_{k-1}A' + Q = h(P_{k-1}).
\end{aligned}$$

Next, we verify the measurement update for the following two cases.

1) $\gamma_k = 1$: According to (26) and (28)

$$\begin{aligned}
\hat{x}_k &= \hat{x}_k^- + L_kz_k \\
P_k &= \tilde{g}\left(P_k^-\right).
\end{aligned}$$

2) $\gamma_k = 0$: the sensor does not send y_k to the remote estimator which computes $\hat{x}_k$ as

$$\begin{aligned}
\hat{x}_k &= \mathbb{E}\left[\mathrm{x}_k|\hat{\mathrm{I}}_k\right] \\
&= \frac{1}{p_\delta}\int_\Omega \mathbb{E}\left[\mathrm{x}_k|\mathrm{I}_{k-1}, \mathrm{z}_k = F_k'^{-1}\epsilon\right]f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon \\
&= \frac{1}{p_\delta}\int_\Omega \left(\hat{x}_k^- + L_kF_k'^{-1}\epsilon\right)f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon \\
&= \hat{x}_k^- + \frac{L_kF_k'^{-1}}{p_\delta}\int_\Omega \epsilon f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon \\
&= \hat{x}_k^-
\end{aligned}$$

where the last equality is due to

$$\int_\Omega \epsilon f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})\mathrm{d}\epsilon = 0$$

since being a pdf of Gaussian distribution, $f_{\epsilon_k}(\epsilon|\mathrm{I}_{k-1})$ is even and Ω defined in (12) is symmetric and centered in the origin.

Now from (23), (25), Lemmas 3.3, and A.1, the corresponding error covariance matrix P_k can be computed as

$$\begin{aligned} P_k =& \mathbb{E}\left[(\mathrm{x}_k - \hat{x}_k)(\mathrm{x}_k - \hat{x}_k)'|\hat{\mathrm{I}}_k\right] \\ =& \mathbb{E}\left[\left(\mathrm{x}_k - \hat{x}_k^-\right)\left(\mathrm{x}_k - \hat{x}_k^-\right)'|\hat{\mathrm{I}}_k\right] \\ =& \mathbb{E}\left[\left\{\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)+L_k\mathrm{z}_k\right\}\left\{\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)+L_k\mathrm{z}_k\right\}'|\hat{\mathrm{I}}_k\right] \\ =& \mathbb{E}\left[\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)'+\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)\mathrm{z}_k'L_k'\right. \\ & \left. +L_k\mathrm{z}_k\left(\mathrm{e}_k^- - L_k\mathrm{z}_k\right)'+L_k\mathrm{z}_k\mathrm{z}_k'L_k'|\hat{\mathrm{I}}_k\right] \\ =& \tilde{g}\left(P_k^-\right)+L_k\mathbb{E}\left[\mathrm{z}_k\mathrm{z}_k'|\hat{\mathrm{I}}_k\right]L_k' \\ =& \tilde{g}\left(P_k^-\right)+[1-\beta(\delta)]\,L_k\left(F_kF_k'\right)^{-1}L_k' \\ =& \tilde{g}\left(P_k^-\right)+[1-\beta(\delta)]\,L_k\left(CP_k^-C'+R\right)L_k' \\ =& \tilde{g}_{\beta(\delta)}\left(P_k^-\right). \end{aligned}$$

■

From Theorem 3.2, we can write the update for P_k in a compact form as

$$P_k = \begin{cases} g(P_{k-1}), & \text{if } \gamma_k = 1 \\ g_{\beta(\delta)}(P_{k-1}), & \text{if } \gamma_k = 0. \end{cases} \tag{29}$$

Remark 3.5: P_k is a function of $\{\gamma_t\}_{t=0}^k$ and $\beta(\delta)$, both of which depend on δ. By properly tuning δ, we can achieve a desired trade-off between the sensor communication rate γ and the estimation quality in terms of P_k. For example, if we wish to have a small γ, then picking a large δ would serve the purpose. The optimal choice of δ depends on the available communication resources.

Lemma 3.6: Let $\delta \geq 0$. Then

$$\Pr\left(\|\epsilon_k\|_\infty \leq \delta|I_{k-1}\right) = [1-2Q(\delta)]^m.$$

Proof: Note that $\|\epsilon_k\|_\infty = \max\{|\epsilon_k^1|,\ldots,|\epsilon_k^m|\} \leq \delta$ iff $|\epsilon_k^i| \leq \delta, \forall\, 1 \leq i \leq m$. Therefore

$$\begin{aligned} \Pr\left(\|\epsilon_k\|_\infty \leq \delta|I_{k-1}\right) &= \prod_{i=1}^m \Pr\left(\left|\epsilon_k^i\right| \leq \delta|I_{k-1}\right) \\ &= [1-2Q(\delta)]^m. \end{aligned}$$

■

The following result is on the average sensor-to-estimator communication rate γ.

Proposition 3.7: Consider the remote state estimation in Fig. 1 with the event-based sensor scheduler (8). Under the assumption (16), the average sensor-to-estimator communication rate γ in (9) is given by

$$\gamma = 1 - [1-2Q(\delta)]^m. \tag{30}$$

Proof: Note that γ_k is a random variable taking value in {0,1} with $\Pr(\gamma_k = 1|\mathrm{I}_{k-1}) = \Pr(\|\epsilon_k\|_\infty > \delta|\mathrm{I}_{k-1})$. From Lemma 3.6, $\Pr(\|\epsilon_k\|_\infty \leq \delta|I_{k-1}) = [1-2Q(\delta)]^m$ whatever value I_{k-1} takes. Therefore, δ solely determines the distribution of γ_k and can be described by

$$\Pr(\gamma_k=0)=[1-2Q(\delta)]^m \text{ and } \Pr(\gamma_k=1)=1-[1-2Q(\delta)]^m.$$

Equation (30) is then proved from the definition of expectation of γ_k and the definition of γ. ■

IV. Simulation Example

In this example, we consider the scheduling of two sensors measuring Process 1 and Process 2 (see Fig. 2). Let Process 1 be the stable process with $A^1 = \begin{bmatrix} 0.9 & 0.1 \\ 0 & 0.9 \end{bmatrix}$, $C^1 = [1\ 0]$, $Q^1 = 5I_2$, $R^1 = 2$ and

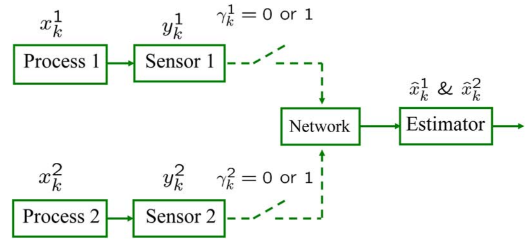

Fig. 2. Scheduling of sensors communication for Process 1 and Process 2.

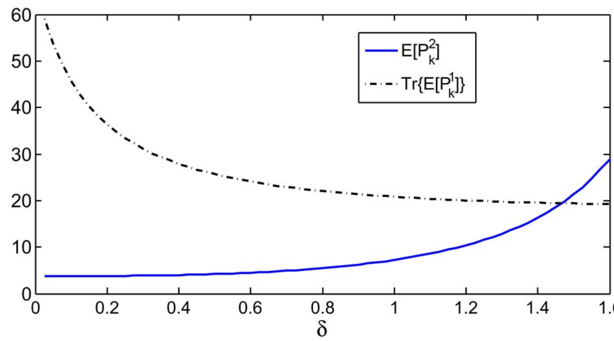

Fig. 3. Estimation quality of Process 1 and Process 2 versus δ.

Process 2 be the unstable process with parameters $A^2 = 1.2$, $C^2 = 1$, $Q^2 = 10$, $R^2 = 5$. Assume at each time only one of the sensors is able to communicate its measurement to the remote estimator due to a shortage of communication bandwidth. Since Process 1 is stable, a trivial sensor scheduler ϕ that guarantees a bounded estimation error covariance for both processes is that sensor 2 occupies the channel all the time, while sensor 1 is idle and the remote estimator predicts the state of Process 1 at each time. Let the estimation error covariances of Process 1 and Process 2 under the scheduler ϕ be $P_k^1(\phi)$ and $P_k^2(\phi)$, respectively, which are given by the following two recursions:

$$P_k^1 = h\left(P_{k-1}^1\right) \quad \text{and} \quad P_k^2 = g\left(P_k^2\right).$$

The steady-state values of $\mathrm{Tr}\{P_k^1\}$ and P_k^2 under ϕ are given by $\mathrm{Tr}\{\overline{P}_1\} = 59.00$ and $\overline{P}_2 = 3.77$. By using the event-based scheduler proposed in this paper, we can reduce the estimation error for Process 1 significantly while letting the estimation error for Process 2 grow only slightly. The idea is simple: let sensor 2 follow the event-based scheduler (8); whenever sensor 2 does not send data due to $\|\epsilon_k^2\|_\infty \leq \delta$, let sensor 1 communicate with the remote estimator. The resulting errors are plotted in Fig. 3 as a function of the parameter δ being used by Process 2, which clearly demonstrates the advantage adopting the event-based scheduler. For example, when $\delta = 0.4$, the values of $\mathrm{Tr}\{\mathbb{E}[P_k^1]\}$[2] and $\mathbb{E}[P_k^2]$ are 27.89 and 3.99, respectively, corresponding to a 52.7% decrease of estimation error for Process 1 and a 5.83% increase of estimation error for Process 2. In Fig. 4, we plot the empirical average sensor communication rate and the theoretical average sensor communication rate (30) for Process 2 under different values of δ. The two curves match almost indistinguishable and demonstrate that the approximated MMSE estimator is very close to the exact MMSE estimator.

V. Conclusion

We propose an event-based sensor data scheduler for state estimation over a network. The MMSE estimator is derived together with an

[2]Since P_k is a stochastic process due to the randomness of γ_k, we will consider $\mathrm{Tr}\{\mathbb{E}[P_k]\}$ as a performance measure for the remote estimator under the event-based sensor scheduler (8), which is obtained in this example via Monte Carlo simulations.

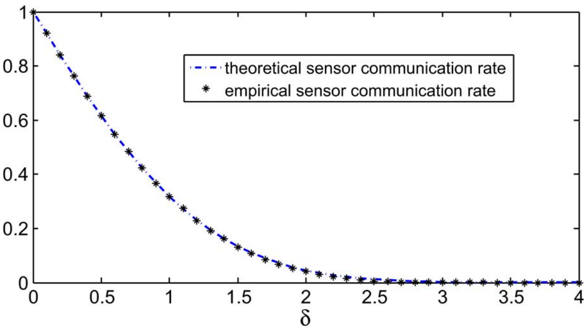

Fig. 4. Sensor communication rate γ versus scheduling parameter δ.

approximate estimator. It is shown that by tolerating a small amount of increase of the estimation error, a significant reduction of the sensor-to-estimator communication rate can be achieved. In many applications of networked control systems, multiple sensors may be involved. Constructing appropriate event-based schedules at each sensor and estimating the process state based on the received data and the additional information inferred by the events are more difficult than the one we have considered. This will be pursued in our future work.

APPENDIX

Lemma A.1: Let $\mathrm{x} \in \mathbb{R}$ be a Gaussian r.v. with zero mean and variance $\mathbb{E}[\mathrm{x}^2] = \sigma^2$. Denoting $\Delta = \delta\sigma$, then $\mathbb{E}[\mathrm{x}^2 || \mathrm{x}| \leq \Delta] = \sigma^2(1-\beta(\delta))$.

Proof: The property $f_{\mathrm{x}}(x||\mathrm{x}| \leq \Delta) = f_{\mathrm{x}}(x)/\int_{-\Delta}^{\Delta} f_{\mathrm{x}}(t)\mathrm{d}t$ yields

$$\mathbb{E}\left[\mathrm{x}^2 || \mathrm{x}| \leq \Delta\right] = \frac{1}{\int_{-\Delta}^{\Delta} f_{\mathrm{x}}(t)\mathrm{d}t}\int_{-\Delta}^{\Delta}\frac{t^2}{\sqrt{2\pi}}e^{-\frac{t^2}{2\sigma^2}}\mathrm{d}t$$

$$= \frac{\sigma^2}{1-2Q(\delta)}\int_{-\delta}^{\delta}\frac{y^2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\mathrm{d}y$$

where $\int_{-\delta}^{\delta}(y^2/\sqrt{2\pi})e^{-(y^2/2)}\mathrm{d}y$ can be calculated as:

$$\int_{-\delta}^{\delta}\frac{y^2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\mathrm{d}y = -\frac{1}{\sqrt{2\pi}}ye^{-\frac{y^2}{2}}\Big|_{-\delta}^{\delta} + \int_{-\delta}^{\delta}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\mathrm{d}y$$

$$= 1 - 2Q(\delta) - \frac{2}{\sqrt{2\pi}}\delta e^{-\frac{\delta^2}{2}}$$

Then $\mathbb{E}[\mathrm{x}^2 || \mathrm{x}| \leq \Delta] = \sigma^2(1-\beta(\delta))$. ■

Lemma A.2 [10, pp. 24–25]: Let $\mathrm{x} \in \mathbb{R}^n$ and $\mathrm{y} \in \mathbb{R}^m$ be jointly Gaussian with mean and variance

$$m = \begin{bmatrix}\bar{x}\\ \bar{y}\end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix}\Sigma_{\mathrm{xx}} & \Sigma_{\mathrm{xy}}\\ \Sigma_{\mathrm{yx}} & \Sigma_{\mathrm{yy}}\end{bmatrix}.$$

Then x is conditionally Gaussian given $\mathrm{y} = y$ with $f_{\mathrm{x}|\mathrm{y}}(x|y) = \mathcal{N}(\mu, \Sigma_{\mathrm{xx}} - \Sigma_{xy}\Sigma_{\mathrm{yy}}^{-1}\Sigma_{\mathrm{yx}})$ where $\mu = \bar{x} + \Sigma_{\mathrm{xy}}\Sigma_{\mathrm{yy}}^{-1}(y-\bar{y})$.

REFERENCES

[1] M. Athans, "On the determination of optimal costly measurement strategies for linear stochastic systems," *Automatica*, vol. 8, no. 4.
[2] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and dynamic sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.
[3] X. Cao, *Stochastic Learning and Optimization: A Sensitivity-Based Approach*. New York: Springer–Verlag, 2007.
[4] O. C. Imer and T. Basar, "Optimal estimation with limited measurements," in *Proc. 44th IEEE Conf. Decision and Control, European Control*, Dec. 2005, pp. 1029–1034.
[5] R. Cogill, S. Lall, and J. P. Hespanha, "A constant factor approximation algorithm for event-based sampling," in *Proc. 2007 American Control Conf.*, New York, 2007, pp. 305–311.
[6] R. Ambrosino, B. Sinopoli, and K. Poolla, "Optimal sensor scheduling for remote estimation over wireless sensor networks," in *Modeling, Estimation and Control of Networked Complex Systems*. New York: Springer, 2009, pp. 127–142.
[7] L. Li, M. Lemmon, and X. Wang, "Event-triggered state estimation in vector linear processes," in *Proc. American Control Conf.*, 2010, pp. 2138–2143.
[8] A. Ribeiro, G. B. Giannakis, and S. I. Roumeliotis, "SOI-KF: Distributed Kalman filtering with low-cost communications using the sign of innovations," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4782–4795, Dec. 2006.
[9] E. J. Msechu, S. I. Roumeliotis, A. Ribeiro, and G. B. Giannakis, "Decentralized quantized Kalman filtering with scalable communication cost," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3727–3741, Aug. 2008.
[10] B. Anderson and J. Moore, *Optimal Filtering*. Upper Saddle River, NJ: Prentice-Hall, 1979.
[11] J. Kotecha and P. Djuric, "Gaussian particle filtering," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2602–2612, Oct. 2003.

Geometric Criteria for the Quasi-Linearization of the Equations of Motion of Mechanical Systems

Dong Eui Chang and Raymond G. McLenaghan

***Abstract*—A linear transformation of velocity for a mechanical system is said to quasi-linearize the equations of motion of the system if it eliminates all terms quadratic in the velocity. It is well-known that controller/observer synthesis becomes tractable when the dynamics of a mechanical system are in quasi-linearized form. In this technical note, we show that the quasi-linearization property is equivalent to the property that the Lie algebra of Killing vector fields is pointwise equal to the tangent space to the configuration manifold with the Riemannian metric induced by the mass tensor of the mechanical system. A sufficient condition for this property is that the Riemannian manifold be locally symmetric. We further show that a necessary and sufficient condition for quasi-linearizability on 2-D Riemannian manifolds is that the scalar curvature is constant. The above results extend the zero Riemannian curvature condition that has been extensively applied since its introduction in 1992. Moreover, the local symmetricity condition and the constant scalar curvature condition can be easily verified using differentiation.**

***Index Terms*—Killing vector fields, mechanical systems, quasi-linearization.**

I. INTRODUCTION

In the Lagrangian formulation of mechanics, the equations of motion of a mechanical system on a configuration space M are normally written in the coordinate system $(x, \dot{x})$ on the tangent bundle TM that is induced by a coordinate system x on M [10]. In $(x, \dot{x})$ coordinates, the equations of motion contain quadratic terms in velocity $\dot{x}$. By introducing a quasi-velocity v defined by $v = A(x)\dot{x}$ [3], where $A(x)$ is an

Manuscript received October 03, 2011; revised April 12, 2012 and August 02, 2012; accepted September 06, 2012. Date of publication September 13, 2012; date of current version March 20, 2013. This work was supported in part by Discovery Grants from the Natural and Engineering Research Council of Canada. Recommended by Associate Editor A. Loria.

The authors are with the Department of Applied Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: dechang@math. uwaterloo.ca; rgmclena@math. uwaterloo.ca).

Digital Object Identifier 10.1109/TAC.2012.2218671