Innovations-Based State Estimation with Wireless Sensor Networks

Daniel E. Quevedo, Anders Ahlén, Jan Østergaard, and Graham C. Goodwin

***Abstract*— We study a state estimation architecture for sensor networks, where several sensors transmit quantized innovations to a central estimator. Transmission is via a wireless channel, which is prone to fading leading to random packet loss. State estimation is carried out at the gateway via a time-varying Kalman filter which accounts for packet loss and quantization effects. To form the innovations at the sensors, the estimator transmits information regarding its current state estimate to the sensors. This information could be dedicated to each sensor or broadcast to all sensors. In addition, the gateway also decides upon power levels and quantization step-sizes to be used by each sensor node. Here, we adopt elements of predictive control to trade off estimation performance versus energy use.**

I. Introduction

The use of wireless sensor networks (WSNs) for estimation and control has attracted significant attention; see, e.g., [1], [2]. Clearly, the use of wireless communications has distinct advantages. For example, installation and maintenance costs are lower than when wired sensors are used. Furthermore, wireless sensors can be placed where wires cannot go, or where power sockets are unavailable.

Perhaps, the main drawback of using wireless communications lies in the fact that wireless channels are inevitably affected by time-varying fading and interference, which frequently cause packet-loss with consequent performance degradation. As in other wireless communications technology, such as cellular (or mobile) phone systems, the effects of time-varying channel fading and interference can be alleviated by adjusting power levels. However, the use of high transmission power is rarely an option for WSNs, since in most applications, sensors are expected to be operational for several years without the replacement of batteries. Another way to diminish packet-loss probabilities lies in reducing the size of the packets transmitted. This, however, needs to be done with utmost care to avoid excessive quantization effects.

The main motivation for the present work lies in achieving enhanced performance versus energy use trade-offs when performing state estimation via WSNs. For that purpose, we will consider an LTI n_x-dimensional system:

$$x(k+1) = Ax(k) + w(k), \quad k \in \mathbb{N}_0 \triangleq \{0, 1, \dots\}, \quad (1)$$

where the initial system state $x(0) \in \mathcal{N}(0, P_0)$, P_0 is $n_x \times n_x$, and $w = \{w(k)\}_{k\in\mathbb{N}_0}$ is i.i.d., where each $w(k) \in \mathcal{N}(0, Q)$.

G. C. Goodwin and D. E. Quevedo are with the School of Electrical Engineering and Computer Science, The University of Newcastle, NSW 2308, Australia, graham.goodwin@newcastle.edu.au, dquevedo@ieee.org

A. Ahlén is with Signals and Systems, Uppsala University, SE-751 21, Uppsala, Sweden, anders.ahlen@signal.uu.se

Jan Østergaard is with the Department of Electronic Systems, Aalborg University, Aalborg, Denmark, janoe@ieee.org

To remotely estimate the state sequence $x = \{x(k)\}_{k\in\mathbb{N}_0}$, we will study a WSN with M sensor nodes. Each sensor m takes scalar noisy measurements, say $y_m = \{y_m(k)\}_{k\in\mathbb{N}_0}$:

$$y_m(k) = C_m x(k) + v_m(k), \quad m \in \{1, 2, \dots, M\}, \quad (2)$$

where $v_m = \{v_m(k)\}_{k\in\mathbb{N}_0}$ is i.i.d. with $v_m(k) \in \mathcal{N}(0, R_m)$.

The values in (2) are processed locally and then, if deemed important, sent to a single gateway (GW). In our approach, in order to make parsimonious use of the sensor energies, as much processing as possible is performed at the GW. In particular, the GW sends predictions of the output values in (2) to the corresponding sensors. The latter then send quantized innovations to the GW for state estimation. Sensor nodes do not communicate with each other.

Optimal state estimation, given quantized innovations, inherently leads to a difficult nonlinear estimation problem involving the numerical solution of Chapman-Kolmogorov equations; see, e.g., [3], [4] and also [5], [6], where estimation with quantized measurements, rather than innovations is studied. In the present work, we will adopt a powerful and widely-used design model, where quantization effects are approximated as additive noise; see, e.g., [7], [8]. State estimates are, thus, provided by a simple time-varying Kalman Filter (KF), which operates on received data and take into account quantization noise.

Within the above setting, see also Fig. 1, our contribution lies in developing a centralized predictive controller for the design of sensor transmission power levels and sensor quantization schemes. The current paper complements our recent work documented, for example, in [9], [10], by allowing for coding of innovations rather than output values. The main advantage of using coded innovations is that, in general, the average variances of the innovations are significantly lower than the variances of the output values. Thus, fewer bits can be used to represent the changes in the outputs of the sensors. Fewer bits require less energy, and will result in fewer packet losses. This in turn gives significant state estimation performance gains, as will become apparent in the simulation study documented in Section VI.

II. Processing at the sensor nodes

As foreshadowed in the introduction, in our formulation, each of the sensor nodes can adapt to changing operating conditions by using, at each time instant k, a different quantizer from within a pre-designed set. Which quantizer to choose is determined by the GW as described in detail in Section V. To avoid transmitting redundant data and, hence, be energy efficient, the sensors only send quantized innovation values to the GW. A key aspect is that, innovations

ISBN 978-963-311-369-1

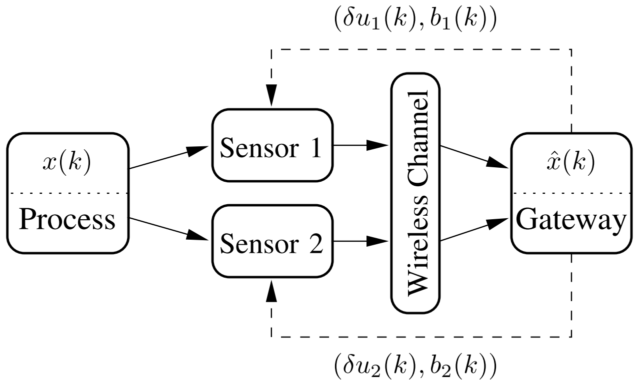

Fig. 1. State Estimation Architecture with a WSN having $M = 2$ sensors.

are computed with respect to a centralized state estimate. For that purpose, at each time $k \in \mathbb{N}$, the GW transmits a predicted sensor output estimate to each of the sensors.[1]

To be more specific, at each time k each Sensor m receives (from the GW) either the state prediction $\hat{x}(k|k-1)$ or the corresponding output prediction

$$\hat{y}_m(k|k-1) \triangleq C_m\hat{x}(k|k-1).$$

and is also told which quantizer to use. Sensor m then takes the measurement $y_m(k)$ and computes the quantized innovation value $\eta_m(k)$ in

$$\eta_m(k) = q_{\Delta_m(k)}\big(\tilde{y}_m(k)\big), \tag{3}$$

where:

$$\tilde{y}_m(k) \triangleq y_m(k) - C_m\hat{x}(k|k-1).$$

In (3), $q_{\Delta_m(k)}(\cdot)$ is a mid-tread quantizer, i.e., satisfies

$$q_{\Delta_m(k)}(z) = 0, \quad \forall z \in [-\Delta_m(k)/2, \Delta_m(k)/2)$$

for some stepsize $\Delta_m(k) > 0$; see, e.g., [11]. Stepsizes and expected bit-rates $b_m(k) \geq 0$ (after entropy coding) depend upon each other and can be used interchangeably to characterize the quantizer. Using smaller stepsizes leads to smaller quantization effects, but requires larger bitrates than using larger stepsizes. In the appendix we given more details about quantizers and also present an iterative algorithm for their off-line design.

It is worth emphasizing that the quantizer output can be calculated efficiently by simply setting:

$$\eta_m(k) = \Delta_m(k)\big\lfloor \tilde{y}_m(k)/\Delta_m(k)\big\rceil,$$

where $\lfloor\cdot\rceil$ denotes rounding to the nearest integer.

If $\eta_m(k)$ in (3) is nonzero, then Sensor m sends this value (or the associated quantization region index codeword) to the GW at the desired power level. If $\eta_m(k) = 0$, then Sensor m remains silent.[2] The situation is depicted in Fig. 2. It amounts to a specific type of *Lebesgue Sampling* [12].

[1] Alternatively, the GW could also broadcast its predicted state estimate $\hat{x}(k|k-1)$ to all M sensors.

[2] Alternatively, Sensor m could also transmit an associated codeword, which serves to confirm that the sensor can communicate with the GW.

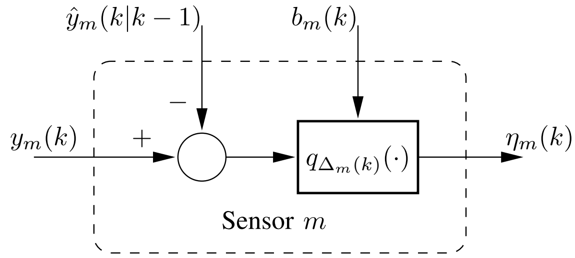

Fig. 2. Processing at Sensor m.

Given the above logic, we see that each sensor has the advantage of using a state estimate, which is built upon past information regarding measurements made by the other sensors. It is important to note that, in our scheme, sensors do not need to communicate with each other. In particular, no local state estimates are computed or broadcast.

The idea of quantizing the innovation processes $\tilde{y}_m$ is certainly not new. It underlies various predictive coding schemes schemes, see, e.g., [13], and has also been studied more recently in, e.g., [3], [4], [14]. The method also bears similarities to the *send-on-delta* concept explored, e.g., in [15], [16]. Our contribution lies in using the idea within a state estimation problem for WSNs over fading channels and where power levels and bitrates are determined by the GW.

III. Communication Issues

Transmission of the quantized innovations is through wireless links. Thus, transmission errors are likely to occur.[3] Following [9], [10], we will model transmission effects by introducing the M binary stochastic arrival processes $\gamma_m = \{\gamma_m(k)\}_{k\in N_0}$, $m \in \{1, 2, \dots, M\}$, where:

$$\gamma_m(k) = \begin{cases} 1 & \text{if } \eta_m(k) \text{ arrives error-free at time } k, \\ 0 & \text{if } \eta_m(k) \text{ does not arrive error-free at time } k. \end{cases}$$

The associated success probabilities

$$\lambda_m(k) \triangleq \mathcal{P}\big\{\gamma_m(k) = 1\big\}, \quad m \in \{1, \dots, M\},$$

depend on the propagation environment, on $b_m(k)$, and on the transmission power used by the sensor radio power amplifiers, which we denote by $u_m(k)$. Indeed, we have:

$$\lambda_m(k) = \Big(1 - \beta_m\big(u_m(k)g_m(k)\big)\Big)^{b_m(k)}, \tag{4}$$

where $g_m(k)$ denotes the channel (power) gain[4], and $\beta_m(\cdot)\colon [0,\infty) \to [0,1]$ denotes the bit-error rate (BER). The latter is a monotonically decreasing function, which depends on the modulation scheme employed.

Example 1 (AWGN Channel and BPSK): If Binary Phase Shift Keying is used over an additive white Gaussian noise

[3] In the present work, we will assume that sensor data is not affected by delays or multiple access interference. Extensions of our framework to include these issues, does not present any conceptual difficulties.

[4] Note that $g_m(k)$ is here defined to include also path-loss, power amplifier efficiency, antenna gain and noise figure.

channel with constant signal-to-noise ratio SNR, then the BER is given by $f_Q\left(\sqrt{2\text{SNR}}\right)$, where

$$f_Q(z) \triangleq \frac{1}{\sqrt{2\pi}}\int_z^{\infty} \exp\left(-\zeta^2/2\right) d\zeta$$

is the Q-function, see [17]. Although the above model is strictly only valid in the time-invariant case, we shall adopt it also for time-varying channels, power levels and bitrates. For that purpose, we introduce the instantaneous signal-to-noise ratio for each channel m via[5]

$$\text{SNR}_m(k) = \frac{g_m(k)u_m(k)}{rk_BT}, \quad m \in \{1,\dots,M\},$$

where k_B is the Boltzmann constant, T is the temperature and r is the channel bit rate. Furthermore, we adopt a block fading model, where the channel is constant over the duration of one packet, but may be subject to fading between packets. Expression (4) then gives:

$$\lambda_m(k) = \left(1 - f_Q\left(\sqrt{\frac{2g_m(k)u_m(k)}{rk_BT}}\right)\right)^{b_m(k)}. \quad (5)$$

We will use this model in Section VI. △

It follows from (4), see also (5), that one can improve transmission reliability and, thus, state estimation accuracy for a given wireless propagation environment, by transmitting shorter packets and/or by simply increasing the power used by the transmitter. Unfortunately, and as already noted in Section II, smaller values of packet lengths $b_m(k)$ will lead to larger quantization distortion. Furthermore, in WSNs, it is of fundamental importance to save energy: Sensor nodes are expected to be operational for several years without maintenance. This motivates us to use the available energy resources with care.

Before proceeding, we note that one can quantify the energy used by each sensor $m \in \{1,\dots,M\}$ at a given (discrete) time instant, k, via $E_m(b_m(k)u_m(k))$, where

$$E_m(b_m(k)u_m(k)) \triangleq \begin{cases} \dfrac{b_m(k)u_m(k)}{r} + E_{\text{P}} & \text{if } u_m(k) > 0,\\ 0 & \text{if } u_m(k) = 0. \end{cases} \quad (6)$$

Here, E_{P} denotes the processing cost, i.e., the energy needed for wake-up, circuitry and sensing.

IV. Forming the State Estimates

The GW receives quantized innovation values. In the present work, we will assume that the data transmitted incorporates error detection coding [17]. Hence, the gateway knows, whether packets received from the sensors contain errors or not. Faulty packets will be discarded when estimating the system state.

From a state estimation perspective, the overall system can be described via the state transition equation (1) together with the measurement equation:

$$y(k) = C(k)x(k) + v(k),$$

where

$$y(k) \triangleq \begin{bmatrix} y_1(k)\\ y_2(k)\\ \vdots\\ y_M(k)\end{bmatrix}, C(k) \triangleq \begin{bmatrix} \gamma_1(k)C_1\\ \gamma_2(k)C_2\\ \vdots\\ \gamma_M(k)C_M\end{bmatrix}, v(k) \triangleq \begin{bmatrix} v_1(k)\\ v_2(k)\\ \vdots\\ v_M(k)\end{bmatrix}.$$

To achieve mean square optimality, one would like to find the expected value of the state $x(k)$ conditioned upon the received quantized data. Unfortunately, due to the quantization aspect, a difficult nonlinear estimation problem needs to be solved, see, e.g., [3], [6]. In the present work, we will use a simple, though powerful, estimation strategy. It is based upon an additive noise model for the quantizers, which we describe next.

A. Quantization model

We will model quantization effects via (recall (3)):

$$\eta_m(k) = \tilde{y}_m(k) + n_m(k) = y_m(k) - C_m\hat{x}(k|k-1) + n_m(k), \quad (7)$$

where $n_m(k)$ is the quantization noise at Sensor m, which we assume independent and with variance $D_m(k)$. The latter depends upon the associated bitrate $b_m(k)$ and is proportional to the variance of $\tilde{y}_m$ and not the variance of y_m. This illustrates the performance gain achieved when quantizing output innovations rather than output values.

If we now introduce the overall quantization noise n via:

$$n(k) = \begin{bmatrix} n_1(k) & \dots & n_M(k)\end{bmatrix}^T,$$

then (7) can be re-written as

$$\eta(k) = C(k)x(k) + v(k) + n(k) - \hat{y}(k|k-1),$$

where

$$\eta(k) \triangleq \begin{bmatrix} q_{\Delta_1(k)}(\tilde{y}_1(k)) & \dots & q_{\Delta_M(k)}(\tilde{y}_M(k))\end{bmatrix}^T \quad (8)$$

contains the quantized innovations, see (3).

We conclude that the effect of quantization is captured by the equivalent observation noise process

$$\nu(k) \triangleq v(k) + n(k), \quad \forall k \in \mathbb{N}_0,$$

whose variance is given by

$$R(k) = \operatorname{diag}\big(R_1 + D_1(k), \dots, R_M + D_M(k)\big). \quad (9)$$

B. Kalman Filter with Intermittent Innovations

Within the quantization model adopted, the KF with intermittent output innovations described below gives the linear minimum variance estimate of the system state; see also [18], [19] which study KF with intermittent output observations.

In our formulation, the updated state estimate, which uses the prediction

$$\hat{x}(k|k-1) = A\hat{x}(k-1|k-1)$$

and the received quantized innovation $\eta_m(k)$, satisfies

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k)\eta(k),$$

[5] $\text{SNR}_m(k)$ denotes the signal-to-noise ratio at the receiver, after the matched filter.

where:

$$K(k) \triangleq P(k)C(k)^T\big(C(k)P(k)C(k)^T + R(k)\big)^{-1} \quad (10)$$

and where

$$P(k+1) \triangleq AP(k)A^T + Q - AK(k)C(k)P(k)A^T$$

is the associated (posterior) covariance matrix. In (10), $R(k)$ is the equivalent noise covariance given in (9), with

$$D_m(k) = \Big(\frac{\pi e}{6}2^{-2b_m(k)}\Big)\,\Sigma_m(k),\ \Sigma_m(k) = C_m P(k) C_m^T\ . \quad (11)$$

By using high-rate approximations, i.e. by replacing $\Sigma_m(k)$ with a constant matrix Σ_m, we can avoid quantization to be performed at the sensors. Instead the quantizers can be designed off-line, for efficient implementation as a table in the sensor memories, see the appendix. The recursion in (10) is initialized with values $P(0) = P_0$ and $\hat{x}(0) = 0$.

It is important to note that the estimator referred to above operates utilizing only received data. This gives a degree of robustness with respect to sensors not being able to transmit due to, e.g., being damaged or out of battery power.

V. Predictive Control of Transmission Power Levels and Quantization Stepsizes

As we have seen, the design of sensor power signals $\{u_m\}$ and quantizer bitrates $\{b_m\}$ involves a trade-off between transmission error probabilities (and state estimation accuracy) and energy consumption at the sensor nodes. We will next present a predictive control strategy which minimizes a cost function, which quantifies this trade-off over a future prediction horizon. To keep processing at the sensors to a minimum, the controller is located at the GW, see Fig. 1.

A. Constraints

In order to save energy required to process the received command signal at the sensors, we would like to keep the signaling from the GW to the sensors as low as possible. In particular, the control signal will contain the bitrates and information on the power levels. Here, and arguing as in [9], [10], we will use coding ideas frequently used in power control architectures for cellular networks, see, e.g., [20] (and compare also to work on Networked Control in [6]). We, thus, send coarsely quantized power increments, say $\delta u_m(k)$, rather than actual power values, to each sensor $m \in \{1, 2, \dots, M\}$. All signals $\{\delta u_m\}$ are constrained according to:

$$\delta u_m(k) \in \mathbb{U}_m, \quad \forall k \in \mathbb{N}_0,\ \forall m \in \{1, 2, \dots, M\}, \quad (12)$$

where $\{\mathbb{U}_m\}$ are given finite sets, each having a small number of elements.

Upon reception of the pair $(\delta u_m(k), b_m(k))$, each sensor m chooses the associated quantizer and reconstructs the power level to be used by its radio power amplifier by simply setting

$$u_m(k) = u_m(k-1) + \delta u_m(k). \quad (13)$$

The quantization constraint (12) imposes

$$\delta u(k) \in \mathbb{U} \triangleq \mathbb{U}_1 \times \mathbb{U}_2 \times \cdots \times \mathbb{U}_M, \quad \forall k \in \mathbb{N}_0,$$

where

$$\delta u(k) \triangleq \begin{bmatrix}\delta u_1(k) & \dots & \delta u_M(k)\end{bmatrix}^T, \quad k \in \mathbb{N}_0. \quad (14)$$

In addition, due to physical limitations of the radio power amplifiers, the actual sensor power levels are constrained in magnitude according to:

$$0 \le u_m(k) \le u_m^{\max}, \quad \forall k \in \mathbb{N}_0,\ \forall m \in \{1, \dots, M\},$$

for given values $\{u_m^{\max}\}$.

B. Cost Function

We will quantify state estimation accuracy via the trace of the matrix $\bar{P}(k)$ defined as:

$$\bar{P}(k) = P(k) - K(k)C(k)P(k),$$

which, if $n(k)$ was Gaussian, would correspond to the posterior covariance of the estimation error.

At each time instant $k \in \mathbb{N}_0$, the predictive controller first calculates the value of $\bar{P}(k)$, which results from iterating the KF equations of Section IV-B for the (known) past realizations of reconstruction outcomes, say γ^k. It also uses channel gain predictions over a finite horizon of fixed length N.

With this information, the controller minimizes the finite-set constrained cost function

$$V(\delta U, B) \triangleq \mathbb{E}_{\Gamma(k)}\{V_1(\Gamma(k), \delta U, B)\} + \rho V_2(\delta U), \quad (15)$$

where $\rho \ge 0$ is a parameter which allows the designer to trade estimation accuracy for energy consumption.

The stochastic aspect of the power and bitrate control problem, namely the possibility of packet-loss, is captured in (15) by the discrete stochastic matrix

$$\Gamma(k) \triangleq \begin{bmatrix} \gamma_1(k+1) & \gamma_1(k+2) & \dots & \gamma_1(k+N) \\ \gamma_2(k+1) & \gamma_2(k+2) & \dots & \gamma_2(k+N) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_M(k+1) & \gamma_M(k+2) & \dots & \gamma_M(k+N) \end{bmatrix}.$$

Accordingly, $\mathbb{E}_{\Gamma(k)}$ denotes expectation taken with respect to the probability mass distribution of $\Gamma(k)$.[6] Note that as summarized in (5), this distribution depends upon the power levels, bitrates, and channel gains. Thus, $\Gamma(k)$ can be regarded as a controlled stochastic disturbance to the system.

The cost function in (15) measures estimation quality and energy use respectively via the terms

$$\begin{aligned} V_1(\Gamma(k), \delta U, B) &\triangleq \sum_{\ell=k+1}^{k+N} \operatorname{trace}\big(\bar{P}'(\ell)\big), \\ V_2(\delta U, B) &\triangleq \sum_{\ell=k+1}^{k+N}\sum_{m=1}^{M} E_m(b'_m(\ell)u'_m(\ell)). \end{aligned} \quad (16)$$

[6] Compare to the scenario based approach taken in [21].

The decision variables are collected in[7]

$$\delta U = \{\delta u'(k+1), \delta u'(k+2), \dots, \delta u'(k+N)\}$$
$$B = \{b'(k+1), b'(k+2), \dots, b'(k+N)\},$$

where $\delta u'(\ell)$ is as in (14) and:

$$b'(\ell) \triangleq \begin{bmatrix} b'_1(\ell) & b'_2(\ell) & \dots & b'_M(\ell) \end{bmatrix}^T.$$

In accordance with (13), δU yields the tentative future power levels $\{u'_m(\ell)\}$ in (16) via

$$u'_m(\ell) = u'_m(\ell-1) + \delta u'_m(\ell),$$
$$\ell \in \{k+1, k+2, \dots, k+N\},\ m \in \{1, 2, \dots, M\}$$

starting from the current levels, i.e., $u'_m(k) = u_m(k)$. The term $E_m(b'_m(k)u'_m(\ell))$ is the energy function (6) evaluated for the tentative values $u'_m(\ell)$ and $b'_m(\ell)$. For a given realization of $\Gamma(k)$, $\operatorname{trace}(\bar{P}'(\ell))$ is obtained from (16) after iterating the KF equations of Section IV-B with initial value $P(k+1)$.

C. The Resultant Controller

At each time instant $k \in \mathbb{N}_0$, and given channel gain predictions, the controller finds the optimizing sequences

$$(\delta U^{\text{opt}}, B^{\text{opt}}) \triangleq \arg\min V(\delta U, B), \tag{17}$$

subject to the constraints:

$$\delta U \in \mathbb{U}^N, \quad B \in \mathcal{B}^N$$
$$0 \leq u'_m(\ell) \leq u_m^{\max}, \quad \forall \ell \in \{k+1, \dots, k+N\},\ \forall m,$$

where $\mathbb{U}^N \triangleq \mathbb{U} \times \mathbb{U} \times \dots \times \mathbb{U}$, and $\mathcal{B}$ is a given (finite) constraint set for the bitrates of all sensors.

Following the moving horizon principle, see, e.g., [22], at each time k, the proposed controller sends only the M bitrates and power updates contained in[8]

$$\delta u(k+1)^{\text{opt}} \triangleq \begin{bmatrix} I_M & 0_M & \dots & 0_M \end{bmatrix} \delta U^{\text{opt}}$$
$$b(k+1)^{\text{opt}} \triangleq \begin{bmatrix} I_M & 0_M & \dots & 0_M \end{bmatrix} B^{\text{opt}}$$

to the corresponding sensors. At the next time instant, namely $k+1$, the optimization procedure is repeated, giving rise to power control increments $\delta u(k+2)^{\text{opt}}$ and bitrates $b(k+2)^{\text{opt}}$. This procedure is repeated *ad infinitum*.

The proposed controller optimizes the resulting performance via dynamic selection of power levels and bitrates. Note that to calculate future success probabilities, one requires channel gain predictions. For that purpose one can use techniques described, e.g., in [23], [24].

It is worth emphasizing that minimization of $V(\delta U, B)$ in (17) is carried out on-line at the GW, where computational issues play less of a role than at the sensors. Interestingly, despite the fact that we are dealing with a stochastic nonlinear optimization problem, solving (17) in real-time is surprisingly simple. In particular, due to the finite-set nature of $\Gamma(k)$, B, and δU, taking expectation in (15) reduces to evaluating a finite sum.

[7]In the sequel, primed variables refer to tentative values of the corresponding physical variables.

[8]I_M denotes the $M \times M$ identity matrix and 0_M the all zeros $M \times M$ matrix.

VI. Simulation Study

In the simulations we will consider the system model (1) with

$$A = \begin{bmatrix} 1.6718 & -0.9948 \\ 1 & 0 \end{bmatrix}, \tag{18}$$

$Q = 1/2I$ and $P_0 = 0.3I$. We simulate a WSN having $M = 2$ sensors with $C_1 = [1, 0], C_2 = [0, 1]$, and variances $R_1 = R_2 = 1/100$. We use power levels in the range $0 \leq u_m(k) \leq 3 \cdot 10^{-4}$ with increments of $\delta u_m(k) \in \{0, \pm 3 \cdot 10^{-5}\}$ and the bit-rates satisfy $b_m(k) \in \{2, \dots, 7\}$.

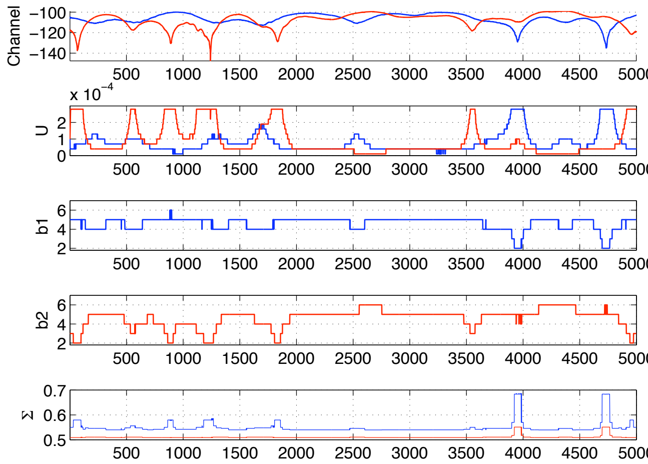

Fig. 3. Simulation results when sending quantized innovations.

At each time instant we minimize the weighted cost given by (15) with $\rho = 10^6$. The minimization is carried out as a brute-force search over all possible combinations of bit-rates and power increments for the two sensors. The complete simulation, running through 5000 data samples, is shown in Fig. 3. In Fig. 3 the top diagram depicts the channel gains for the wireless communication channels between the sensors and the GW. The graph starting at the highest gain at $t = 0$ refers to Sensor 1 and the other one to Sensor 2. The second diagram shows the chosen power levels. Here Sensor 2 starts with a saturated power level caused by the corresponding channel fading dip, cf the top diagram, whereas the power level of Sensor 1 is more cautious since the corresponding channel gain is relatively good. Notice that whenever the sensor channel gains are dropping, the corresponding power levels are increased.

The third and forth diagrams from the top, illustrate the chosen bit-rates for Sensor 1 and Sensor 2, respectively. Finally, in the bottom diagram the two graphs show the variances of the respective sensor innovation sequences.

From the different diagrams of Fig. 3 we observe the following: When the channel gains are dropping the power levels are frequently saturating. If the allocated power does not manage to lift the gain to an acceptable level, then the corresponding bit-rate is decreased. See. for example, b_2 at the time instances 900 and 1250, or, b_1 at time instances 3950 and 4750, respectively. The reason why the power control algorithm decreases the bit-rate is that the low channel gain makes packet loss likely to occur, and therefore it is better not to waist energy.

From the two graphs of the bottom diagram we observe that whenever there is a dip in any of the channel gains, see the top diagram, the variance of innovations for Sensor 1 increases (uppermost graph of the bottom diagram). The deeper the fade, the higher the variance. This is due to the strong coupling between $x_1(k+1)$ and $x_1(k)$, $x_2(k)$, see (18). It is only at the deep fades of Channel 1, at the time instances 3950 and 4750, respectively, that the innovation variance of Sensor 2 is increasing.

In order to compare the benefits of sending coded versions of the innovations $\tilde{y}$ we compare the simulation depicted in Fig. 3 with the situation where we encode the output values y. The results are summarized in Table I. From Table I we note a significant improvement in both state estimation accuracy and energy expenditure, as measured by V_1 and V_2, respectively, see (16). The gain in total cost by using coded innovations, as measured by V, is a noteworthy $(0.1219-0.0948)/0.1219 = 22.23\%$.[9] This gain is primarily due to the reduction in variances, which is, on average, about a factor of 40.

TABLE I

PERFORMANCE COMPARISON BETWEEN USING CODED INNOVATIONS AND CODED OUTPUT VALUES.

	Avg. variances	V_1	V_2	V
y	$\sigma^2_{y_1} = 21.49$ $\sigma^2_{y_2} = 21.93$	0.0692	0.0527	0.1219
$\tilde{y}$	$\Sigma_1 = 0.5492$ $\Sigma_2 = 0.5104$	0.0521	0.0426	0.0948

VII. CONCLUSIONS

We have studied state estimation via sensor networks over fading wireless channels where sensors send coded innovations instead of output values. A model predictive controller at the gateway allocate the power levels and bit-rates to be used by each sensor to optimize a cost function which trades off state estimation accuracy for energy expenditure. A simulation study shows that a performance gain of a noteworthy 22% can be attained if coded innovations are used instead of coded outputs. The performance gain is due to a lower average variance of the innovations as compared to the outputs.

APPENDIX
QUANTIZER DESIGN

In the WSN architecture under consideration, the GW commands the sensors to use quantizers from a pre-designed set. These quantizers are designed off-line and stored at the sensors. Our designs are based upon high-rate approximations, as detailed below:

[9] It should be remarked that we have neglected the processing cost at the sensors, which is caused by receiving the output predictions $\hat{y}(k|k-1)$ from the GW. This cost will, however, be small for high path losses, or long distances, as the transmission cost will then dominate over the processing cost.

A. Linear Quantization Model

It is well known that given a stationary input process $\tilde{y}_m$ with variance Σ_m and under high-resolution assumptions, the expected distortion of an entropy-constrained scalar uniform quantizer satisfies[10]:

$$D_m(k) \approx \Big(\frac{\pi e}{6} 2^{-2b_m(k)}\Big) \Sigma_m. \tag{19}$$

The expected bitrates $b_m(k)$ are related to the stepsizes $\Delta_m(k)$ according to [7], [25]:

$$b_m(k) \approx H(\eta_m(k)) \approx h(\tilde{y}_m(k)) - \log_2(\Delta_m(k)),$$

where $H(\eta_m(k))$ denotes the discrete entropy of $\eta_m(k)$, and $h(\tilde{y}_m(k))$ denotes the differential entropy of $\tilde{y}_m(k)$. Under Gaussian assumptions, the latter is given by:[11]

$$h(\tilde{y}_m(k)) = \big(1/2\big) \log_2(2\pi e \Sigma_m). \tag{20}$$

Consequently, bitrates and step-sizes are related via:

$$\big(\Delta_m(k)\big)^2 = \big(2\pi e \Sigma_m\big) 2^{-2b_m(k)}. \tag{21}$$

In the situation at hand, due to random packet loss and variable bitrates, the quantizer inputs in (3) will, in general, not be stationary. This makes deriving expressions which relate the quantizer parameters a formidable task.

B. Iterative Algorithm

To obtain the variances Σ_m, which are needed in the quantizer design equations included in the previous section, we will neglect the effect of packet-dropouts and assume stationarity. This gives

$$\Sigma_m = C_m \tilde{P} C_m^T,$$

where $\tilde{P}$ solves:

$$\tilde{P} = A\tilde{P}A^T + Q - A\tilde{P}C^T\big(C\tilde{P}C^T + R(k)\big)^{-1} C\tilde{P}A^T \tag{22}$$

with

$$C \triangleq \big[(C_1)^T \quad (C_2)^T \quad \dots \quad (C_M)^T\big]^T.$$

We note that in (22), the equivalent noise covariance matrix $R(k)$ is as in (9) and, thus, depends on Σ_m, see (19). This makes solving for Σ_m difficult.

To overcome this problem, we adopt an iterative method for finding Σ_m whereby bitrates are fixed at representative values, say $\bar{b}_m$. Thus, (19) reduces to:

$$D_m(k) \approx \Big(\frac{\pi e}{6} 2^{-2\bar{b}_m}\Big) C_m \tilde{P} C_m^T. \tag{23}$$

In the first iteration, we set $R(k) = \text{diag}\big(R_1, \dots, R_M\big)$ and solve for $\tilde{P}$ in (22). Equations (23) and (9) then give the value of $R(k)$ to be used in the next iteration, etc.

[10] Expressions (19) and (21) are valid under high-resolution conditions for any smooth source, i.e., any source with finite variance and finite differential entropy.

[11] In addition, Expression (20) constitutes an upper bound for all smooth sources with the same variance Σ_m. Thus, if (20) is used for other sources, then one slightly over-estimates the actual bit-rate for a given distortion or, vice-versa, if one fixes b_m, then the given $\Delta_m(k)$ will yield a slightly smaller distortion $D_m(k)$ than what is predicted by (23).

The values $\{\Sigma_m\}_{m\in\{1,2,\dots,M\}}$ obtained after a few iterations are then replaced in expression (21) to characterize the quantizers for *any* set of bitrates. The values are then used for off-line quantizer design. Extensive simulations suggest that the algorithm converges quickly. We note that, for any fixed value of $R(k)$, expression (22) amounts to an algebraic Riccati Equation, for which efficient solvers exist.

REFERENCES

[1] H. Gharavi and S. P. Kumar, "Special section on sensor networks and applications," *Proc. IEEE*, vol. 91, no. 8, pp. 1151–1152, Aug. 2003.

[2] X. Shen, Q. Zhang, and R. Caiming Qiu, "Wireless sensor networking [guest ed.]," *IEEE Wireless Commun.*, vol. 14, no. 6, pp. 4–5, Dec. 2007.

[3] E. J. Msechu, S. I. Roumeliotis, A. Ribeiro, and G. B. Giannakis, "Decentralized quantized Kalman filtering with scalable communication cost," *IEEE Trans. Signal Processing*, vol. 56, no. 8, pp. 3727–3741, Aug. 2008.

[4] K. You, L. Xie, S. Sun, and W. Xiao, "Multiple level quantized innovation Kalman filter," in *Proc. IFAC World Congr.*, Seoul, Korea, 2008, pp. 1420–1425.

[5] R. Curry, W. Vander Velde, and J. Potter, "Nonlinear estimation with quantized measurements–PCM, predictive quantization, and data compression," *IEEE Trans. Inform. Theory*, vol. 16, no. 2, pp. 152–161, Mar. 1970.

[6] G. C. Goodwin, H. Haimovich, D. E. Quevedo, and J. S. Welsh, "A moving horizon approach to networked control system design," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1427–1445, Sept. 2004.

[7] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic, Boston, MA, 1992.

[8] G. C. Goodwin, D. E. Quevedo, and E. I. Silva, "Architectures and coder design for networked control systems," *Automatica*, vol. 44, no. 1, pp. 248–257, Jan. 2008.

[9] D. E. Quevedo, A. Ahlén, and G. C. Goodwin, "Predictive power control of wireless sensor networks for closed loop control," in *Nonlinear Model Predictive Control: Towards New Challenging Applications*, L. Magni, D. M. Raimondo, and F. Allgöwer, Eds., vol. 384 of *LNCIS*, pp. 215–224. Springer-Verlag, Berlin Heidelberg, 2009.

[10] D. E. Quevedo and A. Ahlén, "A predictive power control scheme for energy efficient state estimation via wireless sensor networks," in *Proc. IEEE Conf. Decis. Contr.*, Cancún, México, Dec. 2008.

[11] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[12] K. J. Åström and B. M. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proc. IEEE Conf. Decis. Contr.*, 2002, pp. 2011–2016.

[13] P. Noll, "On predictive quantizing schemes," *The Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1499–1532, 1978.

[14] M. Fu and C.E. de Souza, "State estimation using quantized measurements," in *Proc. IFAC World Congr.*, Seoul, Korea, 2008, pp. 12492–12497.

[15] J. K. Yook, D. M. Tilbury, and N. R. Soparkar, "Trading communication for bandwidth: Communicatioin in distributed control systems using state estimators," *IEEE Trans. Contr. Syst. Technol.*, vol. 10, no. 4, pp. 503–517, July 2002.

[16] Y. S. Suh, V. H. Nguyen, and Y. S. Ro, "Modified Kalman filter for networked monitoring systems employing a send-on-delta method," *Automatica*, vol. 43, pp. 332–338, 2007.

[17] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, N.Y., 3rd edition, 1995.

[18] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1453–1464, Sept. 2004.

[19] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, pp. 598–607, 2007.

[20] H. Holma and A. Toskala, Eds., *WCDMA for UMTS*, John Wiley & Sons, West Sussex, England, third edition, 2004.

[21] D. E. Quevedo, E. I. Silva, and G. C. Goodwin, "Control over unreliable networks affected by packet erasures and variable transmission delays," *IEEE J. Select. Areas Commun.*, vol. 26, no. 4, pp. 672–685, May 2008.

[22] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Optimality and stability," *Automatica*, vol. 36, no. 6, pp. 789–814, June 2000.

[23] T. Ekman, M. Sternad, and A. Ahlén, "Unbiased power prediction of Rayleigh fading channels," in *Proc. Vehicular Technology Conf.*, 2002, vol. 1, pp. 280–284.

[24] T. Ekman, *Prediction of Mobile Radio Channels: Modeling and Design*, Ph.D. thesis, Uppsala University, Oct. 2002.

[25] V. K. Goyal, "High-rate transform coding: How high is high, and Does it matter?," in *Proc. IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, 2000, p. 207.