Power Control and Coding Formulation for State Estimation with Wireless Sensors

Quevedo, Daniel; Østergaard, Jan; Ahlen, Anders

Published in:
I E E E Transactions on Control Systems Technology

DOI (link to publication from Publisher):
10.1109/TCST.2013.2253464

Publication date:
2014

Document Version
Early version, also known as pre-print

A Power Control and Coding Formulation for State Estimation with Wireless Sensors

Daniel E. Quevedo*, *Member, IEEE,* Jan Østergaard, *Senior Member, IEEE,* and Anders Ahlén, *Senior Member, IEEE*

***Abstract*— Technological advances have made wireless sensors cheap and reliable enough to be brought into industrial use. A major challenge arises from the fact that wireless channels introduce random packet dropouts. Power control and coding are key enabling technologies in wireless communications to ensure efficient communications. In the present work, we examine the role of power control and coding for Kalman filtering over wireless correlated channels. Two estimation architectures are considered: In the first, the sensors send their measurements directly to a single gateway. In the second scheme, wireless relay nodes provide additional links. The gateway decides on the coding scheme and the transmitter power levels of the wireless nodes. The decision process is carried out on-line and adapts to varying channel conditions in order to improve the trade-off between state estimation accuracy and energy expenditure. In combination with predictive power control, we investigate the use of multiple-description coding, zero-error coding and network coding and provide sufficient conditions for the expectation of the estimation error covariance matrix to be bounded. Numerical results suggest that the proposed method may lead to energy savings of around 50%, when compared to an alternative scheme, wherein transmission power levels and bit-rates are governed by simple logic. In particular, zero-error coding is preferable at time instances with high channel gains, whereas multiple-description coding is superior for time instances with low gains. When channels between the sensors and the gateway are in deep fades, network coding improves estimation accuracy significantly without sacrificing energy efficiency.**

***Index Terms*— wireless sensors, Kalman filtering, power control, multiple-description coding, distributed source coding, network coding, relays**

I. Introduction

Wireless sensors (WSs) have become an important alternative to wired sensors [1]–[3]. WSs are equipped with a sensing component (to measure e.g., temperature), a processing device (to perform simple computations on the measured raw data), and a communication device. WSs are cheap and reliable and offer several advantages, such as, flexibility, low cost, and fast deployment. In addition, with WSs electrical contact problems are no longer an issue. Furthermore, WSs and actuators can be placed where wires cannot go, or where power sockets are unavailable.

One major drawback of using WSs is that wireless communication channels are subject to fading and interference, causing random packet errors [4]. The time-variability of the fading channel can be alleviated by adjusting the power levels and the transmitted packet lengths [5], [6]. To keep packet error rates low, short packet lengths and high transmission power should be used. However, the use of high transmission power is rarely an option, since in most applications WSs are expected to be operational for several years without the replacement of batteries; cf., [7]. In addition, short packets may require coarse quantization which may lead to large quantization effects unless careful coding is used [8], [9]. It is safe to assume that, as in other wireless communication applications, power control and coding will become key enabling technologies whenever WSs are used. In particular, due to their wide applicability, including nonlinear constrained MIMO systems (see, e.g., [10]–[14] for recent application studies), the use of predictive control methods is worth investigating.

In this work, we study two architectures having M WSs and a single gateway (GW) for Kalman-filter based state estimation of linear time-invariant (LTI) systems of the form:

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

where $x(0) \in \mathbb{R}^n$ is zero-mean Gaussian distributed with covariance matrix P_0 and the driving noise process $\{w(k)\}_{k\in\mathbb{N}_0}$ is independent and identically distributed (i.i.d.) zero-mean Gaussian distributed with covariance matrix Q. The measurement of sensor m, at time k, is given by

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

where $\{v_m(k)\}$ is i.i.d. zero-mean Gaussian measurement noise with covariance matrix R_m.

The first estimation architecture examined is depicted in Fig. 1 for the particular case of having $M = 2$ WSs. The measurements given by (2) are encoded and transmitted at an appropriate power level over a fading channel (generating random packet loss) to the GW. Received packets are then used to estimate $x(k)$ by means of a time-varying Kalman filter (KF) which takes into account packet loss. As depicted in Fig. 1, in addition to performing state estimation, the GW also controls the power levels and the coding method (including bit-rate) used by the sensors at each time. One of the main purposes of the present work is to show how predictive control methods can be used for this purpose. To keep the sensors simple and energy efficient, the sensor nodes are not allowed to communicate with each other. Thus, joint encoding of the measurements taken by different sensors is not possible. However, in the case where several measurements are received by the GW, joint decoding is possible. By

Daniel Quevedo is with the School of Electrical Engineering & Computer Science, The University of Newcastle, Australia; dquevedo@ieee.org. Jan Østergaard is with the Department of Electronic Systems, Aalborg University, Denmark; janoe@ieee.org. Anders Ahlén is with Signals and Systems, Uppsala University, Sweden; Anders.Ahlen@signal.uu.se.

This research was supported under Australian Research Council's Discovery Projects funding scheme (DP0988601).

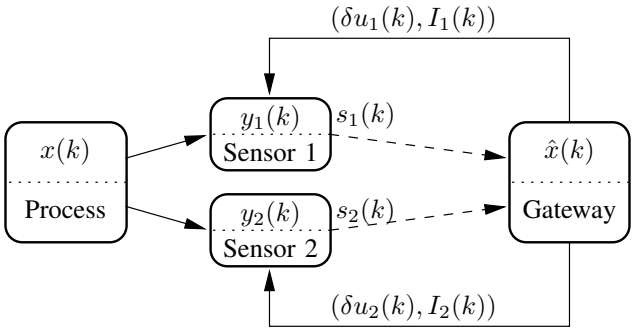

Fig. 1. State Estimation with $M = 2$ wireless sensors. The dashed lines denote fading channels which introduce random transmission errors. The gateway performs state estimation. It also controls the power level updates, $\delta u_m(k)$, and the coding method, as described in the codebook index $I_m(k)$.

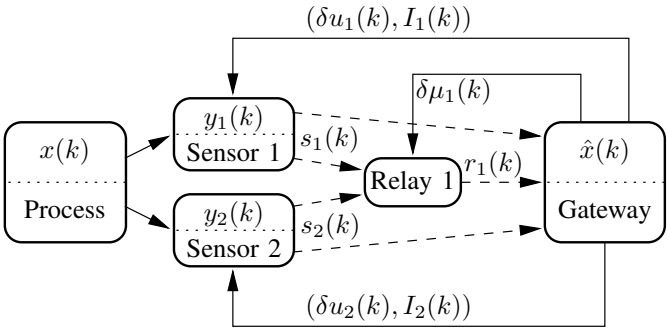

Fig. 2. State Estimation with two wireless sensors and $L = 1$ relay node. Here the GW calculates $\hat{x}(k)$, the power level updates of the sensors and relays, and the bit-rate, $I_m(k)$, of each sensor node.

allowing separate encoding to be followed by joint decoding, it is possible to take advantage of distributed source coding techniques [15], [16]. In the present work, we will focus on a particular distributed source coding technique known as zero-error coding (ZEC) [17]. In addition, to achieve robustness in the presence of packet loss, we allow the sensors to use multiple-description coding (MDC) [18].

In the second estimation architecture studied, the incorporation of L relay nodes allows for additional communication links, see Fig. 2. Here, the measurements in (2) are quantized (with a uniform quantizer) and transmitted at an appropriate power level over fading channels to the GW and relays. The latter perform network coding and forward processed sensor measurements whenever appropriate to the GW. To avoid interference between nodes, the communication channel is accessed in a TDMA fashion with a pre-designed protocol. At the GW, received packets from the sensors and relays are then used to estimate $x(k)$ via Kalman filtering. For this second architecture, the sensors do not perform MDC or ZEC. Thus, the codebook indices $I_m(k)$ amount to the bit-rates to be used by the sensors.

The main contribution of the present work is to investigate the role of dynamic power control and coding for state estimation with WSs through use of predictive control. The objective of the controller is to counteract channel variability and to trade-off battery use for estimation accuracy. It is located at the GW and decides upon the transmission power level and coding scheme to be used by each node. Our results indicate that it is advantageous that power levels approximately invert channel gains provided sufficient power is available and that MDC be used at the sensors when the channel conditions are poor. When good channel conditions are expected, it pays off to use ZEC across the sensors. If relays are available, then it turns out that, when channels between sensors and the GW are subject to severe fading, the use of network coding will improve the estimation performance significantly without increasing energy expenditure compared to the case with no relays. Hence, network coding is an attractive alternative/complement to MDC.

The present work extends our recent work documented in [19]–[23]. The papers [19], [20] introduced the idea of using predictive control of WS power levels for dynamic state estimation and control applications. In [21], the combination of power control and ZEC was considered, our conference contribution [23] examined network coding for architectures with relays, whereas in [22], the combination of power control and MDC was considered. The results reported in [21]–[23] indicate that with simple coding techniques, significant improvements over the uncoded case [19] can be achieved. This motivates the present paper, which combines power control, with either ZEC and MDC, or network coding.

Notation: The trace of a matrix A is denoted by $\operatorname{tr} A$, and its spectral norm by $||A|| \triangleq \sqrt{\max \operatorname{eigs}(A^T A)}$, where $\operatorname{eigs}(A^T A)$ are the eigenvalues of $A^T A$ and the superscript T refers to transposition. The Euclidean norm of a vector x is denoted $|x|$; $\mathbf{Pr}\{\cdot\}$ refers to probability, and $\mathbf{E}\{\cdot\}$ to expectation. Discrete entropy is denoted $H(\cdot)$; for differential entropy we use $h(\cdot)$.

II. Coding Aspects

In this section, we revise some basic aspects on source coding. Throughout this work, we will use standard high-resolution source coding results; see, e.g., [24].

A. Scalar Quantization, Entropy Coding, and High-Resolution Source Coding

Each sensor m encodes its measurement $y_m(k) \in \mathbb{R}$ into a quantized version $\hat{y}_m(k)$, which is further represented by a sequence of bits $s_m(k)$ to be transmitted over the channel, see Fig. 3. The average bit-rate of $s_m(k)$ is denoted $b_m(k)$. The encoder consists of a (time-varying) uniform scalar quantizer $\mathcal{Q}_m$ having step-size $\Delta_m(k)$, which is followed by an entropy encoder $\mathcal{E}_m$. A scalar uniform quantizer can be efficiently implemented by simply scaling $y_m(k)$ by $\Delta_m(k)$ followed by rounding, i.e., by forming $\lfloor y_m(k)/\Delta_m(k) \rceil$ where $\lfloor \cdot \rceil$ denotes rounding to the nearest integer. If the GW receives $s_m(k)$, it reconstructs $\hat{y}_m(k)$ by simply applying the inverse scaling,

$$\hat{y}_m(k) = \lfloor y_m(k)/\Delta_m(k) \rceil \Delta_m(k).$$

Under high-resolution assumptions, the bit-rate is given by[1]

$$b_m(k) \approx H(\hat{y}_m(k)) \approx h(y_m(k)) - \log_2(\Delta_m(k)), \quad (3)$$

[1]The approximation becomes exact in the limit as the distortion tends to zero [24]. However, it is also known that these high-resolution results are approximately true even at rates as low as 2 bit/dimension; cf. [25].

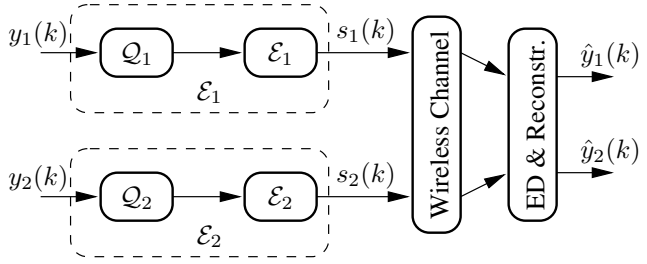

Fig. 3. Coding with $M = 2$ WSs. Measurements $y_1(k)$ and $y_2(k)$ are quantized, entropy coded and transmitted over fading channels. At the receiver, entropy decoding (ED) and reconstruction yields $\hat{y}_1(k)$ and $\hat{y}_2(k)$.

where $\Delta_m(k)$ denotes the step-size and the expected distortion $D_m(k)$ satisfies

$$D_m(k) \triangleq \mathbf{E}\{|y_m(k) - \hat{y}_m(k)|^2 \,|\, b_m(k) = b\} \approx \frac{1}{12} 2^{-2(b-h(y_m(k)))}. \quad (4)$$

In the sequel, we assume that $\{y_m(k)\}_{k\in\mathbb{N}_0}$ is a stationary process with $y_m(k)$ being zero-mean Gaussian with variance $\sigma^2_{y_m}$.

The entropy coder consists of a codebook[2], which due to memory considerations cannot be arbitrarily large. In practice, we choose the size of the codebooks so that the probability of falling outside the support of the entropy coder is very small and the impact of the outliers on the distortion is negligible. Since a codebook is needed for every possible $\Delta_m(k)$, it is necessary to discretize the alphabet of $\Delta_m(k)$ or, equivalently, to discretize the set of possible bit-rates $b_m(k)$. In the following, we will assume that an appropriate discretization for a given system (A, C) is found offline through, e.g., computer simulations. Thus, the controller uses the constraint

$$b_m(k) \in \mathcal{B}_m, \quad \forall m \in \{1, 2 \ldots, M\} \quad (5)$$

for given finite sets $\mathcal{B}_m \subset (0, \infty)$. In particular, in our results documented in Section VII, we confined the bit-rates $b_m(k)$ to the sets $\mathcal{B}_m = \{3, \ldots, 8\}$.

B. Zero-Error Coding

As mentioned in the introduction, the M WSs are separated and cannot communicate with each other. Encoding of the measurements can therefore not be done jointly. However, the GW sees all the received measurements and thereby can perform centralized joint decoding. Thus, we are facing the *distributed source coding* problem, i.e., separate encoding of M correlated variables followed by joint decoding [9], [15], [16]. In this work, the GW will, at times, command the WSs to adopt a distributed source coding technique, known as zero-error coding (ZEC) [17]. With ZEC, the measurements are quantized independently using the same scalar quantizers as previously designed for the case of independent coding. The only change is with regard to the entropy coder: rather than employing independent entropy coding on the quantized measurements, with ZEC the WSs use dependent entropy coders. More specifically, they adopt an asymmetric strategy, where one *dominant* sensor, say sensor $m^\star$, performs independent coding, i.e., independent scalar quantization followed by independent entropy coding. Hereafter, another sensor, say sensor m, performs independent scalar quantization followed by entropy coding with respect to the entropy code of sensor m. With this strategy, if the GW receives both $s_{m^\star}(k)$ and $s_m(k)$, then it is possible to reconstruct $\hat{y}_{m^\star}(k)$ and $\hat{y}_m(k)$. If only $s_{m^\star}(k)$ is received, then the GW can still obtain $\hat{y}_{m^\star}(k)$, but of course not $\hat{y}_m(k)$. However, if only $s_m(k)$ is received, then the GW cannot reconstruct neither $\hat{y}_{m^\star}(k)$ nor $\hat{y}_m(k)$.

[2]Entropy coding can be done by a simple table-lookup since the rounding (quantization) operation directly gives the index of the codeword in the table.

C. Multiple-Description Coding

The idea behind MDC is to create separate descriptions, which are individually capable of reproducing a source to a specified accuracy and, when combined, are able to refine each other [18]. For that purpose, when using MDC, the source vector $y_m(k)$ is mapped to $J_m(k)$ descriptions

$$s^i_m(k), \quad i \in \{1, \ldots, J_m(k)\},$$

which are independently entropy coded and transmitted separately to the GW.

In this work, we will consider MDC based on index-assignments and lattice vector quantization [22], [26], [27]. We will assume that for any $y_m(k)$, the packet-loss probabilities for the $J_m(k)$ descriptions are i.i.d. and equal. Furthermore, we will focus on the symmetric situation where the bit-rates of each description formed at the mth sensor are equal, given by $b_m(k)/J_m(k)$, and where the distortion observed at the GW depends only upon the number of received descriptions and not on which descriptions are received.

D. XOR-based Network Coding

In the second estimation architecture under study relays are used to enhance estimation performance. In this setup, sensor data is sent by using simple independent coding, as described in Section II-A. The relays act as intermediate network nodes and are able to perform simple XOR-based network coding on the data [28]. As illustrated in Fig. 2, the relay nodes are overhearing broadcast communication from the sensors to the GW, and are therefore able to aid the GW with additional information about the sensors' data. In particular, the relays will XOR the incoming data at a bit level, i.e., without decoding [29]. Here one simply zero pads the shortest symbols in order to make them all of equal length [23].

Example 1: Consider the scheme in Fig. 2 and assume that the GW has received either only $s_1(k)$ or only $s_2(k)$. If the relay receives both $s_1(k)$ and $s_2(k)$, then it transmits

$$r_1(k) = s_1(k) \oplus s_2(k) \quad (6)$$

to the GW. If $r_1(k)$ is successfully received, then the GW is able to recover both $s_1(k)$ and $s_2(k)$ and thereby reconstruct both values $\hat{y}_1(k)$ and $\hat{y}_2(k)$ by use of $r(k)$ and its own message $s_1(k)$ or $s_2(k)$, see Table I. □

Data successfully received	Values reconstructed
$s_1(k)$, $s_2(k)$, $r_1(k)$	$\hat{y}_1(k)$, $\hat{y}_2(k)$
$s_1(k)$, $s_2(k)$	$\hat{y}_1(k)$, $\hat{y}_2(k)$
$s_1(k)$, $r_1(k)$	$\hat{y}_1(k)$, $\hat{y}_2(k)$
$s_2(k)$, $r_1(k)$	$\hat{y}_1(k)$, $\hat{y}_2(k)$
$s_1(k)$	$\hat{y}_1(k)$
$s_2(k)$	$\hat{y}_2(k)$
$r_1(k)$	none
none	none

TABLE I

RECONSTRUCTED VALUES AT THE GW WHEN USING THE ESTIMATION ARCHITECTURE IN FIG. 2 WITH NETWORK CODING AS DESCRIBED IN SECTION II-D.

E. Key properties and complexity issues of the proposed coding schemes

1) Independent coding: This is the simplest of the proposed architectures and is furthermore a fundamental part of zero-error coding as well as XOR-based network coding. Since the set of possible bit-rates $\mathcal{B}_m$ for the mth sensor is discrete (but actually not limited to integer valued elements due to entropy coding), the optimization over bit-rates is non-linear and non-convex. Fortunately, the cardinality of $\mathcal{B}_m$ can usually be chosen small in practice, and we therefore simply let the GW perform a brute-force search over all possible candidate bit-rates.

2) Zero-error coding: The advantage of ZEC over independent coding is that one can reduce the rate of any given entropy coder by making it dependent upon another entropy coder without increasing the complexity at the sensor nodes. The complexity at the GW is, however, increased, since the GW has to decide upon whether ZEC should be used or not, see Section VI-A. If the channels from sensors $m^\star$ and m to the GW are both reliable, or if at least one of them is, then it is beneficial to exploit ZEC as is also evident from the simulations in Section VII-A.4. It is worth emphasizing that a reduction of the number of bits to transmit, immediately translates into an energy reduction for a fixed transmission power.

3) Multiple-description coding: When the channels are unreliable and causing packet losses, it is advantageous to use MDC and thereby sent multiple packets, see e.g., Fig. 5, which illustrates the reconstruction accuracy[3] due to using MDC as a function of the channel quality. With the chosen approach to MDC, which is based upon index-assignments [27] (i.e., table-lookups), the complexity at the sensor nodes is not increased over that of independent coding. Moreover, since closed-form solutions for the best choice of MDC parameters and codebooks exist, the complexity at the GW is only slightly increased. The bit-rates of the individual packets are generally smaller than those used for the single packet case. Moreover, the transmission power can often be reduced when using MDC, since it is more likely that at least one small packet out of several packets is received than one particular large packet is received.

4) XOR-based network coding: If there is a relay available (e.g., one of the sensor nodes could act as a relay node), which overhears broadcast messages, then it is advantageous to exploit, e.g., simple XOR-based network coding whenever a subset of the channels are experiencing fading. Since the individual sensors simply perform independent coding, their complexity is not increased. The complexity at the relay is determined by the XOR operations, which can be efficiently executed on most hardware architectures. Due to help of the relay, the individual sensors can reduce their transmission powers, which in turn saves energy.

[3]The accuracy is measured before the GW applies its KF.

III. TRANSMISSION EFFECTS AND POWER ISSUES

We will model transmission effects by introducing the binary stochastic arrival processes

$$\gamma_m^i(k) = \begin{cases} 1 & \text{if } s_m^i(k) \text{ arrives error-free at time } k, \text{ when} \\ & \text{transmitted directly from sensor } m \text{ to the GW,} \\ 0 & \text{otherwise.} \end{cases}$$

Transmission effects when using the estimation architecture with relays and where no MDC or ZEC is used, see Fig. 2, are modeled in a similar manner. Here, we introduce the binary stochastic arrival processes $\zeta_m^\ell = \{\zeta_m^\ell(k)\}_{k\in\mathbb{N}_0}$ and $\tilde{\gamma}_\ell = \{\tilde{\gamma}_\ell(k)\}_{k\in\mathbb{N}_0}$, see Fig. 4 and where

$$\zeta_m^\ell(k) = \begin{cases} 1 & \text{if } s_m^1(k) \text{ arrives error-free at time } k, \text{ when} \\ & \text{transmitted from sensor } m \text{ to the } \ell\text{-th relay,} \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{\gamma}_\ell(k) = \begin{cases} 1 & \text{if } r_\ell(k) \text{ arrives error-free at time } k \text{ at the GW,} \\ 0 & \text{otherwise.} \end{cases}$$

A. Channel Power Gains

In the sequel, we denote by $g_m(k)$ the complex channel gain (at time k) between the sensor m and the GW, by $g_m^\ell(k)$ the channel gain between the m-th sensor and the ℓ-th relay, and by $\tilde{g}_\ell(k)$ the channel gain between the ℓ-th relay and the GW, see Fig. 4 The transmission power used by the radio power amplifier of the m-th sensor is denoted $u_m(k)$, whereas that of the ℓ-th relay is $\mu_\ell(k)$. If we assume that the packet length is equal to the bit-rate, and that the bit errors are independent of one another at a given time k, then the conditional success probabilities

$$\begin{aligned} \lambda_m^i(k) &\triangleq \mathbf{Pr}\{\gamma_m^i(k) = 1 \,|\, u_m(k), g_m(k), b_m(k), J_m(k)\} \\ &= \big(1 - \beta\big(u_m(k)|g_m(k)|^2\big)\big)^{\frac{b_m(k)}{J_m(k)}} \\ \rho_m^\ell(k) &\triangleq \mathbf{Pr}\{\zeta_m^\ell(k) = 1 \,|\, u_m(k), g_m^\ell(k), b_m(k)\} \\ &= \big(1 - \beta\big(u_m(k)|g_m^\ell(k)|^2\big)\big)^{b_m(k)} \\ \tilde{\lambda}_\ell(k) &\triangleq \mathbf{Pr}\{\tilde{\gamma}_\ell(k) = 1 \,|\, \mu_\ell(k), \tilde{g}_\ell(k), \tilde{b}_\ell(k)\} \\ &= \big(1 - \beta\big(\mu_\ell(k)|\tilde{g}_\ell(k)|^2\big)\big)^{\tilde{b}_\ell(k)}, \end{aligned} \tag{7}$$

for $u_m(k), \mu_\ell(k) > 0$ and where $\tilde{b}_\ell(k)$ is the largest packet length received by the ℓ-th relay at time k, see Section II-D. In (7), the function $\beta(\cdot)\colon [0,\infty) \to [0,1]$ is the bit-error rate (BER). It is monotonically decreasing function and depends on the modulation scheme employed.

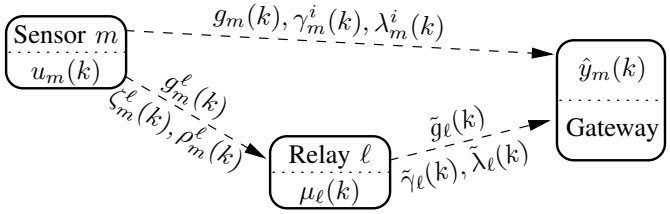

Fig. 4. Transmission power levels, channel gains, transmission outcomes and conditional success probabilities for the estimation architecture with relays and sensors. Only one sensor and relay are shown.

It follows from (7), that one can improve transmission reliability and, thus, state estimation accuracy for given channel gains, by transmitting shorter packets and/or by increasing the power levels used by the transmitters. However, as we have seen in Section II-C, smaller values of packet lengths $b_m(k)$ will lead to larger quantization distortion. Furthermore, the success probabilities are affected by the channel power gains of the different channels. Hence, the statistical properties of the channels will have an impact on the transmit power devised by the controller.

We shall assume a block (flat) fading channel model where the complex gains $g_m(k), g_m^\ell(k), \tilde{g}_\ell(k)$ are all constant over the duration of a packet, and fading between packets. In many wireless sensor network applications the complex channel gains are assumed to be i.i.d. which would make sense if measurements are transmitted rarely as compared to the fading speed. Here we shall, however, adopt a more general model where complex channel gains are correlated. In particular, when the fading channel taps are subject to Rayleigh fading, it is convenient to adopt a first order Markov model [30] of the form[4]

$$\begin{aligned} g_m(k) &= a_{g_m} g_m(k-1) + e_{g_m}(k) \\ g_m^\ell(k) &= a_{g_m^\ell} g_m^\ell(k-1) + e_{g_m^\ell}(k), \\ \tilde{g}_\ell(k) &= a_{\tilde{g}_\ell} \tilde{g}_\ell(k-1) + e_{\tilde{g}_\ell}(k) \end{aligned} \tag{8}$$

where $a_{g_m}, a_{g_m^\ell}, a_{\tilde{g}_\ell}$ determines the amount of correlation and where $e_{g_m}(k), e_{g_m^\ell}(k), e_{\tilde{g}_\ell}(k)$ are mutually independent zero mean circular symmetric complex Gaussian white noises with appropriate covariances. To reduce complexity, the GW discretizes the instantaneous fading gains of these channels into N intervals $[\Gamma_n, \Gamma_{n+1}]$, $n \in \{0, \ldots, N-1\}$, $\Gamma_0 = 0$, $\Gamma_N = \infty$, and adopts a homogeneous finite state Markov chain (FSMC) model with associated states ς_n, $n \in \{0, \ldots, N-1\}$ [33], [34]. The probability that the channel gain switches from state ς_n to state ς_j within a single time step is denoted $p_{n,j}$. We assume that channel states switch only between neighbors, thus, $p_{n,j} = 0$, for all $|n - j| > 1$.

B. Energy Use

When using WSs it is of fundamental importance to save energy. We thus have to find a suitable balance between the transmit power used and the estimation accuracy obtained. The energy used by each sensor $m \in \{1, \ldots, M\}$ to transmit $s_m(k)$ can be quantified via

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

where E_{P} denotes the processing cost, i.e., the energy needed for wake-up, circuitry and sensing, and r is the channel bit-rate.

Due to physical limitations of the radio power amplifiers, the power levels are constrained, for given saturation levels $\{u_m^{\max}\}$, according to:

$$0 \le u_m(k) \le u_m^{\max}, \quad \forall k \in \mathbb{N}_0, \quad \forall m \in \{1, 2, \ldots, M\}. \tag{9}$$

The energy consumption of the relays can be quantified similarly by introducing energy functions $\tilde{E}_\ell(\mu_\ell(k)\tilde{b}_\ell(k))$. For simplicity, we will focus on relays operating in on-off mode, with pre-determined transmission power levels $\{\mu_\ell\}$, thus

$$\mu_\ell(k) \in \{0, \mu_\ell^{\max}\}, \quad \forall \ell \in \{1, 2, \ldots, L\} \tag{10}$$

The relays transmit only if the controller has assigned $\mu_\ell(k) = \mu_\ell^{\max}$ and sensor data which is needed to perform network coding has been successfully received, see Section II-D.

IV. Kalman Filtering with Multiple Intermittent Sensor Links and Coding

State estimation is performed at the GW, which also governs the code-book selection of the sensors, see Figs. 1 and 2. We will assume that the GW knows, whether packets received from the sensors contain errors or not. [5] Thus, at time k, past and present realizations of all transmission processes $\{\gamma_m^i(k-t)\}_{t\in\mathbb{N}_0}$, $i \in \{0, \ldots, J_m(k)-1\}$, $m \in \{1, \ldots, M\}$, and, in case of the estimation architecture with relays, the transmission outcome processes associated to links from relays to the GW, namely, $\{\tilde{\gamma}_\ell(k-t)\}_{t\in\mathbb{N}_0}$, $\ell \in \{1, \ldots, L\}$, are available at the GW to form the state estimate $\hat{x}(k)$.

A. The Reconstruction Processes

To elucidate the situation we introduce the discrete reconstruction processes $\{\theta_m(k)\}_{k\in\mathbb{N}_0}$, via

$$\theta_m(k) = \begin{cases} 1 & \text{if } \hat{y}_m(k) \text{ can be reconstructed at time } k, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly these processes depend on the transmission outcomes, as dictated by the coding schemes employed. More precisely, if the m-th sensor uses only independent coding and no network coding, then $\theta_m(k) = \gamma_m^1(k)$. If, in the estimation architecture of Fig. 1, at time k the sensor m uses MDC, then $\theta_m(k) = 1$, if and only if at least one of the $J_m(k)$ descriptions of $y_m(k)$ is successfully received at the GW and we have:

$$\theta_m(k) = 1 - \prod_{i\in\{1,2,\ldots,J_m(k)\}} \big(1 - \gamma_m^i(k)\big).$$

[4]The use of higher order models, as adopted for example in [31], [32] is straightforward.

[5]This can be handled by the use of a simple cyclic redundancy check.

$\gamma_1^1(k)$	$\gamma_2^1(k)$	$\zeta_1^1(k)$	$\zeta_2^1(k)$	$\tilde{\gamma}_1(k)$	$\theta_1(k)$	$\theta_2(k)$
1	1	1	1	1	1	1
1	1	1	0	0	1	1
1	1	0	1	0	1	1
1	1	0	0	0	1	1
1	0	1	1	1	1	1
0	1	1	1	1	1	1
1	0	1	0	0	1	0
1	0	0	1	0	1	0
1	0	0	0	0	1	0
0	1	1	0	0	0	1
0	1	0	1	0	0	1
0	1	0	0	0	0	1
0	0	1	1	1	0	0
0	0	1	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	0	0

TABLE II

RECONSTRUCTION PROCESSES OF THE ARCHITECTURE IN FIG. 2 WITH NETWORK CODING AS DESCRIBED IN SECTION II-D. NOTE THAT THE RELAY TRANSMITS ONLY IF IT HAS RECEIVED BOTH SYMBOLS $s_1(k)$ AND $s_2(k)$.

If ZEC with dominant coder $m^\star$ is used, then

$$\theta_m(k) = \begin{cases} \gamma_m(k)\gamma_{m^\star}(k) & \text{if } u_m(k)u_{m^\star}(k) > 0, \\ 0 & \text{if } u_m(k)u_{m^\star}(k) = 0, \end{cases} \tag{11}$$

for all $m \in \{1, 2, \ldots, M\}$, see [21, Eq. (17)]. On the other hand, if in the setup with relays, see Fig. 2, network coding is used at time k, then $\theta_m(k)$ also depends upon the transmission outcomes involving the relays, namely, $\tilde{\gamma}_\ell(k)$ and $\zeta_m^\ell(k)$. For example, for the case given in Table I, the processes $\theta_1(k)$ and $\theta_2(k)$ are determined as per Table II.

B. System Model

The GW knows which coding method was used at current time k and also has access to the M reconstruction process realizations. Thus, for state estimation purposes, the overall system amounts to sampling (1)-(2) only at the successful transmission instants of each sensor link. It is convenient to model the overall estimation architectures by introducing the discrete stochastic output matrix process $\{C(k)\}_{k\in\mathbb{N}_0}$ and associated measurements $\{y(k)\}_{k\in\mathbb{N}_0}$ as in:[6]

$$C(k) = \mathcal{C}(\theta(k)) \triangleq \begin{bmatrix} \theta_1(k)C_1 \\ \theta_2(k)C_2 \\ \vdots \\ \theta_M(k)C_M \end{bmatrix}, \; y(k) \triangleq \begin{bmatrix} \theta_1(k)\hat{y}_1(k) \\ \theta_2(k)\hat{y}_2(k) \\ \vdots \\ \theta_M(k)\hat{y}_M(k) \end{bmatrix}, \tag{12}$$

where

$$\theta(k) \triangleq \begin{bmatrix} \theta_1(k) & \theta_2(k) & \ldots & \theta_M(k) \end{bmatrix}.$$

The following time-varying KF gives the best linear estimates of the system state in (1) given the information available at the GW, for both estimation architectures under study:

$$\begin{aligned} \hat{x}(k) &= \hat{x}(k|k-1) + K(k)\big(y(k) - C(k)\hat{x}(k|k-1)\big) \\ \hat{x}(k+1|k) &= A\hat{x}(k|k-1) + AK(k)\big(y(k) - C(k)\hat{x}(k|k-1)\big) \\ P(k+1|k) &= A(I - K(k)C(k))P(k|k-1)A^T + Q \end{aligned} \tag{13}$$

where the gain $K(k)$ and equivalent measurement noise covariance $R(k)$ are given by

$$\begin{aligned} K(k) &\triangleq P(k|k-1)C(k)^T\big(C(k)P(k|k-1)C(k)^T + R(k)\big)^{-1} \\ R(k) &\triangleq \operatorname{diag}\big(R_1 + D_1(k), \ldots, R_M + D_M(k)\big). \end{aligned} \tag{14}$$

Here, $D_m(k)$ are the distortions due to quantization, whereas Q and R_m are the driving noise and measurement noise covariances, respectively. The recursion (13) is initialized with $P(0) = P_0$ and $\hat{x}(-1) = 0$, see (1). In the linear Gaussian case, $P(k+1|k)$ corresponds to the prior covariance of the estimation error. The posterior error covariance matrix is then given by [35]

$$P(k|k) = (I - K(k)C(k))P(k|k-1). \tag{15}$$

Remark 1: Since $C(k)$ is known at the GW, the above Kalman filter uses all successfully reconstructed measurements to form the state estimate $\hat{x}(k)$. Those measurements where $\theta_m(k) = 0$ are not taken into account. This property is reflected in the expression for the filter gain $K(k)$, where the time-varying matrix $C(k)^T$ pre-multiplies the term $\big(C(k)P(k|k-1)C(k)^T + R(k)\big)^{-1}$. Thereby, $\hat{x}(k)$ is not updated based on those values, see (13). □

V. ON-LINE DESIGN OF CODING AND POWER LEVELS

We have seen that transmission power and bit-rate design involves a trade-off between transmission error probabilities (and, thus, state estimation accuracy) and energy use. We will next present a predictive controller which optimizes this trade-off. To keep the sensors simple, the controller is located at the GW. For the first estimation architecture, the controller output contains information on the power levels, and the codebooks to be used by the sensors. In case of the second estimation architecture, the controller also updates the power levels used by the relays.

A. Signaling

To save signal processing energy at the sensors and relays, we would like to limit power control signaling as much as possible. The command signal for each sensor m will contain, in addition to the codebook index $I_m(k)$, a finitely quantized power increment, say $\delta u_m(k)$. Upon reception of $(\delta u_m(k), I_m(k))$, the sensor chooses the codebook $I_m(k)$ (to be used for encoding $y_m(k)$), and reconstructs the power level to be used by its radio power amplifier by simply setting $u_m(k) = u_m(k-1) + \delta u_m(k)$. For the second estimation architecture, see Fig. 2, only one bit is needed to convey each of the relay power levels $\mu_\ell(k)$, see (10).

[6] We note that, if MDC is used, then $\hat{y}_m(k)$ in (12) denotes the reconstruction of $y_m(k)$ based on the $0 \le j \le J_m(k)$ received descriptions. This value differs from $y_m(k)$ due to the measurement noise $v_m(k)$ and the quantization noise.

B. Cost Function

To trade energy consumption for estimation cost, at each instant k, the controller determines the power level increments $\delta u_m(k+1)$ and codebook indices $I_m(k+1)$ and, for the estimation architecture in Fig. 2 also the relay power levels $\mu_\ell(k+1)$, by minimizing the cost function

$$\begin{aligned} &V(S(k+1)) \\ &\quad \triangleq \mathbf{E}\big\{ \operatorname{tr} P(k+1|k+1) \,\big|\, g(k), P(k+1|k), S(k+1)\big\} \\ &\qquad + \varrho V_E(S(k+1)), \end{aligned} \tag{16}$$

where

$$\begin{aligned} S(k+1) = \{&I_m(k+1), \delta u_m(k+1), \mu_\ell(k+1)\} \\ &m \in \{1,2,\dots,M\},\ \ell \in \{1,2,\dots,L\}. \end{aligned} \tag{17}$$

The first term in (16), quantifies the estimation quality. The conditional expectation in (16) is taken for given $P(k+1|k)$, and channel gains $g(k)$, which are available at the GW at time k. Averaging is with respect to the set of possible reconstructions and transmission outcomes due to receiving different subsets of descriptions for each sensor. The probability distribution of these sets depends upon the decision variables, i.e., the bit-rates, coding schemes and power levels.

The second term in $V(S(k+1))$, quantifies the energy use at the next time step.[7] Thus, $\varrho \geq 0$ is a tuning parameter which allows the designer to trade-off estimation accuracy for energy use at the sensors and relays. For the architecture without relays the latter is given by:

$$V_E(S(k+1)) = V_E^S(S(k+1)) \triangleq \sum_{m=1}^{M} E_m(b_m(k+1)u_m(k+1)),$$

whereas, for the estimation architecture with relays,

$$V_E(S(k+1)) = V_E^S(S(k+1)) + \sum_{\ell=1}^{L} \tilde{E}_\ell(\tilde{b}_\ell(k+1)\mu_\ell(k+1)).$$

On-line optimization of the cost function in (16), gives rise to the desired power levels and coding strategy,

$$S(k+1)^{\text{opt}} = \arg \min_{S(k+1)\in\mathbb{S}_{k+1}} V(S(k+1)), \tag{18}$$

where the finite set $\mathbb{S}_{k+1}$ represents the constraints on the decision variables $I_m(k+1)$, $\delta u_m(k+1)$, and $\mu_\ell(k+1)$. In particular, power levels and their increments are finite-set constrained such that (9) and (10) are satisfied. The resulting controller is non-linear, constrained, stochastic and adapts to channel conditions and current estimation quality, i.e., we have

$$S(k+1)^{\text{opt}} = \kappa\big(g(k), P(k+1|k)\big) \tag{19}$$

for some mapping $\kappa(\cdot,\cdot)$. It is worth noting that, in general, due to the cost function being nonlinear and constraints finite, no closed form expression for $\kappa(\cdot,\cdot)$ exists. Instead, the optimization in (19) needs to be performed numerically. To keep computations low, the cost in (16) looks ahead at only one step.[8] In Section VI, we discuss some computational issues.

C. Performance Bound

The effect on packet drops on Kalman filter stability and performance has received considerable attention in the recent literature [36]–[38]. Despite the fundamental importance of power control and coding in wireless communications, it is somewhat surprising that these techniques have received so little attention in this context. In fact, to the best of the authors' knowledge, the only works examining power control for Kalman estimation with packet dropouts are our own [19], [21], [23], [39]. Theorem 1, stated below, establishes sufficient conditions for the expected value of the state estimation error covariance $P(k+1|k)$ to be exponentially bounded. Whilst the main interest of the current work is on state estimation for stable systems, Theorem 1 also applies to unstable systems, extending our recent results documented in [39]–[41] to controller structures such as (19) which are allowed to use the covariance matrices $P(k+1|k)$.

Theorem 1: Consider the stochastic process

$$\eta(\theta) \triangleq \begin{cases} 1 & \text{if } \mathcal{C}(\theta) \text{ is full rank,} \\ 0 & \text{otherwise} \end{cases}$$

and define[9]

$$\nu(P,g) \triangleq \mathbf{Pr}\{\eta(\theta(k+1)) = 0 \,|\, P(k+1|k) = P, g(k) = g\}.$$

Suppose that there exists a uniform bound $\rho \in [0,1)$ such that

$$\nu(P,g) \leq \frac{\rho}{\|A\|^2}, \quad \forall (P,g) \in \mathbb{R}^{n\times n} \times \Omega, \tag{20}$$

where Ω is the support of the (complex) channel gains. Then

$$\mathbf{E}\big\{\|P(k|k-1)\|\big\} \leq \rho^k \operatorname{tr} P_0 + \frac{\varpi c + \operatorname{tr} Q}{1-\rho}(1-\rho^k), \quad \forall k \in \mathbb{N}_0, \tag{21}$$

where

$$\begin{aligned} c &\triangleq \max_{\theta\in\{0,1\}^M:\, \eta(\theta)=1} \|\mathcal{C}^\dagger(\theta)^T \mathcal{C}^\dagger(\theta)\|, \\ \mathcal{C}^\dagger(\theta) &\triangleq (\mathcal{C}(\theta)^T\mathcal{C}(\theta))^{-1}\mathcal{C}(\theta)^T \\ \varpi &\triangleq ||A||^2\Big((\pi e/6)\sigma_{y_m}^2 \sum_{m=1}^{M} 2^{-2\check{b}_m} + \sum_{m=1}^{M} R_m\Big) \end{aligned} \tag{22}$$

and $\check{b}_m \triangleq \min\{\mathcal{B}_m\}$, see (5).

Proof: See appendix. ∎

Theorem 1 establishes a bound on the expected value of $\|P(k|k-1)\|$ which decays with exponential rate ρ. The latter quantity depends on the spectral norm of the system matrix A and the conditional probabilities of the observation matrix $\mathcal{C}(\theta)$ being full-rank. Thus, our result allows one to infer upon estimation accuracy from conditional channel dropout probabilities.

[7] If desired, a term which penalizes the size of the power control signal $\delta u_m(k+1)$ (or of $S(k+1)$), which is transmitted from the gateway to the sensors can be readily included in the cost function.

[8] Multi-step extensions can be easily formulated, following as in [21]

[9] It follows from the analysis in the appendix that this conditional probability is independent of time k.

VI. Computational Issues

In the estimation architecture considered, the sensors do not need to carry out significant computations. In particular, optimizations and Kalman filter recursions are performed by the gateway, see Fig. 1. In relation to Kalman filtering, it is necessary to invert the matrix $C(k)P(k|k-1)C(k)^T + R(k)$ given in (14). This matrix is of size $M \times M$ and therefore scales with the number of sensors. Since it is positive semidefinite and symmetric, its inverse can be obtained with efficient algorithms based upon Cholesky factorizations and having complexity on the order of $\mathcal{O}(M^3)$.

To analyze computational cost of the optimizations in (18), we first note that, from (15), we have $P(k+1|k+1) = \big(I - K(k+1)C(k+1)\big)P(k+1|k)$. Thus, to evaluate the cost function in (16) and find the constrained optimizer $S(k+1)^{\text{opt}}$, the controller first uses $P(k|k-1)$ and $\theta(k)$ to calculate $C(k)$ and $P(k+1|k)$ using (12), (13) and (14). Clearly, $K(k+1)$ and $C(k+1)$ depend upon $P(k+1|k)$, $\theta(k+1)$ and $R(k+1)$. The latter quantity depends upon the decision variable $S(k+1)$. On the other hand, the reconstruction processes at time $k+1$ depend upon the transmission outcomes processes at time $k+1$, as per the mapping induced by the codebook choice; see, e.g., Tables I and II. Thus, for the architecture in Fig. 1, the conditional expectation in (16) can be evaluated by using the conditional probabilities $\mathbf{Pr}\{\gamma_m^i(k+1) = 1 \,|\, g(k), P(k+1|k), S(k+1)\}$ as required. Note that, by using (7) and the law of total probability, one obtains that

$$\begin{aligned}
&\mathbf{Pr}\{\gamma_m^i(k+1) = 1 \,|\, g(k), P(k+1|k), S(k+1)\} \\
&\quad = \mathbf{Pr}\{\gamma_m^i(k+1) = 1 \,|\, g(k), S(k+1)\} \\
&\quad = \mathbf{E}\big\{\mathbf{Pr}\{\gamma_m^i(k+1) = 1 \,|\, g_m(k+1), g(k), S(k+1)\}\big\} \\
&\quad = \mathbf{E}\big\{\mathbf{Pr}\{\gamma_m^i(k+1) = 1 \,|\, g_m(k+1), S(k+1)\}\big\} \\
&\quad = \mathbf{E}\big\{\lambda_m^i(k+1)\big\},
\end{aligned}$$

where expectation is taken with respect the conditional channel gain distribution $\mathbf{Pr}\{g_m(k+1) \,|\, g(k)\}$, see the FSMC model in Section III-A.

A. Two-stage search strategy

As outlined above, at every instant k, the proposed controller first finds the optimal set of power value increments and bit-rates. In the estimation architecture with relays, the sensors simply perform independent coding. In contrast, in the estimation architecture depicted in Fig. 1, each sensor can either do independent coding, ZEC, or MDC. Moreover, in the case of MDC, for any given bit-rate, the controller also needs to decide upon the number of descriptions $J_m(k+1)$ and the level of redundancy between the $J_m(k+1)$ descriptions, see [22].

To develop a simple but efficient method to select the coding scheme for the estimation architecture without relays, we propose a two-stage search strategy. In the first stage, the GW only evaluates whether independent coding or MDC should be used at each sensor. In the second stage, if independent coding is chosen for more than one sensor, then the GW further evaluates whether ZEC should be used across these sensors (or a subset of them). The second stage uses exhaustive search. The first stage can be implemented very efficiently. Here it is important to note that the energy consumption at a given sensor is independent of the number of descriptions chosen. Furthermore, under high-resolution assumptions, it is reasonable to assume that the quantization error at any given sensor does not contain significant information about current and past measurements of any of the sensors.[10] Thus, from the KF point of view, $\hat{y}_m(k)$ amounts to a noisy version of $y_m(k)$; the smaller the variance of this noise, the better the estimate $\hat{x}(k)$. This motivates us to adopt a method where, having chosen the optimal $(\delta_m(k), b_m(k))$ and, thus $u_m(k)$, the controller selects the quantization scheme which results in the minimum expected distortion on $y_m(k)$.[11] Furthermore, given the bit-rate $b_m(k)$ and by considering $\{\lambda_m^i : i = 0, \ldots, J_m(k)\}, \forall J_m(k)$ as weights, the simple method to find the optimal number of descriptions as well as the optimal amount of redundancy between the descriptions proposed in [27] will be used. The following example illustrates the procedure:

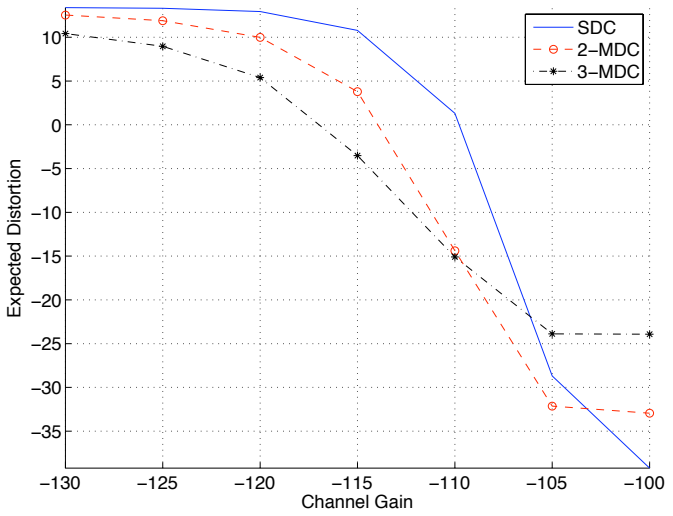

Fig. 5. MDC vs. SDC as a function of channel power gain, $|g_m|^2$. The total bit-rate is fixed at 9 bits/sample; $u_m = 5 \cdot 10^{-5}$.

B. Example of the First Stage Search

We consider a second order linear time-invariant system with transfer function

$$\frac{5.2978(s + 19.46)}{(s^2 + 0.05214s + 33.3)},$$

which upon sampling with a sampling period of 0.1 [s] can be written in the form (1) with

$$A = \begin{bmatrix} 1.6718 & -0.9948 \\ 1 & 0 \end{bmatrix}. \tag{23}$$

The system poles are oscillatory, located at $0.8359 \pm 0.5441\imath$. The driving noise covariance matrix is chosen as $Q = 1/2I$,

[10] Unless one utilizes substractively dithered quantizers, the quantization error will generally not be independent of the input signal. However, the quantization error can be made uncorrelated with the input signal.

[11] Our results in Section VII, use $x(k)$ and rather than $y(k)$ and therefore take the effect of the Kalman filter into account when finding the coding scheme that yields the minimum expected distortion.

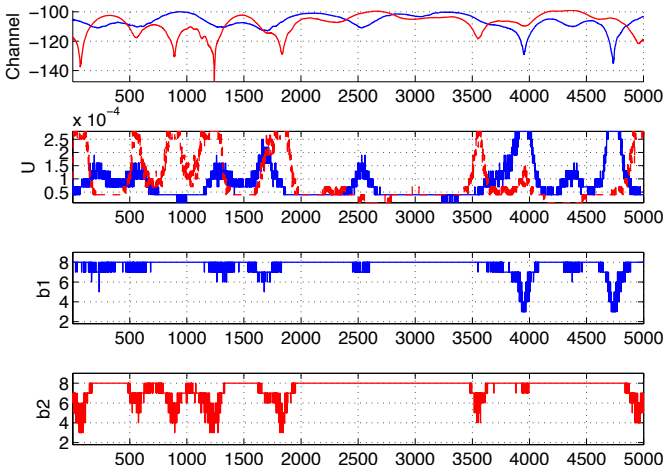

Fig. 6. Channel power gains: $|g_1(k)|^2$ (blue) and $|g_2(k)|^2$ (red); Predictive controller with SDC and constraints as in (24).

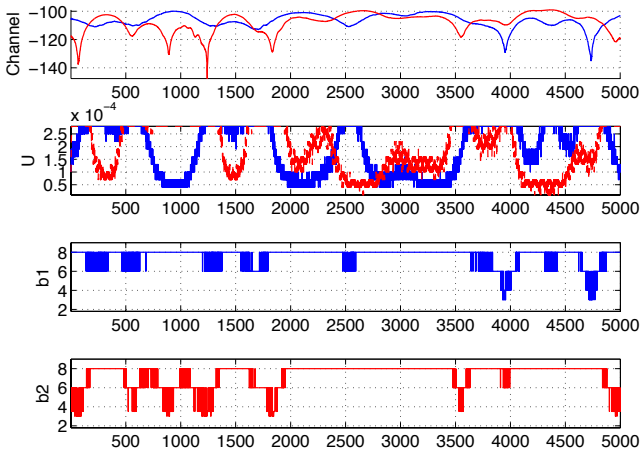

Fig. 7. Simple logic-based scheme.

whereas $P_0 = 0.3I$. We study an estimation architecture having no relays, $M = 2$ WSs with noise variances $R_1 = R_2 = 1/100$. The individual observation matrices are $C_1 = [1\ 0]$, and $C_2 = [0\ 1]$, thus, $c = 1$, see (22).

We analyze the first stage of the two-stage search strategy proposed in Section VI. In other words, we show that the GW can decide up-front whether MDC or independent coding should be used for any given bit-rate. We focus on a single sensor and fix the bit-rate $b_m(k) = 9$ bits/sample and the power level $u_m(k) = 5 \cdot 10^{-5}$. Fig. 5, shows the expected distortion as a function of the channel power gain, $|g_m|^2$ as observed by the GW before applying the KF. In this plot, the expected distortion depends upon the coding scheme.

In the case of independent coding (denoted SDC for "single-description coding"), the encoded description $s_m(k)$ is either received error-less or considered lost with probability $\lambda_m(k)$ and $1 - \lambda_m(k)$, respectively. Thus, the expected distortion is given by $\lambda_m(k)D_m(k) + (1 - \lambda_m(k))\sigma^2_{y_m}(k)$, since, if $s_m(k)$ is lost, we have $y_m(k) - \hat{y}_m(k) = y_m(k)$. Similarly, with MDC the expected distortion is a weighted sum over the distortion due to receiving subsets of descriptions. In Fig. 5, 2-MDC refers to the case of using 2 descriptions at the given sensor and 3-MDC refers to the case with 3 descriptions.

The analysis suggests that when the channel is in a deep fade, it it better to use three descriptions. In the mid-range of channel SNR, it is better to use two descriptions. When the channel is very good, SDC is preferable. (Recall that when MDC is used, the total bit-rate is fixed at $b_m(k) = 9$ bits/sample and evenly split across the descriptions. Hence, in the case of 3 descriptions, each description is encoded at 3 bits/sample. Since the packet length (bit-rate) is reduced, it becomes more probable that a description will be received without errors.)

VII. Results and Discussion

We consider the system described in Section VI-B and use measured channel gain data obtained at the 2.4 GHz ISM band within an office space area at the Signals and Systems group at Uppsala University, Sweden. The top diagram of Fig. 6 illustrates the channel power gains of two realizations, one with horizontal and one with vertical polarization. All our subsequent results use the same channel data, thus facilitating performance comparisons.

A. Estimation Architecture without Relays

We first examine the performance of the controller of Section V for the architecture in Fig. 1.

1) Independent Coding: A simple instance of the controller proposed is where only independent coding (i.e., SDC) is allowed. Power levels and increments are constrained as per

$$0 \leq u_m(k) \leq 3 \times 10^{-4}, \quad \delta u_m(k) \in \{\pm 3 \times 10^{-5}\}. \quad (24)$$

We allow each sensor to use a scalar quantizer at a bit-rate of $b_m(k) \in \{3, \ldots, 8\}$ bits/sample. The second plot in Fig. 6 shows the resulting power levels for the given channel data. Notice that when the channel power gain is dropping, the power level is increased until reaching saturation. Then the bit-rate is decreased to compensate for the increased expenditure; see e.g., $g_2(1800)$ and $g_1(4750)$. The third and forth plots in Fig. 6, illustrate the chosen bit-rates $b_1(k)$ and $b_2(k)$.

2) Simple logic-based controller: The *independent coding* scheme described above constitutes our baseline predictive controller. It is interesting to note that it provides a vast improvement over a simpler algorithm, where the choice of the sensors transmission power and bit-rates are based only on the predicted channel gains. In particular, let us compare the product of the estimated channel gain $\hat{g}_m(k+1)$ and the sensor transmission power $u_m(k)$ (from the previous time instance) to a pre-defined threshold T_u. If it is above the threshold, then the power is decreased by $\delta u_m(k)$ and if it is less than the threshold, then the power is increased by $\delta u_m(k)$. If the choice of power is above its maximum allowable power or below its minimum, it is saturated at its previous level. This simple decision logic, approximately inverts the channel. Let us define the threshold so that when the channel gain is at -110 dB, the transmission power is 0.2 mW, which gives $T_u = 10^{-110/10} \times 2 \times 10^{-4} = 2 \times 10^{-15}$.

When the channel is poor, the probability that a long bit sequence is received error-free is low. On the other hand, in

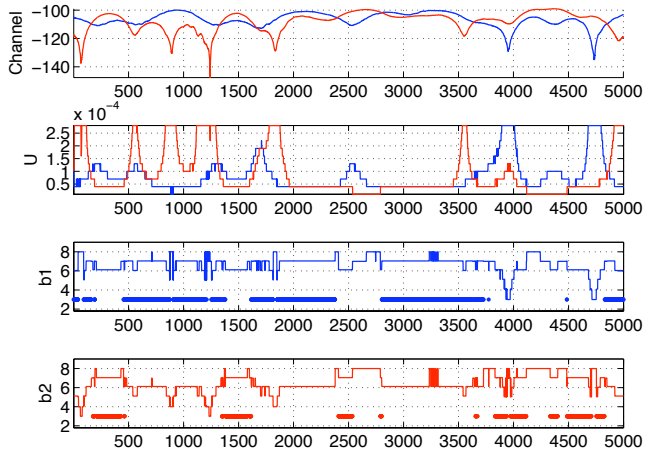

Fig. 8. Predictive controller with SDC and ZEC. The dots indicate that ZEC is used and that the given sensor is dominating.

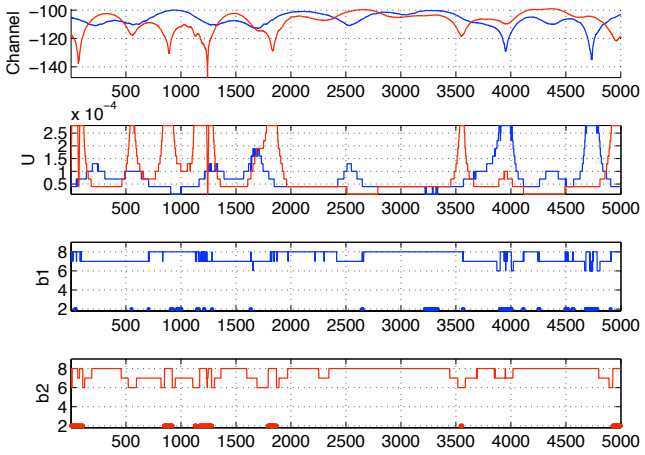

Fig. 9. Predictive controller with SDC and MDC. The dots indicate that MDC is used.

good channel conditions, it makes sense to use longer bit sequences and thereby improve the estimation accuracy of the KF. With this in mind, the channel gain is split into four regions and the bits are allocated using the following rules:

$$b_m(k+1) = \begin{cases} 8, & \text{if } \hat{g}_m(k+1) \geq -110 \text{ dB}, \\ 6, & \text{if } -110 > \hat{g}_m(k+1) \geq -120 \text{ dB}, \\ 4, & \text{if } -120 > \hat{g}_m(k+1) \geq -130 \text{ dB}, \\ 3, & \text{if } \hat{g}_m(k+1) < -130 \text{ dB}. \end{cases} \quad (25)$$

The performance of this simple architecture is presented in Fig. 7. This simple control algorithm, when used on the given channel data, leads to an estimation accuracy of $D = 0.0637$ with a total energy usage of 98.5 nJ. For comparison, the baseline predictive controller with independent coding, uses only 45.5 nJ at the same estimation accuracy. Thus, energy savings above 50% are possible with the predictive controllers proposed in the present work.

3) Robustness towards uncertainties in model dynamics: The proposed controller relies upon knowledge of the dynamics of the underlying process. In particular, the update of the Kalman gain requires knowledge of A, and proper scaling of the quantizers require knowledge of the variances of the processes $y_m, m = 1, \ldots, M$, which depend upon A. To illustrate robustness of the scheme to model uncertainties, let us consider the independent coding scheme described above and assume that the controller uses A given in (23), whereas the actual system dynamics is characterized by

$$\tilde{A} = \begin{bmatrix} 1.68 & -0.99 \\ 1 & 0 \end{bmatrix}.$$

Then, the resulting variances using $\tilde{A}$ are $\sigma^2_{\tilde{y}_1} = 19.95$ and $\sigma^2_{\tilde{y}_2} = 21.93$, which are very close to those obtained using A, i.e., $\sigma^2_{y_1} = 21.48$ and $\sigma^2_{y_2} = 21.93$. Indeed, based on simulations we have measured the average empirical entropies (bit rates) to differ by only 0.02 bits/sample in the two cases. On the other hand, since the Kalman gain is not correctly updated in each iteration due to model mismatch, the state estimator becomes suboptimal and the estimation accuracy is reduced. Indeed, simulations show that the estimation accuracy measured by the average squared error $D = \frac{1}{K}\sum_{k=1}^{K} \|x(k) - \hat{x}(k)\|^2$ is $D = 0.2056$ when A is not perfectly known, as compared to $D = 0.0637$ when A is known. In both cases, the energy expenditure is the same (45.5 nJ).

4) Independent Coding and ZEC: Fig. 8 shows results based on the same system and channel data as in Fig. 6; however, this time we allow the proposed controller to use SDC as well as ZEC, see Section II-B. In the plots showing the bit-rates, we have used dots at the bottom to indicate when ZEC is being used. In particular, a dot indicates that ZEC is being used across the two sensors and that the given sensor is the dominant sensor, i.e., the given sensor is using independent coding whereas the other sensor is using dependent coding. As an example, $|g_1(k)|^2$ is dropping at $k \approx 4000$. The controller decides upon power saturation for that channel and a subsequent decrease in the bit-rate b_1. Here Sensor 2 is dominant, whereas Sensor 1 uses dependent coding.

5) Independent Coding and MDC: We next examine a situation where the controller is allowed to use SDC and MDC with two descriptions. Contrary to the case of ZEC, the sensors can use MDC independently of each other. The total bit-rate is restricted to $b_m(k) \in \{6, 7, 8\}$ bits, so that the side description rates are restricted to $\{3, 3.5, 4\}$ bits. The results are shown in Fig. 9. The dots at the bottom of the bit-rate plots show when MDC is used at the particular sensor. Notice that especially when the channel is weak, e.g., $|g_2(1800)|^2$ and $|g_1(4000)|^2$, MDC is utilized.

6) Independent Coding, ZEC, and MDC: We finally combine all of the above methods and thereby allow the predictive controller to choose SDC, ZEC, as well as MDC. We adopt the two-stage strategy as described in Section VI. Thus, we first evaluate whether SDC or MDC should be used at the different bit-rates. If MDC is chosen, then the controller also finds the optimal trade-off between side and central distortions. If SDC is chosen for both sensors, then the controller applies the second stage involving a brute-force search over all the bit-rates in order to find out whether it is beneficial to use ZEC instead of SDC. The results are shown in Fig. 10, where the dots at the bit-rate level 3 indicate that ZEC is used and that the given sensor is the dominating one; the dots at bit-rate

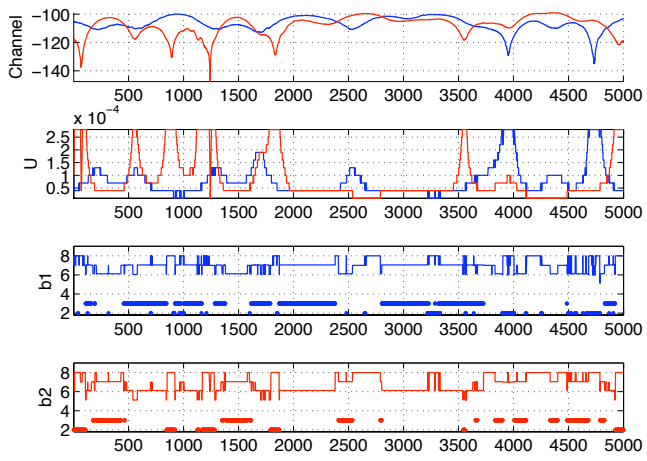

Fig. 10. Predictive controller with independent coding, MDC, and ZEC. The dots at bit-rate level 3 indicate that ZEC is used and that the given sensor is the dominating one. The dots at bit-rate level 2 indicate that MDC is used.

Setup	V	Performance Gain
Indep. coding	0.1256	—
Indep. coding + ZEC	0.1238	1.43%
Indep. coding + MDC	0.1154	8.12%
Indep. coding + ZEC + MDC	0.1135	9.63%

TABLE III

PERFORMANCE OF THE DIFFERENT PREDICTIVE CONTROLLERS FOR CONSTRAINTS AS IN (24).

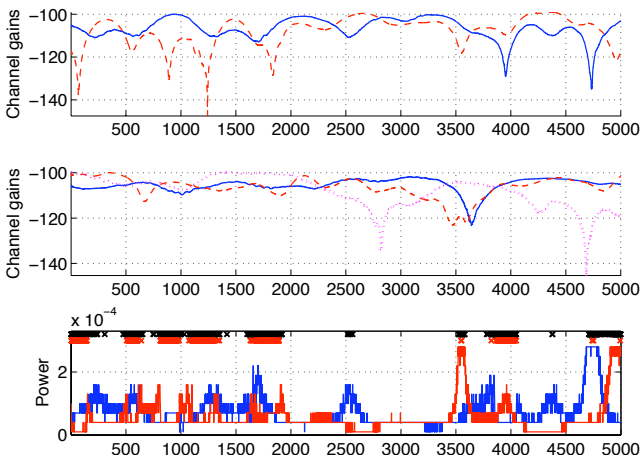

Fig. 11. Predictive controller for the estimation architecture in Fig. 2, with known power channel gains. The GW successfully received $r_1(k)$ in 764 out of 1291 instances where $\mu_1(k) > 0$, thus, the relaying scheme has 59.18% efficiency.

Setup	V	Performance Gain
Indep. coding	0.1391	—
Indep. coding + ZEC	0.1363	2.01%
Indep. coding + MDC	0.1326	4.67%
Indep. coding + ZEC + MDC	0.1300	6.54%

TABLE IV

PERFORMANCE OF THE DIFFERENT PREDICTIVE CONTROLLERS FOR CONSTRAINTS AS IN (27).

level 2 refer to MDC. From Fig. 10 we note, for example, that for Sensor 1 ZEC is used at time 3400 with Sensor 1 being dominant, whereas, at time $k = 4000$, MDC is used for the same sensor.

B. Estimation Accuracy vs. Energy Usage

An underlying theme of our present work is the trade-off between estimation accuracy and energy usage. In particular, the cost function in (16), quantifies this trade-off. To illustrate the potential gains which can be obtained by allowing sensors to perform coding which goes beyond independent coding, Table III shows the empirical average cost,

$$V \triangleq \frac{1}{5000} \sum_{k=0}^{4999} V^\star(k) \tag{26}$$

where $V^\star(k) \triangleq \min_{S(k)} V(S(k))$, for the different control methods examined so far. Notice that a gain of about 1.4% is possible simply by replacing the entropy encoders at the sensors by entropy encoders designed based on the principle of ZEC. It is worth emphasizing that this strategy does not rely upon an increase in the online complexity at the sensors. The offline design of the entropy encoder is of course more complicated than the design of traditional entropy encoders.

With MDC a gain of about 8.1% is possible. Here, it is important to recall that MDC is in our case implemented using a single scalar quantizer followed by a table lookup, which maps the index of the quantized measurement to indices in side codebooks. Thus, the online complexity at the sensors is not increased by this method either. Furthermore, the complexity at the GW is not increased significantly, due to adopting the two-stage search strategy of Section VI. The offline design of the MDC encoders is, however, a lot more complicated than the design of traditional scalar quantizers, see [27] for details.

Finally, if the controller is allowed to choose between independent coding, ZEC, and MDC, then the overall gain is about 9.6%. This shows that the two more advanced coding schemes, ZEC and MDC, complement each other.

If we increase the maximum allowable power level, then it becomes more beneficial to use ZEC and less beneficial to use MDC. This is illustrated in Table IV, where we have shown the results corresponding to the constraints

$$0 \le u_m(k) \le 5 \times 10^{-4}, \quad \delta u_m(k) \in \{\pm 5 \times 10^{-5}\}. \tag{27}$$

C. Estimation Architecture with Relays

We next examine the performance of the predictive controller when used for the estimation architecture in Fig. 2. Here, the controller decides upon power level increments of sensors, the *on-off* state of the relay, and upon bit-rates used by both sensors. The power levels and increments of the sensors are restricted as per (24). The relay performs network coding as in (6) with transmission power $\mu_1^{\max} = 6 \times 10^{-5}$, see (10).

In Fig. 11, the top diagram shows the sensor to GW channel power gains $|g_1(k)|^2$ and $|g_2(k)|^2$ (corresponding to those used in Section VII-A); the middle diagram show the channel power

ϱ	ϕ	$V_E[nJ]$	Relay Channel Models	Reduction of ϕ	System
10^6	0.0707	63.21	–	–	Baseline (no relay)
10^8	0.2021	63.77	Sensor-Relay (predicted), Relay-GW (predicted)	–	Relay always on
655000	0.0341	63.11	Sensor-Relay (known), Relay-GW (known)	51.77%	Relay on/off
680000	0.0382	63.11	Sensor-Relay (predicted), Relay-GW (predicted)	45.87%	Relay on/off
430000	0.0682	63.12	Sensor-Relay (fixed at -100 dB), Relay-GW (predicted)	3.54%	Relay on/off
560000	0.0391	63.11	Sensor-Relay (fixed at -105 dB), Relay-GW (predicted)	44.70%	Relay on/off
1003000	0.0375	63.11	Sensor-Relay (fixed at -110 dB), Relay-GW (predicted)	46.96%	Relay on/off
2080000	0.0429	63.11	Sensor-Relay (fixed at -115 dB), Relay-GW (predicted)	39.32%	Relay on/off

TABLE V

PERFORMANCE GAINS ACHIEVED BY USING THE RELAY AND USING NETWORK CODING GOVERNED BY THE PROPOSED CONTROLLERS.

State	Gain [dB]	$p_{k,k-1}$	$p_{k,k}$	$p_{k,k+1}$
1	-117.77	0.0000	0.9990	0.0010
2	-112.88	0.0010	0.9978	0.0013
3	-110.50	0.0013	0.9973	0.0014
4	-108.83	0.0014	0.9971	0.0015
5	-107.49	0.0015	0.9970	0.0015
6	-106.33	0.0015	0.9970	0.0015
7	-105.30	0.0015	0.9971	0.0014
8	-104.31	0.0014	0.9973	0.0013
9	-103.32	0.0013	0.9976	0.0011
10	-102.29	0.0011	0.9981	0.0008
11	-101.08	0.0008	0.9986	0.0005
12	-99.41	0.0005	0.9995	0.0000

TABLE VI

STATE TRANSITION PROBABILITIES AND THE CHANNEL GAINS THAT THE CONTROLLER USES TO REPRESENT THE STATES.

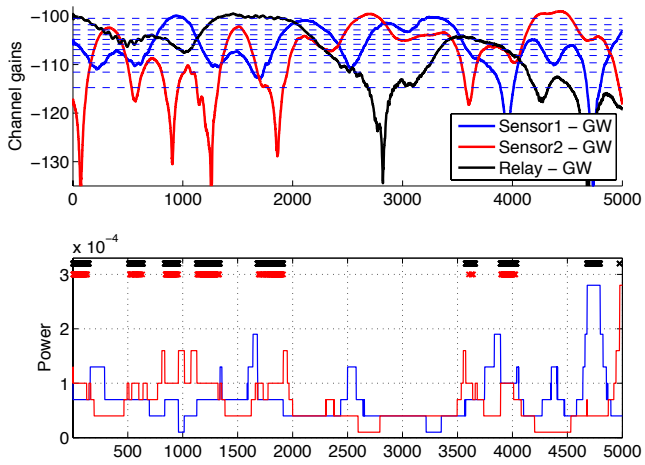

Fig. 12. The three channels that the GW sees are discretized into $K = 12$ intervals as shown by the horizontal dashed lines. The GW successfully received $r_1(k)$ in 495 out of 1196 instances where $\mu_1(k) > 0$, thus, giving 41.39% efficiency.

gains from the sensors to the relay ($|g_1^1(k)|^2$: blue solid line, $|g_2^1(k)|^2$: red dashed line) as well as the channel power gain from the relay to the GW (dotted line); the bottom diagram illustrates the chosen power levels of the two sensors. The black crosses indicate the time slots where the controller has decided to turn on the relay. The red crosses, on the other hand, indicate when the relay operation was successful, i.e., when the relay received $s_1(k)$ and $s_2(k)$ without errors and also successfully transmitted $r_1(k)$, see (6) to the GW. It is clear from Fig. 11, that the controller trades off energy spent on the sensors for energy spent on the relay. Only at the deepest drops in $|g_1(k)|$ and $|g_2(k)|$ (occurring after $k = 3500$) the controller chooses to saturate the sensor power levels. Note that it is beneficial to rely on the relay and network coding most of the time.

To compare different scenarios, we next fix the total energy used by the sensors and relay by adjusting the weighting term ϱ in (16), see Table V. Therefore, the controller seeks to distribute the available energy between the sensor nodes and the relay to minimize the state estimation error variance. As a performance measure, we adopt the empirical value

$$\phi \triangleq \frac{1}{5000} \sum_{k=0}^{4999} \mathrm{tr}\, P(k+1|k+1).$$

Our baseline system uses SDC and no relay; sensor power levels are governed by the predictive controller. In a second configuration, a relay which is always *on* is used. Since the relay uses most of the available energy leaving very little for the sensors to spend, the performance is significantly worse than that of the baseline system, see Table V.

Significantly better performance can be obtained if the controller decides whether the relay shall be *on* or *off*. In this case Table V indicates a performance gain in reduction of ϕ of almost 52%, provided the GW has exact knowledge of the sensor-relay channel gains $g_1^1(k)$ and $g_2^1(k)$. The performance gain is almost 46% if the GW uses simple channel power gain predictions. Table V also illustrates results of situations where the relay uses only a constant estimate for the power gains of the sensor-relay channels. Here, we conclude that it is safer to underestimate the sensor-relay channel power gains than to overestimate them.

D. Estimation Architecture with Relays and Markov Channel Models

In the previous study we assumed that the GW was able to obtain noisy predictions of the future channel gains between the sensors and the GW as well as the relay and the GW. In the following, the GW models the instantaneous fading gains of these three channels by FSMCs, where the possible fading gains are discretized into $K = 12$ intervals (states), see Section III for details. Using the approach described in [33], the resulting state transition probabilities are shown in Table VI. The intervals Γ_k associated with the states are

ϱ	ϕ	$V_E[nJ]$	Relay Channel Models	Reduction of ϕ	System
$1.12 \cdot 10^6$	0.0658	63.08	–	–	Baseline (no relay)
$1.07 \cdot 10^6$	0.0366	63.00	Sensor-Relay (fixed at -110 dB), Relay-GW (modeled)	44.38%	Relay on/off

TABLE VII

PERFORMANCE GAINS ACHIEVED BY USING THE RELAY AND USING NETWORK CODING GOVERNED BY THE PROPOSED CONTROLLERS AND BY MODELING THE CHANNELS AS FIRST-ORDER MARKOVIAN.

illustrated by horizontal dashed lines in Fig. 12 and the corresponding state channel gains are shown in Table VI.

The baseline system has two sensors but no relay. The proposed system has access to one relay. The channels between the two sensors and the relays are shown in the middle plot of Fig. 11. These channels are unknown to the GW and simply modeled by a fixed channel gain of -110 dB. The three other channels, which are connected to the GW and shown in Fig. 12, are modeled by the FSMC approach outlined above. The resulting performance levels are shown in Table VII and are, to some extent, comparable to those obtained when using predictions, see Table V. This observation strengthens the case of using the Markov model in practice.

VIII. CONCLUSIONS

We have studied state estimation with wireless sensors over correlated fading channels. Our work shows that performance gains can be obtained by the use of different coding schemes, when governed by a predictive controller which also determines power levels. Through use of a stochastic Lyapunov function argument we have established sufficient conditions for exponential boundedness of the covariance of the resulting state estimation error. Numerical results revealed that energy savings of more than 50% were possible, when compared to an alternative algorithm, wherein power levels and bit-rates are determined by simple logic which solely depends upon channel power gains and not the estimation error covariance. It is worth noting that the coding schemes examined in the present work do not require significant additional on-line complexity, when compared to direct quantization of the measurements. It is also apparent that the use of relays with simple network coding has the potential to give notable estimation performance gains, with essentially no additional on-line complexity at the sensor and relay nodes. Future work may include the study of more general network topologies and also distributed estimation architectures where individual nodes have additional processing capabilities; see, e.g., [42].

APPENDIX
PROOF OF THEOREM 1

We shall consider the more general system when the system matrix A is unstable, and then evaluate the expressions for stable A. To proceed, we adopt a stochastic Lyapunov function approach, as presented, e.g., in [43], [44] and first prepare the following result:

Lemma 1: The process $\{Z(k)\}_{k\in\mathbb{N}_0}$, where $Z(k) = \big(P(k|k-1), g(k-1)\big)$, is a Markov chain.

Proof: With the model in Section III, $\{g(k)\}_{k\in\mathbb{N}_0}$ are Markovian and

$$\begin{aligned}&\mathbf{Pr}\{g(k)\,|\,Z(k), Z(k-1), Z(k-2), \dots\}\\ &\quad = \mathbf{Pr}\{g(k)\,|\,g(k-1)\} = \mathbf{Pr}\{g(k)\,|\,Z(k)\}.\end{aligned} \tag{28}$$

On the other hand, when using the controller of Section V, the power levels, bit-rates and coding method used at time k depend only on $P(k|k-1)$ and $g(k-1)$ (and deterministic quantities). Thus, (7) and (14) give that the distribution of the term $K(k)C(k)$ used in (13) satisfies $\mathbf{Pr}\{K(k)C(k)\,|\,Z(k), Z(k-1), \dots\} = \mathbf{Pr}\{K(k)C(k)\,|\,Z(k)\}$, which implies that

$$\begin{aligned}&\mathbf{Pr}\{P(k+1|k)\,|\,Z(k), Z(k-1), Z(k-2), \dots\}\\ &\quad = \mathbf{Pr}\{P(k+1|k)\,|\,Z(k)\}.\end{aligned}$$

Use of (28) shows $\mathbf{Pr}\{Z(k+1)\,|\,Z(k), Z(k-1), Z(k-2), \dots\} = \mathbf{Pr}\{Z(k+1)\,|\,Z(k)\}$. ∎

Having established that $\{Z(k)\}_{k\in\mathbb{N}_0}$ is Markovian, we now adopt the procedure used to prove Theorem 1 in [41] and introduce $V_k \triangleq \operatorname{tr} P(k|k-1) \geq 0$.

Lemma 2: Consider $\nu(P,g)$ and ϖ and c defined in (22). Then,

$$\begin{aligned}\mathbf{E}\big\{V_1\,|\,Z(0) = (P,g)\big\} &\leq \operatorname{tr} Q + (1-\nu(P,g))\varpi c\\ &\quad + \nu(P,g)\|A\|^2 \operatorname{tr} P, \quad \forall (P,g) \in \mathbb{R}^{n\times n} \times \Omega.\end{aligned}$$

Proof: We use the total probability formula to write:

$$\begin{aligned}&\mathbf{E}\big\{V_1\,|\,Z(0) = (P,g)\big\}\\ &= \mathbf{E}\big\{V_1\,|\,Z(0) = (P,g), \eta(0) = 0\big\}\nu(P,g)\\ &\quad + \mathbf{E}\big\{V_1\,|\,Z(0) = (P,g), \eta(0) = 1\big\}(1-\nu(P,g))\end{aligned} \tag{29}$$

Following as in the proof of [41, Lemma 2], for $\eta(0) = 0$ we consider the worst case, where $\theta_m(0) = 0$, for all $m \in \{1, \dots, M\}$. This gives:

$$\mathbf{E}\big\{V_1\,|\,Z(0) = (P,g), \eta(0) = 0\big\} \leq \|A\|^2 \operatorname{tr} P + \operatorname{tr} Q. \tag{30}$$

To study the case where $\eta(0) = 1$, we consider the simple state predictor $\bar{x}(k+1) = AC^\dagger(k)y(k)$, where $C^\dagger(k) \triangleq (C(k)^T C(k))^{-1} C(k)^T$ is the pseudo-inverse of $C(k)$. This estimator yields the estimation error $x(k+1) - \bar{x}(k+1) = w(k) + AC^\dagger(k)v(k)$, thus,

$$\begin{aligned}\operatorname{tr} \bar{P}(k+1|k) &\triangleq \operatorname{tr} \mathbf{E}\{(x(k+1) - \bar{x}(k+1))\\ &\qquad \times (x(k+1) - \bar{x}(k+1))^T\}\\ &= \operatorname{tr} Q + \operatorname{tr}\big(AC^\dagger(k)R(k)(C^\dagger(k))^T A^T\big)\\ &\leq \operatorname{tr} Q + ||A||^2 \operatorname{tr}\big(C^\dagger(k)R(k)(C^\dagger(k))^T\big)\\ &\leq \operatorname{tr} Q + ||A||^2 \|(C^\dagger(k))^T C^\dagger(k)\| \operatorname{tr} R(k) \leq \varpi c + \operatorname{tr} Q,\end{aligned}$$

where we have used [45, Fact 5.12.7]. Since the Kalman filter gives the minimum conditional estimation error among all linear estimators we obtain that

$$\mathbf{E}\big\{V_1 \,|\, Z(0) = (P, g), \eta(0) = 1\big\} \leq \varpi c + \operatorname{tr} Q \tag{31}$$

The result follows by substitution of (31) and (30) into (29). ■

To prove Theorem 1, we use V_k as a candidate Lyapunov function. Lemma 2 and (20) give that

$$\begin{aligned} 0 \leq \mathbf{E}\big\{V_1 \,|\, Z(0) = (P, g)\big\} \leq \operatorname{tr} Q + (1 - \nu(P, g))\varpi c \\ + \nu(P, g)\|A\|^2 \operatorname{tr} P \leq \nu(P, g)\|A\|^2 V_0 + \bar{\beta} \leq \rho V_0 + \bar{\beta}, \end{aligned}$$

for all (P, g), and where $\bar{\beta} \triangleq \operatorname{tr} Q + (1 - \nu(P, g))\varpi c \leq \operatorname{tr} Q + \varpi c \in [0, \infty)$. Since $\{Z(k)\}_{k\in\mathbb{N}_0}$ is Markovian, we can use [44, Prop. 3.2] (see also [41]) to conclude that

$$0 \leq \mathbf{E}\big\{V_k \,|\, Z(0) = Z\big\} \leq \rho^k V_0 + \bar{\beta} \sum_{i=0}^{k-1} \rho^i = \rho^k V_0 + \bar{\beta} \frac{1 - \rho^k}{1 - \rho}, \tag{32}$$

for all $k \in \mathbb{N}_0$. Since $P(k|k-1) \succeq 0$, it holds that $V_k \geq \|P(k|k-1)\|$, for all $k \in \mathbb{N}_0$. Thus, upon noting that $P(0|-1) = P_0$ is given, (32) establishes (21). □

Daniel E. Quevedo (S'97–M'05) received Ingeniero Civil Electrónico and Magister en Ingeniería Electrónica degrees from the Universidad Técnica Federico Santa María, Valparaíso, Chile in 2000. In 2005, he received the Ph.D. degree from The University of Newcastle, Australia, where he is currently an Associate Professor. He has been a visiting researcher at various institutions, including Uppsala University, Sweden, KTH Stockholm, Sweden, Aalborg University, Denmark, Kyoto University, Japan, INRIA Grenoble, France, University of Notre Dame, USA, and The Hong Kong University of Science and Technology.

Dr. Quevedo was supported by a full scholarship from the alumni association during his time at the Universidad Técnica Federico Santa María and received several university-wide prizes upon graduating. He received the IEEE Conference on Decision and Control Best Student Paper Award in 2003 and was also a finalist in 2002. In 2009, he was awarded a five-year Australian Research Fellowship. His research interests include several areas of automatic control, signal processing, and power electronics.

Jan Østergaard (S'98-M'99-SM'11) received the M.Sc.E.E. from Aalborg University, Aalborg, Denmark, in 1999 and the PhD degree (cum laude) from Delft University of Technology, Delft, The Netherlands, in 2007. From 1999 to 2002, he worked as an R&D Engineer at ETI A/S, Aalborg, Denmark, and from 2002 to 2003, he worked as an R&D Engineer at ETI Inc., Virginia, United States. Between September 2007 and June 2008, he worked as a post-doctoral researcher at The University of Newcastle, NSW, Australia. From June 2008 to March 2011, he worked as a post-doctoral researcher/Assistant Professor at Aalborg University. Since 2011 he has been an Associate Professor at Aalborg University.

He has been a visiting researcher at Tel Aviv University, Tel Aviv, Israel, and at Universidad Técnica Federico Santa María, Valparaíso, Chile. He has received a Danish Independent Research Council's Young Researcher's Award and a post-doctoral fellowship from the Danish Research Council for Technology and Production Sciences. Dr. Østergaard is currently an associate editor for EURASIP Journal on Advances in Signal Processing.

Anders Ahlén (S'80–M'84–SM'90) is full professor and holds the chair in Signal Processing at Uppsala University where he is also the head of the Signals and Systems Division of The Department of Engineering Sciences. He was born in Kalmar, Sweden, and received the PhD degree in Automatic Control from Uppsala University. He was with the Systems and Control Group, Uppsala University from 1984-1992 as an Assistant and Associate Professor in Automatic Control. During 1991 he was a visiting researcher at the Department of Electrical and Computer Engineering, The University of Newcastle, Australia. He was a visiting professor at the same university in 2008. In 1992 he was appointed Associate Professor of Signal Processing at Uppsala University. During 2001-2004 he was the CEO of Dirac Research AB, a company offering state-of-the-art audio signal processing solutions. He is currently the chairman of the board of the same company. Since 2007 he has been a member of the Uppsala VINN Excellence Center for Wireless Sensor Networks, WISENET. His research interests, which include Signal Processing, Communications and Control, are currently focused on Signal Processing for Wireless Communications, Wireless Sensor Networks, Wireless Control, and Audio Signal Processing.

From 1998 to 2004 he was the Editor of Signal and Modulation Design for the IEEE Transactions on Communications.

REFERENCES

[1] M. Ilyas, I. Mahgoub, and L. Kelly, *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems*. Boca Raton, FL, USA: CRC-Press, Inc, 2004.

[2] H. Gharavi and S. P. Kumar, "Special section on sensor networks and applications," *Proc. IEEE*, vol. 91, pp. 1151–1152, Aug. 2003.

[3] A. Willig, "Recent and emerging topics in wireless industrial communications: A selection," *IEEE Trans. Ind. Inf.*, vol. 4, pp. 102–124, May 2008.

[4] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[5] S. V. Hanly and D.-N. Tse, "Power control and capacity of spread spectrum wireless networks," *Automatica*, vol. 35, pp. 1987–2012, 1999.

[6] F. Gunnarsson and F. Gustafsson, "Control theory aspects of power control in UMTS," *Contr. Eng. Pract.*, vol. 11, pp. 1113–1125, 2003.

[7] M. Johansson, E. Björnemo, and A. Ahlén, "Fixed link margins outperform power control in energy-limited wireless sensor networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 3, (Honolulu, HI), pp. 513–516, 2007.

[8] N. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice Hall, 1984.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley–Interscience, second ed., 2006.

[10] W. Qi, J. Liu, X. Chen, and P. D. Christofides, "Supervisor predictive control of stand-alone wind-solar energy generation systems," *IEEE Trans. Contr. Syst. Technol.*, vol. 19, pp. 199–207, Jan. 2011.

[11] C. Vermillion, J. Sun, and K. Butts, "Predictive control allocation for a thermal management system based on an inner loop reference model—design, analysis, and experimental results," *IEEE Trans. Contr. Syst. Technol.*, vol. 19, pp. 772–781, July 2011.

[12] A. Alessandri, M. Gaggero, and F. Tonelli, "Min-max and predictive control for the management of distribution in supply chains," *IEEE Trans. Contr. Syst. Technol.*, vol. 19, pp. 1075–1089, Sept. 2011.

[13] S. Di Cairan, D. Yanakiev, A. Bemporad, I. Kolmanovsky, and D. Hrovat, "Model predictive idle speed control: Design, analysis, and experimental evaluation," *IEEE Trans. Contr. Syst. Technol.*, vol. 20, pp. 84–97, Jan. 2012.

[14] D. E. Quevedo, R. P. Aguilera, M. A. Pérez, P. Cortés, and R. Lizana, "Model predictive control of an AFE rectifier with dynamic references," *IEEE Trans. Power Electron.*, vol. 27, pp. 3128–3136, July 2012.

[15] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, July 1973.

[16] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, pp. 1–10, Jan. 1976.

[17] P. Koulgi, E. Tuncel, S. L. Regunathan, and K. Rose, "On zero-error coding of correlated sources," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2856–2873, Nov. 2003.

[18] A. A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. 28, pp. 851–857, Nov. 1982.

[19] 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.

[20] 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, Berlin Heidelberg: Springer-Verlag, 2009.

[21] D. E. Quevedo, A. Ahlén, and J. Østergaard, "Energy efficient state estimation with wireless sensors through the use of predictive power control and coding," *IEEE Trans. Signal Processing*, vol. 58, pp. 4811–4823, Sept. 2010.

[22] J. Østergaard, D. E. Quevedo, and A. Ahlén, "Predictive power control and multiple-description coding for wireless sensor networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, (Taipei), 2009.

[23] J. Østergaard, D. E. Quevedo, and A. Ahlén, "Predictive power control for dynamic state estimation over wireless sensor networks with relays," in *Proc. 2010 Eur. Sig. Proc. Conf.*, (Denmark), 2010.

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

[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), p. 207, 2000.

[26] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 39, pp. 821–834, May 1993.

[27] J. Østergaard, J. Jensen, and R. Heusdens, "n-channel entropy-constrained multiple-description lattice vector quantization," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1956–1973, May 2006.

[28] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1204–1216, 2000.

[29] C. Fragouli, J.-Y. Le Boudec, and J. Widmer, "Network coding: an instant primer," *SIGCOMM Comput. Commun. Rev.*, vol. 36, pp. 63–68, Jan. 2006.

[30] H. S. Wang and P.-C. Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," *IEEE Trans. Veh. Technol.*, vol. 45, pp. 353–357, May 1996.

[31] L. Lindbom, A. Ahlén, M. Sternad, and M. Falkenström, "Tracking of time-varying mobile radio channels–part II: A case study," *IEEE Trans. Commun.*, vol. 50, pp. 156–167, Jan. 2002.

[32] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1650–1662, July 2005.

[33] H. S. Wang and N. Moayeri, "Finite state Markov channel – a useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 163–171, Feb. 1995.

[34] A. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state markov channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 868–886, May 1996.

[35] B. D. O. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall, 1979.

[36] 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, pp. 1453–1464, Sept. 2004.

[37] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proc. IEEE Conf. Decis. Contr.*, (Paradise Island, Bahamas), pp. 4180–4186, 2004.

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

[39] D. E. Quevedo, A. Ahlén, A. S. Leong, and S. Dey, "On Kalman filtering over fading wireless channels with controlled transmission powers," *Automatica*, vol. 48, pp. 1306–1316, July 2012.

[40] D. E. Quevedo, A. Ahlén, and K. H. Johansson, "Stability of state estimation over sensor networks with Markovian fading channels," in *Proc. IFAC World Congr.*, 2011.

[41] D. E. Quevedo, A. Ahlén, and K. H. Johansson, "State estimation over sensor networks with correlated wireless fading channels," *IEEE Trans. Automat. Contr.*, vol. 58, pp. 581–593, Mar. 2013.

[42] A. Chiuso and L. Schenato, "Information fusion strategies and performance bounds in packet-drop networks," *Automatica*, vol. 47, pp. 1304–1316, July 2011.

[43] H. Kushner, *Introduction to Stochastic Control*. New York, N.Y.: Holt, Rinehart and Winston, Inc., 1971.

[44] S. P. Meyn, "Ergodic theorems for discrete time stochastic systems using a stochastic Lyapunov function," *SIAM Journal on Control and Optimization*, vol. 27, pp. 1409–1439, Nov. 1989.

[45] D. S. Bernstein, *Matrix Mathematics*. Princeton, N.J.: Princeton University Press, 2nd ed., 2009.