PREDICTIVE POWER CONTROL AND MULTIPLE-DESCRIPTION CODING FOR WIRELESS SENSOR NETWORKS

Jan Østergaard [1], Daniel E. Quevedo [2], and Anders Ahlén [3]

[1]Department of Electronic Systems, Aalborg University, Aalborg, Denmark
[2]School of Electrical Engineering and Computer Science, The University of Newcastle, Australia
[3] Department of Engineering Sciences, Signals and Systems, Uppsala University, Uppsala, Sweden
janoe@ieee.org, dquevedo@ieee.org, anders.ahlen@signal.uu.se

ABSTRACT

We study state estimation via wireless sensor networks over fading channels affected by random packet loss. In the configuration examined, the sensors send their measurements to a single gateway, which decides upon the source coding scheme and the sensor transmitter power levels. The decision process is carried out on-line and adapts to changing channel conditions to achieve an optimal trade-off between estimation quality and sensor energy expenditure. In particular, if some channel conditions are poor, then the gateway commands the corresponding sensors to increase power levels and use multiple-description coding. Simulations based on measured channel data illustrate that the proposed scheme gives excellent results.

Index Terms— wireless sensor networks, power control, multiple-description coding

1. INTRODUCTION

The use of wireless sensor networks (WSNs) has attracted significant interest; see, e.g., [1–3]. The driving force behind this evolution from wired to wireless is the low deployment cost: There is no need for extensive wiring, either in new installations or for upgrading old systems. Furthermore, wireless sensors can be placed where wires cannot go, or where power sockets are not available.

A drawback of wireless channels lies in that they are subject to fading and interference, which frequently causes packet errors. The time-variability of the fading channel can be alleviated by adjusting the power levels and the transmitted packet lengths. 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 in WSNs: in most applications sensor nodes are expected to be operational for several years without the replacement of batteries; see, e.g., [4, 5]. In addition, sending short packets may imply large quantization effects unless careful coding is used.

In the present work, we 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 $x(0) \in \mathcal{N}(0, P_0)$, and P_0 is $n_x \times n_x$. In (1), $w = \{w(k)\}_{k \in \mathbb{N}_0}$ is i.i.d., where each $w(k) \in \mathcal{N}(0, Q)$.

To remotely estimate the state sequence $x = \{x(k)\}_{k \in \mathbb{N}_0}$, a network of M sensors is used. Each sensor m provides a noisy measurement signal, say $y_m = \{y_m(k)\}_{k \in \mathbb{N}_0}$:

$$y_m(k) = C_m x(k) + v_m(k), \quad m \in \{1, 2, \dots, M\}, \quad (2)$$
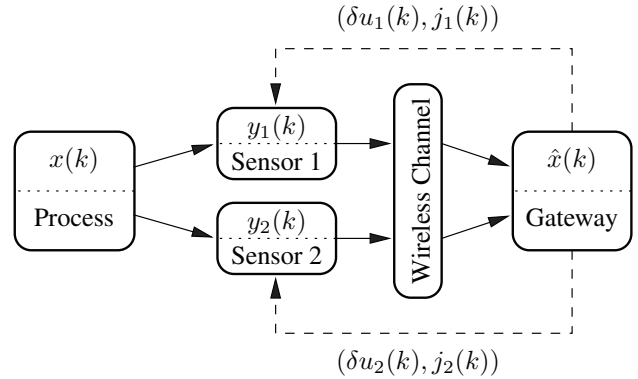
Fig. 1. State Estimation with a WSN having $M = 2$ sensors.

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

The values in (2) are coded and transmitted at an appropriate power level over a fading channel (generating random packet loss) to a single gateway. Received packets are then used to estimate $x(k)$ by means of a time-varying Kalman Filter (KF) which takes into account packet loss. To keep the sensors simple and energy efficient, sensor nodes are not allowed to communicate with each other. In particular, coding is carried out independently. To achieve robustness in the presence of packet loss, the sensors are allowed to use multiple-description coding (MDC) at a set of predefined bit-rates.

Within the above context, the main contribution of the present paper lies in developing a centralized dynamic controller. It is located at the gateway and decides upon the transmission power level and coding scheme to be used by each sensor. The controller uses elements of nonlinear (stochastic) predictive control [6, 7] and trades sensor battery use for estimation accuracy. To further conserve energies, a pre-designed set of codebooks is stored at each sensor, and the controller sends only codebook indices and coarsely quantized power increments to the sensor nodes, whenever deemed necessary. Fig. 1 depicts the overall configuration of the system under study.

The present work extends our recent paper [8] where only power control was examined.

2. PRELIMINARIES

In this section, we present background material on the components of the scheme in Fig. 1, namely, the coding schemes at the sensors, the wireless channels, and the state estimator used at the gateway.

2.1. Multiple-Description Coding

The idea behind MDC is to create separate descriptions, individually capable of reproducing a source to a specified accuracy and, when combined, being able to refine each other. In this work, we will consider MDC based on index-assignments (IAs) and lattice vector quantization [11].

In IA-based MDC, the source vector $y_m(k)$ is first quantized using a central quantizer $\mathcal{Q}^c$ resulting in the central reconstruction $\hat{y}_m^c(k)$ yielding a distortion $D_m^c = \mathbb{E}\|y_m(k) - \hat{y}_m^c(k)\|^2$ given by

$$D_m^c \approx G(\Lambda_c) 2^{2(h(y_m(k)) - b_m^c(k))}, \tag{3}$$

where $G(\Lambda_c)$ is the dimensionless normalized second moment of the lattice Λ_c, $h(y_m(k))$ is the differential entropy of the random vector $y_m(k)$ and $b_m^c(k)$ is the bit-rate of the central quantizer at time k, see [12] for details.[1] We note that (3) is simply the distortion of a single-description (SD) lattice quantizer operating at $b_m^c(k)$ bits. The central reconstruction $\hat{y}_m^c(k)$ is mapped to $J_m(k)$ descriptions by an invertible IA function α, i.e., $\alpha(\hat{y}_m^c(k)) \mapsto (\hat{y}_m^0(k), \cdots, \hat{y}_m^{J_m(k)-1}(k))$. These descriptions are independently entropy coded and transmitted separately to the gateway. We will assume that the packet-loss probabilities $p_m^i(k)$, $i = 0, \ldots, J_m(k) - 1$, for the $J_m(k)$ descriptions, are i.i.d. with probabilities $p_m^i(k) = p_m(k)$. Furthermore, we also restrict attention to the symmetric situation where the bit-rate $b_m^i(k)$ of the ith description formed at the mth sensor is the same for all i (i.e., we have $b_m^i(k) = b_m(k)/J_m(k)$) and the distortion observed at the gateway depends only upon the number of received descriptions and not on which descriptions are received. In particular, it was shown in [12], that the expected distortion $D_m^{(j)}(k)$ due to reconstruction using any subset of $1 \leq j < J_m(k)$ descriptions is approximately given by

$$D_m^{(j)}(k) \approx \frac{J_m(k) - j}{2j J_m(k)} G(S)\psi^2 \tag{4}$$
$$\times 2^{\frac{2J_m(k)}{J_m(k)-1}(b_m^c(k) - b_m(k)/J_m(k))} 2^{2(h(y_m(k)) - b_m^c(k))}$$

where $G(S)$ is the dimensionless normalized second moment of a hypersphere and ψ is a function of the vector dimension and number of descriptions, see [12] for details. If all descriptions are received, the distortion is given by (3). In the scalar case, we have $G(S) = G(\Lambda_c) = 1/12$ and if $J_m(k) = 2$, we have $\psi = 1$. Thus, in this case, the side distortion $D_m^s(k) = D_m^{(1)}(k)$ is given by

$$D_m^s(k) \approx \frac{1}{4 \cdot 12} 2^{2h(y_m(k))} 2^{-2(b_m(k) - b_m^c(k))}. \tag{5}$$

For a fixed bit-rate $b_m(k)$, the side description bit-rate is fixed at $b_m^i(k) = b_m(k)/J_m(k)$. However, the central bit-rate $b_m^c(k)$ can be chosen within the range $b_m(k)/J_m(k) < b_m^c(k) < b_m(k)$. A large central rate $b_m^c(k)$ leads to a small central distortion $D_m^c(k)$ and a large side distortion $D_m^s(k)$, and vice-versa, see (3) and (4). Thus, there exists an MDC design trade-off, which is governed by the ratio $1 < \phi \triangleq \mathcal{V}_s/\mathcal{V}_c$, where $\mathcal{V}_s$ and $\mathcal{V}_c$ denote the volume of the Voronoi cells of the side quantizers and central quantizer, respectively.

---

[1]The quantizer $\mathcal{Q}^c$ is a lattice quantizer meaning that the partition cells of the quantizer are given by the Voronoi cells of the underlying lattice Λ_c. The expression (3) is derived under a high-resolution assumption, but will in practice be accurate for bit rates as low as three bits.

2.2. Transmission Effects

We will model transmission effects by introducing the binary stochastic arrival processes $\gamma_m^i = \{\gamma_m^i(k)\}_{k \in \mathbb{N}_0}$:

$$\gamma_m^i(k) = \begin{cases} 1 & \text{if } s_m^i(k) \text{ arrives error-free at time } k, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where $s_m^i(k)$ denotes the ith encoded description of the mth sensor at time k. The success probabilities $\lambda_m^i(k) \triangleq \mathcal{P}\{\gamma_m^i(k) = 1\}$ satisfy:

$$\lambda_m^i(k) = \left(1 - \beta_m\big(u_m(k)g_m(k)\big)\right)^{b_m^i(k)}, \tag{7}$$

where $b_m^i(k)$ denotes the packet length (which we take equal to the bit-rate), $g_m(k)$ refers to the channel power gain, i.e., the square of the magnitude of the complex channel[2], $u_m(k)$ is the transmission power used by the mth sensor radio power amplifier, and $\beta_m(\cdot) \colon [0, \infty) \to [0, 1]$ is the bit-error rate (BER). The latter is a monotonically decreasing function, which depends on the modulation scheme employed. Notice that the product of probabilities, i.e., $\prod_{i=0}^{J_m(k)-1} \lambda_m^i(k)$, is independent of the number of descriptions $J_m(k)$, since with symmetric MDC, $b_m^i(k) = b_m(k)/J_m(k)$.

It follows from (7), 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 levels used by the transmitters. Unfortunately, as we have seen in Section 2.1, 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.

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

$$E_m(b_m(k)u_m(k)) \triangleq \begin{cases} \dfrac{b_m(k)u_m(k)}{r} + E_\mathrm{P} & \text{if } u_m(k) > 0, \\ 0 & \text{if } u_m(k) = 0, \end{cases} \tag{8}$$

where $b_m(k) = \sum_{i=0}^{J_m(k)-1} b_m^i(k) = J_m(k)b_m^i(k)$, and E_P denotes the processing cost, i.e., the energy needed for wake-up, circuitry and sensing. Thus, the total energy consumption at sensor m at time k, is independent of $J_m(k)$, the number of descriptions chosen.

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

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

2.3. State Estimation with Intermittent Sensor Links

We will assume that the gateway knows, whether packets received from the sensors contain errors or not. Thus, at any time k, past and present realizations of the overall transmission process, say

$$\bar{\gamma}^k \triangleq \{\gamma_m^i(k-\ell)\}_{\ell \geq 0, \, i \in \{0, \ldots, J_m(k)-1\}, \, m \in \{1, \ldots, M\}}, \tag{10}$$

are available at the gateway. Thus, for state estimation purposes, the system amounts to sampling (1)-(2) only at the successful transmission instants of each sensor link. Consequently, the time-varying KF for the system (1) with output matrix

$$C(k) \triangleq \begin{bmatrix} \gamma_1(k)(C_1)^T & \cdots & \gamma_M(k)(C_M)^T \end{bmatrix}, \tag{11}$$

---

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

where $\gamma_m(k) = 1$, if at least one of the $J_m(k)$ descriptions is successfully received at sensor m (and zero otherwise), gives the best linear state estimates; see, e.g., [9]. These are given by:

$$\hat{x}(k+1) = A\hat{x}(k) + K(k+1)\big(\hat{y}(k+1) - C(k+1)A\hat{x}(k)\big),$$

where $\hat{y}_m(k)$ denotes the reconstruction at sensor m based on the $0 \le j \le J_m(k)$ received descriptions, and where

$$
\begin{aligned}
\hat{y}(k+1) &\triangleq \begin{bmatrix} \hat{y}_1(k) & \hat{y}_2(k) & \dots & \hat{y}_M(k) \end{bmatrix}^T \\
K(k) &\triangleq P(k)C(k)^T\big(C(k)P(k)C(k)^T + R(k)\big)^{-1} \\
P(k+1) &\triangleq AP(k)A^T + Q - AK(k)C(k)P(k)A^T \\
R(k) &\triangleq \mathrm{diag}\big(R_1 + D_1(k), \dots, R_M + D_M(k)\big),
\end{aligned}
\tag{12}
$$

with $\{D_m(k)\}$ being the distortions. If at least one but less than $J_m(k)$ descriptions are received, then $D_m(k)$ is given by (4). If all $J_m(k)$ descriptions are received, then $D_m(k)$ is given by (3). The recursion in (12) is initialized with $P(0) = P_0$ and $\hat{x}(0) = 0$.

3. ON-LINE DESIGN OF CODING AND POWER LEVELS

We have seen that sensor 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 controller which optimizes this trade-off over a future prediction horizon. To keep the sensors simple, the controller is located at the gateway. Its output contains information on the power levels and the codebook indices to be used by the M sensors, see Fig. 1.

3.1. Signaling

To save signal processing energy at the sensors, we would like to limit signaling from the gateway to the sensors as much as possible. In particular, the command signal for each sensor m will contain, in addition to the codebook index $j_m(k)$, a finitely quantized power increment, say $\delta u_m(k) \in \mathbb{U}_m$, rather than the actual power value, $u_m(k)$. Upon reception of the pair $(\delta u_m(k), j_m(k))$, the mth sensor chooses the codebook $j_m(k)$ and reconstructs the power level to be used by its radio power amplifier by simply setting

$$u_m(k) = u_m(k-1) + \delta u_m(k). \tag{13}$$

3.2. Cost Function

In order to trade energy consumption for estimation cost, at each time instant k, the proposed controller examines the finite horizon cost function

$$V(k) = V_1(k) + \rho V_2(k), \tag{14}$$

where $\rho \ge 0$ is a design parameter and

$$V_2(k) \triangleq \sum_{\ell=k+1}^{k+N} \sum_{m=1}^{M} E_m(b_m(\ell)u_m(\ell)) \tag{15}$$

is the predicted energy consumption over a horizon N. In (14),

$$V_1(k) \triangleq \sum_{\ell=k+1}^{k+N} \mathbb{E}\{\mathrm{trace}\{\bar{P}(\ell)\}\} \tag{16}$$

quantifies the estimation cost through the expected trace of[3]

$$\bar{P}(k) = P(k) - K(k)C(k)P(k). \tag{17}$$

---

[3] If the quantization noise was Gaussian and i.i.d., then $\bar{P}(\ell)$ would correspond to the conditional posterior covariance of $\hat{x}(k)$; see, e.g., [8,9].

Expectation in (16) is with respect to $\{\hat{y}_m(k)\}$, i.e., the set of possible reconstructions due to receiving different subsets of descriptions for each sensor. The probability distribution of this set depends upon the decision variables, i.e., bit-rates and power levels.

3.3. The Resultant Controller

At every time instant k, the predictive controller first finds the optimal set of pairs $\{\delta u_m(k), b_m^i(k)\}$ through a brute-force search strategy, where the cost function $V(k)$ is evaluated for all possible combinations of bit-rates and power levels over the prediction horizon.

Each sensor can either do standard single-description coding (SDC) or MDC. Moreover, for any given bit-rate, the controller also needs to decide upon the number of descriptions $J_m(k)$ and the level of redundancy between the $J_m(k)$ descriptions. To develop a simple method to select the coding scheme, we recall 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 process at any given sensor does not contain significant information about current and past measurements of any of the sensors.[4] 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 $(\delta_m(k), b_m^i(k))$ and, thus $u_m(k)$, see (13), the controller picks the quantization scheme which results in the minimum expected distortion on $y_m(k)$ by using the expressions in Section 2.1. Furthermore, given the bit-rate $b_m(k) = J_m(k)b_m^i(k)$ and by considering $\{\lambda_m^i : i = 0, \dots, J_m(k)\}, \forall J_m(k)$ 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 [13] will be used.

4. SIMULATIONS

We consider a system model (1) with $A = \begin{bmatrix} 1.6718 & -0.9948 \\ 1 & 0 \end{bmatrix}$, $Q = 1/2I$, and $P_0 = 0.3I$. We simulate a WSN having $M = 2$ sensors with $C_1 = [1\ 0], C_2 = [0\ 1]$, and variances $R_1 = R_2 = 1/100$. Power levels are constrained to the interval $0 \le u_m(k) \le 3 \times 10^{-4}$, increment values are restricted to $\pm 3 \times 10^{-5}$. Each sensor may use either an SD scalar quantizer at a bit-rate of $b_m(k) \in \{3, \dots, 8\}$ bits, or an MD scalar quantizer with two descriptions. In the MDC case the total bit-rate is restricted to $b_m(k) \in \{6, 7, 8\}$ bits, so that the side description rates are $b_m^i(k) \in \{3, 3.5, 4\}$ bits.[5]

We use measured channel data obtained at the 2.4 GHz ISM band within an office space area at the Signals and Systems Group of Uppsala University, Sweden. The top diagram of Fig. 3 illustrates the channel gains of two realizations, one with horizontal and one with vertical polarization.

In Fig. 2, we have shown the expected distortion (before the KF) when coding $y_m(k)$ as a function of the channel gain, for $J_m(k) \in \{1, 2\}$. For that purpose, we have fixed the power levels to $u_m(k) = 4 \times 10^{-5}$ and the bit-rates to $b_m(k) = 7$. Notice that, when the channel gain is weak, e.g., below -105 dB, it is better to use MDC than SDC.

---

[4] Unless we utilize 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.

[5] The Voronoi cell volume ratio ϕ of the MD quantizers, is restricted to $\phi = 3, 5, 7$, see Section 2.1 and Fig. 2.
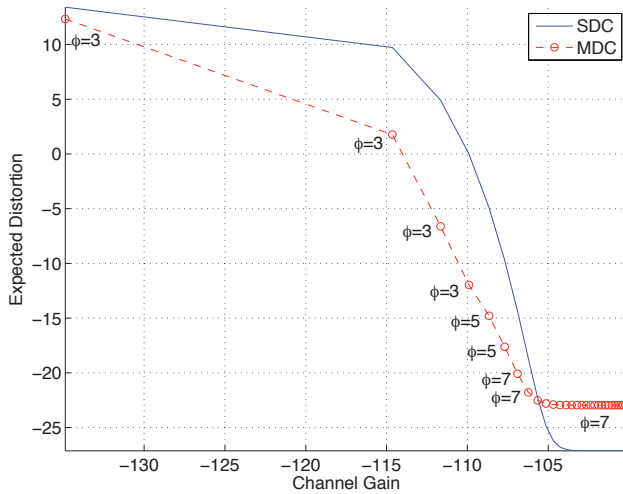
Fig. 2. Expected distortion (in dB) as a function of the channel gain. $\phi = \mathcal{V}_s/\mathcal{V}_c$ indicates the ratio of the volumes of the Voronoi cells for MDC.

A complete simulation of the proposed algorithm with $N = 1$ and $\rho = 10^6$ for $s = 5000$ samples is depicted in Fig. 3. In the bottom two diagrams of Fig. 3, we have used dots to illustrate when MDC is used. Notice that, for example, whenever the channel gain is dropping to low levels, the proposed controller chooses MDC.

We quantify our results via $V \triangleq \frac{1}{s} \sum_{k=0}^{s-1} V(k)$, see (14), resulting in $V = 0.1113$. For comparison, we also carried out simulations of the controller under the same conditions, but where only SDC was allowed, yielding $V = 0.1220$. Thus, in the present case, the use of MDC gives a performance improvement of 8.77%.

5. CONCLUSIONS

We have shown that state estimation via WSNs over fading channels can be improved by the use of a novel multiple-description quantizer and power controller. A simulation study revealed that gains of more than 8% was possible by using simple MDC schemes at the sensors and without significantly increasing the complexity at the gateway. This gain can be further improved by reducing the maximum power level available at the sensors.

6. REFERENCES

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

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

[4] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Processing*, vol. 56, no. 2, pp. 757–770, Feb. 2008.

Fig. 3. System performance.

[5] 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.*, Honolulu, HI, 2007, vol. 3, pp. 513–516.

[6] 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.

[7] G. C. Goodwin, J. Østergaard, and D. E. Quevedo, "A vector quantization approach to scenario generation for stochastic NMPC," in *Proc. Int. Workshop on Assessment and Future Directions of NMPC*, Pavia, Italy, 2008.

[8] 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.

[9] T. Söderström, *Discrete-Time Stochastic Systems*, Prentice Hall, 1994.

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

[11] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 821 – 834, May 1993.

[12] J. Østergaard, J. Jensen, and R. Heusdens, "n-channel entropy-constrained multiple-description lattice vector quantization," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1956 – 1973, 2006.

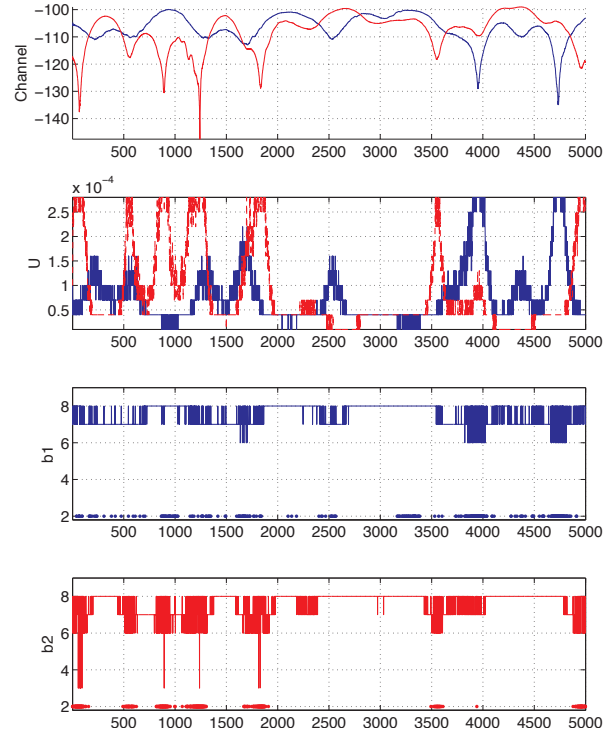[13] J. Østergaard, *Multiple-description lattice vector quantization*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, June 2007.