ENERGY-EFFICIENT JOINT ESTIMATION IN SENSOR NETWORKS: ANALOG VS. DIGITAL

Shuguang Cui[1], Jin-Jun Xiao[2], Andrea J. Goldsmith[1], Zhi-Quan Luo[2], and H. Vincent Poor[3]

[1]Wireless System Lab, Department of Electrical Engineering, Stanford University.

[2]Department of Electrical and Computer Engineering, University of Minnesota.

[3]Department of Electrical Engineering, Princeton University.

ABSTRACT

Sensor networks in which energy is a limited resource so that energy consumption must be minimized for the intended application are considered. In this context, an energy-efficient method for the joint estimation of an unknown analog source under a given distortion constraint is proposed. The approach is purely analog, in which each sensor simply amplifies and forwards the noise-corrupted analog observation to the fusion center for joint estimation. The total transmission power across all the sensor nodes is minimized while satisfying a distortion requirement on the joint estimate. The energy efficiency of this analog approach is compared with previously proposed digital approaches with and without coding. It is shown in our simulation that the analog approach is more energy-efficient than the digital system without coding, and in some cases outperforms the digital system with optimal coding.

1. INTRODUCTION

A typical Wireless Sensor Network (WSN), as shown in Fig. 1, consists of a fusion center and a large number of geographically distributed sensors. The sensors typically have limited energy resources and communication capability. Each sensor in the network makes an observation of the target of interest, generates a local signal (either analog or digital), and then sends it to the fusion center where the received sensor signals are combined to produce a final estimate of the observed signal. Sensor networks of this type are well-suited for situation awareness applications such as environmental monitoring and smart factory instrumentation.
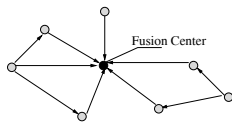

Fig. 1. Sensor network with a fusion center.

Decentralized estimation has been studied first in the context of distributed control [1], distributed tracking [2], and most recently in wireless sensor networks [3]. Among these studies, it is usually assumed that the joint distribution of the sensor observations is known. In practical systems, the probability density function (pdf) of the observation noise is hard to characterize, especially for a large scale sensor network. This motivates us to devise signal processing algorithms that do not require the knowledge of

the sensor noise pdf. Recently, universal decentralized estimation schemes (DES) without the knowledge of noise distribution have been proposed in [7] and [8]. In [7], the author considered the universal DES in a homogeneous sensor network where sensors have observations of the same quality, while in [8], the universal DES in an inhomogeneous sensing environment was considered. These proposed DESs require each sensor to send to the fusion center a short discrete message with length decided by the local Signal to Noise Ratio (SNR), while the performance is guaranteed to be within a constant factor of that achieved by the Best Linear Unbiased Estimator (BLUE). An assumption in these proposed schemes is that the channels between sensors and the fusion center are perfect, and all messages are received by the fusion center without any distortion. However, due to power limitations and channel noise, the signal sent by each individual sensor to the fusion center will be corrupted. Therefore, the goal of the transmission system design for the joint estimation problem is to minimize the effect of the channel corruption while consuming the minimum amount of power at each node.

Historically, if the sensor observation is in analog form, we have two main options to transmit the observation from the sensors to the fusion center: analog or digital communication. For the analog approach, we keep the observation signal analog and further use analog modulation schemes to transmit the signal, which is also called an amplify-and-forward approach. In the digital approach, we digitize the observation into bits, possibly apply channel coding, then use digital modulation schemes to transmit the data. It is well known ([4], [5]) that for a single Gaussian source with an AWGN channel, the amplify-and-forward approach is optimal. However, for an arbitrary source with multiple observations and multiple transmission channels, it is unclear which approach will lead to a smaller power consumption while meeting the distortion requirements at the fusion center.

Another severe challenge facing sensor networks is the hard energy constraint. Since each sensor is equipped with only a small-size battery, for which replacement is very expensive if not impossible, the power consumption must be minimized to increase the network lifetime [6]. Therefore, energy efficiency can be used as a performance criterion to evaluate different sensor network designs under the same distortion requirement. In [9], a digital system is proposed to minimize the total power consumption for the joint estimation problem, where each sensor quantizes the analog observation into digital bits and transmits the bits to the fusion center with uncoded MQAM. The number of quantization bits for each sensor is optimized with the target of minimizing the total transmission power across all the sensor nodes to achieve a given distortion. In this paper, we propose an analog counterpart for the same system and compare the energy efficiency between the analog approach and the digital one. As in [9], we assume that the observed signal

is analog and bounded, the fusion center deploys the best unbiased linear estimator, the observation noise is uncorrelated across different sensors, and only the variance of the observation noise is known.

Our paper is organized as follows. Section II discusses the problem formulation. Section III compares the energy efficiency between the analog approach and the digital approach via some numerical examples. Section IV summaries our conclusions.

2. MINIMUM POWER ANALOG INFORMATION COLLECTION

We assume that there are K sensors and the observation $x_k(t)$ at sensor k is represented as a random signal $\theta(t)$ corrupted with the observation noise $n_k(t)$: $x_k(t) = \theta(t) + n_k(t)$. Each sensor transmits the signal $x_k(t)$ to the fusion center where $\theta(t)$ is estimated from the $x_k(t)$'s, $k = 1, \cdots, K$. We further assume that $n_k(t)$ is of unknown statistics and the amplitude of $x(t)$ is bounded within $[-W, W]$, which is defined by the sensing range of each sensor. For simplicity we assume $W = 1$, but our analysis can be easily extended to any values. We also assume that the network is synchronized, which may be enabled by utilizing beacon signals in a separate control channel.

We assume that $x_k(t)$ is also band-limited and the information is contained within the frequency range $[-B/2, B/2]$. We consider an analog Single Side-Band (SSB) system with a coherent receiver [11]. The transmitted signal is given by

$$y_t(t) = 2\sqrt{\alpha}\cos(\omega_c t)x(t) + 2\sqrt{\alpha}\sin(\omega_c t)\hat{x}(t), \quad (1)$$

for which the average transmission power is

$$P = 4\alpha P_x \le 4\alpha W^2 \quad (2)$$

where 4α is the transmitter power gain and P_x is the peak power of $x(t)$.

The received signal at the fusion center is given by

$$y_r(t) = 2\sqrt{\alpha}\sqrt{g}\cos(\omega_c t)x(t) + 2\sqrt{\alpha}\sqrt{g}\sin(\omega_c t)\hat{x}(t) + n_c(t) \quad (3)$$

where g is the channel power gain, $\hat{x}(t)$ is the Hilbert transform of $x(t)$, and $n_c(t)$ is the channel AWGN. Hence, at the output of the coherent detector the signal is

$$y(t) = \sqrt{\alpha}\sqrt{g}x(t) + \frac{1}{2}n_c^I(t)\cos(\pi\frac{B}{2}t) + \frac{1}{2}n_c^Q(t)\sin(\pi\frac{B}{2}t), \quad (4)$$

where $n_c^I(t) + jn_c^Q(t)$ is the complex envelope of $n_c(t)$. After passing $y(t)$ through a low-pass filter and sampling the baseband signal at a sampling rate B, we can obtain an equivalent discrete-time system. Since we have K such sensors, the overall system is shown in Fig. 2. The K transmitters share the channel via Frequency Division Multiple Access (FDMA), which has the same spectral efficiency as the Time Division Multiple Access (TDMA) that is used in [9].
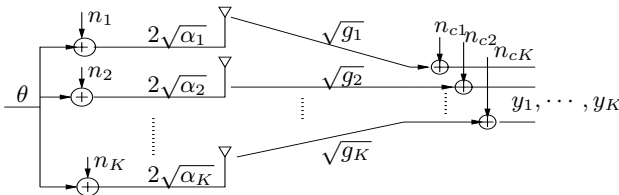

Fig. 2. Amplify and Forward

The received signal vector at each time instance is given by

$$\mathbf{y} = \mathbf{h}\theta + \mathbf{v}, \quad (5)$$

where
$$\mathbf{y} = [y_1, y_2, \cdots, y_K]^\dagger,$$
$$\mathbf{h} = [\sqrt{\alpha_1 g_1}, \sqrt{\alpha_2 g_1}, \cdots, \sqrt{\alpha_K g_K}]^\dagger,$$
$$\mathbf{v} = [\sqrt{\alpha_1 g_1}n_1 + n_{c1}, \cdots, \sqrt{\alpha_K g_K}n_k + n_{cK}]^\dagger,$$

and $\dagger$ means transpose.

According to [10], the best linear unbiased estimator (BLUE) for θ is given by

$$\hat{\theta} = [\mathbf{h}^\dagger \mathbf{R}^{-1}\mathbf{h}]^{-1}\mathbf{h}^\dagger \mathbf{R}^{-1}\mathbf{y}$$
$$= \left(\sum_{k=1}^{K}\frac{\alpha_k g_k}{\sigma_k^2\alpha_k g_k + \xi_k^2}\right)^{-1}\sum_{k=1}^{K}\frac{\sqrt{\alpha_k g_k}y_k}{\sigma_k^2\alpha_k g_k + \xi_k^2}, \quad (6)$$

where the noise variance matrix $\mathbf{R}$ is a diagonal matrix with $R_{kk} = \sigma_k^2\alpha_k g_k + \xi_k^2$ with σ_k^2 the variance of the sensor observation noise $n_k(t)$, $k = 1, \cdots, K$. The channel noise variance ξ_k^2 is defined by the noise power spectral density and the bandwidth B.

The mean squared error of this estimator is given as [10]

$$Var[\hat{\theta}] = [\mathbf{h}^\dagger \mathbf{R}^{-1}\mathbf{h}]^{-1}$$
$$= \left(\sum_{k=1}^{K}\frac{\alpha_k g_k}{\sigma_k^2\alpha_k g_k + \xi_k^2}\right)^{-1}. \quad (7)$$

According to Eq. (2), the transmit power for node k is bounded by $4W^2\alpha_k$. Therefore, the minimum power analog information collection problem can be cast as

$$\min \quad \sum_{k=1}^{K}W^2\alpha_k$$
$$\text{s. t.} \quad \left(\sum_{k=1}^{K}\frac{\alpha_k g_k}{\sigma_k^2\alpha_k g_k + \xi_k^2}\right)^{-1} \le D_0$$
$$\alpha_k \ge 0, \quad k = 1, \cdots, K$$

where D_0 is the distortion target. However, this problem is not convex over the α_k's.

Let us define

$$r_k = \frac{\alpha_k g_k}{\sigma_k^2\alpha_k g_k + \xi_k^2} = \frac{1}{\sigma_k^2 + \frac{\xi_k^2}{g_k\alpha_k}}.$$

Then the above optimization problem is equivalent to

$$\min \quad \sum_{k=1}^{K}W^2\alpha_k$$
$$\text{s. t.} \quad \sum_{k=1}^{K}r_k \ge \frac{1}{D_0}$$
$$r_k = \frac{1}{\sigma_k^2 + \frac{\xi_k^2}{g_k\alpha_k}}, \quad \alpha_k \ge 0, \quad \forall k,$$

where we see that the variable α_k can be completely replaced by a function of r_k. Therefore, the problem can be transformed into a problem with variables $\{r_1, r_2, \ldots, r_K\}$ shown as follows:

$$\min \quad \sum_{k=1}^{K}\frac{W^2\xi_k^2}{g_k}\left(\frac{r_k}{1 - r_k\sigma_k^2}\right)$$
$$\text{s. t.} \quad \sum_{k=1}^{K}r_k \ge \frac{1}{D_0}; \quad 0 \le r_k < \frac{1}{\sigma_k^2}, \quad \forall k \quad (8)$$

which is convex over r_k. The upper limit on r_k in the second constraint is due to the fact that $r_k = 1/\left(\sigma_k^2 + \frac{\xi_k^2}{g_k\alpha_k}\right)$ and $\frac{\xi_k^2}{g_k\alpha_k} \geq 0$.

Now we solve Eq. (8). Its Lagrangian G is given as

$$G(L, \lambda_0) = \sum_{k=1}^{K} \frac{W^2\xi_k^2}{g_k}\left(\frac{r_k}{1 - r_k\sigma_k^2}\right) + \lambda_0\left(\frac{1}{D_0} - \sum_{k=1}^{K} r_k\right) \tag{9}$$

for $0 \leq r_k \leq \frac{1}{\sigma_K^2}, \forall k$, which leads to the following Karush-Kuhn-Tucker (KKT) conditions [12]:

$$\frac{W^2\xi_k^2}{g_k}\frac{1}{(1 - r_k\sigma_k^2)^2} - \lambda_0 = 0, \qquad \forall k$$

$$\sum_{k=1}^{K} r_k - \frac{1}{D_0} = 0$$

for $0 \leq r_k \leq \frac{1}{\sigma_K^2}, \forall k$. Without loss of generality, we rank the channel quality such that $\frac{\xi_1^2}{g_1} \leq \frac{\xi_2^2}{g_2} \leq \ldots \leq \frac{\xi_K^2}{g_K}$ where we call the quantity $\frac{g_k}{\xi_k^2}$ the channel SNR, and we define

$$f(M) = \frac{\sqrt{\frac{\xi_M^2}{g_M}}A(M)}{B(M)}, \quad \text{for } 1 \leq M \leq K, \tag{10}$$

where $A(M) = \sum_{m=1}^{M} \frac{\sqrt{\frac{\xi_m^2}{g_m}}}{\sigma_m^2}$ and $B(M) = \sum_{m=1}^{M} \frac{1}{\sigma_m^2} - \frac{1}{D_0}$.

Let us find K_1 such that $f(K_1) < 1$ and $f(K_1 + 1) \geq 1$. Using the same techniques as in [9], we can show that this K_1 is unique unless $f(M) < 1$ for all $1 \leq M \leq K$, in which case we take $K_1 = K$. Then the KKT conditions give $\lambda_0 = \left(\frac{WA(K_1)}{B(K_1)}\right)^2$ and

$$r_k^{opt} = \frac{1}{\sigma_k^2}\left(1 - W\frac{\sqrt{\frac{\xi_k^2}{g_k}}}{\sqrt{\lambda_0}}\right)^{+}, \quad \forall k \tag{11}$$

where $(x)^+$ equals 0 when $x < 0$, and otherwise is equal to x.

Hence, by definition, we have

$$\alpha_k^{opt} = \frac{\xi_k^2}{g_k}\frac{r_k^{opt}}{1 - \sigma_k^2 r_k^{opt}}$$

$$= \frac{\xi_k^2}{g_k\sigma_k^2}\left(\sqrt{\frac{g_k}{\xi_k^2}}\eta_0 - 1\right) \quad k = 1, \cdots, K_1, \tag{12}$$

and $\alpha_k = 0$ otherwise, where $\eta_0 = A(K_1)/B(K_1)$.

Therefore, the optimal power allocation strategy is divided into two steps. In the first step, a threshold for k is obtained according to Eq. (10). For channels with SNR worse than this threshold, the corresponding sensor is shut off and no power is wasted. For the remaining active sensors, power should be assigned according to Eq. (12). From Eq. (12) we see that when the channel is fairly good, i.e., $\sqrt{\frac{g_k}{\xi_k^2}}\eta_0 \gg 1$, we have $\alpha_k^{opt} \propto \sqrt{\frac{\xi_k^2}{g_k}}\frac{\eta_0}{\sigma_k^2}$, which means the optimal solution is inversely proportional to the square root of the channel SNR. When $\sqrt{\frac{g_k}{\xi_k^2}}\eta_0$ is close to one, the optimal solution may no longer have such properties. For all channel

conditions, the power is scaled by the factor $\frac{1}{\sigma_k^2}$, which means that more power is used to transmit the signals from the sensors with better observation quality.

We now solve the optimization problem for some specific examples. We assume that the channel power gain $g_k = \frac{G_0}{d_k^{3.5}}$ where d_k is the transmission distance from sensor k to the fusion center and $G_0 = -30$ dB is the gain at $d = 1$ m. As in [9], we generate σ_k^2 uniformly within the range $[0.01, 0.08]$. We take $B = 10$ KHz and $\xi_k^2 = -90$ dBm, $k = 1, \cdots, K$. For an example with 100 sensors, Fig. 3 (a) shows the relative power savings compared with the uniform transmission strategy where all the sensors use the same transmission power to achieve the given distortion target. The relative power savings is plotted as a function of $R = \frac{\sqrt{\mathbf{Var}(d)}}{\mathbf{E}(d)}$, the distance deviation normalized by the mean distance. For each value of R, we average the relative power savings over 100 random runs where in each run the d_k's are randomly generated according to the given R. As expected, a larger variation of distance and corresponding channel quality leads to a higher power savings when this variation is exploited with optimal power allocation.

In Fig. 3 (b), the number of active sensors over R is shown for an example with 10 sensors. We see that when the transmission distances for different sensors span a wide range of values (i.e., R is large), more sensors can be shut off to save energy, since the remaining sensors have very good channels.
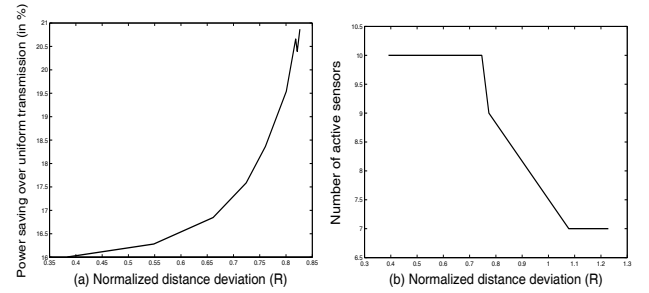

Fig. 3. (a) Power savings of optimal power allocation vs. uniform power; (b) Number of active sensors versus distance deviation

In our model we minimize the power sum $\sum_k P_k$, i.e., the L^1-norm of the transmission power vector $\mathbf{P} = (P_1, P_2, \ldots, P_K)$. If the channel gain and the variance of the observation noise for each sensor are ergodically time-varying on a block-by-block basis, minimizing the L^1-norm of $\mathbf{P}$ in each time block minimizes $E\{\sum_k P_k\}$ with $E\{\}$ the expectation operation. In other words, it maximizes the average node lifetime, which is defined as $\frac{1}{K}\sum_k \frac{E_0}{E\{P_k\}}$ with E_0 the battery energy available to each sensor (we assume that E_0 is the same for all the sensors). This can be proved by the fact that $\frac{1}{K}\sum_k \frac{E_0}{E\{P_k\}} \geq \frac{E_0}{E\{\frac{1}{K}\sum_k P_k\}}$. However, when the channel is static and the variance of the observation noise is time-invariant, minimizing the L^1-norm may lead some individual sensors to consume too much power and die out quickly. In this case minimizing the L^∞-norm, i.e., minimizing the maximum of the individual power values, is the most fair for all sensors, but the total power consumption can be high. As in [9], we can make a compromise to minimize the L^2-norm of $\mathbf{P}$. In this way, we can penalize the large terms in the power vector while still keeping the total power consumption reasonably low. For the L^2-norm minimization, the problem formulation becomes

$$\min \quad \sum_{k=1}^{K} \frac{W^4 \xi_k^4}{g_k^2} \left(\frac{r_k}{1 - r_k \sigma_k^2} \right)^2$$

$$\text{s. t.} \quad \sum_{k=1}^{K} r_k \geq \frac{1}{D_0}; \quad 0 \leq r_k < \frac{1}{\sigma_k^2}, \quad \forall k, \quad (13)$$

which we can solve using interior point methods [12]. In the next section, we compare the power efficiency of the analog and the digital approaches previously discussed, where we minimize the L^2-norm of $\mathbf{P}$. We use the L^2-norm since it simplifies the comparison, but the comparison can be made for any power norms.

3. ANALOG VS. DIGITAL

In order to transmit the observed analog signal with frequency range $[-B/2, B/2]$ to the fusion center, each sensor in the digital system proposed in [9] first samples the signal at a sampling rate B, then quantizes each sample into b_k bits, and finally uses uncoded MQAM to transmit the b_k bits with a symbol rate B and constellation size $M = 2^{b_k}$. Therefore, the total transmission bandwidth in the passband is approximately equal to KB for the digital system where TDMA is used for the multiple access. For the SSB scheme used in our analog system, each sensor only occupies $B/2$ in the passband such that the total bandwidth requirement is $KB/2$ when FDMA is used for the multiple access. Therefore, under the assumption of orthogonal channel usage, the analog system can support double the number of sensors that the digital system supports in the same amount of bandwidth. This is mainly caused by the fact that the digital approach proposed in [9] forces the transmission symbol rate to be equal to the sampling rate in the source coding part.

The power efficiency comparison between the analog approach and the digital approach is shown in Fig. 4, where we deploy 10 sensors for the digital system and we plot the power curves for both the case where uncoded MQAM is used (the dotted line) and the case where single-user capacity-achieving channel codes are applied to each transmitter, which gives the fundamental Shannon limit (the dashed line) for the digital system with orthogonal channel usage. From the figure we see that the analog approach with 10 nodes, which have the same observation quality and transmission distances as in the digital system, is more energy efficient than the digital approach with uncoded MQAM, but not necessarily more energy efficient than the fundamental limit curve. However, since we can support 20 nodes in the analog system with the same bandwidth requirement, with the 20 analog nodes we can achieve a better power efficiency (the solid line) than the digital system with 10 nodes under the optimal single-user channel coding. This is true even when the extra ten nodes in the analog system have worse observation quality and larger transmission distances than the first ten nodes.

4. CONCLUSIONS

In this paper, we have shown that for a bounded source with unknown statistics and a fusion center equipped with the best linear unbiased estimator, we can minimize the total power consumption across all the sensor nodes under a certain distortion requirement. The information collection can be implemented with an analog approach, which may be more energy efficient than the digital approach when only orthogonal multiple access schemes such as TDMA and FDMA are used.
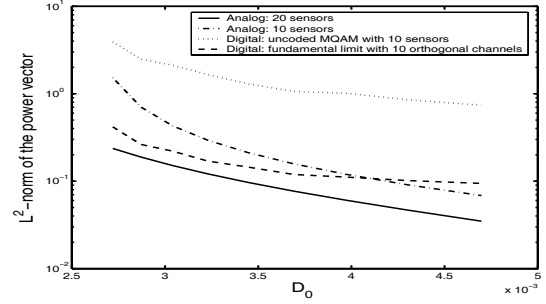

Fig. 4. Comparison of the L^2-norm of the power vector

5. REFERENCES

[1] D.A. Castanon and D. Teneketzis, "Distributed Estimation Algorithms for Nonlinear Systems," *IEEE Transactions on Automatic Control*, Vol. AC-30, pp. 418–425, 1985.

[2] A.S. Willsky, M. Bello, D.A. Castanon, B.C. Levy, and G. Verghese, "Combining and Updating of Local Estimates and Regional Maps Along Sets of One-dimensional Tracks," *IEEE Transactions on Automatic Control,* Vol. AC-27, pp. 799–813, 1982.

[3] H.C. Papadopoulos, G.W. Wornell, and A.V. Oppenheim, "Sequential Signal Encoding from Noisy Measurements Using Quantizers with Dynamic Bias Control," *IEEE Trans. on Information Theory*, Vol. 47, pp. 978–1002, 2001.

[4] T. J. Goblick, "Theoretical Limitations on the Transmission of Data from Analog Sources," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 558567, Oct. 1965.

[5] M. Gastpar, B. Rimoldi and M. Vetterli, "To Code, or Not To Code: Lossy Source-channel Communication Revisited," *IEEE Trans. on Information Theory*, Vol. 49 pp. 1147-1158, May 2003.

[6] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained Modulation Optimization," to appear at *IEEE Transactions on Wireless Communications*, 2003. Also available at http://wsl.stanford.edu/Publications.html.

[7] Z.-Q. Luo, "Universal Decentralized Estimation in a Bandwidth Constrained Sensor Network," submitted to *IEEE Transactions on Information Theory*. Also available at http://www.ece.umn.edu/users/luozq/.

[8] Z.-Q. Luo, J.-J., Xiao, "Decentralized Estimation in an Inhomogeneous Sensing Environment," at *International Syposium on Information Theory*, Chicago, 2004. Also available at http://www.ece.umn.edu/users/luozq/.

[9] J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Joint Estimation in Sensor Networks under Energy Constraint," to appear at the *IEEE first conference on Sensor and Ad Hoc Communications and Networks*, Santa Clara, CA, October, 2004.

[10] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[11] S. Haykin, *Communication Systems*, 3rd edition, John Wiley & Sons, New York, 1994.

[12] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, Cambridge, U.K., 2003.