Energy-Harvesting for Source-Channel Coding in Cyber-Physical Systems

P. Castiglione
FTW
Wien, Austria
Email: castiglione@ftw.at

O. Simeone
NJIT
Newark, NJ, USA
Email: osvaldo.simeone@njit.edu

E. Erkip
Poly of NYU
Brooklyn, NY, USA
Email: elza@poly.edu

T. Zemen
FTW
Wien, Austria
Email: thomas.zemen@ftw.at

***Abstract*—The overall energy required to digitize a given physical source can be comparable to the energy required for communication of the produced information bits, especially in cyber-physical sensing systems where radio links are short. When energy is at a premium, this fact calls for energy management solutions that are able to properly allocate the available energy over time between source and channel coding tasks. Energy management is particularly challenging for devices that operate via energy-harvesting, since the controller has to operate without full knowledge of the energy availability in the future. This work addresses the problem of energy allocation over source digitization and communication for a single energy-harvesting sensor. First, optimal policies that minimize the average distortion under constraints on the stability of the data queue connecting source and channel encoders are derived. It is shown that such policies perform independent resource optimizations for the source and channel encoders. The drawback of these policies is that they require an arbitrarily large battery to counteract the variability of the harvesting process and an infinite data queue to mitigate temporal variations in source and channel qualities. Suboptimal policies that do not have such drawbacks are then investigated as well, along with the optimal trade-off distortion vs. delay, which is addressed via dynamic programming tools.**

I. Introduction

Cyber-physical systems (CPS), broadly speaking, enable the combination and coordination of computational and physical components of the "smart world". A major bottleneck in the development of CPSs is well-recognized to be energy consumption and storage, due to the difficulty to provide a continuous or sporadic energy source in situ for the operation of devices, especially wireless nodes. A natural component of any comprehensive solution to the bottleneck problem identified above is to leverage *energy-harvesting technologies,* whereby the energy necessary for the operation of the devices is collected from the environment by converting different forms of energy, such as solar, elastic or radio frequency, into electrical power. The regime where a device is solely powered by the energy that is able to scavenge from the environment is typically referred to as *energy neutral* (EN). With energy harvesting, one can generally only guarantee energy availability in a statistical sense, apart from very special cases where the energy harvested can be accurately predicted. As also demonstrated by a number of recent works (see, e.g., [1][2] and references therein), the main challenges in the design of EN systems is hence the need to balance the use of the available energy so as to attain the desired trade-off between *short-term* performance guarantees, such as maximum delay constraints, with *long-term* requirements, such as throughput or average distortion.

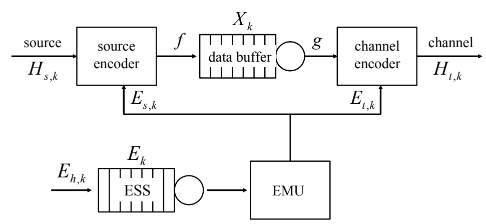

Figure 1. An energy-harvesting device encompassing an Energy Management Unit (EMU) that performs energy allocation between source encoder and channel encoder.

While the previous work mentioned above on EN systems focuses solely on the energy required for transmission, here we observe that, as reported in [3], the energy spent for *source digitization* can out-weight that used for communication. This was shown in [3] via measurements, by accounting only for the energy needed for compression of simple web files and using a Wireless Local Area Network (WLAN) interface for transmission. The energy allocation problem between source and channel coding components identified above is even more pronounced in EN-CPS, where communication distances are often such that the energy needed for communication can be one or two orders of magnitude less than in a WLAN. The issue of energy allocation between source digitization and transmission for *power-limited networks* has received some attention [3]-[6]. However, EN-CPS pose new research challenges due to the randomness of energy harvesting.

In this paper, we tackle some of these research challenges by focusing on a single-link EN-CPS, where a single transmitter, illustrated in Fig. 1, communicates to a single receiver. We focus first on the long-term performance of the system in terms of average distortion, assuming that delay is not an issue and that there is enough energy storage as required in Sec. III, and then introduce short-term limitations in terms of finite battery and/or delay in Sec. III-A and Sec. IV.

II. System Model

We consider a system in which a single sensor shown in Fig. 1 communicates with a single receiver. Time is slotted.

The energy $E_{h,k} \in \mathbb{R}^+$ harvested in time-slot k is stored in an Energy Storage System (ESS), which is assumed at first to be of infinite size. For convenience, all energy measures are normalized to the number N of channel discrete-time symbols available for communication in each time slot, also referred to as channel uses (c.u.). The energy arrival $E_{h,k}$ is a assumed to be a stationary ergodic process with known probability density function (pdf) $p_{E_h}(e)$. The energy E_{k+1} available for use at slot $k+1$ is the residual energy from the previous slot plus the energy arrival at time-slot $k+1$. This evolves as

$$E_{k+1} = [E_k - (E_{s,k} + E_{t,k})]^+ + E_{h,k+1}, \quad (1)$$

where $E_{s,k}$ and $E_{t,k}$ account for the energy spent in slot k per channel use for source digitization (i.e., by the source encoder) and data transmission (i.e., by the channel encoder), respectively, as discussed below. Notice that the energy arriving at time slot $k+1$ is immediately available for use in that slot.

The sensor measures M samples of a given source during each slot. The quality of such observation in slot k depends on a parameter $H_{s,k} \in \mathcal{H}_s$, where $\mathcal{H}_s$ is assumed to be discrete and finite, which is a stationary ergodic process over the time slots k with known probability mass function (pmf) $p_{H_s}(h_s)$. For instance, the sensor may observe the phenomenon of interest with some additive noise whose variance $H_{s,k}$ changes across blocks k due to source movement or environmental factors affecting the measure quality. The source encoder at the sensor acquires the source in a lossy fashion. The loss, due to sampling, analog-to-digital compression and compression, is characterized by distortion $D_k \in \mathbb{R}^+$, as measured with respect to some distortion metric such as the mean square error (MSE).

The number of bits generated by the source encoder at the sensor at slot k is $f(D_k, E_{s,k}, H_{s,k})$, where f is a given function of the distortion level D_k, of the energy per channel use allocated to the source encoder $E_{s,k}$ and on the observation state $H_{s,k}$. The resulting bit stream is buffered in a first-input-first-output (FIFO) data queue with queue length X_k. The function $f(D_k, E_{s,k}, H_{s,k})$ is assumed to be separately continuous convex and non-increasing in D_k and $E_{s,k}$. For simplicity, we will denote such functions also as $f^{h_s}(D_k, E_{s,k}) = f(D_k, E_{s,k}, H_{s,k} = h_s)$. Some examples for function f can be found in [4][5] (see also [7]) and will be further discussed in Sec. V.

The fading channel between sensor and destination is characterized by a process $H_{t,k}$, assumed to be stationary ergodic, where $H_{t,k} \in \mathcal{H}_t$, with set $\mathcal{H}_t$ being for simplicity discrete and finite, and pmf of H_k is given by $p_{H_t}(h_t)$. The *channel encoder* uses energy per channel use $E_{t,k}$ and transmits over N channel uses per slot. A maximum number $g(H_{t,k}, E_{t,k})$ of bits per slot can be delivered successfully to the destination. The channel rate function $g(H_{t,k}, E_{t,k})$ is assumed to be continuous, concave and non-decreasing in $E_{t,k}$. We also use the notation $g^{h_t}(E_{t,k}) = g(H_{t,k} = h_t, E_{t,k})$. The channel encoder takes $\min[X_k, g(H_{t,k}, E_{t,k})]$ bits from the data buffer, according to the selected transmission energy $E_{t,k}$. Based on this discussion, the data queue evolves as

$$X_{k+1} = [X_k - g(H_{t,k}, E_{t,k})]^+ + f(D_k, E_{s,k}, H_{s,k}). \quad (2)$$

A. Problem Definition

At each time slot k, the energy management unit (EMU), see Fig. 1, must determine the distortion D_k and the energies $E_{s,k}$ and $E_{t,k}$ to be allocated to the source and channel encoder, respectively. The decision is taken according to a policy $\pi := \{\pi_k\}_{k\geq 1}$, where $\pi_k := \{D_k(S^k), E_{s,k}(S^k), E_{t,k}(S^k)\}$ determines parameters $(D_k, E_{s,k}, E_{t,k})$ as a function of the present and past states $S^k = \{S_1, \ldots, S_k\}$ of the system, where the $S_i = \{E_i, X_i, H_{s,i}, H_{t,i}\}$ accounts for the state of the available energy E_i , for the data buffer X_i , for the the source observation state $H_{s,i}$ and the channel state $H_{t,i}$. Policies can be optimized according to different criteria, as discussed below.

III. Minimum Distortion Under Stability Constraint

In this section, we adopt as performance criterion the minimization of the long-term average distortion[1]

$$\bar{D} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[D_k] \quad (3)$$

under the constraint that the data queue is stable, that is, that the distribution of X_k is asymptotically *stationary* and *proper* (so that $\Pr(X_k = \infty) \to 0$) (see, e.g., [1]). We have the following result.

Proposition 1: The minimum distortion $\bar{D}$ under stability constraints is lower bounded by the solution of the convex problem

$$\bar{D} \geq \bar{D}^* = \arg\min \sum_{h_s} p_{H_s}(h_s) D^{h_s} \quad (4)$$

$$\text{s.t.} \quad \begin{array}{l} \sum_{h_s} p_{H_s}(h_s) f^{h_s}\left(D^{h_s}, E_s^{h_s}\right) \leq \sum_{h_t} p_{H_t}(h_t) g^{h_t}\left(E_t^{h_t}\right), \\ \sum_{h_s} p_{H_s}(h_s) E_s^{h_s} \leq (1-\alpha)\mathbb{E}[E_{h,k}], \\ \sum_{h_t} p_{H_t}(h_t) E_t^{h_t} \leq \alpha \mathbb{E}[E_{h,k}], \end{array} \quad (5)$$

where minimization is done over the parameters $D^{h_s} \geq 0$, $E_s^{h_s} \geq 0$ for $h_s \in \mathcal{H}_s$, $E_t^{h_t} \geq 0$ for $h_t \in \mathcal{H}_t$, and $0 < \alpha < 1$. Moreover, a policy π that achieves a distortion arbitrarily close to optimal is given by

$$\begin{cases} D_k = D^{h_s} & \text{for } H_{s,k} = h_s \\ E_{s,k} = \min\left[(1-\alpha)E_k,\ E_s^{h_s}\right] & \text{for } H_{s,k} = h_s \\ E_{t,k} = \min\left[\alpha E_k,\ E_t^{h_t}\right] & \text{for } H_{t,k} = h_t \end{cases}, \quad (6)$$

where parameters D^{h_s}, $E_s^{h_s}$, $E_t^{h_t}$ and $0 < \alpha < 1$ are obtained by solving (4)-(5) with the three constraints modified by subtracting a parameter $\epsilon > 0$ arbitrarily small to the right-hand sides.

Proof: Follows from [7], to which we refer for further details.

[1] It can be shown that the initial condition is immaterial for our results.

Parameters D^{h_s}, $E_s^{h_s}$, $E_t^{h_t}$ and $0 < \alpha < 1$ in Proposition 1 have a simple interpretation in terms of the close-to-optimal policy (6). In particular, $E_s^{h_s}$,D^{h_s} are respectively the energy and the distortion that the EMU selects for the source encoder at the times k when the observation state is $H_{s,k} = h_s$, whereas $E_t^{h_t}$ is the energy selected for the channel encoder when the channel state is $H_{t,k} = h_t$. Note that *the source encoder parameters* D^{h_s}, $E_s^{h_s}$ *are adapted only to the source quality* $H_{s,k}$, while the *channel encoder parameter* $E_t^{h_t}$ *is adapted only the channel quality* $H_{t,k}$.

Regarding the constraints in (5), the first ensures stability of the data queue following Loynes theorem. The second and third constraints in (5) instead impose that the average energy used for the source encoder, $\sum_{h_s} p_{H_s}(h_s)E_s^{h_s}$, is no larger than a fraction $1-\alpha$ the average harvested energy $\mathbb{E}\left[E_{h,k}\right]$, and that the average energy used by the channel encoder, $\sum_{h_t} p_{H_t}(h_t)E_t^{h_t}$, is no larger than $\alpha\mathbb{E}\left[E_{h,k}\right]$. In practice, the optimal policy (6) may not be able to use always energies $E_s^{h_s}$, $E_t^{h_t}$ obtained from the optimization (4) due to ESS energy shortages. However, a strictly positive ϵ in (6) guarantees, as k gets large, the energy in the ESS will grow unbounded, thus allowing the EMU to use the optimal energy allocation dictated by $E_s^{h_s}$, $E_t^{h_t}$ from (4)-(5) (with the last two constraints modified as per Proposition 1). This shows the critical importance of leveraging an ESS in order to mitigate the random variations in energy availability due to harvesting, as already shown by previous work [1] by accounting only for transmission energy.

A. Suboptimal Policies

We have seen above that the optimal policy, in the sense of minimizing the long-term distortion $\bar{D}$ under data queue stability constraints makes a heavy use of the ESS, which is allowed to grow unbounded so as to provide an unlimited reservoir of energy and thus smooth out the variations in energy availability. It is thus interesting, for performance comparison, to consider policies that do not use the ESS, or use it only partially. The first suboptimal policy we consider does not use the energy buffer but allocates all the energy arrival $E_{h,k}$ to source and channel coding according to an optimized fraction $0 \leq \alpha^{h_s,h_t} \leq 1$ that depends *jointly* on both source and channel states, $H_{s,k} = h_s$ and $H_{t,k} = h_t$, as:

$$\begin{cases} D_k = D^{h_s,h_t} \\ E_{s,k} = \alpha^{h_s,h_t} E_{h,k} & \text{for } H_{s,k} = h_s \text{ and } H_{t,k} = h_t \ . \\ E_{t,k} = (1-\alpha^{h_s,h_t}) E_{h,k} \end{cases} \tag{7}$$

We refer to this policy as *no ESS with adaptation* (to both source and channel states). A special case whereby adaptation is not enabled (*no ESS with no adaptation*) is obtained by setting $\alpha^{h_s,h_t} = \alpha$. Alternatively, one could use the ESS *only for source or only for channel coding* (*source-only ESS* and *channel-only ESS*). For instance, for the source-only ESS strategy, a fraction of the incoming energy $E_{h,k}$ is used directly for channel coding, while the remaining part of the incoming energy is stored in the battery for possible use by the source encoder. This fraction is adapted to the state of channel $H_{t,k}$ (no gains are possible by adapting also to the state of the source $H_{s,k}$). Definition of these policies and analysis of all suboptimal policies mentioned above follow from [7], similar to Proposition 1 and will not be reported here.

IV. Delay-Distortion Optimization

In Sec. III, we have discussed the optimization of the policies π, as defined in Sec. II-A, with the objective of minimizing the long-term distortion $\bar{D}$, under the constraint of stabilizing the data queue. This criterion does not provide any guarantee on the delay experienced by the reconstruction of the source in a certain time-slot. In fact, in general, at the stability limit, the delay becomes arbitrarily large. This, along with the use of the ESS, allows the system to mitigate the variations in source and channel states and in the harvested energy. In this section, we briefly address a scenario with finite ESS and data queue and look for policies that minimize a weighted sum of distortion and delay. In particular, we propose to minimize the *expected total discounted cost*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \lambda^k \left[\gamma \mathbb{E}[D_k] + (1-\gamma)\,\mathbb{E}[X_k]\right], \tag{8}$$

where $0 \leq \lambda < 1$ is the discount factor and $0 \leq \gamma \leq 1$. The latter parameter weights the importance of distortion versus delay in the optimization criterion. Notice that with $\gamma = 0$ one minimizes the average length of the data queue, which, by Little's theorem, is the same as minimizing the average delay. We tackle the minimization of (8) over the policies π defined in Sec. II-A using dynamic programming tools assuming a finite ESS and data queue. Notice that, due to the finiteness of the ESS, buffer overflow may happen, in which case the compression bits are lost and a maximum distortion D_{max} is accrued for the current slot. Numerical results are discussed in the next section.

V. Numerical Results

At each block k, the transmitter observes M samples of an i.i.d. source $U_{k,i} \sim \mathcal{N}(0, D_{max})$ with $i = 1, ..., M$ over Gaussian noise with (source) Signal-to-Noise Ratio (SNR) $H_{s,k}$ as $\sqrt{H_{s,k}}U_{k,i} + Z_{k,i}$ where $Z_{k,i} \sim \mathcal{N}(0,1)$ is an i.i.d. sequence. Using the model in [4], the rate-distortion-energy function is given by

$$f^{h_s}(D_k, E_{s,k}) = \frac{1}{b}\log_2\left(\frac{D_{max}-D_{mmse}}{D_k - D_{mmse}}\right)\xi(E_{s,k}), \tag{9}$$

where the first term is the (indirect) rate-distortion function for this source in bit/c.u. with $D_{mmse} = (h_s + 1/D_{max})^{-1}$, while function $\xi(T_{s,k}) = \zeta \max\left[(bE_{s,k}/E_{s,max})^{-1/\eta}, 1\right]$ accounts for the loss due to energy limitations, with $\zeta > 1$, $\eta > 1$ and $E_{s,max}$ being design parameters. Communication takes place over an AWGN with channel SNR $H_{t,k}$ so that the number of transmitted bits is $g^{h_t}(E_t) = \log(1 + h_t E_t)$. The SNRs $H_{s,k}$ and $H_{t,k}$ can take two possible values in $\mathcal{H}_s = \mathcal{H}_t = \{1, 10\}$ independently and with equal probability.

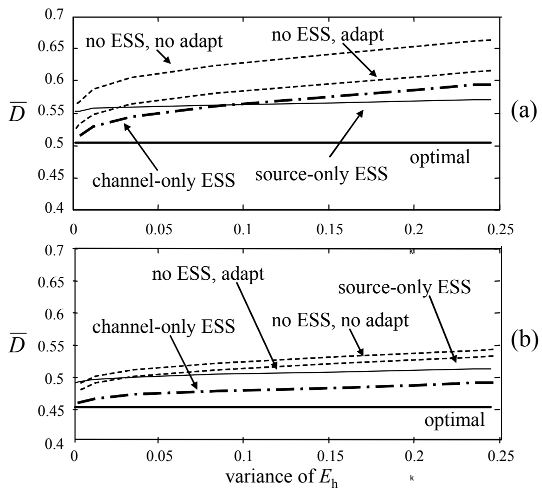

Figure 2. Long-term average distortion versus the variance of the harvesting process.

To assess the role of the ESS management, Fig. 2 shows the minimum distortion attained by the optimal policy and the suboptimal policies with no adaptation for increasing variance of the distribution $p_{E_h}(e)$ of the harvested energy, which is assumed to be Beta distributed with mean $\mathrm{E}[E_{h,k}] = 0.5$. We have parameters $E_{s,max} = 1$, $b = 1$ and $\eta = 3/2$ in Fig. 2-(a) and $\eta = 3$ in Fig. 2-(b). The advantages of the optimal policies with respect to the strategies that do not fully leverage the ESS are clear. Moreover, it is seen that, even using the ESS for either source or channel encoder leads to relevant gains. With $\eta = 3/2$ as in Fig. 2-(a), function (9) has a more pronounced convexity as a function of $E_{s,k}$ than in Fig. 2-(b). By Jensen's inequality, a more convex function (9) implies a larger performance loss in case the encoder is not able to operate at the average energy level. This can be seen by observing that performance loss of the channel-only ESS policy in Fig. 2-(a).

Another interesting aspect is the relative performance of joint or separate adaptation to source and channel SNRs. As seen, with full use of the ESS, a separate approach is optimal. Fig. 2 shows instead that the distortion achieved with no ESS but with adaptation to the joint state of both source and channel SNRs has relevant performance gains with respect to the policy with no adaptation, especially in the scenario of Fig. 2-(a) with a function (9) with more pronounced convexity.

With delay and/or ESS constraints, even the optimal policy needs to adapt to both source and channel SNRs. We study this issue by comparing the optimal result obtained from (8) with a separate policy that uses two distinct ESSs, one for the source encoder and one for the channel encoder, and designs both encoders assuming that the other provides a constant and optimized rate in each slot (see [7] for details). We assume data buffer length equal to 6, unitary battery capacity, harvesting distribution $p_{E_h}(e)$ with $e = 0.5$ or $e = 1$, $\lambda = 0.5$, the distortion takes values $D_k \in \{0.55, 0.75, 1\}$, and other parameters as in the example above. We show the the trade-

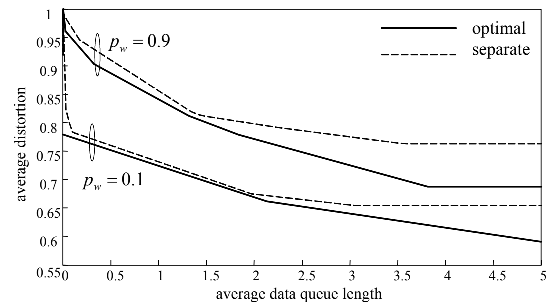

Figure 3. Trade-off between average distortion and average delay for the optimal policy and the policy that performs separate source-channel coding optimization.

off between average distortion $\mathbb{E}[D_k]$ and average queue length $\mathbb{E}[X_k]$, which accounts for delay, for two different values of the probabilities of the worst harvesting, source and channel states $p_{E_h}(0.5) = p_{H_t}(1) = p_{H_s}(1) = p_w$. The curve is obtained by varying parameter γ in (8). We observe that the joint source-channel optimal policy allows to obtain better delay-distortion trade-off compared to the separate policy.

VI. Conclusions

Cyber-physical systems, such as sensor networks, and green communications put a new emphasis on the task of energy management for wireless communications and favor the use of energy-harvesting technologies. In this paper, we have argued that the energy allocation between source and channel coding is a key aspect of the problem of energy management in energy-harvesting sensor networks. Our analysis sheds light on the optimal system design and on the impact of system parameters such as energy buffer size and average delay.

Acknowledgment

This work was supported by the Austria Science Fund (FWF) through grant NFN SISE (S10607) and by U.S. NSF under grant CCF-0914899.

References

[1] V. Sharma, U. Mukherji, V. Joseph and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," , *IEEE Tran. Wireless Commun.*, vol. 9, no. 4, pp. 1326-1336, April 2010.

[2] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," submitted [arXiv:1106.1595].

[3] K. C. Barr and K. Asanović, "Energy-aware lossless data compression," *ACM Transactions on Computer Systems*, vol. 24, no. 3, Aug. 2006.

[4] Z. He, and D. Wu, "Resource allocation and performance analysis of wireless video sensors," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 5, May 2006.

[5] X. Lu, E. Erkip, Y. Wang, and D. Goodman, "Power efficient multimedia communication over wireless channels," *IEEE Journ. Sel. Areas Commun*., vol. 21, no. 10, Dec. 2003.

[6] M. J. Neely and A. Sharma, "Dynamic data compression with distortion constraints for wireless communication over a fading channel," submitted [arXiv:0807.3768v1].

[7] P. Castiglione, O. Simeone, E. Erkip and T. Zemen, "Energy management policies for energy-neutral source-channel coding," submitted [arXiv:1103.4787].