Energy-Aware Compress and Forward systems for Wireless Monitoring of Time-Varying Fields

S. Savazzi, U. Spagnolini
Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy;
e-mail: (savazzi, spagnoli)@elet.polimi.it

Abstract—Dense monitoring of time-varying 2D field by Wireless Sensor Networks (WSN) needs to define new strategies to aggregate and encode data according to 2D sensor deployments. In this paper optimal design of communication protocol for WSN is conveniently cast for a set of linear sensor networks (sensor arrays) monitoring a correlated 2D time-varying field. Data is synchronously acquired by each sensor and routed towards the sink node for data fusion. Delay constraints are imposed by real-synchronously acquired by each sensor and routed towards the sink node for data fusion. Delay constraints are imposed by real-time data delivery (e.g., for tracking time variations). Sensors aggregate data hop-by-hop by compressing the new samples gathered from the sensor board and forwarding towards the next sensor in the route. Compression is based on a linear predictive encoding to exploit the correlation properties of the field. Cross-layer design of source, channel coding and medium access control are jointly analyzed for optimal resource allocation and interference management to minimize energy consumption. The proposed design approach is corroborated by extensive numerical analysis tailored for dense deployment as in seismic monitoring applications.

Index Terms—Signal Processing, Wireless Sensor Networks

I. INTRODUCTION

Pervasive monitoring through Wireless Sensor Network (WSN) systems requires sensors densely deployed over regular, quasi-regular or even fully random known positions to measure and collect samples of a correlated 2D field. The final goal is to obtain an estimate of the entire field with some tolerable distortion. The main design criteria for these systems are meant to lower the hardware complexity to facilitate low-power solutions that can be embedded into low-cost microprocessors, and to extend the lifetime of the network without penalizing the communication reliability [1].

Dense monitoring calls for a joint design of efficient data compression algorithms, low-power communication strategies at physical layer (PHY) and medium access control (MAC) protocols according to the specifics of the application, the statistical properties of the monitored field, and the propagation environment. In many cases, real-time monitoring is necessary to track time-varying field to identify any possible suspicious conditions that ask for a quick response with minimum delay. This additional requirement introduces severe limitations for network planning.

Motivation of this work is the 2D array geometry of sensors illustrated in Fig. 1. The 2D array geometry over regular \( d \times d \) grid is assumed to be decomposed into a sequence of linear arrays (sensor arrays) each containing \( N \) regularly deployed battery-powered sensors delivering data through multi-hop communication towards a common sink node (SN). Sink node is not energy limited as equipped with an external power supply. Even if fairly simple, this topology can be adopted to provide insights to the performance of multihop networks for several monitoring applications as acoustic sensor arrays, temperature monitoring, and seismic acquisition systems. Seismic exploration is one of the most relevant applications where the proposed analysis can be assessed: a short duration seismic pulse (source) is transmitted from the surface, reflected and diffracted from any sub-surface discontinuity and received by the 2D array of sensors densely placed on the surface. The correlated 2D field that is measured is the sub-surface reflectivity. The final goal is to characterize the geological structure of the sub-surface and its evolution over the time to identify new reservoirs [2], or to locate sites for CO\(_2\) sequestration applications [3]. Sensors are typically geophones and should fed back digital measurements within stringent delay constraints (real-time constraint). Range limitations of sensor radio devices force data to be fed back to the SN device(s) by multi-hop communication.
A. Related works and original contributions

Much recent works in cross layer design for wireless monitoring networks typically consider the joint optimization of PHY and MAC layers [4], without taking into account the impact of the source coding and interference for large scale systems. Impact of source coding on WSN routing was considered in [5], while the problem of optimizing the number of hops in linear topologies was taken into account in [6]. It was demonstrated therein that multihop transmission becomes advantageous (in terms of power efficiency) for low throughput applications thanks to the power gains that can counterbalance path loss and noise. Instead, multi-hop fails [6] when the aggregated data rate is high, due to interference limitations, packet losses and high power consumption.

The general cross-layer problem is formulated in this paper by jointly optimizing the encoding for (lossy) compression of the sensed data, the radio design (modulation and RF power), the medium access control (MAC) and the routing strategy. The multi-hop cooperative transmission approach is analyzed for each linear array out of the 2D array in Fig. 1. Along the 1D array, a number of bits are generated by each sensor after source encoding, these bits are aggregated to form a packet that is sequentially routed from source to destination through a series of hops. Source coding of information is based on the knowledge of the statistical properties of the random field samples. Each sensor compress and forward [7] its own measurements by encoding only the innovation introduced by the new samples compared to the predicted field [8], see Fig. 1-(b). At communication layer, variable-length time division access (TDMA) framing structure [4] is optimized to evaluate the optimal device duty cycle (i.e., transmission duration).

The paper is organized as follows. Sect. II gives an overview of the system model for performance optimization. In Sect. III it is proposed and analyzed a differential source coding architecture that exploits a joint linear prediction over the space and time domains (Space Time-Linear Predictive Coding, ST-LPC). Next, cross layer design of the communication system is evaluated in Sect. IV by focusing on 1D linear sensor array. The chosen optimization metric is the (per-frame) energy consumption. All the results valid for isolated 1D sensor array are set together in Sect. V by considering the interference arising from multiple lines of the 2D array for varying reuse factors. As a case study, in Sect. VI system design is tailored for seismic exploration by analyzing the performance of seismic data quality from a real seismic survey.

II. System model

In this section it is outlined the framework for system optimization. The basic network structure is the linear array of \( N \) sensors extracted from 2D array, these nodes are uniformly deployed toward the SN with inter-node spacing \( d \). Nodes communicating along any route of the linear array are labelled as \( i = 1, 2, 3, \ldots \) with inter-node distance \( Q \times d \) with \( Q = 1, 2, 3, \ldots \) indicating the routing strategy adopted, see Fig. 2 and motivation below). Wireless link between a pair of devices \((i, i+1)\) is impaired by frequency-flat fading with baseband complex-valued channel gain \( h_{i,i+1} \), the average power \( \mathbb{E}[|h_{i,i+1}|^2] = g_{i,i+1} = (Q \times d)^{-\nu} \) accounts for path loss with exponent \( \nu \). Information about the channel state \( h_{i,i+1} \) is assumed to be perfectly available at the receivers. Assuming that the interference contribution from co-channel interferers is negligible compared to background noise\(^1\), the instantaneous SNR for a direct transmission from device \( i \) to device \( i+1 \) is

\[
\gamma_{i,i+1}(P_i) = \frac{P_i}{N_0} |h_{i,i+1}|^2
\]

where \( P_i \) is the RF transmit power for device \( i \), while \( N_0 \) is the noise power term.

About the application framework, let \( s(x, y, t) \) be the 2D field in space \((x, y)\) that is in general time-varying \((t)\), the purpose of the analysis is to retrieve the field from a set of time-samples from a deployment of spatial measurements points over a rectangular grid. Sensors deployment depends on space and/or the time sampling conditions to recover the time-varying field [9]. Let the \( i \)-th sensor collect \( M \geq 2B_s \times T_s \) time samples (with \( B_s \) the signal bandwidth of interest), the data-set \( s_i = [s_{i,1}, \ldots, s_{i,M}] \) with \( s_{i,j} = s(x_i, y_j, t) \) for \( j = 1, \ldots, M \) contains the \( M \) time samples of the random field generated during an observation time window of \( T_s = t_M - t_0 \) sec. Data-sets \( \{s_i\}_{i=1}^N \) have to be forwarded periodically to the Sink Node (SN), while SN is in charge of aggregating the data to retrieve the field. A real-time constraint prescribes that all data-sets \( \{s_i\}_{i=1}^N \) acquired by each sensor during the \( T_s \) sec. time window must be forwarded to the SN within a delay constraint of \( T_{sec} \).

Sensor data model. Let us consider any routing path with inter-node distance \( Q \times d \) (see also Fig. 2). Samples of data traces \( s_i \) acquired by each sensor over the route are both spatially (across the sensor position) and temporally correlated according to an AR-1 model:

\[
s_i = \rho_s s_{i-1} + w_i, \quad \mathbb{E}[w_i w_i^T] = (1 - \rho_s^2) \sigma^2 I \tag{2a}
\]

\[
s_{i,j+1} = \rho_t s_{i,j} + n_j, \quad \mathbb{E}[n_j^2] = (1 - \rho_t^2) \sigma^2 \tag{2b}
\]

where \( \rho_s = \rho_s^2 \) is the correlation among the samples \( s_i \) and \( s_{i-1} \) measured by sensors \( i \) and \( i-1 \) (belonging to the same route) at distance \( Q \times d \), this depends on the correlation \( \rho_s \) among the samples generated by sensors at (reference) distance \( d \). \( \rho_t \) models the correlation over time\(^2\) between the local samples from sensor \( i \). For all samples it is \( \mathbb{E}[s_{i,j}^2] = \sigma^2 \), \( w_i \) and \( n_j \) are zero-mean Gaussian random variables. Specific values for correlations \((\rho_s, \rho_t)\) are application-dependent; a detailed analysis tailored for seismic applications is in Sect. VI.

Resource allocation. As illustrated in Fig. 1, time division multiple access (TDMA) is used to guarantee to each device an interference-free channel that can be used for propagating the aggregated measurements towards the SN. TDMA avoids collisions and idle listening that could cause energy waste in random access networks, mainly with dense nodes [10]. Transmission is organized into (logical) frames of length \( T \) consisting of a large number of time-slots of length \( \Delta \). The real-time constraint should guarantee the availability of at least

\(^1\)This assumption is consistent with the case of neighboring arrays operating on different bandwidths, see Sect. V.

\(^2\)Time-correlation model is assumed to be the same for each sensor of the array as it refers to the same stationary random field. Correlation over time depends on the sampling interval \( T_s/M \).
one time-slot to any sensors for periodically uploading the data-set on every \(T\) sec. Reserved transmission period for device \(i\) is \(\alpha_i T \geq \Delta\), where \(\alpha_i\) is the device duty cycle (the fraction of time the transmitter is active) such that \(0 \leq \alpha_i \leq 1\). The reserved transmission period \(\alpha_i T\) is optimally chosen by SN and translates into an integer number of reserved slots of length \(\Delta\). For analytical purposes we assume here that \(\Delta \ll T\) so that the duty cycle \(\alpha_i\) can take real values (even if in practice it is rounded to nearest integer of \(\alpha_i T/\Delta\)).

**Synchronization.** It is assumed that the SN provides the reference clock to all the \(N\) devices of the array through a periodic transmission of beacon slots of length \(\Delta_{\text{synch}}\) [10]. Even if this issue is not covered here, beacon slots can be multiplexed within the data stream to periodically recover clock offsets. Moreover, they can be also exploited to convey information about the routing policy and device duty cycles.

**Channel coding and power allocation.** An uncoded M-ary Quadrature Amplitude Modulation is adopted herein [4]. Device \(i\) is transmitting \(b_i\) bits during the reserved period of duration \(\alpha_i T\); for channel bandwidth \(B_w\), the required spectral efficiency \(r_i\) [bit/symbols] is

\[
r_i = \frac{R_i}{\alpha_i \log_2(N)} \quad \text{with} \quad R_i = \frac{b_i}{T \times B_w}.
\]

To simplify the reasoning, we assume that the rate \(r_i\) can take any value for \(r_i > 0\), although in practice the linear M-QAM modulation employed is constrained to a finite set of possible rates. For AWGN model, the Bit Error Rate (BER) \(P_{b_{i,i}}\), given the SNR \(\gamma_{i,i+1}\) in (1) scales as [11]

\[
P_{b_{i,i}} \leq 2 \times \exp\left[\frac{-3}{2} \times \frac{\gamma_{i,i+1}}{2^{\gamma_{i,i+1}} - 1}\right],
\]

this approximation is valid for \(\gamma_{i,i+1} < 30dB\). For a required BER level \(P_{b_{i,i}}\) the outage is declared when the instantaneous SNR due to the fading impairments is \(\gamma_{i,i+1} < (2^{\gamma_{i,i+1}} - 1)/K_i\) where, from (4), the gap is \(K_i = 3/(2 \times \ln(2 \times P_{b_{i,i}}))\). The outage probability for the link between device \(i\) and \(i + 1\) is \(P_{\text{out}} = 1 - \exp(-[N_0/P_i] \times (2^{\gamma_{i,i+1}} - 1)/K_i)\) while RF power \(P_i = P_i(P_{\text{out}})\) is chosen to achieve the required probability \(P_{\text{out}}\) [12].

**Routing policy for data aggregation.** Routing decisions are taken by the SN controlling the linear array. A routing path \(\mathcal{R}\) consists of a set of devices \(\mathcal{R} := \{i - Q, i, i + Q, \ldots\}\) that sequentially aggregate and forward the data to the SN. As shown in Fig. 2, the routing policy is defined here as a set of \(Q\) disjoint routes \(\{R_1, R_2, \ldots, R_Q\}\) with \(R_h \cap R_h = \emptyset \forall r \neq h\) such that all the devices can reach the SN without interfering with each other. Each \(r\)-th route \(R_h\) contains \(K_h\) devices where sensors are relabeled so that \(R_h = \{1, 2, \ldots, i, i + 1, \ldots K_h\}\) with mutual distance \(Q \times d\) and \(\sum_{r=1}^Q K_r = N\). Sensors of \(R_h\) have exclusive access to the medium for the time-fraction \(T/Q = \sum_{r=1}^Q \alpha_r T\). Both single path routing (\(Q = 1\) and interleaved routing cases (see Fig. 2) are competitive options in system design optimization.

**Source coding for compress and forward.** Let us consider any routing path \(R_h\); before transmitting during the reserved time \(\alpha_i T\), sensor \(i\) in \(R_h\) is receiving \(b_{i-1}\) bits from the previous device in the route \(R_h\) (sensor \(i - 1\)), adding (aggregating) the data received with its own information \(\Delta b_i\) to form a new sequence of bits of length \(b_i = b_{i-1} + \Delta b_i\). Notice that the specific values for bits \(b_i\) depend on the routing strategy \(R_h\) as it influences the spatial correlation properties of the encoded samples as \(\rho_c = \rho_c^2\). The incremental encoding (compression) of the new data set to compute \(\Delta b_i\) is based only on the bits \(\Delta b_{i-1}\) that encode the previous trace in the route \(R_h\), see Fig. 1-(b). This choice provides an acceptable trade-off between complexity and efficiency as it makes the compression strategy suitable for on-the-fly processing. With all the bits correctly received, the SN must provide a representation of the samples of the random field \(\{s_{\text{fl}}(n)\}_{n=1}^N\) with a predefined distortion \(D = \mathcal{D}\) so that \(\mathbb{E}[|s_{\text{fl}}(n) - \hat{s}_{\text{fl}}(n)|^2] \leq \mathcal{D}\), where \(\hat{s}_{\text{fl}}(n)\) is the discrete-valued representation of the sample.

**Energy consumption model.** For each frame, the energy absorbed by \(i\)-th node is

\[
E_i(\alpha_i, P_i) = \frac{TP_i}{\eta} + E_{\text{RX}}(\alpha_i + \gamma_{\text{synch}}/T) + E_{\mu P} + E_{\text{sc},i}
\]

where \(E_{\text{RX}}\) is the energy consumption for receiving, \(E_{\mu P}\) is the energy consumption for basic processing (i.e., phase-lock loop recovery, pulse-shaping, microcontroller and digital modulation). Power efficiency is \(\eta\) with \(\eta \leq 1\). \(E_{\text{sc},i} = E_{\text{sc}}(b_i)\) is the energy consumed for source encoding. In the following analysis it is assumed \(\eta = 1\) as it simply acts as a scaling factor for minimum energy, moreover consumption for receiving is larger than the energy spent for processing data, so that \(E_{\text{RX}} \gg E_{\mu P} + \max_i E_{\text{sc},i}\).

### III. Space-Time Linear Predictive Coding

The main focus of this section is to summarize the fundamental limits to distributed (lossy) data compression instrumental to the compress and forward transmission model illustrated in previous section.

To simplify the reasoning, assume that a stationary source \(i\) is generating random samples \(s_{i,j} \sim p_i(s_{i,j})\) with probability density function (pdf) \(p_i(s_{i,j}) = N(0, \sigma^2)\). According to AR-1 model (2), the encoding strategy for the new samples \(s_i\) (generated by sensor \(i\)) is based on the decoding of the discrete value set \(\hat{s}_{i-1}\) so that \(b_i = b_{i-1} + \Delta b_i(s_{i,j}|\hat{s}_{i-1})\), where \(\Delta b_i\) is the incremental set of coding bits relative only to the information added by sensor \(i\). Bits \(\Delta b_i\) provide a representation of the prediction error experienced after the optimal combination of the discrete value signals \(\hat{s}_{i-1}\) and the
samples \( s_i \). The number of bits \( b_i \) to represent the aggregated samples \( \{s_k\}_{k=1}^i \) with \( E[|s_i - \hat{s}_i|]^2 \leq D \) \( \forall k \leq i \) is

\[
b_i(D) = b_{i-1} + M \times I_i(D|b_{i-1}) = iM \times I_i(D|\rho_t, \rho_s),
\]

(6)

it depends on rate-distortion function \( I_i(D|b_{i-1}) = \Delta b_i(s_i|\hat{s}_i-1)/M \) [bit/sample] and thus on the temporal \( \rho_t \) and spatial \( \rho_s \) correlations.

The architecture is herein referred to as Space-Time Linear Predictive Coding (ST-LPC) and is suitable for the linear multi-hop network (depicted in Fig. 1) as it exploits the sample correlation of the field while routing the data. The predictive encoding quantizes (and transmit) the difference \( \Delta s_{i,j} = s_{i,j} - \hat{s}_{i,j} \) between the prediction \( \hat{s}_{i,j} \) (over space and time dimensions) and the true sample \( s_{i,j} \) generated by the sensing device. Quantized prediction error \( \Delta s_{i,j} = Q(\Delta s_{i,j}) \) after the quantizer non-linear operator \( Q(\cdot) \) is coded into binary digits using an entropy coder at \( I_i(D) = \mathcal{H}(\Delta \hat{s}_{i,j}) \approx \frac{1}{2} \log_2 \left( \frac{\pi e^{1/6}}{D} \times E[|\Delta \hat{s}_{i,j}|^2] \right) \),

(7)

where \( E[|\Delta \hat{s}_{i,j}|^2] \) is the predictor noise mean square error (MSE) for large \( M \). \( E[|\Delta \hat{s}_{i,j}|^2] = \text{var}(s_{i,j}|\hat{s}_{i,j-1}) \approx \sigma^2(1 - \rho^2) + \rho^2 \times D \) and \( E[|\Delta \hat{s}_{i,j}|^2] = \text{var}(s_{i,j}|s_{i,j-1}) \approx \sigma^2 \phi(\rho_s, \rho_t) + D \times \left| \Psi(\rho_s, \rho_t) + \Psi(\rho_t, \rho_s) \right| \). Here it is \( \phi(\rho_s, \rho_t) = (1 - \rho^2) \left( 1 - \rho^2 \right) \right) \left( 1 - \rho^2 \right)^{-1} \) and \( \Psi(\rho_s, \rho_t) = \rho^2 \left( 1 - \rho^2 \right)^2 \left( 1 - \rho^2 \right)^{-2} \). Proof is not given here due to limitations in space.

The number of bits \( b_i \) is found by substituting (7) into (6).

IV. CROSS-LAYER OPTIMIZATION

Focus here is the joint optimization of source encoding, PHY, MAC, and routing for the 1D linear array of \( N \) devices without any interference from neighboring arrays (see Sect. V for extension). Transmission quality and source encoding are constrained by the outage probability \( P_{\text{out}} \) and the distortion \( D \), respectively. Power \( (P_t) \), modulation \( (r_i) \), device duty cycle \( (\alpha_i) \) and routing \( (Q) \) are optimally allocated based on the channel statistics, the space-time correlation profile of the field \( (\rho_s, \rho_t) \) and the real-time constraint \( (T) \).

A. Optimal design for energy efficiency

Given \( P_{\text{out}} \) and \( D \), the sum-energy consumption is

\[
E_{\text{tot}} = \sum_{i=1}^{N} E_i(\alpha_i, P_t, Q | P_{\text{out}}, D) = \sum_{r=1}^{Q} E_{R_r} \]

(8)

where \( E_i \) is the energy consumed by the device \( i \) defined in (5) while \( E_{R_r} = \sum_{i \in R_r} E_i \) is the sum-energy consumption for all the subset of devices in the route \( R_r \). The energy expenditure of \( i \)-th device \( E_i \) is ruled by the routing policy \( Q \), the RF power \( (P_t) \) and duty cycle \( (\alpha_i) \). The RF power \( P_t = P_{i}(P_{\text{out}}, D, \alpha_i) \) is optimally allocated to achieve the outage probability \( P_{\text{out}} \).

The problem we now tackle is two-fold. In Sect. IV-B time fractions \( \alpha_i \) (device duty cycles) are optimized for a given routing strategy \( Q \) towards the SN (Problem 1), whereas in Sect. IV-C the optimal routing policy \( Q \) and time fractions \( \alpha_i \) are derived jointly (Problem 2).

B. Problem 1 - Optimization of device duty cycles \( (\alpha_i) \)

For any routing path \( R_r \) with \( K_r = N/Q \) devices the optimal time-fraction \( \hat{\alpha}_i \) \( \forall i \in R_r \) follows by minimizing the total energy:

\[
\hat{\alpha}_i(Q) = \arg \min_{\alpha_i} E_{R_r}(\alpha_i \mid Q) \quad \text{s.t.} \quad \sum_{i \in R_r} \alpha_i = 1/Q, \quad \forall i \in R_r, \quad \alpha_i > 0.
\]

(9)

Approximation 1 - For small enough \( R_i \) in (3)

\[
\hat{\alpha}_i \approx Q^{-1} \times R_i \sqrt{\beta_i} \times \left( \sum_{r=1}^{K_r} R_i \sqrt{\beta_i} \right)^{-1}
\]

(10)

with \( \beta_i = (K_i \times g_{i+1})^{-1} \). Proof is obtained by solving (9) with \( E_{R_r} \) approximated for small \( R_i \).

Notice that optimal duty cycles are a function of the routing strategy: factor \( Q \) rules both the time-period \( T/Q \) reserved for route \( R_r \) and the throughput \( R_i \) (3).

C. Problem 2 - Optimization of routing \( (Q) \)

By substituting the optimal duty cycles in (9) for route \( R_r \), \( \hat{\alpha}_i = \hat{\alpha}_i(Q) \) with \( i \in R_r \), the joint duty cycles and routing optimization reduces to the selection of the optimal number of routes \( Q \in \mathbb{N} \) such that

\[
Q = \arg \min_{Q \in \mathbb{N}} \sum_{r=1}^{R_r} E_{R_r}(Q \mid \hat{\alpha}_i(Q))
\]

(11)

\text{s.t.} \quad K_r = |R_r| = N/Q, \quad \forall r,

notation \(| \cdot | \) refers to the cardinality of \( R_r \). Optimal duty cycles follow as \( \hat{\alpha}_i = \hat{\alpha}_i(Q) \) \( \forall i \).

The following approximations provide further insights to routing optimization: it is assumed that for all nodes \( \Delta b_i = \Delta b \) (so that \( I_i = I \) \( \forall i \), moreover \( \beta_i = \beta', K_i = K \).

Approximation 2 - Optimal solution to (11) can be approximated for small enough \( R_i \) as

\[
\hat{Q} \approx \arg \min_{Q \in \mathbb{N}} |Q^\nu| \times \left( \frac{2^{D^R \times \hat{Q}^\nu}}{\hat{Q}^\nu + 1} - 1 \right)
\]

(12)

where \( \Delta b = \Delta b/(B_{\text{path}} T) \) while \( \nu \) is the path loss exponent.

Metric in (12) allows for an intuitive interpretation of Problem 2: first term \( Q^\nu \) accounts for the energy scaling due to larger distance \( Q \times d \) to be covered by route with interleaving factor \( Q \), on the other hand term \( 2^{D^R \times \hat{Q}^\nu} \) models the energy reduction due to the lower aggregated throughput that is experienced for large enough \( Q \). Optimal routing accounts for the trade-off between these two contrasting effects.

Approximation 3 - Optimal routing strategy solution to (11) with \( \hat{\alpha}_i(Q) = 1/N \) \( \forall i, Q \) can be approximated as

\[
\hat{Q} \approx \arg \min_{Q \in \mathbb{N}} \left[ \frac{2^{D^R \times N \hat{Q}^\nu}}{\hat{Q}^\nu + 1} - 1 \right] \times \frac{N}{Q}.
\]

(13)

Solution for \( Q \) in (13) provides an effective design rule for optimal routing in practical scenarios where duty cycles could not be optimized (see the analysis in Sect. VI).
A linear sub-network array as shown in Fig. 3, let from the specification reliability constraint assigned to each linear sub-network.

iii) interfering devices are using the same transmission power as all sensors at the same positions in every sub-network.

Mutual exchange of beacon slots and they are propagating a field with the same statistical properties (see Fig. 3).

Medium access design for interference coordination.

V. OPTIMIZATION IN INTERFERENCE-LIMITED SCENARIOS

The problem of efficiently assigning transmission resources for TDMA coordination [14] of multiple linear sensor arrays is addressed in this section according to the network topology depicted in Fig. 3. In the proposed scenario, arrays of sensors (linear sub-networks) are sharing the same bandwidth and are coordinated by time division (TDMA). Frame (of length T) is further divided into sub-frames of length T/\(I_{int}\) to be used by each linear array for delivering information in real-time to their respective SN. Sink nodes are synchronized by mutual exchange of beacon slots and they are propagating a common framing structure to coordinate the use of resources among the linear arrays by time reservations [15]. Linear sub-networks at distance \(d \times I_{int}\) are allowed to re-use the same fraction of the frame (but paying in terms of increased interference level). By choosing the reuse factor \(I_{int}\) large enough, the impact of interference from neighboring arrays is minimized at the price of a reduction of the available transmission resources (the length of the sub-frame \(T/I_{int}\)) assigned to each linear sub-network.

Let us consider the transmission of a sensor \(i\) belonging to a linear sub-network array as shown in Fig. 3, let \(J_i\) denotes the set of interfering devices belonging to neighboring linear sub-networks, furthermore let \(h_{k,i+1} \sim CN(0, g_{k,i+1})\) with \(g_{k,i+1} = d_{k,i+1}^{-\alpha}\) and \(k \in J_i\). It is assumed that each sub-network is delivering the same number of aggregated bits \(b_i\) as all sensors at the same positions in every sub-network are monitoring a field with the same statistical properties \((\rho_s, \rho_f)\). Further limiting assumptions to the analysis are drawn from the specific problem illustrated in Fig. 3: i) outage reliability constraint \(P_{out}\), required distortion \(D\) and fading statistics are the same for each linear sub-network; ii) co-channel interference is mainly originated by the closest pair of interferers that are carrying the same amount of traffic \(r_i\) so that \(SIR(Q, I_{int}) = g_{i,i+1} \times g_{k,i+1}^{-\alpha} \approx [1 + (P_{out}/Q)]^{\alpha/2}\) and iii) interfering devices are using the same transmission power \(P_k = P_t(\bar{P}_{out}, \bar{D}, \alpha)\) \(\forall k \in J_i\) while \(P_t\) is now computed to account for the increased interference level.

For any sensor belonging to the 2D array the duty cycles are first optimized given the reuse factor \(I_{int}\) and the routing configuration \(Q\) (Problem 1), next the routing strategy and reuse factor are optimized jointly (Problem 2).

A. Problem 1 - Optimization of device duty cycles (\(\alpha_i\))

For any choice of reuse factor \(I_{int}\) and for routing path \(R_{r}\) with interleaving factor \(Q\), the duty cycle optimization (9) is now restated as

\[
\hat{\alpha}_i(I_{int}, Q) = \arg \min_{\alpha_i} E_{R_r}(\alpha_i | Q, I_{int}) \quad \text{s.t.} \quad \alpha_i > \hat{\alpha}_i \forall i \in R_{r}, \sum_{i \in R_r} \alpha_i = 1/(Q \times I_{int}),
\]

constraint \(\sum_{i \in R_{r}} \alpha_i = (I_{int} \times Q)^{-1}\) now indicates that the available fraction of the frame allocated to devices belonging to the same route over one line is \(1/Q \times T/I_{int}\) (see Fig. 3).

Approximation 4 - For large SIR level \(SIR(Q, I_{int})\) the feasible set for duty cycles is

\[
\alpha_i > \hat{\alpha}_i \approx R_t \left[\log_2 \left(1 + \frac{K \bar{P}_{out} \times SIR(Q, I_{int})}{|\{i\}|}\right)\right]^{-1}\]

\(\forall i \in R_{r}\). Proof is not given here due to limitations in space.

Approximation 5 - For \(SIR(Q, I_{int}) \gg 2\bar{P}_{out}\) then \(\hat{\alpha}_i\) solution to (14) is

\[
\hat{\alpha}_i(Q, I_{int}) \approx \hat{\alpha}_i(Q)/I_{int} \quad \forall i
\]

with \(\hat{\alpha}_i(Q)\) in (10). Proof is trivial as (16) is a straightforward generalization of (10) now over the fraction \(T/(Q \times I_{int})\).

B. Problem 2 - Optimization of reuse (\(I_{int}\)) and routing (\(Q\))

By substituting the optimal duty cycles solution to Problem 1 in (14) \(\hat{\alpha}_i = \hat{\alpha}_i(I_{int}, Q)\), reuse factor \(I_{int}\) and routing strategy \(Q\) are jointly optimized as

\[
\hat{I}_{int}, \hat{Q} = \arg \min_{Q, I_{int} \in \mathbb{N}} \sum_{i} E_{R_r}[\alpha_i] \quad \text{s.t.} \quad SIR(Q, I_{int}) > \hat{S}(Q, I_{int}) \quad \forall r,
\]

optimal duty-cycles are found as \(\hat{\alpha}_i = \hat{\alpha}_i(\hat{I}_{int}, \hat{Q})\).

Approximation 6 - Feasible set \(SIR(Q, I_{int}) > \hat{S}\) can be restated for large SIR as

\[
T > T_{min}(Q, I_{int}) = \frac{B^{-1} \left[\sum_{i=1}^{N/Q} b_i\right] \times Q \times I_{int}}{\log_2 \left(1 + \frac{\bar{P}_{out}}{|\{i\}|} \times SIR(Q, I_{int})\right)}
\]

with \(T_{min}\) being a lower bound to delay in data delivery.

VI. WSN FOR SEISMIC EXPLORATION: A CASE STUDY

In seismic exploration one (or more) energy source(s) are placed on the surface to generate short duration seismic pulses (with excitation waveform \(w(t)\)) creating elastic waves that propagate over the sub-surface, see Fig. 1-(a). Back-scattered wavefield is measured by the sensor arrays placed on the surface: these are regularly (or quasi-regularly) deployed on the surface to form 2D arrays with typical spacings \(d = \ldots\)
per sensor with an average deployment of $10^4 - 10^5$ nodes [2]), the use of data compression techniques to reduce the large data-rate becomes mandatory. Propagation of elastic wavefield over deep geological structures makes signals to be severely attenuated, still carrying precious information from deep reflections. This means that attenuated signals (or low-level seismic signals) are severely impaired by noise, and this requires an high dynamic range in acquisition system. Conventional systems are cable-based and thus there is no real need to compress the 24 [bit/sample] coding routinely adopted. Purpose here is to prove that in multi-hop network, a meaningful reduction is feasible by accounting on each node for statistical properties of signals.

Compress and forward system is analyzed herein using actual seismic signals acquired from a land seismic acquisition: uncompressed seismic data (used as performance benchmark) has been acquired using 24 [bit/sample] with sample interval of 4ms, data trace length is $T_q = 5\text{ sec}$ so that the number of samples per trace is $M = 1250$ [samples/trace]. In the proposed experiment sensor arrays are composed by $N = 120$ devices (seismic channels) that are delivering data towards the SN by single-path routing ($Q = 1$). Top Fig. 4-(a) shows the grey scale image of the so called Common Shot Gather (CSG) with 24 [bit/sample]: vertical axes depict arrival travel times measured by one sensor during the time window $T_q$ (after the shot), while horizontal axis gathers all the traces $s_i$ recorded sequentially ($Q = 1$) from all the sensors belonging to one linear array. The same CSG is now acquired using the proposed ST linear predictive coding (ST-LPC, Sect. III) with $I_1 = 3$ [bit/sample]. The CSG is shown for comparison purpose in Fig. 4-(b) together with the observed quantization error in Fig. 4-(c) multiplied by a factor of 5 for visualization. The proposed ST-LPC is particularly suited for seismic signals as these exhibit high time correlation, moreover it allows low complexity implementation compared to Transform Coding (TC) approaches [16].

Application to actual experiment needs a number of adjustments to lossy encoding analysis in Sect. III. To cope with the highly non-stationary behavior of seismic signals over time, trace $s_i$ of $M$ samples is divided into a number of $J = 5$ (overlapped to avoid artifacts) sub-blocks, AR-1 parameters $(\rho_{i,j}^{(1)}, \rho_{i,j}^{(2)})$, $j = 1, ..., J$ are estimated over each block and used for prediction within the block. Seismic signals are amplified before quantization to compensate for attenuation of low-amplitude seismic events [9]. In same Fig. 4 (at bottom) the rate-distortion functions are evaluated using the correlations obtained from seismic data sets, rate-distortion results are also shown for a standard implementation of DPCM (without prediction over time or space) with and without compensation of attenuation. Theoretical rate-distortion function $I_1 = (1/J) \sum_{j=1}^{J} I(D|s_i^{(j)}, \rho_{i,j}^{(2)})$ with $I(D|s_i^{(j)}, \rho_{i,j}^{(2)})$ derived in (7) is also reported in dashed lines. For ST-LPC the use of linear prediction with $I_1 = 3$ [bits/sample] guarantees an (average) $SNR_q = 22\text{dB}$ for each trace. For DPCM only the difference between the current and the previous sample over time is encoded: for $I_1 = 3$ [bits/sample] $SNR_q$ scales down to $SNR_q = 17\text{dB}$ (and even lower $SNR_q = 14\text{dB}$ if removing compensation of attenuation). Based on the available data-sets, the influence of the routing strategy ($Q$) on the

Fig. 4. Top figure: (a) CSG with 24 [bit/sample]; (b) CSG encoded using ST-LPC with $I_1 = 3$ [bit/sample] and $Q = 1$; (c) observed quantization error multiplied by a factor of 5 for visualization. Bottom figure: simulated rate-distortion functions for ST-LPC with $Q = 1, 2$ and DPCM. Theoretical rate-distortion function is in dashed lines as reference. Data courtesy of eni. $5 \div 30\text{m}$. Each $i$-th sensor at position $(x_i, y_i)$ on the 2D grid spatially sample a correlated field

$$s(x_i, y_i, t) = w(t) * h(t|x_i, y_i) + n(t),$$

being the convolution between the excitaw waveform $w(t)$ and the reflectivity sequence $h(t|x_i, y_i)$ with $n(t)$ the noise term. Reflectivity is obtained after cumbersome seismic imaging processing [9] and it is the input data-set for estimating the sub-surface structure [13]. Each sensor (or geophone) collects $M$ time samples while the sampled data-set $s_i$ generated during an observation time window of $T_s \text{ sec.}$ is referred to as seismic trace. After the data is successfully received by each SN, a new shot pulse is generated from a new location, say every $10 - 60\text{ sec}$. Real-time constraint prescribes that the seismic data traces $s_i$ (from all the sensors) are forwarded within a delay constraint of $T_s$ sec that is typically comparable with data-trace duration $T_q$.

Numerical analysis is illustrated in this section for the case study of data acquisition in seismic exploration. Space-Time Linear Predictive Coding is adapted to seismic signals where rate-distortions on actual signals and acquisition deployments are compared with the theoretical analysis (Sect. III). Next, communication system design is carried out according to the analysis proposed in previous sections (Sect. IV-V).

A. Space-Time Linear Predictive Coding for seismic signals

Due to the large amount of data to be aggregated and transmitted during a seismic acquisition (approx. $20 \div 50\text{kbps}$ per sensor with an average deployment of $10^4 - 10^5$ nodes [2]), the use of data compression techniques to reduce the large data-rate becomes mandatory. Propagation of elastic wavefield over deep geological structures makes signals to be severely attenuated, still carrying precious information from deep reflections. This means that attenuated signals (or low-level seismic signals) are severely impaired by noise, and this requires an high dynamic range in acquisition system. Conventional systems are cable-based and thus there is no real need to compress the 24 [bit/sample] coding routinely adopted. Purpose here is to prove that in multi-hop network, a meaningful reduction is feasible by accounting on each node for statistical properties of signals.

Compress and forward system is analyzed herein using actual seismic signals acquired from a land seismic acquisition: uncompressed seismic data (used as performance benchmark) has been acquired using 24 [bit/sample] with sample interval of 4ms, data trace length is $T_q = 5\text{ sec}$ so that the number of samples per trace is $M = 1250$ [samples/trace]. In the proposed experiment sensor arrays are composed by $N = 120$ devices (seismic channels) that are delivering data towards the SN by single-path routing ($Q = 1$). Top Fig. 4-(a) shows the grey scale image of the so called Common Shot Gather (CSG) with 24 [bit/sample]: vertical axes depict arrival travel times measured by one sensor during the time window $T_q$ (after the shot), while horizontal axis gathers all the traces $s_i$ recorded sequentially ($Q = 1$) from all the sensors belonging to one linear array. The same CSG is now acquired using the proposed ST linear predictive coding (ST-LPC, Sect. III) with $I_1 = 3$ [bit/sample]. The CSG is shown for comparison purpose in Fig. 4-(b) together with the observed quantization error in Fig. 4-(c) multiplied by a factor of 5 for visualization. The proposed ST-LPC is particularly suited for seismic signals as these exhibit high time correlation, moreover it allows low complexity implementation compared to Transform Coding (TC) approaches [16].

Application to actual experiment needs a number of adjustments to lossy encoding analysis in Sect. III. To cope with the highly non-stationary behavior of seismic signals over time, trace $s_i$ of $M$ samples is divided into a number of $J = 5$ (overlapped to avoid artifacts) sub-blocks, AR-1 parameters $(\rho_{i,j}^{(1)}, \rho_{i,j}^{(2)})$, $j = 1, ..., J$ are estimated over each block and used for prediction within the block. Seismic signals are amplified before quantization to compensate for attenuation of low-amplitude seismic events [9]. In same Fig. 4 (at bottom) the rate-distortion functions are evaluated using the correlations obtained from seismic data sets, rate-distortion results are also shown for a standard implementation of DPCM (without prediction over time or space) with and without compensation of attenuation. Theoretical rate-distortion function $I_1 = (1/J) \sum_{j=1}^{J} I(D|s_i^{(j)}, \rho_{i,j}^{(2)})$ with $I(D|s_i^{(j)}, \rho_{i,j}^{(2)})$ derived in (7) is also reported in dashed lines. For ST-LPC the use of linear prediction with $I_1 = 3$ [bits/sample] guarantees an (average) $SNR_q = 22\text{dB}$ for each trace. For DPCM only the difference between the current and the previous sample over time is encoded: for $I_1 = 3$ [bits/sample] $SNR_q$ scales down to $SNR_q = 17\text{dB}$ (and even lower $SNR_q = 14\text{dB}$ if removing compensation of attenuation). Based on the available data-sets, the influence of the routing strategy ($Q$) on the
encoder performance is minimal and can be neglected for performance analysis so that $\Delta b_i \simeq \Delta b$. Notice that this conclusion is not valid in general for seismic applications as it largely depends on the acquisition geometry.

B. Communication system design

The communication system design focus on $5 \times 5m$ 2D array as typical setting for very high-density land survey. To corroborate analysis in Sect. IV for 1D linear array, it is assumed that interference is negligible (e.g., say that neighboring array as typical setting for very high-density land survey. To of

Each value of sum-throughput corresponds to a different delay constraint $T$ for data delivery. Transmission delay $T$ is ranging here from $T = 1$ sec. to $T = T_i$ = 5 sec and corresponds in this example to Rate $\simeq 0.09$ bit/s/Hz and Rate $\simeq 0.018$ bits/s/Hz, respectively. As benchmark case (A) shows the

energy required by a linear array without any optimization of duty cycle (i.e., $\alpha_i = 1/N$, $\forall i$) and single path routing connecting all the $N$ devices for $Q = 1$. The case for optimized duty cycles (using Approximation 1 in (10)) and single path routing is referred as case (B) while in case (C) routing only is optimized while the duty cycle is not. In the last case (D) routing policy $Q$ and duty cycles $\alpha_i$ are jointly optimized: this provides the largest lifetime uniformly over all the rate requirements. Optimal interleaving factors $Q$ follow Approximation 2 (for case D) and Approximation 3 (for case C). Routing optimization provides significant benefits for large rates (small delays) thus making the choice of routing strategy as crucial in high data-rate applications with stringent real-time constraints. For smaller rates (i.e., larger delays), routing optimization does not offer substantial gain compared to the case (B) where only duty cycle is optimized.

A more detailed comparative analysis between case (B) and case (C) is shown at the bottom of Fig. 5: the maximum bit-rate $\max_i B_q x_i = \max_i b_i/(\alpha_i T)$ [Mbit/s] and the maximum RF power $\max_i P_i(\alpha_i)$ [dBm] over the sensor array is evaluated for varying sum-throughput (or delay $T$). Both parameters are of particular interest as they influence the hardware choices for radio transceiver design. Optimizing the routing (case C) reduces the requirement on the maximum bit-rate by limiting the amount of data to be aggregated over each hop. The price is a larger transmission power to cover longer hops when $Q > 1$. On the other hand, optimal duty cycle optimization (case B) poses more stringent limitations (compared to case C) on device maximum supported bit-rate while it is more conservative with respect to the required maximum RF power.

Different routing schemes (with varying number of routes $Q = 1, 3, 10$) are analyzed for case C as a function of the network length $N$ in Fig. 6-(a) by evaluating the required $E_b/N_0$ for $P_{out} = 10^{-2}$. For the same scenario, in Fig. 6-
for $T$ a comparable link reliability with respect to interference-free $I$ the same frequency, Fig. 7 (on top) depicts the achievable
the maximum hop distance $d$ reuse factor $T$ different propagation environments (path-loss exponents $\nu$)
achievable sum-energy versus the reuse factor $I$ of signals between sensors close coupled to ground [17] as for
large $\nu$ exponent ($\nu = 2$) makes more inconvenient the long routes with small $Q$, on the contrary larger exponents ($\nu = 4$) limits
by penalizing routing choices with large $Q$. In all cases, larger network sizes ($N > 120$) would make routing optimization crucial for efficient design.
Assuming now that in 2D array multiple linear (1D) sensor arrays of $N = 120$ devices are operating simultaneously over the same frequency, Fig. 7 (on top) depicts the achievable
sum-energy scenario versus the delay constraint $T$ for various choices of reuse factor $I_{int}$ of the TDMA frame and for different propagation environments (path-loss exponents $\nu$ =
3, 4). Dot marks indicate the values for the delay lower-bound $T_{min}(Q, I_{int})$ obtained using Approximation 6 and (18) for $Q = Q = 1$ and $\sum_{i=1}^{N} b_i = N(N+1)/2$. To guarantee a comparable link reliability with respect to interference-free cases (in this examples $P_{out} = 10^{-3}$), the presence of interference from co-channel transmissions is counter-balanced by relaxing the real-time constraint $T$ used here as design parameter. Large path-loss exponents ($\nu = 4$) allows for a more efficiently exploitation of spatial reuse so that delivery delay is smaller if compared to milder fading scenarios ($\nu = 3$). Path loss exponent $\nu = 4$ typically models the propagation of signals between sensors close coupled to ground [17] as for seismic exploration applications. At bottom of Fig. 7 reuse factor $I_{int}$ is optimized for minimum sum-energy and for varying delay constraints and propagation environments. Duty cycles are computed by using Approximation 5 while optimal routing is found numerically as $\dot{Q} = 1$ for the analyzed network settings. Larger path loss exponents allow for larger SIR $SIR \approx (1 + I_{int}^2/Q^2)^{\nu/2}$ and smaller delays.

VII. CONCLUSIONS
In this work we evaluated the design problem for energy efficient WSN by considering a cross layer PHY-MAC optimization for linear arrays adopting data aggregation through compress and forward multi-hop transmission. An incremental source encoding is proposed based on the statistical properties of the monitored field over the space and time domains. Space-Time LPC is tailored for monitoring time-varying random fields over a linear sensor array. At communication layer, optimal time assignment (device duty cycle) between sensors, routing scheme and interference coordination among arrays are evaluated jointly to minimize the network energy consumption. Analysis is conveniently cast into the framework of convex optimization while energy consumption metric is computed for varying propagation scenarios, network deployments and statistical properties of the random field. Numerical analysis illustrates the relevant case study of seismic exploration.

REFERENCES