

# DISTRIBUTED COMPRESSION FOR CONDITION MONITORING OF WIND FARMS

*Shuang Wang, Samuel Cheng*

School of Electrical and Computer Engineering  
University of Oklahoma  
Tulsa, OK 74135-2512, USA

*Vladimir Stanković, Lina Stanković*

Dept of Electronic and Electrical Engineering  
University of Strathclyde  
Glasgow, UK

## ABSTRACT

In order to estimate the amount of energy that will be generated by a wind farm and provide efficient power distribution planning, it is necessary to deliver information of wind speed at all wind turbines. This paper proposes a scheme for compressing wind speed measurements exploiting both temporal and spatial correlation between the turbine readings via distributed source coding. The proposed scheme relies on a correlation model based on true measurements. A compression scheme proposed is of low encoding complexity and uses a particle-filtering based belief propagation decoder that adaptively estimates the nonstationary noise of the correlation model. Simulation results using realistic models show significant performance improvements compared to the scheme that does not dynamically refine correlation.

**Index Terms**— Wind Farms, Distributed Compression, Adaptive Decoding

## 1. INTRODUCTION

Wind energy generation depends on temporal wind speed, which varies in time and space. To enable power distribution system planning, precise and timely information of current wind speed is required. This information must be conveyed in near real time to a central substation from each of the wind turbine sites. After collecting the data from all wind turbines of a wind farm, the central substation can estimate temporal wind energy generation. There have been studies in the literature on modeling wind speed, but to the best of our knowledge, nothing has been reported on effective communications of the readings exploiting correlation between wind turbine measurements.

This paper focuses on effectively communicating wind speed information from multiple wind turbines of one wind farm to the central substation. The main communication challenges are related to the fact that the wind turbines are energy-constrained and wireless transmission channels are bandwidth limited. This calls for efficient low-complexity compression mechanisms that must operate in real time. Moreover, due to the proximity of neighboring wind turbines, proposed compression techniques should capture spatial correlation (besides temporal) of the signals measured without requiring communications between the turbines. We resolve the above challenges through the use of distributed source coding (DSC). DSC is an information-theoretical concept introduced in [1, 2, 3] that refers to separate compression and joint decompression of mutually correlated sources. DSC is particularly suitable for the wind turbine measurement compression due to power friendly encoding at the turbine and high coding efficiency due to exploitation of spatial correlation among turbine sites. Indeed, the effective DSC design can capture both temporal correlation between successive readings at a single

turbine site and spatial correlation between the readings of closely located sites.

A prerequisite to efficient DSC code design lies in estimating well the correlation between the measured data. Indeed, a usual assumption in the design of distributed compression codes is that all encoders and the decoder have perfect knowledge of the statistics of the measured data, i.e., the correlation noise among sources. However, often the correlation model is unknown, or the statistics of the sources vary over time. Indeed, in the targeted wind farm scenario, we are dealing with compression of non-stationary sources whose intensity and direction changes irregularly. This poses a huge challenge on the underlying compression scheme since correlation unpredictably varies over time. Recently, in [4, 5], particle filtering was integrated into the DSC decoding process to estimate and track correlation changes over time. The scheme maintains low encoding complexity, and introduces changes to conventional DSC coding only at the decoder side.

In [6], using a vector autoregression (VAR) approach, a spatio-temporal correlation model between wind speed readings in neighbouring wind turbines is proposed based on true measurements in Scotland. The model captures both the time varying nature of the wind speed at each site as well as the spatial correlation between the readings at neighboring sites. The additive correlation noise is in general nonstationary which is an additional challenge for the code design calling for effective correlation tracking.

In this paper we apply the information-theoretical concepts of [4] to distributed compression of wind speed information from neighbouring wind turbines. The proposed multiterminal source coding scheme [3] is of low encoding complexity, but still able to effectively exploit spatio-temporal correlation within measured data. The turbines compress their measurements using DSC and send the compressed data to the central substation, where the joint decoder estimates and tracks the change of correlation using particle-filtering based belief propagation. The scheme can be integrated in the IEC 61400-25 international standard for monitoring of wind power plants and control information exchange.

## 2. DISTRIBUTED SOURCE CODING (DSC)

DSC considers independent compression of correlated data, where correlation is exploited at the decoder side. Let  $X$  and  $Y$  be two correlated sources that should be compressed and sent to a central point for decoding. The compression must be done independently, that is,  $X$  and  $Y$  do not communicate, whereas decompression is joint. For discrete  $X$  and  $Y$  and lossless compression, in 1973, Slepian and Wolf [1] showed, surprisingly, that asymptotically, separate compression can be as efficient as joint compression, as long as  $X$  and  $Y$  are decompressed jointly. The DSC problem was introduced in [1]

and was followed by intense information-theoretical research into developing theoretical compression bounds and quantifying coding gains for similar setups.

There are two types of Slepian-Wolf coding setups investigated in the literature [7]. The first one is called non-asymmetric Slepian-Wolf coding, where both sources need to compress their information exploiting correlation. In the second setup, called asymmetric Slepian-Wolf (SW) coding or source coding with decoder side information, one source, say  $X$ , needs to be compressed assuming that another source,  $Y$ , called side information, is available, uncompressed, at the decoder as side information. An interesting extension of Slepian and Wolf's result is reported in 1976 by Wyner and Ziv [2] who considered a lossy version of the asymmetric SW coding. In the Wyner-Ziv (WZ) setup,  $X$  needs to be recovered under a certain distortion constraint and can be discrete or analogue. The extension of general, non-asymmetric SW coding to the lossy case is referred to as multiterminal (MT) source coding [3]. There are two types of MT source coding schemes. One, called direct MT source coding, where two correlated sources need to be compressed independently and recovered jointly under distortion criteria. Another type is indirect MT source coding, where a single source is observed by two (or more) sensors, where each sensor observes only a noisy version of the sources.

Motivated by the need for distributed compression in wireless sensor networks, in 1999, the first practical DSC scheme, or more precisely WZ coding scheme, based on channel codes appeared [8]. It was followed by many improved and more flexible solutions (see [9, 10] for reviews). Practical code designs for non-asymmetric SW coding and direct and indirect MT source coding are reported in [7, [11] and references therein.

The main assumption in the proof of SW and WZ theorems and in the majority of developed code designs, is the knowledge of statistics at the encoder and the decoder. That is, both encoder and decoder must know correlation between  $X$  and  $Y$  before coding takes place. In many cases this is an unrealistic assumption since correlation varies over time. In [4], a WZ coding scheme is proposed that unifies the process of online temporal correlation estimation and SW decoding into a single iterative process providing better statistics estimate and consequently improved performance. The correlation model, assumed in [4], is based on simple additive white Gaussian noise. The contributions of this paper are: (i) tracking of both spatial and temporal correlation statistics, (ii) a direct MT source codec design, (iii) a decoder based on particle-based belief propagation, whose algorithm operates on a factor graph where the correlation noise is modelled as [6], (iv) performance bounds of the proposed direct MT source coding scheme.

### 3. THE PROPOSED SCHEME

In the proposed scheme, the measurements from each wind turbine are first quantized, then independently SW-encoded for compression and possibly entropy coded for compression of side information that will be generated at the decoder. Both SW-encoded and entropy coded measurements are sent to the central substation for decoding. That is, closely located wind turbines are measuring wind speed, compressing the readings and sending the compressed data to the central substation for decoding. It is assumed that the relative position of the wind turbines is known and that their compression units are synchronized. For simplicity, we consider the case of two turbines only. In the case of multiple turbines, to keep complexity low, one would group them in pairs and independently perform coding on each pair.

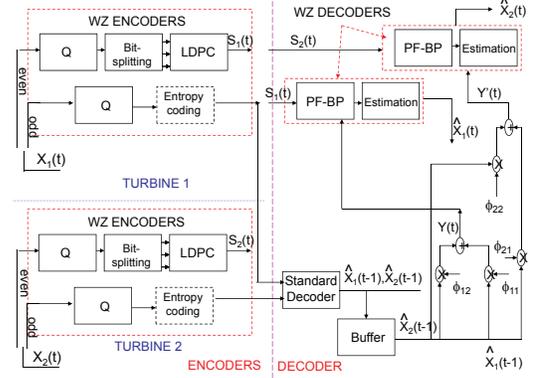


Fig. 1. The proposed scheme: encoder and decoder look.

Let  $X_1(t)$  and  $X_2(t)$  be the wind speed measured at wind turbines 1 and 2, respectively, at time  $t$ . Let  $\mathbf{X}(t) = [X_1(t) X_2(t)]^T$ . It was shown in [6] that a good model relating the measurements in the two turbines is given by:

$$\mathbf{X}(t) = \Phi_1 \mathbf{X}(t-1) + \Phi_2 \mathbf{X}(t-2) + \mathbf{n}(t), \quad (1)$$

where  $\mathbf{n}(t) = [n_1(t) n_2(t)]^T$  is white noise with  $n_1(t)$  and  $n_2(t)$  being nonstationary noises at turbines 1 and 2, respectively;  $\Phi_i$ ,  $i = 1, 2$  is a matrix that depends on the relative position of the turbines and is known at the encoder and decoder.

Each turbine needs to compress its readings and send the data to the central substation, which collects the data from both turbines before attempting to jointly decompress them. We assume that communication is always error free (via effective physical-layer protection) and focus on distributed compression next. For simplicity in the following, we assume that  $\Phi_2$  is a zero matrix, and attempt to exploit only first-order correlation dependency between the two random variables, that is, we simplify (1) to:

$$\mathbf{X}(t) = \Phi \mathbf{X}(t-1) + \mathbf{n}(t). \quad (2)$$

#### 3.1. The proposed solution

The proposed scheme is shown in Figure 1. Both turbines conventionally compress all odd measurements ( $X_1(2\tau-1)$ ,  $X_2(2\tau-1)$ ) using scalar quantization (Q) followed by entropy coding, such as Huffman coding. These measurements, after decompression, are used at the decoder as side information. All even measurements ( $X_1(2\tau)$ ,  $X_2(2\tau)$ ) at both turbines are compressed using DSC with scalar quantization followed by bitplane-by-bitplane low density parity check (LDPC) encoding for syndrome forming [4]. This way, syndrome vectors,  $S_1$  and  $S_2$  are formed, at turbine 1 and 2, respectively. Bitplane-by-bitplane compression enables more flexible compression rate selection since different bitplanes will be correlated in different ways.

The decoder first recovers odd measurements from both turbines as  $\hat{X}_1(2\tau-1)$  and  $\hat{X}_2(2\tau-1)$  using conventional decompression. Then, to recover  $X_1(2\tau)$ , the decoder employs SW decoding using the correlation channel model:

$$\hat{X}_1(2\tau) = \phi_{11} \hat{X}_1(2\tau-1) + \phi_{12} \hat{X}_2(2\tau-1) + n_1(2\tau), \quad (3)$$

with  $Y_1(\tau) = \phi_{11} \hat{X}_1(2\tau-1) + \phi_{12} \hat{X}_2(2\tau-1)$  as side information, which follows directly from (2).

Then,  $X_2(2\tau)$  is recovered as:

$$\hat{X}_2(2\tau) = \phi_{21}\hat{X}_1(2\tau - 1) + \phi_{22}\hat{X}_2(2\tau - 1) + n_2(2\tau), \quad (4)$$

using  $Y_2(\tau) = \phi_{21}\hat{X}_1(2\tau - 1) + \phi_{22}\hat{X}_2(2\tau - 1)$  as side information. Note that side information captures both spatial and temporal correlation. Two SW decodings are necessary, which can be performed in parallel. Note that the proposed scheme essentially employs direct MT source coding. It is possible to tradeoff the compression rates at the two turbines using time-sharing. Also, it is possible to design an asymmetric scheme where one turbine performs only DSC and the other only conventional compression.

Next, we estimate the required compression rate assuming two zero-mean Gaussian memoryless sources, ideal quantization, and use the mean squared error (MSE) distortion metric. Let the required rate to compress measurements of turbine 1 to achieve MSE  $D_1$  be:

$$R_{X_1}(D_1) = R_{odd1}(D_1) + R_{even1}(D_1), \quad (5)$$

where  $R_{odd1}(D_1)$  and  $R_{even1}(D_1)$  are the required rates for odd and even measurements, respectively.

Since conventional compression is done on even measurements, the required rate, at turbine 1, is:

$$R_{odd1}(D_1) \approx \frac{1}{2} \log^+ \frac{\sigma_{X_1}^2}{D_1}, \quad (6)$$

where  $\sigma_{X_1}^2$  is the variance of  $X_1(t)$  (see [9] and references therein).

To find  $R_{even1}(D_1)$ , note that the correlation channel is a Gaussian channel. Assuming that  $n_1(2\tau)$  is a Gaussian memoryless source independent of  $\hat{X}_1(2\tau - 1)$  and  $\hat{X}_2(2\tau - 1)$ , then  $R_{even1}(D_1)$  is:

$$R_{even1}(D_1) = \frac{1}{2} \log^+ \left( \frac{\sigma_{N_1}^2}{D_1} \right). \quad (7)$$

where  $\sigma_{N_1}^2$  is variance of  $n_1(t)$ . Here  $\log^+ p = \log p$  for  $p > 0$  or 0, otherwise. For derivation, see [9] and references therein. Similar sum-rate (5) expression can be derived for turbine 2.

### 3.1.1. Adaptive Decoding

The decoding procedure is an adapted version of [4]. The compressed stream (syndromes) is sent to a graph-based SW decoder, which uses the belief propagation (BP) algorithm with particle filtering (PF) to estimate current correlation noise  $n_1(t)$  and decompress the source. The PF-BP-based algorithm operates on a joint 3D factor graph that represents the probabilistic relationship among SW coding, bit-plane coding and correlation tracking - see Figure 2 and [4] for details of the factor graph construction and the PF-BP algorithm. These are mapped into appropriate variables nodes and factor nodes, where variables nodes denote unknown variables such as SW coded bits and correlation variance and factor nodes represent the algebraic connection among multiple variable nodes. In this paper, correlation variable nodes are modelled as Gaussian and vary slowly over time.

In the PF-BP algorithm, messages are passed iteratively between connected variable nodes and factor nodes in the different regions of the graph (region 1: bipartite graph connecting correlation variable and factor nodes, region 2: 2D SW factor-sub-graph representing SW code used for each bit-plane) until the algorithm converges or until a fixed number of iterations is reached. These messages (inferences or beliefs on source bits and correlation) will represent the influence that one variable has on another. Standard BP (the sum-product algorithm), generally used for SW LDPC decoding, can

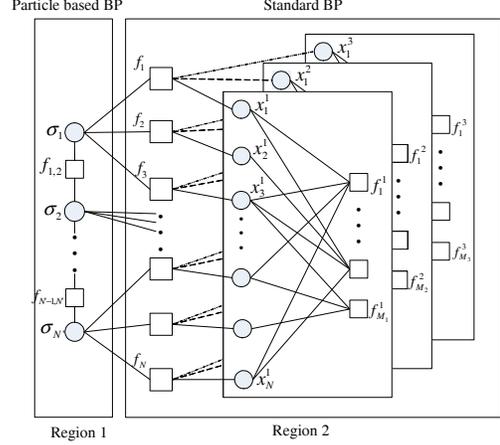


Fig. 2. Factor graph for the PF-BP-based adaptive decoder.

handle only discrete variables. The correlation variance, however, is not a discrete variable, since it varies continuously over time. We therefore resort to PF, which is integrated within the standard BP algorithm in order to handle continuous variables.

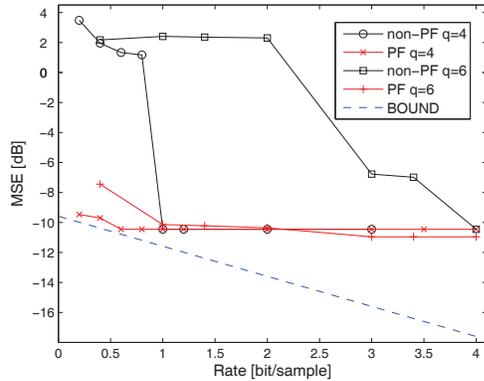
PF estimates the a posteriori probability distribution of the correlation variable node by sampling a list of random particles with associated weights. Systematic resampling is applied once all weights have been updated for all particles in each variable node to discard particles with negligible weight and concentrate on particles with larger weights. To maintain diversity of the particles, the new particle locations are perturbed by applying the random walk Metropolis-Hastings algorithm. The weight of each particle is then reset to a uniform weight for each particle. A new codeword is generated at the end of each iteration until the BP algorithm finds a valid codeword or until it reaches a maximum number of iterations.

## 4. SIMULATION RESULTS

In this section we report results of our simulations for the case of two neighboring turbines measuring wind speed and sending their compressed readings to the central substation. We show the results as coding/compression rate required for DSC versus MSE between the original samples and reconstructed ones. All results are averaged over 100 simulations, and MSE contains both “fine” quantization distortion as well as “coarse” distortion due to SW decoding errors [9]. In all our experiments we set matrix  $\Phi$  to contain all 0.5, which puts equal weight on spatial and temporal signal component. The code length is always 5000 symbols/samples, and each sample is quantized into  $q$  bits.

The number of particles tracking each correlation variable is set to 10 and the variance of the correlation between particles is  $\sigma_\gamma^2 = 0.07$ . These values were heuristically found to provide the best results. The maximum number of decoder iterations was set to 500 in case the PF-BP algorithm did not converge, and the scalar Lloyd-Max quantizer, trained using the first block of the decoder side information, is used for quantization. To keep the overall complexity low and for proof-of-concept, we use low-complexity regular LDPC codes with variable node degree 4 for SW coding [4, 5]. More complex irregular codes would result in improved overall performance. As a benchmark scheme, we use a DSC scheme that uses the same LDPC code and the same code length for compression, but exploits standard BP decoding without correlation tracking.

Figure 3 shows the obtained results as the required coding rate



**Fig. 3.** Code rate vs. MSE for  $q = 4$  and  $q = 6$  and  $\sigma^2 = 0.01$ .

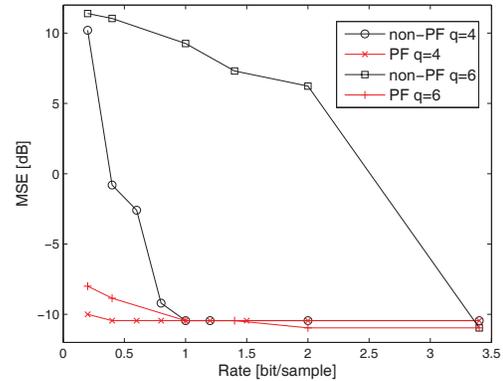
vs. MSE. Each sample is quantized using scalar quantization into either  $q = 4$  or  $q = 6$  bits, and SW coding is done bitplane-by-bitplane. Thus,  $q$  different LDPC codes are used. The number of syndrome bits per bitplane/layer is determined heuristically to minimize the residual bit error rate. We initialise and maintain the correlation noise unchanged as white Gaussian with variance  $\sigma^2 = 0.01$ .

It can be seen from the figure that the proposed PF scheme outperforms the benchmark scheme. This is despite the fact that the correlation noise statistics do not vary over time (similar conclusions are reported in [4]). As expected, at low rates, it is better to use  $q = 4$  quantization levels, whereas at the high rates,  $q = 6$  provides slightly better performance. At 1 bits/sample and  $q = 4$  the benchmark scheme reaches the performance of the PF scheme and the remaining noise is only the quantization noise. For  $q = 6$ , the proposed scheme shows error-free SW decoding performance at rate 2 bits/samples, while the benchmark scheme only at 4 bits/sample. The figure also shows the theoretical limit derived in the previous section. Note that the additional estimation step is applied after SW decoding. Since the resulting estimation gain diminishes as quantization step size decreases (see [9], [11]) a large gap to the bound occurs at higher rates.

The next figure shows results when the noise is white Gaussian with variance that follows Gaussian distribution and has mean of  $\sigma^2 = 0.01$ . That is, the variance slowly varies over time. As expected, the proposed PF-based scheme significantly outperforms the benchmark scheme. Thus, the proposed PF-based scheme successfully tracks the changes in the correlation and “adapts” BP decoding. Note that DSC theorems only hold when the source statistics is assumed to be constant. Therefore, no theoretical bound is shown in Figure 4. One can see that the trends are very similar to the previous example. A small gap between  $q = 4$  and  $q = 6$  PF curves for high rates is due to a lower quantization loss when 6 bits are used for quantization.

## 5. CONCLUSION

In this paper we proposed a scheme for compressing wind speed measurements in a wind farm. Wind speed between turbines provides important information necessary to estimate the amount of energy that can potentially be generated by the wind farm. The proposed solution exploits both temporal and spatial correlation between the turbine readings via distributed source coding. Moreover, the nonstationarity of the correlation model is taken care of with the particle-filtering based belief propagation at the decoder. The resulting scheme has low encoding complexity while being able to exploit



**Fig. 4.** Code rate vs. MSE for  $q = 4$  and  $q = 6$  and varying correlation noise.

correlation between the turbines and dynamically track the changes in correlation noise. Simulation results using realistic models show impressive performance improvements compared to the scheme that does not track correlation.

## 6. REFERENCES

- [1] D. Slepian and J.K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, July 1973.
- [2] A.D. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 1–10, Jan. 1976.
- [3] T. Berger, “Multiterminal source coding,” *The Information Theory Approach to Communications*, G. Longo, Ed., New York: Springer-Verlag, 1977.
- [4] S. Wang, L. Cui, and S. Cheng, “Adaptive Wyner-Ziv decoding using particle-based belief propagation,” in *Proc. IEEE GlobeCom 2010*, Miami, FL, Dec. 2010.
- [5] S. Wang, L. Cui, L. Stankovic, V. Stankovic, and S. Cheng, “Adaptive correlation estimation with particle filtering for distributed video coding,” *IEEE Trans. CSVT*, accepted.
- [6] K.R.W. Bell, D.C. Hill, D. McMillan, F. Li, R.W. Dunn, D.G. Infield, and G.W. Ault, “Spatio-temporal correlations of available wind power and impact on transmission power flows,” *2010 CIGRE Technical Exhibition*, Paris, France, Aug. 2010.
- [7] V. Stanković, A. Liveris, Z. Xiong, and C. Georghiadis, “On code design for the Slepian-Wolf problem and lossless multiterminal networks,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 1495–1507, April 2006.
- [8] S.S. Pradhan and K. Ramchandran, “Distributed source coding using syndromes (DISCUS): design and construction,” *Proc. Data Compression Conference (DCC)*, Snowbird, UT, March 1999.
- [9] Z. Xiong, A. Liveris, and S. Cheng, “Distributed source coding for sensor networks,” *IEEE Signal Processing Magazine*, vol. 21, pp. 80–94, Sept. 2004.
- [10] V. Stanković, L. Stanković, and S. Cheng, “Distributed source coding: Theory and applications,” in *Proc. Eusipco 2010 European Signal Proc. Conf.*, Aalborg, Denmark, Aug. 2010.
- [11] Y. Yang, V. Stanković, Z. Xiong, and W. Zhao, “On multiterminal source code design,” *IEEE Trans. Inform. Theory*, vol. 54, pp. 2278–2302, May 2008.