Goal-oriented Data Compression: Application to an Lp-norm type performance metric

Yifei SUN[1], Hang ZOU[1], Samson LASAULCE[2], Chao ZHANG[3], Michel KIEFFER[1]

[1]Université Paris-Saclay, CNRS, CentraleSupélec, L2S, F-91192 Gif-sur-Yvette, France

[2]Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France

[3]School of Computer Science and Engineering, Central South University, Changsha 410083, China

Résumé – Le nouveau paradigme adopté dans cet article consiste, dans un schéma de compression, à adapter l'étage de transformation linéaire et celui de quantification à l'utilisation finale du signal compressé. La métrique de performance finale retenue est une norme de type Lp qui apparait dans de nombreuses applications et notamment pour le problème fondamental d'ordonnancement de consommation électrique flexible (par exemple pour la charge de véhicules électriques). En appliquant l'approche proposée à des données réelles issues de la base de données Ausgrid, nous montrons que celle-ci permet d'avoir une perte d'optimalité due à la compression qui est bien moindre qu'un étage linéaire standard suivi d'un quantificateur uniforme.

Abstract – In this paper, a goal-oriented coding scheme comprising both precoding and quantization is considered in smart grids. Instead of searching a trade-off between compression ratio and distortion, our objective is to transmit the most relevant information to the decision-making entity under few bits constraint so that the optimality loss is minimized for a given utility function. Numerical results show that our approach could reduce the optimality loss tremendously compared to conventional compression approach, even when the target bitrate is extremely limited.

1 Introduction

The growing size of the electricity transport and distribution infrastructure leads to increasing needs to characterize its state. For that purpose, the states of the electricity network are measured by many sensors, generating a massive volume of data. Therefore, data compression is required to limit transmission and storage resource requirements [1].

Data compression techniques are either lossless or lossy [2]. For instance, for the target application of this paper, which is the problem of flexible power consumption in smart grids, [3] uses a lossless compression scheme for power quality data. Lossy compression achieves usually a higher compression ratio, at the price of some distortion introduced in the compressed data. The two first stages of a lossy compression scheme are typically a (linear) transform and a (uniform) quantizer. In [4] and [5], the discrete wavelet transform (DWT) is exploited for compressing power quality disturbance data.

The conventional data compression paradigm aims at reaching a tradeoff between the compression ratio and the resulting distortion. In the present paper, we revisit this paradigm by tailoring the compression stages to the final use of the compressed data. For that purpose, we rely on recent results developed in [6, 7, 8]. The concept of goal-oriented quantization has been introduced in [6, 7] and goal-oriented precoding schemes have been proposed in [8]. In the present paper, we consider a relevant utility function, which can be applied into the key problem such as the total energy consumption, the peak power and Joule losses minimization, and design a compression scheme which includes both the precoder and the quantizer and is adapted to the considered performance metric.

The paper is structured as follows. In Section 2, we introduce the coding scheme and formulate the problem to be solved. In particular, a utility function is considered to be maximized and an optimality loss, which need to be reduced, is introduced. To solve the considered problem, goal-oriented precoding and quantization schemes are proposed in Section 3. Section 4 shows the numerical results. Section 5 concludes this paper.

2 Problem formulation

Consider a decision-making entity with the following performance metric or utility function

$$u\left(x;\ell\right) = -||x+\ell||_p \tag{1}$$

where $x = [x_1, x_2, \ldots, x_P]^T \in \mathbb{R}_+^P$ corresponds to the decision to be made, $\ell = [\ell_1, \ell_2, \ldots, \ell_P]^T \in \mathbb{R}^P$ is a vector of non-controllable but observable parameters, and $||x||_p$ stands for the Lp-norm : $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$. In the case of a home power consumption scheduler (*e.g.*, for an electric vehicle), x would represent the flexible power consumption and ℓ can represent the non-flexible part. Assume that $E > 0$ indicates the total flexible power consumption budget. The optimization problem (OP) consisting in maximizing the utility with respect to x can be written as

$$\begin{aligned} & \underset{x}{\text{maximize}}\ u\left(x;\ell\right) \\ & \text{s.t.}\ \sum_{j=1}^{P} x_j - E = 0 \\ & \quad x_j \geqslant 0,\ j = 1, \ldots, P, \end{aligned} \tag{2}$$

and $x^{\star}\left(\ell\right)$ is one solution for this OP. Due to limitations in terms of volume of data that may be transmitted, the exact/actual value of ℓ may be typically unavailable to the decision maker. We consider a situation where a lossy compression technique is used to reduce the amount of data to be transmitted to the decision maker. The latter has then only access to an approximate version $\widehat{\ell} \in \mathbb{R}^P$ of ℓ. The decision is made based on $\widehat{\ell}$ while the utility function still experience the real parameter ℓ, leading to the *de facto* utility of the system expressed as

$$u(x^{\star}(\widehat{\ell});\ell) = -||x^{\star}(\widehat{\ell}) + \ell||_p. \tag{3}$$

One always has $u(x^{\star}(\widehat{\ell});\ell) \leqslant u\left(x^{\star}\left(\ell\right);\ell\right)$ due to the approximation of ℓ by $\widehat{\ell}$ which leads to an optimality loss.

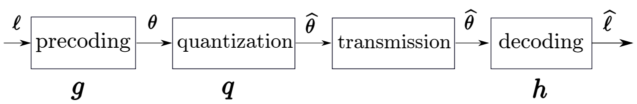

FIGURE 1 – Coding scheme

Figure 1 illustrates the coding scheme that we consider to minimize the expected optimality loss

$$\mathbb{E}_{\ell}\left[\left|u\left(x^{\star}\left(\ell\right);\ell\right) - u(x^{\star}(\widehat{\ell});\ell)\right|^2\right], \tag{4}$$

where the expectation is performed with respect to ℓ. The precoding function g maps a given parameter ℓ to its encoded version θ

$$\begin{array}{cccc} g: & \mathbb{R}_+^P & \rightarrow & \mathbb{R}^N \\ & \ell & \mapsto & \theta \end{array}$$

We always have $N \leqslant P$ to reduce the dimension of ℓ. Then θ quantized with

$$\begin{array}{cccc} q: & \mathbb{R}^N & \rightarrow & \mathbb{R}^N \\ & \theta & \mapsto & \widehat{\theta} \end{array}$$

We assume that $\widehat{\theta}$ is transmitted without error to the decision maker which can reconstruct the signal $\widehat{\ell}$ through a decoding function

$$\begin{array}{cccc} h: & \mathbb{R}^N & \rightarrow & \mathbb{R}^P \\ & \widehat{\theta} & \mapsto & \widehat{\ell} \end{array}$$

Our objective is to find a coding scheme (precoder, quantizer, and decoder) that minimizes the expected optimality loss

$$(g^{\star}, q^{\star}, h^{\star}) \in \arg\min_{(g,q,h)} \mathbb{E}_{\ell}\left[\left|u\left(x^{\star}\left(\ell\right);\ell\right) - u(x^{\star}(\widehat{\ell});\ell)\right|^2\right]. \tag{5}$$

3 Proposed solution

It is rather difficult to jointly find the precoder, quantizer, and decoder satisfying (5). We thus propose a suboptimal solution which optimize them separately.

3.1 Linear transformation

A signal can be projected to a chosen basis by linear transform. Such transform can be used to design a precoder and a decoder. For simplicity, we consider an linear transform without quantizer. Then $\theta = g\left(\ell\right) = \boldsymbol{B}\ell$, $\widehat{\ell} = h\left(\theta\right) = \boldsymbol{B}^T\boldsymbol{B}\ell$, where $\boldsymbol{B} \in \mathbb{R}^{N\times P}$. Then (4) becomes

$$E_{\ell}\left[\left|u\left(x^{\star}\left(\ell\right),\ell\right) - u(x^{\star}(\boldsymbol{B}^T\boldsymbol{B}\ell),\ell)\right|^2\right]. \tag{6}$$

Since $x^{\star}\left(\ell\right)$ is usually a nonlinear function of ℓ, by tayloring $x^{\star}\left(\ell\right)$, a piecewise linear approximation

$$x^{\star}\left(\ell\right) \approx \boldsymbol{H}\left(\ell\right)\ell + b\left(\ell\right). \tag{7}$$

of $x^{\star}\left(\ell\right)$ is considered. The way $\boldsymbol{H}\left(\ell\right)$ and $b\left(\ell\right)$ are evaluated is detailed in [8].

Considering a set $\mathcal{L} = \{\ell^{(1)}, \ell^{(2)}, \ldots, \ell^{(T)}\}$ of realizations of the parameter ℓ, one evaluates the empirical optimality loss

$$\begin{aligned} & \Gamma\left(\boldsymbol{B}\right) \\ = & \frac{1}{T}\sum_{i=1}^{T}\left|u(x^{\star}(\ell^{(i)}),\ell^{(i)}) - u(x^{\star}(\boldsymbol{B}^T\boldsymbol{B}\ell^{(i)}),\ell^{(i)})\right|^2 \\ = & \frac{1}{T}\sum_{i=1}^{T}\left|u(\boldsymbol{H}\left(\ell^{(i)}\right)\ell^{(i)} + b\left(\ell^{(i)}\right),\ell^{(i)}) - \right. \\ & \left. u(\boldsymbol{H}\left(\boldsymbol{B}^T\boldsymbol{B}\ell^{(i)}\right)\boldsymbol{B}^T\boldsymbol{B}\ell^{(i)} + b\left(\boldsymbol{B}^T\boldsymbol{B}\ell^{(i)}\right),\ell^{(i)})\right|^2. \end{aligned}$$

A value $\boldsymbol{B}^*$ of $\boldsymbol{B}$ minimizing $\Gamma\left(\boldsymbol{B}\right)$ can be found using a gradient descent, see Algorithm 1. Then, $\widehat{\ell}$ is obtained as $\widehat{\ell} = \boldsymbol{B}^{\star^T}\boldsymbol{B}^{\star}\ell$.

Algorithm 1 Gradient descent search of $\boldsymbol{B}^*$

Require: Initial matrix $\boldsymbol{B}$
Input : Learning rate ϵ
Output : Matrix $\boldsymbol{B}^{\star}$
  while $\text{iter}_{\max}$ not reached and optimality loss reduced more than 0.01% **do**
    Compute gradient : $\boldsymbol{G} \leftarrow \nabla_{\boldsymbol{B}}\Gamma\left(\boldsymbol{B}\right)$
    Apply update : $\boldsymbol{B} = \boldsymbol{B} - \epsilon\boldsymbol{G}$
  end while

A local minimum of $\Gamma\left(\boldsymbol{B}\right)$ can be found numerically, even if the considered problem is not convex.

3.2 Goal-oriented quantizer

In this part, we assume that the precoding function $g\left(\cdot\right)$ and decoding function $h\left(\cdot\right)$ are given, we focus on the goal-oriented quantizer. A quantizer partitions the space $\mathbb{R}^N$ of the encoded parameter $\theta = g(\ell)$ into several disjoint quantization regions $\mathcal{C}_1, \ldots, \mathcal{C}_M$, *i.e.*,

$$\forall i \neq j,\ \mathcal{C}_i \bigcap \mathcal{C}_j = \oslash \text{ and } \bigcup_{i=1}^{M} \mathcal{C}_i = \mathbb{R}^N,$$

with the quantization rule

$$q(\theta) = r_i \iff \theta \in \mathcal{C}_i, \tag{8}$$

where $\mathcal{R} \triangleq \{r_1, \ldots, r_M\}$ is the set of representatives of the set of quantization regions $\mathcal{C} \triangleq \{\mathcal{C}_1, \ldots, \mathcal{C}_M\}$ respectively. Different from conventional quantizers, in what follows, we search a pair $(\mathcal{R}^\star, \mathcal{C}^\star)$ that minimizes the expected optimality loss

$$\mathbb{E}_\ell\left[\left|u\left(x^\star(\ell);\ell\right) - u(x^\star(\widehat{\ell});\ell)\right|^2\right] =$$
$$\sum_{m=1}^{M} \int_{\ell \in \mathcal{L}_m} \left|u\left(x^\star(\ell);\ell\right) - u\left(x^\star\left(h\left(r_m\right)\right);\ell\right)\right|^2 \phi(\ell)\, \mathrm{d}\ell, \tag{9}$$

where $\phi(\ell)$ is the probability density function of ℓ and

$$\mathcal{L}_m \triangleq \{\ell \in \mathbb{R}^P | g(\ell) \in \mathcal{C}_m\}, 1 \leq m \leq M.$$

Again, finding jointly $(\mathcal{R}^\star, \mathcal{C}^\star)$ is not trivial, a practical algorithm is proposed which is similar to the decisional quantizer used in [7] :

1. Determine the optimal quantization region for a given $\mathcal{R}$

$$\mathcal{C}_m^\star = \{g(\ell) \in \mathbb{R}^N | \mathcal{E}(r_m;\ell) = \min_i \mathcal{E}(r_i;\ell)\} \tag{10}$$

with

$$\mathcal{E}(r_m;\ell) = \left|u\left(x^\star(\ell);\ell\right) - u\left(x^\star\left(h\left(r_m\right)\right);\ell\right)\right|^2 \tag{11}$$

2. Determine the set of optimal representatives for given quantization regions $\mathcal{C}$

$$r_m^\star \in \arg\min_r \int_{\ell \in \mathcal{L}_m} \mathcal{E}(r;\ell)\, \phi(\ell)\, \mathrm{d}\ell \tag{12}$$

Like Lloyd-Max algorithm [9], Algorithm 2 performs these two steps iteratively to find the optimal quantizer.

Algorithm 2 Goal-oriented quantizer design algorithm

Require: Utility function $u(x;\ell)$
Require: Initial $\mathcal{R}^{(0)} = \{r_1^{(0)}, \ldots, r_M^{(0)}\}$
Require: Initial $\mathcal{C}^{(0)} = \{\mathcal{C}_1^{(0)}, \ldots, \mathcal{C}_M^{(0)}\}$
Output : $(\mathcal{R}^\star, \mathcal{C}^\star)$
  while $i \leqslant \text{iter}_{\max}$ not reached and optimality loss reduced more than 0.01% **do**
    Update $\mathcal{C}_k^{(i)}$ from $\mathcal{R}^{(i-1)}$ using (10), $m = 1, \ldots, M$
    Update $r_m^{(i)}$ from $\mathcal{C}_m^{(i)}$ using (12), $m = 1, \ldots, M$
  end while

4 Numerical results

The following simulation results consider the energy consumption data set [10]. We compared the performance of a compression scheme implementing the proposed linear transformation and a uniform quantizer to a scheme implementing the proposed linear transformation and goal-oriented quantizer.

In the considered scenario, ℓ represents daily energy consumption of one user. The dataset is constructed by daily energy consumption of one user during one year. The data are measured each half hour and transmitted each day, consequently, $P = 48$ and $T = 365$. Set $N = 1$. We compare the proposed linear transformation with Karhunen-Loève Transformation (KLT), the optimal approximation in the sense of mean-square error [11].

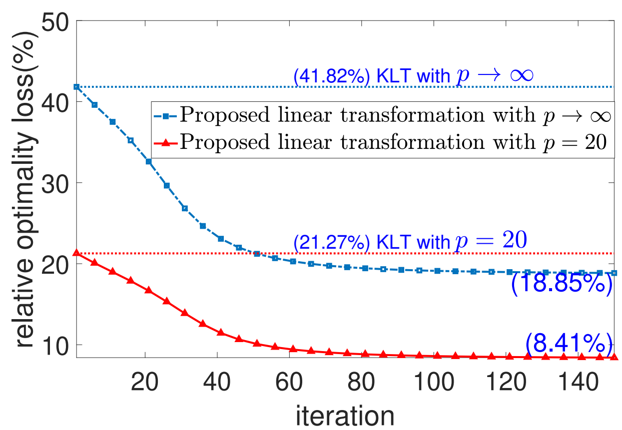

FIGURE 2 – Evolution of the relative optimality loss $(\%)$ as a function of the iteration of Algorithm 1 ; Proposed linear transform outperforms the KLT.

Figure 2 illustrates the evolution of relative optimality loss (ROL) as a function of iteration of Algorithm 1 for proposed linear transformation and for the KLT with different value of p in (1). The relative optimality loss is defined as

$$\frac{\sum_{i=1}^{T}\left|u(x^\star(\ell^{(i)}), \ell^{(i)}) - u(x^\star(\widehat{\ell^{(i)}}), \ell^{(i)})\right|^2}{\sum_{i=1}^{T}\left|u(x^\star(\ell^{(i)}), \ell^{(i)})\right|^2} \times 100\%.$$

One could observed that for both $p = 20$ and $p = \infty$, the proposed linear transform largely reduces the ROL compared to the KLT.

The goal-oriented quantization process is implemented after the precoding. The goal-oriented quantizer is compared to a dead-zone uniform quantizer. We assume that the quantized parameters are transmitted without error to simplify the problem. Figures 3 and 4 illustrate the performance of the two different quantizers when $p = 20$ and $p \to \infty$ as a function of number of bits used to represent the quantizer output.

In both Figure 3 and Figure 4, the proposed linear transformation with the goal-oriented quantizer performs better than when combined with the uniform quantizer in the sense of rela-

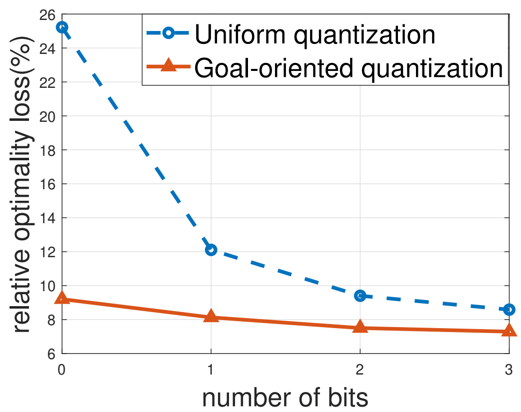

FIGURE 3 – Relative optimality loss v.s. the number of bits for goal-oriented quantizer and uniform quantizer when $p = 20$

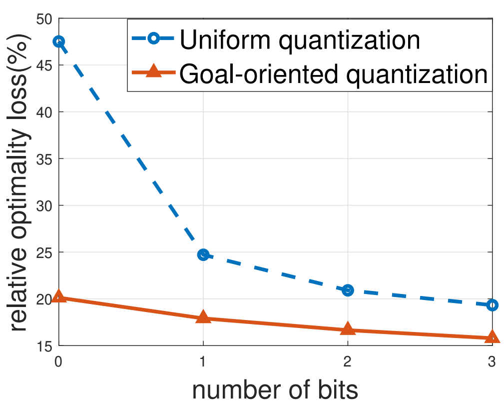

FIGURE 4 – Relative optimality loss v.s. the number of bits for goal-oriented quantizer and uniform quantizer when $p \to \infty$

tive optimality loss. Even with 0 bit, *i.e.*, when there is a single representative to optimize ($M = 1$), the goal-oriented quantizer provides a relatively small ROL which evidences the benefit of our approach.

5 Conclusions

In this paper, we design a goal-oriented compression scheme combining a precoder, a quantizer, and a decoder for the parameters of a parametrized Lp-norm-based utility function. For instance, such a utility function is well suited as a final performance metric in smart-grid applications. This function has to be minimized considering approximate parameter values due to compression. A utility loss is obtained compared to a situation where the actual parameter values are available. The joint design of the precoding and quantization stages, minimizing the expected utility loss, is generally complicated. Consequently, we have resorted to a separation assumption. Significant gains are observed when comparing the proposed goal-oriented linear transformation to the Karhuenen-Loeve transformation. Similarly, the goal-oriented quantizer yields very significant gains compared to a uniform quantizer. These results show the high potential of formulating the final task as an optimization problem and then tailoring the compression stages to this optimization problem.

Acknowledgements : This work was fully supported by the RTE-CentraleSupelec Chair. We acknowledge Prof. Sorin Olaru, Lucas Saludjian, and Patrick Panciatici for their support.

Références

[1] M. P. Tcheou et al., ”The Compression of Electric Signal Waveforms for Smart Grids : State of the Art and Future Trends,” in IEEE Trans. Smart Grid, vol. 5, no. 1, pp. 291-302, Jan. 2014

[2] K. Sayood, Introduction to Data Compression, 2nd ed. San Francisco, CA, USA : Morgan Kaufman, 2000.

[3] D. Zhang, Y. Bi and J. Zhao, ”A new data compression algorithm for power quality online monitoring,” In Proc. Int. Conf. on Sustainable Power Generation and Supply, 2009, pp. 1-4

[4] S. Santoso, E. J. Powers and W. M. Grady, ”Power quality disturbance data compression using wavelet transform methods,” in IEEE Trans. on Power Delivery, vol. 12, no. 3, pp. 1250-1257, July 1997

[5] J. Ning, J. Wang, W. Gao and C. Liu, ”A Wavelet-Based Data Compression Technique for Smart Grid,” in IEEE Trans. Smart Grid, vol. 2, no. 1, pp. 212-218, March 2011

[6] C. Zhang, N. Khalfet, S. Lasaulce, V. Varma and S. Tarbouriech, ”Payoff-oriented quantization and application to power control,” in Proc. Int. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2017.

[7] H. Zou, C. Zhang, S. Lasaulce, L. Saludjian and P. Panciatici, “Decision-Oriented Communications : Application to Energy-Efficient Resource Allocation”, in Proc. WINCOM (invited paper), Marrakesh, Morocco, 2018.

[8] Y. Sun, H. Zou, S. Lasaulce, M. Kieffer and L. Saludjian, ”A New Approach of Data Pre-processing for Data Compression in Smart Grids : Invited Paper,” In Proc. Int. Conf. on Wireless Networks and Mobile Communications, 2019, pp. 1-6

[9] S. Lloyd, ”Least squares quantization in PCM,” in IEEE Trans. Information Theory, vol. 28, no. 2, pp. 129-137, March 1982

[10] Ausgrid, “Solar home half-hour data - 1 july 2012 to 30 june 2013,” 2014, `https://www.ausgrid.com.au/Industry/Innovation-and-research/Data-to-share`.

[11] S. Mallat, “Chapter 9 - approximations in bases,” in *A Wavelet Tour of Signal Processing (Third Edition)*, pp. 435 – 480. Academic Press, Boston, 2009.