Goal-Oriented Quantization: Applications to Convex Cost Functions with Polyhedral Decision Space

Hang Zou*, Yifei Sun*, Chao Zhang†, Samson Lasaulce‡, Michel Kieffer* and Lucas Saludjian§
* Université Paris-Saclay, CNRS, CentraleSupélec, L2S, F-91192 Gif-sur-Yvette, France
†School of Computer Science and Engineering, Central South University, Changsha 410083, China
‡Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France
§RTE, 92073 Cedex La Defense, France

***Abstract*—In this paper, the situation in which a receiver has to execute a task from a quantized version of the information source of interest is considered. The task is modeled by the minimization problem of a general cost function $f(x;g)$ for which the decision x has to be taken from quantized parameters g. Especially, we focus on the particular scenario where the decision space is a convex polyhedron with cost function being convex. Furthermore, we propose a new goal-oriented quantization algorithm by combining the procedure of iteratively expanding and reinstating decision set together with Jensen's inequality. Proposed method could also be extended to some non-convex scenarios, namely, weakly convex cost function whose eigenvalues of Hessian matrix w.r.t decision x are lower-bounded. Numerical results show that proposed algorithm can considerably reduce the optimality loss (OL) compared to conventional approaches or the required number of quantization bits to achieve a certain relative optimality loss.**

I. Introduction

Most advances in wireless communications are deeply based on the idea "reproducing at one point either exactly or approximately a message selected at another point" proposed in Shannon's *locus classicus* on information theory [1]. Its concentration on merely the reliability requirement of communication (e.g., the bit error rate, the packet error rate, the outage probability) starts to show its limitations in meeting the ambitious goals set for the sixth generation (6G) system [2]–[4] in which the ignored semantic aspects of transmitted information are believed to bring significant gain. Besides classical paradigm could be highly inefficient in situations where communication systems are merely designed to execute a given task known both for transmitter and receiver. For instance, transmitting an image of 1 Mbyte to a receiver whose utility only depends on the absence/presence of a given object in the image might be extremely extravagant since the transmission of that particular one bit carrying the information of existence is sufficient to accomplish the task. This simple but insightful example motivates us to make a communication task- or goal-oriented (GO).

Different aspects of goal-oriented communication (GOC) are worth exploring and remain open. In this paper, we restrict our attention to the signal quantization problem being crucial for the design of a signal transmitter. The goal-oriented quantization (GOQ) problem is of great importance for many applications in various domains. First, it appears in controlled networks that are built on a communication network. For example, a smart meter may have to quantize or cluster the measured series for complexity or privacy reasons [5] in smart grid. Second, GOQ could also be formulated for some important wireless resources allocation (RA) problems, e.g., the limited-rate feedback quantization problem in [6]–[8] where transmitter allocates resources based on some quantized information from the receivers/sensors through a feedback channel. The idea of sending minimum number of bits to execute the task could be a competitive candidate for the upcoming 6G system to increase largely the spectral efficiency (SE). Third, there exist some early works on the problem of adapting the quantizer to the task. By combining the system task with the quantization process, [10] [11] investigated the influence of scalar quantization on specific tasks and characterized the limiting performance of recovering a lower dimensional linear transformation of the analog signal and reconstruction of quadratic function of received signals. Finally the conventional quantization paradigm in [9] could also be regarded as a special case of GOQ by taking the minimization of some distortion measures between the original signal and its representation as the system task. Therefore goal-oriented quantization is also a general framework for communication systems with or without being goal-oriented.

Nevertheless, the drawback of almost all existing works is that either the impact of quantization on a given performance metric is studied or a very specific performance metric is considered. In contrast with this line of research works, we introduce a general framework for GOQ. The task or goal of the receiver is chosen to be modeled by a generic optimization problem (OP) which contains both decision variables x and parameters g. One fundamental difference of this framework compared to existing works is that both for the performance analysis and the design, the goal function is a generic cost function $f(x;g)$, x being the decision to be made based on a quantized version of the parameters g. The goal-oriented quantization paradigm is illustrated in Fig. 1. Another possibility of making quantization goal-oriented is related to recent works on semantic communications [16]–[24] where semantics are employed here with its etymological meaning, that of significance. It can be seen as a measure of the usefulness/importance of messages with respect to the system task [16]. Reference [18] indicated that by properly recognizing and extracting the relevant information to the system task,

ISBN 978-3-903176-49-2 © 2022 IFIP

the communication efficiency and reliability can be enhanced without using more bandwidth. By introducing intrinsic states and extrinsic observations, [19] uses indirect rate-distortion theory to characterize the reconstruction error of semantic information induced by lossy source coding schemes. Learning tools have also been implemented to extract important attributes in semantic communications in [22]–[24].

The closest contributions to the present work have been produced by the authors through [26]–[29]. To the best of the authors knowledge, the concept of GOQ has been introduced for the first time in [26] and applied in other contexts in [27]–[29]. Numerical results are provided and the focus is on a Lloyd-Max (LM)-type algorithm [30] [31] in these references. Despite the fact that proposed goal-oriented quantizers could reduce the number of quantization bits tremendously compared to conventional goal-ignorant quantizer for some cost functions, the works supported by theoretical derivations are rare. It is observed that hardness of goal-oriented quantization varies from cost functions to cost functions. Therefore it is natural to study the impact of regularity properties of cost functions on GOQ problem, e.g., different types of generalized convexity (pseudo-convexity, quasi-convexity and weak convexity) or Lipschitz continuity. For instance, it is shown that a $\log$-type cost function is remarkably easier to quantize compared to a $\exp$-type cost function in [27]. To start with, in this paper we will focus on a special case of GOQ problem where the decision space is a convex polyhedron with cost function being convex. This type of GOQ problem is of great importance in the various domains, e.g., power allocation problems in [32], [33].

The rest of paper is organized as the following. In Sec. II, the goal-oriented quantization problem with convex cost function and convex polyhedral decision space is formulated. A new algorithm by iteratively expanding and reinstating the decision set is proposed and generalized to weakly convex cost functions as well in Sec. III. Numerical results are presented in Sec. IV to show the potential benefit of proposed method. Finally Sec. V concludes this paper.

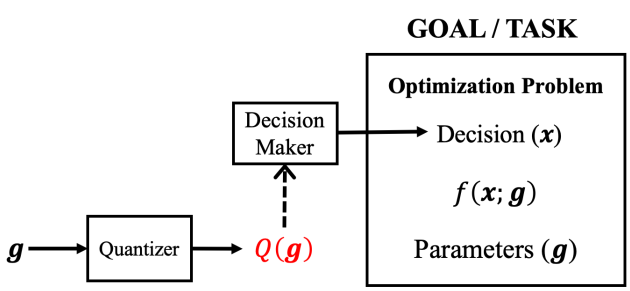

Fig. 1. Goal-oriented quantization paradigm

II. Problem Formulation

Definition II.1. *Let $M \geq 1$ be an integer. An $M-$quantizer $\mathcal{Q}_M$ is fully determined by a piecewise constant function $Q_M : \mathcal{G} \to \mathcal{G}$ that is defined by $Q_M(g) = z_m$ for all $z_m \in \mathcal{G}_m$ where: $m \in \{1, ..., M\}$, the sets $\mathcal{G}_1, ..., \mathcal{G}_M$ are called the quantization regions and define a partition of $\mathcal{G}$, and the points $z_1, ..., z_M$ are called the region representatives.*

Without loss of generality we assume that $\mathcal{G} \subset \mathbb{R}^{\ell_2}$ with ℓ_2 representing the dimension of the parameter space. Goal-oriented quantization consists in assuming that the task to be performed by the decision-making entity (e.g., the transmitter) can be represented by a standard optimization problem (OP), that is, a given function has to be minimized under some constraints. Therefore, the objective is to minimize a certain function or performance metric $f(x; g)$ (e.g., some cost or expense functions) with respect to the decision variable $x \in \mathcal{X} \subset \mathbb{R}^{\ell_1}$, where $\ell_1 \geq 1$ represents the dimension of the decision space. This mathematically writes as the following standard form OP:

$$\underset{x \in \mathcal{X}}{\operatorname{minimize}} \quad f(x; g) \tag{1}$$

Denote $\chi(g)$ an optimal solution of the above OP, i.e.,

$$\chi(g) \in \arg\min_{x \in \mathcal{X}} \quad f(x; g), \tag{2}$$

For a fixed probability density function (p.d.f.) $\phi(g)$ of the parameter g, to evaluate the performance of the goal-oriented quantization, we can assess the absolute optimality loss (OL) induced by quantization error of quantizer $\mathcal{Q}_M$ as follows:

$$\begin{aligned} \mathrm{L}(\mathcal{Q}_M) = \; & \mathbb{E}_g \left[f\left(\chi\left(\mathcal{Q}_M(g)\right); g\right) - f\left(\chi(g); g\right) \right] \\ = & \int_{g \in \mathcal{G}} \left[f\left(\chi\left(\mathcal{Q}_M(g)\right); g\right) - f\left(\chi(g); g\right) \right] \phi(g) \, \mathrm{d}g \end{aligned} \tag{3}$$

Furthermore, we denote $d_i \triangleq \chi(z_i)$ the decision for i-th quantization region with $\forall 1 \leq i \leq M$ and the decision set $\mathcal{D}_M \triangleq \{d_1, \ldots, d_M\}$ associated with quantizer $\mathcal{Q}_M$. To this end, the GOQ problem is to solve the OP formulated in Eq. 4, i.e., to find the optimal goal-oriented quantizer minimizing the OL:

$$\min_{\{d_m\},\{\mathcal{G}_m\}} \sum_{m=1}^{M} \int_{g \in \mathcal{G}_m} \left[f(d_m; g) - f(\chi(g); g) \right] \phi(g) \mathrm{d}g, \tag{4}$$

where the m-th quantization region is defined as:

$$\mathcal{G}_m = \left\{ g \in \mathbb{R}^{d_2} : f(d_m; g) \leq f(d_n; g), \; \forall n \neq m \right\} \tag{5}$$

This OP can be interpreted to jointly find the optimal quantization region with its corresponding decision (or its representative). Interestingly, it can be checked that the conventional quantization approach can be treated as a special case of the OP defined by (4) by specializing f as $f(x; g) = \|x - g\|^2$, $\|.\|$ standing for the Euclidean norm.

However, in many practical applications, it could be frequently met that the probability distribution of parameter is unknown or too difficult to obtain. In this paper, we assume only some realizations of parameters are available instead of its statistical model. Therefore we consider an empirical version of OL as our performance metric for a set of parameter samples $\mathcal{T} = \{g^{(t)}\}_{t=1}^{T}$:

$$\begin{aligned}&\overline{\mathrm{L}}\left(\mathcal{D}_M;\mathcal{T}\right)\\ =&\frac{1}{T}\sum_{t=1}^{T}\sum_{m=1}^{M}\left[f\left(d_m;g^{(t)}\right)-f\left(\chi\left(g^{(t)}\right);g^{(t)}\right)\right]\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_m\right\}}\end{aligned} \tag{6}$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function. Different from OP in (4) which could be equivalent to solve a NP-hard problem, one only needs to find the decision set minimizing the empirical OL since the corresponding quantization region for parameter samples could be obtained by exhaustive comparison between different decisions. Therefore, it is sufficient to solve the following OP taking empirical OL as our performance metric:

$$\min_{\mathcal{D}_M}\ \frac{1}{T}\sum_{t,m}\left[f\left(d_m;g^{(t)}\right)-f\left(\chi\left(g^{(t)}\right);g^{(t)}\right)\right]\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_m\right\}} \tag{7}$$

Additionally we assume some extra assumptions to make our GOQ problem more traceable:

1) decision space $\mathcal{X}$ is a **convex polyhedron** represented by a graph (V,E). $V=\{v_1,\dots,v_P\}$ with v_i is called the vertex of the polyhedron.
2) cost function is a **convex** function w.r.t. decision x.

III. Proposed Algorithm

We first present our general idea about how to find the optimal decision set minimizing the empirical OL. For a decision set $\mathcal{D}_M$, we first expand it to a larger decision set $\mathcal{D}_{M+N}$ with $N\in\mathbb{N}_+$. Then we select the optimal subset $\mathcal{D}_M^{'}$ of $\mathcal{D}_{M+N}$ which minimizes the empirical optimality loss. It is obvious that we always have $\overline{\mathrm{L}}\left(\mathcal{D}_M^{'};\mathcal{T}\right)\le\overline{\mathrm{L}}\left(\mathcal{D}_M;\mathcal{T}\right)$. Then the convergence of this method is guaranteed by repeatedly expanding and then selecting the optimal subset of decisions. Therefore it only remains to solve the problem of expanding a decision set $\mathcal{D}_M$ to $\mathcal{D}_{M+N}$ in an efficient way. For notation convention we denote the new expanded decision set as:

$$\begin{aligned}\mathcal{D}_{M+N}&=\{d_1,\dots,d_M,\zeta_1,\dots,\zeta_N\}\\&=\mathcal{D}_M\bigcup X_N,\end{aligned} \tag{8}$$

where $X_N=\{\zeta_1,\dots,\zeta_N\}$ is obviously the set of new decisions. Since the decision space is a convex polyhedron, each new decision ζ_n can be expressed as the convex combination of vertices: $\zeta_n=\sum_{i=1}^{P}\mathrm{A}_{in}v_i$ with $\sum_{i=1}^{P}\mathrm{A}_{in}=1,\forall 1\le n\le N$ with matrix $\mathrm{A}\in\mathbb{R}_+^{P\times N}$. Define the following partition of the parameter space $\mathcal{G}$:

$$\mathcal{G}_{mn}\triangleq\left\{g\,|\widehat{\chi}\left(g|\mathcal{D}_M\right)=d_m,\widehat{\chi}\left(g|\mathcal{D}_{M+N}\right)=\zeta_n\right\} \tag{9}$$

where

$$\widehat{\chi}\left(g|\mathcal{D}\right)\in\arg\min_{x\in\mathcal{D}}f\left(x;g\right) \tag{10}$$

The region $\mathcal{G}_{mn}$ is the set of all parameter so that its corresponding optimal decision switches from d_m to ζ_n once the decision set $\mathcal{D}_M$ is replaced by $\mathcal{D}_{M+N}$. For $\mathcal{D}_{M+N}$, one can easily have:

$$\begin{aligned}&\overline{\mathrm{L}}\left(\mathcal{D}_M;\mathcal{T}\right)-\overline{\mathrm{L}}\left(\mathcal{D}_{M+N};\mathcal{T}\right)\\=&\frac{1}{T}\sum_{t,m,n}\left[f\left(d_m;g^{(t)}\right)-f\left(\zeta_n;g^{(t)}\right)\right]\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_{mn}\right\}}\end{aligned} \tag{11}$$

Applying Jensen's inequality, one thus has

$$\begin{aligned}&\overline{\mathrm{L}}\left(\mathcal{D}_M;\mathcal{T}\right)-\overline{\mathrm{L}}\left(\mathcal{D}_{M+N};\mathcal{T}\right)\\=&\frac{1}{T}\sum_{t,m,n}\left[f\left(d_m;g^{(t)}\right)-f\left(\zeta_n;g^{(t)}\right)\right]\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_{mn}\right\}}\\=&\frac{1}{T}\sum_{t,m,n}\left[f\left(d_m;g^{(t)}\right)-f\left(\sum_i\mathrm{A}_{in}v_i;g^{(t)}\right)\right]\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_{mn}\right\}}\\\ge&\frac{1}{T}\sum_{t,m,n}\left[f\left(d_m;g^{(t)}\right)-\sum_i\mathrm{A}_{in}f\left(v_i;g^{(t)}\right)\right]\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_{mn}\right\}}\\=&\frac{1}{T}\left[\sum_{t,m,n}f\left(d_m;g^{(t)}\right)-\sum_{t,m,n,i}\mathrm{A}_{in}f\left(v_i;g^{(t)}\right)\right]\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_{mn}\right\}}\end{aligned} \tag{12}$$

Introduce matrix $\mathrm{W}\in\mathbb{R}^{T\times M}$ with $\mathrm{W}_{tm}=f\left(d_m;g^{(t)}\right)$ and a constant matrix $\mathrm{U}\in\mathbb{R}^{T\times P}$ with $\mathrm{U}_{ti}=f\left(v_i;g^{(t)}\right)$. We further define a tensor function $\mathrm{B}\left(\mathrm{A}\right)\in\mathbb{R}^{T\times M\times N}$ with $\mathrm{B}_{tmn}\left(\mathrm{A}\right)=\mathbf{1}_{\left\{g^{(t)}\in\mathcal{G}_{mn}\right\}}$. By writing $\mathrm{A}=\left(\mathrm{A}_1,\dots,\mathrm{A}_M\right)$, $\mathrm{B}=\left(\mathrm{B}_{11},\dots,\mathrm{B}_{MN}\right)$ and $\mathrm{U}=\left(\mathrm{U}_1,\dots,\mathrm{U}_P\right)$ in form of columns, the optimality loss decay $\Delta\overline{\mathrm{L}}\left(\mathrm{A};\mathcal{D}_M\right)$ by expanding $\mathcal{D}_M$ to $\mathcal{D}_{M+N}$ with X_N representing by A could be lower bounded:

$$\begin{aligned}&\Delta\overline{\mathrm{L}}\left(\mathrm{A};\mathcal{D}_M\right)\\=&\ \overline{\mathrm{L}}\left(\mathcal{Q}_M;\mathcal{T}\right)-\overline{\mathrm{L}}\left(\mathcal{Q}_{M+N};\mathcal{T}\right)\\\ge&\ \underbrace{\frac{1}{T}\sum_{n,m}\mathrm{W}_m^{\mathrm{T}}\mathrm{B}_{mn}\left(\mathrm{A}\right)-\frac{1}{T}\sum_{i,n}\left[\mathrm{A}_{in}\sum_m\mathrm{U}_i^{\mathrm{T}}\mathrm{B}_{mn}\left(\mathrm{A}\right)\right]}_{\Delta\widehat{\mathrm{L}}\left(\mathrm{A};\mathcal{D}_M\right)}\end{aligned} \tag{13}$$

It is clear that one should make $\Delta\widehat{\mathrm{L}}\left(\mathrm{A};\mathcal{D}_M\right)$ as large as possible to expand the decision set efficiently. The search space for A is the set of N-fold unit simplex of dimension P denoted as $\triangle_P^N$:

$$\triangle_P^N\triangleq\left\{\mathrm{A}_{in}:\sum_{i=1}^{P}\mathrm{A}_{in}=1,\forall 1\le n\le N\right\} \tag{14}$$

Before explain how to maximize $\Delta\widehat{\mathrm{L}}\left(\mathrm{A};\mathcal{D}_M\right)$, we introduce an important conception for matrix A:

Definition III.1. *(Equivalent Relation) For a given parameter sample set $\mathcal{T}$ and $A,A'\in\triangle_P^N$, we denote $A'\sim_{\mathcal{T}}A$ if and only if $B\left(A\right)=B\left(A'\right)$; otherwise we denote $A\nsim A'$. The equivalent class of A is denoted as $[A]_{\mathcal{T}}$.*

One can easily prove the defined operation $\sim_{\mathcal{T}}$ is an equivalent relation for any parameter sample set $\mathcal{T}$. Two decision sets (matrix) are equivalent for a given parameter sample set $\mathcal{T}$ means that their images under mapping B are exactly the same. If the parameter sample set is unchanging, we will omit it for both equivalent operator and equivalent class to simplify the notation. One can easily prove that

$$\begin{aligned}&\Delta\widehat{\mathrm{L}}\left(\mathrm{C};\mathcal{D}_M\right)-\Delta\widehat{\mathrm{L}}\left(\mathrm{A};\mathcal{D}_M\right)\\ =&\frac{1}{T}\sum_{i,n}\left[\left(\mathrm{A}_{in}-\mathrm{C}_{in}\right)\sum_m \mathrm{U}_i^{\mathrm{T}}\mathrm{B}_{mn}\left(\mathrm{A}\right)\right],\ \text{for } \forall\mathrm{C}\sim\mathrm{A}. \quad (15)\end{aligned}$$

The meaning of Eq. 15 is that $\Delta\widehat{\mathrm{L}}\left(\cdot\,;\mathcal{D}_M\right)$ is a simple linear function within a given equivalent class [A] for tensor B being invariant. Therefore maximizing $\Delta\widehat{\mathrm{L}}\left(\cdot\,;\mathcal{D}_M\right)$ is relatively easy in each equivalent class of matrix A. Besides, we will show that it is sufficient to consider singleton-expansion for decision set. For a matrix A within a given equivalent class, we denote $y=\left[y_1,\dots y_P\right]^{\mathrm{T}}\in\mathbb{R}^{P\times 1}$ with $y_i=\sum_m \mathrm{U}_i^{\mathrm{T}}\mathrm{B}_{mn}\left(\mathrm{A}\right)$, then one has

$$\Delta\widehat{\mathrm{L}}\left(\mathrm{A};\mathcal{D}_M\right)=\frac{1}{T}\sum_{n,m}\mathrm{W}_m^{\mathrm{T}}\mathrm{B}_{mn}\left(\mathrm{A}\right)-\frac{1}{T}\sum_i y_i\left(\sum_n \mathrm{A}_{in}\right) \quad (16)$$

By introducing a vector $\alpha=\left[\alpha_1,\dots,\alpha_P\right]^{\mathrm{T}}$ with $\alpha_i=\frac{1}{N}\sum_n \mathrm{A}_{in}$, one can easily find that the decision represented by α plays the same role as N decisions represented by matrix A. The reason for this degeneration is due to the independence of new decisions maximizing $\Delta\widehat{\mathrm{L}}\left(\cdot\,;\mathcal{D}_M\right)$. Thus we will only consider single-decision expansion for decision set in the rest of paper. Therefore we use vector α to represent the single new decision ζ_1 instead of matrix A. With a little abuse of notation, tensor B degenerates to a matrix defined as $\mathrm{B}\left(\alpha\right)\in\mathbb{R}^{T\times M}$ with $\mathrm{B}_{tm}\left(\alpha\right)=\mathbf{1}\left\{g^{(t)}\in\mathcal{G}_{m1}\right\}$. And we have

$$\begin{aligned}&\Delta\widehat{\mathrm{L}}\left(\alpha;\mathcal{D}_M\right)\\ =&\frac{1}{T}\sum_m \mathrm{W}_m^{\mathrm{T}}\mathrm{B}_m\left(\alpha\right)-\frac{1}{T}\sum_i\alpha_i\sum_m \mathrm{U}_i^{\mathrm{T}}\mathrm{B}_m\left(\alpha\right) \quad (17)\end{aligned}$$

Even we have shown that singleton expansion is sufficient, it is still cumbersome to maximize $\Delta\widehat{\mathrm{L}}\left(\alpha;\mathcal{D}_M\right)$ globally for two reasons. First, it ts obvious that the number of equivalent class for α depends on the number of parameter samples T. In worst case, we could have $|[\alpha]|=2^T$ which entails that the direct exhaustive search for matrix B leads to an exponential complexity of $O\left(2^T\right)$. Second, the optimal solution within equivalent class is always on the boundary of equivalent class which is hard to determine or at least implicitly given. So instead of directly searching the optimal solution, we introduce the conception of greedy improvement to give a sub-optimal direction.

Definition III.2. *(Greedy Improvement) Define* $\alpha^{\dagger}$ *for* α *with* $y_i=\sum_m U_i^{\mathrm{T}}B_m\left(\alpha\right)$ *s.t.* $\alpha^{\dagger}-\alpha=\nu e_{i^\star}$ *with* $\nu=\sup\left\{c\geq 0\,|\alpha\sim\alpha+ce_{i^\star}\right\}$ *and* $i^\star\in\arg\min_i y_i$*;* e_i *is the* P*-tuple with all components equal to 0, except the* i*-th.*

The meaning of greedy improvement is that this improvement minimizes $\Delta\widehat{\mathrm{L}}\left(\alpha;\mathcal{T}\right)$ along the deepest-gradient-descent direction within equivalent class of α. However, it is still cumbersome to determine the exact value of ν since the boundary of equivalent is generally implicitly given. Fortunately, this fact never stops us to construct an algorithm to find the solution for OP in (7). The basic idea is that if α is over-improved, i.e., $\exists\nu'>\nu$ and $\alpha'=\alpha+\nu\mathrm{e}_{i^\star}$ s.t. $\Delta\widehat{\mathrm{L}}\left(\alpha';\mathcal{T}\right)<\Delta\widehat{\mathrm{L}}\left(\alpha;\mathcal{T}\right)$ and $\alpha'\nsim\alpha$, then we do find a solution α' better than α outside of the equivalent class of α which means that one can further improve α along the greedy direction; Otherwise, the result of further move along the greedy direction is not clear, then a new search direction should be created for α'. Performance (empirical OL) for all obtained candidates is evaluated finally to find the best solution for each iteration. Combining this technique with the idea of expanding and then reinstating the decision results in proposed algorithm summarized in alg. 1.

Algorithm 1: Expanding-Reinstating Algorithm

Initialization: Number of decisions M; initialize $\mathcal{D}_M^{(0)}$; choose number of expanding decisions N, step rate ε and error tolerance δ.
for $t=1$ *to* T_{ex} **do**
  for $m=M$ *to* $M+N-1$ **do**
    Find $\alpha^{(0)}$ s.t. $\Delta\overline{\mathrm{L}}\left(\alpha^{(0)};\mathcal{D}_m^{(t-1)}\right)\geq 0$
    $\beta^{(0)}\leftarrow\alpha^{(0)}$;
    $\mathrm{G}^{(0)}\leftarrow\Delta\overline{\mathrm{L}}\left(\alpha^{(0)};\mathcal{D}_m^{(t-1)}\right)$;
    for $i=0$ *to ITER* **do**
      for $j=1$ *to* P **do**
        $y_j\leftarrow\sum_m \mathrm{U}_j^{\mathrm{T}}\mathrm{B}_m\left(\alpha^{(i)}\right)$;
      $j^\star\in\arg\min_{1\leq k\leq P}\ y_k$;
      $\gamma^{(0)}\leftarrow\alpha^{(i)}$;
      for $j=1$ *to* J **do**
        $\gamma^{(j)}\leftarrow\gamma^{(j-1)}+\varepsilon\mathrm{e}_{j^\star}$;
        if $\gamma^{(j)}\nsim\gamma^{(j-1)}$ **then**
          $\alpha^{(i+1)}\leftarrow\gamma^{(j)}$;
          if $\Delta\widehat{\mathrm{L}}\left(\gamma^{(j-1)};\mathcal{D}_m^{(t-1)}\right)>\Delta\widehat{\mathrm{L}}\left(\gamma^{(j)};\mathcal{D}_m^{(t-1)}\right)$ **then**
            $\mathrm{G}^{(i+1)}\leftarrow\Delta\overline{\mathrm{L}}\left(\gamma^{(j-1)};\mathcal{D}_m^{(t-1)}\right)$;
            $\beta^{(i+1)}\leftarrow\gamma^{(j-1)}$;
          else
            $\mathrm{G}^{(i+1)}\leftarrow\Delta\overline{\mathrm{L}}\left(\gamma^{(j)};\mathcal{D}_m^{(t-1)}\right)$;
            $\beta^{(i+1)}\leftarrow\gamma^{(j)}$;
    $k\in\arg\max_{0\leq i\leq\mathrm{ITER}}\ \mathrm{G}^{(i)}$;
    $\mathcal{D}_{m+1}^{(t-1)}\leftarrow\mathcal{D}_m^{(t-1)}\bigcup\left\{\sum_p\alpha_p^{(k)}v_p\right\}$;
  $\mathcal{D}_M^{(t+1)}\in\arg\min_{\mathcal{D}'\subset\mathcal{D}_{M+N}^{(t)},|\mathcal{D}'|=M}\overline{\mathrm{L}}\left(\mathcal{D}';\mathcal{T}\right)$
  if $\overline{\mathrm{L}}\left(\mathcal{D}_M^{(t)};\mathcal{T}\right)-\overline{\mathrm{L}}\left(\mathcal{D}_M^{(t+1)};\mathcal{T}\right)<\delta$ **then**
    Break;

Output: Required decision set is $\mathcal{D}_M^{(t)}$;

Before ending the section we will show that proposed method could be extended to some cases of non-convex cost function. Without loss of generality, we assume that the cost function $f(x;g)$ is twice-differentiable w.r.t. to decision variable x.We denote the Hessian matrix $\mathrm{H}_f(x_0,g)$ of f w.r.t. x at point (x_0,g) as

$$\mathrm{H}_f(x_0,g) \triangleq \left.\frac{\partial^2 f(x;g)}{\partial x^i \partial x^j}\right|_{(x_0,g)} \tag{18}$$

and its smallest eigen-value given parameter g for all possible $x \in \mathcal{X}$: $\lambda_{\min}(g) = \min_{x\in\mathcal{X}} \lambda\{\mathrm{H}_f(x,g)\}$. To generalize the Jensen's inequality to non-convex function, we introduce so called weakly convex function [35]:

Definition III.3. *(r-weakly convex function) Given a continuous function $u : \mathbb{R}^K \to \mathbb{R}$ defined on a convex set S, consider the function $h : \mathbb{R}^{K+1} \to \mathbb{R}$ with $r \in \mathbb{R}$ defined by $h(x,r) = u(x) - \frac{1}{2} r x^{\mathrm{T}} x$. If function $h(x,r)$ is a convex function on S for some r, then $h(x,r)$ is called a convexfication of u and r is a convexifier of u on S. function u is said to be r-weakly convex if it has a convexifier r.*

Weakly convex functions are important for various optimization problem (see examples in [36], [37]). Furthermore, we can generalized Jensen's inequality from strictly convex functions to weakly convex functions as the following proposition:

Proposition III.4. *(Generalized Jensen's inequality) For any r-weakly convex function u, it holds that for $\forall a \in \triangle_P$*

$$u\left(\sum_{i=1}^{P} a_i x_i\right) \le \sum_{i=1}^{P} a_i u(x_i) - \frac{r}{2}\left[\sum_{i,j=1}^{P} a_i a_j x_i^{\mathrm{T}}(x_i - x_j)\right] \tag{19}$$

Proof: The proof is omitted here. For more details, see [38]. ∎

Equipped with generalized Jensen's inequality, similar analysis is possible for the optimality loss. One can easily have:

$$\begin{aligned}&\overline{\mathrm{L}}(\mathcal{Q}_M;\mathcal{T}) - \overline{\mathrm{L}}(\mathcal{Q}_{M+N};\mathcal{T}) \\ \ge& \frac{1}{T}\sum_m \mathrm{W}_m^{\mathrm{T}} \mathrm{B}_m(\alpha) - \frac{1}{T}\sum_i \alpha_i \sum_m \mathrm{U}_i^{\mathrm{T}} \mathrm{B}_m(\alpha) \\ &+ \frac{1}{2T}\sum_{t,m,i,j} \alpha_i \alpha_j v_i^{\mathrm{T}}(v_i - v_j)\mathrm{B}_{tm}(\alpha)\rho_t \end{aligned} \tag{20}$$

where ρ_t is a convexfier of cost function f given $g^{(t)}$ w.r.t. x. Obviously one would like to minimize the optimality loss introduced by expanding decision set, then the optimal choice for vector ρ_t should be $\rho_t = \lambda_{\min}\left(g^{(t)}\right)$ since all $\rho_t' < \lambda_{\min}\left(g^{(t)}\right)$ is also a convefixier of cost function for parameter $g^{(t)}$. Similar analysis and algorithm could be applied to the lower bound obtained in Eq. 20. Due to limit of space, it is omitted to avoid duplicated materials.

IV. Numerical Results

In this section, we aim at showing the benefits of our proposed methods by comparing with conventional quantization technique.

► We first consider the spectral efficiency (SE) function $f^{\mathrm{SE}}(x;g) = -\sum_{i=1}^{S} \log\left(\sigma^2 + x_i g_i\right)$ under maximum power constraint $\sum_{i=1}^{S} x_i \le P_{\max}$. One can easily verify that its Hessian matrix is $\mathrm{H}_{f^{\mathrm{SE}}}(x;g) = \mathrm{diag}\left\{\frac{g_i^2}{(\sigma^2 + x_i g_i)^2}\right\}_i$. Therefore the cost function is convex function w.r.t. decision variable x. Meanwhile the decision space is a convex polyhedron. For parameter setting, we choose number of bands $S = 4$, power budget $P_{\max} = 10\mathrm{mW}$, variance of noise $\sigma^2 = 1\mathrm{mW}$, number of parameter samples $N_{\text{sample}} = 10000$, number of iteration changing the equivalent class ITER $= 1000$; iteration number for decision set expansion $T_{\text{ex}} = 10$ and largest cardinality of decision set is chosen as $N = M + 1$. In Fig. 1 , the relative optimality loss (ROL) in percentage (relatively to the idea case):

$$\text{Relative OL}(\%) = 100 \times \left(\frac{f(\chi(Q_M(g)); g) - f(\chi(g); g)}{f(\chi(g); g)}\right) \tag{21}$$

v.s. number of decisions is illustrated for decision set found by alg. 1 and Lloyd-Max algorithm which is distortion-based or goal-ignorant. One can observe that the performance of the proposed algorithm always dominates the Lloyd-Max algorithm. Moreover algorithm 1 is easier to implement for general cost function it requires merely simple linear computation of matrices compared to Lloyd-Max algorithm which is an alternating algorithm in essence.

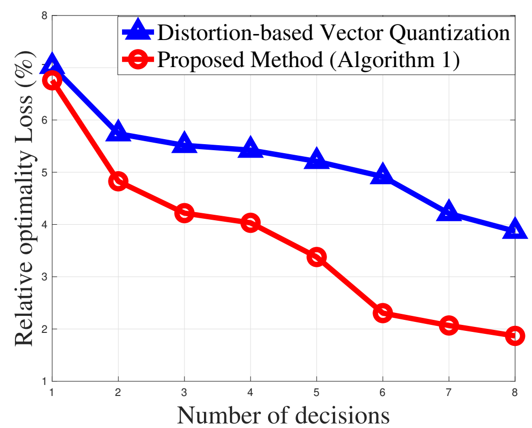

Fig. 2. Relative optimality loss (%) v.s. number of decisions for distortion-based vector quantization (Lloyd-Max algorithm) and proposed algorithm 1. This figure clearly shows the importance of adapting the quantizer to the goal instead of using the conventional distortion-based quantization paradigm.

► Now we assume $\ell_1 = \ell_2 = 24$. We consider a performance metric which is relevant for a communication problem in the smart grid. Indeed, we consider that the goal function $f^{\mathrm{PCS}}(x;g) = \|x+g\|_{\mathrm{P}}$, P being the exponent power parameter of the L_{P} norm, and PCS stands for power consumption scheduling. Here the vector $x = (x_1, ..., x_{\ell_1})$ represents the chosen flexible power consumption scheduling strategy; we impose that $x_i \ge 0$ and $\sum_{i=1}^{\ell_1} x_i \ge E$, $E > 0$ being the desired energy level. The parameter vector g represents the non-controllable part of the power. When P becomes large, the problem amount to limiting the peak power. For a given relative OL of 5% one then looks at the number of required clusters (that is, M) versus the exponent power parameter of the $\mathrm{L_P}$ (that is, P). In Fig. 3, the performance of the $k-$means algorithm with the proposed algorithm for the Pecanstreet database [39] with $E = 40kWh$ is illustrated. For large exponent power, e.g., $\mathrm{P} = 20$, proposed goal-oriented approach reduce the number of cells from 50 to 14 to achieve a ROL of 5% by taking advantage of the convexity of the $\mathrm{L_P}$ norm.

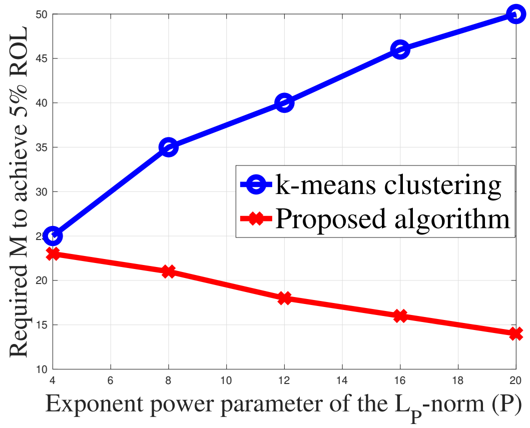

Fig. 3. Required number of cells (M) against the exponent power parameter of the $\mathrm{L_P}$-norm (P) for the k-means and proposed algorithm. Proposed goal-oriented approach could largely reduce the number of cells especially for large exponent power.

V. Conclusions

In this paper, we focus on a special scenario of general goal-oriented quantization problem where the decision space is a convex polyhedron and the cost function is (weakly) convex function with respect to decision variable. An algorithm based on expanding and reinstating the decision set and taking advantage of Jensen's inequality is proposed. We have shown that it is sufficient to consider single-decision expansion for proposed methods. Besides, numerical results show that proposed algorithm could reduce the optimality loss or the required number of quantization bits comparing to conventional task-ignorant quantizer. One future work of current framework is to extend proposed method for general scenario where cost functions are no longer convex with arbitrary shape of decision space. Techniques such as deep learning could be of great interest to explore for the situation where the optimal solution of the targeted optimization problem is unknown or its closed form is missing.

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

References

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, Jul. 1948.

[2] W. Saad, M. Bennis and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems", *IEEE Network*, vol. 34, no. 3, pp. 134-142, May/June 2020.

[3] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan and M. Zorzi, "Toward 6G Networks: Use Cases and Technologies", *IEEE Communications Magazine*, vol. 58, no. 3, pp. 55-61, March 2020.

[4] K. B. Letaief, W. Chen, Y. Shi, J. Zhang and Y. -J. A. Zhang, "The Roadmap to 6G: AI Empowered Wireless Networks", *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84-90, August 2019.

[5] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V Poor, "Smart meter privacy: A utility-privacy framework", in *Proceedings of IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 190-195, Oct. 2011.

[6] J. Zheng, E. R. Duni and B. D. Rao, "Analysis of Multiple-Antenna Systems With Finite-Rate Feedback Using High-Resolution Quantization Theory", *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1461-1476, April 2007.

[7] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao and M. Andrews, "An overview of limited feedback in wireless communication systems", *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341-1365, October 2008.

[8] C. K. Au-Yeung, S. Y. Park and D. J. Love, "A Simple Dual-Mode Limited Feedback Multiuser Downlink System", *IEEE Transactions on Communications*, vol. 57, no. 5, pp. 1514-1522, May 2009.

[9] R. M. Gray, and D. L. Neuhoff, "Quantization", *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325-2383,1998.

[10] N. Shlezinger, Y. C. Eldar and M. R. D. Rodrigues, "Hardware-Limited Task-Based Quantization", *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5223-5238, 2019.

[11] S. Salamtian, N. Shlezinger, Y. C. Eldar, and M. Médard, "Task-based quantization for recovering quadratic functions using principal inertia components", in *Proceedings of IEEE Information Theory Workshop (ITW)*, 2019, pp. 390-394.

[12] V. Misra, V. K. Goyal and L. R. Varshney, "Distributed Scalar Quantization for Computing: High-Resolution Analysis and Extensions", *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5298-5325, Aug. 2011.

[13] P. Fleischer, "Sufficient conditions for achieving minimum distortion in a quantizer", *IEEE Innt. Conv. Rec.*, pp. 104-111, 1964.

[14] R. Cabral Farias and J. Brossier, "Scalar Quantization for Estimation: From An Asymptotic Design to a Practical Solution", *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2860-2870, June 1, 2014.

[15] P. Li, N. Shlezinger, H. Zhang, B. Wang and Y. C. Eldar, "Graph Signal Compression via Task-Based Quantization," *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP'21), pp. 5514-5518.

[16] M. Kountouris, N. Pappas, "Semantics-empowered communication for networked intelligent systems", *IEEE Communications Magazine*, vol. 59, no. 6, pp. 96-102, 2021.

[17] G. Shi, Y. Xiao, Y. Li and X. Xie, "From Semantic Communication to Semantic-Aware Networking: Model, Architecture, and Open Problems", *IEEE Communications Magazine*, vol. 59, no. 8, pp. 44-50, August 2021.

[18] E. Calvanese Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications", *Computer Networks*, vol. 190, 2021.

[19] J. Liu, W. Zhang, H. V. Poor, "A Rate-Distortion Framework for Characterizing Semantic Information", in *Proceedings of IEEE International Symposium on Information Theory (ISIT'21)*, pp. 2894-2899, July 2021.

[20] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems", *IEEE Transactions on Signal Processing*, vol. 69, pp. 2663-2675, 2021.

[21] M. Sana and E. C. Strinati, "Learning semantics: An opportunity for effective 6G communications", in *Proceedings of IEEE 19th Annual Consumer Communications and Networking Conference (CCNC'19)*, 2022.

[22] W. J. Yun, B. Lim, S. Jung, Y.-C. Ko, J. Park, J. Kim, and M. Bennis, "Attention-based reinforcement learning for real-time uav semantic communication", in *Proceedings of 17th IEEE International Symposium on Wireless Communication Systems (ISWCS'17)*, 2021, pp. 1–6.

[23] M. K. Farshbafan, W. Saad, and M. Debbah, "Common Language for Goal-Oriented Semantic Communications: A Curriculum Learning Framework", arXiv preprint arXiv:2111.08051.

[24] H. Seo, J. Park, M. Bennis, and M. Debbah, "Semantics-Native Communication with Contextual Reasoning", arXiv preprint arXiv:2108.05681.

[25] S. Lloyd, "Least Squares Quantization in PCM", *IEEE Transactions on Information Theory*, 28(2), 129-137, 1982.

[26] C. Zhang, N. Khalfet, S. Lasaulce, V. Varma, and S. Tarbouriech, "Payoff-oriented quantization and application to power control", in *Proceeding*

of 15th IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (Wiopt'17), Paris, France, 2017.

[27] H. Zou, C. Zhang, S. Lasaulce, L. Saludjian, and P. Panciatici, "Decision-Oriented Communications: Application to Energy-Efficient Resource Allocation", in *Proceeding of 9th IEEE International Conference on Wireless Networks and Mobile Communications (WINCOM'18)*, Marrakech, Morroco, 2018.

[28] H. Zou, C. Zhang, S. Lasaulce, L. Saludjian and P. Panciatici, "Decision Set Optimization and Energy-Efficient MIMO Communications", in *Proceeding of 30th IEEE International Symposium on Personal,Indoor and Mobile Radio Communications (PIMRC'19)*, Istanbul, Turkey, 2019.

[29] C. Zhang, S. Lasaulce, M. Hennebel, L. Saludjian, P. Panciatici, and H. V. Poor, "Decision-making oriented clustering: Application to pricing and power consumption scheduling", *Applied Energy*, vol. 297, 2021.

[30] S. Lloyd, "Least squares quantization in PCM", *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129-137, 1982.

[31] J. Max, "Quantizing for minimum distortion", *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7-12, 1960.

[32] E.V. Belmega, S. Lasaulce, and M. Debbah, "Power allocation games for MIMO multiple access channels with coordination", IEEE Trans. on Wireless Communications, vol. 8, no. 6, pp. 3182-3192, Jun. 2009.

[33] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination", *IEEE Transactions on Wireless Communications*, 14(1):1–14, January 2015.

[34] J. L. W. V. Jensen, "Sur les fonctions convexes et les inegalites entre les valeur moyennes", *Acta Mathematica* 30(1906), 175–193.

[35] D. Drusvyatskiy, "The proximal point method revisited", arXiv preprint arXiv:1712.06038.

[36] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A.S. Willsky, "Rank-sparsity incoherence for matrix decomposition", *SIAM Journal on Optimization*, 21(2):572-596, 2011.

[37] E. Abbe, A. S. Bandeira, A. Bracher and A. Singer, "Decoding Binary Node Labels from Censored Edge Measurements: Phase Transition and Efficient Recovery," *IEEE Transactions on Network Science and Engineering* , vol. 1, no. 1, pp. 10-22, 1 Jan.-June 2014.

[38] S. Zlobec "Jensen's inequality for non-convex functions", *Mathematical Communications*, 9(2004), 119-124.

[39] Pecan street inc. dataport. [Online]. Available: https://dataport.pecanstreet.org/data.

[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K. Cambridge Univ. Press, 2004.