An Efficient Lloyd-Max Quantizer for Matching Pursuit Decompositions

Lisandro Lovisolo, Eduardo A B da Silva and Paulo S R Diniz

***Abstract*— Several applications are using the Matching Pursuit algorithm for signal and video compression. The Matching Pursuit approximates signals iteratively using linear combinations of pre-defined atoms of a dictionary. In compression applications Matching Pursuits coefficients, which multiply the atoms in the linear combination, need to be quantized. The Lloyd-Max quantizer is known to be the best quantizer for a given source. However, to design a Lloyd-Max quantizer the statistics of the source need to be known. The statistics of Matching Pursuit coefficients are difficult to model. In this paper, starting from the observation that the statistics of the angles between the residues and the atoms present little variation along Matching Pursuit iterations, we propose to use these statistics to model the ones of Matching Pursuit coefficients. This permits the design of Lloyd-Max quantizers for Matching Pursuit coefficients. The Lloyd-Max quantize is compared to a state-of-the-art off-loop Matching Pursuit quantization scheme. Results show that the proposed scheme has good rate-distortion performance, specially at low rates.**

***Index Terms*— Matching Pursuit, Quantization, Compression.**

I. Introduction

The Matching Pursuit algorithm [1] (MP) approximates signals iteratively. The approximation is obtained using atoms (pre-defined signals) $\mathbf{g}_k$ from a dictionary $\mathcal{D}$ (the collection of possible atoms). The MP is being used in compression schemes for 1-dimensional [2], [3] and 2-dimensional signals [2], [4].

The MP works as follows. Let $\mathcal{D} = \{\mathbf{g}_k\}$ with $k \in [1, \ldots, C(\mathcal{D})]$, such that $\|\mathbf{g}_k\| = 1 \ \forall k$, $C(\mathcal{D})$ is the dictionary cardinality (number of elements in $\mathcal{D}$). At each iteration $n \geq 1$, the MP searches for the atom $\mathbf{g}_{i(n)}$, $i(n) \in [1 \ldots C(\mathcal{D})]$, with largest inner product with the residual signal $\mathbf{r}_{\mathbf{x}}^{n-1}$ [1], [5], which is given by

$$\gamma_n = \langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{i(n)} \rangle. \tag{1}$$

Observe that the first residue is equal to the signal to be decomposed, that is $\mathbf{r}_{\mathbf{x}}^{0} = \mathbf{x}$. The inner product γ_n, which is a measure of "how much" of $\mathbf{g}_{i(n)}$ is included in $\mathbf{r}_{\mathbf{x}}^{n-1}$, is used to compute the next residue

$$\mathbf{r}_{\mathbf{x}}^{n} = \mathbf{r}_{\mathbf{x}}^{n-1} - \gamma_n \mathbf{g}_{i(n)}. \tag{2}$$

Lisandro Lovisolo (lisandro@uerj.br) is with DETEL-FEN-UERJ, Eduardo A B da Silva and Paulo S R Diniz (eduardo, diniz@lps.ufrj.br) are with POLI/COPPE-UFRJ.

After M iterations, the MP obtains the M-term approximation $\hat{\mathbf{x}}$, or simply M-term, given by

$$\hat{\mathbf{x}} = \sum_{n=1}^{M} \gamma_n \mathbf{g}_{i(n)}, \tag{3}$$

which has distortion $\mathbf{x} - \hat{\mathbf{x}} = \mathbf{r}_{\mathbf{x}}^{M}$, the M^{th} residue.

The performance of the MP for signal compression applications depends heavily on two aspects:

- **i**) Dictionary – it should include atoms that are good matches to the possible components of the signals to compress. In addition, the dictionary cardinality affects the data rate.
- **ii**) Quantization – for compression the coefficients γ_n must be quantized.

Using a coefficient quantization rule $Q[\cdot]$ the compressed signal is retrieved by the quantized M-term

$$\hat{\mathbf{x}}_q = \sum_{n=1}^{M} Q[\gamma_n] \mathbf{g}_{i(n)}. \tag{4}$$

In order to design efficient quantizers $Q[\cdot]$, one needs a statistical model for MP coefficients. However, MP coefficients are difficult to model. For example, in [6] it has been observed that MP residues have a chaotic behavior. Here, instead of searching for a good model for MP coefficients, our approach is based on a statistical model for the MP angles. It relies on the observation that the angles between the atom chosen and the residue being decomposed along MP iterations have well behaved statistics.

Define, at each MP iteration, the angle between the residue $\mathbf{r}_{\mathbf{x}}^{n-1}$ and the selected atom $\mathbf{g}_{i(n)}$ as

$$\theta_n = \arccos\left(\frac{|\langle \mathbf{r}_{\mathbf{x}}^{n-1}, \mathbf{g}_{i(n)} \rangle|}{\|\mathbf{r}_{\mathbf{x}}^{n-1}\|}\right). \tag{5}$$

We have verified, experimentally (see section II), that the statistics of θ_n are approximately independent of n. Here, it is conjectured that the angles between the residues and the atoms in MP iterations can be statistically modeled as independent and identically distributed.

II. Angles in Matching Pursuit Iterations

Using the definition of θ_n, eq. (5), the MP algorithm is such that

$$|\gamma_1| = \|\mathbf{x}\| \cos(\theta_1), \tag{6}$$
$$|\gamma_2| = \|\mathbf{x}\| \sin(\theta_1) \cos(\theta_2), \tag{7}$$
$$\vdots$$
$$|\gamma_n| = \|\mathbf{x}\| \prod_{i=1}^{n-1} \sin(\theta_i) \cos(\theta_n). \tag{8}$$

However, if $\mathcal{D}$ includes the elements $\mathbf{g}_k$ and $-\mathbf{g}_k$ for any k, forcing the inner products to be positive, then

$$\gamma_n = \|\mathbf{x}\| \prod_{i=1}^{n} \sin(\theta_i) \cos(\theta_n). \tag{9}$$

This work assumes that $\mathbf{g}_k$ and $-\mathbf{g}_k$ belong to $\mathcal{D}$ obtaining always positive coefficients, and all subsequent references to dictionaries consider that. That is, if $\mathcal{D}$ does not include $-\mathbf{g}_k$, then $-\mathbf{g}_k$ is included in $\mathcal{D}$ and $C(\mathcal{D})$ is updated accordingly.

A. Statistics of the Angles in MP Iterations

In [6], it has been observed that MP residues have a chaotic behavior, thus it would be reasonable to assume that, after some MP iterations, the residues can have any orientation. In addition, one may assume that these orientations may have uniform probability density function on the unit-ball. This is equivalent to assume that the residues are realizations from a memoryless independent and identically distributed (iid) Gaussian source. That is, the signal source is such that if an outcome is $\mathbf{x} = [x(1), x(2), \ldots, x(N)]$, then the $x(j)$ have the same Gaussian distribution $\mathcal{N}(0, \sigma^2)$.This source does not privilege any signal orientation. Therefore, we use this source to investigate MP angle statistics.

Fig. 1 shows the relative frequency histograms of the RVs Θ_n, which correspond to the angles θ_n that result from decompositions of realizations of a Gaussian source using a dictionary composed of 16 normalized random atoms also drawn from a Gaussian source in $\mathbb{R}^4$. These histograms were obtained using an ensemble of 50,000 MP decompositions of random signals from a Gaussian source. Such dictionaries, composed of $C(\mathcal{D})$ normalized signals drawn from an N-dimensional Gaussian source, are referred here as $GSND(C(\mathcal{D}), N)$ and thus the former dictionary is denoted by $GSND(16, 4)$. It should be noted that this dictionary has its cardinality doubled in order to obtain just positive coefficients. In Fig. 1 one notes that the histograms of the RVs Θ_n have very similar shape for all n. This leads to the conjecture that the pdfs $f_{\Theta_n}(\theta_n)$ are independent and identically distributed at any iteration. The results presented in [6] corroborate this assumption, where it is shown that, under specific conditions, MP residues have chaotic behavior.

Fig. 2 shows the mean and the variance of Θ_n for several n. Fig. 2 depicts also the covariances between MP angles in

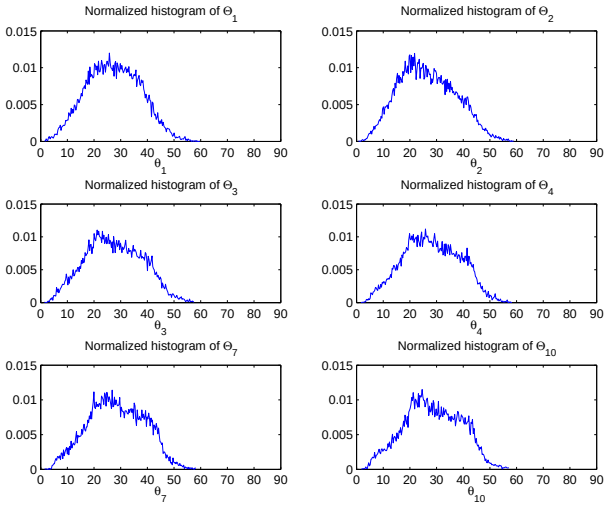

Fig. 1. Relative frequency histograms of Θ_n for a Gaussian source in $\mathbb{R}^4$ using the $GSND(16, 4)$.

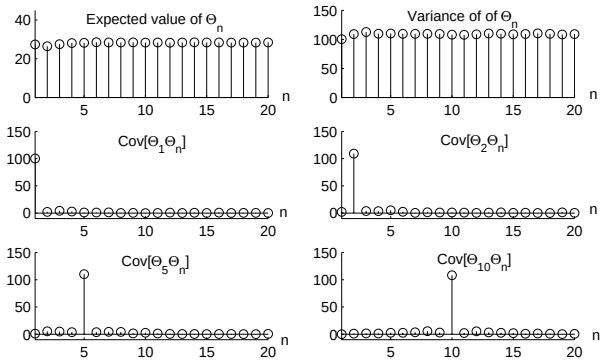

Fig. 2. Mean, variance and covariance of Θ_n for a Gaussian source in $\mathbb{R}^4$ using the $GSND(16, 4)$.

different steps. From it one observes that $\mathrm{Cov}[\Theta_i \Theta_k] = 0$, $\forall i \neq k$, that is, the angles are uncorrelated. In this work we assume that the angles are independent RVs. This is not an unreasonable assumption, since it does not contradicts the behavior observed in Fig. 2.

The results presented so far make use of a dictionary of relatively low dimension and cardinality $C(\mathcal{D})$. In practice, the MP is commonly used in large dimensional spaces with dictionaries such as the Gabor one [5]. The elements of this dictionary are defined by translations, modulations and dilations of a prototype signal. The most common choice for the prototype $f(n)$ is the Gaussian window. We analyze here a real Gabor dictionary composed of atoms with predefined phases in multiples of $\frac{\pi}{V}$. Each atom is then given by [5]

$$g(n) = \begin{cases} \delta(n), & j = 0 \\ K_{(j,p,v)} f\left(\dfrac{n - p2^j}{2^j}\right) \cos\left(nk\pi 2^{1-j} + \dfrac{\pi v}{V}\right), & j \quad (0, L) \\ \frac{1}{\sqrt{N}}, & j = L \end{cases}, \tag{10}$$

where $f(n) = 2^{\frac{1}{4}} e^{-\pi n^2}$, n is the sample, $K_{(j,p,v)}$ provides unit-norm atoms, and $v \in [0, \ldots, V-1]$. Above, j defines the atom scale, p defines the time shift, and k defines the atom modulation, and for $L = \log_2(N)$ scales their ranges are [1] $j \in [0, L]$, $p \in [0, N2^{-j})$, $k \in [0, 2^j)$, and $v \in [1, V]$.

Fig. 3 shows $f_{\Theta_n}(\theta_n)$, for some n, obtained for an ensemble of 128,000 decompositions of Gaussian signals in $\mathbb{R}^{64}$ using the Gabor dictionary with four phases. Fig. 4 shows $f_{\Theta_n}(\theta_n)$, for some n, obtained for an ensemble of 128,000 decompositions of signals driven from a source that has gamma distributed coordinates in $\mathbb{R}^{64}$ for the same dictionary. Note that the angle statistics shown for each signal source differ strongly only at the first MP iteration, being visually very similar for the other iterations. It can be noted that for this dictionary the angles in different MP steps $n \geq 2$ have similar statistics even for very different sources. Note that, although the statistics of Θ_n are not exactly equal to the statistics obtained in the first MP iteration for a Gaussian source, they are reasonably close to these; therefore $f_{\Theta_n}(\theta_n)$, for $n > 1$, can be reasonably approximated by the $f_{\Theta_1}(\theta_1)$ obtained for a memoryless Gaussian source. This is a reasonable assumption, since the memoryless Gaussian source does not privilege any orientation, what seem to be the case for the residues $\mathbf{r}_{\mathbf{x}}^{n-1}$ for $n > 1$.

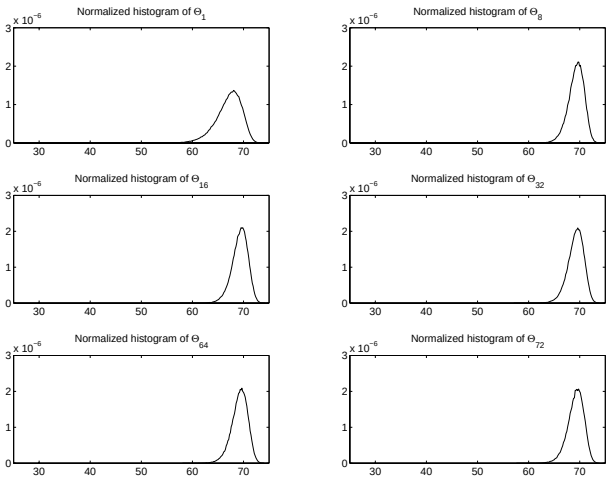

Fig. 3. Normalized histograms of MP angles for a Gaussian source in $\mathbb{R}^{64}$, using 100 bin, at $n = \{1, 8, 16, 32, 64, 72\}$, for the 4-phase Gabor dictionary in $\mathbb{R}^{64}$.

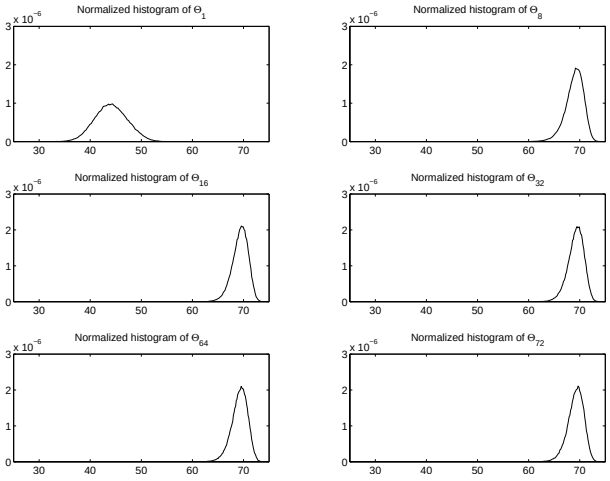

Fig. 4. Normalized histograms of MP angles for a source with coordinates driven from a gamma distribution, using 100 bin, at $n = \{1, 8, 16, 32, 64, 72\}$, for the 4-phase Gabor dictionary in $\mathbb{R}^{64}$.

B. Discussion of Results

It was verified that the statistics of the angles, after the first MP iteration, can be considered to be invariant with respect to the step number. For some dictionaries, for example the Gabor one, the angle pdf, after some steps, is slightly different from the angle pdf obtained for a memoryless Gaussian source. However, one can still use a Gaussian source to obtain good estimates of the pdf of the angle in MP iterations. The difference in the statistics means that, although the residues of MP iterations have similar orientations, these orientations are not uniformly distributed. Nevertheless, the iid statistical model is not too bad an assumption. In the sequel, this model is used to design Lloyd-Max quantizers of MP coefficients. This permits to verify the validity of the statistical model for the angles in MP iterations. It should be noted that the statistics of the first MP angle, and therefore of the first coefficient, are much more source dependent than they are at further MP steps. Therefore, in order, to make an appropriate use of the presented angle model, the first coefficient, γ_1, will be quantized with negligible error and encoded as side information. Note that this value replaces $\|\mathbf{x}\|$, the value of the signal norm which is usually transmitted as side information in practical MP based compression schemes [2], [4].

III. Quantized Matching Pursuit Decompositions

In compression applications, one encodes the quantized versions $Q[\gamma_n]$ of the coefficients γ_n of M-terms. One way to obtain the $Q[\gamma_n]$ is to quantize the coefficients off-loop, i.e. first the whole decomposition is obtained and then the coefficients are quantized. Another strategy is in-loop quantization [4], in which the quantized coefficient is used to compute the residue, and due to that in-loop quantization might result in atoms in M-term that depend on the quantizer used.

We choose off-loop quantization since it allows a simple rate×distortion (RD) optimization procedure, in which different quantizers are tried out in order to find one that meets a prescribed RD criterion. In contrast, in-loop quantization would require, for RD optimization, several MP signal decompositions, what would greatly increase the computational demands.

Since in off-loop quantization the residues are computed without knowledge of the quantizer, the quantization error in one step cannot be compensated in subsequent steps. However, these quantization errors can be minimized by designing appropriate quantizers [2]. In this work, MP coefficient quantizers design is achieved by employing the iid statistical model of MP angles to design Lloyd-Max quantizers for the MP coefficients.

The quantizers designed make use of dead-zones because some advantage in RD performance can be gained by quantizing small coefficients to zero [2], [4]: neither the coefficients quantized to zero nor their atoms indices are

sent, reducing the data rate.

For transmission, it is important that if coefficients and/or indices are lost, the decoder can just successfully ignore the lost terms when reconstructing the signal. Therefore, it would be highly desirable for the quantizer for a given γ_n to be independent of the quantized values of other γ_m ($m \neq n$).

A. Distortion Due to Off-Loop Quantization

When M-terms are quantized off-loop, each coefficient γ_n is replaced by its quantized version $Q[\gamma_n]$ and the signal approximation is retrieved using eq. (4). The distortion criterion employed for MP quantizer design can be:

i) the energy of the error relative to the actual signal

$$\mathbf{d} = \mathbf{x} - \hat{\mathbf{x}}_q; \tag{11}$$

ii) the energy of the error relative to the M-term

$$\mathbf{d}_M = \hat{\mathbf{x}} - \hat{\mathbf{x}}_q = \sum_{n=1}^{M} (\gamma_n - Q[\gamma_n])\mathbf{g}_{i(n)}. \tag{12}$$

In this work, the second one is used, since we design one Lloyd-Max quantizer for each number of coded terms M. The distortion per sample of the quantized M-term is given by $d_M^2 = \frac{1}{N}\|\mathbf{d}_M\|^2$, that is,

$$d_M^2 = \frac{1}{N}\sum_{n=1}^{M}\sum_{m=1}^{M}(\gamma_n - Q[\gamma_n])(\gamma_m - Q[\gamma_m])\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle, \tag{13}$$

where N is the signal length. $\mathcal{D}$ is composed of unit norm vectors ($\|\mathbf{g}_{i(n)}\| = 1$). Thus, defining the quantization error

$$e_q(\gamma_n) = \gamma_n - Q[\gamma_n], \tag{14}$$

it follows that

$$d_M^2 = \frac{1}{N}\left[\sum_{n=1}^{M} e_q^2(\gamma_n) + \sum_{n=1}^{M}\sum_{m=1,\neq n}^{M} e_q(\gamma_n)e_q(\gamma_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]. \tag{15}$$

For an M-term $\hat{\mathbf{x}}$, the minimization of eq. (15) implies also the minimization of the distortion per sample (see eq. (11)) given by

$$d^2 = \frac{1}{N}\|\mathbf{d}\|^2 = \frac{1}{N}\|\mathbf{x} - \hat{\mathbf{x}}_q\|^2. \tag{16}$$

Note that if d_M^2 is equal to zero then $d^2 = \|\mathbf{r}_{\mathbf{x}}^M\|^2/N$.

1) Distortion for a Given Source: For a given signal source $\mathcal{X}$, one may consider the expected value of d_M^2

$$E[d_M^2] = \frac{1}{N}\left\{\sum_{n=1}^{M} E\left[e_q^2(\Gamma_n)\right] + \sum_{n=1}^{M}\sum_{m=1,m\neq n}^{M} E\left[e_q(\Gamma_n)e_q(\Gamma_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]\right\}. \tag{17}$$

Each Γ_n is a random variable (RV) that corresponds to γ_n, for $1 \leq n \leq M$, for signals drawn from $\mathcal{X}$. In eq. (17), $E\left[e_q^2(\Gamma_n)\right]$ stands for the expected value of the squared quantization errors of the RV Γ_n.

B. MP Coefficient Quantization Using MP Angles

In eq. (9) the value of $\|\mathbf{x}\|$ is required to compute the MP coefficients from the MP angles. According to the discussion at the end of section II we use the first coefficient γ_1 instead of $\|\mathbf{x}\|$, then eq. (9) can be rewritten as

$$\gamma_n = \gamma_1\delta_n,\ \delta_n = \tan(\theta_1)\prod_{i=2}^{n-1}\sin(\theta_i)\cos(\theta_n),\ n \geq 2. \tag{18}$$

Thus, the pdfs of the coefficients γ_n can be computed from the pdfs of the angles Θ_n, assuming that the Θ_n are independent. The pdfs of the θ_n, $n > 1$, can be obtained as explained in section II. For a known γ_1, the pdf of the RV Γ_n, for $n \geq 2$, is given by $f_{\ _n}(\gamma_n|\gamma_1) = f_{\Delta_n}(\gamma_1\delta_n|\gamma_1)$, where Δ_n is the RV whose outcome is δ_n, see eq. (18).

If an optimal quantizer Q is designed for the RV Y, then the optimal quantizer for $Z = cY$ (c is a constant) is simply a scaled version of Q. Thus, for γ_1 known, the quantization of δ_n instead of γ_n, from eq. (17), gives

$$E[d_M^2|\gamma_1] = \frac{{\gamma_1}^2}{N}\left\{\sum_{n=2}^{M} E\left[e_q^2(\Delta_n)\right] + \sum_{n=2}^{M}\sum_{m=2,m\neq n}^{M} E\left[e_q(\Delta_n)e_q(\Delta_m)\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)}\rangle\right]\right\}. \tag{19}$$

Once the pdfs of the RVs Δ_n are known $E[d_M^2|\gamma_1]$ can be computed for any quantization rule applied to Δ_n. Since the quantization is applied to δ_n instead of γ_n, the value of γ_1 is required at the decoder. In this work we use γ_1 to define γ_n, for $n \geq 2$. The use of γ_1 to compute γ_n ($n \neq 1$) guarantees its correct value at the decoder. For quantizer design using eq. (19) the pdfs of Δ_n are needed which can be computed from the RVs Θ_i, $i \in [1, \ldots, n]$, corresponding to the angles between the residues and the atoms selected in MP steps, as in eq. (18).

IV. Lloyd-Max Quantizers for MP Coefficients

For a given first coefficient γ_1 the quantization of MP coefficients should aim to minimize the distortion in eq. (19). In this case, one should design quantizers for Δ_n, $n \geq 2$. Therefore, if just one quantizer is to be used then it shall be designed for $\Delta = \cup_{n=2}^{M}\Delta_n$, the RV given by the union of the Δ_n ($n \geq 2$). Since MP iterations are disjoint events the pdf of $\Delta = \cup_{n=2}^{M}\Delta_n$ is given by $f_\Delta(\delta) = \frac{1}{M-1}\sum_{n=2}^{M} f_{\Delta_n}(\delta_n)$.

A. Distortion for an Optimal Quantizer

The distortion per sample in eq. (19) has two terms. The first term is the sum of the squared quantization errors of Δ_n, whereas the second contains a sum of inner products between dictionary atoms weighted by the products of the quantization errors of the atoms involved in the inner products. As verified, in section II, the RVs of the angles Θ_n can be assumed to be uncorrelated. Although Δ_n and Δ_m may be correlated, when designing a quantizer for $\Delta = \cup_{n=2}^{M}\Delta_n$ the assumption that the quantization errors $e_q(\Delta_n)$ are uncorrelated is reasonable. It is also

reasonable to assume that the quantization errors products $e_q(\Delta_n)e_q(\Delta_m)$ are not correlated to the inner products $\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)} \rangle$. Using the assumptions above, the second term in eq. (19) becomes

$$\sum_{n=2}^{M} \sum_{m=2, m\neq n}^{M} E\left[e_q(\Delta_n)\right] E\left[e_q(\Delta_m)\right] E\left[\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)} \rangle\right]. \tag{20}$$

The atoms selected at different MP steps may be correlated. But due to the invariant nature of the angles statistics in different MP steps, one can consider the expected value of the inner product between the atoms selected in any two different MP steps n and m to be also invariant. That is $E\left[\langle \mathbf{g}_{i(n)}, \mathbf{g}_{i(m)} \rangle\right] = c$. Therefore, noting that

$$E\left[e_q(\Delta)\right] = \frac{1}{M-1} \sum_{n=2}^{M} E\left[e_q(\Delta_n)\right], \tag{21}$$

eq. (20) yields

$$\sum_{n=2}^{M} E\left[e_q(\Delta_n)\right] \sum_{m=2, m\neq n}^{M} E\left[e_q(\Delta_m)\right] c = \tag{22}$$
$$(M-1) E\left[e_q(\Delta)\right] \sum_{n=2}^{M} \sum_{m=2, m\neq n}^{M} E\left[e_q(\Delta_m)\right] c.$$

The optimal quantization of Δ leads to $E\left[e_q(\Delta)\right] = 0$, so that the expression above vanishes. As a result eq. (19) becomes

$$E[d_M^2|\gamma_1] = \frac{{\gamma_1}^2}{N} \sum_{n=2}^{M} E[(\Delta_n - Q[\Delta_n])^2], \tag{23}$$

This result is a sum of terms $E[(\Delta_n - Q[\Delta_n])^2]$, thus

$$E[d_M^2|\gamma_1] = \frac{\gamma_1^2}{N} \sum_{n=2}^{M} \int (\delta_n - Q[\delta_n])^2 f_{\Delta_n}(\delta_n) d\delta_n. \tag{24}$$

Since the same quantizer $Q[\cdot]$ is applied to all Δ_n, $2 \leq n \leq M$, one has that

$$E[d_M^2|\gamma_1] = \gamma_1^2 \frac{M-1}{N} \int (\delta - Q[\delta])^2 f_{\Delta}(\delta) d\delta, \tag{25}$$

$$\text{with } f_{\Delta}(\delta) = \frac{1}{M-1} \sum_{n=2}^{M} f_{\Delta_n}(\delta_n) \tag{26}$$

If $\Delta = \cup_{n=2}^{M} \Delta_n$ and defining

$$\text{MSE}(\Delta) = E\left[(\Delta - Q[\Delta])^2\right] \tag{27}$$

eq. (25) becomes

$$E[d_M^2|\gamma_1] = \frac{{\gamma_1}^2 (M-1)}{N} \text{MSE}(\Delta). \tag{28}$$

The design of an optimal quantizer for $\Delta = \cup_{n=2}^{M} \Delta_n$, that minimizes eq. (28), is accomplished by Lloyd-Max quantizers [7] (LMQ). Note that since Lloyd-Max quantizers are unbiased estimates of the input then $E[e_q(\Delta)] = 0$, and thus the term in eq. (20) vanishes validating eq. (23).

B. Lloyd-Max Quantizer Design

The design of Lloyd-Max quantizers requires $f_{\Delta_n}(\delta_n)$, for $2 \leq n \leq M$; these are estimated from $f_{\Theta_1}(\theta_1)$. In turn, $f_{\Theta_1}(\theta_1)$ is estimated by applying one MP decomposition step to a large set of signals drawn from a memoryless Gaussian source. As discussed in section II, this is an acceptable procedure since all Θ_n, $n \geq 2$, have statistics that are similar to the ones obtained for a memoryless Gaussian source. The estimated $f_{\Delta_n}(\delta_n)$ are then used to calculate $f_{\Delta}(\delta)$, and $f_{\Delta}(\delta)$ is then used to obtain the LMQ of b_{coef} bits ($L = 2^{b_{\text{coef}}}$ levels). The quantizer thresholds and reconstruction levels are calculated using an iterative algorithm [7]. The $f_{\Delta_n}(\delta_n)$, that are used to compute $f_{\Delta}(\delta)$, are obtained recursively from an estimate of $f_{\Theta_1}(\theta_1)$ that is in turn obtained using the angle in the first MP step for an ensemble of $C(\mathcal{D})N^2$ realizations from a Gaussian source.

The same quantizer law is used for all coefficients γ_n. Note that $f_{\Delta}(\delta)$ varies with M, therefore each M leads to a different quantizer. The quantizer design is independent of γ_1 and it suffices to design quantizers for γ_1=1, storing copies of the quantizers in both encoder and decoder. The encoder sends γ_1, the number of bits of the quantizer, and the number of terms of the decomposition (M), in a header, to the decoder. The parameter γ_1 is used to scale the quantizer in both coder and decoder, a simple strategy that makes good use of resources.

V. Lloyd-Max Quantizers Performance

The state-of-the-art for off-loop quantization of MP coefficients is presented in [2]. There, uniform quantizers whose number of quantization levels and range adapt according to the coefficients of the M-term are used. Here, this quantization scheme is referred to as adaptive bounded uniform quantizer (ABUQ). The ABUQ implements a bit-allocation per coefficient that relies on the known result that MP coefficients magnitudes decrease on average at each MP iteration. In the ABUQ, previous to quantization all the coefficients need to be sorted in decreasing magnitude and the ABUQ is fed with the coefficients in this order. For each coefficient, the ABUQ employs a uniform quantizer of different range and number of levels; the quantizer range for the l^{th} coefficient depends on the quantized value of the $(l-1)^{th}$ coefficient, and the number of levels of each coefficient quantizer is decided using a criterion based on a Lagrangian multiplier – a bit-allocation procedure. For decoding the quantized decomposition the number of bits used to quantize the second coefficient as well as the larger coefficient are sent as side information.

The total rate of MP quantized decompositions is given by

$$R = S\left[\log_2\left(C(\mathcal{D})\right)\right] + r_{\text{coef}}, \tag{29}$$

where S is the number of terms that remain after quantization, and r_{coef} is the rate incurred in coding the quantized coefficients. Therefore, the rate in bits per sample is given by R/N, where N is the signal space dimension. The

strategy employed to generate the ABUQ coded bitstream is to entropy code the differences between the quantization indices of successive coefficients. In the comparisons between Lloyd-Max quantization and ABUQ presented here, this strategy is also employed to code the Lloyd-Max quantized M-terms.

Fig. 5 shows the RD curves of quantized MP expansions arriving from three different random sources in $\mathbb{R}^{10}$ (a memoryless Gaussian, a memoryless uniform and a memoryless Gamma distributed source) using both the Lloyd-Max quantization and the ABUQ and a $GSND(128, 10)$. For each distinct source the results are averages over an ensemble of 100 quantized MP decompositions of signals from each source. For this experiment the LMQs were designed with bit-depth ranging from 1 to 8. It can be seen in Fig. 5 that both quantizers have similar performance for the three signal sources; however the LMQ tends to be slightly better at low rates (below 8 bits/sample).

Fig. 6 shows the RD curves of quantized MP expansions arriving from three different random sources in $\mathbb{R}^{64}$ for the LMQ and the ABUQ for the Gabor dictionary of 4 phases in $\mathbb{R}^{64}$, see eq. (10). For that purpose the LMQs were designed with bit-depth ranging from 1 to 6. The decompositions to be coded allowed a maximum of 256 terms. For each distinct source the results are averages over an ensemble of 200 quantized MP decompositions of signals from each source. The decompositions quantized with the ABUQ require a larger bitrate because its bit-allocation scheme does not manages to control the rate as the shape of $f_{\Theta_1}(\theta_1)$ for this dictionary implies a low decay rate of the coefficients involved in M-terms obtained using this dictionary.

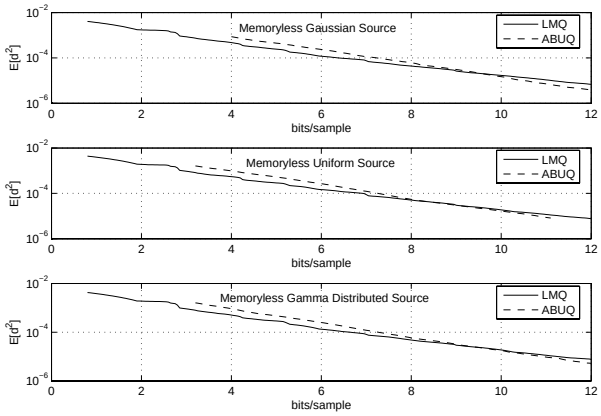

Fig. 5. LMQ and ABUQ RDs for three different random sources using the $GSND(128, 10)$, the distortion is defined in eq. (16).

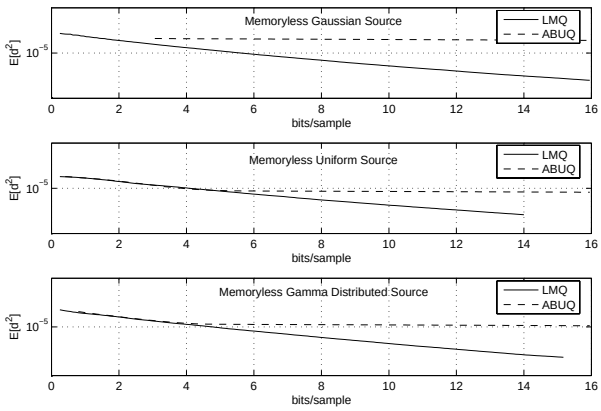

Fig. 6. LMQ and ABUQ RDs for three different random sources using the 4-phase Gabor dictionary in $\mathbb{R}^{64}$, the distortion is defined in eq. (16).

VI. Conclusion

The angles between the residues and the selected atoms in Matching Pursuit iterations can be statistically modeled as independent and identically distributed (iid) at each iteration. That is, the angles in Matching Pursuit steps can be considered statistically invariant with respect to the decomposition step. As a result, the statistics of Matching Pursuit angles can be obtained from the statistics of the first Matching Pursuit angle for signals drawn from a memoryless Gaussian source.

Based on the iid statistical model for Matching Pursuit angles, Lloyd-Max quantizers for Matching Pursuit coefficients were designed. The Lloyd-Max quantizer is designed to be the same for all the coefficients in an M-term, and its design requires both the number of terms to be coded and the probability density function of the first Matching Pursuit angle for a memoryless Gaussian source, resulting in dictionary dependent quantizers.

It is important to point out that if the source is not Gaussian the first MP angle has different statistics than the MP angles in other steps. However, even in this case, the Lloyd-Max quantizers obtained for the Gaussian source can still be employed since the first coefficient is not quantized but transmitted with negligible error as side information.

The Lloyd-Max quantization presented was compared to the state-of-the-art off-loop quantization scheme in [2] and both quantization schemes showed similar rate–distortion performance.

References

[1] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, December 1993.

[2] P. Frossard, P. Vandergheynst, R. M. Figueras I Ventura, and M. Kunt, "A posteriori quantization of progressive matching pursuit streams," *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 525–535, February 2004.

[3] L. Lovisolo, M. A. M. Rodrigues, E. A. B. da Silva, and P. S. R. Diniz, "Efficient coherent decompositions of power systems signals using damped sinusoids," *IEEE Trans. on Signal Processing*, vol. 53, pp. 3831–3846, 2005.

[4] R. Neff and A. Zakhor, "Modulus quantization for matching-pursuit video coding," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, pp. 895–912, 2000.

[5] S. Mallat, *A Wavelet Tour of Signal Processing*, 1st ed. San Diego, California, USA: Academic Press, 1998.

[6] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *Journal of Constructive Approximation*, vol. 13, pp. 57–98, 1997.

[7] A. K. Jain, *Fundamentals of Digital Image Processing*, 13rd ed., ser. Prentice Hall Information And System Sciences Series. Englewood Cliffs, NJ 07632: Prentice Hall, 1989.