Error Correction for Erasures of Quantized Frame Coefficients

Bernhard G. Bodmann[(1)], Peter G. Casazza[(2)], Gitta Kutyniok[(3)] and Steven Senger[(2)]
(1) Department of Mathematics, University of Houston, Houston, TX 77204, USA.
(2) Department of Mathematics, University of Missouri, Columbia, MO 65211, USA.
(3) Institute of Mathematics, University of Osnabrück, 49069 Osnabrück, Germany.
bgb@math.uh.edu, pete@math.missouri.edu, kutyniok@math.uni-osnabrueck.de, senger@math.missouri.edu

Abstract:

In this paper we investigate an algorithm for the suppression of errors caused by quantization of frame coefficients and by erasures in their subsequent transmission. The erasures are assumed to happen independently, modeled by a Bernoulli experiment. The algorithm for error correction in this study embeds check bits in the quantization of frame coefficients, causing a possible, but controlled quantizer overload. If a single-bit quantizer is used in conjunction with codes which satisfy the Gilbert Varshamov bound, then the contributions from erasures and quantization to the reconstruction error is shown to have bounds with the same asymptotics in the limit of large numbers of frame vectors.

1. Introduction

The versatility of redundant systems, in particular frames, has been demonstrated by their resilience to erasures and by their usefulness to suppress quantization errors. In the context of finite frames, the statistical error estimates by Goyal, Kovačević, Vetterli and Kelner [7, 6] were to the authors' knowledge the first instance of a combined analysis of erasures and quantization.

In recent years, the robustness of finite frames against erasures has been more extensively studied, for instance, in [5] or [2]. These studies typically provide estimates for the (average and worst case) blind reconstruction error, meaning all erased (unknown) coefficients are set to zero and the reconstruction relies on a fixed synthesis operator. It is well-known that if the frame vectors related to the non-erased coefficients still form a spanning set, then the frame operator of those can be inverted, leading to perfect reconstruction. However, the latency caused by the wait until all coefficients have been transmitted and the computational cost of inverting the frame operator make perfect reconstruction less practicable.

On the other hand, Benedetto, Powell and Yilmaz [1] investigated an easily implementable, active error correction for the compensation of quantization errors with so-called sigma-delta algorithms, which provide highly accurate reconstruction.

Recently, Boufounos, Oppenheim and Goyal [4] introduced an erasure correction scheme with strong similarities to quantization-noise shaping, offering the possibility of a combined treatment of both types of errors. The idea of pre-compensation and error-forward projection deserves to be explored further, but the algorithm by Boufounos, Oppenheim and Goyal is computationally still more costly than a simple application of sigma-delta quantization.

The need for results on low-complexity quantization-and-erasure correcting algorithms motivated the present study, which investigates a rather simple strategy for error compensation, a modified sigma-delta algorithm with embedded check bits. The error correction algorithm we present allows precise bounds on quantization errors and also on the effect of erasures from unreliable transmissions of frame coefficients.

2. PCM quantization and blind reconstruction

We first revisit erasure-averaged error bounds for PCM quantization of frame coefficients and blind reconstruction after transmission.

Definition. Let $\mathcal{H}$ be a d-dimensional Hilbert space. A frame $\mathcal{F} = \{f_1, f_2, \dots f_N\}$ for $\mathcal{H}$ is a spanning set. If all vectors in the frame have the same norm, we call $\mathcal{F}$ equal-norm. If $x = \frac{1}{A}\sum_{j=1}^{N}\langle x, f_j\rangle f_j$ for all $x \in \mathcal{H}$, then we say that $\mathcal{F}$ is A-tight.

Quantizing frame coefficients simply means mapping them to a finite set of values.

Definition. A function Q on $\mathbb{R}$ is called *a quantizer with accuracy* $\epsilon > 0$ *on the interval* $[-L, +L]$ if it has a finite range $\mathbb{A}$ and for any $x \in [-L, +L]$, $Q(x)$ satisfies $|x - Q(x)| \le \epsilon$. The range $\mathbb{A}$ of the quantizer Q is also called the alphabet. If this alphabet consists of all integer multiples of a fixed step-size δ contained in the interval $[-L - \delta/2, +L + \delta/2]$ and the quantizer assigns to $x \in [-L, +L]$ the unique value $m\delta, m \in \mathbb{Z}$, satisfying $(m - \frac{1}{2})\delta < x \le (m + \frac{1}{2})\delta$ then we call Q the *uniform mid-tread quantizer with step-size* δ [3]. Alternatively, if the alphabet is $\mathbb{A} = (\mathbb{Z} + \frac{1}{2})\delta \cap [-L - \delta/2, +L + \delta/2]$ and if Q assigns to $x \in [-L, +L]$ the value $(m + \frac{1}{2})\delta$ such that $m\delta < x \le (m + 1)\delta$, then we speak of the so-called *uniform mid-riser quantizer with step-size* δ. In the latter part of this study, we focus on the single-bit mid-riser quantizer which rounds the input to $\mathbb{A} = \{-\delta/2, +\delta/2\}$. We want to apply this quantizer to frame coefficients.

Definition. Given a quantizer Q, the *PCM quantization* of a vector x in a real Hilbert space $\mathcal{H}$ of dimension

$\dim(\mathcal{H}) = d$, equipped with an A-tight frame $\mathcal{F} = \{f_j\}_{j=1}^N$, is defined by

$$Q_{\mathcal{F}}(x) = \frac{1}{A}\sum_{j=1}^N Q(\langle x, f_j\rangle) f_j\,.$$

Remark. We recall that the PCM quantization error resulting from a uniform quantizer Q with accuracy $\epsilon > 0$ on $[-L, +L]$, and a N/d-tight equal-norm frame $\mathcal{F}$ applied to any input vector $x \in \mathcal{H}$ satisfying $\|x\| \le L$ is in norm bounded by

$$\|Q_{\mathcal{F}}(x) - x\| \le \max_{\|v\|=1}\max_{u_j\in\{\pm 1\}} \frac{d}{N}|\sum_{j=1}^N u_j\langle f_j, v\rangle|$$
$$\le \frac{d}{N}(\sqrt{N}\epsilon)(\sum_{j=1}^N |\langle f_j, v\rangle|^2)^{1/2} = \sqrt{d}\epsilon\,.$$

This is in contrast to erasures, where the bound on the reconstruction error depends on the norm of the input vector.
Definition. Given a probability measure $\mathbb{P}$ on the set of erasures, and the analysis operator V belonging to an A-tight frame, we define the erasure-averaged reconstruction error to be

$$e(V, \mathbb{P}) = \mathbb{E}[\|\frac{1}{A}V^* E(\omega) V - I\|]\,.$$

Hereby, $\mathbb{E}[\cdot]$ is the expectation with respect to the probability measure $\mathbb{P}$ on $\Omega = \{0,1\}^N$, and $E : \Omega \to \mathbb{R}^{N\times N}$ is a random diagonal matrix with entries $E_{j,j} = \omega_j$.
Theorem. Let $\mathcal{H}$ be a real Hilbert space of dimension d, equipped with an A-tight equal-norm frame $\mathcal{F}$. If all the frame coefficients are erased with a probability $0 \le p \le 1$, independently of each other, then the erasure-averaged reconstruction error is bounded by

$$p \le \mathbb{E}[\|\frac{1}{A}V^* E(\omega) V - I\|] \le pd\,.$$

Proof. The lower bound uses Jensen's inequality and the convexity of the norm on the real vector space of Hermitian operators [8]. The upper bound relies on the identity for the operator norms $\|V^*(I-E)V\| = \|(I-E)VV^*(I-E)\|$ and on the bound for entries in the Grammian, $|(VV^*)_{j,k}| \le \|f_j\|\|f_k\| = 1$, derived from the Cauchy-Schwarz inequality, which implies $\mathbb{E}[\|(I - E(\omega))VV^*(I-E(\omega))\|] \le Np$. □
Thus, for a vector x for which $p\|x\|$ is bigger than $(\delta/2)\sqrt{d}$, the bound on the worst case error due to erasures dominates that of PCM quantization.
A similar phenomenon happens when the quantization is obtained with first and higher-order sigma delta quantization. For sufficiently large N, the bound for the worst-case quantization error, see e.g. [3], is smaller than the worst-case erasure error. This motivates investigating active error correction for erasures.

3. Sigma-delta quantization with embedded check bits

Our main goal is to make the two error bounds for erasures and quantization comparable. To this end, we use systematic binary error-correcting codes for packets of quantized coefficients, and replace a portion of the output from the sigma-delta quantizer by the check bits.
Definition. A binary (n, k)-code is an invertible map

$$C : \mathbb{Z}_2^k \to \mathbb{Z}_2^n\,.$$

The *minimum distance* of this code is the minimal number of bits by which any two code words (elements in the range of C) differ. A *systematic* (n, k)*-code* simply appends check bits, meaning $q = (q_1, q_2, \dots q_k)$ maps to $C(q) = (q_1', q_2', \dots q_n')$ such that $q_j' = q_j$ for all $j \in \{1, 2, \dots k\}$.
The relevance of this definition is that among any block of n transmitted bits, the minimum distance is the number of bit erasures that cannot be corrected any more.
As already mentioned, we will exploit a particular accompanying quantization strategy, which we briefly explain.
Definition. Let Q be the binary mid-riser quantizer with stepsize $\delta > 0$ and let $\mathcal{F} = \{f_1, f_2, \dots f_N\}$ be an N/d-tight frame for a d-dimensional real Hilbert space $\mathcal{H}$. Also, assume that C is a binary (n, k)-code. Given an input vector $x \in \mathcal{H}$, then the C*-embedded sigma-delta quantization of* x *is* $Q_{\mathcal{F},C}(x) = \frac{d}{N}\sum_{j=1}^N q_j f_j$, *where the sequence* $\{q_j\}_{j=1}^\infty$ associated with the initialization value $u_0 = 0$ is defined by

$$q_{m+j} := \begin{cases} Q(\langle x, f_{m+j}\rangle + u_{m+j-1}), & 1 \le j \le k\,,\\ C((q_{m+1}, q_{m+2}, \dots q_{m+k}))_j, & \text{else}\,,\end{cases}$$

for any $m \in \{0, n, 2n, \dots\}$, and $j \in \{1, 2, \dots n\}$, and the map for updating the internal variable is

$$u_{m+j} := \langle x, f_{m+j}\rangle - q_{m+j} + u_{m+j-1}\,.$$

Our first main theorem is the stability of this modified sigma-delta algorithm.
Theorem. Let Q be a binary mid-riser quantizer with stepsize $\delta > 0$, let $\mathcal{F} = \{f_1, f_2, \dots f_N\}$ be an N/d-tight equal-norm frame for a d-dimensional real Hilbert space $\mathcal{H}$, and let C be a systematic binary (n, k)-code, such that n divides N. If $\|x\| \le \alpha\delta/2$, $\alpha < 1$, and

$$k \ge \frac{n}{2}(1+\alpha)$$

then in the course of the C-embedded first-order sigma-delta quantization, the internal variable is bounded by

$$|u_j| \le \frac{\delta}{2}((n-k+1) + (n-k)\alpha) \le \delta(k - \frac{k^2}{n} + \frac{1}{2})$$

for all $j \in \{1, 2, \dots N\}$.
Proof. We proceed by induction. At the end of the first block of n bits, if all $n-k$ check bits were chosen incorrectly and the input is taken to be the worst case, then u_N reaches the maximum magnitude stated in the theorem. In the course of quantizing the next block, due to the bound on the input, each bit allows the quantizer to recover at least $\delta/2 - \alpha\delta/2$. With the inequality $k \ge \frac{n}{2}(1+\alpha)$ we deduce

$$k(\frac{1}{2} - \frac{\alpha}{2}) \ge \frac{1}{2}(n-k)(1+\alpha)$$

which means u_j is contained in $[-\delta/2, \delta/2]$ before the next check bit is encountered. □

Similarly as in [1] and [3], we deduce an error estimate from the bound on the internal variable.
The relevant quantity in this estimate is derived from the frame geometry, as in [3],

$$T(\mathcal{F}) = \|(f_1 - f_2) \pm (f_2 - f_3) \pm \cdots \pm (f_{N-1} - f_N) \pm f_N\| \,.$$

We define the maximal error caused by quantization to be

$$eq(V, \delta, \alpha) = \max_{\|x\| \le \alpha\delta/2} \|Q_{\mathcal{F},C}(x) - x\| \,,$$

where V is the analysis operator of the frame $\mathcal{F}$.
Theorem. Under the same assumptions as in the preceding theorem,

$$eq(V, \delta, \alpha) \le \frac{d}{N}(\frac{\delta}{2}((n-k)(1+\alpha)+1)T(\mathcal{F}) \,.$$

In comparison with the unmodified first order sigma-delta quantization, we have a bound that is worse by at most a factor of $2(n-k)$. However, the advantage of the embedded check bits is the ability to correct erasures in each block.
Assume the initial probability measure applies an erasure with a probability of p to each coefficient. Assume that the code C has minimal distance $np + t$ with $t > 0$. Let $\mathbb{P}'$ denote the probability measure governing the erasures remaining after the error correction has been applied in each block of length n.
Definition. The combination of quantization, erasures and error correction gives the reconstruction error

$$ec(V, \delta, \alpha, \mathbb{P}') = \mathbb{E}[\max_{\|x\| \le \alpha\delta/2} \|\frac{1}{A}\sum_j \omega_j q_j f_j - x\|] \,,$$

where $\omega_j = 0$ means that the j-th coefficient is erased.
The following lemma helps bound the probability of erasures remaining, if the weight of the code is larger than the expected number of erasures before correction.
Lemma. (Hoeffding). Let $\mathbb{E}[\omega_j] = 1 - p$ and assume that the minimum distance of C is bounded below by $n(p+\epsilon)$, $\epsilon > 0$. The probability $p'$of an individual coefficient being erased after the error correction is applied is bounded by

$$p' \le \exp(-2n\epsilon^2) \,.$$

Now we can combine the two error estimates for quantization and erasures.
Theorem. Let $\epsilon > 0$, assume C has minimal distance $n(p+\epsilon)$. Let $\mathbb{P}'$ be the probability measure governing the erasures after the error correction has been applied. Under the additional assumptions of the preceding theorem,

$$ec(V, \delta, \alpha, \mathbb{P}') \le eq(V, C, \delta, \alpha) + d\delta \exp(-2n\epsilon^2) \,.$$

Proof. First we apply Minkowski's inequality to separate the error caused by quantization and by erasures. The expected number of erasures is Np', with p' bounded in accordance with the preceding lemma. Each erased coefficient has magnitude $\delta/2$, so the norm of the vectors which are omitted in the reconstruction can at most be $\delta dp'/2$. □

The remaining question is which asymptotics can be achieved for the minimum distance with a suitable sequence of codes.
To this end, we quote a version of the Gilbert-Varshamov bound.
Lemma. Let $0 \le q \le 1/2$, then there exist infinitely many systematic linear (n,k)-codes with minimum distance at least nq and rate

$$\frac{k}{n} \ge 1 - H_2(q) \,,$$

where $H_2(q) = -q\log_2 q - (1-q)\log_2(1-q)$ is the binary entropy.
Proof. The usual form of the Gilbert Varshamov bound for linear codes can be re-stated as a bound for the maximal number of erasures that can be corrected by certain codes. In this form, it states the existence of linear codes for which any $n - d + 1$ rows of the generator matrix have rank k if $d \ge nq$, meaning up to $d - 1$ erasures can be corrected. Permuting the rows so that the first k have maximal rank and right-multiplying by the inverse of this $k \times k$ block gives the generator matrix for a systematic code that can correct the same number of erasures.
We are ready to state the final result.
Theorem. Let $0 \le p < q \le 1/2$, $H_2(q) \le (1-\alpha)/2$, $0 < \alpha < 1$ and denote $\epsilon = q - p$. Consider the sequence of systematic linear codes provided by the Gilbert-Varshamov bound for minimum distance bounded below by nq and let $N \ge ne^{2n\epsilon^2}$, then

$$\begin{aligned} ec(V, \delta, \alpha, \mathbb{P}') \le & \frac{d\delta}{2N}((2\ln N H_2(q)/\epsilon^2 + 1)T(\mathcal{F}) \\ & + \frac{1}{2\epsilon^2}\ln N) \,. \end{aligned}$$

Proof. From the assumption, we have $e^{2n\epsilon^2} \le N$ and thus $n \le \frac{1}{2\epsilon^2}\ln N$. By the Gilbert-Varshamov bound,

$$n - k \le nH_2(q) \le \frac{1}{2\epsilon^2}\ln N H_2(q) \,.$$

Using the Hoeffding inequality on the error due to the remaining erasures gives

$$e^{-2n\epsilon^2} \le \frac{n}{N} \le \frac{1}{2\epsilon^2}\frac{\ln N}{N} \,.$$

Thus, the two error terms have the same asymptotic behavior.
We note that this error bound is worse by a term logarithmic in N compared to the quantization error without erasures.

Acknowledgment

This work was partially supported by NSF DMS 07-04216, NSF DMS 08-07399 and by the Deutsche Forschungsgemeinschaft (DFG) under Heisenberg Fellowship KU 1446/8-1.

References:

[1] John J. Benedetto, Alexander M. Powell, and Ozgur Yilmaz, Sigma-Delta quantization and finite frames, *IEEE Trans. Inform. Theory* 52:1990–2005, 2006.

[2] Bernhard G. Bodmann and Vern I. Paulsen. Frames, graphs and erasures. *Linear Algebra Appl.* 404:118–146, 2005.

[3] Bernhard G. Bodmann and Vern I. Paulsen. Frame Paths and Error Bounds for Sigma-Delta Quantization. *Appl. Comput. Harmon. Anal.* 22:176–197, 2007.

[4] Petros Boufounos, Alan V. Oppenheim, and Vivek K. Goyal. Causal Compensation for Erasures in Frame Representations. *IEEE Trans. Signal Proc.* 3:1071–1082, 2008.

[5] Peter Casazza and Jelena Kovačević, Equal-norm tight frames with erasures. (English summary) Frames. *Adv. Comput. Math.* 18:387–430, 2003.

[6] Vivek K. Goyal, Jelena Kovačević, and Jonathan A. Kelner. Quantized frame expansions with erasures. *Appl. Comp. Harm. Anal.*, 10:203–233, 2001.

[7] Vivek K. Goyal, Jelena Kovačević, and Martin Vetterli. Quantized frame expansions as source-channel codes for erasure channels. In: Proc. Data Compr. Conf., Snowbird, UT, Mar. 1999.

[8] D. Petz, Spectral scale of self-adjoint operators and trace inequalities, J. Math. Anal. Appl. **109** (1985), 74–82.