Angle-preserving Quantized Phase Embeddings

Boufounos, P.T.

TR2013-075 September 2013

Abstract

We demonstrate that the phase of randomized complex-valued projections of real-valued signals preserves information about the angle, i.e., the correlation, between signals. This information can be exploited to design quantized angle-preserving embeddings, which represent such correlations using a finite bit-rate. These embeddings generalize known results on binary embeddings and 1-bit compressive sensing and allow us to explore the trade-off between number of measurements and number of bits per measurement, given the bit rate. The freedom provided by this trade-off, which has also been observed in quantized Johnson-Lindenstrauss embeddings, can improve performance at reduced rate in a number of applications.

SPIE Conference on Wavelets and Sparsity

Angle-preserving Quantized Phase Embeddings

Petros T. Boufounos
Mitsubishi Electric Research Laboratories, Cambridge, MA, USA

ABSTRACT

We demonstrate that the phase of randomized complex-valued projections of real-valued signals preserves information about the angle, i.e., the correlation, between signals. This information can be exploited to design quantized angle-preserving embeddings, which represent such correlations using a finite bit-rate. These embeddings generalize known results on binary embeddings and 1-bit compressive sensing and allow us to explore the trade-off between number of measurements and number of bits per measurement, given the bit rate. The freedom provided by this trade-off, which has also been observed in quantized Johnson-Lindenstrauss embeddings, can improve performance at reduced rate in a number of applications.

Keywords: Quantized angle embeddings, phase measurements, phase-only compressive sensing

1. INTRODUCTION

Randomized embeddings play an increasingly important role in signal processing. Embeddings map signals to a space that is computationally favorable, while preserving some aspect of signal geometry. Computationally efficient operations on the embedded data can, therefore, directly map to operations in the original space. In many cases, a stable embedding is also invertible, given the class of signals being embedded, thus providing a mechanism for robust signal acquisition and measurement.

The most commonly used embeddings are due to Johnson and Lindenstrauss (J-L).[1] J-L embeddings preserve the Euclidean, i.e., ℓ_2, distance between signals in a finite signal set, while representing them in a lower dimensional space. J-L embeddings were recently rediscovered by the signal processing community in the context of compressive sensing, due to their strong connection with the Restricted Isometry Property (RIP).[2–5] The RIP is, in fact, a reformulation of the J-L embedding property, established for sets of sparse signals.

In this paper we describe quantized angle-preserving embeddings. In other words, we describe embeddings that encode signals into a quantized representation that approximately preserves the angles between pairs of signals in the original signal space. The embeddings we describe can be considered generalizations of binary ϵ-stable embeddings (BeSE), which preserve the angles between pairs of signals using binary representations.

Angle-preserving embeddings are useful in applications where the appropriate signal metric is correlation of signals rather than euclidean distance. This is the case, for example, in detection and estimation applications that require invariance to signal normalization, or use signals that always have the same norm.

This work also extends earlier results on the importance of phase information in recovering signals, with a number of practical applications.[6–9] In summary, the phase of a fully sampled Fourier transform of a signal contains, under a variety of conditions, sufficient information to uniquely specify the signal and enable its reconstruction within a scaling factor. Our results further demonstrate that the phase incorporates correlation information about the signals and preserves their correlation. Combined with our work on reconstruction from those measurements,[10] our results transfer these classical results to the compressive sensing framework. While we establish the results using random matrices with i.i.d. normal entries, we conjecture that a large variety of distributions could be used, including subsampled Fourier transforms.

In the next section we provide a brief background on embeddings, which also helps in establishing notation. Section 3 describes the proposed angle-preserving embeddings and their properties. Section 4 presents experimental results demonstrating the performance of the embeddings under a variety of conditions. Section 5 discusses our results and concludes.

Further author information: E-mail: petrosb@merl.com, Telephone: +1-617-621-7575.

2. BACKGROUND

2.1 Embeddings Preserving the Euclidean distance

The best known embedding family is Johnson-Lindestrauss (J-L) embeddings, which can reduce dimensionality of a finite signal set while approximately preserving the ℓ_2 distance between points in the set.[1] Specifically, they are functions $f : \mathcal{S} \to \mathbb{R}^K$ that map a finite set of L signals, $\mathcal{S} \subset \mathbb{R}^N$, to an M-dimensional vector space such that the images of any two signals $\mathbf{x}$ and $\mathbf{x}'$ in $\mathcal{S}$ satisfy:

$$(1-\epsilon)\|\mathbf{x}-\mathbf{x}'\|_2^2 \le \|f(\mathbf{x})-f(\mathbf{x}')\|_2^2 \le (1+\epsilon)\|\mathbf{x}-\mathbf{x}'\|_2^2. \tag{1}$$

Since the original work of Johnson and Lindenstrauss, which demonstrated the existence of such embeddings in $M = O(\log L)$ dimensions, a large body of work has established that a variety of randomly generated linear maps can be used to generate J-L embeddings.[11,12]

More recently, in the context of compressed sensing, the restricted isometry property (RIP) has been established for K-sparse signal sets $\mathcal{S}$. A randomized linear map $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \mathbb{R}^{M\times N}$ satisfies (1) with overwhelming probability as long as $M = O(K \log N/K)$.[3,4] The RIP is sufficient to guarantee signal recovery in compressive sensing (CS).[2] The same guarantees have also been extended to other models such as signals in a union of subspaces or in a manifold.[13,14]

When such embeddings are used for signal storage or transmission, their performance under a finite bit-rate is important. Specifically, when these ℓ_2 embeddings are uniformly quantized to B bits per dimension, the embedding guarantee becomes[15,16]

$$(1-\epsilon)\|\mathbf{x}-\mathbf{x}'\|_2 - 2^{-B+1}S \le \|f(\mathbf{x})-f(\mathbf{x}')\|_2 \le (1+\epsilon)\|\mathbf{x}-\mathbf{x}'\|_2 + 2^{-B+1}S,$$

where S is the quantizer saturation level (assumed set to ensure negligible saturation). Quantized embeddings are essential in applications, such as augmented reality, that require efficient transmission for pattern matching. These embeddings preserve the distances between all pairs of signals. In addition, recent work has demonstrated that if the application does not require accurately preserving the full range of distances but only a small range, then the bit-rate can be further reduced.[17]

2.2 Embeddings Preserving Correlations

While ℓ_2 distance is the appropriate metric in a number of applications, correlation, i.e., angle between signals, is often preferable. Of course, inner product can always be inferred from distances using

$$\|\mathbf{x}-\mathbf{x}'\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{x}'\|_2^2 - 2\langle \mathbf{x},\mathbf{x}'\rangle \Longrightarrow \langle \mathbf{x},\mathbf{x}'\rangle = (\|\mathbf{x}\|_2^2 + \|\mathbf{x}'\|_2^2 - \|\mathbf{x}-\mathbf{x}'\|_2^2)/2. \tag{2}$$

When ℓ_2 embeddings are used to estimate this angle we can trivially obtain the following embedding guarantee

$$(1-\epsilon)\langle \mathbf{x},\mathbf{x}'\rangle - \epsilon\|\mathbf{x}-\mathbf{x}'\|_2^2 \le \langle f(\mathbf{x}), f(\mathbf{x}')\rangle \le (1+\epsilon)\langle \mathbf{x},\mathbf{x}'\rangle + \epsilon\|\mathbf{x}-\mathbf{x}'\|_2^2. \tag{3}$$

Note that this guarantee contains a term in the true distance of the signals. This implies that as signals become larger, even if their correlation is the same, the embedding guarantee becomes looser. While similar, and often tighter, bounds can and have been shown,[18,19] they still suffer from the same issue: the error guarantee becomes looser as the signal magnitude increases.

Correlations are also equivalent to angles; the normalized acute angle between signals is defined as

$$d_\angle(\mathbf{x},\mathbf{x}') = \frac{1}{\pi}\arccos\frac{\langle \mathbf{x},\mathbf{x}'\rangle}{\|\mathbf{x}\|\|\mathbf{x}'\|} \tag{4}$$

Using this definition, it is possible to show that RIP-type embeddings preserve angles between signals, without dependence on the signal magnitude, as long as the signals are sparse.[20] In contrast to earlier results, the proof explicitly uses the sparsity of the signals and their projection to remove the dependency on the signal magnitude.

We should note that since angles are related to correlations through a cosine transformation and its inverse, the error guarantees provided on angle embeddings are distorted through the same transformation when transferred to correlations. Existing techniques in the literature enable us to characterize the ambiguity in this transformation,[17] although we do not attempt this in this paper.

Another approach to remove the dependency on the signal magnitude is to use a transformation explicitly designed invariant to signal scaling. This was first established in the context of 1-bit compressed sensing.[21,22] Specifically, the transformation $f(\mathbf{x}) = \text{sign}(\mathbf{Ax})$, where $\text{sign}(\cdot)$ is applied element-wise to preserve only the sign of its argument, has the property $f(\mathbf{x}) = f(c\mathbf{x})$ for all $c > 0$. The embedding does not preserve signal distances but it can preserve angles between them. In particular, if the elements of $\mathbf{A}$ are chosen from an i.i.d. normal Gaussian distribution then $f(\cdot)$ is a binary ϵ-stable embedding (BeSE), i.e., the hamming distance $d_H(\cdot,\cdot)$ between the embedded signals $\mathbf{q} = f(\mathbf{Ax})$, and $\mathbf{q}' = f(\mathbf{Ax}')$ satisfies

$$|d_H(\mathbf{q},\mathbf{q}') - d_\angle(\mathbf{x},\mathbf{x}')| \leq \epsilon. \tag{5}$$

The embedding holds with probability greater than $1 - 2e^{2\log L - 2\epsilon^2 M}$ on a set of signals of size L. As with most randomized embedding proofs,[3] this can be shown using a union bound on the probability that this is satisfied for a pair of signals. With some more care, a similar guarantee can be extended to infinite sets, such as sparse signals.[21]

3. GENERALIZED ANGLE EMBEDDINGS

3.1 Continuous Phase Embeddings

Instead of binary embeddings, this work considers their generalization to phase embeddings, i.e., embeddings in which only the phase of a linear projection is measured. The phase function is a generalization of the sign function which, in addition, takes a continuous range of values. However, to measure a phase we require the measurements to be complex, which will not be the case if the signal and the measurement matrix are real-valued. Thus, since the signal is real-valued, we expose phase and obtain phase measurements by first projecting it to a complex-valued space through a complex linear transformation, i.e., using

$$\mathbf{z} = \mathbf{Ax},\ \mathbf{y} = \angle(\mathbf{z}) = \angle(\mathbf{Ax}), \tag{6}$$

where $\angle(\cdot)$ measures the principal phase of a complex number and is applied element-wise to its vector argument. The measurement matrix $\mathbf{A} \in \mathbb{C}^{K\times N}$ is randomly drawn with i.i.d. elements drawn from the standard complex normal distribution. These embeddings generalize the BeSE in the same sense that the phase of complex numbers generalizes the sign of real numbers.

To show that this transformation preserves the angles between signals we first use the expected value of the absolute phase difference of a single pair of projection coefficients given their angle. This can be computed by appropriately integrating the density of the phase difference.[23] Given a pair of signals $\mathbf{x}$, $\mathbf{x}'$, we obtain

$$E\left\{\left|\angle\left(\frac{z_m}{z'_m}\right)\right|\right\} = E\left\{\left|\angle\left(e^{i(y_m - y'_m)}\right)\right|\right\} = \pi d_\angle(\mathbf{x},\mathbf{x}'), \tag{7}$$

where $\angle(z_m/z'_m) = \angle\left(e^{i(y_m-y'_m)}\right)$ computes the phase difference of the m^{th} measurement of the two signals, taking phase wrapping appropriately into account. Next, we use Hoeffding's inequality to bound the probability that the average of M random variables $|\angle(e^{i(y_m-y'_m)})|$ deviates from (7). Using the union bound on L^2 point pairs, a property reminiscent of Johnson-Lindenstrauss (JL) embeddings[1] follows.

Theorem 3.1. *Consider a finite set $\mathcal{S} \subset \mathbb{R}^N$ of L points measured using (6), with $\mathbf{A} \in \mathbb{C}^{M\times N}$ consisting of i.i.d elements drawn from the standard complex normal distribution. With probability greater than $1 - 2e^{2\log L - 2\epsilon^2 M}$ the following holds for all $\mathbf{x},\mathbf{x}' \in \mathcal{S}$ and corresponding measurements $\mathbf{y},\mathbf{y}' \in \mathbb{R}^M$.*

$$|d_E(\mathbf{y},\mathbf{y}') - d_\angle(\mathbf{x},\mathbf{x}')| \leq \epsilon, \tag{8}$$

where the embedding distance $d_E(\mathbf{y},\mathbf{y}')$ is defined as

$$d_E(\mathbf{y},\mathbf{y}') \triangleq \frac{1}{M}\sum_m \left|\frac{1}{\pi}\angle\left(e^{i(y_m-y'_m)}\right)\right| \tag{9}$$

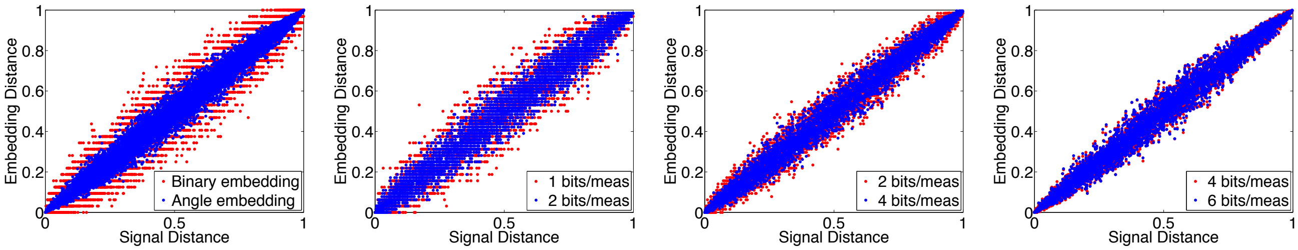

Figure 1. Angle angle embeddings capture the true angle between signals with uncertainty depending to quantization rate. Left to Right: Unquantized angle embedding vs. Binary embedding, quantized angle embedding: 1 vs. 2 bits/measurement, 2 vs. 4 bits/measurement, 4 vs. 6 bits/measurement. As the quantization rate increases, performance improves. However, beyond 4 bits/measurement performance is does not improve noticeably.

Thus, similar to J-L embeddings, $K = O(\log L)$ dimensions are sufficient to embed a set of L points. Using a well-established covering argument[3] and some care to take discontinuities due to phase wrapping into account, the theorem can be extended to infinite signal sets such as sparse signals.[10]

Note that, similar to all guarantees of such nature, the embedding guarantee is probabilistic on the measure of $\mathbf{A}$. In other words, the embedding guarantees that once $\mathbf{A}$ is drawn, with high probability it satisfies (8) for all pairs of signals in $\mathcal{S}$. However, a failure of the embedding guarantee only requires that there is a single pair of signals in $\mathcal{S}$ for which (8) is not satisfied. In practice we observe a gradual behavior, i.e., as M decreases and the embedding starts to fail, a small relaxation of the error bound ϵ is sufficient to make the embedding hold again. In other words, the failure of the inequality is not catastrophic but gradual. Thus the guarantees provided by the theorem are often too conservative in practice.

While the embeddings can be used to measure and reconstruct signals within a scaling factor using a simple convex program,[10] such reconstruction is not discussed in this paper. However, embedding and reconstruction experiments do suggest that the bounds of Thm. 3.1 are not tight; it should be possible to provide multiplicative instead of additive ambiguity, similar to J-L embeddings[1] and angle to embeddings of sparse signals through linear projections.[20]

3.2 Quantized Phase Embeddings

In a number of applications, embeddings are used not only for dimensionality reduction during processing but also for data reduction during transmission. In this case, the embedding should produce a quantized finite bit-rate output; the total bit-rate is an important part of the design.

In the remainder we consider scalar quantization of the phase measurements

$$\mathbf{y} = \angle(\mathbf{A}\mathbf{x}),\ \mathbf{q} = Q(\mathbf{y}) = Q(\angle(\mathbf{A}\mathbf{x})), \tag{10}$$

where $Q : \mathbb{R} \longrightarrow \mathbb{Q} \subset \mathbb{R}$ is a scalar quantizer which is applied element-wise to each coefficient in its argument and uses B bits per coefficient, and $\mathbb{Q}$ is the 2^B element quantizer alphabet. Thus, the total rate used by the quantized representation $\mathbf{q}$ is equal to $R = MB$.

In designing the quantizer, it is important to note that each measurement y_m is uniformly distributed in $(-\pi, \pi]$. Using Lloyd-Max optimality conditions,[24,25] it is straightforward to show that the optimal quantizer is uniform with quantization interval $2\pi/2^B = 2^{-B+1}\pi$. In other words, the worst case error per coefficient is $2^{-B}\pi$ and the total quantization error is bounded by

$$\|\mathbf{q} - \mathbf{y}\|_2 \leq 2^{-B}\pi\sqrt{M} \tag{11}$$

Thus, using the triangle inequality, it is straightforward to extend Thm. 3.1 to quantized phase embeddings.

Theorem 3.2. *Consider a finite set $\mathcal{S} \subset \mathbb{R}^N$ of L points measured using* (10)*, with $\mathbf{A} \in \mathbb{C}^{M\times N}$ consisting of i.i.d elements drawn from the standard complex normal distribution and $Q(\cdot)$ a uniform B-bit scalar quantizer* *in the interval* $[-\pi, \pi)$. *With probability greater than* $1 - 2e^{2\log L - 2\epsilon^2 M}$ *the following holds for all* $\mathbf{x}, \mathbf{x}' \in \mathcal{S}$ *and corresponding measurements* $\mathbf{q}, \mathbf{q}' \in \mathbb{Q}^M \subset \mathbb{R}^M$.

$$|d_E(\mathbf{q}, \mathbf{q}') - d_\angle(\mathbf{x}, \mathbf{x}')| \le \epsilon + 2^{-B+1}\pi \tag{12}$$

Of course, as with Thm. 3.1, we can generalize Thm. 3.2 to infinite sets, such as sparse signals. In fact, for any set $\mathcal{S}$ for which we can demonstrate the unquantized embedding properties of Thm. 3.1 we can also trivially demonstrate the quantized embedding properties of Thm. 3.2, with the same probability of failure.

These quantized embeddings enable the generalization of binary stable embeddings to multi-bit ones by exploiting the mechanics of continuous phase embeddings. Thus, similar to binary stable embeddings, quantized phase embeddings eliminate magnitude information from signals but preserve sufficient information to allow angle computations.

However, the proposed embeddings allow greater design freedom, compared to binary ones. In particular, given a fixed available rate $R = MB$, the designer can allocate more bits per coefficient B and use fewer coefficients M, or vice versa. For a fixed probability that the embedding is guaranteed, as M increases the additive error ϵ decreases as $O(1/\sqrt{M})$. On the other hand, as B increases, the quantization error decreases exponentially. Of course, when $B = 1$ and $M = R$ we obtain binary stable embeddings as a special case of quantized phase embeddings.

This, or similar, trade-offs have also been shown in the literature for quantized J-L embeddings[15,16] and universal embeddings.[17] Unfortunately, to our knowledge, there is no principled method to explore this trade-off theoretically, only experimentally. One of the main issues in attempting theoretical analysis is that the probabilistic guarantees provided with embedding theorems, such as Thm. 3.2, are not sufficiently tight. Optimizing a bound on the performance does not necessarily optimize the performance in practice. Furthermore, as we discuss in the previous section, even if the bounds were tight, the failure is not catastrophic but gradual. Therefore, the nature of the guarantee is too conservative for practical purposes and, thus, not appropriate for practical design guidance.

The next section presents experimental results on quantized phase embeddings. We also examine the trade-off between bits and number of measurements experimentally and show that finding the optimal bit allocation given the rate is not straightforward.

4. EXPERIMENTAL RESULTS

To experimentally verify the performance trade-off, we tested our results on randomly generated pairs of signals with different angles between them. The results shown here are for signals of dimension $N = 1024$ and 10000 pairs of signals per experiment. However, experiments were also performed for a variety of conditions with consistent results.

Figure 1 plots the signal distance, as measured by (4), the normalized angle between each pair, versus the embedding distance, as measured by (9), for an $M = 32$-dimensional embedding and a variety of quantization conditions. Each plot compares two different quantization conditions and each point corresponds to a randomly generated pair of signals. The leftmost plot compares unquantized phase embeddings with binary stable embeddings, i.e., phase embedding quantized to 1 bit per measurement. The remaining plots compare quantized phase embeddings. From left to right: 1 bit vs. 2 bits per measurement, 2 bits vs. 4 bits per measurement and 4 bits vs. 6 bits per measurement.

Ideally, if $\epsilon = 0$, all points should be on the diagonal of the plot. Since the embeddings are approximate, the point clouds have some thickness. The thinner the point cloud, the better the performance of the embedding. As evident from the plots, the experiments confirm our intuition. The finer the quantization per measurement, the better the performance of the embedding. However, as shown in the rightmost plot, the performance does not improve noticeably beyond 4 bits per measurement. We should also note that as the signals become more correlated, i.e., have normalized angle closer to 0 or 1, the embedding becomes tighter. This suggests that the additive form of the error in Thm. 3.1 provides a loose bound. We conjecture that a multiplicative error bound should be possible.

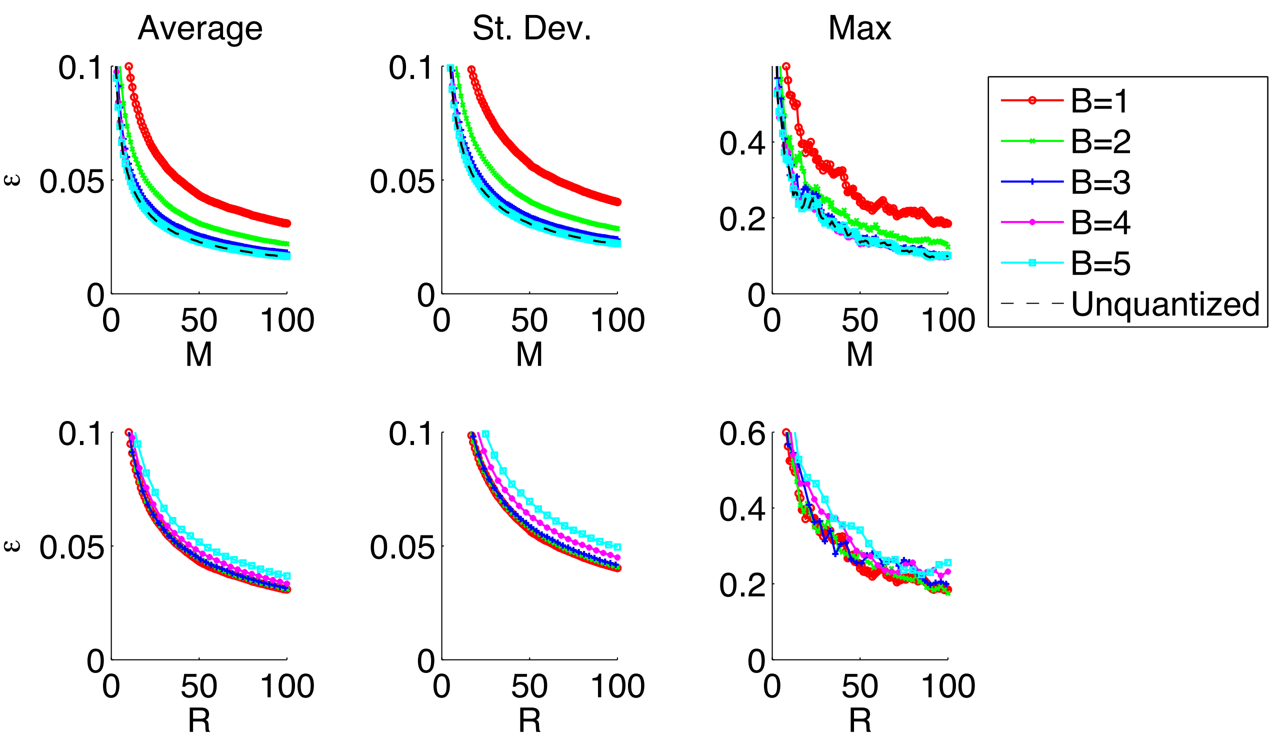

Figure 2. Error performance results on quantized embeddings for 10000 pairs of signals drawn uniformly with angles in the full range $[0, 1]$. Left to right: average, standard deviation and maximum absolute error between $d_\angle$ and d_E. Top row: performance as a function of the total number of measurements M. Bottom row: performance as a function of the total rate $R = MB$ used. Binary ($B = 1$) embeddings perform best in this usage scenario.

Since the embedding dimension is fixed at $M = 32$, as quantization becomes finer, the rate R used by the embedding increases. Thus, to perform a fair comparison, we performed experiments keeping the rate $R = MB$ fixed. Furthermore, as we discuss above, the worst case error can be a very conservative estimate of the embedding performance. Thus, to get a better sense of the performance, in our experiments we also examine the mean and standard deviation of the embedding error $|d_E(\mathbf{q}, \mathbf{q}') - d_\angle(\mathbf{x}, \mathbf{x}')|$.

Figure 2 shows experimental results plotting, from left to right, the mean, variance and maximum embedding error among all 10000 signal pairs generated for the experiment. The top row shows how these error metrics change as the embedding dimensionality M increases, for $B = 1, \ldots, 5$ bits per dimension. The bottom row shows how the error metrics change as the total embedding rate $R = MB$ increases.

The top row of plots confirms our previous findings: As the embedding dimensionality M increases, the performance improves for all error metrics. Furthermore, as the number of bits per dimension increases, for a fixed number of dimensions, the performance increases. Note that the performance at $B = 5$ bits per dimension is almost indistinguishable from the performance of unquantized embeddings, plotted using a dashed black line on the top row. More refined quantization does not significantly improve performance.

The bottom row of plots, however, demonstrates that if we consider the total rate, $R = MB$ used by the embedding, then the performance is better if we use fewer bits per dimension and higher dimensional embeddings. Specifically, they suggest that binary embeddings (i.e., BeSE) perform optimally compared to all other quantization methods. Still, the performance difference between $B = 1, 2$, and 3 bits per measurement, is very small, almost indistinguishable. Beyond $B = 4$ bits per measurement and $M = R/4$ measurements the performance begins to degrade.

The plots in Fig. 2 demonstrate results for pairs of signals with angles drawn uniformly from all possible angles, i.e., $d_\angle(\mathbf{x}, \mathbf{x}') \in [0, 1]$. The performance behavior changes if we are only interested in a smaller range

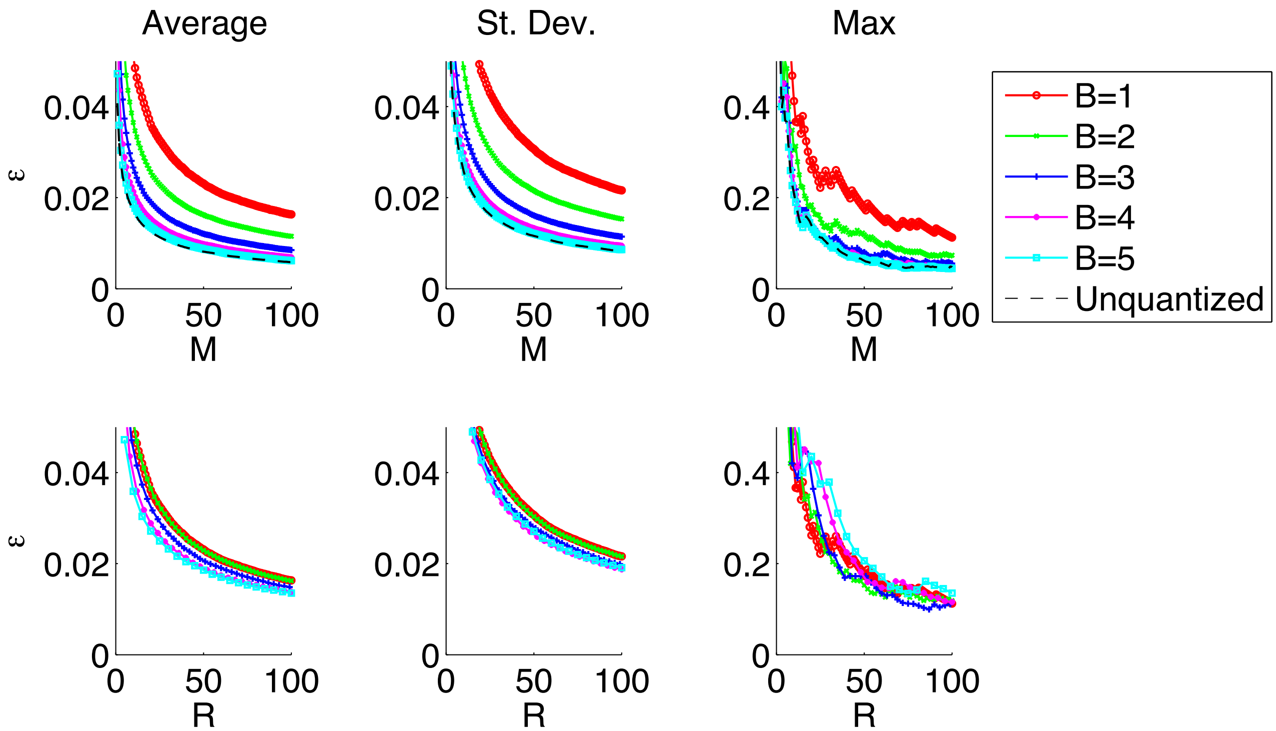

Figure 3. Error performance results on quantized embeddings for 10000 pairs of signals drawn uniformly with small angles in the range $[0, .2) \cup (.8, 1]$. Left to right: average, standard deviation and maximum absolute error between $d_\angle$ and d_E. Top row: performance as a function of the total number of measurements M. Bottom row: performance as a function of the total rate $R = MB$ used. Higher-rate ($B = 4$ or 5) embeddings perform best in this usage scenario.

of signal angles. For example, Fig. 3 shows simulation results for signal pairs with angles drawn uniformly in $d_\angle(\mathbf{x}, \mathbf{x}') \in [0, .2) \cup (.8, 1]$, i.e., only for pairs of similar signals.

As evident in the top row of Fig. 3, for a fixed number of measurements, the performance improvement with respect to the number of bits per measurement B is more noticeable, compared to Fig. 2. The implications are evident in the bottom row of Fig. 3. In contrast to the previous results, average performance in this case is better at $B = 4$ or 5 bits per coefficient.

This interesting behavior occurs because, as we discuss above and demonstrate in Fig. 1 the embedding is tighter on smaller angles. Thus, the embedding error, even with lower embedding dimension, is small and the quantization error dominates the performance. In this case it is beneficial to reduce the embedding dimension M and increase the bits per dimension B while maintaining the same rate R.

5. DISCUSSION

Our work generalizes binary ϵ-stable embeddings (BeSE) and enables multi-bit embeddings explicitly designed to preserve angles between signals. This allows us to explore the trade-off between quantization rate and embedding dimensionality. As we show experimentally, if we are only interested in preserving the angles between similar signals then it is better to use lower-dimensional embeddings, quantized at a higher rate per dimension. On the other hand, if the application requires preserving a large range of angles, even for nearly-orthogonal signals, then the design should use more embedding dimensions and fewer bits per dimension. In this case the BeSE seems to perform best.

Although we develop theoretical bounds on the performance of phase embeddings, the experimental results suggest that our bounds are loose. It should be possible to develop a multiplicative bound for unquantized phase embeddings, similar in form to the Johnson-Lindenstrauss lemma and the restricted isometry property.

We should also note that one can formulate convex programs, similar to the ones for 1-bit CS,[26] for sparse reconstruction using only the phase of complex projections of a signal, i.e., enable phase-only CS.[10] These phase measurements may also be quantized. It is straightforward to develop consistent reconstruction algorithms for quantized phase measurements. Of course, a similar performance trade-off in number of measurements vs. bits per measurement exists. Exploring this trade-off is something we defer to future work.

REFERENCES

[1] Johnson, W. and Lindenstrauss, J., "Extensions of Lipschitz mappings into a Hilbert space," *Contemporary Mathematics* **26**, 189 –206 (1984).

[2] Candès, E., Romberg, J., and Tao, T., "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Appl. Math.* **59**(8), 1207–1223 (2006).

[3] Baraniuk, R., Davenport, M., DeVore, R., and Wakin, M., "A simple proof of the restricted isometry property for random matrices," *Const. Approx.* **28**(3), 253–263 (2008).

[4] Candès, E., "The restricted isometry property and its implications for compressed sensing," *Comptes rendus de l'Académie des Sciences, Série I* **346**(9-10), 589–592 (2008).

[5] Krahmer, F. and Ward, R., "New and improved Johnson-Lindenstrauss embeddings via the Restricted Isometry Property," *SIAM J. Math. Anal.* **43**(3), 1269–1281 (2011).

[6] Hayes, M., Lim, J., and Oppenheim, A., "Signal reconstruction from phase or magnitude," *IEEE Trans. Acoustics, Speech and Signal Processing* **28**(6), 672 – 680 (1980).

[7] Quatieri, T., J. and Oppenheim, A., "Iterative techniques for minimum phase signal reconstruction from phase or magnitude," *IEEE Trans. Acoustics, Speech and Signal Processing* **29**(6), 1187 – 1193 (1981).

[8] Oppenheim, A. and Lim, J., "The importance of phase in signals," *Proc. IEEE* **69**(5), 529 – 541 (1981).

[9] Oppenheim, A. V., Hayes, M. H., and Lim, J. S., "Iterative procedures for signal reconstruction from fourier transform phase," *Optical Engineering* **21**(1), 122–127 (1982).

[10] Boufounos, P. T., "Sparse signal reconstruction from phase-only measurements," in [*Proc. Int. Conf. Sampling Theory and Applications (SampTA)*], (July 1-5 2013).

[11] Achlioptas, D., "Database-friendly Random Projections: Johnson-lindenstrauss With Binary Coins," *Journal of Computer and System Sciences* **66**, 671–687 (2003).

[12] Dasgupta, S. and Gupta, A., "An elementary proof of a theorem of Johnson and Lindenstrauss," *Random Structures & Algorithms* **22**(1), 60–65 (2003).

[13] Baraniuk, R. G. and Wakin, M. B., "Random projections of smooth manifolds," *Foundations of computational mathematics* **9**(1), 51–77 (2009).

[14] Blumensath, T. and Davies, M. E., "Sampling theorems for signals from the union of finite-dimensional linear subspaces," *IEEE Trans. Info. Theory* **55**(4), 1872–1882 (2009).

[15] Li, M., Rane, S., and Boufounos, P. T., "Quantized embeddings of scale-invariant image features for mobile augmented reality," in [*Proc. IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*], (Sept. 17–19 2012).

[16] Rane, S., Boufounos, P. T., and Vetro, A., "Quantized embeddings: An efficient and universal nearest neighbor method for cloud-based image retrieval," in [*Proc. SPIE Applications of Digital Image Processing XXXVI*], (August 25-19 2013).

[17] Boufounos, P. T. and Rane, S., "Efficient coding of signal distances using universal quantized embeddings," in [*Proc. Data Compression Conference (DCC)*], (March 20-22 2013).

[18] Davenport, M. A., Boufounos, P. T., Wakin, M. B., and Baraniuk, R. G., "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing* **4**, 445–460 (April 2010).

[19] Achlioptas, D., McSherry, F., and Scholkopf, B., "Sampling techniques for kernel methods," in [*Proc. Neural Information Processing Systems (NIPS) 14*], **1**, 335 (2002).

[20] Haupt, J. and Nowak, R., "A generalized restricted isometry property," tech. rep., University of Wisconsin-Madison Tech. Rep. ECE-07-1 (2007).

[21] Jacques, L., Laska, J. N., Boufounos, P. T., and Baraniuk, R. G., "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Trans. Info. Theory* **59** (April 2013).

[22] Plan, Y. and Vershynin, R., "Dimension reduction by random hyperplane tessellations," *Arxiv preprint arXiv:1111.4452* (Nov. 2011).

[23] Vo, A., Oraintara, S., and Nguyen, T. T., "Statistical image modeling using distribution of relative phase in the complex wavelet domain," in [*Proc. 16th European Signal Processing Conference (Eusipco)*], **8** (August 25-29 2008).

[24] Max, J., "Quantizing for minimum distortion," *IRE Trans. Info. Theory* **6**(1), 7–12 (1960).

[25] Lloyd, S., "Least squares quantization in PCM," *IEEE Trans. Info. Theory* **28**, 129 – 137 (Mar. 1982).

[26] Plan, Y. and Vershynin, R., "One-bit compressed sensing by linear programming," *Comm. Pure and Applied Math.* **66**(8), 1275–1297 (2013).