Repairing Multiple Description Quantizers in Distributed Storage Systems

Symeon Chatzinotas

SnT - securityandtrust.lu, University of Luxembourg

Email: Symeon.Chatzinotas@uni.lu,schatzin@ieee.org

Abstract—Distributed storage systems have been receiving increasing attention lately due to the developments in cloud and grid computing. Furthermore, a major part of the stored information comprises of multimedia, whose content can be communicated even with a lossy reconstruction. In this context, Multiple Description Quantizers (MDQ) can be employed to encode such sources for distributed storage. However, a question which naturally arises is how to repair lost descriptions which are due to node failures. In this paper, we employ MDQs based on translated lattices and a common decoding method through averaging over the available descriptions. The descriptions of failed nodes are repaired by quantizing the estimate of common decoding and then by reusing the same side codebook. Based on simulations, we study the effect of system size and number of failures on the distortion of the reconstructed source. As expected, the distortion deteriorates with the number of failures but the degradation is graceful especially for large systems.

Fig. 1. Graphical representation of a distributed storage system which suffers a failure followed by a repair.

I. INTRODUCTION

Distributed storage refers to storing an information source over a number of data units stored in multiple (probably geographically distributed) nodes. The objective is to improve reliability, since the information source can be reconstructed using the data units of a subset of nodes. This area has gained importance due to the recent developments in cloud networks and applications [1], [2].

Starting from this point, we can distinguish two different cases depending on the type of information source: lossless and lossy reconstruction. In the former case, the information source has to be perfectly reconstructed in order to be useful, e.g. documents, software etc. (n, k) Maximum Distance Separable (MDS) codes (e.g. [3], [4]) can be used in order to ensure that any k data units can perfectly reconstruct the source. However, this means that obtaining access to more than k data units provides no improvement in reconstruction. This phenomenon is also known as the "cliff" effect. In the latter case, lossy reconstruction can still convey some useful message, e.g. music, photos, video. This is due to the way humans perceive media files, meaning that the message can still be conveyed from imperfect representations of the source. In this direction, Multiple Description (MD) codes [5], [6] and quantizers [7] can be employed in order to ensure that the reconstruction quality is an increasing function of the available nodes k.

Independently of the coding technique, a common problem in distributed storage systems is how to repair the nodes that are bound to fail from time to time. When a node fails, the information it stored is 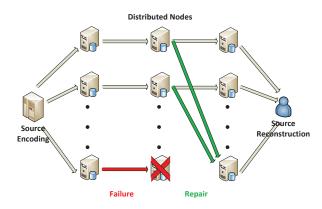lost forever but it may be possible to perfectly of approximately recover it using the redundancy information stored across the distributed storage system. In lossless reconstruction, this problem has been tackled in the literature taking also into account the repair network traffic [8] and exploiting the interference alignment concept inspired by communication networks [9], [10].

A natural question that arises now is how to repair nodes in distributed multimedia storage systems which can afford lossy reconstruction. Fig 1 depicts such a system and follows a set of nodes through a failure and repair. In this paper, we address this problem by proposing a repair technique for the MD scalar quantization proposed in [7]. More specifically, in section II a review of preliminaries and related work is presented. In section III, the system model is described and the repair problem is formally defined. In section III-D, the repair technique and the rationale behind it are presented. Finally, section IV presents numerical results for the evolution of source distortion and section V concludes the paper.

II. RELATED WORK

Multiple Description Coding (MDC) can convert a source into multiple messages, each containing an elementary form (description) of the source. Although an estimation of the source can be derived independently from each description, combining more descriptions yields a monotonically improving estimate. Obviously, this coding approach is advantageous is some setups, such as erasure channels (e.g. packet networks, Internet), where only a random subset of transmitted messages is received correctly. A unique characteristic of MDC is that

the order of the received descriptions is not important in contradiction to layered coding. More specifically, taking into account that all descriptions are not received at the same time, layered coding techniques have to wait for the next codeword according to a predefined sequence in order to be able to refine the reconstruction quality. If this codeword is lost due to an erasure the refinement can no longer proceed. On the other hand, MDC-based techniques can refine quality using any received description. Furthermore, MDC is especially suitable for encoding multimedia files since the human brain is able to convey some information even from highly-distorted source reconstructions which can result from a small number of available descriptions.

A. Multiple Description Codes

Multiple description coding has been studied in the literature both from an information-theoretic point of view, as well as from a constructive point of view. In [5], (n, k) source-channel erasure codes are employed to show that the reconstruction improvement is nearly linear with the number of available descriptions. A rate region is derived using random binning arguments inspired by related multiuser problems, such as distributed source coding. Furthermore, it is shown that such codes can outperform approaches based on source-channel separation. In [6], an achievable rate region is presented based on a concatenation of source-channel erasure codes. In addition, a constructive approach for such codes is presented in [11].

B. Multiple Description Quantizers

A number of constructive solutions have been also investigated in the literature using multiple description quantizers. In [12], dithered delta-sigma quantization is employed taking advantage of its inherent redundancy due to oversampling in order to construct practical symmetric MD scheme. A noise-shaping filter is used to provide a trade-off between central and side distortion. Furthermore, it is shown that this scheme approaches the symmetric two-channel MD rate-distortion function for the Gaussian case asymptotically with lattice dimension and filter order.

In [13], it is shown that the product of a function of the side distortions and the central distortion is asymptotically independent of the redundancy between the descriptions, if the central distortion is much smaller than the side distortion. More importantly, an expression for the side distortions is derived using the normalized second moment of a sphere of higher dimensionality than the quantization space. This high-rate distortion product bound can be used to evaluate the performance of any MD system.

In [14], a single parameter is employed to regulate the trade-off between central and side distortions. Assuming high resolution, the expected distortion is minimized based on entropy constraints for side distortions given packet-loss probabilities. Furthermore, the optimal number of descriptions for minimum expected distortion is found using as input the source, target rate and packet-loss probabilities. A similar study for the asymmetric case is presented in [15].

C. Repairing Distributed Storage Nodes

To the best of the author's knowledge, this problem has not been considered for MD systems, but a fair amount of work has been found in the literature for distributed storage systems that provide lossless reconstruction. A survey of the state of the art can be found in [8], focusing on repair bandwidth and related trade-offs.

In distributed storage systems, nodes can fail at any time. In order to preserve data integrity, the data unit in a failed node has to be reconstructed and stored in a new node. Since this is a common phenomenon, the repair effort has to be minimized. In other words, it should be possible to reconstruct the lost data unit with minimum delay, minimum bandwidth usage and minimum access to other nodes. In this context, exact repair refers to reconstructing the exact lost data unit while functional repair refers to reconstructing a data unit which is not identical to the lost one but it preserves the properties of the distributed storage system.

The problem of minimizing repair bandwidth for lossless reconstruction was addressed in the literature, showing that a single failed node can be repaired by accessing at least $\frac{n-1}{n-k} < k$ data units. This was shown using cut set bounds and it is an improvement over the trivial solution of accessing k data units [16].

Erasure and MDS codes include redundancy so that the information can be retrieved as long as a minimum number of packets k is available. The receiver has to retrieve k packets for lossless reconstruction while additional packets are useless. Source-channel separation applies, since the different packets include independent chunks (data units) of the source.

As mentioned in the introduction, an alternative approach for lossy reconstruction is MD where even a single packet can convey useful information. The reconstruction quality is a monotonically increasing function of the number of retrieved packets/descriptions. Source-channel separation does not apply here, since the different descriptions are correlated because they are representations of the same source. Therefore, a natural question that arises is how to repair failed nodes in MD distributed storage systems.

III. SYSTEM MODEL

Let as assume that a single scalar random variable v has to be quantized and stored in K nodes. If the receiver has access to all K nodes, it should be able to retrieve the exact quantization point $\bar{v}$. If not, he should be able to reconstruct an estimate of the variable $\tilde{v}$. The objective of this model is to study how failures and subsequent repairs affect the reconstruction quality.

A. Encoding

In order to encode v in K data units of rate R (symmetric descriptions), we need a central initial quantizer and then K codebooks that will produce the data units. A quantizer can be

formally defined through its lattice $\mathcal{L}$. Its operation $\bar{v} = \mathcal{Q}(v)$ is to assign its input v to the nearest point $\bar{v}$ in lattice $\mathcal{L}, \bar{v} \in \mathcal{L}$. Subsequently, a labeling function λ is used to assign $\bar{u}$ to a vector $\mathbf{v} = [v_1 \ v_2 \ldots v_k]$

$$\mathbf{v} = \lambda(\bar{v}) \tag{1}$$

where $v_k \in \mathcal{L}_k, \forall k = 1 \ldots K$ and $\mathcal{L}_k$ is the kth sidebook.

B. Decoding

When all K data units are available, the quantized variable $\bar{v}$ can be uniquely identified using the inverse labeling function (index decoding):

$$\bar{v} = \lambda^{-1}(\mathbf{v}). \tag{2}$$

However, when the receiver can only access N out of K data units, the source can be estimated by averaging over the available descriptions [14]:

$$\tilde{v} = \sum_{i \in \mathcal{S}^{(K,N)}} v_i / N. \tag{3}$$

where $\mathcal{S}^{(K,N)}$ denotes a set including K by N index combinations. This is a low-complexity decoding method and its performance depends on the number N of available nodes. For the remainder of this paper, we will refer to this method as common decoding.

C. Codebook & Labeling Design

In order to complete the definition of the codec, two additional elements have to be designed, namely the K codebooks and the labeling function λ. Adopting the approach of translated lattice codebooks [7], the kth codebook is defined as

$$\mathcal{L}_k = \{u_k | u_k = Kwx_k + (2k - K + 1)w/2, x_k \in \mathbb{Z}\} \tag{4}$$

where Kw is the lattice cell width and $\mathbb{Z}$ is the integer set. It can be easily seen that $\mathcal{L}_{k+1}$ can be produced by moving the points of $\mathcal{L}_k$ rightwards by w ("translating"). By accumulating the points of all $\mathcal{L}_k$ codebooks, a denser lattice can be produced and define as:

$$\dot{\mathcal{L}} = \left\{ \dot{u} | \dot{u} = \sum_{k=1}^{K} u_k / K, u_k \in \mathcal{L}_k, k = 1 \ldots K \right\}$$
$$= \{\dot{u} | \dot{u} = w\dot{x}, \dot{x} \in \mathbb{Z}\}. \tag{5}$$

Finally, the lattice of the central quantizer can be defined as:

$$\mathcal{L} = \{\bar{u} | \bar{u} = wx/M + w/2(1/M - 1), x \in \mathbb{Z}\} \tag{6}$$

where $M \in \mathbb{Z}$ represents the number of central points in a Voronoi cell of $\dot{\mathcal{L}}$ and can be used as a tradeoff parameter between central and side distortion. Figure 2 depicts an example of the defined lattices for $M = 2$ and $K = 4$.

Due to the symmetry and regularity of translated lattices, we only have to design the labeling function for the central Voronoi cell and the indices of other cells can be easily derived by appropriate shifting. In order to label the M points of the central Voronoi cell, we first find K-tuples $\mathbf{v}$ which have
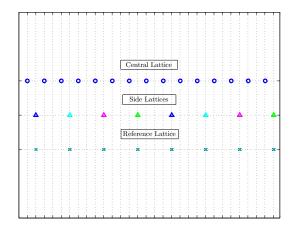

Fig. 2. Structure of central, side and reference lattice. Parameters: $M = 2, K = 4$.

zero centroid, namely $u_c = \sum_{k=1}^{K} v_k = 0$. This constraint guarantees that the squared distance (SD) between the central lattice point and the centroid of $\mathbf{v}$ is minimized. Subsequently, we select M K-tuples that also minimize the squared distances (SSD) between the individual components $u_k, k = 1 \ldots K$ and the centroid of the tuple u_c. As shown in [7], this procedure can be expressed as a geometric minimization problem in a $K - 1$ dimensional space.

D. Repair technique

Now let us assume that a single node fails and the corresponding description u_i is lost. Subsequently, the node may be repaired or a new node may join the cluster. This entails that the lost description has to be reconstructed using a subset of the remaining descriptions $u_{k,k \neq i}$ and stored again based on the same sidebook, so that the same decoding process can be used. Building on the low-complexity decoding method, an estimate for the lost description can be obtained using eq. (3), when the new node has access to a subset of N descriptions.

In order to exploit the same side codebook, the estimate has to be quantized using the lattice of the lost description:

$$v_i = \mathcal{Q}_i(\tilde{v}) \tag{7}$$

Although this is a low complexity repair method, it should perform well when the percentage of available descriptions is high. In the next section, its performance is investigated using numerical simulations.

E. Performance metrics

In order to measure the reconstruction quality, the squared error is used as a distortion measure:

$$\mathrm{D}(v, \tilde{v}) = (v - \tilde{v})^2. \tag{8}$$

In case N out of K symmetric descriptions are available, the distortion is denoted by $\mathrm{D}^{(K,N)}$.

Parameter	Symbol	Value
Number of System Nodes	K	$4-10$
Number of Available Nodes	N	9
Non-consecutive Repairs	r	$1-8$
Description Rate	R	6
Central-side Distortion Tradeoff	M	2
Source Realizations		10^4

TABLE II
LABELING EXAMPLE $M = K = 4$

$\bar{v}(\cdot)$	$v_1(\cdot)$	$v_2(\cdot)$	$v_3(\cdot)$	$v_4(\cdot)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
0	1	1	1	0
1	1	1	0	1
2	1	0	1	1
3	0	1	1	1
4	1	1	1	1
5	2	1	1	0
6	2	1	0	1
7	2	0	1	1
8	2	1	1	1
9	1	2	1	1
10	2	2	1	0
11	2	2	0	1
12	2	2	1	1
13	2	1	2	1
14	1	2	2	1
15	2	2	2	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

IV. NUMERICAL RESULTS

In the simulation setup, it is assumed that a single node fails and is being repaired at each step. Obviously, if the same node fails for a second time, this has no effect on the distortion since the repair method will yield exactly the same lattice point. Therefore, the parameter r denotes the number of non-consecutive node failures that take place in the distributed storage system. The purpose of the simulation is to investigate the reconstruction degradation or equivalently the distortion increase with r. Furthermore, the effect of system size K and node availability N is studied. It should be noted that the source variable v is uniformly distributed across the quantization range and the distortion is calculated by averaging over 10^4 realizations of the source variable. Table I shows an overview of the simulation parameters.

A. Labeling

Before we present the distortion results, some examples of side codebooks and labeling functions are described. Assuming $M = 4$ and $K = 4$, Table II shows the labeling function for the central Voronoi cell of the central quantizer. The first column indicates the index of a lattice point of the central quantizer, while the following four columns indicate the corresponding sidebook indices. As explained before, this pattern is repeated in order to extend the labeling function to other cells.
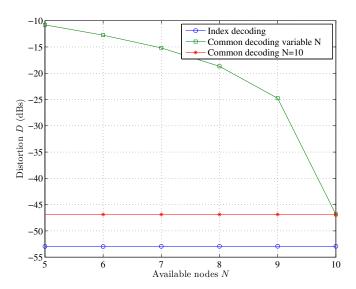

Fig. 3. Distortion $\mathrm{D}^{(K,N)}$ vs available nodes N for $N = 5 \ldots 10$. Parameters: $M = 2, R = 6, K = 10$ bits.

B. Distortion

Figure 3 shows the distortion degradation when the receiver has access to $N \leq K$ nodes. It can be seen that the distortion $\mathrm{D}^{(K,N)}$ of common decoding sharply increases when nodes are unavailable or have failed. Perfect index and common decoding have been also plotted for reference. Let us now focus on the optimistic scenario, where the new node has access to all other nodes $N = K - 1$ for the repair and the receiver has access to all K nodes for decoding. In this case, figure 4 depicts the distortion $\mathrm{D}^{(K,K)}$ versus the system size K for $r = 1 \ldots 4$. It can be seen that using common instead of index decoding yields a 6 dB distortion increase for $M = 2$, which should increase monotonically for higher values of M (higher side distortion). Repairing error yields approximately an additional 5 dB distortion increase, as long as the system size is larger that the number of node failures $K > r$. This distortion gap decreases as K grows large. Similarly, figure 5 shows the distortion $\mathrm{D}^{(K,K)}$ versus the number of repairs r for $K = 4 \ldots 10$, while the point $r = 0$ corresponds to common decoding without repairs. It can be seen that the distortion gracefully degrades with r and the degradation is slower as K increases. Obviously, as r grows large the reconstruction quality may become unacceptable and then all nodes will have to be updated with the original descriptions. This is different from the case of lossless repair (exact or functional) [8], where once the node has been repaired it is as if the failure/erasure never happened.

V. CONCLUSION

In this paper, the problem of repairing MDQs in distributed storage systems was addressed. MDQs are a constructive method of creating multiple descriptions when lossy reconstruction is acceptable, e.g. in multimedia distributed storage systems. The reconstruction quality increases monotonically with the number of available descriptions. This is in contrast
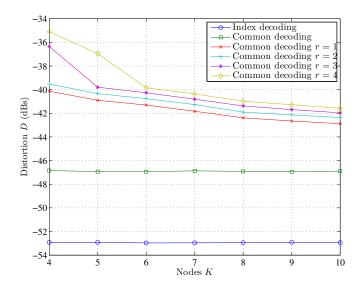
Fig. 4. Distortion $D^{(K,K)}$ vs system size K for $r = 1 \ldots 4$. Parameters: $M = 2, R = 6$ bits.
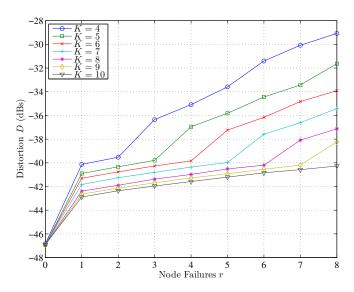

Fig. 5. Distortion $D^{(K,K)}$ vs number of repairs r for $K = 4 \ldots 10$. Parameters: $M = 2, R = 6$ bits.

to distribute storage which requires lossless reconstruction and where the ability to reconstruct is binary based on a predefined threshold of available descriptions. In this direction, a simple method has been proposed for estimating an erased description based on the remaining ones with the purpose of repairing a failed node. More specifically, by assuming translated lattice codebooks for generating the multiple descriptions, an estimate of the failed descriptions can be acquired by averaging over the remaining descriptions. Subsequently, this estimate is quantized based on the codebook of the repaired node. In this context, the distortion degradation was investigated versus number of available nodes, system size and number of failures. It was shown that the degradation is graceful especially for large systems and for a number of non-consecutive failures which is smaller than the system size.

Future work in this area will be focused on studying the described scenarios using upper-bounding analytical tools and imposing constraints on the number of available nodes or the volume of repair traffic. At the same time, improved achievable schemes will be investigated targeting both correct repair and low distortion during source reconstruction.

REFERENCES

[1] W. Zhu, C. Luo, J. Wang, and S. Li, "Multimedia cloud computing," *Signal Processing Magazine, IEEE*, vol. 28, no. 3, pp. 59 –69, may 2011.
[2] A. Marinos and G. Briscoe, "Community cloud computing," in *Cloud Computing*, ser. Lecture Notes in Computer Science, M. Jaatun, G. Zhao, and C. Rong, Eds. Springer Berlin Heidelberg, 2009, vol. 5931, pp. 472–484. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-10665-1_43
[3] I. S. Reed and G. Solomon, "Polynomial Codes Over Certain Finite Fields," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, pp. 300–304, 1960. [Online]. Available: http://dx.doi.org/10.1137/0108018
[4] M. O. Rabin, "Efficient dispersal of information for security, load balancing, and fault tolerance," *J. ACM*, vol. 36, no. 2, pp. 335–348, Apr. 1989. [Online]. Available: http://doi.acm.org/10.1145/62044.62050
[5] S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions - part i: (n, k) source-channel erasure codes," *Information Theory, IEEE Transactions on*, vol. 50, no. 1, pp. 47 – 61, jan. 2004.
[6] R. Puri, S. Pradhan, and K. Ramchandran, "n-channel symmetric multiple descriptions-part ii:an achievable rate-distortion region," *Information Theory, IEEE Transactions on*, vol. 51, no. 4, pp. 1377 – 1392, april 2005.
[7] G. Zhang, J. Klejsa, and W. Kleijn, "Optimal index assignment for multiple description scalar quantization with translated lattice codebooks," *Signal Processing, IEEE Transactions on*, vol. 60, no. 8, pp. 4444 –4451, aug. 2012.
[8] A. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 476 –489, march 2011.
[9] S. Chatzinotas and B. Ottersten, "Cognitive interference alignment between small cells and a macrocell," in *Telecommunications (ICT), 2012 19th International Conference on*, april 2012, pp. 1 –6.
[10] ——, "Interference mitigation techniques for clustered multicell joint decoding systems," *EURASIP Journal on Wireless Communications and Networking, Multicell Cooperation for Next Generation Communication Systems Series*, vol. 132, 2011.
[11] R. Puri, S. Pradhan, and K. Ramchandran, "n-channel multiple descriptions: theory and constructions," in *Data Compression Conference, 2002. Proceedings. DCC 2002*, 2002, pp. 262 – 271.
[12] J. Ostergaard and R. Zamir, "Multiple-description coding by dithered delta-sigma quantization," *Information Theory, IEEE Transactions on*, vol. 55, no. 10, pp. 4661 –4675, oct. 2009.
[13] G. Zhang, J. Ostergaard, J. Klejsa, and W. Kleijn, "High-rate analysis of symmetric l-channel multiple description coding," *Communications, IEEE Transactions on*, vol. 59, no. 7, pp. 1846 –1856, july 2011.
[14] J. Ostergaard, J. Jensen, and R. Heusdens, "n-channel entropy-constrained multiple-description lattice vector quantization," *Information Theory, IEEE Transactions on*, vol. 52, no. 5, pp. 1956 – 1973, may 2006.
[15] J. Ostergaard, R. Heusdens, and J. Jensen, "n -channel asymmetric entropy-constrained multiple-description lattice vector quantization," *Information Theory, IEEE Transactions on*, vol. 56, no. 12, pp. 6354 – 6375, dec. 2010.
[16] A. Dimakis, P. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4539 –4551, sept. 2010.