It's Easier to Approximate

Plenary talk presented at the 2009 IEEE International Symposium on Information Theory, Seoul, South Korea

David Tse

Abstract

Shannon provided an exact characterization of the fundamental limits of point-to-point communication. After almost 40 years of effort, meeting the same goal for networks proved to be far more difficult. In this talk, we argue that much broader progress can be made in network information theory when instead one seeks *approximate* solutions with a guarantee on the gap to optimality. We discuss a specific approach focusing on the practically important models of linear Gaussian channels and Gaussian sources.

I. Introduction

In his seminal paper [1], Shannon provided a complete solution to the fundamental limits of point-to-point communication. Since the coding schemes allowed are of arbitrary block lengths, the original design problem is an infinite-dimensional optimization problem. Yet, the optimal solution can be expressed as that of a finite-dimensional optimization problem ("single-letter" characterization). Moreover, for many specific channels and sources, this finite-dimensional optimization problem can be solved explicitly in closed form. This desirable state of affairs is remarkable and almost unique among engineering fields, but it also sets a high standard for the information theory field.

A holy grail of information theory is to extend Shannon's point-to-point result to the network setting. The general network information theory problem is to analyze the fundamental limits of communication when multiple senders want to communicate with multiple receivers with the help of intermediate nodes. The first success came in the earlier 1970's, when Ahlswede [2] and Liao [3] independently provided a single-letter characterization of the capacity region of the multiple access channel. In this network, K users want to send information to a common receiver across a noisy channel. This result is rather general in the sense that it holds for arbitrary number of users as well as arbitrary channel statistics. It led to much excitement in the field at that time. However, as it turned out, there have been essentially no other network information theory results of such generality since then. Most of the other results, for example, hold for only two users (such as the degraded message set problem for broadcast channels) or for specific class of channel or source statistics (such as degraded broadcast channels). Even these results are few in number. So despite almost forty years of effort, it is fair to say that we are still very far from solving the general network information theory problem.

A class of channels and a class of sources that have received much attention are linear Gaussian channels with quadratic cost constraint and Gaussian sources with quadratic distortion measure respectively. Not only are these models practically relevant for applications such as wireless and sensor networks, the physically meaningfulness of their structures give some hope that Gaussian problems are easier to solve than the general case. Indeed, as is well-known, the capacity of the point-to-point Gaussian channel and the rate-distortion function of the Gaussian source are known in closed form. Can this luck help us make more progress in Gaussian network problems than in the general case? The answer is yes for broadcast channels. While the capacity region of the general broadcast channel is open even in the case of two users, the capacity region of Gaussian broadcast channels with arbitrary number of users is known. However, it seems that the luck ran out rather quickly as most Gaussian network problems are still open. Examples are interference channels (even the two-user case is open), relay networks (even the single-relay channel is open), multiple description and distributed lossy source coding (both open for more than 2 users). So it seems that Gaussian network problems are not too much easier than the general ones.

In this talk, we outline a recent approach to make progress in Gaussian network information theory problems. The idea is to *approximate*. Rather than asking for *exactly optimal* solutions for network problems, we recognize that network problems are far more difficult than point-to-point problems and are willing to settle for *approximate* solutions. Not any old approximate solutions however, but approximate solutions with a hard guarantee on the gap to optimality.

Approximate solutions to information theory problems are not new. However, they are by and far isolated results each with its own proof technique. What distinguishes the approach we advocate here with these results is that it is a systematic approach that can be applied to *many* problems.

The approach consists of four steps:

- *Noisy* channel coding problems are approximated by *noiseless* problems. Lossy source coding problems are approximated by *lossless* problems.
- Analyze the simplified problem.
- Use insights to find new schemes and/or outer bounds to the original Gaussian problem.
- Derive a worst-case gap of the performance of the proposed scheme to optimality, universal for all values of the channel parameters.

What is the rationale for this approach? Take channel coding problems for example. In the point-to-point case, the noise is the central object of interest and it occupies the sole attention of Shannon's point-to-point theory. In networks, however, in addition to the noise there is also the interaction between the signals of multiple users. To try to solve the problem in one shot is fighting two battles at the same time. Approximating the noisy problem by a noiseless (deterministic) one allows us to first focus on the signal interaction. Noiseless problems are often easier than noisy problems. For example, while the general noisy broadcast channel problem is open, the deterministic broadcast channel is solved (independently by Pinsker [4] and Marton [5]). Similarly, lossless source coding problems are often easier than lossy ones. For example, while the general lossy distributed source coding problem is open (even for two users), the lossless distributed source coding problem is solved (the celebrated Slepian-Wolf Theorem [6]).

Because this approach in effect decouples the effect of the noise and the signal interaction, it does not in general yield exactly optimal solution. (Although sometimes one can get lucky, as we will see.) The approximation becomes relatively more accurate when the noise is small compared to the signals (interference-limited or low-noise regime). So while the worst-case gap holds for all parameter ranges, the performance gap is more meaningful in the low-noise regime where the achievable rates are high. The dual statement for source coding is that the approximation using this approach becomes relatively more accurate when the target distortion levels are small and the required rates are high.

In the rest of the talk, we will illustrate this approach using the four open problems mentioned above.

II. Interference Channels

A. Strong Interference

The capacity region of the 2-user Gaussian interference channel (IC) (Fig. 1(a)) is one of the long-standing open problems in network information theory. Two users interfere with each other through cross talk. The problem is to determine the set of all rates (R_1, R_2) that are simultaneously achievable by the two users. This channel was first considered in the 1970's and the capacity region of the Gaussian IC in the *strong interference* regime was quickly figured out (independently by Sato [7] and Han and Kobayashi [8]). In this regime, transmitter 1(2) has a better channel to receiver 2(1) than to receiver 1(2). In any working system for this channel, receiver 2 can decode its own message m_2, and therefore can cancel off m_2's contribution. Now, receiver 2 has a clear view of transmitter 1's signal, and since receiver 2 has a better channel than receiver 1 from transmitter 1 and receiver 1 can decode its own message m_1, then receiver 2 can decode the message m_1 as well. Similarly, receiver 1 can decode the message m_2. So although the communication system is designed only to deliver the message m_1 to receiver 1 and the message m_2 to receiver 2, these messages are automatically *public*, i.e. decodable at the other receiver. This converts the strong interference channel to a *compound* multiple access channel, i.e. both messages have to be decodable at each of the receivers, and the capacity region of the Gaussian IC is simply the intersection of the capacity regions of the two multiple access channels, one at each receiver.

B. El-Gamal-Costa Deterministic IC

The strong interference regime was quickly solved, but very little progress has been made on the other parameter regimes for many years since then. When the channel to the other receiver is weaker than to your own receiver, requiring the other receiver to decode your message is obviously sub-optimal. But what is the right strategy? In fact, the only non-trivial IC whose capacity region is fully solved is the *deterministic* IC studied by El Gamal and Costa [9]. This is shown in Fig. 1(b). The channel output Y_1 is a function of the input X_1 from transmitter 1 and V_2, which in turn is a function of the input X_2 from transmitter 2. This would just have been a general deterministic IC but for an important property they assumed: that V_2 is a function of X_1 and Y_1 (and similar for V_1). What is the optimal strategy for this channel? In any working system, receiver 2 can decode its own message. Therefore, receiver 2 knows X_2. From X_2 and Y_2, it has a clear view of V_1. So the part of the message from transmitter 1 that is on V_1 will also be decodable by receiver 2, i.e. is public. This argument is similar to that used in the strong interference regime, except that now only a part of the message, the part on V_1, is public. The rest is private. This strategy is a special case of the Han-Kobayashi achievable scheme [8] with a specific prescription on how to do the private-public split.

C. Connection with Gaussian IC

The El-Gamal-Costa channel seems to have nothing to do with the Gaussian IC, but in 2006, Raul Etkin, Hua Wang and myself observed a connection. The key is to *approximate* how the Gaussian IC behaves. Consider an example of a Gaussian IC where $h_{11} = 2^n$, $h_{12} = 2^m$, and $m < n$ so that we are not in the strong interference regime. Suppose X_1 has a binary expansion $0 \cdot b_1 b_2 b_3 \ldots$. The resulting signal at receiver 1, before adding noise and interference, is $b_1 b_2 \ldots b_n . b_{n+1} \ldots$, and the corresponding signal at receiver 2 is $b_1 b_2 \ldots b_m . b_{m+1} \ldots$. The noises at both receivers are normalized to have unit variance, so the decimal point in the above expansions is the "noise level". One can divide the transmitted bits $b_1, b_2, b_3 \ldots$ into three groups:

- $b_1, b_2, \ldots b_m$, which appear above the noise level at both receivers.
- $b_{m+1}, b_{m+2}, \ldots b_n$, which appear above the noise level at receiver 1 but below noise level at receiver 2.

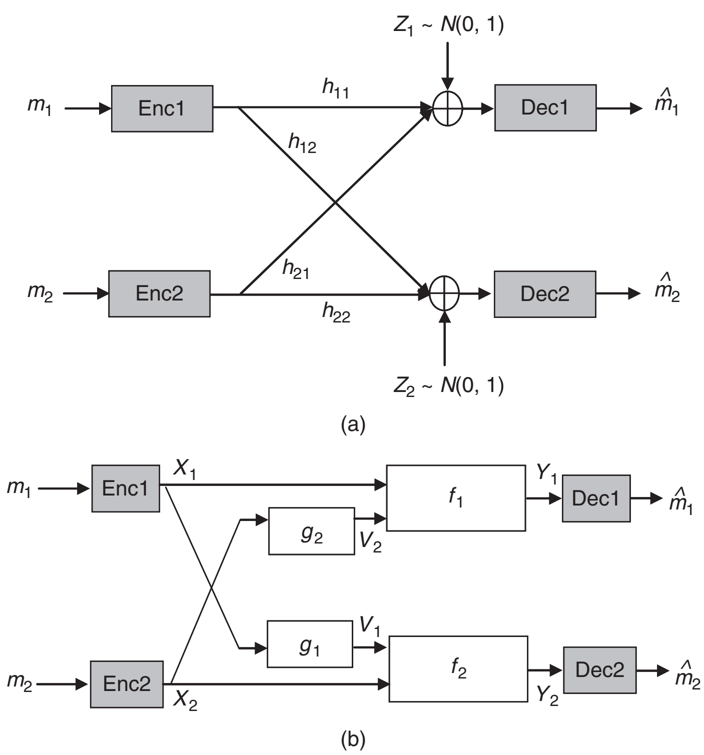

Fig. 1 (a) Gaussian IC; (b) Deterministic IC.

- $b_{n+1}, b_{n+2}, \ldots$ appear below the noise level at both receivers.

This decomposition suggests a way to approximate the Gaussian IC by a El-Gamal-Costa IC. Since $b_{n+1}, b_{n+2}, \ldots$ are below the noise level at both receivers, they convey little information but also have little interfering effect, being masked by the noise. Let's ignore them entirely and assume the transmit signal X_1 is just $(b_1, b_2, \ldots b_n)$. Of these bits, $b_1, \ldots, b_m$ are observed at both receivers while the rest appear below noise level at receiver 2 and so also have little interfering effect. So we can let $V_1 = (b_1, \ldots, b_m)$. The key property of the El-Gamal-Costa IC is approximately satisfied: given the input X_2 and the output Y_2, the interference V_1 above the noise level can be (approximately) determined. The El-Gamal-Costa result then tells us that user 1 should split its message into a public and a private message, the public message conveyed in $b_1, \ldots, b_m$ while the private message conveyed in $b_{m+1}, \ldots, b_n$.

D. Gaussian Capacity to within 1 Bit

Once this correspondence is established, it is clear what is the natural scheme to try on the Gaussian IC. Split each transmitter's message into a public message and a private message. Allocate power to the private message such that it is received just below the noise level at the other receiver. The rest of the power is allocated to the public message. Use independently generated Gaussian codebooks to convey the public and the private messages. In [10], it was shown that this strategy can achieve to within 1 bit/s/Hz (i.e. 0.5 bit per real dimension) of the capacity region. This gap holds for all values of the channel parameters. To show this result, new outer bounds are obtained for the Gaussian IC to match (approximately) the performance of the proposed scheme. Like the scheme, the outer bounds were also inspired by the corresponding outer bounds of the El-Gamal-Costa IC.

The correspondence between the Gaussian IC and the deterministic IC described above is approximate but not exact. In the deterministic IC, bits are either perfectly observed or are completely invisible. In the noisy Gaussian IC, such is not the case. This accounts for why there is a gap between the performance of the proposed scheme and the outer bound. Somewhat surprisingly, subsequently works [11], [12], [13] showed that by further tightening one of the new outer bounds in [10], an *exact* characterization of the sum rate of the Gaussian IC can be obtained in a certain very weak interference regime.

E. Lattice Codes for Interference Alignment

The within-1-bit strategy is a special case of the general Han-Kobayashi scheme with randomly generated Gaussian codebooks for both the private and the public messages (and a specific power split). Since Han-Kobayashi allows arbitrary input distributions for the private and the public messages, what we showed is that Gaussian input distribution is "nearly" optimal for the 2-user IC. This is consistent with the folklore in information theory that "Gaussian inputs are good for Gaussian problems". But does this continue to hold true for IC with more users?

Consider an example of a many-to-one Gaussian IC in Figure 2. Here, there are three users and the top user is interfered by the other two. We consider a particular operating point, and show in the figure the binary expansion of the signals in the deterministic approximation of this Gaussian IC. Both transmitter 2 and transmitter 3 are sending two bits above noise level at their respectively receivers. In this example, the channels from both these transmitters to receiver 1 are stronger than their own direct channels and so the bits from these transmitters are shifted upwards at receiver 1 relative to receiver 2 and 3. We make two observations:

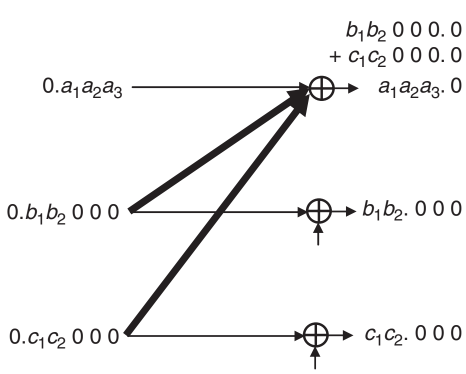

Fig. 2 Many-to-one IC.

- The two most significant bits at receiver 1 are un-usable for transmitter 1 as long as one of the other two transmitters send information at those levels. So if one is sending, the other might as well send as well. This is the phenomenon of *interference alignment.*
- The next three significant bits are left empty by both users 2 and 3. So now user 1 can send 3 bits on those levels.

How can we translate this picture back to the Gaussian world? A natural strategy would be for both users 2 and 3 to use a capacity-achieving Gaussian code on their own link. Because of the strong channel to receiver 1, the codewords in each of the Gaussian codes will be spaced far apart there. However, the *summed* codewords will be close together. This is because the summed codewords will be all distinct and so the size of the summed codebook is the square of the size of each user's codebook. This means that while the *individual* interference is confined within the most significant two bits, the *aggregate* interference leaks to the next three bits, making these levels unusable for user 1. But if instead we use the same *lattice* code for both users 2 and user 3, then interference alignment can be achieved. This is because the summed codewords will remain on the lattice. Now the space in between the codewords is preserved for user 1 to transmit information. Thus, unlike in the two-user case, Gaussian codes are no longer good when there are more users. Generalizing from this example, it is shown in [14] that lattice codes can achieve the capacity of the many-to-one Gaussian to within constant gap universal of the values of the channel gains.

III. Relay Networks

Consider a relay network with a single sender node who wants to transmit information to a single destination node with the help of a number of relay nodes in between. The received signal at a node is a superposition of the (attenuated) signals transmitted at other nodes plus Gaussian noise. What is the capacity, the maximum rate of information transfer from the sender to the destination?

This problem has a very long history, but even the simplest case with a single relay node (the so-called relay channel) is open. The best known achievable strategies were obtained by Cover and El Gamal in 1979 [15].

If instead of Gaussian channels, the nodes are connected via noiseless, orthogonal links, then we have a wireline network and the capacity is given by the famous max-flow min-cut theorem of Ford and Fulkerson. The Gaussian problem is significantly more complex due to the superposition of the signals as well as the additive noise at each of the node. A natural generalization of the min-cut of wireline networks to general networks is the cutset bound:

$$C_{\text{cutset}} = \max_{p_{X_1,\dots,X_n}} \min_{\Omega} I(X_\Omega; Y_{\Omega^c} \mid X_{\Omega^c}) \quad (1)$$

where Ω is a cut in the network (a set of nodes including the sender) and the maximization is over all joint distributions on the transmit signals at the nodes. $I(X_\Omega; Y_{\Omega^c} \mid X_{\Omega^c})$ is the information flow across the cut assuming full cooperation of the nodes in Ω to send information and full cooperation of the nodes outside of Ω to decode the information. In the wireline network, the cutset bound evaluates to the minimum cut of the network and yields the capacity. In the Gaussian network, however, the cutset bound only provides an upper bound. How tight is this bound?

In the Fall of 2006, Salman Avestimehr, Suhas Diggavi and myself started to study this question. Fresh after the 1-bit gap result on the 2-user Gaussian IC, we naturally seek a constant-gap result for relay networks as well. Our first observation is that for the (single) relay channel, the decode-forward strategy proposed by Cover and El Gamal in 1979 actually achieves within 1 bit/s/Hz of the cutset bound, universally for all values of the channel gains. But what about for general networks with more than 1 relay? It is clear that requiring each relay to decode the entirety of the sender's information is not the right thing to do in general. So what to forward? How to forward?

In the work on the two-user Gaussian IC [10], we found a good scheme and new outer bounds by drawing an approximate analogy with the El Gamal-Costa deterministic IC. However, this analogy was of a heuristic nature and in that work we actually never introduced a specific deterministic channel to approximate the Gaussian IC. In the relay work, we took this approach one step further and introduced a specific deterministic channel model as a bridge between the Gaussian and the wireline models. This allows us to leverage off insights from the wireline network to solve the Gaussian relay problem.

One insight from our earlier discussion on interference channels is that bits received above noise level can be approximated as clean and bits below the noise level as useless. This insight can be converted into a deterministic channel model as follows. In the Gaussian model, the received signal at a relay node j is:

$$Y_j^G = \sum_i h_{ij} X_i + W_j, \qquad W_j \sim N(0,1),$$

where X_i is the signal sent at node i and h_{ij} is the gain from node i to node j. Since noise is normalized to be unit variance, the integer part of the received signal can be considered as the part above noise level. This yields the following deterministic channel:

$$Y_j^D = \left[\sum_i h_{ij} X_i \right].$$

The next step is the analysis of this deterministic network. In earlier works, Aref [16] and Ratner and Kramer [17], had looked at deterministic networks but with broadcast only and no superposition of the signals at the nodes. They showed that a random forwarding strategy at each relay (randomly mapping the received signal to a transmit codeword) is sufficient to achieve the cutset bound. This strategy is reminiscent of the random network coding strategy of Ahlswede et al [18] for wireline networks and in fact deterministic networks with broadcast only is a generalization of the wireline model. But in our deterministic channel model, there is superposition of signals as well. This led us to prove the following generalization of these results to *general* deterministic relay networks: the rate

$$R = \max_{p_{X_1} \cdot p_{X_2} \cdot, \dots, p_{X_n}} \min_{\Omega} I(X_\Omega; Y_{\Omega^c} \mid X_{\Omega^c}) \quad (2)$$

is achievable [19], [20]. We see that this is identical to the cutset bound (1) *except* that in the maximization, the input distribution is constrained to be independent across the nodes. For wireline and deterministic networks with broadcast only, it is optimal to have independent inputs at each node and (2) matches the cutset bound. For the deterministic network derived from the Gaussian network, we showed that this achievable rate is a constant gap from the cutset bound, irrespective of channel gain parameters. So not allowing the correlation of inputs results only in bounded loss.

Finally, we brought back these insights to the original Gaussian relay network. The answers to our earlier questions is now clear: 1) What to forward at each relay? The received signal quantized at the noise level; 2) How to forward? Each relay randomly maps the quantized received signal into a Gaussian codeword to transmit. In [20] we showed that this strategy achieves within a constant gap κ from the cutset bound for the Gaussian network.

Among all the schemes proposed in [15], our scheme is philosophically most similar to compress-forward. There is one important difference, however. In the compress-forward scheme discussed in [15], the destination is required to decode the quantized signal at the relay and then, with the help of the quantized signal and the direct reception from the sender, decodes the sender's message. In our scheme, the quantized signals are never decoded anywhere. at the relays or the final destination. Instead, the sender's message is decoded directly at the final destination based on all the forwarded information. These two approaches yield identical performance on single relay networks, but the latter approach is superior for more than 1 relay nodes. In fact, it is not even clear how the first approach can be naturally generalized to more than 1 relay. In recent work, Lim et al [21] generalized our scheme from the Gaussian case to general noisy networks and coined it "network compress-forward".

Finally, a comment about the gap κ to the cutset bound. This gap does not depend on the values of the channel gains, but unfortunately it depends on the number of nodes n in the networks. It grows like $n \log n$, and so our result is not very good when the network is large. Basically, each quantized signal at a relay contains noise, and with increasing number of relay stages, noise gets accumulated more and more and the performance of the scheme degrades. An interesting open question is to either find another scheme that has a network-size-independent gap to the cutset bound, or find a better upper bound than the cutset bound.

IV. Source Coding

We approximate the Gaussian channel by a deterministic channel by assuming the bits of the received signal above the noise level are completely clean and bits below are completely useless. In a dual way, we can approximate a Gaussian lossy source coding problem with quadratic distortion measure by viewing a source sample X in terms of its binary expansion $0.b_1b_2\dots$, and the goal of the source decoder as recovering the first n most significant bits, where $n = \frac{1}{2}\log_2 d$ and d is the distortion requirement. Hence, the source encoder only has to focus on the first n bits and the lossy problem of recover X to distortion d is replaced by a *lossless* problem of recovering $(b_1, b_2, \ldots, b_n)$ exactly. This approximation is applied to two source coding problems below.

A. Multiple Description (MD)

A source has to be described using K descriptions, such that the decoder that receives a subset S of the descriptions can recover the source to within distortion d_S. Given distortion requirement for every subset $S \subset \{1, \ldots, K\}$, what are the set of rates $(R_1, \ldots, R_K)$ needed to generate the descriptions? Let us focus on Gaussian sources with squared error distortion. In the case of 2 descriptions, Ozarow [22] showed that an achievable rate region by El Gamal and Cover for general sources [23] is tight. The problem for more than 2 descriptions is open.

For simplicity, let us focus on the *symmetric* MD problem, where the same distortion d_m is required for *any* subset of m descriptions, and $d_1 < d_2 \ldots < d_K$. The approximating lossless source coding problem is as follows [26]. Let $n_i = \frac{1}{2}\log d_i$, $i = 1, \ldots K$. The source X is $(b_1, b_2, \ldots b_{n_K})$, and any decoder that receives i descriptions have to recover $b_1, b_2, \ldots b_{n_i}$. Note that the bits that needed to be recovered at different "levels" are nested. This lossless source coding problem had been considered before: it is called "multilevel diversity coding" [24], [25]. The optimal coding strategy breaks up the source into $V_1, V_2, \ldots V_K$, where $V_i = (b_{n_{i-1}}, \ldots b_{n_i})$ are the additional bits in level i beyond those in level $i-1$, codes V_i using a (K, i) MDS code, and constructs the descriptions as shown in Figure 3. This ensures that whenever one receives i descriptions, $V_1, \ldots V_i$ can be recovered.

The V_i's can be thought of as successive refinement layers of the source: V_1 is the base layer (most significant bits), V_1 are additional refinement bits, and so forth. Thus, the above lossless approximation suggests a natural strategy for the original Gaussian MD problem: use a successive refinement code to generate layers $V_1, V_2, \ldots V_K$, such that with $V_1, \ldots V_i$ the source can be reconstructed with distortion d_i, and then apply multilevel diversity coding to generate the descriptions as above. Using the successive refinability of Gaussian sources, it is shown in [26] that this strategy achieves within 1.48 bits/sample of the symmetric rate point for any number of descriptions. A more sophisticated scheme by Puri et al [27] has a gap of 0.92 bits/sample.

B. Distributed Lossy Source Coding

K sources $Y_1, \ldots Y_1$ are distributedly encoded at rates $R_1, \ldots R_K$ respectively. Using the encodings, a central decoder has to reconstruct these sources with distortions $d_1, \ldots d_K$ respectively. What is the achievable rate region? In the case when the sources are correlated Gaussian and the distortion measure is quadratic, this problem for 2 sources was recently solved by Wagner et al [28], building on earlier work by Oohama [29]. The optimal strategy is Gaussian quantization of the sources followed by Slepian-Wolf binning. The problem is wide open for three or more sources, but progress can be made using the approximation approach.

Consider an example of 3 tree sources Y_1, Y_2, Y_3, i.e. there exists a Gaussian $X \sim N(0,1)$, such that $Y_i = X + Z_i$, $i = 1, 2, 3$ with $Z_i \sim N(0, \sigma_i^2)$ and X, Z_1, Z_2, Z_3 are independent. Approximately, we can think of the Z_i's as "noises" which make the less significant bits of the Y_i's independent while keep the more significant bits identical. For example:

$$X = 0.a_1a_2a_3a_4 \ldots,$$

$$Y_1 = 0.a_1a_2b_1b_2 \ldots,$$

$$Y_2 = 0.a_1a_2a_3c_1c_2 \ldots,$$

$$Y_3 = 0.a_1a_2a_3a_4d_1d_2 \ldots,$$

for the case when $\sigma_1^2 > \sigma_2^2 > \sigma_3^2$.

In the approximating lossless problem, each encoder has to deliver the significants bits of its Y_i up to the target distortion level. But because there is correlation (like the a_1 bit that appears in all of the Y_i's in the above example), rate can be saved by only sending one copy of each independent bit. In the lossless problem, this can be pre-arranged by making sure each independent bit is delivered only by one encoder. Alternatively, all the encoders can do random binning into bins of appropriate size to remove the redundancy in the encodings.

This latter strategy naturally yields a strategy for the original lossy problem. First, each encoder does Gaussian quantization up to the distortion requirement of its observation Y_i. This in effect extract the significant bits that the decoder needs. Then, the index of the quantized vector is randomly binned. It is shown in [30] that this strategy is within 2.4 bit/sample of the optimal rate region.

The strategy above is exactly the same as the Gaussian-quantize-and-bin strategy that is optimal for the 2-source case. So what was shown is that this strategy is within a constant gap to optimality for *tree* sources. Is this strategy good for *any* jointly Gaussian sources?

Consider the follow example. Y_1, Y_2 are correlated and $Y_3 = Y_1 - Y_2$, and our goal is to recover Y_3 at a certain distortion d_3 with encodings from Y_1 and Y_2 only. We can write out the binary expansions:

$$Y_1 = 0.a_1a_2a_3b_1b_2b_3 \ldots$$

$$Y_2 = 0.a_1a_2a_3c_1c_2c_3 \ldots$$

$$Y_1 - Y_2 = 0.000e_1e_2e_3 \ldots$$

where $e_i = a_i - b_i$, $i = 1, 2, 3$. Suppose we want to recover Y_3 up to the 5th significant bit. The Gaussian-quantize-and-bin strategy will first yield the first 5 bits from each of the sources via Gaussian

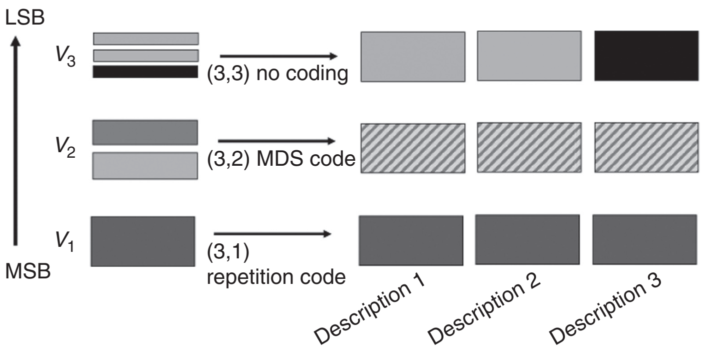

Fig. 3 Multi-level coding.

quantization, and use Slepian-Wolf binning to remove the redundancy in the encodings. So in effect, only one copy of a_1, a_2, a_3 are sent.

But this is still wasteful! The decoder does not actually need any copy of a_1, a_2, a_3; it only needs b_1, b_2, c_1, c_2 to compute the difference bits e_1, e_2. So what is needed is a quantizer for Y_1 to extract only the less significant bits b_1, b_2 and a quantizer for Y_2 to extract c_1, c_2. A (random) Gaussian quantizer will not do; the five significant bits are all mixed up in the representation. Rather, what is needed is a lattice quantizer, consisting of a coarse lattice representing the most significant bits (a_1, a_2, a_3) and a fine lattice representing the less significant bits (b_1, b_2 for Y_1, and c_1, c_2 for Y_2). Each encoder only needs to send the fine lattice index of the quantized vector. This scheme was proposed by Krithivasan and Pradhan [31] and shown to be within 1 bit/sample to optimality by Wagner [32].

V. Conclusion

Traditionally, *exact* analysis of Gaussian network information theory proceeds by finding a good Gaussian scheme and then proving a converse using an extremal information inequality for which Gaussian is tight. This approach is problematic because: 1) we don't have too many such inequalities in our arsenal (basically entropy-power inequality and its variants) and inventing new ones is difficult; 2) Gaussian schemes may be very far away from being optimal (as we saw); 3) the analysis is very much tied to the details of the Gaussian noise/source model. The approximation approach tries to circumvent these difficulties. Moreover, it has the added bonus of connecting Gaussian problems with other problems such as network coding and lossless source coding, and thus helps to shed more light into the structure of the network information theory field as a whole.

References

[1] C. E. Shannon, "A Mathematical Theory of Communication", BSTJ, July, Oct. 1948.

[2] R. Ahlswede, "Multi-way communication channels", IEEE International Symposium on Information Theory, Tsahkadsor USSR, pp. 103–135, 1971.

[3] H. Liao, "A coding theorem for multiple access communications", International Symposium on Information Theory, Asilomar CA, 1972.

[4] M. Pinsker, "Capacity of Noiseless Deterministic Channels", Prob. Perdachi Inform., 1978.

[5] K. Marton, "The capacity region of deterministic broadcast channels", Proc. ISIT, 1979.

[6] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources",IEEE Trans. IT, July 1973.

[7] H. Sato, "The capacity of the Gaussian interference channel under strong interference", IEEE Trans. IT, Nov. 1981.

[8] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," IEEE Trans. IT, Jan. 1981.

[9] A. El Gamal and M. Costa, "The Capacity Region of a Class of Deterministic Interference Channels," IEEE Trans. IT, March 1982.

[10] R. Etkin, D. Tse, and H. Wang, "Gaussian Interference Channel Capacity to within One Bit," IEEE Trans. IT, Dec. 2008.

[11] V. Annaprueddy and V. Veeravalli, "Gaussian interference networks: sum capacity in the low-interference regime and new outer bounds on the capacity region", IEEE Trans. IT, July 2009.

[12] A. S. Motahari and A. K. Khandani, "Capacity Bounds for the Gaussian Interference Channel", IEEE Trans. IT, Feb. 2009.

[13] Shang, X., G. Kramer, and Chen B., "A New Outer Bound and the noisy interference sum rate capacity for the Gaussian interfrence channel", IEEE Trans. IT, Feb. 2009.

[14] G. Bresler, A. Parekh, and D. Tse, "The Approximate Capacity of an One-Sided Gaussian Interference Channel", Proc. Allerton Conf, Sep. 2007.

[15] T. Cover and A. El Gamal, "Capacity Theorems for the Relay Channel," IEEE Trans. IT, Sept. 1979.

[16] M. R. Aref, "Information Flow in Relay Networks", Ph.D. dissertation, Stanford Univ., Stanford, CA, 1980.

[17] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference", IEEE Trans. IT, June 2006.

[18] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow", IEEE Trans. IT, July, 2000.

[19] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless Network Information Flow," in Allerton Conference on Communication, Control, and Computing, (Monticello, IL), Sept. 2007.

[20] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless Network Information Flow: A Deterministic Approach", submitted to IEEE Trans. IT, arXiv:0906.5394v2

[21] S. H. Lim, Y.H. Kim, A. El Gamal, S.Y. Chung, "Noisy Network Coding", Proc. of ITW, Cairo, Jan. 2010.

[22] L. Ozarow, "On a source-coding problem with two channels and three receivers" , BSTJ, Dec. 1980.

[23] A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions", IEEE Trans. IT, Nov. 1982.

[24] J. R. Roche, R. W. Yeung, and K. P. Hau, "Symmetrical multilevel diversity coding", IEEE Trans. IT, May 1997.

[25] R. W. Yeung and Z. Zhang, "On symmetrical multilevel diversity coding" IEEE Trans. IT, Mar. 1999.

[26] C. Tian, S. Mohajer, and S. N. Diggavi, "Approximating the Gaussian multiple description rate region under symmetric distortion constraints," IEEE Trans. IT, Aug. 2009.

[27] R. Puri, S. S. Pradhan, and K. Ramchandran, "n-channel symmetric multiple descriptions – part II: an achievable rate-distortion region", IEEE Trans. IT, vol. 51, Apr. 2005.

[28] A. B. Wagner, S. Tavildar and P. Viswanath, Rate Region of the Quadratic Gaussian Two-Terminal Source-Coding Problem, IEEE Trans. IT, May 2008.

[29] Y. Oohama, "Gaussian multiterminal source coding", IEEE Trans. IT, Nov. 1997.

[30] M. Maddah-Ali, D. Tse Approximating the Rate-Distortion Region of the Distributed Source Coding for Three Jointly Gaussian Tree-Structured Sources, Proc. of ISIT, July 2009.

[31] D. Krithivasan and S. S. Pradhan, "Lattices for distributed source coding: Jointly Gaussian sources and reconstruction of a linear function", Proc. AAECC-17, December 2007.

[32] A. B. Wagner, "An Outer Bound for Distributed Compression of Linear Functions," Proc. CISS, Mar. 2008.