Rematch-and-Forward: Joint Source–Channel Coding for Parallel Relaying with Spectral Mismatch†

Yuval Kochman*, Anatoly Khina, Uri Erez§ and Ram Zamir*

Dept. of EE-Systems, Tel Aviv University, Israel

{yuvalko,anatolyk,uri,zamir}@eng.tau.ac.il

Abstract

The Gaussian parallel relay network, introduced by Schein and Gallager, consists of a concatenation of a Gaussian additive broadcast channel from a single encoder to a layer of relays followed by a Gaussian multiple-access channel from the relays to the final destination (decoder), where all noises are independent. This setup exhibits an inherent conflict between digital and analog relaying; while analog relaying (known as "amplify-and-forward", A&F) suffers from noise accumulation, digital relaying (known as "decode-and-forward") looses the potential "coherence gain" in combining the relay noises at the decoder. For a large number of relays the coherence gain is large, thus analog relaying has better performance; however it is limited to white channels of equal bandwidth. In this work, we present a generalization of the analog approach to the bandwidth mismatch case. Our strategy, coined "Rematch and Forward" (R&F), is based upon applying joint source–channel coding techniques that belong to a certain "maximally analog" class. Using such techniques, R&F converts the bandwidth of the broadcast section to that of the multiple-access section, creating an equivalent matched-bandwidth network over which A&F is applied. It is shown that this strategy exploits the full bandwidth of the individual channels, without sacrificing the coherence gain offered by A&F. Specifically, for given individual-link capacities, R&F remains within a constant gap from the network capacity for any number of relays and any bandwidth ratio between the sections. Finally, the approach is extended to the case of colored channels.

† Parts of this work were presented at ISIT2008, Toronto, ON, July 2008 and IEEEI2008, Eilat, Israel, November 2008.

* This work was supported by the ISF under grant # 1259/07.

§ This work was supported by the ISF under grant # 1234/07.

Index Terms

Parallel relay network, network capacity, amplify-and-forward, joint source–channel coding, hybrid digital–analog, modulo-lattice modulation, broadcast channel, multiple access channel, coherence gain, bandwidth mismatch, ISI channel.

I. Introduction

The emerging field of communication over networks witnesses the collapse of the traditional distinction between channel, source and joint source–channel problems. A special class of problems has a relay structure, in which a message source node wishes to pass information to a destination node, while other nodes act as relays, whose sole purpose is to help in this data transfer. Although the end-to-end problem is a channel coding one, the techniques used to solve it are diverse. Consider the best known relaying techniques (see e.g. [1]), where each one is known to be better under different conditions of network topology and signal-to-noise ratios:

1. A channel coding approach: *Decode-and-Forward* (D&F), where a relay decodes the message, and then re-encodes it.

2. A source and channel coding approach: *Compress-and-Forward* (C&F), where a relay treats its input as an "analog" source, compresses it, and then uses a channel code to forward it.

3. A joint source–channel coding (JSCC) approach: *Amplify-and-Forward* (A&F), where a relay simply forwards its input, only applying power adjustment.

The last is indeed a non-digital JSCC approach. It does not opt to decode the input, but rather treats it as a source; furthermore, the analog processing of this source, reminiscent of analog transmission in Gaussian point-to-point communications [2], relies upon the natural matching that exists in the quadratic-Gaussian case between the "source" statistics, channel statistics and distortion measure.

In this work we concentrate on a simple test-case: The Gaussian parallel relay network, first introduced by Schein and Gallager [3]. In this network, depicted in Figure 1, all the relays are ordered in a parallel manner; the message source node is connected to the relays by a Gaussian broadcast channel (BC), while the relays are connected to the destination node by a Gaussian multiple access channel (MAC). In the original setting, all noises are white and the channels all have the same bandwidth (BW). Surprisingly, the capacity of this relatively simple network is unknown, except when the signal-to-noise ratio (SNR) of the BC section is high enough relative to the SNR of the MAC section (in which case D&F is optimal).

If the network only contained one relay, then D&F would be always optimal, and the resulting network capacity would be the minimum between the capacity of the link to the relay and the link from the relay.

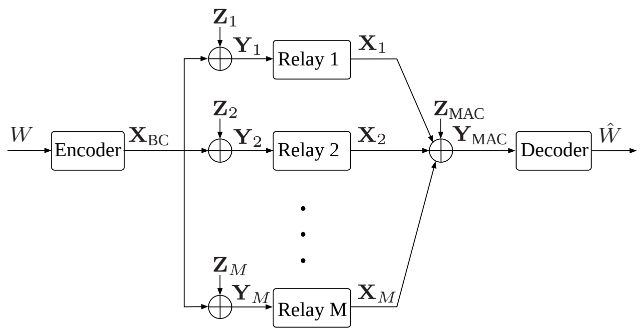


Fig. 1: The Parallel Relay Network.

The advantage of a digital scheme in such a situation was pointed out early in the development of Information Theory [4], and it stems from the ability of the relay to recover the original message of the encoder, thus avoiding *noise accumulation* between network sections. If this relay had multiple antennas, then the scheme could enjoy additional gains: multiple *receive* antennas mean that the relay can perform maximum-ratio combining (MRC) of its inputs, resulting in a gain also known as receive beamforming gain; multiple *transmit* antennas mean that the relay can perform transmit beamforming, resulting in a gain also known as array gain. We refer to these two gains as *coherence gains* with respect to both network sections. Unfortunately, if the multiple-antenna relay is replaced by multiple single-antenna relays, it is not possible to avoid noise accumulation and enjoy coherence gain w.r.t. the BC section at the same time. This is since in a digital scheme the BC noises have to be removed at the relays — but due to the distributed setting, the potential coherence gain in combining the noises is lost. Using the A&F strategy, the task of removing the BC noise is left for the final decoder, where MRC can be performed, with the penalty of noise accumulation. It is not surprising, then, that in the limit of high SNR in the MAC section, as well as in the limit of many relays [5], the A&F approach is optimal.

In this work we extend the view to networks where the noises are colored, and specifically to the important case of BW mismatch between the BC and MAC sections. In such cases, there is no hope for the A&F strategy to be optimal; for instance, in the presence of a BW mismatch, any linear processing in the relays can only result in an end-to-end channel of the minimum BW, so that the extra BW is not utilized. This disadvantage is particularly critical in the case where the section with the higher BW has a worse SNR, thus the extra BW could serve to “open the bottleneck.” Still, in order to enjoy the coherence gain we must preserve as much as possible the analog nature of the transmission. In order to overcome this problem we introduce a new relaying strategy, a “maximally analog” JSCC approach named *Rematch-and-Forward* (R&F). We show that it enjoys the same coherence gains as A&F does, yet it exploits the BW of both sections and the noise memory, beyond what is achievable by A&F.

The R&F approach is based upon treating the channel codeword as a Gaussian process having the same BW as the MAC section. A JSCC scheme is used to transform this "source" into a signal having the BW of the BC section, which allows each relay to obtain an estimate of that codeword. The relays now transmit their estimates (now back in the original BW) over the MAC section. It turns out, that if the overall scheme is to achieve the simultaneous BW utilization and coherence gain properties discussed above, this JSCC scheme needs to satisfy two conditions:

1. The estimation MSE at each relay should be optimal with respect to the capacity of the channel from the transmitter to that relay.

2. The estimation errors at the relays should be mutually uncorrelated.

The first condition cannot be satisfied by analog transmission, which does not exploit the full channel capacity (unless the "source" is transmitted over a white channel of matched BW). The second condition rules out digital transmission, where the estimation errors are all identical, being the error of some quantization applied at the encoder. We show that some JSCC schemes — variants of hybrid digital–analog schemes (HDA, [6],[7]) and of modulo-lattice modulation schemes (MLM,[8]) — satisfy the second property at all SNR, and the first condition in the limit of high SNR.

Based on these conditions, we show that the R&F approach satisfies that for given single-link capacities (i.e., a given capacity of the point-to-point channel from the encoder to a single relay, and of the point-to-point channel from a single relay to the decoder), the R&F rate remains within a *constant gap* from an upper bound on the network capacity, for any number of relays and for any ratio between the BW of the MAC and BC sections.

When using HDA techniques, R&F amounts to using superposition and frequency-division for transmitting digital and analog components over the network sections. Interestingly, a similar structure [9] was recently proposed for the same problem using a different approach. Although both points of view are valid (and lead to similar results, see Section VI), the JSCC approach has the benefit of encapsulation: it treats each link on its own, without resorting to complicated global optimization of the network.

The rest of this paper is organized as follows: We start in Section II by formally introducing the problem and the notation used. In Section III we present the concept of channel coding over JSCC schemes. Section IV contains a new view of the A&F strategy which motivates our approach. In Section V we present the R&F strategy and use it to prove an achievability result for the parallel relay network with BW mismatch. In Section VI we analyze the resulting performance and compare it with alternative strategies and with outer bounds. In Section VII we extend the results to colored noises, using a modulo-lattice modulation implementation. In Section VIII we discuss the use of the R&F strategy as a building block for more general networks.

II. Problem Formulation and Notation

In the network model of interest (recall Figure 1), a message W needs to be recovered at the destination decoder. We describe here the setting where all noises are white, although there may be a mismatch between the bandwidths of the two sections; we shall break with the whiteness assumption in Section VII. We assume a symmetric model, under which the noises in all the M branches of the BC section have the same power spectral density (PSD), and the transmission BW and power constraints are equal for all the M relays.

In continuous-time notation, let the BC noises have a flat (two-sided) PSD of level $P_{\text{BC}}/2$, while the MAC noise has flat PSD of level $P_{\text{MAC}}/2$. We may assume without loss of generality, that each relay and the source encoder are subject to the same power constraint P_X per time unit. Each channel is limited to BW B, that is, no transmission is allowed in frequencies $|f| > B$. The BC and MAC sections have BW $\rho > 0$ and $\rho = 1$, respectively; ρ is thus the BW expansion ratio of the BC section w.r.t. the MAC section. We define the SNRs

$$\text{SNR}_{\text{BC}} = \frac{P_X}{P_{\text{BC}}} \tag{1a}$$

$$\text{SNR}_{\text{MAC}} = \frac{MP_X}{P_{\text{MAC}}} \,, \tag{1b}$$

respectively. Note that the MAC SNR is defined w.r.t. the total power of all the relays.

For convenience, we work with a discrete time model, where each relay transmits an N-dimensional block and the encoder transmits a block of length $\lfloor \rho N \rfloor$. Such blocks are denoted in bold, e.g. the transmitted block of the m-th relay is $\mathbf{X}_m$; the n-th element of that block is $X_{m,n}$. In the discrete-time model, the additive mutually-independent noise sequences $\mathbf{Z}_1, \cdots, \mathbf{Z}_M$ and $\mathbf{Z}_{\text{MAC}}$ are i.i.d. Gaussian, with variances:

$$E\{Z_{m,n}^2\} = P_{\text{BC}}$$

$$E\{Z_{\text{MAC},n}^2\} = P_{\text{MAC}} \,.$$

The channel power constraints are given by:

$$\frac{1}{\rho N} E\{\|\mathbf{X}_{\mathrm{BC}}\|^2\} = \frac{P_X}{\rho}$$
$$\frac{1}{N} E\{\|\mathbf{X}_{\mathrm{MAC}}\|^2\} = P_X \,.$$

A rate R is admissible if a coding scheme exists where each of $\lceil \exp\{NR\} \rceil$ equiprobable messages can be reconstructed with arbitrarily low probability of error, and the network capacity C is the supremum of all such rates.

We denote by $C(\cdot)$ the AWGN channel capacity with SNR A:

$$C(A) \triangleq \frac{1}{2} \log(1 + \max(A, 0)). \tag{2}$$

As stated in the introduction, we shall prove our results using joint source–channel results, thus we will resort to estimation and distortion, where the useful measure turns out to be the mean squared error (MSE). When it comes to MSE distortion, our results are more easily presented in terms of *unbiased* errors, defined as follows.

*Definition 1 (**CUBE estimation**):* Let $\hat{\mathbf{S}}$ be an estimator of a random vector $\mathbf{S}$ of size N. $\hat{\mathbf{S}}$ is a correlation-sense (sample-wise) unbiased estimator (CUBE), if

$$E\{(\hat{S}_n - S_n)S_n\} = 0 \quad \forall n = 1, \ldots, N \,.$$

Note that any estimator which is unbiased in the strong sense, i.e.,

$$E\{\hat{\mathbf{S}} - \mathbf{S} | \mathbf{S}\} = 0 \quad \forall \mathbf{S} \,,$$

is CUBE. However, weak-sense unbiased estimator, where $E\{\hat{\mathbf{S}} - \mathbf{S}\} = 0$, is not necessarily CUBE. Also note that any estimator $\tilde{S}$ of S can be transformed into a CUBE estimator $\hat{S}$ by multiplication by a suitable constant:

$$\hat{S} = \frac{E\{S^2\}}{E\{S\tilde{S}\}} \tilde{S} \,.$$

In the extreme case of an estimator which is uncorrelated with the source, this constant approaches infinity and so does the CUBE MSE, unlike the biased MSE which need not exceed the source variance. We define the signal-to-distortion ratio w.r.t. a CUBE estimator:

$$\mathrm{SDR} \triangleq \frac{E\{\|\mathbf{S}\|^2\}}{E\{\|\hat{\mathbf{S}} - \mathbf{S}\|^2\}} \,. \tag{3}$$

Note that this definition is different than the standard (biased) definition of SDR. In terms of this CUBE SDR, the unbiased quadratic-Gaussian white rate–distortion function (see also [10]):

$$R(D) \triangleq \inf_{w(\hat{S}|S): E\{(\hat{S}-S)^2\} \le D, E\{(\hat{S}-S)S\}=0} I(S; \hat{S}) \tag{4}$$

equals $C(\text{SDR})$. We use this rate–distortion function (rather than the more common biased one) throughout the paper.[1]

III. Data Transmission Over JSCC Schemes

In this section we present some results that will be needed in the sequel, connecting the performance of a JSCC scheme with achievable rates of a channel encoder–decoder pair which encapsulates it.

Proposition 1: Let S be some memoryless source. Let each N-dimensional reconstruction block $\hat{\mathbf{S}}$ be drawn according to some conditional distribution $w(\hat{\boldsymbol{s}}|\boldsymbol{s})$, i.i.d. between blocks. Denote the (scalar) distortion measure by $d(S_i, \hat{S}_i)$ and the corresponding distortion between the blocks $\mathbf{S}$ and $\hat{\mathbf{S}}$ — by

$$d(\mathbf{S}, \hat{\mathbf{S}}) = \frac{1}{N} \sum_{i=1}^{N} d(S_i, \hat{S}_i)\,.$$

If the expected distortion satisfies

$$E\{d(\mathbf{S}, \hat{\mathbf{S}})\} \leq N D_0\,, \tag{5}$$

then there exists a channel coding scheme of any rate

$$R < R(D_0) \tag{6}$$

over the channel from S to $\hat{S}$, where $R(D)$ is the rate–distortion function of $\mathbf{S}$ w.r.t. the distortion measure $d(\cdot,\cdot)$.

Proof: According to the rate–distortion theorem,

$$I(\mathbf{S}, \hat{\mathbf{S}}) \geq N R(D_0)\,. \tag{7}$$

Consider now the "block channel" between $\mathbf{S}$ and $\hat{\mathbf{S}}$, described by the conditional distribution $w(\hat{\boldsymbol{s}}|\boldsymbol{s})$. Let R_N be a coding rate for this channel, i.e., a rate per a "block input" $\mathbf{S}$. According to the channel coding theorem, any

$$R_N < I(\mathbf{S}, \hat{\mathbf{S}})$$

is achievable. Re-normalizing per a single (scalar) input, we have that any rate

$$R = \frac{1}{N} R_N < \frac{1}{N} I(\mathbf{S}, \hat{\mathbf{S}}) \tag{8}$$

[1] This biased RDF is achievable, since it is exactly the SDR obtained when removing the bias from a reconstruction that obeys the optimal (biased) quadratic-Gaussian test channel. To the contrary, it cannot be exceeded, or introducing an MMSE factor would result in surpassing the performance of the quadratic-Gaussian (biased) RDF.

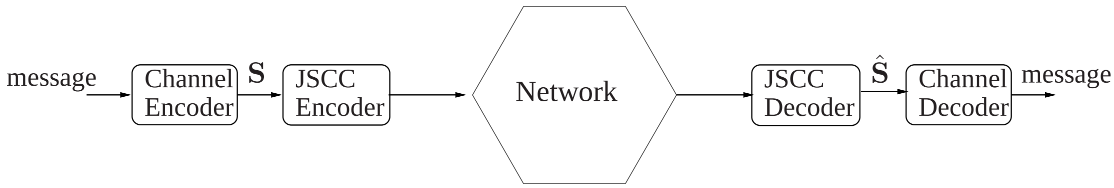


Fig. 2: Channel coding using a JSCC scheme

is achievable. Combining the results of (7) and (8) we conclude that any rate satisfying

$$R \leq R(D_0) \tag{9}$$

is achievable. ■

Remark 1: When working in this encapsulated manner, the input $\mathbf{S}$ is not drawn according to a memoryless distribution, but instead it is a symbol of the block-channel codebook. Nevertheless, considering the distortion $d(\mathbf{S}, \hat{\mathbf{S}})$ averaged over the memoryless source distribution is sufficient.

We now specialize this result to the white Gaussian RDF. In terms of the CUBE performance (see Definition 1) of the JSCC scheme, Proposition 1 becomes the following.

Corollary 1: If the CUBE SDR satisfies

$$\text{SDR} \geq \text{SDR}_0 \,,$$

then there exists a channel coding scheme for any rate

$$R < C(\text{SDR}_0)$$

over the channel from $\mathbf{S}$ to $\hat{\mathbf{S}}$.

We give this result the following interpretation (see Figure 2): For some communication network, if we can find a scheme, containing any block functions (of dimension N) at the network nodes (which we view as JSCC encoders and decoders), such that the expected SDR is at least SDR_0, then the rate $C(\text{SDR}_0)$ can be approached in transmission between these nodes by viewing the whole scheme as an equivalent channel and adding an external channel encoder/decoder pair. The resulting scheme is as follows.

Scheme 1: 1) Use a codebook with 2^{KNR} words; each word is composed of K super-symbols, each super-symbol $\mathbf{S}$ composed of N real numbers.

2) Feed each such super-symbol $\mathbf{S}$ is independently to the JSCC scheme, resulting in a reconstruction $\hat{\mathbf{S}}$.

3) The decoder recovers the message from the JSCC scheme outputs, assuming that the outputs $\hat{\mathbf{S}}_1, \ldots, \hat{\mathbf{S}}_K$ were created from the corresponding $\mathbf{S}_1, \ldots, \mathbf{S}_K$ by a memoryless $w(\hat{\mathbf{S}}|\mathbf{S})$. the (distorted) source-blocks $\hat{\mathbf{S}}$ using the JSCC decoder.

Remark 2: One may construct the codebook using randomness as follows. The codebook is drawn as $2^{KNR} \times (KN)$ i.i.d. Gaussian values. Then, in average over these codebooks, the CUBE SDR of the JSCC scheme is at least SDR_0, thus for any rate below $C(\text{SDR}_0)$ the error probability of the ensemble approaches zero; consequently, at least one codebook has vanishing error probability.

Remark 3: This coding strategy is similar to [11, Theorem 10], although the approach there suggests to eliminate the JSCC encoder and decoder rather than use them as part of the scheme.

Remark 4: In general this approach may require a highly complex decoder which takes into account the (possibly non-white) transition distribution $w(\hat{\mathbf{S}}|\mathbf{S})$. In the Gaussian case, however, super-symbols are not required and a simple Euclidean decoder is always sufficient, see e.g. [12, Theorem 2]. Interestingly, a similar result to Proposition 1 was proven for discrete alphabets in [13] without resorting to super-symbols.

IV. A&F as a Joint Source–Channel Approach

In the simple case where both sections have the same BW ($\rho = 1$) and the noises are white, we can use a codebook which is good for transmission over an AWGN channel in conjunction with the A&F strategy. In this strategy, each relay simply forwards a scaled version of its input:

$$\mathbf{X}_m = \gamma \mathbf{Y}_m = \gamma(\mathbf{X}_{\text{BC}} + \mathbf{Z}_m) \tag{10}$$

where the relay amplification factor is set by the channel power constraint:

$$\gamma = \sqrt{\frac{P_X}{P_X + P_{\text{BC}}}} = \sqrt{\frac{\text{SNR}_{\text{BC}}}{1 + \text{SNR}_{\text{BC}}}} \ .$$

Consequently, the decoder receives

$$\begin{aligned} \mathbf{Y}_{\text{MAC}} &= \sum_{m=1}^{M} \mathbf{X}_m + \mathbf{Z}_{\text{MAC}} \\ &= \gamma(M\mathbf{X}_{\text{BC}} + \sum_{m=1}^{M} \mathbf{Z}_m) + \mathbf{Z}_{\text{MAC}} \ . \end{aligned}$$

From the point of view of the decoder, this is equivalent to a point-to-point AWGN channel with SNR:

$$\begin{aligned} \text{SDR}_{\text{AF}} &= \frac{E\{\|\gamma M \mathbf{X}_{\text{BC}}\|^2\}}{E\{\|\gamma \sum_{m=1}^{M} \mathbf{Z}_m\|^2\} + E\{\|\mathbf{Z}_{\text{MAC}}\|^2\}} \\ &= M \left(\text{SNR}_{\text{BC}} \upharpoonleft\!\upharpoonright \text{SNR}_{\text{MAC}}\right) \ , \end{aligned} \tag{11}$$

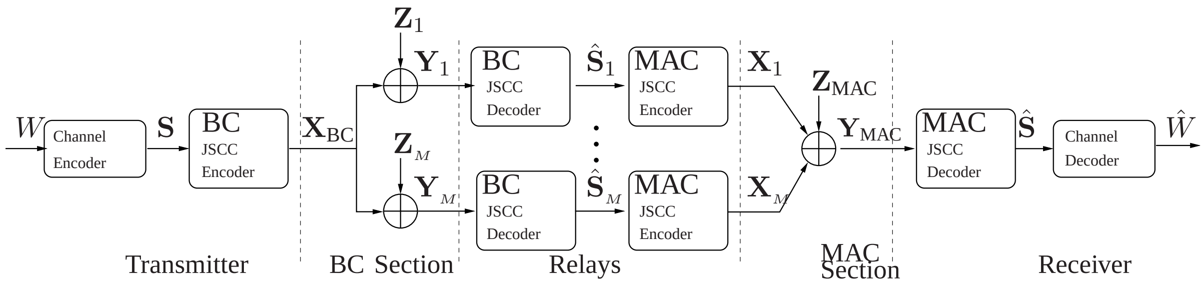


Fig. 3: A&F strategy as the transmission of a codeword over the concatenation of JSCC schemes

where

$$A \upuparrows B \triangleq \frac{AB}{1 + A + B} \tag{12}$$

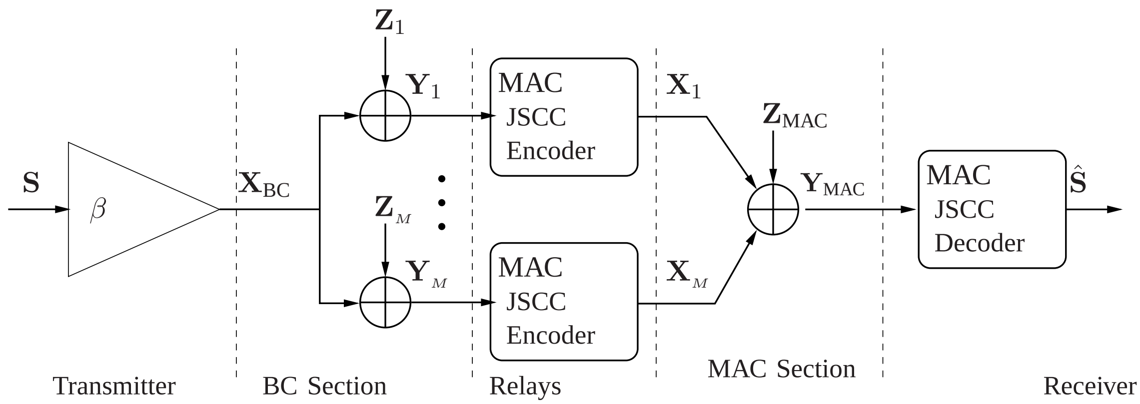
is the equivalent SNR in transmitting a signal through the concatenation of additive noise channels of SNRs A and B.[2] The resulting achievable rate of the network is as in [3]:

$$R_{\text{AF}} = C\left(\text{SDR}_{\text{AF}}\right). \tag{13}$$

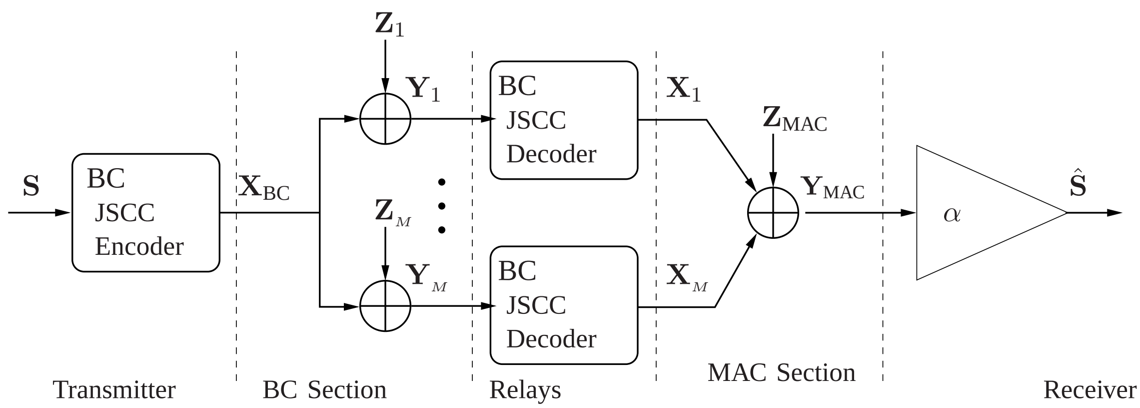
This rate reflects the coherence gain w.r.t. the noises of both sections, as explained in the introduction. However, when we leave the white equal-BW case, the A&F strategy fails to fully exploit the capacity of the individual channels; for example, it is restricted to the minimal BW of the two sections, since it is *fully* analog.

We present here an alternative view of the A&F strategy. Using the analysis in Section III, we think of the scheme as a joint source–channel coding (JSCC) scheme over the relay network, surrounded by a channel encoder–decoder pair. The sole purpose of this JSCC scheme is to produce a "good" estimate of the codeword $\mathbf{S}$ at the decoder; in light of Corollary 1, the relevant measure of goodness is MSE distortion. For the purpose of this scheme, $\mathbf{S}$ is treated as an i.i.d. Gaussian signal (drawn from an i.i.d. source rather than from a codebook). For the parallel relay network topology, the JSCC scheme is a concatenation of two schemes, from the encoder to the relays and from the relays to the decoder. The A&F strategy uses the simplest JSCC schemes, i.e. analog transmission, for both sections. While this choice is not optimal in general, we show in the sequel that if we restrict *one* of the sections to be analog, then the other may be analog as well without further loss.

[2]Here for each channel, the SNR is defined as the ratio of its input to noise power, i.e. for the second channel the "signal" includes the noise of the first channel. This causes the additional term "1" in the denominator, without which this would have been a scaled harmonic mean of the SNRs.



(a) The data collection problem



(b) The data distribution problem

Fig. 4: Sensor network problems and the parallel relay network.

A classic result by Goblick [2] states that analog transmission is optimal in the white quadratic-Gaussian point-to-point setting. Recent works, e.g. [14], [15], extend this property to some network settings. Gastpar [16] shows the optimality of analog transmission in a problem of particular interest to us: a sensor network problem, where agents observe a white Gaussian source contaminated by white Gaussian noise (i.i.d. between the agents), and communicate with a central decoder through a Gaussian MAC; the aim of the decoder is to reconstruct the source, subject to an MMSE criterion. Thus, this is a combination of the quadratic-Gaussian CEO problem with the Gaussian MAC problem.

This setting is equivalent to the joint source–channel problem between $\mathbf{S}$ and $\hat{\mathbf{S}}$ in Figure 3, under the constraint that the JSCC scheme over the BC section must be analog, see Figure 4a. We present the result, then, in the relay network notation.

*Theorem 1 (**Collection / "Sensors Network"**, [16]):* In the white Gaussian relay network, if $\mathbf{S}$ is a white Gaussian source and $\mathbf{X}_{\text{BC}} = \beta\mathbf{S}$ for some β satisfying the power constraint, the CUBE SDR of the reconstruction $\hat{\mathbf{S}}$ satisfies

$$\text{SDR} \leq \text{SDR}_{\text{AF}}\,,$$

where SDR_{AF} was defined in (11).

We now look at a dual problem: Suppose that we can distribute $\mathbf{S}$ to the relays using any JSCC scheme, but the estimation of $\mathbf{S}$ at the final decoder is obtained by an analog operation. This is again equivalent to the joint source–channel problem between $\mathbf{S}$ and $\hat{\mathbf{S}}$ in Figure 3, only now under the constraint that the JSCC scheme over the MAC section must be analog, see Figure 4b. For this problem, we have a similar result.

*Theorem 2 (**Distribution / "Emitters Network"**):* In the white Gaussian relay network, if $\mathbf{S}$ is a white Gaussian source and $\hat{\mathbf{S}} = \alpha\mathbf{Y}_{\text{MAC}}$ for some α, the CUBE SDR of the reconstruction $\hat{\mathbf{S}}$ satisfies

$$\text{SDR} \leq \text{SDR}_{\text{AF}}\,,$$

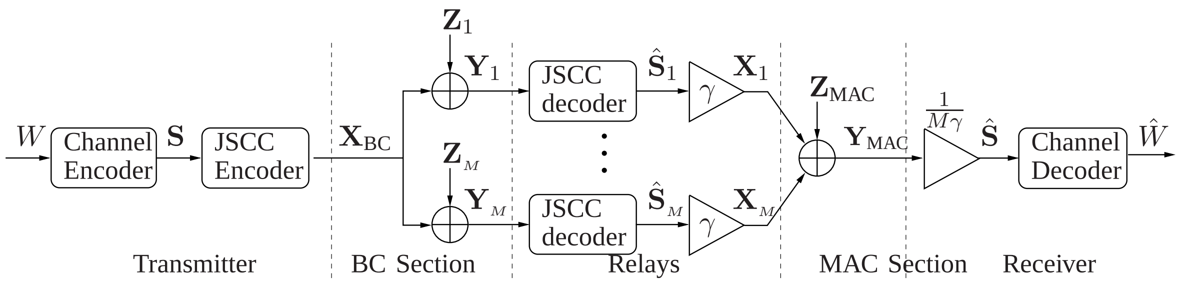
where SDR_{AF} was defined in (11).

The proof is given in Appendix A. It is based upon considering the joint statistics of the CUBE errors at the relays, denoted by $\tilde{\mathbf{Z}}_m = \hat{\mathbf{S}}_m - \mathbf{S}$. The proof shows that errors which are uncorrelated and have fixed variance, both in time and in the spatial dimension, are optimal, with a minimum SDR that equals the BC SNR. Thus, from the point of view of a linear decoder under a quadratic distortion measure, these errors are equivalent to the BC noises. In the white equal-BW case, they may as well be these noises themselves. In the following sections, we define a generalization of this principal: *additive* JSCC schemes.

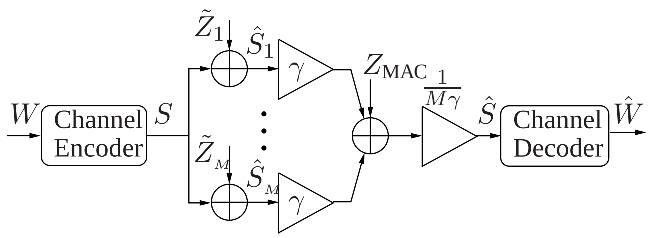
We see, then, that if either the encoder or the decoder are restricted to be scalar and linear, then the whole scheme from $\mathbf{S}$ to $\hat{\mathbf{S}}$ may be scalar and linear as well, without increasing the distortion. The A&F strategy can be thus described as follows.

*Scheme 2 (**Amplify-and-Forward**):*

1) **Channel Encoding**: Choose a codeword $\mathbf{S}$ from an i.i.d. Gaussian codebook.
2) **Distribution:** Over the BC section, use the optimum strategy that obtains reconstructions $\{\hat{\mathbf{S}}_m\}$ at the relays, under the assumptions that $\mathbf{S}$ is a sequence from a white Gaussian source and that the final decoder is linear and scalar.
3) **Collection:** Over the MAC section, use the optimum strategy that obtains a final reconstruction $\hat{\mathbf{S}}$, under the assumptions that $\mathbf{S}$ is a sequence from a white Gaussian source, that the estimates $\{\hat{\mathbf{S}}_m\}$ are the BC channel outputs and the encoder is linear and scalar.



(a) The R&F scheme



(b) Equivalent A&F channel

Fig. 5: The R&F Strategy for the parallel relay network.

4) **Channel Decoding:** Treat the estimation $\hat{\mathbf{S}}$ as the output of a channel, and decode the message.

We shall see in the sequel, that the rematch-and-forward strategy extends these principles beyond the white equal-BW case.

V. Rematch-and-Forward for Bandwidth Mismatch

We now turn to present the Rematch-and-Forward (R&F) strategy for the BW mismatch case. We follow the steps of the A&F strategy, as defined in Scheme 2. We use a random white Gaussian codebook, and choose it to have the BW of the MAC section. Consequently, analog transmission remains optimal for the collection stage (over the MAC section). However, for the distribution stage (over the BC section) we need to replace the analog transmission by an adequate JSCC scheme. Figure 5 shows the resulting structure of the R&F strategy: Some JSCC scheme, to be specified later, is used over the BC section, while analog transmission is used over the MAC section. Seeing the CUBE errors (recall Definition 1) of the JSCC decoders at the relays as "channel noises", we have the equivalent channel of Figure 5b, which in turn is just the AF strategy applied to a white relay network, where all the links have the original BW of the MAC section. The following theorem states our main achievability result, using that approach. Our rate expression makes use of the equivalent, (or "mutual-information preserving") SNR of the BC section, denoted by $\overline{\mathrm{SNR}}_{\mathrm{BC}}$, satisfying

$$C(\overline{\mathrm{SNR}}_{\mathrm{BC}}) = \rho C\left(\frac{\mathrm{SNR}_{\mathrm{BC}}}{\rho}\right), \tag{14}$$

which is equivalent to

$$\overline{\mathrm{SNR}}_{\mathrm{BC}} = \left(1+\frac{\mathrm{SNR}_{\mathrm{BC}}}{\rho}\right)^{\rho} - 1\,. \tag{15}$$

*Theorem 3 (**R&F Performance for BW Mismatch**):* The capacity of the Gaussian parallel relay network (per MAC section use) is lower-bounded by R_{RF}, where for $\rho > \mathbf{1}$:

$$R_{\mathrm{RF}} = C\Big(M(\mathrm{SNR}'_{\mathrm{BC}} \upharpoonleft\!\upharpoonright \mathrm{SNR}_{\mathrm{MAC}})\Big), \tag{16}$$

where the operator $\upharpoonleft\!\upharpoonright$ was defined in (12) and:

$$\mathrm{SNR}'_{\mathrm{BC}} = \left(1+\frac{\mathrm{SNR}_{\mathrm{BC}}}{\rho}\right)^{\rho-1}\frac{\mathrm{SNR}_{\mathrm{BC}}}{\rho} = \frac{\mathrm{SNR}_{\mathrm{BC}}}{\rho+\mathrm{SNR}_{\mathrm{BC}}}(1+\overline{\mathrm{SNR}}_{\mathrm{BC}})\,, \tag{17}$$

while for $\rho \leq \mathbf{1}$:

$$R_{\mathrm{RF}} = \rho\cdot C\Big(M(\overline{\mathrm{SNR}}_{\mathrm{BC}} \upharpoonleft\!\upharpoonright \mathrm{SNR}_{\mathrm{MAC}})\Big) + (1-\rho)\cdot C\Big(\overline{\mathrm{SNR}}_{\mathrm{BC}} \upharpoonleft\!\upharpoonright (M\,\mathrm{SNR}_{\mathrm{MAC}})\Big). \tag{18}$$

The rest of this section is devoted to proving this result. Section V-A and Section V-B contain the proofs of (16) and (18), respectively.

A. Transmission over the BC Section with BW Expansion

For $\rho \geq 1$, the JSCC scheme used to materialize the distribution stage of Scheme 2 must be a "good" scheme for BW expansion. Recall that the coherence gain for A&F was achieved due to the mutual independence of the BC section noises; the following definition and lemma provide us with a sufficient condition for a scheme to produce such an equivalent channel.

*Definition 2 (**Additive JSCC scheme**):* A JSCC scheme for a source $\mathbf{S}$ is *additive with error probability* p_e, if there exists an event $\mathcal{C}$ of probability at least $1-p_e$, such that the CUBE error $\tilde{\mathbf{Z}}$ is statistically independent of $\mathbf{S}$ given the encoding and decoding functions and the event $\mathcal{C}$.

In this definition, the error probability corresponds to a decoding failure of some digital element of the scheme, if such an element exists. It is understood, that by taking a large enough block length this probability can be made arbitrarily small. The additivity in this definition is in a point-to-point setting; the following lemma translates this to the BC setting.

Note that the independence of $\tilde{\mathbf{Z}}$ in $\mathbf{S}$, *given the encoding and decoding functions*, implies that $\tilde{\mathbf{Z}}$ and $\mathbf{S}$ are independent given any common randomness shared by the encoder and the decoder.

*Lemma 1 (**Additive JSCC scheme over a BC**):* A JSCC scheme for a source $\mathbf{S}$ is used over a BC

$$w_M(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M|\mathbf{x}) = \prod_{m=1}^{M} w(\boldsymbol{y}_m|\mathbf{x}) \tag{19}$$

by applying the decoding function of the JSCC scheme to each channel output $\mathbf{Y}_m$. If the scheme is additive with error probability p_e, then there exists an event $\mathcal{C}_M$ of probability at least $1 - Mp_e$ such that the CUBE errors $\{\tilde{\mathbf{Z}}_m\}$ are mutually independent given $\mathcal{C}_M$ and the encoding and decoding functions.

Proof: The probability that *all* $\left\{\tilde{\mathbf{Z}}_m\right\}$ are independent of $\mathbf{S}$ is lower-bounded by the union-bound

$$P(\mathcal{C}_M) \geq 1 - Mp_e \,.$$

Given the event $\mathcal{C}_M$ and the encoding and decoding functions, $f(\cdot)$ and $g(\cdot)$, we have

$$P\left(\left\{\tilde{\mathbf{Z}}_m \leq \mathbf{z}_m\right\}_{m=1}^{M} \middle| \mathcal{C}_M, f, g\right) = \int P\left(\left\{\tilde{\mathbf{Z}}_m \leq \mathbf{z}_m\right\}_{m=1}^{M} \middle| \mathcal{C}_M, f, g, \mathbf{S} = \boldsymbol{s}\right) f_{\mathbf{S}}(\boldsymbol{s}) d\boldsymbol{s} \quad ,$$

where $f_{\mathbf{S}}(\cdot)$ is the probability distribution function of the source $\mathbf{S}$. Under this conditioning, $\tilde{\mathbf{Z}}_m$ is a function of the channel noise $\mathbf{Z}_m$ only, and hence independent of all other CUBE errors $\left\{\tilde{\mathbf{Z}}_m\right\}_{i\neq m}$, due to the BC structure (19). Thus,

$$\begin{aligned} P\left(\left\{\tilde{\mathbf{Z}}_m \leq \mathbf{z}_m\right\}_{m=1}^{M} \middle| \mathcal{C}_M, f, g\right) &= \int \prod_{m=1}^{M} P\left(\tilde{\mathbf{Z}}_m \leq \mathbf{z}_m \middle| \mathcal{C}_M, f, g, \mathbf{S} = \boldsymbol{s}\right) f_{\mathbf{S}}(\boldsymbol{s}) d\boldsymbol{s} \\ &= \prod_{m=1}^{M} P\left(\tilde{\mathbf{Z}}_m \leq \mathbf{z}_m \middle| \mathcal{C}_M, f, g\right) \,, \end{aligned}$$

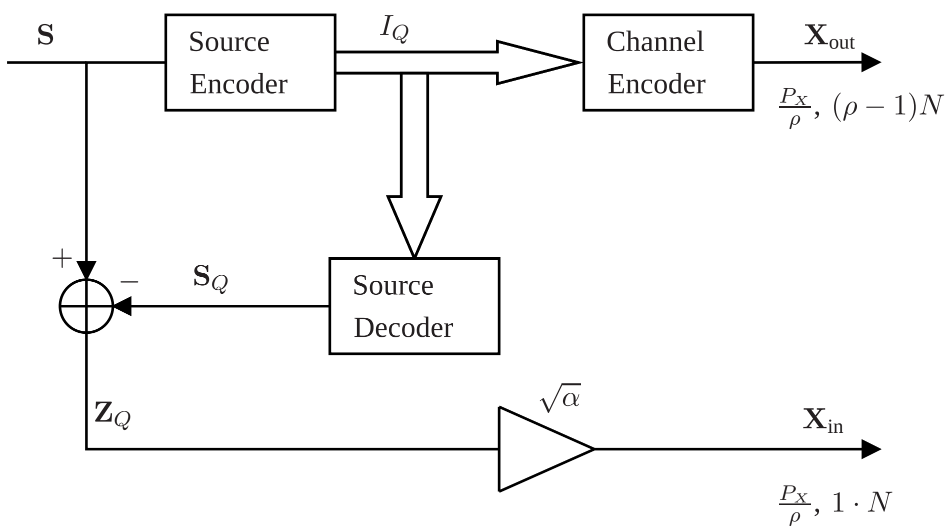
where the last equality holds true since $\left\{\tilde{\mathbf{Z}}_m\right\}$ are independent of $\mathbf{S}$ given $\mathcal{C}_M$ and the decoding and encoding functions. ∎

The additivity allows us to express the SDR of $\hat{\mathbf{S}}$ in terms of the point-to-point performance of the BC section JSCC scheme as follows.

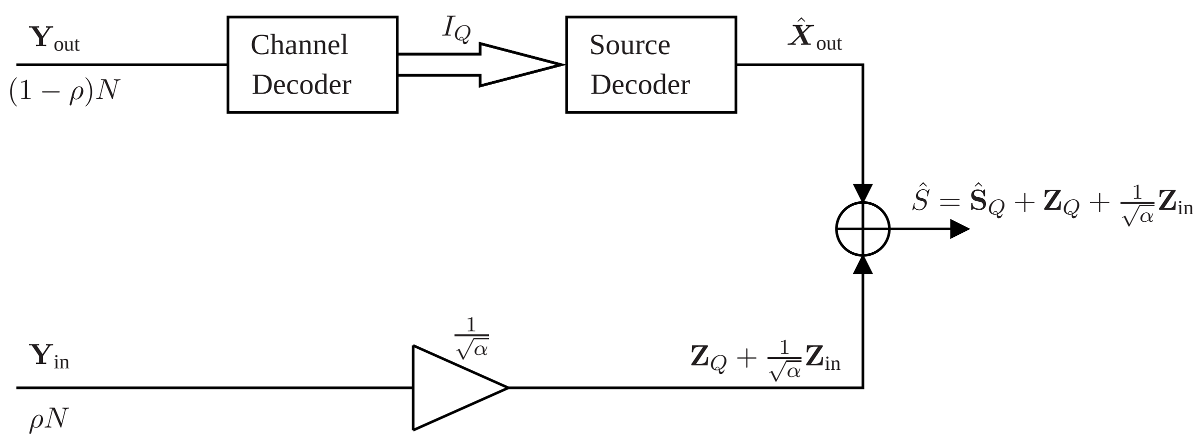
Lemma 2: Consider using over the BC section a sequence of additive JSCC schemes, indexed by the block length N, with error probabilities p_{eN}, and with CUBE SDR of $\overline{\text{SNR}}_{\text{BC}}$ (when there is no error). If $\lim_{N\to\infty} p_{eN} = 0$, then $\lim_{N\to\infty} \text{SDR} = M(\overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}})$.

Proof: For $p_e = 0$, the result follows with an appropriate choice of factors, just as in A&F. It remains to show that the effect of errors is small in the limit of small p_e; this follows from basic probability-theoretic arguments; see the proof of [17, (5.2)]. ∎

Recall that by Corollary 1, any channel rate below $C(\text{SDR}_0)$ is achievable over a general channel that has CUBE SDR of SDR_0 for a white Gaussian input. Thus, Lemma 2 proves the achievability of the



(a) Encoder



(b) Decoder

Fig. 6: HDA Scheme for BW Expansion.

rate of (16), if SNR'_{BC} is the CUBE SDR of an additive JSCC scheme with a white Gaussian input. It remains to demonstrate that there exists a scheme that achieves SNR'_{BC} of (17). To that end, we use an additive variant of a scheme by Mittal and Phamdo, depicted also in Figure 6; later in Section VII we present an alternative which has the same performance in the high-SNR limit.

In this section, we denote the first $(\rho-1)N$ samples of an N-dimensional vector and the rest of the vector by subscripts 'out' and 'in', respectively. For example, $\mathbf{X}_{\text{out}}$ and $\mathbf{X}_{\text{in}}$ denote the first $(\rho-1)N$ and the last N entries of the vector $\mathbf{X}$ of length ρN.

*Scheme 3 (**HDA Scheme for BW expansion**, after [6, System 3]):*

Encoding:

1) Quantize the source vector $\mathbf{S}$ of length N, i.e., split $\mathbf{S}$ into a quantized value $\mathbf{S}_Q$ and a quantization error $\mathbf{Z}_Q$.
2) Use a digital channel code (satisfying the average channel power constraint $P_{\mathbf{X}}/\rho$) to obtain $(\rho-1)N$ channel inputs representing $\mathbf{S}_Q$, denoted by $\mathbf{X}_{\text{out}}$.
3) Apply power adjustment to the quantization error vector $\mathbf{Z}_Q$ (of length N), i.e., multiply it by $\sqrt{\alpha}$, where

$$\alpha \triangleq \frac{P_{\mathbf{X}}/\rho}{\sigma^2_{\mathbf{Z}_Q}}, \tag{20}$$

and send the resulting inflated error signal $\mathbf{X}_{\text{in}} = \sqrt{\alpha}\mathbf{Z}_Q$ over the remaining N samples, in an analog manner.

Decoding:

1) Decode the channel code from the first $(\rho-1)N$ output samples $\mathbf{Y}_{\text{out}}$ (corresponding to $\mathbf{X}_{\text{out}}$), and then the source code, to recover the quantized source representation $\hat{\mathbf{S}}_Q$.
2) Reverse the power adjustment applied to the remaining N channel outputs $\mathbf{Y}_{\text{in}}$, i.e., multiply $\mathbf{Y}_{\text{in}}$ by $\sqrt{1/\alpha}$:

$$\sqrt{\frac{1}{\alpha}}\mathbf{Y}_{\text{in}} = \mathbf{Z}_Q + \sqrt{\frac{1}{\alpha}}\mathbf{Z}_{\text{in}}\,.$$

3) Add $\sqrt{\mathbf{Y}_{\text{in}}/\alpha}$ to $\hat{\mathbf{S}}_Q$ to arrive at the (unbiased) reconstructed signal

$$\hat{\mathbf{S}} = \hat{\mathbf{S}}_Q + \sqrt{\frac{1}{\alpha}}\mathbf{Y}_{\text{in}} = \hat{\mathbf{S}}_Q + \mathbf{Z}_Q + \sqrt{\frac{1}{\alpha}}\mathbf{Z}_{\text{in}}\,. \tag{21}$$

Proposition 2: For a white Gaussian source, Scheme 3 is additive with vanishing error probability as $N \to \infty$. Furthermore,

$$\widetilde{\text{SDR}} = \left(1 + \frac{\text{SNR}}{\rho}\right)^{\rho-1} \frac{\text{SNR}}{\rho}\,, \tag{22}$$

where $\widetilde{\text{SDR}}$ is the scheme CUBE SDR given that no error event occurred, and SNR is the channel SNR.

Proof: By the separation principle (see, e.g., [18]), a distortion (viz. average power of $\mathbf{Z}_Q$) satisfying

$$\log\left(\frac{\sigma_S^2}{\sigma_{Z_Q}^2}\right) = (\rho-1)\log\left(1 + \frac{\text{SNR}}{\rho}\right), \tag{23}$$

is achievable, with vanishing error probability.

Assuming no decoding error was made, the reconstructed signal (21) is equal to

$$\hat{\mathbf{S}} = \mathbf{S} + \frac{1}{\sqrt{\alpha}}\mathbf{Z}_{\text{in}} = \mathbf{S} + \sqrt{\frac{\sigma^2_{\mathbf{Z}_Q}}{P_X/\rho}}\mathbf{Z}_{\text{in}}\,,$$

which implies in turn an unbiased distortion of

$$E\left[\left(\hat{S}-S\right)^2\right] = \frac{\sigma^2_{Z_Q}}{\mathrm{SNR}/\rho}\,. \tag{24}$$

By combining (23) and (24) the desired (unbiased) SDR is achieved:

$$\widetilde{\mathrm{SDR}} = \frac{\sigma_S^2}{E\left[\left(\hat{S}-S\right)^2\right]} = \left(1+\frac{\mathrm{SNR}}{\rho}\right)^{\rho-1}\frac{\mathrm{SNR}}{\rho}\,. \tag{25}$$

■

Since, when substituting $\mathrm{SNR} = \mathrm{SNR}_{\mathrm{BC}}$, the $\widetilde{\mathrm{SDR}}$ of (22) is equal to $\mathrm{SNR}'_{\mathrm{BC}}$ of (17), the proof of Theorem 3 for $\rho > 1$ is completed. Note that $\mathrm{SNR}'_{\mathrm{BC}} < \overline{\mathrm{SNR}}_{\mathrm{BC}}$, where the difference, which can be seen as a penalty for additivity, vanishes in the high-SNR limit. In the original scheme of Mittal and Phamdo [6], the full mutual information of the channel is exploited by multiplying the analog channel outputs by an MMSE factor, before adding them to the quantized values. While this reduces the estimation error, it causes it to depend on the quantization error, thus turning the scheme unto a non-additive one. We conjecture that no additive JSCC scheme over a Gaussian channel with BW expansion can exploit the full mutual information.

B. Transmission over the BC Section with BW Compression

We now show how (18) can be approached. Note, that in this rate expression the coherence gain w.r.t. the BC noises is achieved only over a portion ρ of the BW. This occurs since now the JSCC scheme over the BC section needs to perform BW compression rather than expansion. Intuitively speaking, such a scheme cannot be additive in the sense of Definition 2, since the channel does not supply enough degrees of freedom. Consequently, the equivalent noise seen by the channel coding scheme is not white, and the channel scheme must be chosen accordingly. We start, then, by defining the additivity property which is applicable to this case.

*Definition 3 (ρ-**Additive JSCC scheme**):* A JSCC scheme is *ρ-additive with error probability* p_e, if the last ρN elements of the CUBE error $\tilde{\mathbf{Z}}$ are additive with error probability p_e, in the sense of Definition 2.

In this section, we denote the first $(1-\rho)N$ samples of an N-dimensional vector and the rest of the vector by subscripts ‘out’ and ‘in’, respectively. Imagine that we have a ρ-additive JSCC scheme, which achieves the same error-free CUBE SDR of $\overline{\mathrm{SNR}}_{\mathrm{BC}}$ for both $\mathbf{S}_{\mathrm{out}}$ and $\mathbf{S}_{\mathrm{in}}$. By Lemma 1, $\{\tilde{\mathbf{Z}}_{\mathrm{in},m}\}$ are mutually independent; $\{\tilde{\mathbf{Z}}_{\mathrm{out},m}\}$, on the other hand, may be arbitrarily correlated. This leads to the following.

Lemma 3: Consider using over the BC section a sequence of ρ-additive JSCC schemes with error probabilities p_{eN}, and with CUBE SDR of $\overline{\text{SNR}}_{\text{BC}}$ (when there is no error), equal for both sub-vectors $\mathbf{S}_{\text{out}}$ and $\mathbf{S}_{\text{in}}$. If $\lim_{N\to\infty} p_{eN} = 0$ then

$$\lim_{N\to\infty} \text{SDR}_{\text{in}} = M(\overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}})$$
$$\lim_{N\to\infty} \text{SDR}_{\text{out}} = \overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright (M\,\text{SNR}_{\text{MAC}})\,.$$

Proof: As in the proof of Lemma 2, we only need to verify for $p_e = 0$. The result for SDR_{in} follows from the analysis of A&F; for SDR_{out}, make the worst-case assumption that $\{\tilde{\mathbf{Z}}_{\text{out},m}\}$ are identical to derive the desired result. ■

Again, in order to approach the performance promised by Theorem 3 we must have a scheme that achieves $\overline{\text{SNR}}_{\text{BC}}$ of (15) for a white Gaussian input. Another HDA scheme by Mittal and Phamdo, depicted in Figure 7, qualifies.

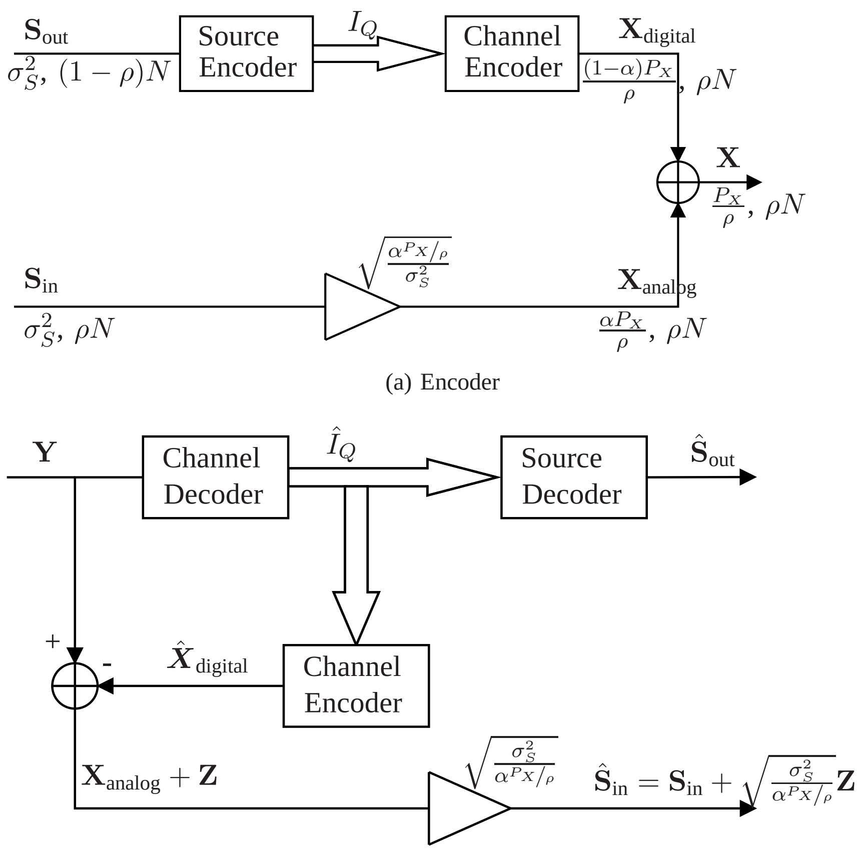
*Scheme 4 (**HDA Scheme for BW compression**, after [6, System 4]):*

Encoding:

1) Quantize the vector composed of the first $(1-\rho)N$ source samples $\mathbf{S}_{\text{out}}$, and denote the corresponding reconstructed signal by $\hat{\mathbf{S}}_{\text{out}}$; then use a digital channel code, to obtain ρN channel inputs denoted by $\mathbf{X}_{\text{digital}}$ representing $\hat{\mathbf{S}}_{\text{out}}$, of average power $(1-\alpha)^{P_{\mathbf{X}}}/_{\rho}$, where
$$\alpha = \frac{(1 + {}^{\text{SNR}}/_{\rho})^{\rho} - 1}{{}^{\text{SNR}}/_{\rho}}, \tag{26}$$
which satisfies $0 \le \alpha \le 1$ for all $\rho \le 1$.
2) Multiply the remaining ρN source samples S_{in} by $\sqrt{\frac{\alpha^{P_X}/_{\rho}}{\sigma_S^2}}$, to form $\mathbf{X}_{\text{analog}}$ of average power $(1-\alpha)^{P_{\mathbf{X}}}/_{\rho}$:
$$\mathbf{X}_{\text{analog}} = \sqrt{\alpha \frac{P_X/\rho}{\sigma_S^2}}\mathbf{S}_{\text{in}}\,.$$
3) Transmit the sum of $\mathbf{X}_{\text{digital}}$ and $\mathbf{X}_{\text{analog}}$, namely $\mathbf{X} = \mathbf{X}_{\text{digital}} + \mathbf{X}_{\text{analog}}$, whose length is ρN and average power — ${}^{P_{\mathbf{X}}}/_{\rho}$.

Decoding:

1) Decode the channel code $\mathbf{X}_{\text{digital}}$, treating $\mathbf{X}_{\text{analog}}$ as noise, then decode the source code to recover $\hat{\mathbf{S}}_{\text{out}}$.
2) Subtract the decoded $\mathbf{X}_{\text{out}}$ from $\mathbf{Y}$ and multiply the result by $\sqrt{\frac{\sigma_S^2}{\alpha^{P_X}/_{\rho}}}$, to attain the unbiased reconstruction of $\mathbf{S}_{\text{in}}$ which equals to (assuming no channel decoding errors were made)
$$\hat{\mathbf{S}}_{\text{in}} = \mathbf{S}_{\text{in}} + \sqrt{\frac{\sigma_S^2}{\alpha^{P_X}/_{\rho}}}Z\,.$$



(a) Encoder

(b) Decoder (assuming $\hat{X}_{\text{out}} = X_{\text{out}}$)

Fig. 7: HDA Scheme for BW Compression.

Proposition 3: For a white Gaussian input, Scheme 4 is ρ-additive with vanishing error probability as $N \to \infty$. Furthermore,

$$C(\widetilde{\text{SDR}}_{\text{out}}) = C(\widetilde{\text{SDR}}_{\text{in}}) = \rho C\left(\frac{\text{SNR}}{\rho}\right), \tag{27}$$

where $\widetilde{\text{SDR}}_{\text{out}}$ and $\widetilde{\text{SDR}}_{\text{in}}$ are the CUBE SDRs of the scheme for both source parts given that no error event occurred, and SNR is the channel SNR.

Proof: Since the worst noise distribution subject to a given power constraint is Gaussian and due to the separation principle (see, e.g., [18] for both), the following CUBE SDR is achievable (assuming no channel decoding errors were made) for the first $(1-\rho)N$ samples, with vanishing decoding error probability:

$$(1-\rho)C\left(\widetilde{\text{SDR}}_{\text{out}}\right) = \rho C\left(\text{SNR}_{\text{digital}}\right), \tag{28}$$

where

$$\mathrm{SNR}_{\mathrm{digital}} \triangleq \frac{(1-\alpha)^{P_X}/\rho}{\alpha\, {}^{P_X}/\rho + P_{\mathrm{BC}}} = \frac{(1-\alpha)^{\mathrm{SNR}}/\rho}{\alpha\, {}^{\mathrm{SNR}}/\rho + 1}\,.$$

Assuming no decoding error of the channel code was made, subtracting $\mathbf{X}_{\mathrm{digital}}$ of $\mathbf{Y}$ and multiplying the result by $\sqrt{\frac{\sigma_S^2}{\alpha^{P_X}/\rho}}$, results the unbiased estimation of $\mathbf{S}_{\mathrm{in}}$, namely

$$\hat{\mathbf{S}}_{\mathrm{in}} = \mathbf{S}_{\mathrm{in}} + \sqrt{\frac{\sigma_S^2}{\alpha^{P_X}/\rho}}\mathbf{Z}\,.$$

Thus, the CUBE SDR (assuming no decoding errors were made) of the remaining ρN samples satisfies:

$$C(\widetilde{\mathrm{SDR}}_{\mathrm{in}}) = C\left(\alpha\,\frac{\mathrm{SNR}}{\rho}\right)\,. \tag{29}$$

The ρ-additivity of the scheme follows since $\mathbf{S}_{\mathrm{in}}$ is effectively transmitted in an analog manner over an AWGN channel (again, assuming no errors were made). Finally, by substituting α (26) in (28) and (29), (27) follows. Note that $\rho \leq 1$ impies

$$1 \leq \left(1+\frac{\mathrm{SNR}}{\rho}\right)^{\rho} \leq 1+\frac{\mathrm{SNR}}{\rho}\,,$$

which is in turn equivalent to $0 \leq \alpha \leq 1$. ■

In order to complete the proof of Theorem 3 for $\rho < 1$, we note that we can compose $\mathbf{S}$ of two codebooks, one which gives the 'out' samples and another — for the 'in' ones. By applying Corollary 1 to each codebook separately, we have that rates approaching

$$(1-\rho)C(\mathrm{SDR}_{\mathrm{out}}) + \rho C(\mathrm{SDR}_{\mathrm{in}})$$

are achievable. Substituting the results of Lemma 3 and Proposition 3, we arrive at (18) as desired.

VI. Strategies for the Parallel Relay Network: Performance Comparison

In this section we consider the rate expression R_{RF} of Theorem 3. We start by presenting a simple upper bound, as well as lower bounds by previously known strategies, on the network capacity. In Section VI-D we compare these bounds, concentrating on limiting cases and showing the asymptotic optimality of the RF strategy. Finally in Section VI-E we discuss variations on the RF strategy, among them time-sharing which allows us to improve the best known performance even for the equal-BW case.

A. Upper Bound on the Network Capacity

The network capacity is upper-bounded by cut-set bounds over both sections, as follows.

Proposition 4: The network capacity C is upper-bounded by R_{outer} where

$$R_{\text{outer}} = \min\Big\{\rho C\left(\frac{M \cdot \text{SNR}_{\text{BC}}}{\rho}\right), C\left(M \cdot \text{SNR}_{\text{MAC}}\right)\Big\} .$$

Proof: We prove that each of the two terms in the minimization is an upper bound on the capacity. The first term follows by a cut-set bound on the BC section, normalizing the rate by the BW ratio ρ (recall that the MAC BW is chosen to be 1). For the second — we use a cut-set bound on the MAC section: Consider the AWGN between $\sum_{m=1}^{M} X_m$ and Y_{MAC}. For a given power, the mutual information is maximized by a Gaussian input, and the maximal input power $M^2 P_X$ is achieved when all the relays transmit the same signal. ■

B. Universality and Asymptotic Optimality of R&F

This outer bound allows us to show that the R&F strategy is universal in the following way: for all the networks that possess the same single-relay capacity over the two sections, i.e., the same $\overline{\text{SNR}}_{\text{BC}}$ and SNR_{MAC}, there exists a uniform bound on the gap from optimality.

Theorem 4: For any parallel relay network with SNRs SNR_{BC} and SNR_{MAC}, number of relays M and BW expansion ratio ρ:

$$R_{\text{outer}} - R_{\text{RF}} \leq \Delta(\overline{\text{SNR}}_{\text{BC}}, \text{SNR}_{\text{MAC}}) < \infty .$$

For $\rho > 1$, this holds provided that $\text{SNR}_{\text{BC}} > \rho$.

Proof: For $\rho > 1$:

$$\begin{aligned} R_{\text{outer}} - R_{\text{RF}} &\leq C\left(M \cdot \text{SNR}_{\text{MAC}}\right) - C\Big(M(\text{SNR}'_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}})\Big) \\ &\leq \frac{1}{2}\log\left(\frac{\text{SNR}_{\text{MAC}}}{\text{SNR}'_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}}}\right) \\ &\leq C\left(2 \cdot \frac{\text{SNR}_{\text{MAC}} + 1}{\overline{\text{SNR}}_{\text{BC}} - 1}\right) \quad , \end{aligned}$$

where the last transition is justified by the assumption $\text{SNR}_{\text{BC}} > \rho$. On the other hand, for $\rho < 1$:

$$\begin{aligned} R_{\text{outer}} - R_{\text{RF}} &\leq \rho C\left(\frac{M \cdot \text{SNR}_{\text{BC}}}{\rho}\right) - \rho C\Big(M(\overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}})\Big) - (1-\rho) C\left(\overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}}\right) \\ &\leq \frac{1}{2}\log\left(\frac{(\text{SNR}_{\text{BC}}/\rho)^{\rho}}{\overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}}}\right) \\ &\leq C\left(\frac{\overline{\text{SNR}}_{\text{BC}}}{\overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}}}\right) + C\left(\frac{1}{\overline{\text{SNR}}_{\text{BC}}}\right) \end{aligned}$$

■

As a direct corollary from the bounds derived in the proof, we have the following asymptotic optimality result:

Corollary 2: In the parallel relay network, for any fixed M and ρ, the following limits are taken by varying SNR_{BC} and SNR_{MAC}. For $\rho > 1$:

$$\overline{\lim}_{\frac{\text{SNR}_{\text{BC}}}{\text{SNR}_{\text{MAC}}} \to \infty} \lim_{\text{SNR}_{\text{MAC}} \to \infty} [R_{\text{outer}} - R_{\text{RF}}] = 0\,.$$

For $\rho < 1$:

$$\overline{\lim}_{\frac{\text{SNR}_{\text{MAC}}}{\text{SNR}_{\text{BC}}} \to \infty} \overline{\lim}_{\text{SNR}_{\text{BC}} \to \infty} [R_{\text{outer}} - R_{\text{RF}}] = 0\,.$$

We see, then, that the R&F strategy is optimal in the limit of high SNR and when the single-relay capacity of the narrower section is much lower than that of the wider section.

C. *Other Relaying Strategies*

In comparison, we present lower bounds on the network capacity given by well known strategies for relaying (see e.g. [1]) applied to this network. We will see that all of these strategies fail to achieve universality, either with respect to the number of relays or to the bandwidth expansion ratio.

Amplify-and-Forward: For the equal-BW case ($\rho = 1$), this strategy was described in Section IV, and the achievable rate is given by (13). It is not obvious what should be considered the extension to $\rho \neq 1$. A natural choice, is to restrict the relays to any linear operation. This means that the excess BW cannot be exploited, and we may as well work with a codebook which has the lower BW of the two sections. The resulting performance is given by

$$R_{\text{AF}} = \begin{cases} C\Big(M \cdot (\text{SNR}_{\text{BC}} \upharpoonleft\upharpoonright \text{SNR}_{\text{MAC}})\Big) & \rho > 1 \\ \rho C\Big(M \cdot \left(\frac{\text{SNR}_{\text{BC}}}{\rho} \upharpoonleft\upharpoonright \frac{\text{SNR}_{\text{MAC}}}{\rho}\right)\Big) & \rho < 1 \end{cases} . \tag{30}$$

This has an unbounded gap from the outer bounds, when either $\rho \to 0$ or $\rho \to \infty$.

Decode-and-Forward: In this strategy, we use a low enough rate such that each relay can reliably decode the codeword. In the second stage all the relays use the same transmission, enjoying the MAC coherence gain. Consequently:

$$R_{\text{DF}} = \min \left\{ \rho C\left(\frac{\text{SNR}_{\text{BC}}}{\rho}\right), C\left(M \text{SNR}_{\text{MAC}}\right) \right\} . \tag{31}$$

The lack of coherence gain with respect to the BC noises causes an unbounded loss as the number of relays M grows, when the first term in the minimization is the limiting one.

Compress-and-Forward: The relays digitally compress their estimations of the codeword, and subsequently send the digital data over the MAC section. The performance is given by comparing the minimal rate of the symmetric quadratic Gaussian CEO problem [19], [20] with the Gaussian MAC capacity, see e.g. [18]. This combination is suboptimal, since by using source–channel separation it fails to achieve the coherence gain (see e.g. [5], [21]). Using this strategy, the achievable rate is $R_{\text{CF}} = \rho C(\overline{\text{SDR}}_{\text{CF}}/\rho)$, where $\overline{\text{SDR}}_{\text{CF}}$ is given by the solution of:

$$C\left(\frac{\overline{\text{SDR}}_{\text{CF}}}{\rho}\right) - MC\left(-\frac{\overline{\text{SDR}}_{\text{CF}}}{M \cdot \text{SNR}_{\text{BC}}}\right) = \frac{1}{\rho}C(\text{SNR}_{\text{MAC}}) \quad , \tag{32}$$

as long as the equation has a positive solution. This rate approaches a finite limit as $M \to \infty$, since the CEO rate has a finite limit, found in [22]. Consequently, we can bound $\overline{\text{SDR}}_{\text{CF}}$ for any number of relays M by the solution of:

$$C\left(\frac{\overline{\text{SDR}}_{\text{CF}}}{\rho}\right) + \frac{\overline{\text{SDR}}_{\text{CF}}}{2 \cdot \text{SNR}_{\text{BC}}} = \frac{1}{\rho}C(\text{SNR}_{\text{MAC}}) \,. \tag{33}$$

Simple algebraic manipulation shows that such $\overline{\text{SDR}}_{\text{CF}}$ must satisfy:

$$\overline{\text{SDR}}_{\text{CF}} < \rho(1 + \text{SNR}_{\text{MAC}})^{\frac{1}{\rho}} \tag{34a}$$

$$\overline{\text{SDR}}_{\text{CF}} < \frac{\log(1 + \text{SNR}_{\text{MAC}}) \cdot \text{SNR}_{\text{BC}}}{\rho} \,. \tag{34b}$$

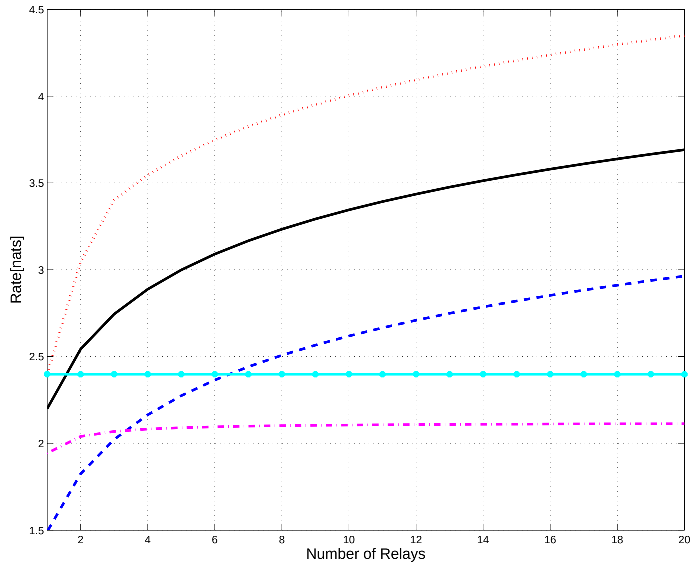
The first inequality shows, that there is no coherence gain w.r.t. the MAC noise. According to the second one, the MAC SNR needs to be very high, if the scheme is to achieve a gain w.r.t. the BC SNR. This causes an unbounded loss as the number of relays grows, when $\rho > 1$.

The graphs of Figure 8 show the different bounds for two cases, as a function of M: in the first case the BC section is wider but of lower SNR, and in the second case the roles change. It is evident, that R&F achieves the coherence gain of A&F for any M, resulting in a similar behavior as a function of the number of relays, but with a higher rate.

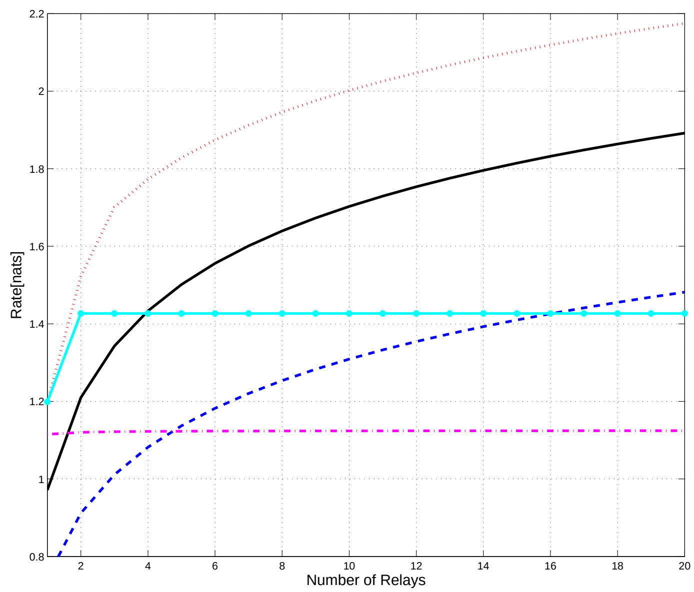
D. Performance Comparison

In order to compare the various bounds for $\rho \neq 1$, we consider the high SNR limit $\overline{\text{SNR}}_{\text{BC}} \gg 1$, $\text{SNR}_{\text{MAC}} \gg 1$. Within this limit, we further consider four limiting cases, in order of rising SNR_{MAC}:

1) Decodable: $\overline{\text{SNR}}_{\text{BC}} \gg M \cdot \text{SNR}_{\text{MAC}}$.
2) MAC-limited: $M\overline{\text{SNR}}_{\text{BC}} \gg M \cdot \text{SNR}_{\text{MAC}} \gg \overline{\text{SNR}}_{\text{BC}}$.
3) BC-limited: $\exp\{\rho \cdot M\} > \text{SNR}_{\text{MAC}} \gg \overline{\text{SNR}}_{\text{BC}}$.
4) Recoverable: $\text{SNR}_{\text{MAC}} > \exp\{\rho \cdot M\} \gg \overline{\text{SNR}}_{\text{BC}}$.



(a) $\rho = 2, \mathrm{SNR}_{\mathrm{BC}} = 20, \mathrm{SNR}_{\mathrm{MAC}} = 300$



(b) $\rho = 0.5, \mathrm{SNR}_{\mathrm{BC}} = 150, \mathrm{SNR}_{\mathrm{MAC}} = 10$

Fig. 8: Rate vs. number of relays. solid = RF, dashed = AF, dash-dotted = CF, starred = DF, dotted = outer bound.

For these cases, the effective SNRs (i.e., the ones satisfying $C(\mathrm{SNR}) = R$ where R is the corresponding rate) according to the different strategies are summarized in the following table, along with the outer bounds of Proposition 4. The expression for C&F in the BC-limited case is according to the bounds (34), so the C&F performance for finite M is worse.

<table>
<tr><th>Bound</th><th>Decodable</th><th>MAC-limited</th><th>BC-limited</th><th>Recoverable</th></tr>
<tr><td>AF</td><td colspan="2">$(M \cdot \text{SNR}_{\text{MAC}}/r)^r$</td><td colspan="2">$(M \cdot \text{SNR}_{\text{BC}}/r)^r$</td></tr>
<tr><td>RF</td><td rowspan="2">$M \cdot \text{SNR}_{\text{MAC}}$</td><td>$\left(M \cdot \text{SNR}_{\text{MAC}} \left(\frac{\text{SNR}_{\text{BC}}}{\rho}\right)^{1-r}\right)^r$</td><td colspan="2">$M^r \left(\frac{\text{SNR}_{\text{BC}}}{\rho}\right)^\rho$</td></tr>
<tr><td>DF</td><td colspan="3">$\left(\frac{\text{SNR}_{\text{BC}}}{\rho}\right)^\rho$</td></tr>
<tr><td>CF</td><td colspan="2">SNR_{MAC}</td><td>$\left(\log(\text{SNR}_{\text{MAC}}^{\frac{1}{\rho}}) \cdot \frac{\text{SNR}_{\text{BC}}}{\rho}\right)^\rho$</td><td>$\left(\frac{M \cdot \text{SNR}_{\text{BC}}}{\rho}\right)^\rho$</td></tr>
<tr><td>outer</td><td colspan="2">$M \cdot \text{SNR}_{\text{MAC}}$</td><td colspan="2">$\left(\frac{M \cdot \text{SNR}_{\text{BC}}}{\rho}\right)^\rho$</td></tr>
</table>

Fig. 9: Comparison of bounds on the effective SNR. $r = \min(\rho, 1)$.

The first and last limits correspond to extremes, where either the SNR of the BC section or the SNR of the MAC section are so high, that no coherence gain is required for that section. In these limits, D&F and C&F are optimal, respectively. The intermediate limits are more interesting from a practical point of view, since the capacities of both sections, given by the cut-set bounds of Proposition 4, are closer to each other. In these limits, it is evident that even A&F is better than both C&F and D&F for large enough M, since these digital methods do not posses the relevant coherence gains. R&F has the same coherence gain as A&F, $M^{\min(1,\rho)}$, but makes better use of the BW, resulting in (generally) better SNR-dependent factors. We also note that in the MAC-limited case where $\rho > 1$, and in the BC-limited case where $\rho < 1$, the asymptotic performance of R&F approaches the outer bound; this corresponds to the optimality claim of Corollary 2.

E. Improvements of R&F

The R&F scheme, as presented here, is not necessarily optimal beyond the asymptotic sense. In fact, we can point out some possible improvements.

- **Global optimization of the estimation:** Consider the R&F scheme when using the HDA JSCC approach of Mittal and Phamdo over the BC section as presented in Section V. We may view the operation of this specific HDA scheme as decomposing the chosen channel codeword (seen as a source) into two JSSC codewords, so that their sum is equal to the codeword: the quantized value is a "digital" word, while the quantization error is an "analog" one. At the relays, the digital codeword is decoded, while the analog one is left with the corresponding channel noise. Alternatively, one may start with a *channel* codebook obtained as a superposition (sum) of two codebooks, which are processed as the digital and analog codebooks above, i.e., only one is decoded. Hence, one codebook is relayed using a D&F approach, while the other — using A&F; these codebooks are superimposed over the narrower section, while over the wider section they occupy separate bands. From this perspective, one is left with the task of determining the power allocation when superimposing the analog and digital layers. Following the R&F approach, these weights are taken so as to produce the optimal reconstruction of of the original codeword (the sum of the two). However, one may ask whether a different power allocation may result in a better performance at the *destination* node. Indeed, as recently proposed by Saeed et al. [9], performance may be enhanced by optimization of the power allocation.

- **Non-white transmission:** While white transmission is optimal over a single white section, there is no guarantee that it is also optimal for the parallel relay network. In fact, if transmission over one band enjoys a coherence gain while transmission over another band does not, it is plausible to increase the transmission PSD in the band that does.

Common to both points above, is that *global* (power) optimization is needed, i.e., the transmission scheme over one section depends upon the parameters of the whole network. Since the simpler local optimization approach is sufficient for understanding the basic gains and tradeoffs in the network, and on the other hand no optimality claim can be made about the scheme even after global optimization, we do not pursue these directions in this work. However, we do point out the potential benefit in combining R&F with D&F and A&F by means of time-sharing, as suggested For A&F and D&F in [23] for the equal-BW case. The sharing strategy allocates different power to the R&F and D&F relay transmissions, such that they effectively function with different SNRs, satisfying the total power constraint. The following specifies rates which are achievable by time-sharing.

Theorem 5: (**R&F–D&F Time-Sharing**) The capacity of the parallel relay network satisfies:

$$C \geq \max\{\lambda R_{\text{RF}} + (1-\lambda) R_{\text{DF}}\} ,$$

where R_{RF} is given by Theorem 3 for signal-to-noise ratios $\text{SNR}_{\text{RF,BC}}$ and $\text{SNR}_{\text{RF,MAC}}$ and BW expansion factor $\tilde{\rho}$; R_{DF} is given by (31) for signal-to-noise ratios $\text{SNR}_{\text{DF,BC}}$ and $\text{SNR}_{\text{DF,MAC}}$ and for BW expansion factor $\frac{\rho-\tilde{\rho}\lambda}{1-\lambda}$. The maximization is performed over all $\tilde{\rho}$, λ, $\text{SNR}_{\text{RF,BC}}$, $\text{SNR}_{\text{RF,MAC}}$, $\text{SNR}_{\text{DF,BC}}$, $\text{SNR}_{\text{DF,MAC}}$

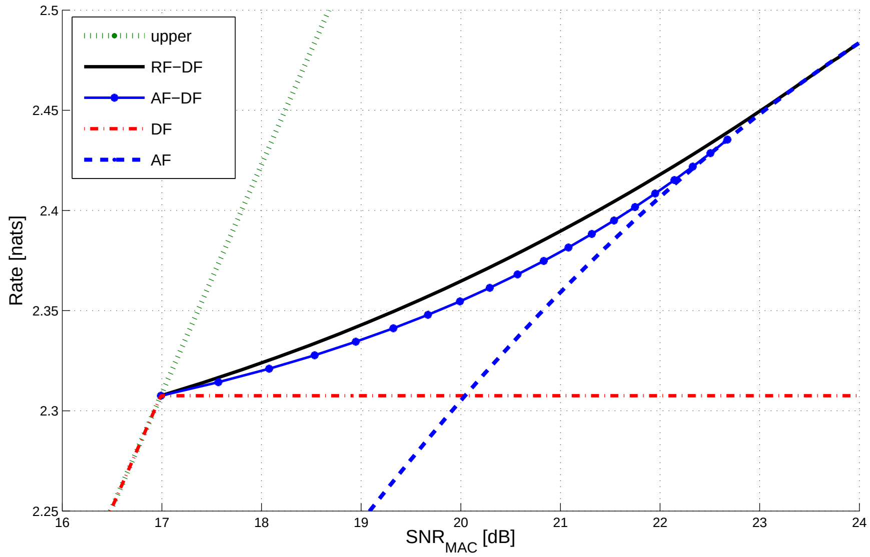


Fig. 10: Time-sharing for the equal-BW case. $M = 2$, $\mathrm{SNR}_{\mathrm{BC}} = 100$.

which satisfy:

$$0 \leq \lambda \leq 1 \tag{35a}$$

$$0 \leq \tilde{\rho}\lambda \leq \rho \tag{35b}$$

$$\mathrm{SNR}_{\mathrm{RF,BC}} + \mathrm{SNR}_{\mathrm{DF,BC}} = \mathrm{SNR}_{\mathrm{BC}} \tag{35c}$$

$$\mathrm{SNR}_{\mathrm{RF,MAC}} + \mathrm{SNR}_{\mathrm{DF,MAC}} = \mathrm{SNR}_{\mathrm{MAC}}\,. \tag{35d}$$

The proof is straightforward: This rate can be achieved by constructing two AWGN codebooks where the first codebook is transmitted on the first $\lambda\tilde{\rho}N$ BC section uses and λN MAC section uses, by applying the R&F strategy, whereas in the remaining $(\rho - \lambda\tilde{\rho})N$ BC section uses and λN MAC section uses, D&F is performed. In each section, we allocate a different power to each transmission scheme, resulting in a different SNR, such that the average power constraints are met. A rather surprising result is, that when substituting the case $\rho = 1$ in the theorem, one gets a slightly improved performance over the time-sharing between A&F and D&F. This happens since the replacement of A&F by R&F gives another degree of freedom in the design: The BW allocated to R&F may change between the transmitter and the relays. In other words, using R&F we may introduce an artificial BW change to the equal-BW case. Figure 10 demonstrates this improvement for two relays and $\mathrm{SNR}_{\mathrm{BC}} = 20dB$. While there is no reason to believe that the new achievable rates are the optimum, this result demonstrates that the known inner bounds of [23] can be improved.

In some other cases, time sharing between R&F and A&F may be plausible. The expressions can be derived in a similar manner, keeping in mind that A&F works with equal BW in both sections.

VII. Colored Noises and Time-Domain Processing

We now abandon the assumption that the noises are white. We keep restricting our analysis to the symmetric setting, thus we assume that all the BC noises have identical spectra. We denote this spectrum and the MAC noise spectrum by $S_{\text{BC}}(\mathrm{e}^{j2\pi f})$ and $S_{\text{MAC}}(\mathrm{e}^{j2\pi f})$, respectively. Without loss of generality, assume that the noise spectra of both sections are monotonically increasing as a function of $|f|$. [3] The BW of a section is defined as the maximal frequency in which the noise spectrum is finite. In practice, these frequencies arise from the sampling frequencies used, thus they are always finite.

We assume equal sampling rate at both sections, which is taken to correspond to the maximum of the two BW; consequently, both the source encoder and the relays use transmission blocks of equal length N. We then define the BW of both sections:

$$B_{\text{BC}} = 2\sup\{f : S_{\text{BC}}(\mathrm{e}^{j2\pi f}) < \infty\} \tag{36a}$$

$$B_{\text{MAC}} = 2\sup\{f : S_{\text{MAC}}(\mathrm{e}^{j2\pi f}) < \infty\} \, . \tag{36b}$$

We denote by ρ the BW ratio:

$$\rho = \frac{B_{\text{BC}}}{B_{\text{MAC}}} \, .$$

By definition, if $\rho > 1$ or $\rho \leq 1$ then $B_{\text{BC}} = 1$ or $B_{\text{MAC}} = 1$, respectively. Note, that for $\rho > 1$ the time units used differ from those used thus far in the paper.

Under this notation, we define the SNRs of the BC and MAC sections w.r.t. the total inband noise power:

$$\begin{aligned} \text{SNR}_{\text{BC}} &\triangleq \frac{P_X}{\int_{2|f|\leq B_{\text{BC}}} S_{\text{BC}}\left(\mathrm{e}^{j2\pi f}\right) df} \\ \text{SNR}_{\text{MAC}} &\triangleq \frac{MP_X}{\int_{2|f|\leq B_{\text{MAC}}} S_{\text{MAC}}\left(\mathrm{e}^{j2\pi f}\right) df} \, . \end{aligned} \tag{37}$$

Finally, we denote by $C(P, S(\mathrm{e}^{j2\pi f}))$ the capacity of an additive Gaussian-noise channel with power input constraint P and noise spectrum $S(\mathrm{e}^{j2\pi f})$, given by the water-filling solution, see e.g. [18].

[3] This is done for convenience of the definition of bandwidth only. Since any node may perform a "frequency-bands swapping" operation, it is not restrictive.

The cut-set outer bounds of Proposition 4 can be easily extended to the colored case, as follows.

Proposition 5: The network capacity C is upper-bounded by R_{outer} where

$$R_{\text{outer}} = \min\Big\{C\left(MP_X, S_{\text{BC}}(\mathrm{e}^{j2\pi f})\right), C\left(M^2 P_X, S_{\text{MAC}}(\mathrm{e}^{j2\pi f})\right)\Big\}.$$

The D&F rate is also easily computed to be:

$$R_{\text{DF}} = \min\Big\{C\left(P_X, S_{\text{BC}}(\mathrm{e}^{j2\pi f})\right), C\left(M^2 P_X, S_{\text{MAC}}(\mathrm{e}^{j2\pi f})\right)\Big\} \quad . \tag{38}$$

We see that, as in the white case, D&F is optimal when the performance is limited by the MAC section, but fails to achieve the coherence gain w.r.t. the BC section noise. Clearly, A&F enjoys both coherence gains even in the colored setting, so for a large enough number of relays M it outperforms digital approaches. However, even for $\rho = 1$, and even if we allow *any* linear relaying function, generally it cannot exploit the full capacity offered by links with colored noise; see e.g. [24] for a discussion of this issue in the point-to-point setting.

As in the white BW-mismatch case, the R&F strategy aims to make better use of the individual link capacities than A&F does, while maintaining the coherence gains. We use the same scheme of Figure 5, adjusting the JSCC and channel encoder/decoder pairs to the colored setting. We now state an achievable rate using this strategy in terms of $C_{\text{white}}(P, S(\mathrm{e}^{j2\pi f}))$, the mutual information over an additive channel with a noise spectrum $S(\mathrm{e}^{j2\pi f})$ of BW B, using a white input of the same BW:

$$C_{\text{white}}(P, S(\mathrm{e}^{j2\pi f})) \triangleq \frac{1}{2}\int_{2|f|<B} \log\left(1 + \frac{P}{BS(\mathrm{e}^{j2\pi f})}\right) \tag{39}$$

and of $\Gamma(S(\mathrm{e}^{j2\pi f}))$, the prediction gain of a spectrum $S(\mathrm{e}^{j2\pi f})$ of BW B: (see e.g. [25]):

$$\Gamma\Big(S(\mathrm{e}^{j2\pi f})\Big) \triangleq \frac{\frac{1}{B}\int_{2|f|<B} S_S(\mathrm{e}^{j2\pi f})df}{\exp\frac{1}{B}\int_{2|f|<B}\log\Big(S_S(\mathrm{e}^{j2\pi f})\Big)df} \quad . \tag{40}$$

Theorem 6: (**R&F Performance for colored noise**) The capacity of the Gaussian parallel relay network with colored noises. is lower-bounded by $R_{\text{RF}} \triangleq C_{\text{white}}\left(P_X, S_{\text{RF}}(\mathrm{e}^{j2\pi f})\right)$, where

$$S_{\text{RF}}(\mathrm{e}^{j2\pi f}) = \left\{\begin{array}{cc} M\left(\overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \frac{MP_X}{S_{\text{MAC}}(\mathrm{e}^{j2\pi f})}\right), & 2|f| \le \min(B_{\text{BC}}, B_{\text{MAC}}) \\ \overline{\text{SNR}}_{\text{BC}} \upharpoonleft\upharpoonright \frac{M^2 P_X}{S_{\text{MAC}}(\mathrm{e}^{j2\pi f})}, & \text{otherwise} \end{array}\right\},$$

with the equivalent BC SNR being:

$$\overline{\text{SNR}}_{\text{BC}} = \left[\Gamma\Big(S_{\text{BC}}(\mathrm{e}^{j2\pi f})\Big)\text{SNR}_{\text{BC}}\right]^{\rho} - 1, \tag{41}$$

as long as $\overline{\text{SNR}}_{\text{BC}} > 0$.

Before proceeding to prove the theorem, we analyze this rate expression. While we show in the sequel that further adjustments of the R&F scheme may improve the achievability result, this closed-form expression already possesses the main gains achievable by the scheme. As in the white case of Theorem 3, it shows coherence gain w.r.t. both sections inside the BW of the BC section, and w.r.t. the MAC section for all the codebook BW. It also exploits most of the gain offered by the noise color, in the high-SNR limit where a white channel input is asymptotically optimal. Taking $\rho = 1$ (no BW mismatch), the following result can be proven by straightforward calculations.

Proposition 6: (**Near-optimality of R&F for colored channels**) Let $\mathrm{SDR}_{\mathrm{RF}}$, $\mathrm{SDR}_{\mathrm{DF}}$ and $\mathrm{SDR}_{\mathrm{outer}}$ be the SNRs corresponding to R_{RF}, R_{DF} and R_{outer} of Theorem 6, (38) and Proposition 5, respectively. For $\rho = 1$, consider the high-SNR bound where all the parameters are held fixed except for P_X. Then:

$$\lim_{P_X \to \infty} \frac{\mathrm{SDR}_{\mathrm{outer}}}{\mathrm{SDR}_{\mathrm{RF}}} \leq 2$$
$$\lim_{P_X \to \infty} \frac{\mathrm{SDR}_{\mathrm{DF}}}{\mathrm{SDR}_{\mathrm{RF}}} \leq 1 + \frac{1}{M} \; . \tag{42}$$

The proof of Theorem 6 relies upon the existence of a JSCC scheme which exploits the full mutual information available over a link (asymptotically in high SNR), while being additive over the minimum between the source and channel BW. Then, Corollary 1 is invoked, as in the proof of Theorem 3. In order to use additivity in the colored setting, we need to extend the definition of ρ-additivity as follows.

Definition 4: (ρ-**Additive JSCC scheme — generalized**) A JSCC scheme is *ρ-additive with error probability* p_e, if for some unitary transformation $\mathbf{S}'$ of the source $\mathbf{S}$, the first ρN elements of the CUBE error $\tilde{\mathbf{Z}}$ of $\mathbf{S}'$ are additive with error probability p_e, in the sense of Definition 2.

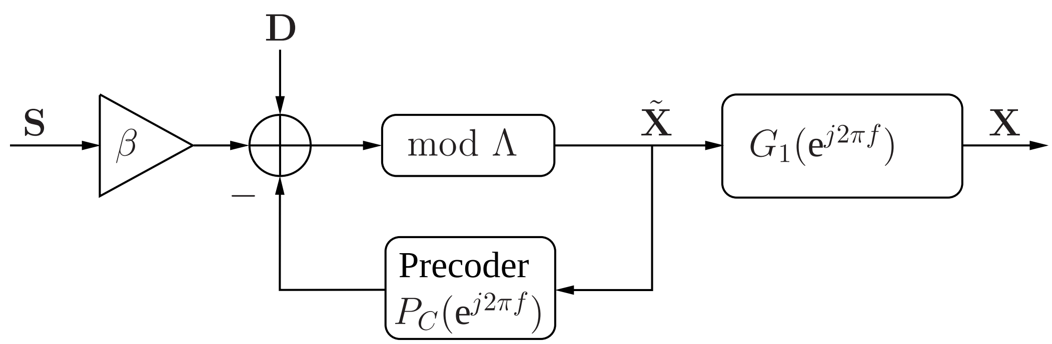
W.l.o.g. we assume that the unitary transformation, under which the JSCC scheme is ρ-additive, is the DFT, since otherwise one could concatenate the DFT with another unitary transformation. Parallel to the exposition in Section V-B, we divide the source signal $\mathbf{S}$ of length N into an "in" part filtered to BW $r = \min(1, \rho)$ and an "out" part with the remaining signal.

Proposition 7: For any white Gaussian source of BW B_S and for any additive Gaussian channel with power constraint P_X and noise spectrum $S(\mathrm{e}^{j2\pi f})$, let B_C be the channel BW. Define $B \triangleq \min(B_S, B_C)$, and

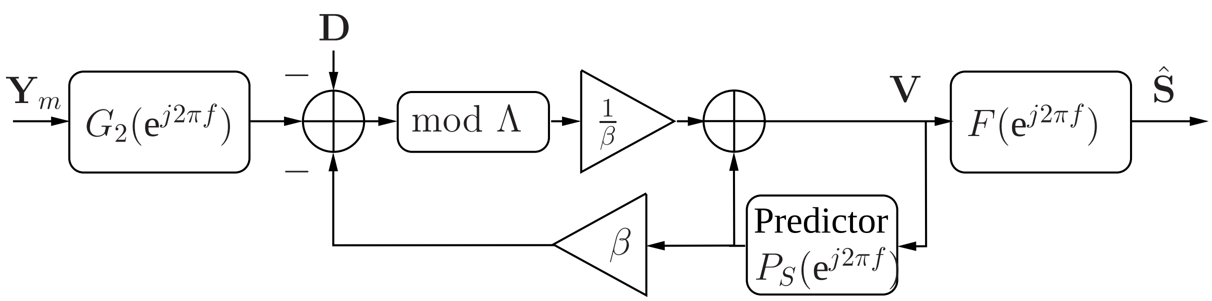
$$\mathrm{SNR} \triangleq \frac{P_X}{\int_{2|f| \leq B_C} S\left(\mathrm{e}^{j2\pi f}\right) df} .$$

Then there exists a B-additive JSCC scheme with vanishing error probability as $N \to \infty$ that satisfies:

$$\overline{\mathrm{SDR}}_{\mathrm{in}} = \overline{\mathrm{SDR}}_{\mathrm{out}} = \left[\Gamma\left(S(\mathrm{e}^{j2\pi f}) \right) \mathrm{SNR} \right]^{\frac{B_C}{B_S}} - 1 \; ,$$



(a) Encoder



(b) Decoder

Fig. 11: Analog Matching encoder and decoder for R&F

where $\overline{\mathrm{SDR}}_{\mathrm{in}}$ and $\overline{\mathrm{SDR}}_{\mathrm{out}}$ are the CUBE SDRs of the scheme for both source parts given that no error event occurred. Furthermore, given that no error event occurred, the samples of the CUBE error are mutually independent.

We prove this using a time-domain approach, based upon the Analog Matching (AM) scheme of [25]; in the sequel we discuss an alternative, namely the application of HDA schemes in the spirit of the exposition in Section V. In our context, the AM scheme consists of predictors (either at the encoder side, the decoder side or both, according to the BW expansion factor and the noise spectrum), modulo-lattice operations at both sides, and a linear filter pair, one at each side, taken to be ideal low-pass filters (LPFs), which may become redundant depending on the BW expansion factor. For simplicity, our choice of filters reflects a "zero-forcing" approach (targeting the high-SNR regime); even when using optimized filters, there is an inherent loss in using the Analog Matching scheme, which vanishes at the limit of high SNR.

The encoder and the decoder are depicted in Figure 11. We use the notation of Section VII; under this notation, either the source signal $\mathbf{S}$ is white, or it is flat inside the MAC BW and zero outside. Note that besides the components in Figure 11, the scheme must employ interleaving; see [25]. We assume that the lattice has a second moment equal to the channel power constraint P_X.

Proof: Following the considerations made in [25], as long as a correct decoding condition holds, we have at each time instant n (see Figure 11):

$$V_n - S_n = \frac{1}{\beta} E_n * (\delta_n - p_{Cn}) \ ,$$

where

$$E_n = [g_{1n} * g_{2n} - \delta_n] * \tilde{X}_n + g_{2n} * Z_n \ ,$$

and where g_{1n}, g_{2n} and p_{Cn} are the time responses of the filters $G_1(\mathrm{e}^{j2\pi f})$, $G_2(\mathrm{e}^{j2\pi f})$ and $P_C(\mathrm{e}^{j2\pi f})$, respectively, δ_n is the unit impulse function, $\tilde{X}_n$ and Z_n are the encoder modulo-lattice output and channel noise, respectively, and $*$ denotes convolution. If the dither $\mathbf{D}$ is independent of the source and uniformly distributed over the basic cell of the lattice Λ then $\tilde{\mathbf{X}}$ has power P_X and is statistically independent of $\mathbf{X}$. As a consequence, we have that the channel power constraint is satisfied, and also that $\mathbf{E}$ is independent of $\mathbf{S}$.

Denote the spectrum of E_n by $S_E(\mathrm{e}^{j2\pi f})$. If we choose $P_C(\mathrm{e}^{j2\pi f})$ to be the optimal predictor of that spectrum, the resulting prediction error $V_n - S_n$ is white with variance

$$\tilde{D} = \frac{\exp \int_{2|f|<1} \log S_E(\mathrm{e}^{j2\pi f})}{\beta^2} \ .$$

Now we take $G_1(\mathrm{e}^{j2\pi f})$ and $G_2(\mathrm{e}^{j2\pi f})$ to be LPFs of height $\sqrt{\frac{1}{B_C}}$ and $\sqrt{B_C}$, respectively, and width B_C, to assert that

$$S_E(\mathrm{e}^{j2\pi f}) = \left\{ \begin{array}{cc} B_C S_Z(\mathrm{e}^{j2\pi f}), & 2|f| < B_C \\ P_X, & \text{otherwise} \end{array} \right\} \ .$$

Thus:

$$\beta^2 \tilde{D} = B_C^{B_C} P^{1-B_C} \exp \int_{2|f|<B_C} \log S_Z(\mathrm{e}^{j2\pi f}) df = \frac{P_X}{(\Gamma(S(\mathrm{e}^{j2\pi f}))\mathrm{SNR})^{B_C}} \ .$$

Following [25] again, if an optimal $P_S(\mathrm{e}^{j2\pi f})$ is used, then for large enough blocks correct decoding holds with arbitrarily small error probability as long as the power of T_n is at most P_X, i.e.,

$$\beta^2 \left(\frac{P}{B_S} + \tilde{D} \right)^{B_S} \tilde{D}^{1-B_S} < P \ .$$

Algebraic manipulation shows that any $\tilde{D}$ satisfying:

$$\frac{P}{B_S \tilde{D}} < \left(\Gamma(S(\mathrm{e}^{j2\pi f}))\mathrm{SNR} \right)^{\frac{B_C}{B_S}} - 1$$

is achievable. Now if we take $F(\mathrm{e}^{j2\pi f})$ to be an LPF of bandwidth B_S and unit height, we arrive at the desired result since $D = B_S \tilde{D}$. We also have that the scheme is B_C-additive as long as correct decoding holds, since the inband component is a filtered version of the channel noise. ■

Note that, for high SNR, this performance approaches the Shannon upper bound on the channel capacity (see e.g. [26]), thus using the AM scheme is optimal in this limit. For general SNR, it has a loss w.r.t. the optimum performance of the AM scheme presented in [25]. This happens since the equivalent noise of the optimum AM scheme contains a "self noise" component, i.e. a filtered version of the channel input; such a component prohibits the scheme from being additive.[4] For additivity, we had to choose here sub-optimal "zero-forcing" filters inside the channel BW.

Calculating the total distortion of both sections, as done in Section V, shows that the decoder can achieve a CUBE error of spectrum $S_{\text{RF}}(\mathrm{e}^{j2\pi f})$. In order to complete the proof, we need to extend Corollary 1 to colored estimation error; the extension follows directly from substituting the colored Gaussian RDF in Proposition 1.

Corollary 3: Let $\mathbf{S}$ be a Gaussian i.i.d. vector of element power P. Denote the CUBE error PSD by $S(\mathrm{e}^{j2\pi f})$. Then there exists a channel coding scheme of any rate

$$R < C\left(P, S(\mathrm{e}^{j2\pi f})\right)$$

over the channel from $\mathbf{S}$ to $\hat{\mathbf{S}}$.

After establishing Theorem 6, we note that this is not the tightest inner bound on capacity that we can give. We restricted the input to both sections to be white, although the water-filling spectrum of both may be colored. However, taking the input of an individual link to have the water-filling spectrum of that link does not guarantee global optimality either. For example, the water-filling solution over the BC section may have a lower BW than B_{BC}. Outside that water-filling BW, there will be no coherence gain w.r.t. the BC noise. Obviously, for large enough M this has greater effect than that of the water-filling gain. The optimum spectra can be computed by a straightforward, though cumbersome, optimization; given such spectra, a modified scheme adding filters at the relays may be materialized. Choosing a white input, as we did, simplifies on matters. Nevertheless, in the limit of large number of relays or of high SNR, the choice becomes optimal.

In Section V we have thus far proven our achievablity results for the white BW-mismatch case using specific hybrid digital–analog (HDA) schemes. In fact, we could use also for the colored case an additive variant of an HDA scheme proposed by Prabhakaran et al. [7], which splits the source and the channel into frequency bands, and then applies to each band the HDA techniques by Mittal and Phamdo that we suggested to use for the white case in Section V. This approach may have slightly better performance than

[4]The self noise is indeed independent of the source $\mathbf{S}$, but is not independent of it given the encoding function, i.e., given the dither vector $\mathbf{D}$.

the modulo-lattice approach for the parallel relay network (as, when substituting in the analog-matching performance the special case of BW mismatch, performance is worse than the HDA performance), but this advantage vanishes at the limit of high SNR. We chose not to pursue it in this work, since computing the performance of the additive variant turns out to be a cumbersome task, and we believe that insight into exploiting the noise memory is better gained using a time-domain approach.

In addition, note that the HDA approach calls for partitioning the frequency band into multiple bins, and using multiple source and channel codes (of different rates) for these bins. From the practical point of view, this may be a drawback when the frequency response exhibits significant variation, corresponding with many bins. This motivates looking at time-domain approaches; the considerations are reminiscent of those concerning the difference between DMT and FFE–DFE in point-to-point channel coding [27].

VIII. Discussion: Extensions to Layered Networks and to MIMO Channels

We conclude the paper by pointing out how the R&F approach can be used beyond the colored parallel relay network scenario. We first look at more complex networks, and then we turn to replacing the BW mismatch by a mismatch in the number of antennas (degrees of freedom) in a MIMO setting.

A. R&F as a Building Stone for Relayed Networks

Turning our view to more complex networks, the ideas presented in this paper are most easily applied to*layered networks*, which are directed acyclic graphs (DAGs) where the nodes can be divided into layers, and nodes in each layer receive the (noisy) sum of transmissions from the adjacent preceding layer only. We index the relay layers as $l = 1, \cdots, L$, where layer L consists of the destinations. Specifically we consider the symmetric case, where the noise spectra, as well as the number of received transmissions ("fan-in") and the number of destinations in the next layer ("fan-out") are identical at all nodes in the same layer. In the white symmetric case with BW mismatch between layers, each layer is characterized by its SNR SNR_l and BW ρ_l. Figure 12 shows two examples of networks which fall under this category. In the sequel, we show how combining the R&F and C&F strategies is beneficial in the first example, while recursive use of the R&F strategy is the key to the treatment of the second one.

Consider the network of Figure 12a. We use a codebook BW according to layer 4, i.e. $\rho_4 = 1$. From layer 1 to layer 3 there are no MACs, thus analog transmission produces no coherence gain. Hence, the noise accumulation in these layers can be avoided by having each of the nodes in layer 1 compress their estimation of $\mathbf{V}$ according to the rate $\min\big(\rho_2 C(\mathrm{SNR}_2), \rho_3 C(\mathrm{SNR}_3)\big)$ and send it digitally, making sure that the resulting quantization errors are mutually independent (c.f. by using mutually-independent

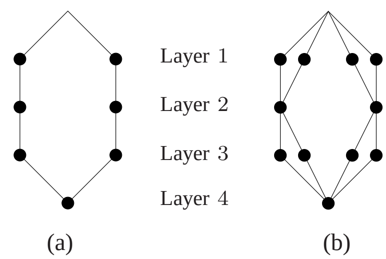


Fig. 12: BW-mismatched symmetric layered networks. Each node denoted by a full circle contains a MAC channel.

dither in each node). For the outer layers, we use the R&F scheme. The decoder "sees" a parallel relay network, where the MAC noise is the sum of the noise of layer 4 and the quantization noises. Thus we achieve the coherence gain w.r.t. all the noises in the network.

Next consider the network of Figure 12b, which contains MACs in both layer 2 and layer 4. On one hand, the coherence gain is only known to be achieved for analog transmission over the MACs, but on the other hand using analog transmission for both does not enable to utilize the full BW if $\rho_2 \neq \rho_4$. This difficulty can be circumvented by using two information-bearing signals of different BW. We use again a codebook BW according to layer 4, applying a JSCC method to re-match it to BW ρ_2. This re-matched codeword is sent using the RF scheme to layer 2, where instead of the codebook decoder, the "source signal" associated with the outer JSCC scheme is reproduced. Note that this way the estimation errors in both relays of layer 2 are mutually independent. Next we use RF again to transmit these reproductions to relay 4. In the overall result, again the full coherence gain is achieved.

B. R&F for MIMO Channels

The BW mismatch framework may be thought of as a model for combining channels with a different number of antennas. For example, it may reflect relays communicating with the end-users using one antenna, while using multiple antennas for the link with the base station. For a recent work regarding parallel relays in the MIMO setting, see [28], which in contrast to the present work, assumes a digital use of the MAC section, leading to a C&F approach.

Note that the applying the schemes developed in this paper to MIMO channels is not straightforward. In a practical scenario we need to abandon the the symmetric assumption and allow each relay to have different channel matrices; but unlike LTI systems, MIMO systems are not diagonalized by the same transform.

APPENDIX A
PROOF OF THEOREM 2

For the sake of simplicity we prove the theorem assuming the use of encoders which produce stationary channel inputs. The general case follows by similar arguments.

Assume w.l.o.g. that $\mathbf{S}$ has per-element variance P. Denote the CUBE estimate at each relay by $\hat{\mathbf{S}}_m = \mathbf{S} + \tilde{\mathbf{Z}}_m$ and define by $\mathrm{SDR}_m \triangleq \frac{P}{E\{\tilde{Z}^2_{m,n}\}}$ the unbiased signal-to-distortion ratio at relay m. Now we have that:

$$C(\mathrm{SDR}_m) \le I(\mathbf{S}; \hat{\mathbf{S}}_m) \le I(\mathbf{X}_{\mathrm{BC}}; \mathbf{Y}_m) \le C(\mathrm{SNR}_{\mathrm{BC}}) \quad ,$$

where the inequalities are justified by the (unbiased) source rate-distortion function (4), the separation principle and the channel capacity, respectively. Consequently, we have that

$$\mathrm{SDR}_m \le \mathrm{SNR}_{\mathrm{BC}} \quad \forall m = 1, \ldots, M \; . \tag{43}$$

We can also consider a joint decoder which observes $\{\mathbf{Y}_m\}$ and obtains an estimate $\hat{\mathbf{S}}_{\mathrm{BC}}$ of $\mathbf{S}$ with CUBE signal-to-distortion ratio $\mathrm{SDR}_{\mathrm{BC}}$. Since this is a point-to-point scenario, we can repeat the above considerations, now replacing the point-to-point capacity by the BC capacity $\frac{1}{2}\log(1 + M \cdot \mathrm{SNR}_{\mathrm{BC}})$, obtaining:

$$\mathrm{SDR}_{\mathrm{BC}} \le M \cdot \mathrm{SNR}_{\mathrm{BC}} \; , \tag{44}$$

where SDR is the CUBE SDR of S. Now we describe a specific joint decoder, which must obey this bound. The decoder first obtains the CUBE estimates at each relay, and then combines them:

$$\hat{S}_{\mathrm{BC}} = \frac{\sum_{m=1}^{M} \gamma_m \hat{S}_m}{\sum_{m=1}^{M} \gamma_m} \; .$$

This is a CUBE estimator, with:

$$\mathrm{SDR}_{\mathrm{BC}} = \frac{(\sum_{m=1}^{M} \gamma_m)^2 P}{E\{(\sum_{m=1}^{M} \gamma_m \tilde{Z}_m)^2\}} \; . \tag{45}$$

Equipped with these bounds, we now turn to the relay functionality in the original setting. Without loss of generality, we assume that the relays transmit:

$$X_m = \gamma_m \hat{S}_m$$

where the gains must satisfy the power constraints:

$$\gamma_m^2 (P + E\{\tilde{Z}_m^2\}) \le P \quad \forall 1 \le m \le M \quad . \tag{46}$$

The optimal CUBE estimate of S from $Y_1, \ldots, Y_M$, where

$$Y_m = X_m + Z_m = \gamma_m \hat{S}_m + Z_{\text{MAC}} = \gamma_m (S_m + \tilde{Z}_m) + Z_{\text{MAC}}$$

is given by maximum-ratio combining (MRC), with performance:

$$\frac{1}{\text{SDR}} = \left\{ \frac{1}{\text{SDR}_{\text{BC}}} + \frac{M}{\left(\sum_{m=1}^{M} \gamma_m\right)^2 \text{SNR}_{\text{MAC}}} \right\} , \tag{47}$$

where SDR_{BC} is given by (45). Any SDR achieved by the scheme is thus according to (47), where the constraints (43)–(46) must be met. Using this, we have:

$$\frac{1}{\text{SDR}} \geq \frac{1}{M \cdot \text{SNR}_{\text{BC}}} + \frac{M}{\left(\sum_{m=1}^{M} \gamma_m\right)^2 \text{SNR}_{\text{MAC}}} .$$

Now we are left with the task of bounding the performance w.r.t. the optimum choice of $\{\gamma_m\}$. This is a convex optimization problem in the parameters $\{\gamma_m^2\}, \{\frac{1}{\text{SDR}_m}\}$ under the $2M$ constraints given by (43), (46); the solution is uniform, with

$$\gamma_m^2 = \frac{\text{SNR}_{\text{BC}}}{1 + \text{SNR}_{\text{BC}}} \; \forall m = 1, \cdots, M .$$

Consequently:

$$\frac{1}{\text{SDR}} \geq \frac{1}{M \cdot \text{SNR}_{\text{BC}}} + \frac{1 + \text{SNR}_{\text{BC}}}{M \cdot \text{SNR}_{\text{BC}} \text{SNR}_{\text{MAC}}} = \frac{1}{M \left(\text{SNR}_{\text{BC}} \upharpoonleft\!\upharpoonright \text{SNR}_{\text{MAC}}\right)} ,$$

and the proof under the stationarity assumption is completed.

REFERENCES

[1] T. M. Cover and A. A. El Gamal. Capacity theorems for the relay channel. *IEEE Trans. Info. Theory*, IT-25:572–584, Sept. 1979.

[2] T.J. Goblick. Theoretical limitations on the transmission of data from analog sources. *IEEE Trans. Info. Theory*, IT-11:558–567, 1965.

[3] B. Schein and R. G. Gallager. The Gaussian parallel relay channel. In *Proc. Int. Symp. Inform. Theory (ISIT), Sorrento, Italy*, page 22, June 2000.

[4] B.M. Oliver, J.R. Pierce, and C.E. Shannon. The philosofy of PCM. *Proc. IRE*, 36:1324–1331, Nov. 1948.

[5] M. Gastpar and M. Vetterli. On the capacity of large Gaussian relay networks. *IEEE Trans. Info. Theory*, IT-51:765–779, March 2005.

[6] U. Mittal and N. Phamdo. Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications. *IEEE Trans. Info. Theory*, IT-48:1082–1103, May 2002.

[7] V. M. Prabhakaran, R. Puri, and K. Ramchandran. Hybrid digital–analog codes for source–channel broadcast of Gaussian sources over Gaussian channels. *IEEE Trans. Info. Theory*, 57:4573–4588, July 2011.

[8] Y. Kochman and R. Zamir. Joint Wyner-Ziv/dirty-paper coding by modulo-lattice modulation. *IEEE Trans. Info. Theory*, IT-55:4878–4899, Nov. 2009.

[9] S. Saeed, S. S. C. Rezaei, S. O. Gharan, and A. K. Khandani. A new achievable rate for the Gaussian parallel relay channel. *IEEE Trans. Info. Theory*, submitted.

[10] M.S. Derpich, J. Ostergaard, and C. Goodwin. The quadratic Gaussian rate-distortion function for source uncorrelated distortions. In *Proc. of the Data Comp. Conf.*, pages 73–82, Snowbird, Utah, March 2008.

[11] B. Nazer and M. Gastpar. The case for structured random codes in network capacity theorems. *Euro. Trans. Telecomm.*, 19:455–474, June 2008.

[12] Y. Lomnitz and M. Feder. Communication over individual channels. *IEEE Trans. Info. Theory*, 57(11):7333–7358, Nov. 2011.

[13] M. Agarwal, A. Sahai, and S. Mitter. Coding into a source: a direct inverse rate-distortion theorem. *available at arXiv:cs/0610142*.

[14] A. Lapidoth and S. Tinguely. Sending a bi-variate Gaussian source over a Gaussian MAC. In *ISIT-2006, Seattle, WA*, pages 2124–2128, 2006.

[15] S. Bross, A. Lapidoth, and S. Tinguely. Broadcasting correlated Gaussians. In *ISIT-2008, Toronto, ON*, pages 1198–1202, July 2008.

[16] M. Gastpar. Uncoded transmission is exactly optimal for a simple Gaussian "sensor" network. *IEEE Trans. Info. Theory*, 54:5247–5251, Nov. 2008.

[17] A.D. Wyner. The rate-distortion function for source coding with side information at the decoder - II: General sources. *Information and Control*, 38:60–80, 1978.

[18] T. M. Cover and J. A. Thomas. *Elements of Information Theory, Second Edition*. Wiley, New York, 2006.

[19] J. Chen, S. Zhang, T. Berger, and S.B. Wicker. An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem. *IEEE Jr. on Selected Areas in Comm.*, 22:977–987, Aug. 2004.

[20] Y. Oohama. Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder. *IEEE Trans. Info. Theory*, IT-51:2577–2593, Jul. 2005.

[21] Y. Kochman and R. Zamir. The joint Gaussian CEO/MAC problem and relaying. In *Proceedings of the 46th Annual Allerton Conference on Communication, Control and Computing*, 2008.

[22] Y. Oohama. The rate-distortion function for the quadratic Gaussian CEO problem. *IEEE Trans. Info. Theory*, IT-44:1057–1070, May 1998.

[23] B. Schein. *Distributed Coordination in Network Informat Theory*. PhD thesis, Massachusetts Institute of Technology, 2001.

[24] T. Berger and D.W. Tufts. Optimum pulse amplitude modulation part I: Transmitter-receiver design and bounds from information theory. *IEEE Trans. Info. Theory*, IT-13:196–208, Apr. 1967.

[25] Y. Kochman and R. Zamir. Analog matching of colored sources to colored channels. *IEEE Trans. Info. Theory*, 57:3180–3195, June 2011.

[26] R. Zamir, Y. Kochman, and U. Erez. Achieving the Gaussian rate distortion function by prediction. *IEEE Trans. Info. Theory*, IT-54:3354–3364, July 2008.

[27] J.R. Barry, E.A. Lee, and D.G Messerschmitt. *Digital Communication*. Springer, 2004.

[28] A. Sanderovich, S. Shamai, and Y. Steinberg. Distributed MIMO receiver - achievable rates and upper bounds. *IEEE Trans. Info. Theory*, 55:4419–4438, Oct. 2009.