

# Broadcast Approach for the Information Bottleneck Channel

1<sup>st</sup> Avi Steiner

Andrew and Erna Viterbi Faculty of Electrical Engineering  
Technion—IIT, Haifa 3200003, Israel  
steiner.avi@gmail.com

2<sup>nd</sup> Shlomo Shamai (Shitz)

Andrew and Erna Viterbi Faculty of Electrical Engineering  
Technion—IIT, Haifa 3200003, Israel  
sshlo@ee.technion.ac.il

**Abstract**—This work considers a layered coding approach for efficient transmission of data over a wireless block fading channel, connected to a limited capacity reliable link, known as the bottleneck channel. Two main approaches are considered, the first is an oblivious approach, where the sampled noisy observations are compressed and transmitted over the bottleneck channel without having any knowledge of the original information codebook. This is compared to a decode-forward (non-oblivious) approach where the sampled noisy data is decoded, and whatever is successfully decoded is reliably transmitted over the bottleneck channel. In both settings it is possible to analytically describe in closed form expressions, the optimal continuous layering power distribution which maximizes the average achievable rate. Numerical results demonstrate the achievable broadcasting rate in the limit of continuous layering.

## I. INTRODUCTION AND PRELIMINARIES

Block fading channel model is commonly used for wireless communications, dominating the cases when mobile endpoints move slow relatively to the block coherence time. In slowly varying fading channels the fading realization is fixed throughout a transmission block, giving rise to the block fading notion. By this model, the receiver can easily learn the channel characterization over the block, thus we can assume perfect Channel State Information (CSI) only at the receiver side. In most practical cases, there is no feedback channel to the transmitter, resulting in its total unawareness of the instantaneous channel, yet it knows the channel statistics.

Consider the problem of transmitting over a block fading channel to a relay node, which has to forward the received signal to a destination over a reliable link with a fixed capacity  $C$ , see Figure 1 for the schematic channel model. For Gaussian channels this is known as the bottleneck channel [1]. This channel model is also applicable for the evolving next generation 5G/6G cellular networks, where the communication with the promising architecture of the Cloud radio access network (C-RAN) introduces stringent requirements on the fronthaul capacity and latency [2], [3], and many other timely contributions.

When transmitting over a block fading channel with receive CSI only, a broadcast approach may be considered on the transmission to maximize the average achievable rate. The broadcast approach, which is essentially a variable-to-fixed

channel coding [4], was studied in [5] for the MIMO fading channel with receiver CSI only. A finite capacity link to base-station subject to random fluctuations was studied in [6] for the case of two users connecting to the same base-station. Another related overview of matrix monotonic optimization is studied in [7]. Broadcast methods for the diamond channel, which is the two parallel relays channel, were studied in [8].

In the classical Gaussian bottleneck problem, depicted in Figure 1, define the random variable triplet  $X - Y - Z$  forming a Markov chain, and related according to

$$Y = h \cdot X + N, \quad (1)$$

where  $X$  and  $N$  are independent variance random variables, with  $N \sim CN(0, 1)$  being complex Gaussian with a unit variance, and the signal to noise ratio (SNR) is  $SNR = P|h|^2$ , where the gain  $|h|^2 = 1$  for a non-fading Gaussian channel, and  $P$  is the transmission power  $E[X^2] = P$ . The bottleneck channel output  $Z$  is a compressed version of  $Y$  adhering to a limited capacity of the bottleneck channel  $C$ . It is of interest to maximize

$$\max_{P(X), P(Z|Y)} \text{s.t. } I(Y; Z) \leq C \quad I(X; Z) \quad (2)$$

Evidently if  $X$  is Gaussian it is well known by Tishby et al [1], and [9], then also  $Y - Z$  is a Gaussian channel, and the maximization result of (2)

$$C_{Obliv} = I(X; Z) = \frac{1}{2} \log(1 + P|h|^2) - \frac{1}{2} \log(1 + P|h|^2 \cdot \exp(-2C)), \quad (3)$$

which follows immediately from the rate distortion approach, that is

$$Z = Y + M, \quad (4)$$

where the variance of  $Y$  is  $P|h|^2 + 1$  as determined by the channel model (1), and the variance of  $M$  is determined by requiring  $I(Z; Y) = C$ , that is

$$E[M^2] = \frac{P|h|^2 + 1}{\exp(2C) - 1} \quad (5)$$

The bottleneck gives reliable information rate that can be transmitted from  $X$  to  $Z$ , when the relay  $Y$  operates in an oblivious manner (it has no knowledge about the codebook and can not decode the message) [10]. For a non oblivious approach the result is immediate, as the relay may decode the

This work was supported by the European Union's Horizon 2020 Research And Innovation Programme, Grant Agreement No. 694630.

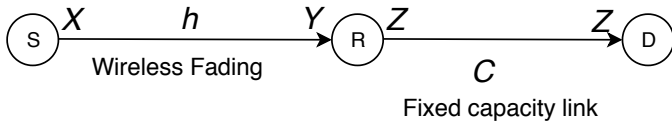


Fig. 1. Information bottleneck fading channel system model block diagram.

data, and then transmit over the limited bandwidth channel  $Y - Z$  at rate  $C$ . Therefore the achievable transmission rate is the minimum of the two channels capacity,

$$C_{non-Obliv} = \min\left\{\frac{1}{2}\log(1 + P|h|^2), C\right\}. \quad (6)$$

### A. Broadcast Approach Preliminaries

Consider a transmitted signal  $X$  is composed of multi-layer coded information, in a continuum of layers, such that each code layer receives an infinitesimal power  $\rho(u)du$ . The broadcast approach was introduced in detail in [5]. We briefly review the principles of the broadcast approach. The incremental rate as function of power allocation, for a Gaussian fading channel, is [5]

$$dR(u) = \frac{1}{2} \frac{\rho(u)udu}{1 + I(u)u} \quad (7)$$

where  $I(u)$  is the residual interference function, such that  $I(0) = P$ , and  $\rho(u) = -I'(u)$  is the power allocation density function. The total allocated rate as function of  $s$  is thus

$$R(s) = \frac{1}{2} \int_0^s \frac{\rho(u)udu}{1 + I(u)u} \quad (8)$$

where  $s$  is a continuous layer index, which also corresponds to the channel fading gain. The maximal average rate is expressed as follows

$$R_{bs,avg} = \max_{I(u)} \frac{1}{2} \int_0^\infty du (1 - F_\nu(u)) \frac{\rho(u)u}{1 + I(u)u} \quad (9)$$

where  $F_\nu(u)$  is the cumulative distribution function (CDF) of the fading gain random variable. It can be shown, [5], that the optimal power allocation is given by

$$I_{opt}(u) = \begin{cases} P & u < u_0 \\ \frac{1 - F_\nu(u) - u \cdot f_\nu(u)}{u^2 f_\nu(u)} & u_0 \leq u \leq u_1 \\ 0 & u > u_1 \end{cases} \quad (10)$$

where  $u_0$  and  $u_1$  are obtained from the boundary conditions  $I_{opt}(u_0) = P$ , and  $I_{opt}(u_1) = 0$ , respectively.

Interestingly, the optimal allocated rate can be expressed in closed form by substituting the optimal power allocation (10) into the cumulative layering rate in (8), by

$$R_{opt}(s) = \begin{cases} 0 & s < u_0 \\ \log(s/u_0) + \frac{1}{2} \log\left(\frac{f_\nu(s)}{f_\nu(u_0)}\right) & u_0 \leq s \leq u_1 \\ \log(u_1/u_0) + \frac{1}{2} \log\left(\frac{f_\nu(u_1)}{f_\nu(u_0)}\right) & s > u_1 \end{cases} \quad (11)$$

## II. BROADCAST APPROACH FOR FADING INFORMATION BOTTLENECK CHANNEL

Consider a fading channel on the wireless link to  $Y$ , where  $s = |h|^2$  is the block fading gain with a unit variance. Under a slowly fading channel, the random variable gain  $s$  changes independently from codeword to codeword, and remains fixed over the codeword. The channel model for  $Z$  can be expressed by its block fading gain, under an oblivious approach

$$Z = \text{sqrt}(SNR_{eq})X + N, \quad (12)$$

where  $N$  is a unit variance Gaussian noise, and

$$SNR_{eq} = \frac{s(1 - \exp(-2C))}{1 + s \cdot P \cdot \exp(-2C)} \quad (13)$$

which is directly obtained from (5).

### A. Oblivious Bottleneck Channel Approach

In the oblivious approach, the received codeword in  $Y$  can not be decoded, as there is no information on the used codebook, therefore, the compression of  $Y$  into  $Z$  is performed accounting for the distribution of  $Y$  only. Under a fading channel model, the ergodic capacity of the bottleneck fading channel is determined by

$$\begin{aligned} C_{Obliv,Erg} &= E_s \left[ \frac{1}{2} \log(1 + P \dot{SNR}_{eq}) \right] \\ &= E_s \left[ \frac{1}{2} \log \frac{1 + s \cdot P}{1 + s \cdot P \cdot \exp(-2C)} \right] \end{aligned} \quad (14)$$

1) *Single Layer Coding*: Using a single layer coding approach for the fading channel, the achievable average rate depends on the allocated rate, and the fading distribution in the following way. In the oblivious communication scheme, let the transmitter allocate a rate  $R_{1,obliv}$  as function of a fading threshold parameter  $s_{th}$ . Then, the decoding of the noisy compressed signal  $Z$  is possible for fading gains  $s \geq s_{th}$ , and the allocated rate corresponds to

$$\begin{aligned} R_{1,obliv} &= \frac{1}{2} \log(1 + SNR_{eq}(s_{th})P) \\ &= \frac{1}{2} \log \left( 1 + \frac{s_{th}(1 - \exp(-2C))P}{1 + s_{th} \cdot P \cdot \exp(-2C)} \right) \end{aligned} \quad (15)$$

Since  $SNR_{eq}(s_{th})$  is a monotonic function of  $s_{th}$ , the rate  $R_{1,obliv}$  can be achieved for any fading gain  $s \geq s_{th}$ , and therefore the average rate with a single layer is

$$R_{1,obliv,avg} = \frac{1}{2} (1 - F(s_{th})) \log \left( 1 + \frac{s_{th}(1 - \exp(-2C))P}{1 + s_{th} \cdot P \cdot \exp(-2C)} \right) \quad (16)$$

and the outage capacity is

$$\max_{s_{th} \geq 0} (1 - F(s_{th})) \frac{1}{2} \log \left( 1 + \frac{s_{th}(1 - \exp(-2C))P}{1 + s_{th} \cdot P \cdot \exp(-2C)} \right) \quad (17)$$

2) *Continuous Broadcast Approach*: We derive here the continuous broadcasting approach, where the transmitted signal  $X$  is multi-layer coded, in a continuum of layers, such that each code layer receives an infinitesimal power  $\rho(u)du$ . The broadcast approach was introduced in detail in [5]. The incremental rate as function of power allocation is defined in (7), for a Gaussian fading channel [5]. The total allocated rate as function of the fading gain  $s$  is thus (8), with an average broadcasting rate  $R_{bs,avg}$  as in (9). The optimal power allocation  $I_{opt}(u)$  (10) is given in closed form. In this oblivious bottleneck channel, the broadcast approach is optimized for a fading distribution  $F_\nu(u)$  where  $\nu = SNR_{eq}$  from (13), which reflects an equivalent channel gain which depends on the channel fading gain  $s$  distribution, and on the bottleneck channel capacity, as well as the transmission power. Clearly for high bottleneck channel capacity  $C \rightarrow \infty$ , then  $SNR_{eq} \rightarrow s$ .

### B. A Decode-Forward Bottleneck Channel Approach

On the decode-forward (non-oblivious) approach, the received codeword in  $Y$  can be decoded, and then all the decoded data up to rate  $C$  can be reliably conveyed to  $Z$ . Under a fading channel model, the decode-forward ergodic capacity of the bottleneck fading channel  $C_{non-Obliv,Erg}$  provides an ergodic upper bound which is not achievable under a block fading model, where each codeword is transmitted over a relatively short duration with a single channel realization, without capturing the full distribution of  $s$ . For the slowly block fading channel it is beneficial to transmit a multi-layered codeword when transmitter has no channel state information (CSI). Under this model, decode-forward ergodic capacity of the bottleneck fading channel is formulated as

$$C_{non-Obliv,Erg} = E_s[\min\{\frac{1}{2}\log(1+sP), C\}]. \quad (18)$$

which corresponds to a block fading model, where the transmission and decoding are done over a single fading realization, due to a slow fading nature of the channel, and the ergodic capacity is obtained over a long period of time by capturing the complete fading distribution. Note that ergodic capacity is achievable only with transmit CSI, which not available according to the channel model in this work (1), which makes it a bound.

1) *Single Layer Coding*: Using a single layer coding approach for the fading channel, the achievable average rate depends on the allocated rate, the bottleneck channel capacity  $C$  and the fading distribution in the following way. In the decode-forward communication scheme, let the transmitter allocate a rate  $R_{1,non-obliv}$  as function of a fading threshold parameter  $s_{th}$ . Then, the decoding of the noisy compressed signal  $Z$  is possible for fading gains  $s \geq s_{th}$ , and the allocated rate corresponds to

$$R_{1,non-obliv} = \frac{1}{2}\log(1+s_{th}P) \quad (19)$$

where the rate  $R_{1,obliv}$  is selected such that  $R_{1,obliv} \leq C$ , and can be achieved for any fading gain  $s \geq s_{th}$ , and conveyed

reliably over the bottleneck channel after decoding. Therefore the average rate with a single layer is

$$R_{1,non-obliv,avg} = (1 - F(s_{th})) \cdot \frac{1}{2}\log(1+s_{th}P) \quad (20)$$

and hence the outage capacity of the decode-forward non-oblivious channel is given by

$$R_{1,non-obliv,avg} = \max_{s_{th} \geq 0} (1 - F(s_{th})) \cdot \min(C, \frac{1}{2}\log(1+s_{th}P)) \quad (21)$$

2) *Continuous Broadcast Approach*: We derive here the continuous broadcasting strategy for the decode-forward approach, where the transmitted signal  $X$  is multi-layer coded, in a continuum of layers. The received signal  $Y$  is decoded layer-by-layer in a successive decoding manner. All the successfully decoded layers with a total rate up to the bottleneck channel capacity  $C$  can be reliably conveyed over the bottleneck channel. The broadcast approach optimization goal is to maximize the average transmitted rate over the bottleneck channel in this block fading channel model. We formulate here the optimization of power density distribution function  $\rho_{opt}(u)$  so that average transmission rate is maximized under the bottleneck channel capacity constraint.

*Proposition 1*: For the decode-forward block fading bottleneck channel, the total *expected average achievable rate of the broadcast approach* is obtained by the following residual power distribution function

$$I_{opt}(u) = \arg \max_{I(u)} \frac{1}{2} \int_0^\infty du (1 - F_\nu(u)) \frac{\rho(u)u}{1 + I(u)u}, \quad (22)$$

$$s.t. \int_0^\infty du \frac{\rho(u)u}{1 + I(u)u} \leq C$$

where  $F_\nu(u)$  is the CDF of the fading gain random variable, and  $C$  is the bottleneck channel capacity. The optimal power allocation  $I_{opt}(u)$  is given by

$$I_{opt}(u) = \begin{cases} P & u < u_0 \\ \frac{1 - F_\nu(u) + \lambda_{opt} - u \cdot f_\nu(u)}{u^2 f_\nu(u)} & u_0 \leq u \leq u_1 \\ 0 & u > u_1 \end{cases} \quad (23)$$

where  $\lambda_{opt} \geq 0$  is a Lagrange multiplier specified by

$$\lambda_{opt} = -u_1 \cdot f_\nu(u_1) - 1 + F_\nu(u_1) \quad (24)$$

and for any  $\lambda_{opt} > 0$ ,

$$u_1^2 \cdot f_\nu(u_1) = \exp(2C) \cdot u_0^2 \cdot f_\nu(u_0) \quad (25)$$

*Proof*: This is a constrained variational problem which can be solved in closed form. The details are omitted due to space limitation.

Fig. 2 demonstrates the achievable rate with a single layer coding oblivious approach, specified in (15), compared to the ergodic oblivious bound, as specified in (14). The multiple curves correspond to different values of the bottleneck channel capacity  $C$ . It is clear from the results here that for small  $C$  the single layer asymptotically achieves the ergodic bound, while

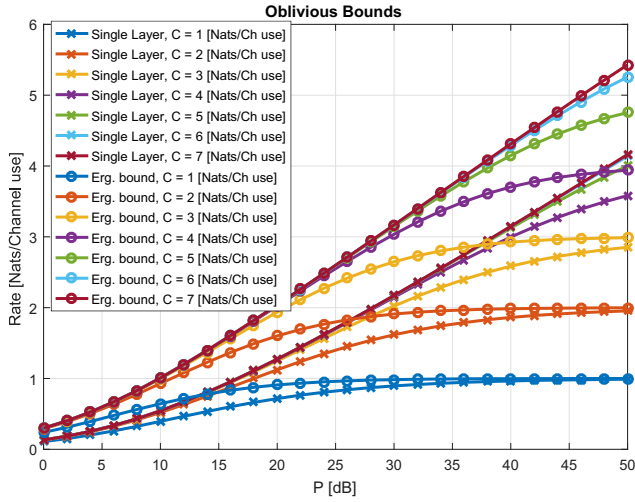


Fig. 2. Oblivious approach of single layer coding vs. ergodic capacity, as function of bottleneck channel capacity.

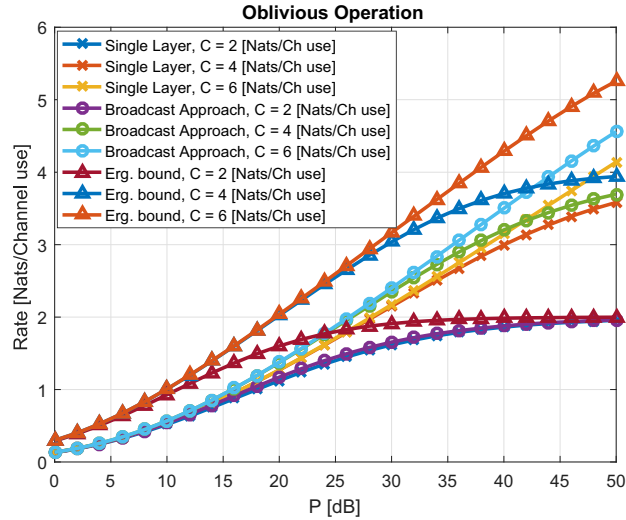


Fig. 3. Oblivious approach of single layer coding and broadcast approach compared to the ergodic capacity, as function of bottleneck channel capacity.

for  $C \geq 3$  there is a large gap of the single layer approach to the ergodic bound, which may be narrowed down by using the broadcast approach.

Fig. 3 demonstrates the achievable rates with a single layer coding oblivious approach, specified in (15), compared to the oblivious broadcast approach, specified in (9)-(10) where  $F_\nu(u)$  is defined by  $\nu = SNR_{eq}$  from (13), and the ergodic oblivious bound (14). The multiple curves correspond to different values of the bottleneck channel capacity  $C$ . It may be observed from the results that the higher the bottleneck channel capacity  $C$ , the higher is the oblivious broadcast approach gain compared to the single layer coding approach. Fig. 4 demonstrates the achievable rates with a single layer coding decode-forward approach, specified in (19), compared

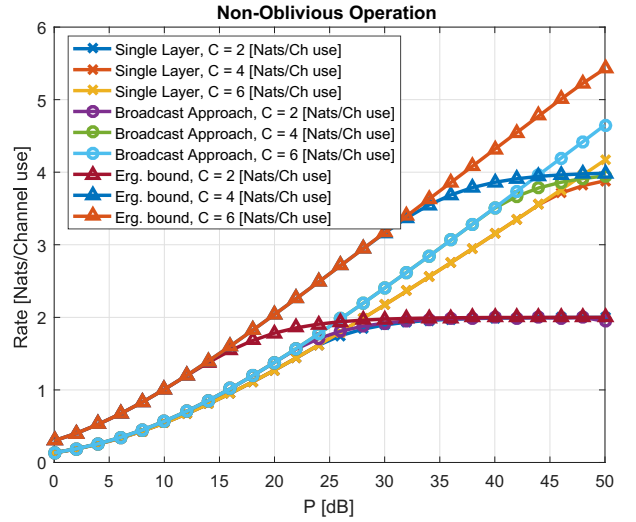


Fig. 4. Non-Oblivious single layer coding and broadcast approach compared to the ergodic capacity, as function of bottleneck channel capacity.

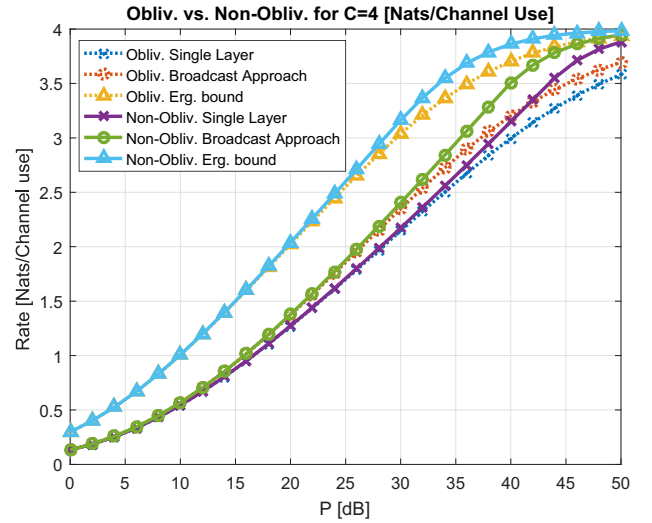


Fig. 5. Oblivious vs. Non-Oblivious single layer coding and broadcast approach compared to the ergodic capacity, for bottleneck channel capacity of  $C = 4$  [Nats/Channel Use].

to the decode-forward broadcast approach, specified in Proposition 1, and the ergodic oblivious bound (18). The multiple curves correspond to different values of the bottleneck channel capacity  $C$ . It may be observed from the results that the higher the bottleneck channel capacity  $C$ , the higher is the decode-forward (non-oblivious) broadcast approach gain compared to the single layer coding decode-forward approach. Fig. 5 demonstrates the achievable rates with a decode-forward approach as compared to an oblivious approach, for a bottleneck channel capacity  $C = 4$  [Nats/Channel use]. It can be observed here that at high SNR region the gain of the broadcast approach compared to single layer coding is higher with a decode-forward approach.

### III. NUMERICAL RESULTS

The following section provides some examples of achievable rates with single layer coding and continuous broadcasting, with comparison to the ergodic bound, for the block fading information-bottleneck channel. Both oblivious, and decode-forward (non-oblivious) approaches are evaluated. The numerical results are calculated for a Rayleigh fading channel, where  $F_\nu(u) = 1 - \exp(-u)$ .

### IV. CONCLUSION

This work considers the problem of efficient transmission over the block fading channel with a bottleneck limited channel. Two main approaches are considered, the first is an oblivious approach, where the sampled noisy observations are compressed and transmitted over the bottleneck channel without having any knowledge of the original information codebook. This is compared to a decode-forward (non-oblivious) approach where the sampled noisy data is decoded, and whatever is successfully decoded is reliably transmitted over the bottleneck channel. As may be observed from the numerical results, the decode-forward approach has a noticeable advantage over the oblivious relay at high SNRs, for both the single layer coding and broadcast approach. In addition, on many cases the broadcast approach gains more compared to the single layer coding under the decode-forward strategy. In both settings it is possible to analytically describe in closed form expressions, the optimal continuous layering power distribution which maximizes the average achievable rate. Future work may include adapting the broadcast MIMO approach for the vector bottleneck channel [11].

### ACKNOWLEDGMENT

#### REFERENCES

- [1] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, "Information bottleneck for gaussian variables," in *Advances in Neural Information Processing Systems 16*, S. Thrun, L. K. Saul, and B. Schölkopf, Eds. MIT Press, 2004, pp. 1213–1220. [Online]. Available: <http://papers.nips.cc/paper/2457-information-bottleneck-for-gaussian-variables.pdf>
- [2] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5g backhaul challenges and emerging research directions: A survey," *IEEE Access*, vol. 4, pp. 1743–1766, 2016.
- [3] K. Zeb, X. Zhang, and Z. Lu, "High capacity mode division multiplexing based mimo enabled all-optical analog millimeter-wave over fiber fronthaul architecture for 5g and beyond," *IEEE Access*, vol. 7, pp. 89 522–89 533, 2019.
- [4] S. Verdú and S. Shamai, "Variable-rate channel capacity," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2651–2667, June 2010.
- [5] S. Shamai and A. Steiner, "A broadcast approach for a single-user slowly fading mimo channel," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2617–2635, 2003.
- [6] R. Karasik, O. Simeone, and S. S. Shitz, "Robust uplink communications over fading channels with variable backhaul connectivity," *IEEE Trans. on Wireless Communications*, vol. 12, no. 11, pp. 5788–5799, November 2013.
- [7] C. Xing, S. Wang, S. Chen, S. Ma, H. V. Poor, and L. Hanzo, "Unified overview of matrix-monotonic optimization for MIMO transceivers," *CoRR*, vol. abs/1810.11244, 2018. [Online]. Available: <http://arxiv.org/abs/1810.11244>.
- [8] M. Zamani and A. K. Khandani, "Broadcast approaches to the diamond channel," *IEEE Trans. on Information Theory*, vol. 60, no. 1, pp. 623–642, Jan. 2014.
- [9] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *Proc. of the 37-th Annual Allerton Conference on Communication, Control and Computing*, 1999, pp. 368–377.
- [10] A. Sanderovich, S. Shamai (Shitz, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Trans. on Information Theory*, vol. 54, pp. 3008–3023, 07 2008.
- [11] A. Winkelbauer, S. Farthofer, and G. Matz, "The rate-information trade-off for gaussian vector channels," in *2014 IEEE International Symposium on Information Theory*, June 2014, pp. 2849–2853.