Broadcast Approach under Information Bottleneck Capacity Uncertainty

Shlomo Shamai (Shitz)

Joint work with Avi Steiner

Viterbi Faculty of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 3200003, Israel

ITA 2020, UCSD, San-Diego

2–7 February, 2020

Overview

- Broadcast approach preliminaries
- Information bottleneck channel model
- Broadcast approach for the information bottleneck channel
  - Known bottleneck channel capacity
  - Uncertainty in bottleneck channel capacity
- Some numerical examples
- Conclusion

The Case for a Broadcast Approach

Ergodic capacity, the codeword is transmitted over many block (many realizations of S) $\rightarrow I(X;Y|S)$. Consider a state dependent channel (Cover '72): $P(Y|X,S)$
Y - channel output $Y \in \mathbb{Y}$.

X - channel input $X \in \mathbb{X}$.

S - channel state parameter, where $S \in \mathbb{S}$.

- Broadcast approach = Variable-to-Fixed channel coding (Verdu-Shamai-IT'10)
- This is viewed as a broadcast setting where each possible state S is associated with a different receiver.
- Given a probability space for $\mathbb{S}$, problems of average type (rate/delay/distortion) performance are motivated.

★ Reliable transmission with rate **adapted** to actual channel realization, where channel state is unavailable at the transmitter (Cover'72), (Equitz-Cover'91), (Rimoldi'94), (Shamai'97), (Shamai-Steiner-IT '03), .

The Block Fading Channel Model

$$\mathbf{y} = h\mathbf{x} + \mathbf{n} \tag{1}$$

$\mathbf{x}_{1\times N}$ - transmitted vector, with power constraint $\frac{1}{N}E[\mathbf{x}\mathbf{x}^{\dagger}] \leq P$.

$\mathbf{y}_{1\times N}$ - received vector.

$\mathbf{n}_{1\times N}$ - additive white Gaussian noise (AWGN), with iid elements $N(0,1)$.

h - fading coefficient, perfectly known by the receiver, fixed over every block, $N(0,1)$ iid distributed over multiple blocks.

$s = |h|^2$ - channel state parameter.

h remains fixed during every transmission block-during which the full message is transmitted, and known at the receiver, but unavailable at the transmitter

Ergodic capacity of the fading channel: $E\frac{1}{2}log(1 + |h|^2P)$

Details in (Shamai '97) (Shamai-Steiner-IT '03).

Broadcast Approach - Definitions

- s - fading gain designating SNR of a virtual receiver.
- Transmitter views a *degraded* broadcast channel.
- $R(s)$ - the reliably conveyed information rate at fading level s.
- Power assigned to the s-th stream $\rho(s)ds$.
- Information streams indexed by $u > s$ are undetectable $\Longrightarrow$ residual interference $I(s)$.

Broadcast Approach - Overview

- Incremental differential rate

$$dR(s) \;=\; \frac{1}{2}\log\left(1+\frac{s\rho(s)ds}{1+sI(s)}\right) \;=\; \frac{1}{2}\cdot\frac{s\rho(s)ds}{1+sI(s)} \tag{2}$$

- The residual interference power is

$$I(s)=\int_s^\infty \rho(u)du \tag{3}$$

- Total transmitted power P is

$$P=\int_0^\infty \rho(u)du = I(0) \tag{4}$$

Broadcast Approach - Overview

- Reliable rate at fading level s

$$R(s) = \frac{1}{2}\int_0^s \frac{u\rho(u)du}{1+uI(u)} \quad (5)$$

- Expected achievable broadcasting rate

$$R_T = \int_0^\infty du\ f_s(u)R(u) = \int_0^\infty du\frac{1}{2}(1-F_s(u))\frac{u\rho(u)}{1+uI(u)} \quad (6)$$

$f_s(u)$ - the pdf of the fading power s, $F_s(u) = \int\limits_0^u da f_s(a)$ - the corresponding cdf.

- Optimization problem: maximize expected rate over the input power distribution

$$R_{T,max} = \max_{I(u)} \int_0^\infty du\frac{1}{2}(1-F_s(u))\frac{u\rho(u)}{1+uI(u)} \equiv \max_{I(u)} \int_0^\infty duS(u,I,I_u) \quad (7$$

where $I_u \equiv \frac{dI(u)}{du} = -\rho(u)$.

Broadcast Approach - Overview I

Proposition 1

The power distribution, which maximizes the expected throughput in (7) is

$$I(x) = \begin{cases} \frac{1-F_s(x)-x\cdot f_s(x)}{x^2 f_s(x)} & , x_0 \le x \le x_1 \\ 0 & , else \end{cases} \quad (8)$$

where x_0 is determined by $I(x_0) = P$, and x_1 by $I(x_1) = 0$. And the broadcasting rate is expressed as function of the fading gain distribution

$$R_{opt}(s) = \begin{cases} 0 & s < x_0 \\ \log(s/x_0) + \frac{1}{2}\log\left(\frac{f_s(s)}{f_s(x_0)}\right) & u_0 \le s \le u_1 \\ \log(x_1/x_0) + \frac{1}{2}\log\left(\frac{f_s(x_1)}{f_s(x_0)}\right) & s > x_1 \end{cases} \quad (9)$$

Broadcast Approach - Overview II

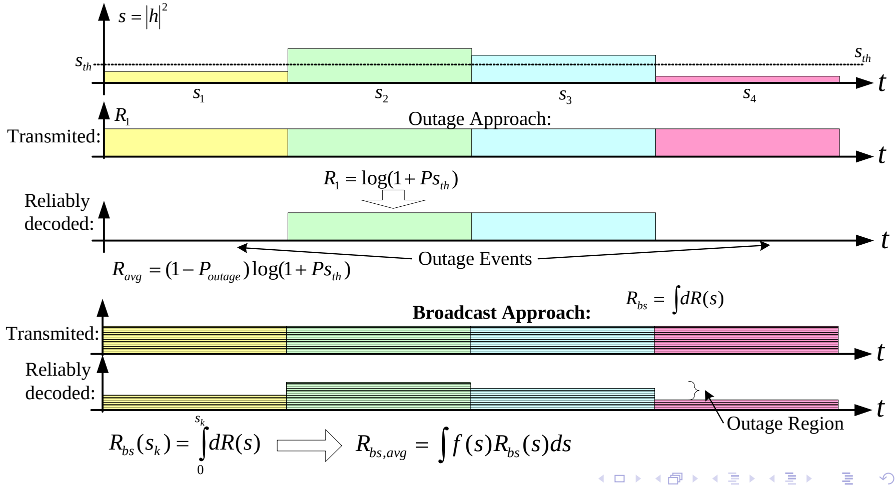
Proof: The extremum condition for the functional (7) is given by the associated Eüler equation

$$S_I - \frac{d}{du} S_{I_u} = 0$$

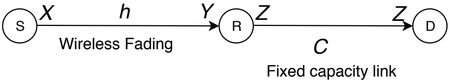
which simplifies from a differential equation into a linear equation by $I(u)$, yielding (8).

Broadcast vs. Outage (a special case)

$\mathbf{y} = h\mathbf{x} + \mathbf{n},$ Channel model: block fading



Information Bottleneck Channel Model I



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

$$Y = h \cdot X + N, \quad s = |h|^2, \tag{10}$$

- Maximize 'relevance': $I(X;Z)$, under the constraint of 'complexity': $I(Y;Z)$
- For information transmission: The bottleneck channel output Z is a compressed version of Y adhering to a limited capacity of the bottleneck channel C, and oblivious to the transmission codebook
- This is equivalent to remote source coding (Dobrushin '62) with logarithmic loss

Information Bottleneck Channel Model II

- For the Gaussian X case (Tishby et. al. '99) $Y - Z$ is a Gaussian channel

$$\begin{aligned} C_{Obliv} \quad &= I(X;Z) = \tfrac{1}{2}\log(1+P|h|^2) \\ &-\tfrac{1}{2}\log(1+P|h|^2\cdot\exp(-2C)), \end{aligned} \tag{11}$$

- Thus the original $F_s(u)$ the cdf of the fading is replaced by $F_v(u)$ the cdf of an equivalent fading gain FPR_{eq}, which equals

$$F_v(u) = F_s\left(\frac{u}{1-(1+Pu)\exp(-2C)}\right),\ 0 \le u \le (exp(2C)-1)/P \tag{12}$$

Oblivious Broadcasting over the Information Bottleneck Channel

- The channel model for Z can be expressed by

$$Z = sqrt(FPR_{eq})X + N, \tag{13}$$

where N is a unit variance Gaussian noise, and

$$FPR_{eq} = \frac{s(1 - \exp(-2C))}{1 + s \cdot P \cdot \exp(-2C)} \tag{14}$$

- Single layer coding (outage), maximal achievable expected rate is

$$R_{1,obliv,avg} = \max_{s_{th} \geq 0} (1 - F_s(s_{th}))\frac{1}{2}log\left(1 + \frac{s_{th}(1-\exp(-2C))P}{1+s_{th} \cdot P \cdot \exp(-2C)}\right) \tag{15}$$

- Continuous broadcast approach is optimized for a fading distribution $F_\nu(u)$ where $\nu = FPR_{eq}$ (14): equivalent channel gain depending on the channel fading gain s, bottleneck channel capacity C, and transmission power P

Decode-Forward Broadcasting over the Information Bottleneck Channel

- Decode-forward (non-oblivious) approach, the relay received codeword in Y can be decoded, and the decoded data rate limited to C can be reliably conveyed to destination Z.
- Single layer coding, maximal achievable expected rate is

$$R_{1,non-obliv,avg} = \max_{0 \leq s_{th} \leq \exp(2C)-1} (1 - F_s(s_{th})) \cdot \frac{1}{2} log(1 + s_{th}P) \quad (16)$$

- Broadcast approach: Layers are successively decoded from Y. Successfully decoded layers are reliably conveyed to Z, provided that total rate $\leq C$
- The broadcast approach maximal achievable expected rate can be obtained in a closed form

Decode-Forward Broadcasting: Expected Rate Maximization I

Proposition 2

The expected average achievable rate of the broadcast approach is obtained by the following power distribution

$$I_{opt}(u) = \arg\max_{I(u)} \frac{1}{2} \int_0^\infty du (1 - F_s(u)) \frac{\rho(u)u}{1 + I(u)u}, \quad s.t. \int_0^\infty du \frac{\rho(u)u}{1 + I(u)u} \leq C$$

The optimal power allocation is

$$I_{opt}(u) = \begin{cases} P & u < u_0 \\ \frac{1 - F_s(u) + \lambda_{opt} - u \cdot f_s(u)}{u^2 f_s(u)} & u_0 \leq u \leq u_1 \\ 0 & u > u_1 \end{cases} \tag{17}$$

Decode-Forward Broadcasting: Expected Rate Maximization II

where $\lambda_{opt} \geq 0$ in Proposition 2 is a Lagrange multiplier specified by

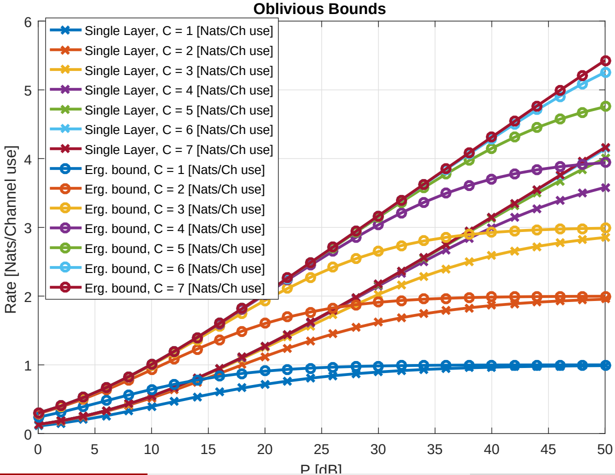
$$\lambda_{opt} = -u_1 \cdot f_s(u_1) - 1 + F_s(u_1) \tag{18}$$

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

$$u_1^2 \cdot f_s(u_1) = \exp(2C) \cdot u_0^2 \cdot f_s(u_0) \tag{19}$$

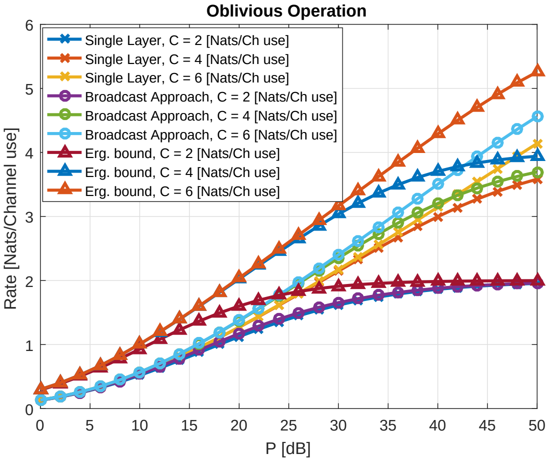
Oblivious Single-Layer-Coding over the Information Bottleneck Channel

Single layer coding vs. ergodic capacity, as function of bottleneck channel capacity



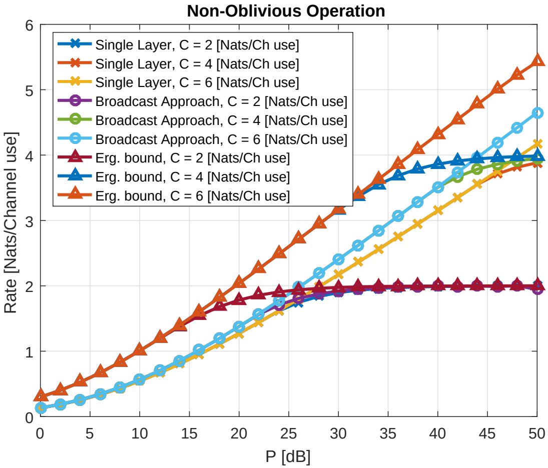
Oblivious Broadcasting over the Information Bottleneck Channel

Single layer vs. Broadcast approach vs. ergodic capacity vs. bottleneck capacities



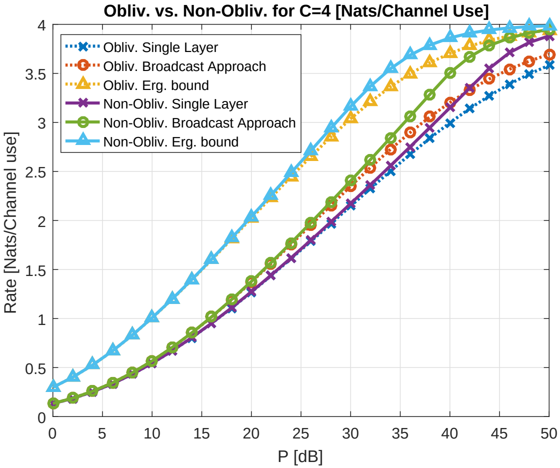
Decode-Forward Broadcasting over the Information Bottleneck Channel

Single layer vs. Broadcast approach vs. ergodic capacity vs. bottleneck capacities



Oblivious and Decode-Forward Broadcasting over the Information Bottleneck Channel

Oblivious vs. Decode-Forward vs. Single layer vs. Broadcast approach



Broadcasting under Information Bottleneck Capacity Uncertainty I

- Bottleneck capacity may dynamically change due to rapidly varying bandwidth demands
- It is assumed that transmitter has no information of the available bottleneck capacity, only its distribution. The relay knows the available capacity for each codeword
- Uncertainty of a bottleneck channel capacity is defined by a discrete random variable C_b, which admits to N capacity values $\{C_i\}_{i=1}^N$ s.t. $C_1 \geq C_2 \geq \cdots \geq C_N$, with corresponding probabilities $\{p_{b,i}\}_{i=1}^N$, s.t. $p_{b,i} \geq 0$ and $\sum_{i=1}^N p_{b,i} = 1$.
- The average bottleneck capacity is denoted by $\bar{C} = E[C_b]$

Oblivious Broadcasting under Information Bottleneck Capacity Uncertainty I

- Under oblivious broadcasting, the combined equivalent channel viewed by the transmitter

$$FPR_{eq}(s, C_b) = \frac{s(1 - \exp(-2C_b))}{1 + s \cdot P \cdot \exp(-2C_b)}, \quad s = |h|^2, \tag{20}$$

- Continuous broadcast approach is optimized for a fading distribution $F_\mu(u)$ where $\mu = FPR_{eq}(s, C_b)$ (20) : equivalent channel gain depending on the fading gain realization s, and bottleneck channel capacity C_b available per codeword
- The cdf of this fading gain is

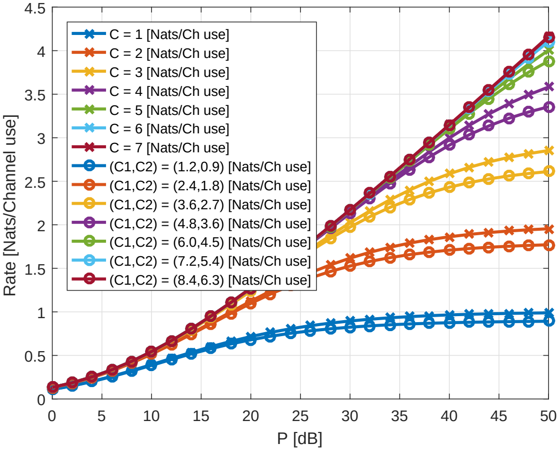
$$F_\mu(u) = \sum_{i=1}^{N} p_{b,i} F_s \left(\frac{u}{1 - (1 + Pu)\exp(-2C_i)} \right) \tag{21}$$

Oblivious Broadcasting under Information Bottleneck Capacity Uncertainty II

- A simple upper bound for the oblivious is given by the fixed backhaul capacity $C = \bar{C}$ (the average), which actually is the exact result in case of ergodic backhaul model (that is there are many backhaul realizations during a single fading state).
- The decode and forward strategy is immediate, as for any fading power s realization the respective number of layers are decoded, and the rate transmitted to the destination can not be larger than the actual backhaul realization C_i.

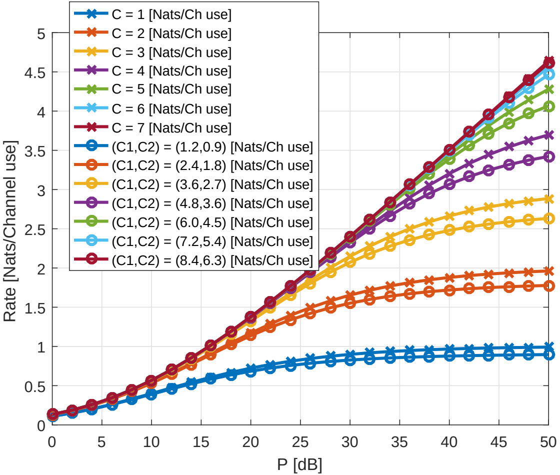
Oblivious Single-Layer-Coding over Variable Capacity Information Bottleneck Channel

Oblivious single layer coding, $C = \sum_{i=1}^{N} p_{b,i} C_i$, and $p_{b,1} = 1/3$



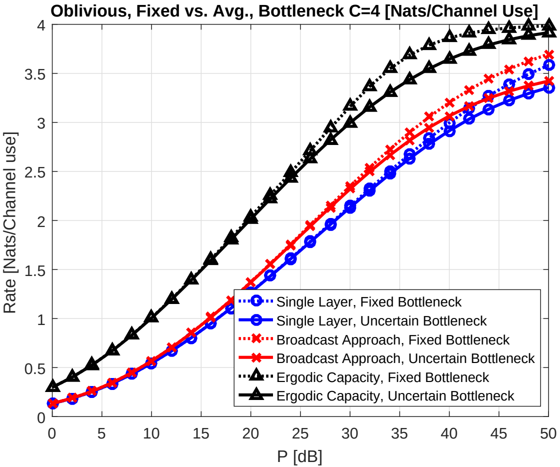
Oblivious Broadcast Approach over Variable Capacity Information Bottleneck Channel

Oblivious Broadcast approach, $C = \sum_{i=1}^{N} p_{b,i} C_i$, and $p_{b,1} = 1/3$



Oblivious Broadcast Approach over Variable Capacity Information Bottleneck Channel

Single Layer vs. Broadcast approach vs. Ergodic Capacity



Conclusion and Future Work I

- Broadcast approach with oblivious and decode-forward bottleneck fading channel was studied. Both **fixed** and **variable** bottleneck channel capacity are considered, where transmitter does not know fading realization and bottleneck capacity, only their distribution
- The optimal continuous layering power distribution maximizing the expected achievable rate was obtained in closed form
- The decode-forward approach has a noticeable advantage over the oblivious relay at high SNRs, for both the single layer coding and the broadcast approach
- The broadcast approach usually gains more compared to the single layer coding under the decode-forward strategy

Conclusion and Future Work II

- Multilayering at the transmitter gives little advantage for mild fading and high SNR (P) regimes (it is seen that $P \cdot FPR_{eq}$ depends at the limit $P \to \infty$, only of the realizations C_i). The performance then of oblivious and decode and forward techniques are similar
- Future work may include
  - adapting the broadcast MIMO approach (Shamai et. al. '03) for the vector bottleneck channel (Winkelbauer et. al. '14)
  - Extending the bottleneck channel model to multiple relays - the diamond channel (Aguerri et. al. 2019)
  - Extending the model to multiple users, relevant to Cloud Radio Access Networks (CRAN) (Aguerri et. al. 2019)

Conclusion and Future Work III

- Extending previous efforts (Karasik et al, 2013) to scenarios where the variable backhaul links capacities $\{C_i\}$ are not available at the relay node, but at the destination only. Successive refinement source coding techniques (Tian et. al. 2008) and (Ng et. al. 2009) adapted to the logarithmic loss, provide the basic tools that are in current study for this application adhering to oblivious processing

References I

[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] A. Steiner and S. Shamai, "Broadcast Approach for the Information Bottleneck Channel," *International Conference on Microwaves, Communications, Antennas & Electronic Systems* (COMCAS2019), 4-6, November 2019, Tel Aviv, Israel.

[3] 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.

[4] 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.

[5] S. Verdu and S. Shamai, "Variable-rate channel capacity," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2651–2667, June 2010.

References II

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

References III

[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.

[12] I. Aguerri, A. Zaidi, G. Caire and S. Shamai (Shitz), ”On the Capacity of Cloud Radio Access Networks with Oblivious Relaying,” IEEE Transactions on Information Theory, vol. 65, no. 7, pp. 4575-4596, July 2019.

[13] X. Wu, A. Ozgur, M. Peleg and S. Shamai (Shitz), ”New Upper Bounds on the Capacity of Primitive Diamond Relay Channels,” in IEEE Information Theory Workshop (ITW2019) Visby, Gotland, Sweden, 25-28 August 2019.

[14] M. Mondelli, S. H. Hassani and R. Urbanke, ”A New Coding Paradigm for the Primitive Relay Channel,” Algorithms, MDPI, Published: 18 October 2019.

[15] R. Karasik, O. Simeone and S. S. Shitz, ”Robust uplink communications over fading channels with variable backhaul connectivity,” 2013 IEEE International Symposium on Information Theory, Istanbul, 2013, pp. 1172-1176.

References IV

[16] R. L. Dobrushin and B. S. Tsybakov, ''Information transmission with additional noise,'' Trans. Inform. Theory, vol. IT-8, pp. 293304, 1962.

[17] C. Tian, A. Steiner, S. Shamai (Shitz) and S. Diggavi, "Successive Refinement Via Broadcast: Optimizing Expected Distortion of a Gaussian Source Over a Gaussian Fading Channel," IEEE Trans. Information Theory, Vol. 54, No. 7, pp. 2903-2918, July 2008.

[18] C. T. K. Ng, D. Gunduz, A.J. Goldsmith adn E. Erkip, "Distortion Minimization in Gaussian Layered Broadcast Coding With Successive Refinement," Trans. Information Theory, Vol. 55, No. 11, pp. 5074-5086, Nov. 2009.

THANK YOU!

ITA 2020 Paper

Title: "Broadcast Approach under Information Bottleneck Capacity Uncertainty"

Authors: Avi Steiner and Shlomo Shamai (Shitz)

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. The work is extended for an uncertain bottleneck channel capacity setting, where transmitter is not aware of the available backhaul capacity per transmission, but rather its capacity distribution. 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.