Distributed Extremization: Fundamental Limits & Practical Schemes

John MacLaren Walsh
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA
jwalsh@ece.drexel.edu

Thanks to AFOSR FA9550-12-1-0086 & NSF CCF-1053702.

Collaborators & Co-Authors – Mentioned Results

Solmaz Torabi,
Ph.D. Candidate
ASPITRG,
Drexel University

Interactive Distributed Function Computation for Massively Parallel Learning

Jie Ren, Ph.D.

Huawei USA

communication efficient control of wireless networks

Gwanmo Ku, Ph.D.
Senior Researcher
LG Electronics

Efficient Control Signaling for Resource Allocation in OFDMA Networks

Steven Weber, Ph.D.
Professor, Dept. of ECE
Drexel University

advisor for Bradford Boyle

Bradford Boyle, Ph.D.
Data Scientist/Developer
Sift Security

Achievable Schemes for Cost/Performance Tradeoffs in Networks

Collaborators & Co-Authors – This Paper

Solmaz Torabi,
Ph.D. Candidate
ASPITRG,
Drexel University

Interactive Distributed Function Computation for Massively Parallel Learning

Jie Ren, Ph.D.

Huawei USA

communication efficient control of wireless networks

Gwanmo Ku, Ph.D.
Senior Researcher
LG Electronics

Efficient Control Signaling for Resource Allocation in OFDMA Networks

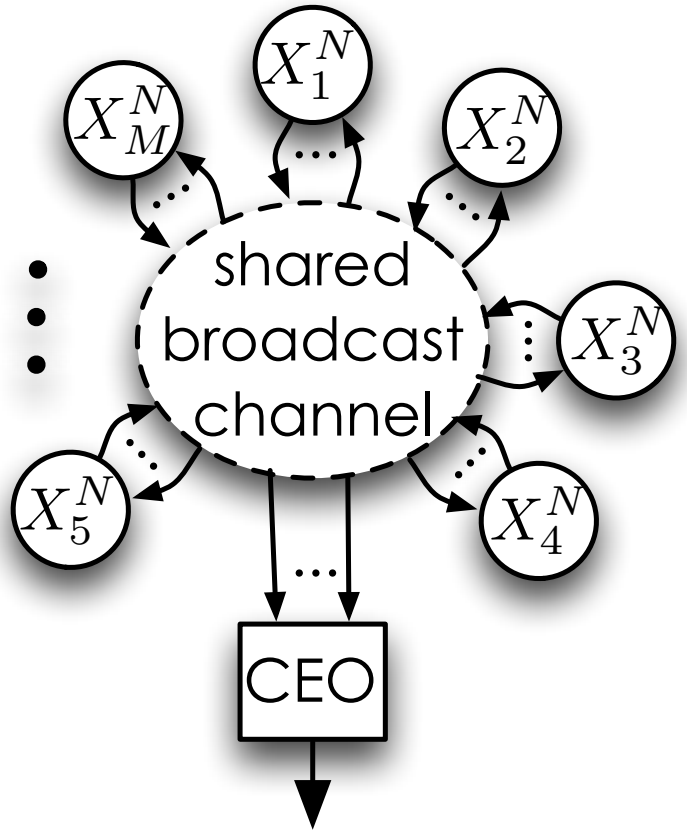
Steven Weber, Ph.D.
Professor, Dept. of ECE
Drexel University

advisor for Bradford Boyle

Bradford Boyle, Ph.D.
Data Scientist/Developer
Sift Security

Achievable Schemes for Cost/Performance Tradeoffs in Networks

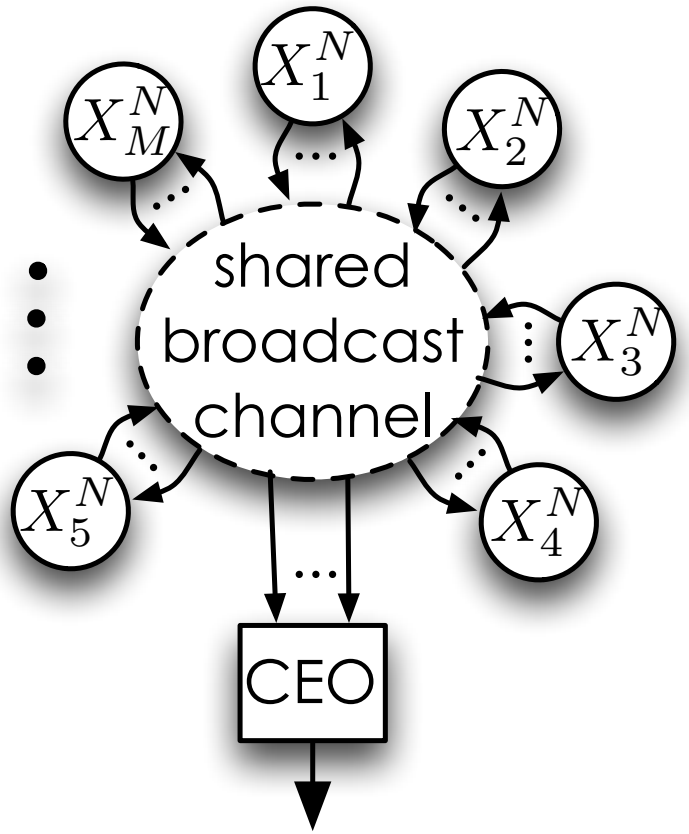
Distributed Extremization Problem



$$\approx f(X_1^{(n)}, \ldots, X_M^{(n)}),$$
$$n \in \{1, \ldots, N\}$$

- $X_m^{(n)}$ IID in m and n
- each user m wants to know 1 of:
  $\arg\max$ with loss $d(i, X_{1:M}^{(n)}) = \max_m X_m^{(n)} - X_i^{(n)}$
  $\max$ with loss $d(\hat{X}, X_{1:M}^{(n)}) =$
  $$\begin{cases} \max_m X_m^{(n)} - \hat{X} & \hat{X} \leq \max_m X_m^{(n)} \\ \max_m X_m^{(n)} & \text{otherwise} \end{cases}$$
- both $(\max, \arg\max)$ w/ loss $d(\hat{X}, i, X_{1:M}^{(n)}) =$
  $$\begin{cases} \max_m X_m^{(n)} - \hat{X} & \hat{X} \leq X_i^{(n)} \\ \max_m X_m^{(n)} & \text{otherwise} \end{cases}$$
- lossless vs. lossy
- 1-shot vs. interactive (colloc. or parallel)
- fundamental limit & achieve practically

Distributed Extremization Problem – ISIT 2016 Paper



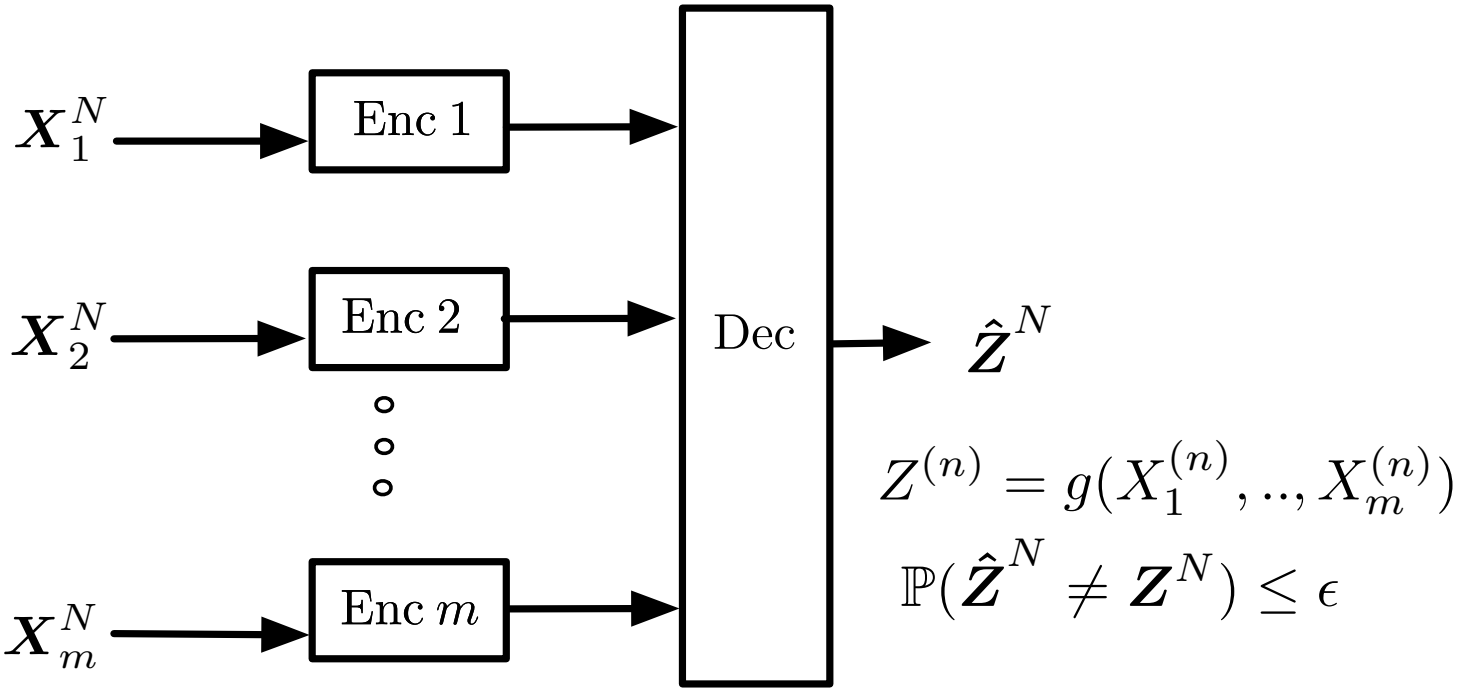
$$\approx f(X_1^{(n)}, \ldots, X_M^{(n)}),$$
$$n \in \{1, \ldots, N\}$$

- $X_m^{(n)}$ IID in m and n
- each user m wants to know 1 of:

  $\arg\max$ with loss $d(i, X_{1:M}^{(n)}) = \max_m X_m^{(n)} - X_i^{(n)}$

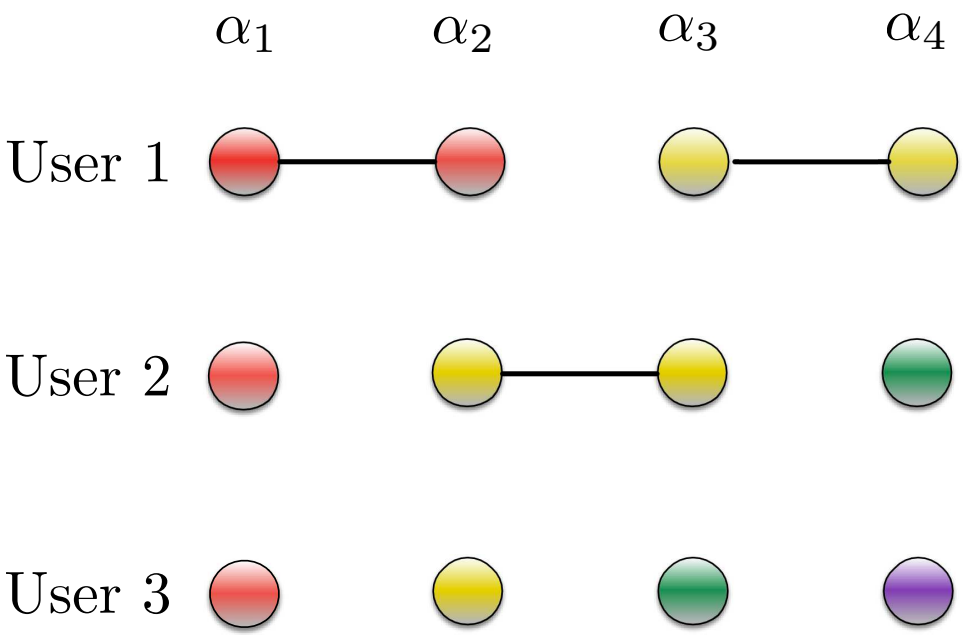
  $\max$ with loss $d(\hat{X}, X_{1:M}^{(n)}) =$
  $$\begin{cases} \max_m X_m^{(n)} - \hat{X} & \hat{X} \leq \max_m X_m^{(n)} \\ \max_m X_m^{(n)} & \text{otherwise} \end{cases}$$
- both $(\max, \arg\max)$ w/ loss $d(\hat{X}, i, X_{1:M}^{(n)}) =$
  $$\begin{cases} \max_m X_m^{(n)} - \hat{X} & \hat{X} \leq \max_m X_m^{(n)} \\ \max_m X_m^{(n)} & \text{otherwise} \end{cases}$$
- lossless vs. lossy
- 1-shot vs. interactive (colloc. or parallel)
- fundamental limit & achieve practically

Lossless One Shot Extremization



- Witsenhausen, Orlitsky & Roche, Sefidgaran & Tchamkerten
- Set up Characteristic Graph for each user
- Independence $\implies$ conditional graph entropy = graph entropy = chromatic entropy

Lossless One Shot Extremization



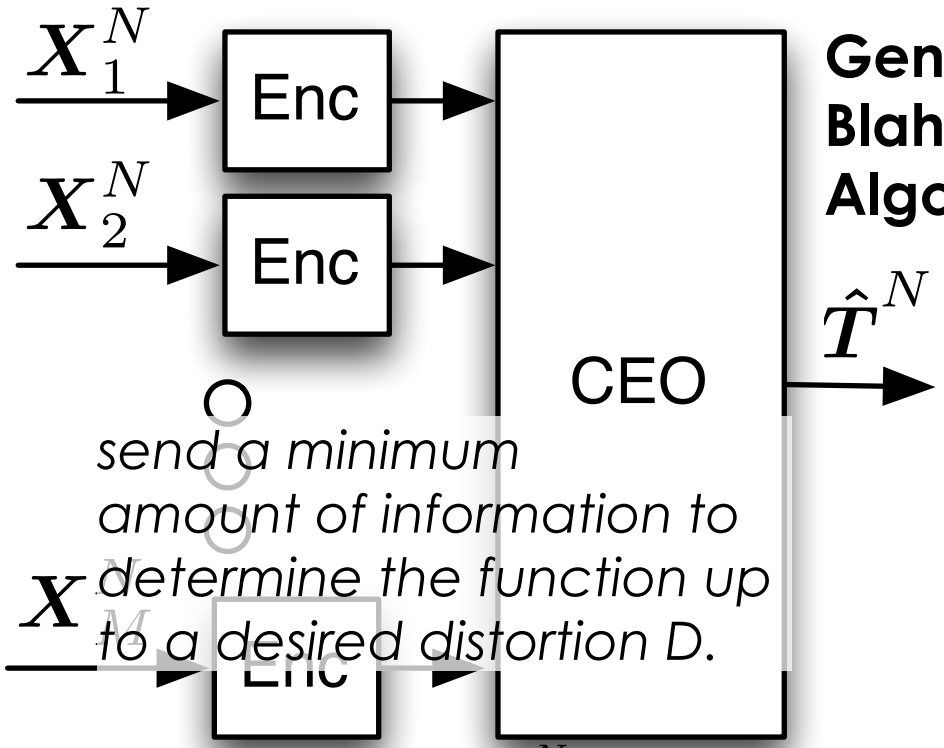
- complement of characteristic graphs
- $\arg\max$: Only adjacent vertices can be given same color at 2 users.
- Can save ≤ 2 bits relative to forwarding observations [Ren, Boyle, Ku, Weber, Walsh TIT '16]

Lossy One Shot

- Lossy variant of one shot problem: CEO problem
- CEO rate distortion region unknown except for
  - independent sources
  - Quadratic Guassian
- CEO for Indep. Sources (R,D) region can be calculated with an adaptation of the Blahut Arimoto Algorithm [Ku Ren Walsh TSP 2015]

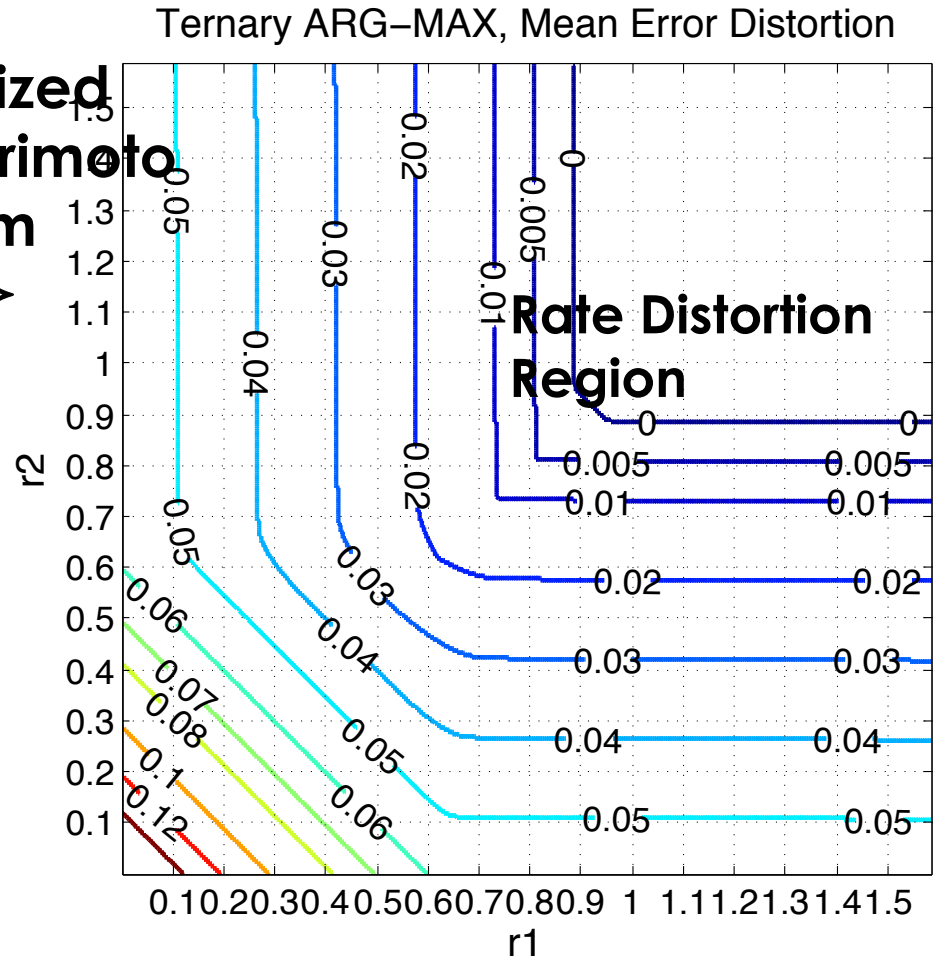
Lossy One Shot

Rate Regions for the CEO Problem with Independent Sources



send a minimum amount of information to determine the function up to a desired distortion D.

Generalized Blahut Arimoto Algorithm $\Rightarrow$

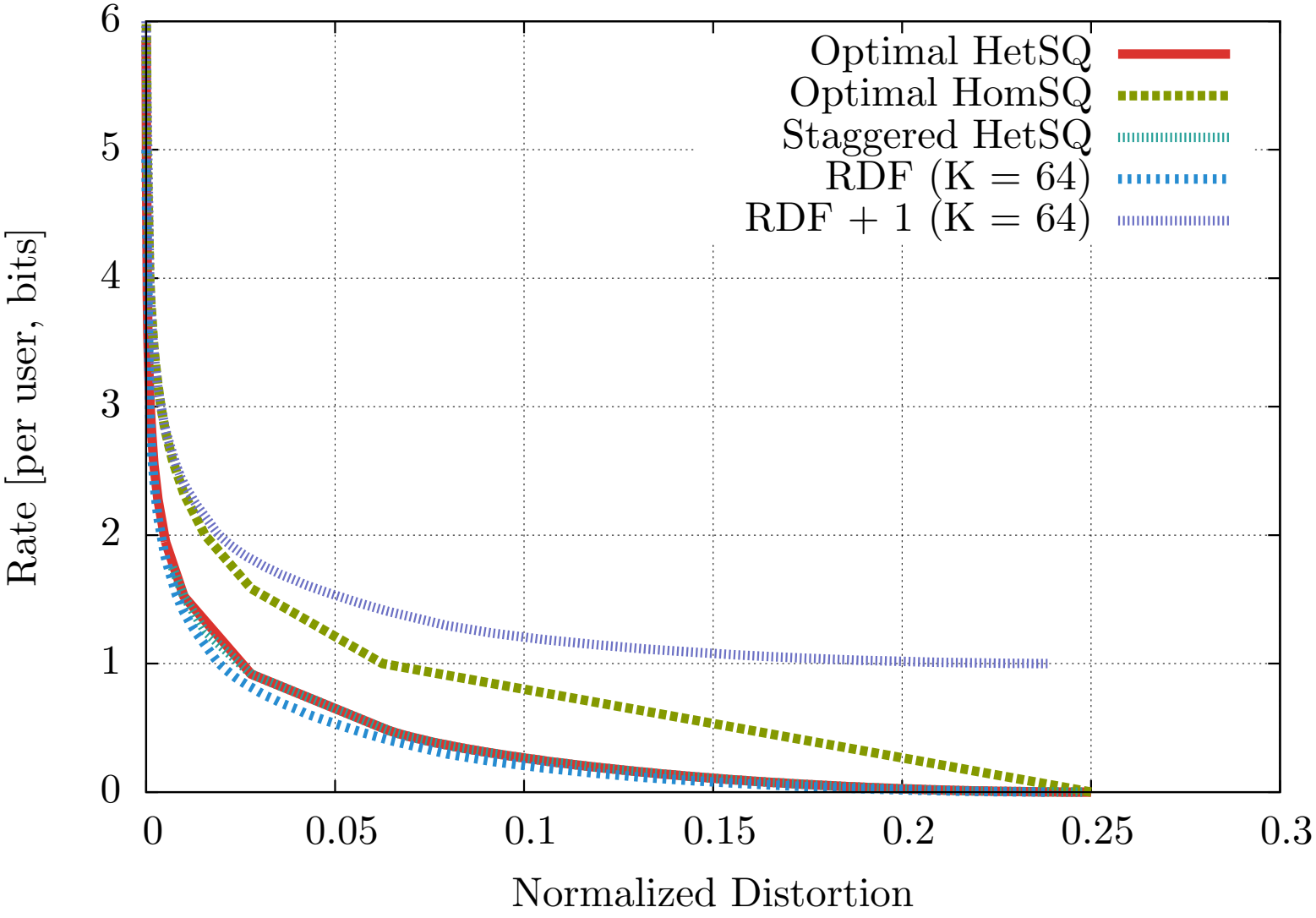


$$\mathbb{E}\left[d(\boldsymbol{T}^N, \hat{\boldsymbol{T}}^{\boldsymbol{N}})\right] = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[d(t^{(n)}, \hat{t}^{(n)})\right] < D$$

$$T_n = f(X_{1,n}, \ldots, X_{M,n})$$

[Ku Ren Walsh TSP 2015]

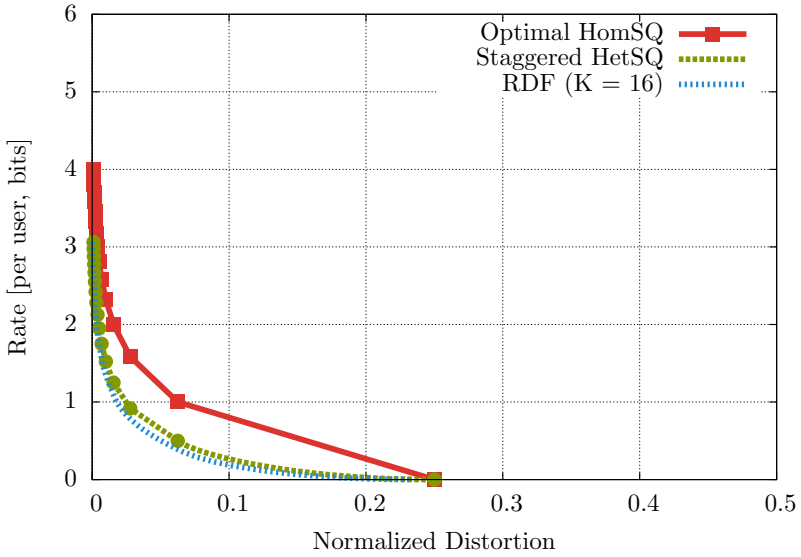
Lossy One Shot Extremization

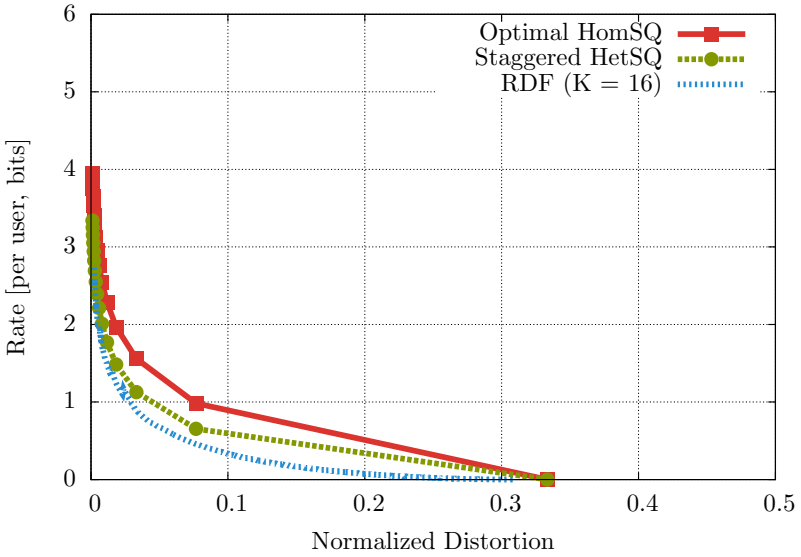


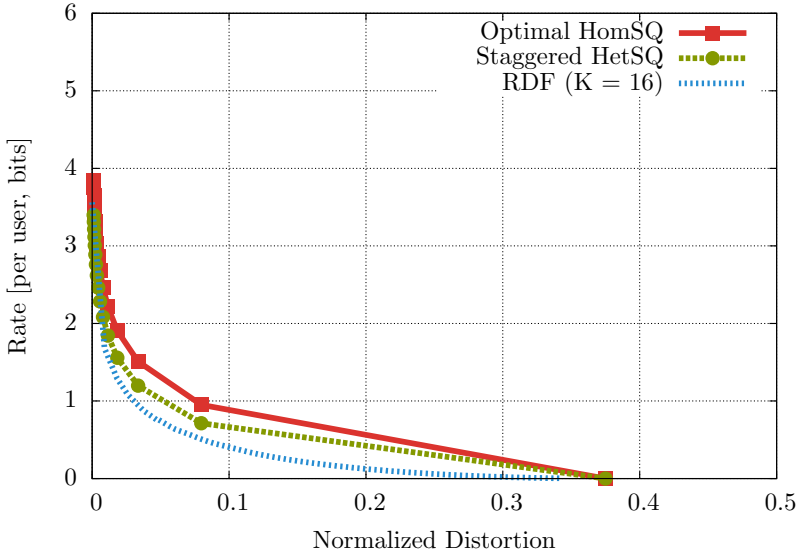
2 users, uniform, arg max

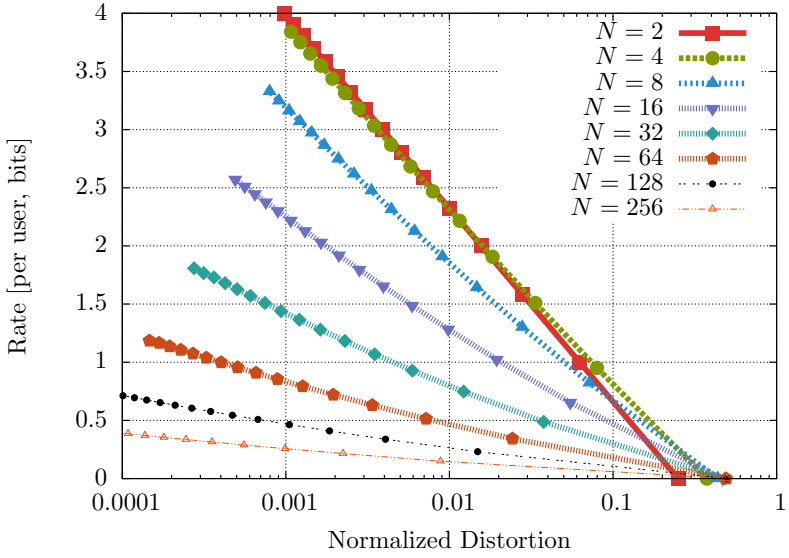
[Ren, Boyle, Ku, Weber, Walsh TIT '16]

Lossy One Shot Extremization





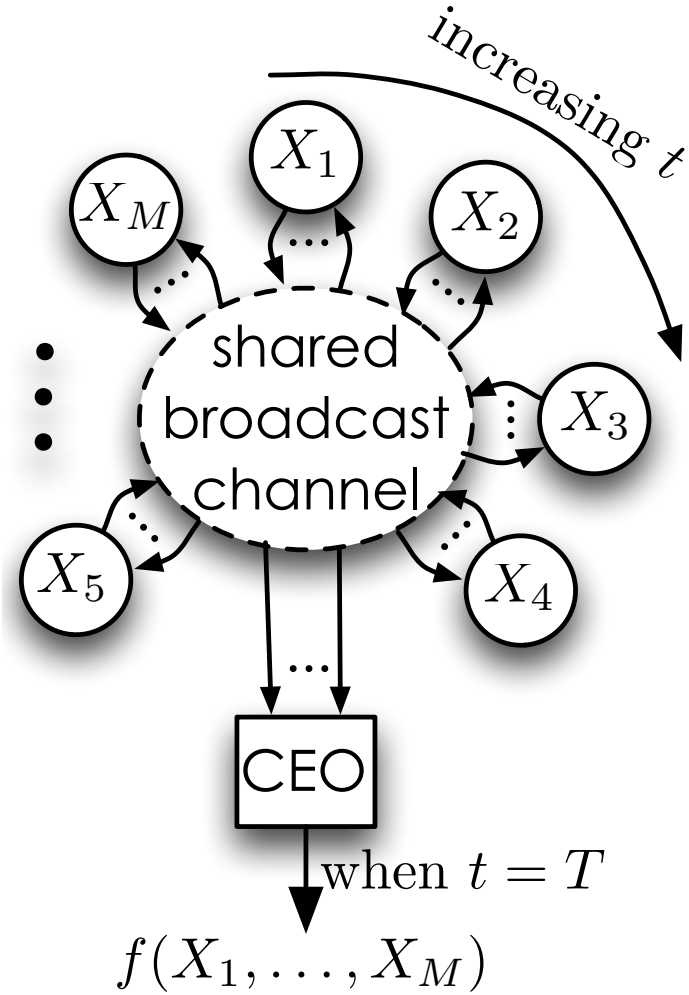


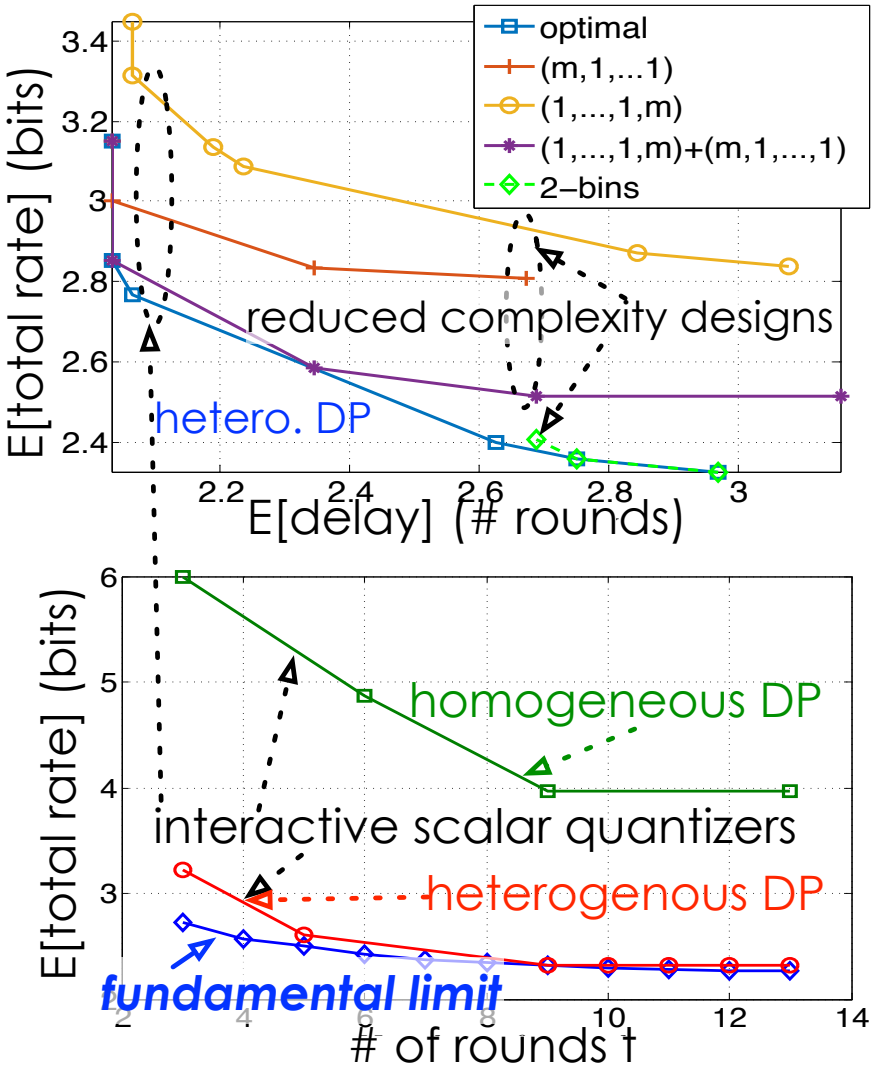


uniform, arg max, 2,3,4 users, varying number

[Ren, Boyle, Ku, Weber, Walsh TIT '16]

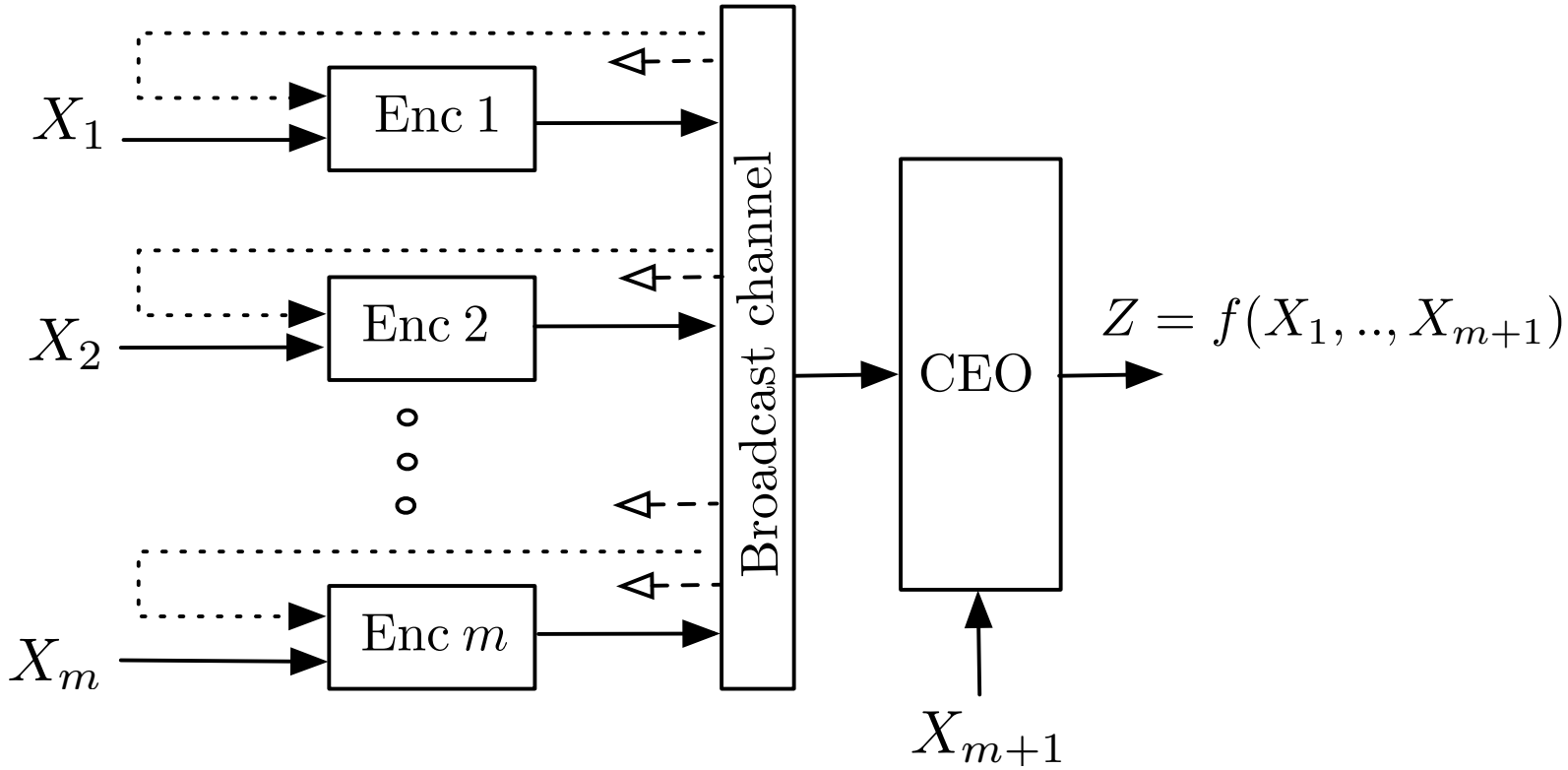
Lossless Interactive Extremization





Collocated Limit: [Ma, Ishwar, Gupta TIT]. Parallel: [Boyle TSP '16], Collocated: [Torabi ISIT '16]

Lossy Interactive



Quadratic Gaussian Interaction doesn't help: [Prabhakaran, Ramchandran, Tse Allerton '04]

Lossy Interactive [Torabi Allerton '16 sub]

Sources conditionally independent given side information.

Theorem 1: A

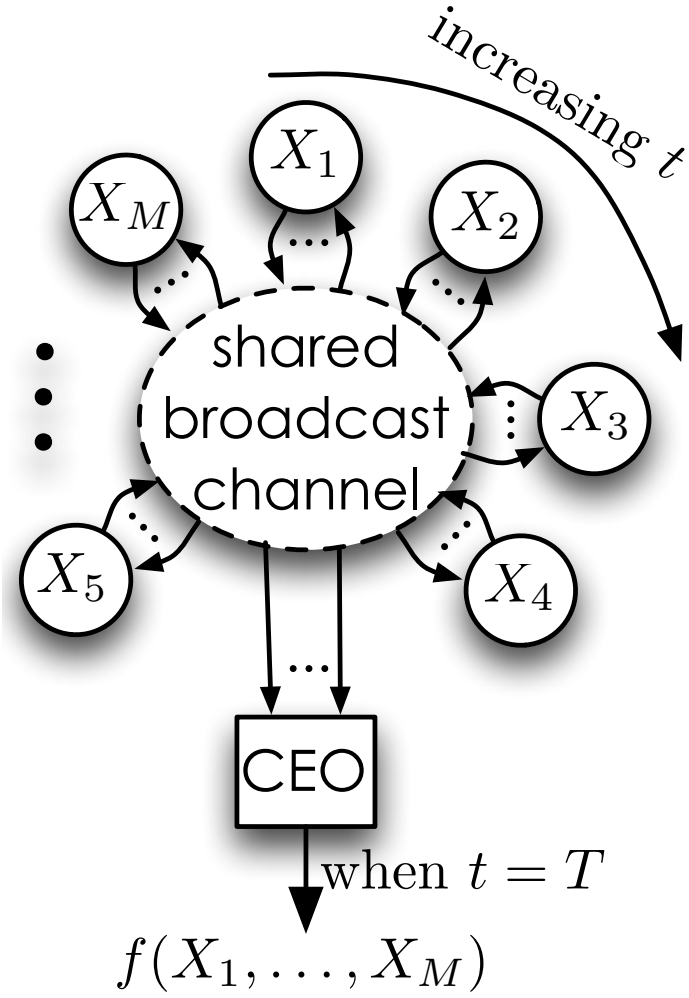
ny achievable $(\boldsymbol{R}, D)$ pair must lie in $\mathcal{RD}^t_{out}$ = the convex hull of the region

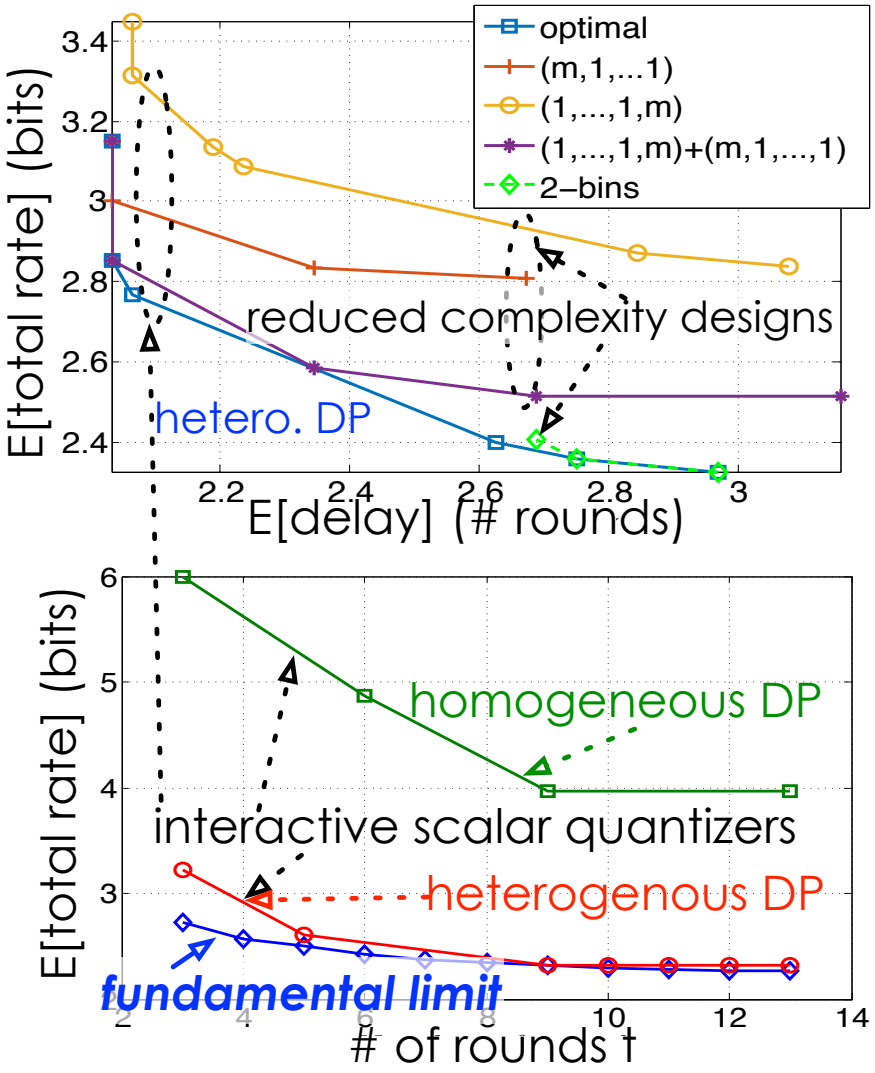
$$\left\{ (\mathbf{R}, D) \;\middle|\; \begin{array}{c} \forall i \in [t] \quad R_i \geq I(X_j; U_i | U_{1:i-1}, X_{m+1}), \\ U_i \leftrightarrow X_j U_{1:i-1} \leftrightarrow X_{[m+1]\setminus\{j\}}, \\ j = ((i-1) \bmod m) + 1, \\ \mathbb{E}\big[d(X_1, ., X_{m+1}, \hat{Z}(U_{1:t}, X_{m+1}))\big] \leq D \end{array} \right\}$$

Matching achievability scheme when:

- sources and side information are jointly independent.
- sources are conditionally independent of side information, and each $X_j = Z_{A_j}$ for some fixed $A_j \subset \{1, \ldots, K\}$ with $Z_k,\ k \in \{1, \ldots, K\}$ independent.

Lossless Interactive Extremization





Collocated Limit: [Ma, Ishwar, Gupta TIT]. Parallel: [Boyle TSP '16], Collocated: [Torabi DCC '16, Torabi ISIT '16 sub]