Reliable Adaptive Recoding for Batched Network Coding with Burst-Noise Channels

Jie Wang†, Talha Bozkus‡, Yao Xie†, Urbashi Mitra‡

† H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

‡ Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California

Personal Website: Preprint:

Introduction

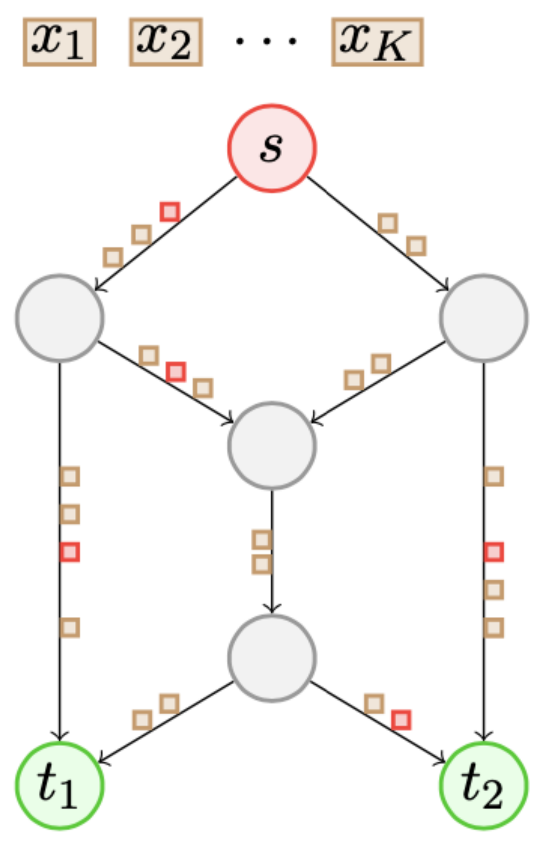
Networks with Packet Loss

A file transmission problem

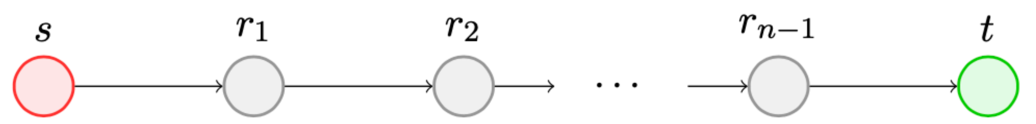
- $\#$ packets $= K$
- E.g., 2MB file $\approx$ 2,000 packets
- Single source node
- One/multiple destination nodes
- Non-negligible packet loss rate

A practical solution

- low computational and storage costs
- high transmission rate
- small protocol overhead



Line Networks of n Hops [yang2021capacity]

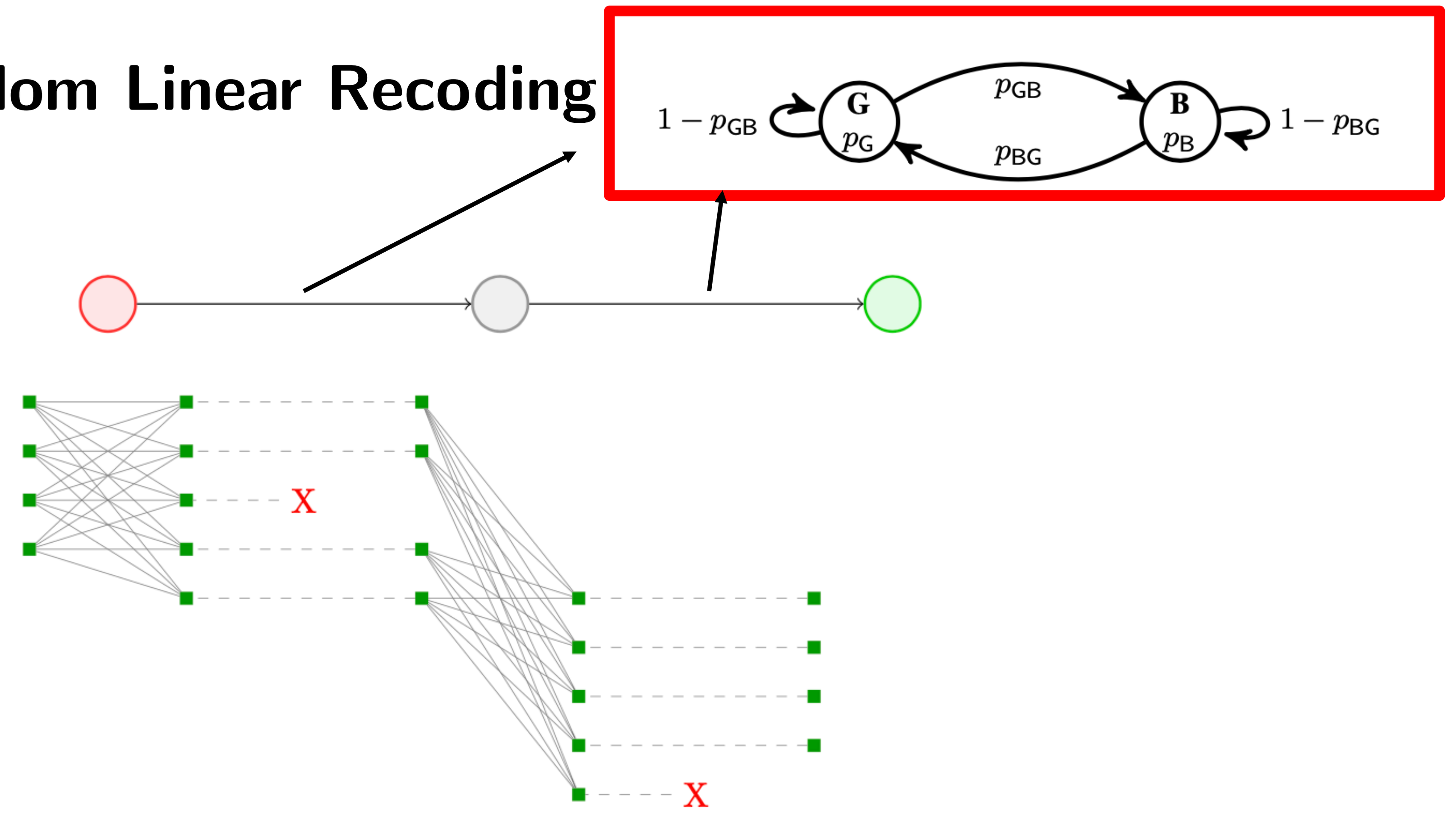


All links have a packet loss rate ϵ.

Intermediate Operation	Maximum Rate
forwarding	$(1-\epsilon)^n \to 0$
network coding	$1-\epsilon$

- [yang2021capacity]: Yang S, **Wang J**, Dong Y, et al. Capacity Scalability of Line Networks with Batched Codes. arXiv preprint arXiv:2105.07669, 2021.
- [yin2021unified]: Yin, H. H., Tang, B., Ng, K. H., Yang, S., Wang, X., & Zhou, Q. (2021). A unified adaptive recoding framework for batched network coding. IEEE Journal on Selected Areas in Information Theory.
- [wang2021small]: **Wang, J.**, Jia, Z., Yin, H. H., & Yang, S. (2021). Small-sample inferred adaptive recoding for batched network coding. In 2021 IEEE International Symposium on Information Theory.

Random Linear Recoding



Pros:

- Nearly optimal expected rank.

Cons:

- Highest recoding computation cost.
- Waiting for packets for recoding.

Adaptive Recoding [yin2021unified, wang2021small]

Idea: generate different number of received packets for batches of different ranks.

MDP Formulation

- Stage: The index of node, denoted as $\ell \in [L]$.
- State: The rank of the received batch, denoted as $s_\ell \in [M]$.
- Action: Number of recoded packets sent to the outcoming link, denoted as N_ℓ.
  Policy: $\pi_\ell(\cdot|s_\ell)$.
- Reward Function: At stage $\ell \in [L-1]$, the reward $r_\ell(s_\ell, N_\ell) = -\eta \cdot N_\ell$.
  At the final stage L, the reward $r_L(s_L) = s_L$.
- Batch-wise packet loss model: for $\ell \in [L-1]$, the probability that node ℓ transmits s packets while node $\ell+1$ receives s' packets equals $q_\ell(s' \mid s)$
- Transition dynamics:

$$P_\ell(s|s_\ell, N_\ell) = \begin{cases} 0, & \text{if } s_\ell < s, \\ \sum_{k=s}^{N_\ell} q_\ell(k|N_\ell)\zeta_s^{s_\ell,k}, & \text{if } s_\ell \geq s. \end{cases}$$

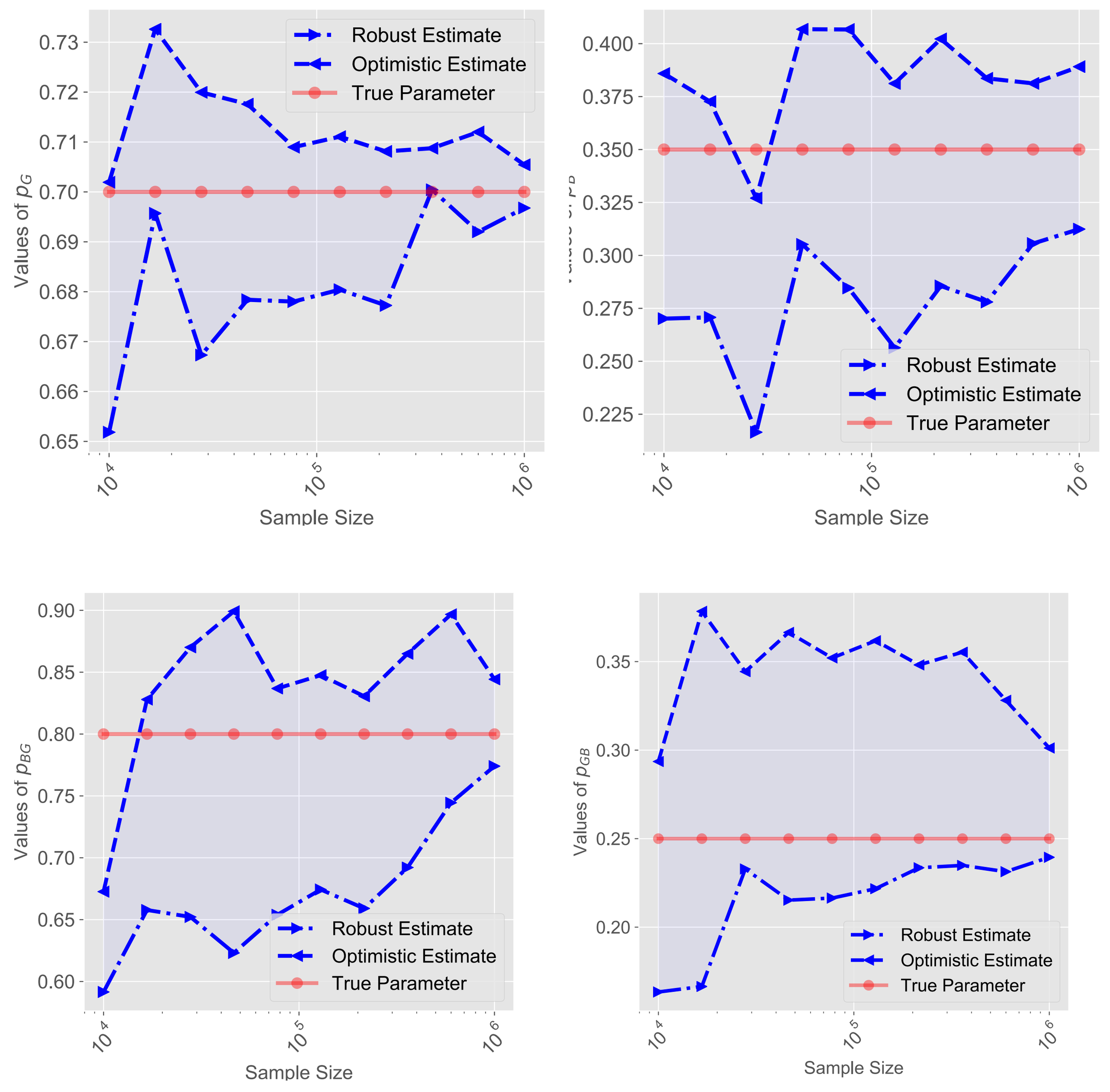
Here $\zeta_j^{i,k}$ is the probability that $i \times k$ size matrix with independent entries uniformly distributed over the field of size q has rank j.

Uncertainty Quantification of Channel Parameters

Unknown: Channel parameters $\theta := (p_G, p_B, p_{GB}, p_{BG})$.
Ground Truth Data: $\mathcal{D} = \{(X_j^i, Y_j^i)\}_{j\in[n], i\in[1:m]}$.
The i-th trajectory: $((X_0^i, Y_0^i), (X_1^i, Y_1^i), \ldots, (X_n^i, Y_n^i))$.
Here X_j^i is a *latent* variable, indicating whether the channel state is good ($X_j^i = 1$) or bad ($X_j^i = 0$), and Y_j^i is a observation variable indicating whether the j-th packet from i-th trajectory is successfully transmitted or not.
Observation: $\mathcal{D}_o = \{Y_j^i\}_{j\in[n], i\in[1:m]}$.
Point Estimation: Let $\widehat{\theta}_i$ be the EM algorithm estimator based on i-th trajectory.
Confidence Set Estimation: Since $n^{1/2}(\widehat{\theta}_i - \theta^*) \to$ Normal, we take $(1-\alpha)$-coverage confidence set

$$\Xi = \left\{\theta : \ (\theta - \overline{\theta})\widehat{\Sigma}^{-1}(\theta - \overline{\theta}) \leq \frac{T_{4,m-4}^2(1-\alpha)}{m}\right\}.$$

Numerical Study



- (a)-(d): Estimation for parameters p_G, p_B, p_{BG}, p_{GB}.
- Sample size: $m = 20, n \in [1e4, 1e6]$.
- Coverage probability: 95%