Noisy Network Coding

Abbas El Gamal
Stanford University

Joint work with Y-H. Kim (UCSD), S. H. Lim and S-Y. Chung (KAIST)

DARPA ITMANET, May 2010

Noisy Multicast Network

- Consider an $N$-node discrete memoryless multicast network (DM-MN) $(\mathcal{X}_1 \times \cdots \times \mathcal{X}_N, p(y_2, \ldots, y_N|x_1, \ldots, x_N), \mathcal{Y}_2 \times \cdots \times \mathcal{Y}_N)$

- Source node 1 wishes to send message $M$ to destination nodes $\mathcal{D}$

\[ p(y_1, \ldots, y_N|x_1, \ldots, x_N) \]
Noisy Multicast Network

- A \((2^{nR}, n)\) code for the DM-MN:
  - Encoder: \(x_1^n(m)\) for each message \(m \in [1 : 2^{nR}]\)
  - Relay encoder \(j \in [2 : N]\): \(x_{ji}(y_{ji}^{i-1})\) for each \(y_{ji}^{i-1} \in \mathcal{Y}_j^{i-1}, i \in [1 : n]\)
  - Decoder \(k \in D\): \(\hat{m}_k(y_k^n)\) for each \(y_k^n \in \mathcal{Y}_k^n\)

- The average probability of error \(P_{e(n)} = P\{\hat{M}_k \neq M \text{ for some } k \in D\}\)

Noisy Multicast Network

- Rate \(R\) achievable if there exists a sequence of codes with \(P_{e(n)} \rightarrow 0\)
- The capacity \(C\) of the DM-MN is supremum of achievable rates
- Capacity is not known in general
- There are upper and lower bounds that coincide in some special cases
Cutset Upper Bound

\[ C \leq \max_{p(x^N)} \min_{k \in D} \min_{S:1 \in S, k \in S^c} I(X(S); Y(S^c)|X(S^c)) \]

- \( X(S) \) inputs in \( S \); \( X(S^c), Y(S^c) \) inputs/outputs in \( S^c \)

Capacity Results and Lower Bounds

- Shannon (1948) established the capacity of noisy point-to-point channel using random coding
- Ford, Fulkerson (1956) (also Elias, Feinstein, Shannon) established the capacity of noiseless unicast network using forwarding
- Ahlswede, Cai, Li, Yeung (2000) established the capacity of noiseless multicast network using network coding
- Network coding extended to multi-source multicast erasure network by Dana, Gowaikar, Palanki, Hassibi, Effros (2006)
- Network coding extended to obtain lower bound on capacity of multicast deterministic network by Avestimehr, Diggavi, Tse (2007)
- In earlier development, Cover, EG (1979) developed compress–forward scheme for the relay channel

- EG, Kim (Lecture Notes on NIT 2009) developed a noisy network coding scheme that unifies and extends above results (Lim, Kim, Chung, EG ISIT 2010, WiNC 2010)
Noiseless Multicast Networks

Consider noiseless network modeled by graph \((\mathcal{N}, \mathcal{E})\)

Node 1 wishes to send \(M\) to set of destination nodes \(D\)
Capacity region coincides with cutset bound

Network Coding Theorem (Ahlswede, Cai, Li, Yeung 2000)

\[
C \leq \min_{k \in D} \min_{S} \min_{1 \in S, k \in S^c} C(S)
\]

Outline of Proof: Acyclic Network

Wolog assume zero node delay
Use block coding (assume \(C_{jk}\) are integer valued)
Random codebook generation:
\[f_{jk} \in [1 : 2^{nC_{jk}}], \ (j, k) \in \mathcal{E}, \text{ and } f_4 \text{ are randomly and independently generated, each according to uniform pmf}
\]
Key step: If \(R < \min_S C(S)\), \(f_4(m)\) is one-to-one with high prob.
Koetter, Medard (2003) showed that cutset bound can be achieved with zero error using linear network coding
Outline of Proof: Cyclic Network

- Cannot assume zero delay nodes. Assume unit delay at each node.
- Unfold to time extended (acyclic) network with \( b \) blocks.
- **Key step:** Min-cut capacity of the new network is \( \approx bC \) for \( b \) large.
- By result for acyclic case, min-cut capacity for the new network is achievable.
- Need to send **same message** \( b \) times using independent mappings.

Multicast Erasure Network

- Source nodes wish to send their messages to destination nodes \( D \).
- Link failure is observed as an erasure symbol; destination nodes have access to network erasure pattern (includes noiseless case).
- **Capacity region coincides with cutset bound** and achieved via network coding (Dana, Gowaikar, Palanki, Hassibi, Effros 2006).
Deterministic Multicast Network

- Generalizes noiseless multicast network with broadcast, interference

Node 1 wishes to send message to subset of nodes $D$
- Node 1 sends $x_{1i}(m)$ and node $j$ sends $x_{ji}(y_{ij}^{i-1})$ at time $i \in [1 : n]$
- Capacity is not known in general
- Cutset upper bound reduces to

$$C \leq \max_{p(x^N)} \min_{S:1 \in S, k \in S^c} \min_{k \in D} H(Y(S^c)|X(S^c))$$

Lower bound on capacity (Avestimehr, Diggavi, Tse 2007)

$$C \geq \max_{\prod_{j=1}^N p(x_j)} \min_{k \in D} \min_{S:1 \in S, k \in S^c} \min_{k \in D} H(Y(S^c)|X(S^c))$$

- Cutset bound:

$$C \leq \max_{p(x^N)} \min_{S:1 \in S, k \in S^c} \min_{k \in D} H(Y(S^c)|X(S^c))$$

- Bounds coincide for:
  - No interference (Ratnakar, Kramer 2006):
    $$Y_k = (y_{k1}(X_1), \ldots, y_{kN}(X_N)), \ k \in [2 : N]$$
  - Finite-field network (Avestimehr, Diggavi, Tse 2007):
    $$Y_k = \sum_{j=1}^N g_{jk} X_j \text{ for } g_{jk}, X_j \in \mathbb{F}_q, \ j \in [1 : N], \ k \in [2 : N]$$

  Used to approximate capacity of Gaussian networks in high SNR
Outline of Proof

- Layered networks:

  - Random codebook generation: Randomly and independently generate $x^n_j(y^n_j)$ for each sequence $y^n_j$
  - Key step: If $R$ satisfies lower bound, end-to-end mapping is one-to-one with high probability

- Non-layered network:
  - Construct time extended (layered) network with $b$ blocks
  - Key step: If $R$ satisfies lower bound, end-to-end mapping is one-to-one with high probability
  - Again send the same message $b$ times using independent mappings

Relay Channel

- The relay channel (van der Meulen 1971) is a 3-node DMN

  - Node 1 wishes to send $M$ to node 3 with help of node 3 (relay)
  - Capacity is not known in general
  - Cutset upper bound reduces to (Cover, EG 1979)

$$C \leq \max_{p(x_1, x_2)} \min \{I(X_1, X_2; Y_3), I(X_1; Y_2, Y_3|X_2)\}$$
Compress–Forward Lower Bound

$C \geq \max_{p(x_1)p(x_2)p(y_2|x_2)} I(X_1; \hat{Y}_2, Y_3|X_2)$

subject to $I(X_2; Y_1) \geq I(Y_2; \hat{Y}_2|X_2, Y_3)$

Outline of Proof

- Send $b - 1$ independent messages over $b, n$-transmission blocks
- At the end of block $j$, relay chooses description $\hat{y}_2^n(j)$ of $y_2^n(j)$
- Since the receiver has side information $y_3^n(j)$ about $\hat{y}_2^n(j)$, we use Wyner–Ziv coding to reduce rate necessary to send $\hat{y}_2^n(j)$
- The bin index is sent to the receiver in block $j + 1$ via $x_2^n(j + 1)$
- At the end of block $j + 1$, the receiver first decodes $x_2^n(j + 1)$ from which it finds $\hat{y}_2^n(j)$
- It then finds unique $\hat{m}_j$ such that $(x_1^n(\hat{m}_j), x_2^n(j), \hat{y}_2^n(j), y_3^n(j))$ are jointly typical
Equivalent Compress–Forward Lower Bound

Compress–forward lower bound (EG, Mohseni, Zahedi 2006)

\[ C \geq \max_{p(x_1)p(x_2)p(\hat{y}_2|y_2,x_2)} \min \{ I(X_1, X_2; Y_3) - I(Y_2; \hat{Y}_2|X_1, X_2, Y_3), I(X_1; \hat{Y}_2, Y_3|X_2) \} \]

- Compress–forward lower bound (Cover, EG 1979):
  \[ C \geq \max_{p(x_1)p(x_2)p(\hat{y}_2|y_2,x_2)} I(X_1; \hat{Y}_2, Y_3|X_2) \]
  subject to \( I(X_2; Y_3) \geq I(Y_2; \hat{Y}_2|X_2, Y_3) \)

- Cutset bound:
  \[ C \leq \max_{p(x_1,x_2)} \min \{ I(X_1, X_2; Y_3), I(X_1; Y_2, Y_3|X_1) \} \]

This characterization generalizes naturally to networks

Generalization yields strictly higher rates than extension of original characterization by Kramer, Gastpar, Gupta (2005)

Noisy Multicast Network

Main Result: Noisy Network Coding Lower Bound

- The alternative characterization of compress–forward lower bound for relay channel generalizes to discrete memoryless multicast network

Theorem (EG, Kim Lecture on NIT 2009)

\[ C \geq \max \min_{k \in \mathcal{D}} \min_{S \subseteq \{1:N\}} \min_{1 \in \mathcal{S}, k \in \mathcal{S}^c} \left( I(X(S); \hat{Y}(S^c), Y_k|X(S^c)) - I(Y(S); \hat{Y}(S)|X^N, \hat{Y}(S^c), Y_k) \right), \]

where the maximum is over \( \prod_{k=1}^N p(x_k)p(\hat{y}_k|y_k, x_k) \)

- Includes as special cases:
  - Capacity of noiseless multicast networks
  - Lower bound on deterministic multicast networks
  - Capacity of wireless erasure multicast networks

- Shows that network coding is a special case of compress–forward 😊
- Simpler and more general proof (deals directly with cyclic networks)
Proof Outline

- Source node sends **same message** \( b \) times; relays use compress–forward; decoders use simultaneous decoding
- No binning; don’t require decoding compression indices correctly!
- For simplicity, consider proof for relay channel

The relay uses independently generated compression codebooks:
\[
\mathcal{B}_j = \{ \hat{y}_2^n(l_j|l_{j-1}) : l_j, l_{j-1} \in [1 : 2^{nR_2}] \}, \ j \in [1 : b]
\]

\( l_{j-1} \) is compression index of \( \hat{Y}_2^n(j-1) \) sent by relay in block \( j \)

The senders use independently generated transmission codebooks:
\[
\mathcal{C}_j = \{(x_1^n(j, m), x_2^n(l_{j-1})) : m \in [1 : 2^{nbR}], l_{j-1} \in [1 : 2^{nR_2}] \}
\]

**Encoding**: Sender transmits \( X_1^n(j, m) \) in block \( j \in [1 : b] \)

Upon receiving \( Y_2^n(j) \) and knowing \( X_2^n(l_{j-1}) \), the relay finds jointly typical \( \hat{Y}_2^n(l_j|l_{j-1}) \), and sends \( X_2^n(l_j) \) in block \( j + 1 \)

**Decoding**: After receiving \( Y_3^n(j), \ j \in [1 : 2^{nR}] \), the receiver finds unique \( \hat{m} \) such that:

\[
(x_1^n(j, \hat{m}), \hat{y}_2^n(l_j|l_{j-1}), x_2^n(l_{j-1}), y_3^n(j)) \text{ are jointly typical for all } j \in [1 : b] \text{ and for some } l_1, l_2, \ldots, l_b
\]
Analysis of the Probability of Error

- Assume $M = 1$, $L_1 = L_2 = \cdots = L_b = 1$, and let
  \[ \mathcal{E}_j(m, l_{j-1}, l_j) = \{(X_1^n(j, m), \hat{Y}_2^n(l_{j-1}), X_2^n(l_{j-1}), Y_3^n(j)) \in \mathcal{T}^{(n)}\} \]

- The average probability of error

\[
P(\mathcal{E}) \leq P\left(\bigcup_{j=1}^b \mathcal{E}_j^c(1, 1, 1)\right) + P\left(\bigcup_{m \neq 1} \bigcap_{j=1}^b \mathcal{E}_j(m, l_{j-1}, l_j)\right)
\]

\[
\leq \sum_{j=1}^b P(\mathcal{E}_j^c(1, 1, 1)) + \sum_{m \neq 1} \prod_{j=2}^b P(\mathcal{E}_j(m, l_{j-1}, l_j))
\]

- If $m \neq 1$ and $l_{j-1} = 1$,
  \[ P(\mathcal{E}_j(m, l_{j-1}, l_j)) \leq 2^{-n(I(X_1; \hat{Y}_2, Y_3|X_2) - \delta(\epsilon))} \]

- If $m \neq 1$ and $l_{j-1} \neq 1$,
  \[ P(\mathcal{E}_j(m, l_{j-1}, l_j)) \leq 2^{-n(I(X_1, X_2; Y_3) + I(\hat{Y}_2; X_1, Y_3|X_2) - \delta(\epsilon))} \]

- The rest of the proof is just algebra 😊

---

Extension: Noisy Multi-source Multicast Network

Noisy network coding generalizes to this case
(Lim, Kim, El Gamal, Chung ITW 2010)

Extends result on erasure networks (Dana, Gowaikar, Palanki, Hassibi, Effros 2006) and deterministic networks (Perron 2009)
Extension: Noisy Multi-unicast Network

- Noisy network coding extends naturally to multi-unicast networks
- For example, consider an $N$-node DMN where node 1 wishes to send $M_j$ to node 3, and node 2 wishes to send $M_j$ to node 4
- Using noisy network coding, we view the network as an interference channel with senders $X_1$ and $X_2$ and respective receivers $(Y_3, \hat{Y}_5, \hat{Y}_6, \ldots, \hat{Y}_N)$ and $(Y_4, \hat{Y}_5, \hat{Y}_6, \ldots, \hat{Y}_N)$
- We use coding strategies for interference channel (ISIT 2010):
  - Each receiver decodes only its message (treats interference as noise)
  - Each receiver decodes both messages
  - One receiver uses the former strategy, the other uses the later
- Each relay can generate different $\hat{Y}$ for each destination node (WiNC 2010)

How Good Is Noisy Network Coding

- In noisy network coding we compress signals and retransmit them
- This has the advantage that relays don’t need to know codebooks
- It is also a good strategy in “high SNR” (e.g., noiseless links)
- However, noisy network coding is not always a good strategy
  - Consider cascade of noisy channels:
    - Optimal strategy is for each relay to use decode–forward
  - This is also what we do in wireless mesh networks

- Another strategy is amplify–forward (analog-to-analog)
Example 1: AWGN Relay Channel

- **Decode-forward**: within $1/2$ bit of cutset bound
- **Compress-forward**: within $1/2$-bit of cutset bound (Chang, Chung, Lee 2008)
- **Amplify-forward**: within $1$-bit of cutset bound (Chang, Chung, Lee 2008)
Example 2: AWGN Two-Way Relay

- Two-way relay channel is a 3-node DMN

\[ \frac{M_1}{M_2} \xrightarrow{(X_1 : Y_1)} p(y_1, y_2, y_3|x_1, x_2, x_3) \xrightarrow{(X_2 : Y_2)} \frac{\hat{M}_1}{\hat{M}_2} \]

- Node 1 wishes to send message \( M_1 \) to node 2
- Node 2 wishes to send message \( M_2 \) to node 1
- AWGN two-way relay:

\[ Y_k = \sum_{j \neq k} g_{jk} X_j + Z_k \text{ for } k = 1, 2, 3, \]

where \( Z_k \sim N(0, 1) \). Power constraint \( P \) on every sender.

Example 2: AWGN Two-Way Relay Channel

- Extensions of decode–forward, compress–forward, and amplify–forward compared by Rankov, Wittneben (2006) and Katti, Maric, Goldsmith, Katabi, Medard (2007) among others

Node 1 to 2 distance: 1; node 1 to 3 distance: \( d \in [0, 1] \): \( g_{13} = g_{31} = d^{-3/2} \), \( g_{23} = g_{32} = (1 - d)^{-3/2} \)
Example 2: AWGN Two-Way Relay Channel

- NNC is within 1 bit of cutset bound for all channel gains
- Gap unbounded for all other schemes

\[ g_{12} = g_{21} = 0.1, \ g_{13} = g_{32} = 0.5, \ g_{23} = g_{31} = 2 \]

![Graph showing sum rate vs. power](image)

Example 3: Multi-source Multicast Gaussian Network

- Channel model: \( Y^N = GX^N + Z^N \)
- The cutset bound yields

\[
\sum_{j \in S} R_j \leq \frac{1}{2} \log \left| I + \frac{P}{2} G(S)G(S)^T \right| + \frac{\min\{|S|, |S^c|\}}{2} \log(2|S|)
\]

- With \( \hat{Y}_j = Y_j + \hat{Z}_j, \ \hat{Z}_j \sim N(0, 1) \), noisy network coding bound yields

\[
\sum_{j \in S} R_j < \frac{1}{2} \log \left| I + \frac{P}{2} G(S)G(S)^T \right| - \frac{|S|}{2}
\]

- Thus, noisy network coding is optimal within \((N/4) \log(2N)\) bits/transaction for \( N > 3 \)
- This generalizes and improves single-source result by Avestimehr, Diggavi, Tse (2007)
- 2-way relay: Unbounded gap for decode–forward / amplify–forward
Example 4: AWGN Interference Relay Channel

- Consider a 2-user pair interference channel with a relay:
  \[ Y_k = g_{1k}X_1 + g_{2k}X_2 + Z_k \quad \text{for} \quad k = 3, 4, 5 \]
  \[ Z_k \sim N(0, 1) \] and power constraint \( P \) on each sender

\[ g_{14} = g_{25} = 1, g_{15} = g_{24} = g_{13} = 0.5, g_{23} = 0.1, R_0 = 1 \]

(a) treating interference as noise  \quad (b) decoding both messages
Conclusion

- Network coding and its generalizations to erasure and deterministic networks are special cases of compress–forward
- Noisy network coding lower bound can be extended to non-multicast messaging requirements
- Many interesting areas to explore
  - Applications of noisy network coding to wireless networks, . . .
  - Combining noisy network coding with (partial) decode (compute)–forward
- To learn more:
  Lecture notes on NIT at: http://arxiv.org/abs/1001.3404
  Papers at: http://arxiv.org/abs/1002.3188, ISIT 2010, WiNC 2010