On the interaction between network coding and the physical layer

Muriel Médard

ISWCS 2013.
Collaborators

- MIT: Georgios Angelopoulos, Anantha Chandrakasan, Flavio Du Pin Calmon, Nadia Fawaz (now Technicolor), Kerim Fouli, Minji Kim (now Oracle), Marie-Jose Montpetit, Arun Paidimarri, Ali ParandehGheibi (now Plexxi), Shirley Shi (now Ropes and Gray), Jay-Kumar Sundararajan (now Qualcomm), Surat Teerapittayanon (now Harvard)
- Caltech: Michelle Effros
- TUM: Ralf Kötter, Mohit Thakur
- Stanford: Andrea Goldsmith, Ivana Maric (now Ericsson)
- Rutgers: Ivan Seskar
- BBN: Abhimanyu Gosain
Regimes of SNR

• Information theoretic insights: we consider three regimes:
  – Low SNR: noise limited
    • Separate PHY coding from network coding
  – High SNR: interference limited
    • Analog network coding
    • Discrete approximations
  – Other SNRs:
    • Use equivalence theory for bounds

• Practical considerations:
  • Consider insight from separation to provide practical approaches – WIMAX case study
  • Low-power chip

• Joint PHY and network coding may be limited in usefulness
Low-SNR Approximation

• **Broadcast:**
  – Superposition coding rates $\sim$ time-sharing rates
  – Common rate received by both destinations rate received only by the most reliable destination

• **Multiple access**
  – No interference, FDMA
  – Both sources achieve same rate as in the absence of the other user
Perils of Virtual MIMO

- SIMO bound is loose in low SNR
- Any given quantization level is insufficient to transmit an uncoded, still noisy version of the data
- Example: in relay network SIMO bound is loose
- At low SNR, network becomes equivalent to a set of edges and hyperedges, with PHY-layer decoding and linear network coding
What Min-cut?

- Open question: Can the gap to the cut-set upper-bound be closed?
- An $\infty$ capacity on the link R-D would be sufficient to achieve the cut like in SIMO [Kramer et al ‘05, ’06]

- In the limit of a large bandwidth, if the relay cannot decode, large noise power and finite R – D link capacity render the relay contribution useless. **Cannot reach the SIMO cut-set upper-bound** [Fawaz, M. ‘11]

- Proof relies on rate-distortion theory and equivalence theory

- For physically degraded BC or when the source uses the channel as such, peaky binning relaying is optimal [Fawaz, M. ‘11]

- Optimum is then selective decode and forward – network is indeed a set of hyperedges
Perils of Virtual MIMO

- In *dense* networks, at high SNRs, SINRs are low.
  - Spectrum segmentation to avoid interference, requires infinite bandwidth
  - Therefore, does not scale
- In *extended* networks, SIMO bound does not hold in low SNR

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Dense</th>
<th>Extended</th>
<th>Stage Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantization &amp; SNR</td>
<td>✓</td>
<td></td>
<td>Stage 3</td>
</tr>
<tr>
<td>Scale Invariance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spectrum Segmentation</td>
<td>✓</td>
<td></td>
<td>Stage 1 &amp; 3</td>
</tr>
</tbody>
</table>

[Zeger, M. ‘13]
Practical Implications

- Achievable hypergraph model: Superposition coding, FDMA.
- Multicommodity flow optimization => Linear program for simple costs (network power, linear cost functions etc.).
- Separable dual => decentralized solutions.
- Hypergraph model can be used to design wireless networks by placing relays [Thakur, M. ‘10, Thakur, Fawaz, M. ’11, ‘12]
- Allows interesting geometric programming with results close to optimum
High SNR

- Open problem: capacity & code construction for wireless relay networks
  - Channel noise
  - Interference

- [Avestimehr et al. ‘07]“Deterministic model” (ADT model)
  - Interference
  - Does not take into account channel noise
  - In essence, high SNR regime of the Cover-Wyner region
  - Separation of network coding and underlying physical channel
  - Loss of 0.5 bits/s/Hz

Model as error free links

\[
Y(e_3) = \beta_1 Y(e_1) + \beta_2 Y(e_2)
\]
ADT Network Model

- Original ADT model:
  - Broadcast: multiple edges (bit pipes) from the same node
  - Interference: additive MAC over binary field – [Effros et al ‘04]

- Algebraic model:

Possible “codes” at $e_{12}$, which represents the MAC constraint

<table>
<thead>
<tr>
<th>$c$</th>
<th>$f$</th>
<th>$c+f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$c$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$f$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$c+f$</td>
</tr>
</tbody>
</table>
• Linear operations
  – Coding at the nodes $V$: $\beta(e_j, e_{j'})$
    – $F$ represents physical structure of the ADT network
    – $F^k$: non-zero entry = path of length $k$ between nodes exists
    – $(I-F)^{-1} = I + F + F^2 + F^3 + \ldots$ : connectivity of the network
      (impulse response of the network)

$$F = \begin{pmatrix}
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}$$

- Broadcast constraint (hyperedge)
- MAC constraint (addition)
- Internal operations (network code)
Algebraic Connection

- [Avestimehr et al. ’07] requires optimization over a large set of matrices
- [Kim and M. ‘10] ADT network can be expressed with Algebraic Network Coding Formulation [Koetter and M. ’01, ‘02, ’03]:
  - Model broadcast constraint with **hyper-edge**
  - Rank of single **system matrix** \( M \) maps to **physical** min-cut of hypergraph
- Prove an algebraic definition of \( \text{min-cut} = \text{rank}(M) \)
- Prove Min-cut Max-flow for unicast/multicast holds
- Extend optimality of linear operations to **non-multicast** sessions
- Show that **random linear network coding** achieves capacity
- Incorporate **failures, random erasures** [Lun et al ‘08, Dana et al ‘05] and **delay** (allows cycles within the network) [Koetter and M. ‘02, ’03]
SNR in Networks

- High SNR in a link
  - Noise → 0
  - Large gain
  - Large transmit power

- Consider diamond network [Schein, Gallager’ 01]

- Gain:
  - increase a [Avestimehr et al ’07]

- Large transmit power
  - Amplify-and-forward in the network, ignorant of topology
  - Asymptotically optimal
Analog Network Coding Optimal at High SNR

[Maric, Goldsmith, M. ‘10, ‘12]
Practical Implications – Zig-Zag Decoding

- Successive interference cancellation is a form of analog network coding in high SNR
- Basis for zig-zag [Gollakota and Katabi ‘08]
- Chunk 1 from user A from 1st copy of collided packet can be decoded successfully
  - Subtract from 2nd copy to decode the Chunk 1 of user B
- Subtract from 1st copy of collided packet to decode Chunk 2 from user A
  - Subtract from 2nd copy of collided packet to decode Chunk 2 from user B
- Can be extended to coded packets

Decode the first chunk of \( x \) and \( y \) using two interference free portions

Combination of algebraic network coding and analog network coding
Innovative reception: At least one undecoded packet is connected
\[ T_i = \text{Time of reception of } i\text{th degree of freedom} \]
\[ T_D = T_n = \sum_{i=1}^{n} X_i, \quad \text{where } X_i = T_i - T_{i-1} \]
\[ D_i = \text{Number of decoded packets at time } T_{i-1} \]
\[ (X_i|D_i) \sim \text{Geom}\left(\frac{1}{1 - p^{n-D_i}}\right) \]

[ParandehGheibi, Sundararajan, M. '10]
What about Low SNR?

- Consider again hyperedges
- At low SNR, noise is the main issue
- Non-coherence is not bothersome [Fawaz, M. ‘10]
What About Other Regimes?

• The use of hyperedges is important to take into account dependencies

• In general, it is difficult to determine how to proceed (see the difficulties with the relay channel)

• **Equivalence** leads to certain bounds for multiple access and broadcast channels, but these bounds may be loose

• Separates the issue of physical layer coding from that of network coding
  – The new network should be composed by **bit pipes**. This allows the abstraction of the stochastic nature of the network.
  – Instead of bounding the entire network, create bounding **components** for different elements (e.g. channels)
Point-to-Point Equivalence

**Theorem** [Kötter, Effros, M. ‘09, ’11, ‘13]: A network composed by discrete memoryless point-to-point links is equivalent to a network where each link is substituted by a noiseless bit pipe with throughput equal to its capacity.

- Consequence: feedback and cooperation cannot increase the achievable rate region.
- How does this extend to networks composed of multi-terminal channels?
Extending to Multi-terminal Channels

- Key idea: create bounding models by using “equivalent” bit pipe components for each channel.

- Achievable region changes if transmitting and/or receiving nodes are allowed to cooperate
- Feedback can increase capacity.
Example: Two User Gaussian MAC

\[ \gamma(x) = \frac{1}{2} \log(1 + x) \]

\[ Y = a_1 X_1 + a_2 X_2 + N \]
\[ \mathbb{E}[X_i^2] \leq P_i \]
\[ N \sim \mathcal{N}(0, \sigma^2) \]

[du Pin Calmon, Effros, M. ‘11]

\[ \Delta \leq \frac{1}{2} \text{ bits/s/Hz (smaller at low SNR)} \]

\[ C_1 = \gamma(\text{SNR}_1/\alpha^*) \]
\[ C_2 = \gamma(\text{SNR}_2/(1 - \alpha^*)) \]
\[ C_S = \gamma((\frac{1}{\text{SNR}_1} + \frac{1}{\text{SNR}_2})^{-1}) \]
\[ \alpha^* = \left(1 + \sqrt{\frac{1}{\text{SNR}_1^{-1} + \text{SNR}_2^{-1}}}ight)^{-1} \]
Network Coding vs. ARQ and HARQ

- **Scenario**: 5-packet block transfer from BS to SS
- **Downlink**: fixed 40% packet error pattern (every 3rd and 5th packet)
- **Uplink**: feedback NACKs not subject to loss
- **ARQ**: repeated transmissions create RTT feedback loops
- **HARQ**: feedback reduced by combining corrupted packet versions
- **Network Coding**: a-priori systematic coding with added redundancy of 3/5
- **Clear delay, throughput, and energy gains**
  - No feedback loop
  - Redundancy cost amortized over block

[Teeratapittayanon et al ‘12]
HARQ and ARQ tests in WiMAX

Global Environment for Network Innovations (GENI) → Indoor experiment at BBN

Average downlink throughput (Mbps) for different HARQ/ARQ configurations and packet sizes under 5Mbps offered load

- Turning off HARQ/ARQ increases the available bandwidth in WiMAX
- WiMAX is NOT in any way crucial to our technology – based on availability, through GENI program, of base station
- ARQ and HARQ mechanisms are close between LTE and WiMAX

<table>
<thead>
<tr>
<th></th>
<th>70Bytes</th>
<th>140Bytes</th>
<th>280Bytes</th>
<th>360Bytes</th>
<th>720Bytes</th>
<th>1440Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>HARQ OFF ARQ OFF</td>
<td>3.03</td>
<td>3.69</td>
<td>3.72</td>
<td>3.84</td>
<td>4.21</td>
<td>4.29</td>
</tr>
<tr>
<td>HARQ OFF ARQ ON</td>
<td>1.42</td>
<td>1.94</td>
<td>1.48</td>
<td>2.28</td>
<td>2.19</td>
<td>1.79</td>
</tr>
<tr>
<td>HARQ ON ARQ OFF</td>
<td>0.60</td>
<td>0.67</td>
<td>0.84</td>
<td>0.88</td>
<td>0.85</td>
<td>0.79</td>
</tr>
<tr>
<td>HARQ ON ARQ ON</td>
<td>0.38</td>
<td>0.51</td>
<td>0.59</td>
<td>0.41</td>
<td>0.67</td>
<td>0.64</td>
</tr>
</tbody>
</table>
Consistency

- BBN
- UCLA
- Rutgers

Opportunity for network coding?

20Mbps

- HARQ off ARQ off 5Mbps
- HARQ on ARQ on 5Mbps

6Mbps (1)

6Mbps (2)
Experimental Setup

- Intra-flow NC modules at the Base Station (BS) and Subscriber Station (SS)
- Toggle ARQ, HARQ, and various NC configurations
- IPERF → application-layer throughput / loss
- UFTP (FTP over UDP) → application-layer file-transfer delay
• WiMAX MAC inaccessible → IP-based implementation
• Performance measurements at the application layer (IPERF and UFTP)
• IP layer: *Netfilter* used to intercept packets, route them to encoder/decoder, then re-inject them to IP layer
• PDCP does not need to be involved, although that may be quite doable
• Network coding included at the e-Node B before handing it to the MAC, and ARQ and HARQ bypassed at the MAC – does not require a proxy, but can be used if convenient.
Block-Based Encoding Process

1. Buffering
   - IP packet
   - IP packet
   - IP packet
   - Coding buffer list

2. Concatenation
   - IP packet
   - IP packet
   - IP packet
   - Coding block

3. Padding
   - IP packet
   - IP packet
   - IP packet
   - Padding
   - Padded coding block

4. Segmentation
   - Segment
   - Segment
   - Segment
   - Segment
   - Segment

5. Coding
   \[ \text{Coded segment} = \sum_{i=1}^{N_s} a_i \text{ Segment} \]

6. Encapsulation
   - NC header
   - Coded segment
   - NC header
   - Coded segment
   - NC header
   - Coded segment
   - Coded IP packets

- Concatenate packets up to size- or time-limit
- Pad to minimum block size
- Systematic NC: only redundant packets are coded
Case Study: File Transfer Delay (UFTP)

- Raw “throughput” unreliable: UFTP runs its own ACK mechanism
- This is not using TCP
Consistency

- NC-Best decreases packet loss from 11-32% to nearly 0%
- NC offers up to 5.9x gain in throughput and 5.5x reduction in file transfer delay
Coding in Sensor nodes

Matlab program on a PC through an FPGA.

Generic commercial transceiver (Texas Instruments)

Transmission data rate of 500 kbps

FSK Modulation

Data transmission and coherent demodulation at receiver

Hard Viterbi decoding and an interleaver of 4 bytes

PC-based packet sniffer software transfers the data from the CC2511 over a USB interface

CC2511 chip provides the Received Signal Strength Indicator (RSSI)

[Angelopoulos, Paidimarri, Chandrakasan, M. ‘13]
The Benefit of Joint Coding

![Graph showing the benefit of joint coding with different FEC and RLNC rates and their corresponding SNR improvements.]

<table>
<thead>
<tr>
<th>FEC Rate</th>
<th>RLNC Rate</th>
<th>SNR Improvement (PER=10⁻¹)</th>
<th>SNR Improvement (PER=10⁻²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>4/5</td>
<td>0dB</td>
<td>1.5dB</td>
</tr>
<tr>
<td>1</td>
<td>4/6</td>
<td>0.625dB</td>
<td>2.5dB</td>
</tr>
<tr>
<td>1</td>
<td>4/8</td>
<td>1.5dB</td>
<td>3.4dB</td>
</tr>
<tr>
<td>1/2</td>
<td>1</td>
<td>2.5dB</td>
<td>2.25dB</td>
</tr>
<tr>
<td>1/2</td>
<td>4/5</td>
<td>2.25dB</td>
<td>4dB</td>
</tr>
<tr>
<td>1/2</td>
<td>4/6</td>
<td>2.75dB</td>
<td>4.25dB</td>
</tr>
<tr>
<td>1/2</td>
<td>4/8</td>
<td>3.5dB</td>
<td>5.6dB</td>
</tr>
</tbody>
</table>
Conclusions

- Joint physical layer and network coding seems in many cases to be limited:
  - Low SNR:
    - Discard noise rather than propagate it
    - Practical implication: simple hypergraph model approximation
  - High SNR:
    - Benefit from performing analog network coding with respect to digitized approach
    - Applications: Zig-Zag generalization
  - In general:
    - Benefits seem to be in general limited

- In practice:
  - WiMax:
    - Remove lower layer approaches by higher-layer coding
  - Low power sensor nodes:
    - Both PHY and network coding are beneficial
    - Coordination between the two may not be necessary

- Suggests an approach that is mostly based on separation – empirical study shows promising results
References

- N. Fawaz and Médard, M., “On the Non-Coherent Wideband Multipath Fading Relay Channel”, *ISIT 2010*


References

- S. Teerapittayanon, Fouli, K., Médard, M., Montpetit, M.-J., Shi, X., Seskar, I., and Gosain, A., “Network Coding as a WiMAX Link Reliability Mechanism”, *MACOM 2012*


- Y. Xu, E. Yeh, M., Médard, “Approaching Gaussian Relay Network Capacity in the High SNR Regime: End-to-End Lattice Codes”, Arxiv 2013

Model

- Denote power at destination
  \[ P_D = \left( \sum_{i \in \mathcal{N}(D)} h_i D \sqrt{P_i} \right)^2 \]

- MAC cut-set
  \[ C_{MAC} = \frac{1}{2} \log (1 + P_D) \]

\[ y_k = \sum_{j \in \mathcal{N}(k)} h_{jk} x_j + z_k \]
A different view of high SNR

- In a layered relay network under high-SNR conditions:

\[
\left( \sum_{j \in N(k)} h_{jk} \sqrt{P_j} \right)^2 \leq \delta \quad \text{for each } k \neq D
\]

- Analog network coding achieves

\[
R = \frac{1}{2} \log \left( 1 + \frac{1}{(1 + \delta)^{L-1}} \frac{P_D}{P_{Z,D} + 1} \right)
\]

- At high SNR ANC achieves capacity:

\[
P_{Z,D} \leq L\delta P_D \quad R = \frac{1}{2} \log (1 + P_D) - O(\delta)
\]

\[
\delta \to 0: \quad P_{Z,D} \to 0
\]