Wireless Network information Theory

P. R. Kumar

Dept. of Electrical and Computer Engineering, and
Coordinated Science Lab
University of Illinois, Urbana-Champaign

Email: prkumar@illinois.edu
Web: http://decision.csl.illinois.edu/~prkumar
What is really the best way to operate wireless networks?

And what are the ultimate limits to information transfer over wireless networks?
Outline

- Reappraising multi-hop transport: 4
- What is information theory?: 11
- Network information theory: 22
- Model for wireless network information theory: 33
- Results when absorption or relatively large path loss: 45
- Order optimality of multi-hop transport: 65
- The effect of fading: 80
- Low path loss: 82
- A quick survey of more recent results: 94
- Remarks: 99
- References: 100
Reappraising multi-hop transport
Reappraising multi-hop transport

- Nodes fully decode packets at each stage
- Treating interference as noise

- But why should nodes *Decode and Forward*?
- Why not just *Amplify and Forward*?

- Why should intermediate nodes be able to decode the packets?
- Why go *digital*?
Why treat interference as noise?

- Interference is not interference
  - Subtract loud signal
  - Interference is information

- Packets do not destructively collide

- Why not use *multi-user decoding*?

- How much benefit can multi-user decoding give for wireless networks?
Should we try to do active interference cancellation?

- Why not reduce the denominator in the SINR rather than increase the denominator?
Why even take small hops?

- Why not use *long range communication* with multi-user decoding?
In fact is the notion of *spatial reuse* appropriate for wireless networks?

- Spatial reuse of frequency

- If spatial reuse of frequency is the goal, then is a sharper path loss better for wireless networks?
  - Is $\frac{1}{r^8}$ better for wireless networks than $\frac{1}{r^4}$?
  - Or worse?

- Are **jungles** better for wireless networking than **deserts**?
Wireless networks are not wired networks …

- Wireless networks are formed by nodes with radios
  - There is no *a priori* notion of “links”
  - Nodes simply radiate energy
  - Maxwell rather than Kirchoff

- Nodes can cooperate in many complex ways

  “There are more things in heaven and earth, Horatio, Than are dreamt of in your philosophy.”
  — Hamlet

- So how should information be transported in wireless networks?
- What should be the architecture of wireless networks?
- What are the limits to information transfer?

- Need an information theory to provide strategic guidance for wireless networks
What is Information Theory?
Model of communication

Information Source → Information Transmitter → Channel → Receiver → Information Sink

Message → Transmitted Signal → Received signal

Noise
Question that Shannon posed and answered

- Given a *noisy* communication channel
- Channel Modeled by $p(y|x)$
  » Called a Discrete Memoryless Channel

Question: How many bits per transmission can be *reliably* sent?

- Call this the capacity of the channel
- How can we *achieve* this capacity over the channel?
Shannon’s formulation

- There are a set of $2^{nR}$ messages
- One message $W$ in $\{1, 2, \ldots, 2^{nR}\}$ is picked by the source out of these $2^{nR}$ messages
- This is encoded as a codeword $\{X_1, X_2, \ldots, X_n\}$
- $X_k$ is transmitted on the $k$-th transmission
- $Y_k$ is received on the $k$-th transmission
- So in $n$ uses of the channel $\{X_1, X_2, \ldots, X_n\}$ is sent, and $\{Y_1, Y_2, \ldots, Y_n\}$ is received
Shannon’s formulation

- There are a set of $2^{nR}$ messages
- One message $W$ in $\{1, 2, \ldots, 2^{nR}\}$ is picked by the source out of these $2^{nR}$ messages
- This is encoded as a codeword $\{X_1, X_2, \ldots, X_n\}$
- $X_k$ is transmitted on the $k$-th transmission
- $Y_k$ is received on the $k$-th transmission
- So in $n$ uses of the channel $\{X_1, X_2, \ldots, X_n\}$ is sent, and $\{Y_1, Y_2, \ldots, Y_n\}$ is received
- The receiver decodes $\{Y_1, Y_2, \ldots, Y_n\}$ as $W$
Definition of Achievable Rate $R$

- Let $P_{\text{error}} = \text{Prob}(W \neq W)$
- Suppose we can make $P_{\text{error}}$ smaller than any $\varepsilon$ we desire by choosing $n$ large
- Then we say that the channel can support a Rate of $R$ bits per transmission
- **Overall scheme**
  - Choose encoder $E$: $\{1, 2, \ldots, 2^{nR}\} \rightarrow X^n$
  - Choose decoder $D$: $X^n \rightarrow \{1, 2, \ldots, 2^{nR}\}$
  - Want $P_{\text{error}}$ smaller than a desired $\varepsilon$
  - Then we can “reliably transmit $R$ bits per transmission”
Shannon’s Answers

- **Capacity Theorem**
  - Given Channel Model \( p(y|x) \)
  
- **Capacity** = \( \max_{p(x)} I(X;Y) \) bits/transmission

  - Where \( I(X;Y) = \sum_{x,y} p(x,y) \log \left( \frac{p(X,Y)}{p(X)p(Y)} \right) \) is called the “mutual information”
  
  - This is the supremum of the achievable rates

- **Shannon’s architecture for digital communication**
Capacity of Gaussian Channel

- Gaussian Channel
  
  \[ Y_i = X_i + Z_i \]

- \( Z_i \sim N(0, \sigma^2) \)
  - Independent, identically distributed noise

- Power constraint \( P \) on transmissions:
  \[
  \frac{1}{n} \sum_{i=1}^{n} X_i^2 \leq P
  \]

- Capacity = \[
  \frac{1}{2} \log \left( 1 + \frac{P}{\sigma^2} \right)
  \] bits per transmission
Capacity of Continuous AWGN Bandlimited Channel

- AWGN Noise $Z(t)$ with Power Spectral Density $\frac{N}{2}$
- Band Limited Channel $[-W,+W]$
- Power constraint $P$ on signal transmitted: $\frac{1}{T} \int_0^T X^2(t) \leq P$

$Z(t)$ White Gaussian Noise with PSD $N$

$X(t)$  \[ + \]  \[ Y(t) \]

Capacity = $W \log \left( 1 + \frac{P}{WN} \right)$ bits per second
Limitations of Shannon’s result

- Does not address the issue of latency
- Delay incurred by block coding

- What is the joint tradeoff between
  - Throughput and Delay (and Error Rate)
The classic references


Network Information Theory
The Multiple Access Channel

- **Model**
  - Node 1 sends
  - Node 2 sends
  - The receiver receives generated as $p(y|x_1, x_2)$

- **Senders and their Rates**
  - Message 1: $W_1 \in \{1, 2, \ldots, 2^{nR_1}\}$
    - Sends $X_1(W)$
  - Message 2: $W_2 \in \{1, 2, \ldots, 2^{nR_2}\}$
    - Sends $X_2(W)$

- **Decoder:** $\hat{X}_1(Y)$ and $\hat{X}_2(Y)$

- **What rate vectors $(R_1, R_2)$ are feasible?**
Solution

- **Capacity region:**

  All rate vectors \((R_1, R_2)\) satisfying

  \[
  R_1 < I(X_1; Y | X_2) \\
  R_2 < I(X_2; Y | X_1) \\
  R_1 + R_2 < I(X_1, X_2; Y)
  \]

  for some distribution \(p(x_1)p(x_2)\) are feasible.
Interpretation and coding strategy

- At point A

\[ Max_{p(x_1)p(x_2)} I(X_1; Y|X_2) \]
\[ = Max_{p(x_1)p(x_2)} \sum_{x_2} p(x_2) I(X_1; Y|X_2 = x_2) \]
\[ = Max_{x_2} Max_{p(x_1)} I(X_1; Y|X_2 = x_2) \]

\{ Node 2 acts as a pure facilitator \}
Interpretation and coding strategy

- At point B
  - Receiver first decodes
  - Possible since $R_2 < I(X_2; Y)$
  - Then decodes
  - Possible since $R_1 < I(X_1; Y|X_2)$

Successive subtraction and decoding strategy (CDMA)
The Scalar Gaussian Broadcast Channel

- **Goal**
  - To send $W_1 \in \{1, 2, \ldots, 2^{nR_1}\}$ to Receiver 1
  - To send $W_2 \in \{1, 2, \ldots, 2^{nR_2}\}$ to Receiver 2
  - Simultaneously
  - Through one broadcast $X(W_1, W_2)$
  - Power constraint

- **Receiver 1 receives** $Y_1 = X + Z_1$ where $Z_1$ is $N(0, \sigma_1^2)$
  - Decodes $\hat{W}_1(Y_1)$

- **Receiver 2 receives** $Y_2 = X + Z_2$ where $Z_2$ is $N(0, \sigma_2^2)$
  - Decodes $\hat{W}_2(Y_2)$

- What rate vectors $(R_1, R_2)$ are feasible?
Solution

- Assume $\sigma_1^2 \leq \sigma_2^2$
  - Receiver 1 is better than Receiver 2
  - So Receiver 1 can decode anything that Receiver 2 can
  - So Receiver 1 can decode

- **Capacity region**: All vectors $(R_1, R_2)$ satisfying

\[
R_1 < \frac{1}{2} \log(1 + \frac{\alpha P}{\sigma_1^2})
\]
\[
R_2 < \frac{1}{2} \log(1 + \frac{(1 - \alpha)P}{\alpha P + \sigma_2^2})
\]

for some $0 < \alpha < 1$

- Sender uses power $P$ for Receiver 1, and power $(1 - \alpha)P$ for Receiver 2

- Receiver 2 has signal strength $(1 - \alpha)P$ and noise $\alpha P + \sigma_2^2$

- Receiver 1 first decodes $\hat{W}_2$ and then subtracts it. So signal in noise
General broadcast channel

- General Broadcast channel capacity unknown
  - Vector Gaussian channel capacity recently established
Max Flow - Min Cut Theorem

- **Theorem (El Gamal Ph. D. Thesis)**

  Suppose \( \{ R_{ij} \} \) is feasible vector of rates.

  Then
  \[
  \sum_{i \in S, j \in S^c} R_{ij} \leq I(X(S); Y(S^c)|X(S^c))
  \]

- **Example: Relay Channel**

  \[
  R \leq \text{Min}\{ I(X; Y_1, Y|X_1), I(X, X_1; Y) \}
  \]
The Slepian-Wolfe Problem: Distributed Source Coding

- To reconstruct \((X, Y)\) at the destination, it is sufficient to have

\[
R_1 > H(X|Y) \\
R_2 > H(Y|X) \\
R_1 + R_2 > H(X, Y)
\]

- So \(X\) and \(Y\) can code separately and still achieve the same result as though they were cooperating
Network information theory

**Triumphs**

- Gaussian broadcast channel
- Multiple access channel

**Unknowns**

- The simplest relay channel
- The simplest interference channel

- Networks being built (ad hoc networks, sensor nets) are much more complicated
Model for Wireless Network Information Theory
Model of system: A planar network

- Introduce distance
  - Node locations
  - Distances between nodes,
  - Attenuation as a function of distance

- $n$ nodes in a plane

- $\rho_{ij}$ = distance between nodes $i$ and $j$

- Signal attenuation with distance $\rho$ is $\frac{e^{-\gamma \rho}}{\rho^\delta}$
  - $\delta > 0$ is the path loss exponent
  - $\gamma \geq 0$ is the absorption constant

  » Generally $\gamma > 0$ since the medium is absorptive unless over a vacuum
  » Corresponds to a loss of $20\gamma \log_{10}e$ db per meter
Transmitted and received signals

- $W_i =$ symbol from $\{1, 2, 3, \ldots, 2^{TR_i}\}$ to be sent by node $i$ in $T$ transmissions
Transmitted and received signals

- \( W_i \) = symbol from \( \{1,2,3,\ldots,2^{TR_k}\} \) to be sent by node \( i \) in \( T \) transmissions
Transmitted and received signals

- $W_i = \text{symbol from } \{1, 2, 3, \ldots, 2^{R_{ik}}\} \text{ to be sent by node } i \text{ in } T \text{ transmissions}$
- $x_i(t) = f_{i,t}(y_{i}^{t-1}, W_i)$ = signal transmitted by node $i$ time $t$
Transmitted and received signals

- $W_i$ = symbol from $\{1, 2, 3, \ldots, 2^{TR_k}\}$ to be sent by node $i$ in $T$ transmissions
- $x_i(t) = f_i, t(y_{i-1}^t, W_i)$ = signal transmitted by node $i$ time $t$
- $y_j(t) = \sum_{i=1}^{n} e^{-\gamma_{ij} \rho_{ij}} \sum_{\delta x_i(t)} + z_j(t)$ = signal received by node $j$ at time $t$

- Transport Capacity: bit-meters/second or bit-meters/slot

Graphical representation:
- Node $x_i$ transmits signals to node $y_j$.
Transmitted and received signals

- \( W_i = \) symbol from \( \{1,2,3,\ldots,2^{TR_{ik}}\} \) to be sent by node \( i \) in \( T \) transmissions

- \( x_i(t) = f_i,y_i(t-1),W_i \) = signal transmitted by node \( i \) time \( t \)

- \( y_j(t) = \sum_{i=1}^{n} \frac{e^{-\gamma_{ij}}}{\rho_{ij}}x_i(t) + z_j(t) \) = signal received by node \( j \) at time \( t \)

- Destination \( j \) uses the decoder \( \hat{W}_i = g_j(y_j, W_j) \)
Transmitted and received signals

- $W_i =$ symbol from $\{1,2,3,\ldots,2^{TR_{ik}}\}$ to be sent by node $i$ in $T$ transmissions

- $x_i(t) = f_{i,t}(y_{i_{t-1}}, W_i)$ = signal transmitted by node $i$ time $t$

- $y_j(t) = \sum_{i=1}^{n} \frac{e^{-\rho_{ij}}}{\delta} x_i(t) + z_j(t) = \text{signal received by node } j \text{ at time } t$

- Destination $j$ uses the decoder $\hat{W}_i = g_j(y_j^T, W_j)$

- Error if $\hat{W}_i \neq W_i$
Transmitted and received signals

- $W_i$ = symbol from $\{1, 2, 3, \ldots, 2^{TR_k}\}$ to be sent by node $i$ in $T$ transmissions
- $x_i(t) = f_{i,t}(y_{i,t}^{t-1}, W_i)$ = signal transmitted by node $i$ time $t$
- $y_j(t) = \sum_{i=1}^{n} e^{-\rho_{ij}} \delta_{ij} x_i(t) + z_j(t)$ = signal received by node $j$ at time $t$
- Destination $j$ uses the decoder $\hat{W}_i = g_j(y_j^T, W_j)$
- Error if $\hat{W}_i \neq W_i$
- $(R_1, R_2, \ldots, R_l)$ is feasible rate vector if there is a sequence of codes with
  $$\max_{W_1, W_2, \ldots, W_l} \Pr(\hat{W}_i \neq W_i \text{ for some } i \mid W_1, W_2, \ldots, W_l) \to 0 \text{ as } T \to \infty$$
Transmitted and received signals

- $W_i =$ symbol from $\{1, 2, 3, \ldots, 2^{TR_i}\}$ to be sent by node $i$ in $T$ transmissions
- $x_i(t) = f_{i,t}(y_{i,t-1}, W_i)$ = signal transmitted by node $i$ time $t$
- $y_j(t) = \sum_{i=1}^{n} e^{-\frac{p_{ij}}{\delta}} x_i(t) + z_j(t)$ = signal received by node $j$ at time $t$
- Destination $j$ uses the decoder $\hat{W}_i = g_j(y_j^T, W_j)$
- Error if $\hat{W}_i \neq W_i$
- $(R_1, R_2, \ldots, R_l)$ is feasible rate vector if there is a sequence of codes with $\max_{W_1, W_2, \ldots, W_l} \Pr(\hat{W}_i \neq W_i \text{ for some } i \mid W_1, W_2, \ldots, W_l) \to 0$ as $T \to \infty$
- Individual power constraint $P_i \leq P_{ind}$ for all nodes $I$. 
Transmitted and received signals

- $W_i$ = symbol from $\{1,2,3,...,2^{TR_{ik}}\}$ to be sent by node $i$ in $T$ transmissions
- $x_i(t) = f_{i,t}(y_{i,t}^{t-1}, W_i)$ = signal transmitted by node $i$ time $t$
- $y_j(t) = \sum_{i=1}^{n} e^{-\gamma \rho_{ij}} \sum_{i \neq j} x_i(t) + z_j(t)$ = signal received by node $j$ at time $t$
- Destination $j$ uses the decoder $\hat{W}_i = g_j(y_j^T, W_j)$
- Error if $\hat{W}_i \neq W_i$
- $(R_1, R_2, ..., R_l)$ is feasible rate vector if there is a sequence of codes with
  \[ \max_{W_1, W_2, ..., W_l} \Pr(\hat{W}_i \neq W_i \text{ for some } i \mid W_1, W_2, ..., W_l) \rightarrow 0 \text{ as } T \rightarrow \infty \]
- **Individual power constraint** $P_i \leq P_{\text{ind}}$ for all nodes $I$.  **Or Total power constraint** $\sum_{i=1}^{n} P_i \leq P_{\text{total}}$
Transmitted and received signals

- $W_i =$ symbol from $\{1,2,3,\ldots,2^{TR_i}\}$ to be sent by node $i$ in $T$ transmissions

- $x_i(t) = f_{i,t}(y_i^{t-1},W_i) =$ signal transmitted by node $i$ time $t$

- $y_j(t) = \sum_{i=1, i \neq j}^{n} e^{-\rho_{ij}} e^{\delta} x_i(t) + z_j(t) =$ signal received by node $j$ at time $t$

- Destination $j$ uses the decoder $\hat{W}_i = g_j(y_j^T, W_j)$

- Error if $\hat{W}_i \neq W_i$

- $(R_1,R_2,\ldots,R_l)$ is feasible rate vector if there is a sequence of codes with
  \[
  \max_{W_1,W_2,\ldots,W_l} \Pr(\hat{W}_i \neq W_i \text{ for some } i \mid W_1,W_2,\ldots,W_l) \to 0 \text{ as } T \to \infty
  \]

- **Individual power constraint** $P_i \leq P_{\text{ind}}$ for all nodes $I$. **Or Total power constraint** $\sum_{i=1}^{n} P_i \leq P_{\text{total}}$

- **Transport Capacity** $C_T = \sup_{(R_1,R_2,\ldots,R_{n(n-1)})} \sum_{i=1}^{n(n-1)} R_i \cdot \rho_i$ bit-meters/second or bit-meters/slot
Results when there is absorption or a relatively large path loss
Total transmitted power bounds the transport capacity

- **Theorem:** Bit-meters per Joule bound (Xie & K ’02)
Total transmitted power bounds the transport capacity

- Theorem: Bit-meters per Joule bound (Xie & K ’02)
  - Suppose $\gamma > 0$, there is some absorption,
Total transmitted power bounds the transport capacity

- **Theorem:** Bit-meters per Joule bound (Xie & K ’02)
  - Suppose $\gamma > 0$, there is some absorption,
  - Or $\delta > 3$, if there is no absorption at all
Total transmitted power bounds the transport capacity

- **Theorem:** Bit-meters per Joule bound (Xie & K ’02)
  - Suppose \( \gamma > 0 \), there is some absorption,
  - Or \( \delta > 3 \), if there is no absorption at all
  - Then for all Planar Networks
Total transmitted power bounds the transport capacity

- Theorem: **Bit-meters per Joule bound** (Xie & K ’02)
  - Suppose $\gamma > 0$, there is some absorption,
  - Or $\delta > 3$, if there is no absorption at all
  - Then for all Planar Networks

$$
C_T \leq \frac{c_1(\gamma, \delta, \rho_{\text{min}})}{\sigma^2} \cdot P_{\text{total}}
$$
Total transmitted power bounds the transport capacity

- **Theorem:** *Bit-meters per Joule bound* (Xie & K ’02)

  - Suppose $\gamma > 0$, there is some absorption,
  
  - Or $\delta > 3$, if there is no absorption at all
  
  - Then for all Planar Networks

  \[
  C_T \leq \frac{c_1(\gamma, \delta, \rho_{\text{min}})}{\sigma^2} \cdot P_{\text{total}}
  \]

  Energy cost of communicating one bit-meter in a sensor network
Total transmitted power bounds the transport capacity

**Theorem:** Bit-meters per Joule bound (Xie & K ’02)

- Suppose \( \gamma > 0 \), there is some absorption,
- Or \( \delta > 3 \), if there is no absorption at all
- Then for all Planar Networks

\[
C_T \leq \frac{c_1(\gamma, \delta, \rho_{\text{min}})}{\sigma^2} P_{\text{total}}
\]

**where**

\[
c_1(\gamma, \delta, \rho_{\text{min}}) = \begin{cases} 
\frac{2^{2\delta+7}}{\gamma^2 \rho_{\text{min}}^{2\delta+1}} e^{\gamma \rho_{\text{min}} / 2} (2 - e^{-\gamma \rho_{\text{min}} / 2}) & \text{if } \gamma > 0 \\
\frac{2^{2\delta+5}}{(\delta - 2)^2 (\delta - 3) \rho_{\text{min}}^{2\delta-1}} (3\delta - 8) & \text{if } \gamma = 0 \text{ and } \delta > 3
\end{cases}
\]

Energy cost of communicating one bit-meter in a wireless network
O(n) upper bound on Transport Capacity

Theorem: Transport capacity is $O(n)$ (Xie & K'02)

- Suppose $\gamma > 0$, there is some absorption,
- Or $\delta > 3$, if there is no absorption at all

Then for all Planar Networks

- Same as square root law based on treating interference as noise since area $A$ grows like $\Omega(n)$

So multi-hop with decode and forward with interference treated as noise is order optimal architecture whenever $\Theta(n)$ can be achieved
O(n) upper bound on Transport Capacity

- Theorem: Transport capacity is $O(n)$ (Xie & K '02)

$\Theta(n) = \Theta(n)$
O(n) upper bound on Transport Capacity

- Theorem: Transport capacity is $O(n)$ (Xie & K ’02)
  - Suppose $\gamma > 0$, there is some absorption,
  - Or $\delta > 3$, if there is no absorption at all

Time for a little bit of math:

\[ C_T \leq c_1(\gamma, \delta, \rho, m, i, n, P) \cdot n = \Theta(n^2) \]
O(n) upper bound on Transport Capacity

- Theorem: Transport capacity is $O(n)$ (Xie & K ’02)
  - Suppose $\gamma > 0$, there is some absorption,
  - Or $\delta > 3$, if there is no absorption at all
  - Then for all Planar Networks
Theorem: Transport capacity is $O(n)$ (Xie & K ’02)

- Suppose $\gamma > 0$, there is some absorption,
- Or $\delta > 3$, if there is no absorption at all
- Then for all Planar Networks

$$C_T \leq \frac{c_1(\gamma, \delta, \rho_{\min})}{\sigma^2} \cdot P_{total}$$
O(n) upper bound on Transport Capacity

- Theorem: \textit{Transport capacity is }\mathcal{O}(n)\textit{ (Xie & K ’02)}
  
  - Suppose \(\gamma > 0\), there is some absorption,
  
  - Or \(\delta > 3\), if there is no absorption at all
  
  - Then for all Planar Networks

\[
C_T \leq \frac{c_1(\gamma, \delta, \rho_{\text{min}})P_{\text{ind}}}{\sigma^2} \cdot n
\]

\[
P_{\text{total}} = P_{\text{ind}} \cdot n
\]
O(n) upper bound on Transport Capacity

- Theorem: Transport capacity is $O(n)$ (Xie & K ’02)
  - Suppose $\gamma > 0$, there is some absorption,
  - Or $\delta > 3$, if there is no absorption at all
  - Then for all Planar Networks
    \[
    C_T \leq \frac{c_1(\gamma, \delta, \rho_{\text{min}})P_{\text{ind}}}{\sigma^2} \cdot n
    \]
    \[
    P_{\text{total}} = P_{\text{ind}} \cdot n
    \]

- Same as square root law based on treating interference as noise
Theorem: Transport capacity is $O(n)$ (Xie & K ’02)

- Suppose $\gamma > 0$, there is some absorption,
- Or $\delta > 3$, if there is no absorption at all
- Then for all Planar Networks

$$C_T \leq \frac{c_1(\gamma, \delta, \rho_{\text{min}})P_{\text{ind}}}{\sigma^2} \cdot n$$

same as square root law based on treating interference as noise

- $\Theta(\sqrt{An}) = \Theta(n)$ since area $A$ grows like $\Omega(n)$
O(n) upper bound on Transport Capacity

- **Theorem:** Transport capacity is $O(n)$ (Xie & K ’02)
  
  - Suppose $\gamma > 0$, there is some absorption,
  
  - Or $\delta > 3$, if there is no absorption at all
  
  - Then for all Planar Networks
    
    $$ C_T \leq \frac{c_1(\gamma, \delta, \rho_{\min})P_{\text{ind}}}{\sigma^2} \cdot n $$
    
    $P_{\text{total}} = P_{\text{ind}} \cdot n$

- Same as square root law based on treating interference as noise
  
  - $\Theta(\sqrt{An}) = \Theta(n)$ since area $A$ grows like $\Omega(n)$

- So multi-hop with decode and forward with interference treated as noise is order optimal architecture whenever $\Theta(n)$ can be achieved
Idea behind proof

- A Max-flow Min-cut Lemma
  - \( N \) = subset of nodes
  - \( P_N^{rec}(T) = \) Power received by nodes in \( N \) from outside \( N \)
    
    \[
    P_N^{rec}(T) = \frac{1}{T} \sum_{t=1}^{T} \sum_{j \in N} \mathbb{E} \left( \sum_{i \in N} \frac{x_i(t)}{\rho_{ij}^S} \right)^2
    \]
  - Then
    
    \[
    \sum_{\{l: d_l \in N \text{ but } s_l \not\in N\}} R_l \leq \frac{1}{2\sigma^2} \liminf_{T \to \infty} P_N^{rec}(T)
    \]
To obtain power bound on transport capacity

- Idea of proof

- Consider a number of cuts one meter apart

- Every source-destination pair \((s_l, d_l)\) with source at a distance \(\rho_l\) is cut by about \(\rho_l\) cuts

- Thus

\[
\sum_{l} R_l \rho_l \leq c \sum_{N_k \{l \text{ is cut by } N_k\}} \sum_{l} R_l \leq \frac{c}{2\sigma^2} \sum_{N_k} \liminf_{T \to \infty} P_{r_{N_k}}^{rec}(T) \leq \frac{cP_{\text{total}}}{\sigma^2}
\]
O(n) upper bound on Transport Capacity

- **Theorem**
  - Suppose $\gamma > 0$, there is some absorption,
  - Or $\delta > 3$, if there is no absorption at all
  - Then for all Planar Networks
    \[
    C_T \leq \frac{c_1(\gamma, \delta, \rho_{\text{min}}) P_{\text{ind}}}{\sigma^2} \cdot n
    \]
    where
    \[
    c_1(\gamma, \delta, \rho_{\text{min}}) = \frac{2^{2\delta+7}}{\gamma^2 \rho_{\text{min}}^{2\delta+1}} \frac{e^{-\rho_{\text{min}}}}{2 \frac{e^{-\rho_{\text{min}}}}{2}} \frac{(2-e^{-\rho_{\text{min}}})}{(1-e^{-\rho_{\text{min}}})^2}
    \]
    if $\gamma > 0$
    \[
    = \frac{2^{2\delta+5}}{(\delta-2)^2(\delta-3)} \rho_{\text{min}}^{2\delta-1}
    \]
    if $\gamma = 0$ and $\delta > 3$
Order optimality of multi-hop transport
Order optimality of multihop transport in a randomly chosen scenario

\[ \Omega \left( \frac{n \log n}{\sqrt{n}} \right) \]
Order optimality of multihop transport in a randomly chosen scenario

- Random traffic

\[\Omega \frac{1}{n \log n} \]
Order optimality of multihop transport in a randomly chosen scenario

- Random traffic
Order optimality of multihop transport in a randomly chosen scenario

- Random traffic

- Multihop can provide $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ bits/second
Order optimality of multihop transport in a randomly chosen scenario

- Random traffic

- Multihop can provide $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ bits/second
  
  - for every source
Order optimality of multihop transport in a randomly chosen scenario

- Random traffic

- Multihop can provide $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ bits/second
  - for every source
  - with probability $\to 1$
Order optimality of multihop transport in a randomly chosen scenario

- Random traffic

- Multihop can provide $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ bits/second
  - for every source
  - with probability $\to 1$
  - as the number of nodes $n \to \infty$
Order optimality of multihop transport in a randomly chosen scenario

• Random traffic

• Multihop can provide \( \Omega\left(\frac{1}{\sqrt{n \log n}}\right) \) bits/second
  
  – for every source
  
  – with probability \( \to 1 \)
  
  – as the number of nodes \( n \to \infty \)

• Nearly optimal since transport capacity achieved is \( \Omega\left(\frac{n}{\sqrt{\log n}}\right) \)
Order optimality of multihop transport in a randomly chosen scenario

- Random traffic

- Multihop can provide $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ bits/second
  - for every source
  - with probability $\to 1$
  - as the number of nodes $n \to \infty$

- Nearly optimal since transport capacity achieved is $\Omega\left(\frac{n}{\sqrt{\log n}}\right)$

- So Random case $\approx$ Best Case
What can multihop transport achieve?

◆ Theorem

– A set of rates \((R_1, R_2, \ldots, R_l)\) can be supported by multi-hop transport if

– Traffic can be routed, possibly over many paths, such that

\[
\rho_{\text{S}}^{-\gamma} - \frac{2^{2+2\delta}}{\rho_{\text{min}}(\delta-1)} \leq \gamma \rho_{\text{min}}^{1+2\delta} + \sigma^2
\]

– No node has to relay more than

\[
S = \frac{e^{-2\gamma \bar{\rho}} P_{\text{ind}}}{\rho^{2\delta}} \frac{\bar{\rho}^{2\delta}}{c_3(\gamma, \delta, \rho_{\text{min}}) P_{\text{ind}} + \sigma^2}
\]

– where \(\bar{\rho}\) is the longest distance of a hop

\[
c_3(\gamma, \delta, \rho_{\text{min}}) = \begin{cases} 
2^{3+2\delta} e^{-\gamma \rho_{\text{min}}} & \text{if } \gamma > 0 \\
\gamma \rho_{\text{min}}^{1+2\delta} & \text{if } \gamma = 0 \text{ and } \delta > 1 \\
\frac{2^{2+2\delta}}{\rho_{\text{min}}(\delta-1)} & \text{if } \gamma = 0 \text{ and } \delta > 1
\end{cases}
\]
Multihop transport can achieve $\Theta(n)$

**Theorem**

- Suppose $\gamma > 0$, there is some absorption,
- Or $\delta > 1$, if there is no absorption at all
- Then in a regular planar network

$$C_T \geq S \left( \frac{e^{-2\gamma} P_{ind}}{c_2(\gamma, \delta) P_{ind} + \sigma^2} \right) \cdot n$$

where

$$c_2(\gamma, \delta) = \begin{cases} 
\frac{4(1+4\gamma)e^{-2\gamma} - 4e^{-4\gamma}}{2\gamma(1-e^{-2\gamma})} & \text{if } \gamma > 0 \\
16\delta^2 + (2\pi - 16)\delta - \pi & \text{if } \gamma = 0 \text{ and } \delta > 1
\end{cases}$$

$\sqrt{n}$ sources each sending over a distance $\sqrt{n}$
Optimality of multi-hop transport

- **Corollary**
  - So if $\gamma > 0$ or $\delta > 3$
  - And multi-hop achieves $\Theta(n)$
  - Then it is optimal with respect to the transport capacity
    - up to order

- **Example**
Multi-hop is almost optimal in a random network

- **Theorem**
  - Consider a regular planar network
  - Suppose each node randomly chooses a destination
    » Choose a node nearest to a random point in the square
  - Suppose $\gamma > 0$ or $\delta > 1$
  - Then multihop can provide $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ bits/time-unit for every source with probability $\to 1$ as the number of nodes $n \to \infty$

- **Corollary**
  - Nearly optimal since transport achieved is $\Omega\left(\frac{n}{\sqrt{\log n}}\right)$
Idea of proof for random source-destination pairs

- Simpler than Gupta-Kumar since cells are square and contain one node each

- A cell has to relay traffic if a random straight line passes through it

- How many random straight lines pass through cell?

- Use Vapnik-Chervonenkis theory to guarantee that no cell is overloaded
The effect of fading
Large path loss: Effect of fading

- $n$ nodes located on the plane
  - Base-band model
    \[ Y_j(t) = \sum_{i \neq j} \frac{\beta e^{-\gamma \rho_{ij}}}{\rho_{ij}^\delta} \left( \sum_{l=0}^{\infty} H_{ijl}(t) \cdot X_i(t - \tau_{ij} - l) \right) + Z_j(t) \]
  - Consider $\delta > 3$ or $\gamma > 0$
  - Then even with full channel state information, $C_T(n) \leq c_1 \cdot n$
  - Even with iid unknown channel, for regular node locations, there is a scheme yielding $C_T(n) \geq c_2 \cdot n$

(Xue, Xie and K ’03)
What happens when the attenuation is very low?
A feasible rate for the Gaussian multiple-relay channel

**Theorem**

- Suppose $\alpha_{ij} =$ attenuation from $i$ to $j$
- Choose power $P_{ik} =$ power used by $i$ intended directly for node $k$
- Then
  
  $R < \min_{1 \leq j \leq n} S \left( \frac{1}{\sigma^2} \sum_{k=1}^{j} \left( \sum_{i=0}^{k-1} \alpha_{ij} \sqrt{P_{ik}} \right)^2 \right)$

  is feasible

**Proof based on coding**
A group relaying version

- Theorem

- A feasible rate for group relaying

\[ R < \min_{1 \leq j \leq M} S \left( \frac{1}{\sigma^2} \sum_{k=1}^{j} \left( \sum_{i=0}^{k-1} \alpha_{NiNj} \sqrt{P_{ik}} n_i n_i \right)^2 \right) \]
A dichotomy: Optimal architecture depends on attenuation by medium

- When $\gamma=0$ and $\delta$ small (XK ‘04)
  - Transport capacity can grow superlinearly like $\Theta(n^\theta)$ for $\theta > 1$
  - Coherent multi-stage relaying with interference cancellation can be optimal

- Unbounded transport capacity for fixed total power
Another strategy

- Coherent multi-stage relaying with interference subtraction (CRIS)

- All upstream nodes coherently cooperate to send a packet to the next node

- A node cancels all the interference caused by all transmissions to its downstream nodes
Another strategy

- Coherent multi-stage relaying with interference subtraction (CRIS)

- All upstream nodes coherently cooperate to send a packet to the next node

- A node cancels all the interference caused by all transmissions to its downstream nodes
Another strategy

- **Coherent multi-stage relaying with interference subtraction (CRIS)**

- All upstream nodes coherently cooperate to send a packet to the next node
- A node cancels all the interference caused by all transmissions to its downstream nodes
Unbounded transport capacity can be obtained for fixed total power

**Theorem**
- Suppose $\gamma = 0$, there is no absorption at all,
- And $\delta < 3/2$
- Then $C_T$ can be unbounded in regular planar networks even for fixed $P_{total}$

**Theorem**
- If $\gamma = 0$ and $\delta < 1$ in regular planar networks
- Then no matter how many nodes there are
- No matter how far apart the source and destination are chosen
- A fixed rate $R_{min}$ can be provided for the single-source destination pair
Idea of proof of unboundedness

- **Linear case**: Source at 0, destination at \( n \)

\[
P_{ik} = \frac{P}{(k-i)^\alpha k^\beta}
\]

- **Choose**

- **Planar case**
Networks with transport capacity $\Theta(n^\theta)$

- **Theorem**
  - Suppose $\gamma=0$
  - For every $1/2 < \delta < 1$, and $1 < \theta < 1/\delta$
  - There is a family of linear networks with
    $$C_T = \Theta(n^\theta)$$
  - The optimal strategy is coherent multi-stage relaying with interference cancellation
Idea of proof

- Consider a linear network

- Choose \( P_{ik} = \frac{P}{(k-i)^\alpha} \) where \( 1 < \alpha < 3 - 2\theta\delta \)

- A positive rate is feasible from source to destination for all \( n \)
  - By using coherent multi-stage relaying with interference cancellation

- To show upper bound
  - Sum of power received by all other nodes from any node \( j \) is bounded
  - Source destination distance is at most \( n^{\theta} \)
Low path loss

◆ Theorem (Unbounded path loss)
  – Suppose $\gamma = 0$ and $\delta < 3/2$
  – Then $C_T$ can be unbounded in regular planar networks even for fixed $P_{total}$

What happens when $\frac{3}{2} \leq \delta \leq 3$?

◆ Theorem (Superlinear scaling)
  – Suppose $\gamma = 0$. Then for every $\frac{1}{2} < \delta < 1$, and $1 < \theta < 1/\delta$
  – There is a family of linear networks with $C_T = \Theta(n^\theta)$

◆ Physically unrealistic

(Xie and K ‘02)
Recent work
Low path loss: Scaling behavior for path loss exponent $\delta < 3$

- For what path loss exponents *smaller* than 3 is $C_T = \Theta(n)$?
  - Jovicic, Viswanath and Kulkarni '04: $\delta > 3$ $\longrightarrow$ $\delta > 5/2$
  - Xie and K '06: $\longrightarrow$ $\delta > 2$
  - So the question remains for $1 < \delta < 2$

- Common per-node throughput in a random network
  - Leveque and Telatar '05: $\lambda(n) = o(1)$ when $\delta > 1$
What is the scaling behavior in the range $1 < \delta < 2$

- Aeron and Saligrama ’07: How to achieve a total throughput of $\Theta(n^{2/3})$ in a dense network
- Ozgur, Leveque and Tse ’07: Lower bound
  
  $n\lambda(n) \geq cn^{2-\delta-\epsilon}$ for $1 \leq \delta \leq \frac{3}{2}$
  
  $n\lambda(n) \geq c'\sqrt{n}$ for $\frac{3}{2} \leq \delta \leq 2$

- Based on cooperation
  - Long range MIMO between blocks of nodes
  - Intra-cluster cooperation
  - Transmit and receive cooperation

- Xie ’08: Exact study of pre-constant and shows it is $o(1)$
- Niessen, Gupta and Shah ‘08: Arbitrarily spaced nodes
Is “channel” the right model for massive cooperation?

- Franceschetti, Migliore, Minero ’08
- Number of information channels is only $O\left(\sqrt{n}\right)$
- Scaling law $O\left(\frac{\log^2 n}{\sqrt{n}}\right)$ per node
  - Limitation in spatial degrees of freedom
  - Not based on empirical path-loss models and stochastic fading models
  - Depends only on geometry
Physical Limits to Communication

Seth Lloyd, 1,2 Vittorio Giovannetti, 1,* and Lorenzo Maccone 1,†

1Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA
2Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

(Received 16 July 2003; revised manuscript received 13 January 2004; published 31 August 2004)

The limits posed by physics to the quantity of information that can be transmitted with a certain amount of power are investigated. The same ultimate limits are found for transmission of information encoded using matter and massless fields.

DOI: 10.1103/PhysRevLett.93.100501

PACS numbers: 03.67.Hk, 04.70.Dy, 89.70.+c

How much information can two parties exchange through a communication channel given certain energy resources? On one hand, quantum mechanics constrains the physical resources of any system in its capacity to store, transmit, and process information [1–4]. On the other hand, general relativity constrains the amount of information that can be stored in a finite volume of space through black hole thermodynamics [5–10]. Here we study the fundamental limits to the transmission of information by joining these two approaches: we provide some arguments which indicate that the ultimate communication rate might be obtained by transmitting infor-

\[ R \approx \frac{\sqrt{AP}}{\sqrt{I_p R}} \]  

where \( I_p = \sqrt{\hbar G/c^3} \approx 1.6 \times 10^{-35} \) m is the Planck length (\( G \) being the gravitational constant). As is customary in information theory, we leave out of the energy balance the energy needed by Alice to prepare the signal and by Bob to decode it.

**Channel capacity bound.**—Consider the scenario of Fig. 1: Alice encodes the information into slabs of material with rest-mass density \( \rho \) and sends them to Bob at speed \( c \). The data can be coded in \( \rho = \rho_{\text{max}} \).
Remarks

- Studied networks with arbitrary numbers of nodes
  - Explicitly incorporated distance in model
    » Distances between nodes
    » Attenuation as a function of distance
    » Distance is also used to measure transport capacity

- Make progress by asking for less
  - Instead of studying capacity region, study the transport capacity
  - Instead of asking for exact results, study the scaling laws
    » The exponent is more important
    » The preconstant is also important but is secondary - so bound it
  - Draw some broad conclusions
    » Optimality of multi-hop when absorption or large path loss
    » Optimality of coherent multi-stage relaying with interference cancellation when no absorption and very low path loss

- Open problems abound
  - What happens for intermediate path loss when there is no absorption
  - The channel model is simplistic, ......
  - ......
References-1

References-2

References-3

References-4

References-5
