Communication, Computation, and Control: Through the Information Lens

Pulkit Grover
Stanford University

CMU, April 16, 2012

Acknowledgments:
C2S2/Interconnect Focus Center (IFC)
NSF Center for Science of Information (CSoI)
NSF-CCF-0917212 (cyber-physical systems)
Flows of information in modern systems
Flows of information in modern systems
Flows of information in modern systems
Flows of information in modern systems
Talk outline:
Understanding Invisible Information Flows

1) Information flows in circuits

\[ M \xrightarrow{\text{Transmitter}} P_T \xrightarrow{\text{Channel}} \tilde{M} \xrightarrow{\text{Receiver}} \hat{M} \]
Talk outline:
Understanding Invisible Information Flows

1) Information flows in circuits

\[ M \rightarrow \text{Transmitter} \xrightarrow{P_T} \text{Channel} \rightarrow \text{Recevier} \rightarrow \hat{M} \]
Talk outline:

Understanding Invisible Information Flows

1) Information flows in circuits

\[ M \xrightarrow{\text{Transmitter}} P_T \xrightarrow{\text{Channel}} \hat{M} \]

2) Information flows in physical systems
Talk outline:
Understanding Invisible Information Flows

1) Information flows in circuits

\[ M \xrightarrow{\text{Transmitter}} \text{Channel} \xrightarrow{PT} \widehat{M} \]

2) Information flows in physical systems
Talk outline: Understanding Invisible Information Flows

1) Information flows in circuits

\[ M \xrightarrow{\text{Transmitter}} \text{Channel} \xrightarrow{\text{Receiver}} \widehat{M} \]

Traditional: Minimize \( P_T \)

Now: Minimize total power

Fundamental limits, experimental results

2) Information flows in physical systems

Implicit channel

Plant

\[ C_1 \xrightarrow{\text{Plant}} \text{Implicit channel} \xrightarrow{\text{Plant}} C_2 \]

\[ C_3 \xrightarrow{\text{Plant}} \text{Implicit channel} \xrightarrow{\text{Plant}} C_4 \]

\[ C_5 \xrightarrow{\text{Plant}} \text{Implicit channel} \xrightarrow{\text{Plant}} C_6 \]
Short-distances: Processing power can exceed transmit power!

Short-distance communication at Gbps (e.g. indoor wireless, data-centers ethernet)

Transmitter

Encoder

PA

80 mW [1]

Receiver

LNA

ADC

Equalizer

Decoder


$M$  \rightarrow  \hat{M}$

Short-districts: Processing power **can exceed** transmit power!

Short-distance communication at Gbps (e.g. indoor wireless, data-centers ethernet)

Choosing a $P_{total}$-efficient code?

Short-distances: Processing power can exceed transmit power!

Short-distance communication at Gbps (e.g. indoor wireless, data-centers ethernet)

Choosing a $P_{total}$-efficient code?

$$P_{total} = P_T + P_{enc} + P_{dec}$$

Minimizing transmit power: Shannon’s waterfall curves

Goal: Rate $R$, error probability $P_e$

$$C = W \log \left( 1 + \frac{\eta P_T}{N_0} \right) > R$$

---

Shannon's waterfall curves diagram:
- $M \rightarrow \mathcal{E}_{P_T} \rightarrow \mathcal{D} \rightarrow \hat{M}$
- Path loss
- Shannon limit (Tx power)
- Transmit power (Watts)
- $\log_{10}(P_e)$
- Fixed $R$, $W$, $\eta$
Minimizing total power:
Initial theoretical approaches
Minimizing total power: Initial theoretical approaches

“Black-box” model of processing power

\[ M \xrightarrow{P_T} \hat{M} \]

[Cui, Goldsmith, Bahai ’05]
[Massaad, Medard, Zheng ’08]
Minimizing total power:
Initial theoretical approaches

“Black-box” model of processing power

Optimal “burstiness”

[Cui, Goldsmith, Bahai ’05]
[Massaad, Medard, Zheng ’08]
Minimizing total power:
Initial theoretical approaches

“Black-box” model of processing power

Optimal “burstiness”
Minimizing total power: Initial theoretical approaches

“Black-box” model of processing power

Optimal “burstiness”
Minimizing total power: Initial theoretical approaches

“Black-box” model of processing power

Optimal “burstiness”

[Cui, Goldsmith, Bahai ’05]
[Massaad, Medard, Zheng ’08]
Minimizing total power: Initial theoretical approaches

“Black-box” model of processing power

Optimal “burstiness”

What are the “greenest” codes?

[Cui, Goldsmith, Bahai ’05]
[Massaad, Medard, Zheng ’08]
“Green” system = Capacity-code + “Green” implementation?

“A 47-Gbps LDPC decoder . . .” (10 GBase-T) [Zhang, Anantharam, Wainwright, Nikolic JSSC ’10]

“On the design of LDPC codes within 0.0045 dB of the Shannon limit.” [Chung, Forney, Richardson, Urbanke IT’01]
“Green” system = Capacity-code + “Green” implementation?

“On the design of LDPC codes within 0.0045 dB of the Shannon limit.”
[Chung, Forney, Richardson, Urbanke IT’01]

“A 47-Gbps LDPC decoder . . .” (10 GBase-T)
[Zhang, Anantharam, Wainwright, Nikolic JSSC ’10]

Wireless LAN and 60 GHz band
“Green” system = Capacity-code + “Green” implementation?

On the design of LDPC codes within 0.0045 dB of the Shannon limit.” [Chung, Forney, Richardson, Urbanke IT’01]

“A 47-Gbps LDPC decoder . . .” (10 GBase-T) [Zhang, Anantharam, Wainwright, Nikolic JSSC ’10]

Wireless LAN and 60 GHz band uncoded transmission!
What can the total power look like?
Some experiments
What can the total power look like?
Some experiments

Circuit-simulation-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]
What can the total power look like?
Some experiments

Circuit-simulation-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]
What can the total power look like?
Some experiments

Circuit-simulation-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]

Regular LDPCs
(3,4) (3,6) (4,8) (5,10)

\[
P_{\text{total}}
\]

Shannon waterfall

\[\log_{10}(P_e)\]

Power (Watts)

-20 -15 -10 -5

0 0.025 0.05 0.075 0.1

R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE
What can the total power look like?

Some experiments

Circuit-simulation-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]

Regular LDPCs
(3,4) (3,6) (4,8) (5,10)

\[ \log_{10}(P_e) \]

Power (Watts)

Shannon waterfall

\[ P_{total} \]

\[ P^* \]

R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE
What can the total power look like?
Some experiments

Circuit-simulation-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]

Regular LDPCs
(3,4) (3,6) (4,8) (5,10)

\[ \log_{10}(P_e) \]

Power (Watts)

Bad code choice?

R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE
What can the total power look like?
Some experiments

**Circuit-simulation-based modeling** of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]

![Graph showing the relationship between log10(P_e) and Power (Watts)]

- Regular LDPCs
  - (3,4)
  - (3,6)
  - (4,8)
  - (5,10)

**Parameters**
- R = 7 Gbps
- W = 7 GHz
- path-loss coeff. = 3
- d = 3 meters
- fc = 60 GHz
- STM CMOS 90 nm
- Cadence Encounter (Layout)
- Synopsis HSPICE

**Questions**
- Bad code choice?
- Bad implementations?
What can the total power look like?
Some experiments

Circuit-simulation-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]

Regular LDPCs
(3,4) (3,6) (4,8) (5,10)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>7 Gbps</td>
</tr>
<tr>
<td>W</td>
<td>7 GHz</td>
</tr>
<tr>
<td>path-loss coeff.</td>
<td>3</td>
</tr>
<tr>
<td>d</td>
<td>3 meters</td>
</tr>
<tr>
<td>fc</td>
<td>60 GHz</td>
</tr>
<tr>
<td>technology</td>
<td>STM CMOS 90 nm</td>
</tr>
<tr>
<td>Cadence Encounter</td>
<td>(Layout)</td>
</tr>
<tr>
<td>Synopsis HSPICE</td>
<td></td>
</tr>
</tbody>
</table>

\[
\log_{10}(P_e) \quad P_{total} \quad P_T^* \quad P_T
\]

Shannon waterfall

Bad code choice?  Bad implementations?  Both?
What can the total power look like?
Some experiments

Circuit-simulation-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]

Regular LDPCs
(3,4) (3,6) (4,8) (5,10)

Bad code choice? Bad implementations? Both? Neither?

$\log_{10}(P_e)$

Power (Watts)

R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE

Power (Watts)
Fundamental limits on total power?
Fundamental limits on total power?

\[ M \rightarrow \mathcal{E}_{P_T} \rightarrow D \rightarrow \hat{M} \]

- Fixed rate, distance
- Shannon limit (Tx power)

\[ \log_{10}(P_e) \]

Total power (watts)
Fundamental limits on total power?

![Diagram](image)

- Shannon limit (Tx power)
- Fixed rate, distance

\[
\log_{10}(P_e) \text{ versus Total power (watts)}
\]
Fundamental limits on total power?

Total power behavior

Shannon limit (Tx power)

fixed rate, distance

\[ \log_{10}(P_e) \]

Total power (watts)
Fundamental limits on total power?

\[ M \rightarrow \mathcal{P}_T \rightarrow \widehat{M} \]

Fixed rate, distance

Shannon limit (Tx power)

Total power behavior this?
Fundamental limits on total power?

$M \xrightarrow{P_T} \hat{M}$

Total power behavior

fixed rate, distance

Shannon limit (Tx power)

$\log_{10}(P_e)$

Total power (watts)

0 0.05 0.1 0.15 0.2

-30 -25 -20 -15 -10 -5
Questions for this talk:

1) Must $P_{total}$ always diverge to infinity?

2) Should we operate close to capacity?

3) Can we use bounded $P_T$?

4) What codes minimize $P_{total}$?
Our approach
Our approach

Circuit Models
Our approach

- Communication complexity
- Decentralized Estimation
- Network information theory

Information flow in circuits
Our approach

- Comm. complexity
- Decentralized Estimation
- Circuit Models
- Network information theory
- Circuit power models

Information flow in circuits

Encoding/decoding power

Decoding power (Watts)

$\log_{10}(P_e)$
Our approach

- Comm. complexity
- Decentralized Estimation
- Network information theory
- Circuit power models
- Circuit Models

Information flow in circuits

Decoding power (Watts)

Encoding/decoding power

Information theory (Transmit power)

$\log_{10}(P_e)$ vs. Decoding power (Watts)

Shannon limit (Tx power)
Our approach

- Shannon limit (Tx power)
- Information theory
- Circuit power models
- Network information theory
- Decentralized Estimation
- Comm. complexity

Information flow in circuits

Encoding/decoding power

Information theory (Transmit power)

Fundamental bounds on total power
Obtaining fundamental limits:
Main difficulty

Code

passing through channel

Transmit power $P_T$

Models:
- AWGN
- fading
  
  ...
Obtaining fundamental limits: Main difficulty

- Code
- passing through channel
- Transmit power $P_T$

Models:
- AWGN fading
- ...

implementation (encoding/decoding)

encoding/decoding power

Models: ??
Obtaining fundamental limits: Main difficulty

Need implementation models that are relevant and analyzable
A “graphical” model for circuit implementation

“VLSI model” of computing [Thompson ’80]
- degree of each node at most $\alpha$ (e.g. 4)
A “graphical” model for circuit implementation

“VLSI model” of computing [Thompson ’80]
- degree of each node at most $\alpha$ (e.g. 4)

Channel model:

Binary Symmetric Channel

\[
p_{ch} = \varnothing \left( \sqrt{\frac{\eta P_T}{N_0}} \right)
\]
Decoding complexity/power: Fundamental limits

**Theorem** [Grover, Woyach, Sahai ISIT ’08, JSAC ’11]

\[
\text{# of clock-cycles, } \tau \geq \frac{1}{\log(\alpha - 1)} \log \left( \frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right)
\]

\ldots \text{for any code and any decoding algorithm.}

\(\alpha\) : Max. node degree

\(V(P_T)\) : “channel dispersion”

* precise result for any \(P_e\) and any gap appears in [Grover, Woyach, Sahai ’11]
Decoding complexity:
Proof Technique
Decoding complexity:
Proof Technique

Information-flow constrained by “fan-out”
Decoding complexity:
Proof Technique

Information-flow constrained by “fan-out” $B_i$
Decoding complexity: Proof Technique

Information-flow constrained by “fan-out”

\[ B_i \]

\[ \tau = 1 \]
Decoding complexity:
Proof Technique

Information-flow constrained by “fan-out”

\[ B_i \]

\[ \tau = 1 \]

\[ \tau = 2 \]
Decoding complexity: Proof Technique

Information-flow constrained by “fan-out”
Decoding complexity:
Proof Technique

Information-flow constrained by “fan-out”

\[ B_i \]

\[ \tau = 1 \]
\[ \tau = 2 \]
\[ \tau = 3 \]

“Sphere-packing” for decoding neighborhoods

[Grover ’07][Grover, Sahai ’08][Grover, Woyach, Sahai ’11]
Decoding complexity:
Proof Technique

Information-flow constrained by “fan-out”

\[ B_i \]

\( \tau = 1 \)
\( \tau = 2 \)
\( \tau = 3 \)

“Sphere-packing” for decoding neighborhoods

\([\text{Grover '07}][\text{Grover, Sahai '08}][\text{Grover, Woyach, Sahai '11}]\)

\( n_{\text{nbd}} \) : number of nodes in decoding neighborhood
Decoding complexity:
Proof Technique

Information-flow constrained by “fan-out”

“Sphere-packing” for decoding neighborhoods

\[ B_i \]

\[ \tau = 1 \]

\[ \tau = 2 \]

\[ \tau = 3 \]

\[ n_{nbd} : \text{number of nodes in decoding neighborhood} \]

\[ P_e \gtrsim e^{-n_{nbd} \frac{(C(P_T)-R)^2}{V(P_T)}} \]
Decoding complexity:
Proof Technique

Information-flow constrained by “fan-out”

“Sphere-packing” for decoding neighborhoods

$B_i$

$\tau = 1$
$\tau = 2$
$\tau = 3$

$P_e \geq e^{-n_{nbd}} \frac{(C(P_T) - R)^2}{V(P_T)}$

$\tau \geq \frac{1}{\log(\alpha - 1)} \log(n_{nbd})$

$n_{nbd}$: number of nodes in decoding neighborhood
Decoding complexity: Proof Technique

Information-flow constrained by “fan-out”

“Sphere-packing” for decoding neighborhoods

\[ B_i \]

\[ \tau = 1 \]
\[ \tau = 2 \]
\[ \tau = 3 \]

\[ n_{nbd} : \text{number of nodes in decoding neighborhood} \]

\[ P_e \gtrsim e^{-n_{nbd}} \frac{(C(P_T) - R)^2}{V(P_T)} \]

\[ \tau \gtrsim \frac{1}{\log(\alpha - 1)} \log(n_{nbd}) \]

Theorem *[Grover, Woyach, Sahai ISIT '08, JSAC '11]

\[ \# \text{ of clock-cycles}, \tau \gtrsim \frac{1}{\log(\alpha - 1)} \log \left( \frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right) \]
Node model:
Fundamental bounds on total power

**Node model**: each node consumes constant amount of energy per clock cycle
Node model:
Fundamental bounds on total power

Each node consumes constant amount of energy per clock cycle

$$P_{total} = \min_{P_T} P_T + P_{decoding}$$
Node model:
Fundamental bounds on total power

**Node model:** each node consumes constant amount of energy per clock cycle

\[
P_{total} = \min_{P_T} P_T + P_{decoding} \\
\geq \min_{P_T} P_T + \gamma \log \left( \frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right)
\]
Node model:
Fundamental bounds on total power

Node model: each node consumes constant amount of energy per clock cycle

\[ P_{total} = \min_{P_T} P_T + P_{decoding} \]
\[ \geq \min_{P_T} P_T + \gamma \log \left( \frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right) \]

Enode = 3 pJ
distance = 10 m
fc = 60 GHz
W = 3 GHz
R = 1.5 Gbps
Node model:

Fundamental bounds on total power

**Node model**: each node consumes constant amount of energy per clock cycle

\[ P_{total} = \min \limits_{P_T} P_T + P_{decoding} \]

\[ \geq \min \limits_{P_T} P_T + \gamma \log \left( \frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right) \]

- **Shannon limit** (Tx power)
- **Total power (Watts)**
- **Enode = 3 pJ**
- **distance = 10 m**
- **fc = 60 GHz**
- **W = 3 GHz**
- **R = 1.5 Gbps**

**Enode** = 3 pJ
**distance** = 10 m
**fc** = 60 GHz
**W** = 3 GHz
**R** = 1.5 Gbps
Node model:

Fundamental bounds on total power

**Node model**: each node consumes constant amount of energy per clock cycle

\[
P_{\text{total}} = \min_{P_T} P_T + P_{\text{decoding}}
\]

\[
\geq \min_{P_T} P_T + \gamma \log \left( \frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right)
\]

**Figure**: Graph showing the relationship between total power and Shannon limit power (Tx power).

- **Shannon limit** (Tx power)
- **Total power (Watts)**
- **Log10(Pe)**

**Parameters**:
- E_node = 3 pJ
- Distance = 10 m
- fc = 60 GHz
- W = 3 GHz
- R = 1.5 Gbps
**Node model:** each node consumes constant amount of energy per clock cycle

\[ P_{total} = \min_{P_T} P_T + P_{decoding} \]

\[ \geq \min_{P_T} P_T + \gamma \log \left( \frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right) \]

**Fundamental bounds on total power**

**Logarithmic Shannon limit**

\[ \log_{10}(P_e) \]

**Shannon limit**

**Total power (Watts)**

\[ 0 \quad 0.05 \quad 0.1 \quad 0.15 \quad 0.2 \quad 0.25 \]

**Moral:** do not operate too close to capacity!

- Node model: each node consumes constant amount of energy per clock cycle
- Enode = 3 pJ
- distance = 10 m
- fc = 60 GHz
- W = 3 GHz
- R = 1.5 Gbps
Node model
What codes are good?
Node model
What codes are good?

Capacity-approaching LDPCs, iterative decoding?
### Node model

**What codes are good?**

<table>
<thead>
<tr>
<th>Capacity-approaching LDPCs, iterative decoding?</th>
<th>Random block codes, ML decoding?</th>
</tr>
</thead>
</table>


Node model

What codes are good?

Capacity-approaching **LDPCs**, iterative decoding?

**Convolutional** codes, Viterbi decoding

Random block codes, ML decoding?
## Node model

### What codes are good?

<table>
<thead>
<tr>
<th>Capacity-approaching LDPCs, iterative decoding?</th>
<th>Random block codes, ML decoding?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convolutional codes, Viterbi decoding</td>
<td>Polar codes, successive cancelation?</td>
</tr>
</tbody>
</table>
Node model
What codes are good?

Capacity-approaching LDPCs, iterative decoding?

Convolutional codes, Viterbi decoding

Random block codes, ML decoding?
Polar codes, successive cancelation?
# Node model

## What codes are good?

<table>
<thead>
<tr>
<th>Capacity-approaching LDPCs, iterative decoding?</th>
<th>Random block codes, ML decoding?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convolutional codes, Viterbi decoding</td>
<td>Polar codes, successive cancelation?</td>
</tr>
</tbody>
</table>

Regular LDPCs, iterative decoding [Lentmaier ’05][Grover, Woyach, Sahai JSAC’11]
Node model
What codes are good?

- Capacity-approaching LDPCs, iterative decoding?
- Convolutional codes, Viterbi decoding
- Random block codes, ML decoding?
- Polar codes, successive cancellation?

Regular LDPCs, iterative decoding [Lentmaier ’05][Grover, Woyach, Sahai JSAC’11]

![Graph showing the performance of (3,4)-LDPC compared to rate 1/4 lower bound. The gap (in log-scale, i.e. dBs) between performance of (3,4)-LDPC and rate 1/4 lower bound is bounded by a constant even as $P_e \to 0$. The Shannon limit, total power (Watts) vs. $\log_{10}(P_e)$ graph is shown, with a gap of 4.8 dB.]
Node model
What codes are good?

- Capacity-approaching LDPCs, iterative decoding?
- Convolutional codes, Viterbi decoding
- Random block codes, ML decoding?
- Polar codes, successive cancelation?

Regular LDPCs, iterative decoding [Lentmaier ’05][Grover, Woyach, Sahai JSAC’11]

Data-centers (10 GBASE-T): regular LDPC!
Talk outline

Information flows: visible and invisible

$M \rightarrow P_T \rightarrow \hat{M}$

**Traditional:** Minimize $P_T$

**Now:** Minimize total (transmit + proc.) power

Fundamental limits

*Node model:* stay away from capacity

regular LDPCs are order optimal
Talk outline

Information flows: visible and invisible

Traditional: Minimize $P_T$

Now: Minimize total (transmit + proc.) power

Fundamental limits

Node model: stay away from capacity

regular LDPCs are order optimal

Wire model: using bounded $P_T$ suboptimal
What did we expect from our experimental results?

Experiment-based modeling of decoding power
[Ganesan, Grover, Rabaey SiPS’11][Ganesan, Wen, Grover, Rabaey’12]

Regular LDPCs
(3,4) (3,6) (4,8) (5,10)

$P_{T}$

$P_{total}$

$R = 7 \text{ Gbps}
W = 7 \text{ GHz}
\text{path-loss coefft} = 3
d = 3 \text{ meters}
f_c = 60 \text{ GHz}
\text{STM CMOS 90 nm}
\text{Cadence Encounter (Layout)}
\text{Synopsis HSPICE}$
Wire model of power consumption
Wire model of power consumption
Wire model of power consumption

Fourier Transform [Thompson ’80]

\[ \text{Power} \propto \frac{A_{\text{wires}} \tau}{n} = \Omega \left( \sqrt{n \log n} \right) \]

for implementing any algorithm.

Similar results for sorting, boolean functions, . . .
Wire model of power consumption

Fourier Transform [Thompson ’80]

\[ \text{Power} \propto \frac{A_{\text{wires}} \tau}{n} = \Omega \left( \sqrt{n \log n} \right) \] for implementing any algorithm.

Similar results for sorting, boolean functions, . . .

Encoding/decoding complexity: [El Gamal et al. ’84]

\[ A_{\text{chip}} \tau^2 \geq \Omega \left( n R^2 \log \frac{n}{P_e} \right) \]
Wire model:
Fundamental limits on encoding/decoding power

**Theorem** *[Grover, Goldsmith, Sahai Allerton’11; ISIT Sub.’12]*

Encoding/decoding *Power* \( \geq \Omega \left( \sqrt{\frac{\log \frac{1}{P_e}}{P_T}} \right) \)

\[ \text{\ldots for any code and any encoding/decoding algorithm.} \]

* precise result for any \( P_e \) and any gap appears in [Grover, Goldsmith, Sahai ’12]
Wire model:

Fundamental limits on total power

\[ P_{\text{total}} = \min_{P_T} P_T + P_{\text{decoding}} + P_{\text{encoding}} \]

\[ \geq \min_{P_T} P_T + 2\beta \sqrt{\log \frac{1}{P_e}} \frac{1}{P_T} \]
Wire model:

Fundamental limits on on total power

\[ P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding} \]

\[ \gtrsim \min_{P_T} P_T + 2\beta \sqrt{\log \frac{1}{P_e}} P_T \]

\( P_{total} \) vs \( \log_{10}(P_e) \)

- Power (Watts)
- \( \log_{10}(P_e) \)

- CMOS 90 nm tech.
- d = 0.5 m
- path-loss coeff = 2
- Capacitance = 1 fF/um
Wire model:
Fundamental limits on total power

\[ P_{\text{total}} = \min_{P_T} P_T + P_{\text{decoding}} + P_{\text{encoding}} \]

\[ \geq \min_{P_T} P_T + 2\beta \sqrt{\log \frac{1}{P_e}} \]

\[ d = 0.5 \text{ m} \]
path-loss coeff = 2
Capacitance = 1 fF/um
CMOS 90 nm tech.
Wire model:
Fundamental limits on on total power

\[ P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding} \]

\[ \geq \min_{P_T} P_T + 2\beta \sqrt{\log \frac{1}{P_e}} \]

\[ P_{total} \approx 3 \log \frac{1}{P_e} \]

\[ d = 0.5 \text{ m} \]
\[ \text{path-loss coeff} = 2 \]
\[ \text{Capacitance} = 1 \text{ fF/um} \]
\[ \text{CMOS 90 nm tech.} \]
Wire model:
Fundamental limits on on total power

\[ P_{\text{total}} = \min_{P_T} P_T + P_{\text{decoding}} + P_{\text{encoding}} \]

\[ \geq \min_{P_T} P_T + 2\beta \sqrt{\log \frac{1}{P_e}} \]

\[ P^*_{T} \]

\[ P_{\text{total}} \]

with bounded \( P_T \)

\( d = 0.5 \) m

path-loss coeff = 2

Capacitance = 1 fF/um

CMOS 90 nm tech.
Wire model:
Fundamental limits on total power

\[ P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding} \]

\[ \geq \min_{P_T} P_T + 2\beta \sqrt{\log \frac{1}{P_e}} \]

\[ \sim \frac{3}{2} \log \frac{1}{P_e} \]

\[ d = 0.5 \text{ m} \]
\[ \text{path-loss coeff} = 2 \]
\[ \text{Capacitance} = 1 \text{ fF/um} \]
\[ \text{CMOS 90 nm tech.} \]
Wire model:
Fundamental limits on on total power

\[ P_{\text{total}} = \min_{P_T} P_T + P_{\text{decoding}} + P_{\text{encoding}} \]

\[ \gtrsim \min_{P_T} P_T + 2\beta \sqrt{\log \frac{1}{P_e}} \]

**Moral:** Bounded \( P_T \) is suboptimal
Suboptimality of bounded transmit power: Experimental results
Information flows in circuits:

Summary

\( M \rightarrow P_T \rightarrow \widehat{M} \)

Information needs to flow at the encoder and decoder

(mixing) (de-mixing)

Fundamental limits:

**Traditional**: Operate close to capacity

**Total power perspective**: **Node model**: stay away from capacity

**Wire model**: use unbounded \( P_T \)

No assumptions on the code structure
Looking forward:
“Greenest” codes for wireless?

**Grand goal:**
Code/encoder/decoder triples within bounded dBs of the optimal.
Looking forward:
“Greenest” codes for wireless?

**Grand goal:**
Code/encoder/decoder triples within bounded dBs of the optimal.

Distance (meters)

\[
\log_{10}(P_e)
\]

<table>
<thead>
<tr>
<th>Distance (meters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

Values for \(P_e\): -5, -10, -15, -20
Looking forward:
“Greenest” codes for wireless?

**Grand goal:**
Code/encoder/decoder triples within bounded dBs of the **optimal**.

Distance (meters)

<table>
<thead>
<tr>
<th>Distance (meters)</th>
<th>Log10(Pe)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-5</td>
</tr>
<tr>
<td>4</td>
<td>-10</td>
</tr>
<tr>
<td>6</td>
<td>-15</td>
</tr>
<tr>
<td>8</td>
<td>-20</td>
</tr>
</tbody>
</table>

(3,6), girth-6, 1-bit msg
(4,8), girth-6, 1-bit msg
(3,6), girth-6, 2-bit msg
(3,6), girth-8, 1-bit msg
(3,6)-LDPC girth-8, 2-bit msg

- R = 7 Gbps
- W = 7 GHz
- path-loss coeff = 3
- fc = 60 GHz
- STM CMOS 90 nm
- Layout: Cadence Enctr
- Synopsis HSPICE

[Ganesan, Grover, Rabaey ’11][Ganesan, Wen, Grover, Rabaey, Goldsmith ’12]
Looking forward:
“Greenest” codes for wireless?

Grand goal:
Code/encoder/decoder triples within bounded dBs of the optimal.

Distance (meters)

---

Uncoded transmission

(3,6), girth-8, 1-bit msg

(3,6)-LDPC girth-8, 2-bit msg

---

R = 7 Gbps
W = 7 GHz
path-loss coeff = 3
fc = 60 GHz
STM CMOS 90 nm
Layout: Cadence Enctr
Synopsis HSPICE

[Ganesan, Grover, Rabaey ’11][Ganesan, Wen, Grover, Rabaey, Goldsmith ’12]
Looking forward:
“Green” data-center communication

Traditional view

Total-power perspective

Code choice (10-Gbase-T):
(6,32) - “RS-LDPC”, length 2048
Looking forward:
“Green” data-center communication

Traditional view

Total-power perspective

Code choice (10-Gbase-T):
(6,32) - “RS-LDPC”, length 2048
Looking forward:
“Green” data-center communication

Traditional view

Total-power perspective

Code choice (10-Gbase-T):
(6,32) - “RS-LDPC”, length 2048

Figure 3.1: FER (dotted lines) and BER (solid lines) performance of the Q4.2 sum-product decoder of the (2048,1723) RS-LDPC code using different number of decoding iterations.

3.1 Characterization of Error Events
Hardware emulations reveal that the errors dominating the error floors do not resemble the errors occurring in the waterfall region of the BER-SNR curve. Errors in the waterfall region are mostly random-like errors with bit error count greater than the minimum distance of the code. At higher SNR levels where the error floors occur, random-like errors are very rare, and instead the dominant errors exhibit rather pronounced characteristics, either oscillatory or absorbing. Both these types of errors appear to start with a small number of bits that are received incorrectly. An oscillation error is illustrated in Fig. 3.2 showing the soft decisions after each decoding iteration. The horizontal axis is for each of the 2048 bits in the code block, and the vertical axis is the soft decision each bit assumes after an iteration of decoding. For simplicity of illustration, it is assumed that all-zeros
Looking forward:
“Green” data-center communication

Traditional view

Code choice (10-Gbase-T):
(6,32) - “RS-LDPC”, length 2048

Total-power perspective

Distance (meters)

Code-choice must depend on distance

Figure 3.1: FER (dotted lines) and BER (solid lines) performance of the Q4.2 sum-product decoder of the (2048,1723) RS-LDPC code using different numbers of decoding iterations.

3.1 Characterization of Error Events

Hardware emulations reveal that the errors dominating the error floors do not resemble the errors occurring in the waterfall region of the BER-SNR curve. Errors in the waterfall region are mostly random-like errors with bit error rate greater than the minimum distance of the code. At higher SNR levels, where the error floors occur, random-like errors are very rare, and instead the dominant errors exhibit rather pronounced characteristics, either oscillatory or absorbing. Both these types of errors appear to start with a small number of bits that are received incorrectly. An oscillation error is illustrated in Fig. 3.2 showing the soft decisions after each decoding iteration. The horizontal axis is for each of the 2048 bits in the code block, and the vertical axis is the soft decision each bit assumes after an iteration of decoding. For simplicity of illustration, it is assumed that all-zeros
Looking forward:
Energy costs of communication
Looking forward:
Energy costs of communication

Other models of computing
- e.g. nodes/wires can sleep
Looking forward:
Energy costs of communication

Other models of computing
- e.g. nodes/wires can sleep

Multiple simultaneously transmitting links
- Code-choice must depend on distance, density
  [Grover, Woyach, Sahai JSAC ’11]
Looking forward:
Energy costs of communication

Other models of computing
- e.g. nodes/wires can sleep

Multiple simultaneously transmitting links
- Code-choice must depend on distance, density
  [Grover, Woyach, Sahai JSAC ’11]

Other components of system power
- Equalizer [Grover, Sahai, Park ’11]
Talk outline

1) Information flows in circuits

\[ M \xrightarrow{\text{Transmitter \ $P_T$}} \text{Channel} \xrightarrow{\text{Receiver}} \hat{M} \]

**Communicating** and **processing** information

2) Information flows in physical systems

**Communicating** and **Using** information

“Implicit” communication

Witsenhausen’s counterexample

- approximately optimal solutions

Beyond the counterexample
Looking forward:
Implicit information flows in control
Looking forward:
Implicit information flows in control
Looking forward: Implicit information flows in control

Communication theory

\[ M \xrightarrow{\mathcal{E}_{P_T}} D \xrightarrow{\hat{M}} \]

Plant

\[ C_1 \]
\[ C_2 \]
\[ C_3 \]
\[ C_4 \]
\[ C_5 \]
\[ C_6 \]
Looking forward:
Implicit information flows in control

Communication theory

Control theory

Plant

Control agent

$M \rightarrow \mathcal{E}_{PT} \rightarrow \mathcal{D} \rightarrow \hat{M}$

$\mathcal{C}_1$  $\mathcal{C}_2$  $\mathcal{C}_3$  $\mathcal{C}_4$  $\mathcal{C}_5$  $\mathcal{C}_6$
Looking forward:
Implicit information flows in control

Communication theory

Control theory

Plant

Control agent

$\mathcal{E}_{PT} \xrightarrow{M} \mathcal{D} \xrightarrow{\hat{M}}$
Looking forward: Implicit information flows in control

“... semantic aspects of communication are irrelevant to the engineering problem.” [Shannon ‘48]
“[comm theory] deals with an essentially simple problem, because the transmission of information is considered independently of its use.” [Witsenhausen ’71]
Looking forward:
Implicit information flows in control

Communication theory

Control theory

Plant

Control agent

$M \xrightarrow{\mathcal{E}} \hat{M}$

$C_1 \xleftarrow{} C_2 \xrightarrow{} C_3$

$C_4 \xleftarrow{} C_5 \xrightarrow{} C_6$
Looking forward:
Implicit information flows in control

Communication theory

Plant

Message?

Control theory

Control agent

Plant

$M \xrightarrow{\mathcal{E}_{PT}} D \xrightarrow{\hat{M}}$
Looking forward:
Implicit information flows in control

Communication theory

Message? a choice

Control theory

Plant
Looking forward:
Implicit information flows in control

Communication theory

Message? a choice

Control theory
Looking forward:
Implicit information flows in control

- Control
- Communicate

Message? a choice

Dual role of actions
- control
- communicate
Looking forward: Implicit information flows in control

Dual role of actions
- control
- communicate

Communication theory

Message? a choice

Implicit channel

Plant

Control theory

\[ M \xrightarrow{\mathcal{E}_{PT}} M \]

\[
\begin{align*}
C_1 & \quad C_2 \\
C_6 & \quad C_5 & \quad C_4 & \quad C_3
\end{align*}
\]

Implicit channel Plant

\[
\begin{align*}
& x_0 \quad x_1 \quad x_2 \\
& C_1 \quad u_1 \quad y \quad C_2 \quad u_2 \\
& z \sim \mathcal{N}(0, 1)
\end{align*}
\]
Looking forward: Implicit information flows in control

- Dual role of actions
  - control
  - communicate

Implicit channel

Communication theory

Message? a choice

Plant

Control theory
Breakdown of traditional theory: Witsenhausen’s counterexample

\[ x_0 \sim \mathcal{N}(0, \sigma_0^2) \quad \text{min} \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\} \]

Linear, Quadratic, Gaussian

\[ z \sim \mathcal{N}(0, 1) \]  

[Witsenhausen ’68]
Breakdown of traditional theory: Witsenhausen’s counterexample

\[ x_0 \sim \mathcal{N}(0, \sigma_0^2) \quad \min \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\} \]

Linear, Quadratic, Gaussian

[Witsenhausen ’68]
Breakdown of traditional theory: Witsenhausen’s counterexample

\[ x_0 \sim \mathcal{N}(0, \sigma_0^2) \quad \text{min} \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\} \]

Linear, Quadratic, Gaussian

\[ z \sim \mathcal{N}(0, 1) \]

[Witsenhausen ’68]
Breakdown of traditional theory: Witsenhausen’s counterexample

\[ x_0 \sim \mathcal{N}(0, \sigma_0^2) \quad \text{min} \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\} \]

Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen '68]
Breakdown of traditional theory: Witsenhausen’s counterexample

$$x_0 \sim \mathcal{N}(0, \sigma_0^2) \quad \text{min} \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\}$$

Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen ’68]

Nonconvex [Witsenhausen ’68], “Intractable” (NP-hard) [Papadimitriou, Tsitsiklis ’86]
Breakdown of traditional theory: Witsenhausens’s counterexample

\[ x_0 \sim \mathcal{N}(0, \sigma_0^2) \quad \min \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\} \]

Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen ’68]

Nonconvex [Witsenhausen ’68], “Intractable” (NP-hard) [Papadimitriou, Tsitsiklis ’86]

Semi-exhaustive searches
[Denk, Ho ’99] [Baglietto et al. ’01]
[Lee, Lau, Ho ’01][Li, Marden, Shamma ’09]
[Karlsson et al. ’10]
Breakdown of traditional theory: Witsenhausen’s counterexample

\[ x_0 \sim \mathcal{N}(0, \sigma_0^2) \quad \text{min} \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\} \]

Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen '68]

Nonconvex [Witsenhausen '68], "Intractable" (NP-hard) [Papadimitriou, Tsitsiklis '86]

Semi-exhaustive searches
[Deng, Ho '99] [Baglietto et al. '01]
[Lee, Lau, Ho '01] [Li, Marden, Shamma '09]
[Karlsson et al. '10]
Our approach: understanding information flows to obtain fundamental limits on cost
Our approach: understanding information flows to obtain fundamental limits on cost

Original problem
[Witsenhausen ’68]
Our approach: understanding information flows to obtain fundamental limits on cost

Original problem
[Witsenhausen '68]

Identify information-flow
[Grover, Sahai '08]
Our approach: understanding information flows to obtain fundamental limits on cost.

Original problem
[Witsenhausen ’68]

Identify information-flow
[Grover, Sahai ’08]

Understand information-flow:
Binary version [Avestimehr et al. ’08][Grover, Sahai ’09]
Our approach: understanding information flows to obtain fundamental limits on cost

Original problem [Witsenhausen '68]

Asymptotic vector version [Grover, Sahai '08]

Identify information-flow [Grover, Sahai '08]

Understand information-flow:
Binary version [Avestimehr et al. '08][Grover, Sahai '09]
Our approach: understanding information flows to obtain fundamental limits on cost

Original problem
[Witsenhausen '68]

Identify information-flow
[Grover, Sahai '08]

Asymptotic vector version
[Grover, Sahai '08]

Large-deviations/
“sphere-packing”
[Grover, Park, Sahai '09, '12]

Understand information-flow:
Binary version [Avestimehr et al. '08][Grover, Sahai '09]
Approx. solutions to Witsenhausen’s counterexample

\[
\begin{align*}
\min \{ k^2 \mathbb{E}[u_1^2] + \mathbb{E}[x_2^2] \}
\end{align*}
\]

**Theorem** [Grover, Park, Sahai IEEE Tran Auto Cont’12]

For all \((k, \sigma_0)\),

\[
\begin{align*}
\text{Upper bound} & \leq 8 \\
\text{Lower bound} & \leq 8
\end{align*}
\]

**Upper bound**
Scalar quantization

**Lower bound**
Info-theory + Large deviations

First approximately-optimal solutions to Witsenhausen’s counterexample
Approx. solutions to Witsenhausen’s counterexample

Theorem [Grover, Park, Sahai IEEE Tran Auto Cont’12]

For all \((k, \sigma_0)\), \[ \frac{\text{Upper bound}}{\text{Lower bound}} \leq 8 \]

\[
\min \left\{ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ x_2^2 \right] \right\}
\]

Upper bound
- Scalar quantization
- “Watermarking” (“dirty-paper” coding)

Lower bound
- Info-theory + Large deviations

First approximately-optimal solutions to Witsenhausen’s counterexample
Looking forward:
Information-flows in Cyber-Physical Systems
Looking forward:
Information-flows in Cyber-Physical Systems
Looking forward:
Information-flows in Cyber-Physical Systems

$C_1$ $\rightarrow$ $y_1$

$C_2$ $\rightarrow$ $u_2$

System
Looking forward:
Information-flows in Cyber-Physical Systems
Looking forward:
Information-flows in Cyber-Physical Systems

\[
\begin{align*}
C_1 & \quad \text{sensor} \\
\text{External channel} & \quad C_2 \\
\downarrow & \quad \downarrow \\
y_1 & \quad u_2
\end{align*}
\]
Looking forward:
Information-flows in Cyber-Physical Systems
Looking forward:
Information-flows in Cyber-Physical Systems
Looking forward:
Information-flows in Cyber-Physical Systems

System

$C_1$

sensor

$y_1$

$C_2$

actuator

$u_2$

External channel

32/33
Looking forward:
Information-flows in Cyber-Physical Systems

Communicating a evolving state
[Bansal, Basar '89][Wong, Brockett '90][Tatikonda, Mitter '01][Sahai, Mitter '06][Sinopoli et al. '04] . . .
Looking forward:
Information-flows in Cyber-Physical Systems

Communicating a evolving state
[Bansal, Basar ’89][Wong, Brockett ’90][Tatikonda, Mitter ’01]
[Sahai, Mitter ’06] [Sinopoli et al. ’04] . . .
Looking forward: Information-flows in Cyber-Physical Systems

Communicating a evolving state
[Bansal, Basar ’89][Wong, Brockett ’90][Tatikonda, Mitter ’01]
[Sahai, Mitter ’06] [Sinopoli et al. ’04] . . .

[Grover, Sahai ’10]
[Grover ’12]
Looking forward:
Information-flows in Cyber-Physical Systems

Communicating a evolving state
[Bansal, Basar ’89][Wong, Brockett ’90][Tatikonda, Mitter ’01]
[Sahai, Mitter ’06] [Sinopoli et al. ’04] . . .
Summary:
Understanding hidden information-flows is useful

Information flows in circuits

Information flows in physical systems

Information flow in computing: fundamental bounds on total power

Implicit information flows: approximately-optimal solutions
Backup slides
The appropriate approximations in the two bounds

Node model

Wire model

Dispersion approximation

Approximation far from capacity

Sphere-packing exponent

\( E_{sp}(R) \)
Total power minimizing density

![Graph showing density vs. total power](image)

- **Optimal code** (if decoding were free)
- **Uncoded transmission**
- **Outer bound** with $E_{\text{node}} = 3 \text{ pJ}$
- **Extrapolated performance of the decoder of [Zhang et al.]**
Wireless information transfer
Wireless information transfer

\[ M \]

\[ L_1 \]

\[ L_2 \]

\[ r_s \]

\[ v_s \]

\[ i_s \]

\[ r_i \]

\[ c_l \]

\[ i_i \]

\[ i_2 \]

\[ \text{Load} \]
Wireless information transfer
Optimal information transfer strategy?
Optimal information transfer strategy?

![Graph showing the relationship between frequency (Hz) in log-scale and N(t)/n(t).](image)
Optimal information transfer strategy?

\[ C(P^{avail}) = \max_{P_1, \ldots, P_n \text{ s.t.} \sum_{i=1}^{n} P_i \leq P^{avail}} \sum_{i=1}^{n} \log \left( 1 + \frac{\eta_i P_i}{N} \right). \]
Optimal information transfer strategy?

\[ C(P_{\text{avail}}) = \max_{P_1, \ldots, P_n \text{ s.t. } \sum_{i=1}^{n} P_i \leq P_{\text{avail}}} \sum_{i=1}^{n} \log \left( 1 + \frac{\eta_i P_i}{N} \right). \]
## Secure Communication Using Inductive Coupling

<table>
<thead>
<tr>
<th>Inductive coupling</th>
<th>Quantum entanglement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement process causes detuning</td>
<td>Measurement process causes waveform collapse</td>
</tr>
<tr>
<td>Eavesdropper detection through spectrum sensing</td>
<td>Eavesdropper detection through physics</td>
</tr>
<tr>
<td>Information destruction through capacity-approaching codes and strong converse of channel coding theorem</td>
<td>Information destruction through physics</td>
</tr>
<tr>
<td><strong>Inductive key distribution</strong></td>
<td>Quantum key distribution</td>
</tr>
</tbody>
</table>
A deterministic abstraction

\[
x_0 \to + \x_1 \to + \x_2 \to - \x_0 \\
\tilde{x}_1 = u_2 = \hat{x}_1 \\
E = x_0 u_1 u_2 \equiv \hat{x}_1 \\
\mathcal{E} \quad \mathcal{D} \\
\mathcal{E} \quad \mathcal{D} \\
\z \sim \mathcal{N}(0, 1) \\
R_{ex} \\
k^2 \mathbb{E} [u_1^2] + \mathbb{E} [ (x_1 - \hat{x}_1)^2 ]
A deterministic abstraction

\[ x_0 \rightarrow \mathcal{E} \rightarrow u_1 \rightarrow x_1 \rightarrow \mathcal{D} \rightarrow u_2 = \hat{x}_1 \]

\[ z \sim \mathcal{N}(0, 1) \]

\[ R_{ex} \]

\[ k^2 \mathbb{E} [u_1^2] + \mathbb{E} [(x_1 - \hat{x}_1)^2] \]
A deterministic abstraction

\[ x_0 \xrightarrow{+} x_1 \xrightarrow{+} x_2 \]

\[ \mathcal{E} \quad u_1 \]

\[ D \quad u_2 = \hat{x}_1 \]

\[ z \sim \mathcal{N}(0, 1) \]

\[ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ (x_1 - \hat{x}_1)^2 \right] \]

\[ \mathcal{D} \quad x_0 \]

\[ b_1 \quad b_2 \quad b_3 \quad b_4 \quad b_5 \]

\[ x_0 \xrightarrow{+} x_1 \xrightarrow{+} \hat{x}_1 \]
A deterministic abstraction

\[ x_0 \xrightarrow{+} x_1 \xrightarrow{+} x_2 \]

\[ u_1 \xrightarrow{+} \hat{x}_1 \]

\[ z \sim \mathcal{N}(0, 1) \]

\[ k^2 \mathbb{E} \left[ u_1^2 \right] + \mathbb{E} \left[ (x_1 - \hat{x}_1)^2 \right] \]

\[ x_0 \xrightarrow{b_1} b_2 \xrightarrow{b_3} b_4 \xrightarrow{b_5} x_1 \]

\[ \hat{x}_1 \xrightarrow{D} z_2 \]

\[ u_1 \xrightarrow{+} x_1 \xrightarrow{+}  \]

\[ \mathcal{E} \xrightarrow{D} u_2 = \hat{x}_1 \]

\[ R_{ex} \]
A deterministic abstraction

\[ x_0 \xrightarrow{+} x_1 \xrightarrow{+} x_2 \]

\[ z \sim \mathcal{N}(0, 1) \]

\[ k^2 \mathbb{E} [u_1^2] + \mathbb{E} \left[ (x_1 - \hat{x}_1)^2 \right] \]
A deterministic abstraction

\[ x_0 \xrightarrow{\mathcal{E}} u_1 \xrightarrow{R_{ex}} z \sim \mathcal{N}(0,1) \]

\[ k^2 \mathbb{E} [u_1^2] + \mathbb{E} [(x_1 - \hat{x}_1)^2] \]

\[ R_{ex} = 2 \]
A deterministic abstraction

\[ x_0 \xrightarrow{+} x_1 \xrightarrow{+} x_2 \]

\[ u_1 \xrightarrow{\mathcal{E}} x_0 \]

\[ z \sim \mathcal{N}(0, 1) \]

\[ k^2 \mathbb{E} [u_1^2] + \mathbb{E} [(x_1 - \hat{x}_1)^2] \]

\[ R_{ex} = 2 \]
Strategy for deterministic abstraction

\[ x_0, b_1, b_2, b_3, b_4, b_5 \rightarrow x_1 \rightarrow D \rightarrow \hat{x}_1 \]

\[ \mathcal{E} \rightarrow u_1 \rightarrow z_2 \rightarrow R_{ex} = 2 \]
Strategy for deterministic abstraction
Strategy for deterministic abstraction
Strategy for deterministic abstraction

$x_0 \rightarrow b_1 \rightarrow b_2 \rightarrow b_3 \rightarrow b_4 \rightarrow b_5 \rightarrow x_1 \rightarrow \hat{x}_1$

$R_{ex} = 2$
Strategy for deterministic abstraction

\[ x_0, b_1, b_2, b_3, b_4, b_5 \]

\[ \mathcal{E} \]

\[ R_{ex} = 2 \]

\[ \hat{x}_1 \]
Strategy for deterministic abstraction

\[ x_0 \quad b_1 \quad b_2 \quad b_3 \quad b_4 \quad b_5 \]

\[ \mathcal{E} \quad u_1 \quad z_2 \]

\[ R_{ex} = 2 \]

\[ \hat{x}_1 \]

\[ D \quad b_1 \quad b_2 \]
Strategy for deterministic abstraction

\[ E \]

\[ D \]

\[ x_0, b_1, b_2, b_3, b_4, b_5 \]

\[ \hat{x}_1 \]

\[ u_1, z_2 \]

\[ R_{ex} = 2 \]
Strategy for deterministic abstraction
Strategy for deterministic abstraction
Strategy for deterministic abstraction

\[ \mathcal{D} \]

\[ R_{ex} = 2 \]
Strategy for deterministic abstraction
Strategy for deterministic abstraction

\[ x_0 b_1 b_2 b_3 b_4 b_5 \]

\[ b_1 b_2 b_3 b_4 0 \]

\[ R_{ex} = 2 \]

\[ \hat{x}_1 \]
Strategy for deterministic abstraction

\[ x_0 \rightarrow b_1 \rightarrow b_2 \rightarrow b_3 \rightarrow b_4 \rightarrow b_5 \rightarrow x_1 \]

\[ R_{ex} = 2 \]
Strategy for deterministic abstraction

Explicit channel: fine information
Strategy for deterministic abstraction

Explicit channel: fine information
Strategy for deterministic abstraction

Explicit channel: fine information
Strategy for deterministic abstraction

Explicit channel: fine information

Implicit channel: high order bits
Strategy for deterministic abstraction

Explicit channel: fine information

Implicit channel: high order bits
Strategy for deterministic abstraction

Explicit channel: fine information

Implicit channel: high order bits

$R_{ex} = 2$
Infinitely better than the “natural” strategies!

Average cost

Nonlinear implicit, $SNR_{ex} = P_{ex} = \sigma_0^2$

linear explicit [Martins '06]

ratio diverges to infinity!

Our synergistic strategy

$\sigma_0^k = 10^{-2}$
Is the counterexample relevant?

Induced norm: linear strategies optimal! [Rotkowitz ’06]

Witsenhausen’s LQG formulation: more assumptions harder!

Q. Why consider the quadratic norm?

[John Doyle, “Paths ahead”, MIT, 2009]
Induced norm formulation too conservative?

Makes no assumptions: noise completely arbitrary, adversarial!

Usually we know something about the noise, e.g. “typical” values, or bounds
say $z \in (-a, a)$

A quadratic cost
$$\min_{x_0, z} \max_{u_1, u_2} k^2 u_1^2 + x_2^2$$

Quantization (signaling): can outperform linear by an unbounded factor
is approximately optimal!
Which is a “right” norm?

\[
\min \max_{x_0, z} \frac{k^2 u_1^2 + x_2^2}{x_0^2 + z^2}
\]

Induced norm: cost normalized by size of state and noise.

Induced norm: cost normalized by size of state and noise.

In many practical cases, small perturbations and small noises do cost less!

Quadratic norm captures how much large perturbations/noises hurt

Moral: signal even in adversarial situations