Communication, Computation, and Control: Through the Information Lens

Pulkit Grover
Stanford University

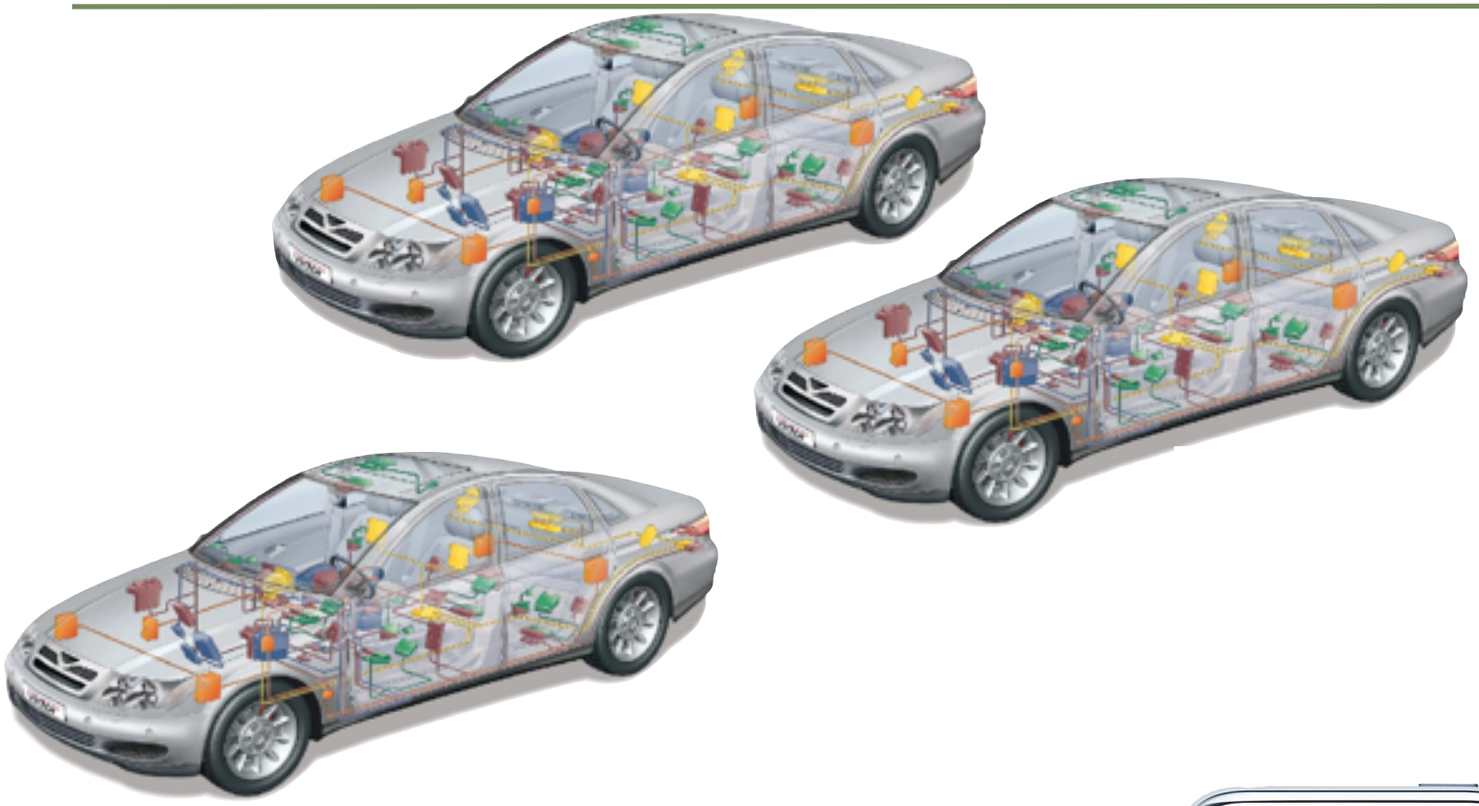
CMU, April 16, 2012

Acknowledgments: C2S2/Interconnect Focus Center (IFC)
NSF Center for Science of Information (CSoI)
NSF-CCF-0917212 (cyber-physical systems)

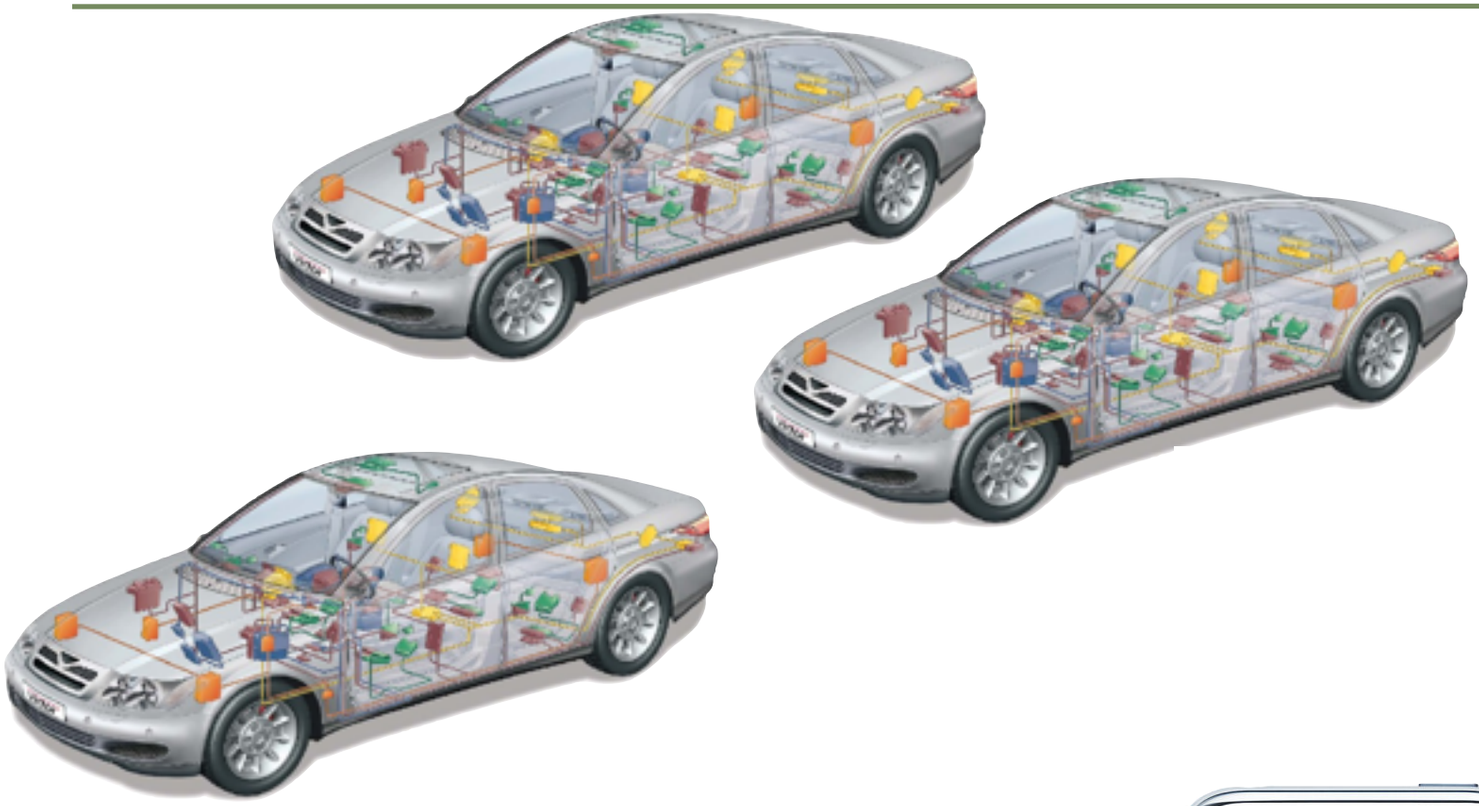
Flows of information in modern systems

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

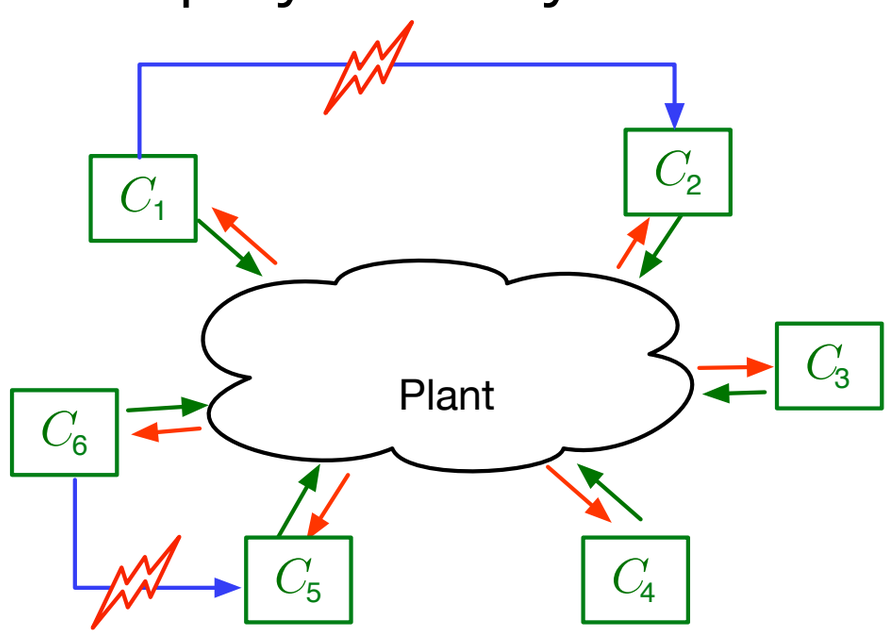
Talk outline:
Understanding Invisible Information Flows

1) Information flows in circuits

Talk outline:
Understanding Invisible Information Flows

1) Information flows in circuits

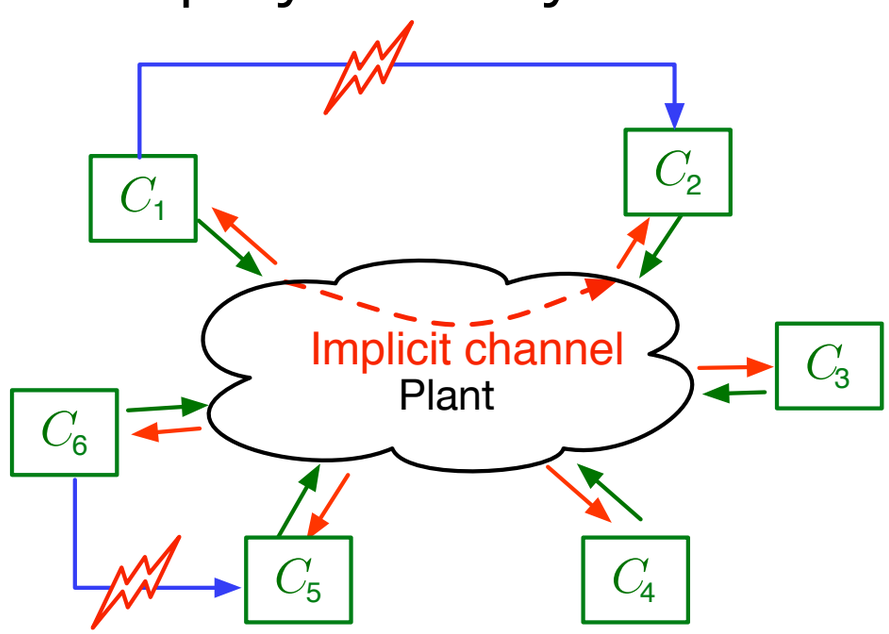
2) Information flows in physical systems



Talk outline:
Understanding Invisible Information Flows

1) Information flows in circuits

2) Information flows in physical systems



Talk outline:
Understanding Invisible Information Flows

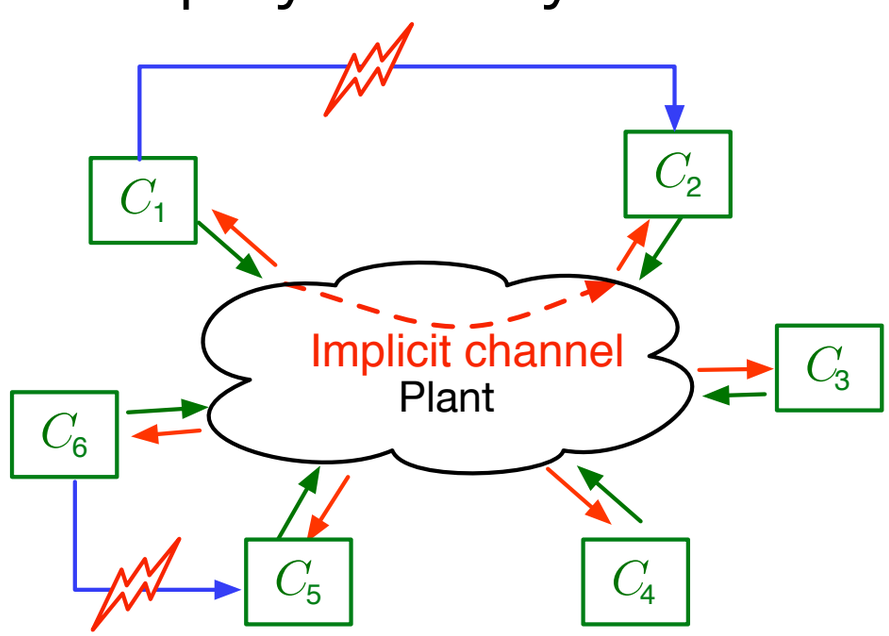
1) Information flows in circuits

Traditional: Minimize P_T

Now: Minimize total power

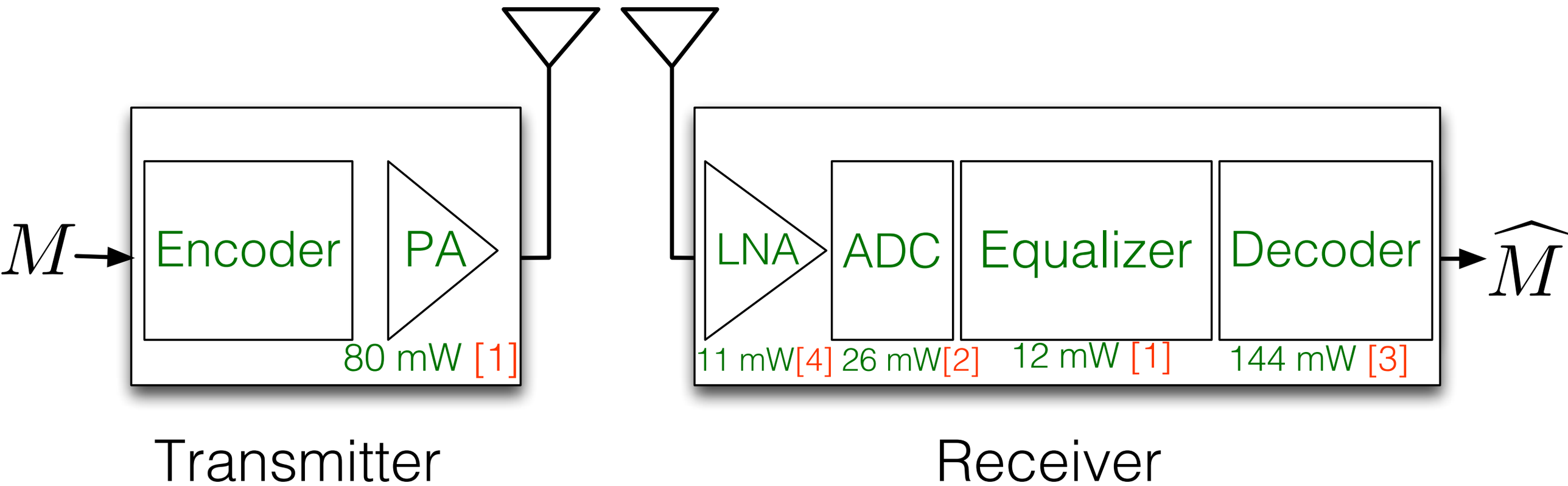
Fundamental limits, experimental results

2) Information flows in physical systems



Short-distances:
Processing power can exceed transmit power!

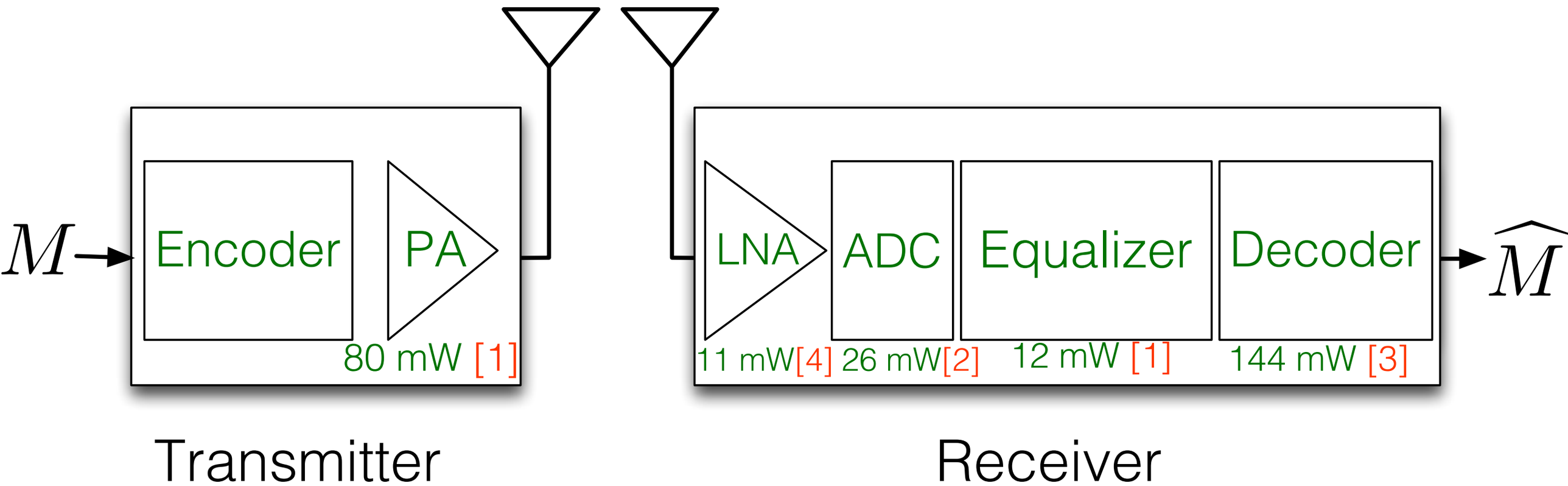
Short-distance communication at Gbps
(e.g. indoor wireless, data-centers ethernet)



[1] Uncoded [Marcu et al '09] [2] 7-bit ADCs [Murmann '08] [3] (6,32)-LDPC [Zhang et al '10]
[4] [Shin '09]

Short-distances:
Processing power can exceed transmit power!

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

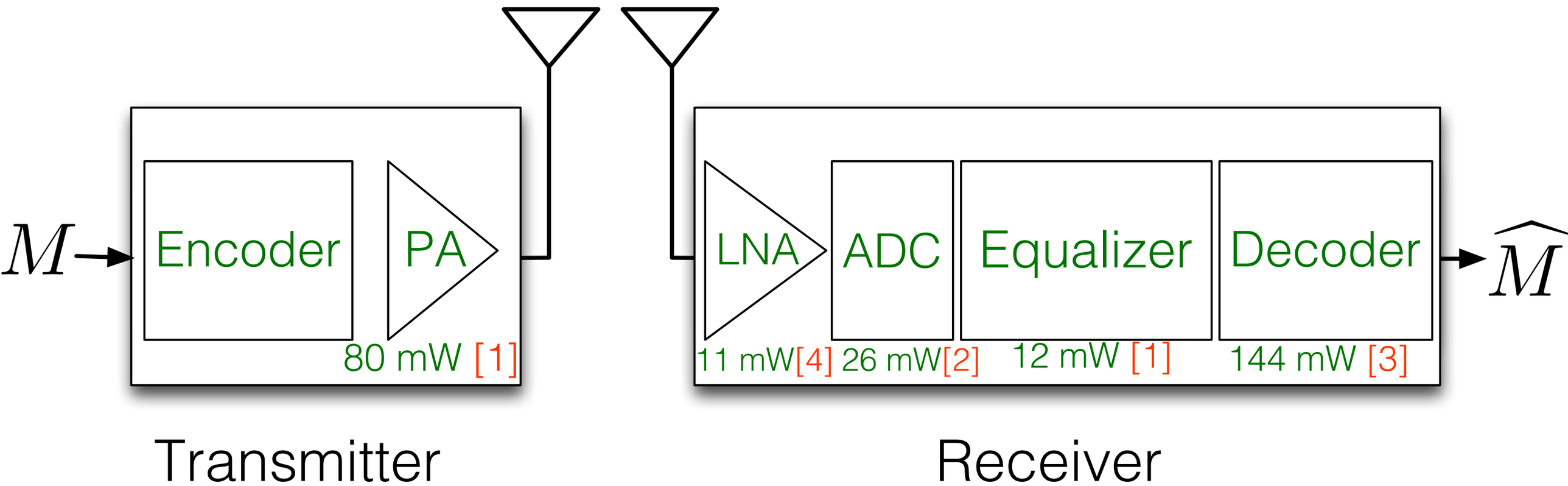


Choosing a P_{total}-efficient code?

[1] Uncoded [Marcu et al '09] [2] 7-bit ADCs [Murmann '08] [3] (6,32)-LDPC [Zhang et al '10]
[4] [Shin '09]

Short-distances:
Processing power can exceed transmit power!

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

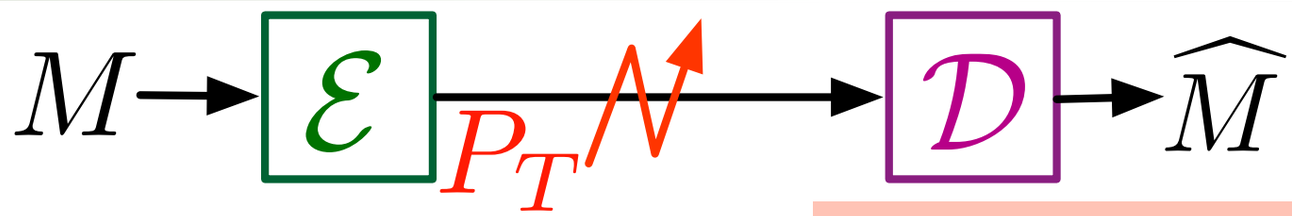


Choosing a P_{total}-efficient code?

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

[1] Uncoded [Marcu et al '09] [2] 7-bit ADCs [Murmann '08] [3] (6,32)-LDPC [Zhang et al '10]
[4] [Shin '09]

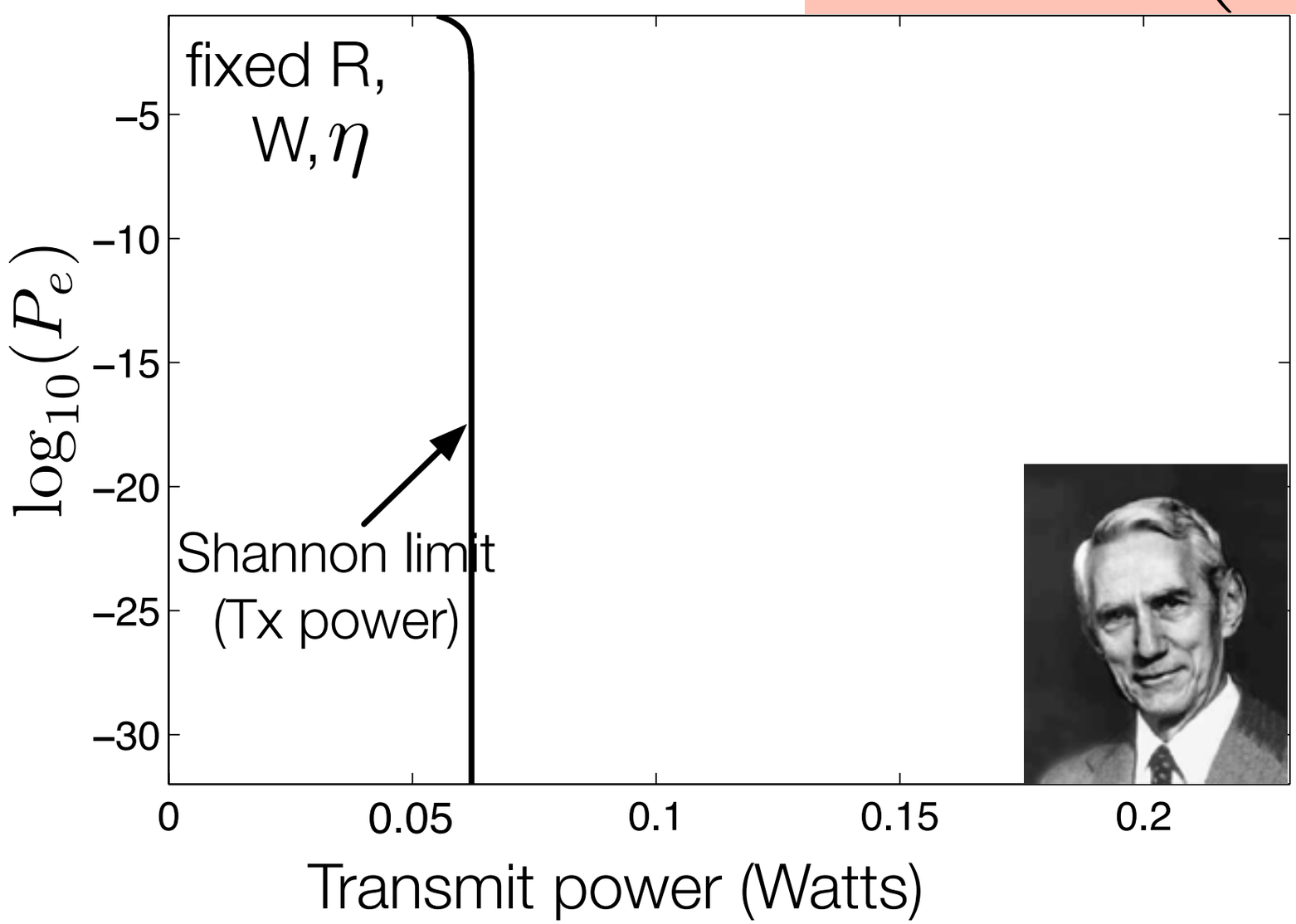
Minimizing transmit power: Shannon's waterfall curves



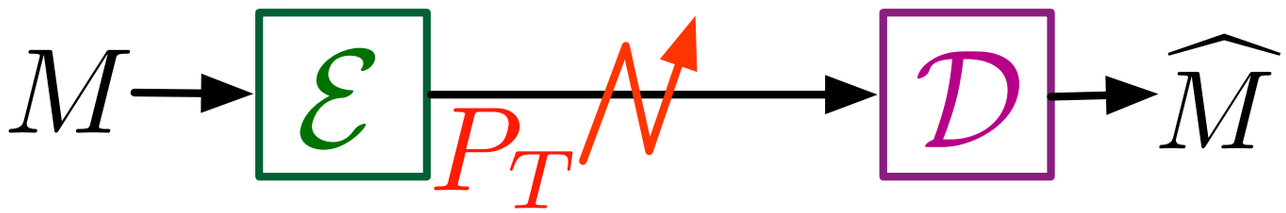
path loss

Goal: Rate R, error probability P_e

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

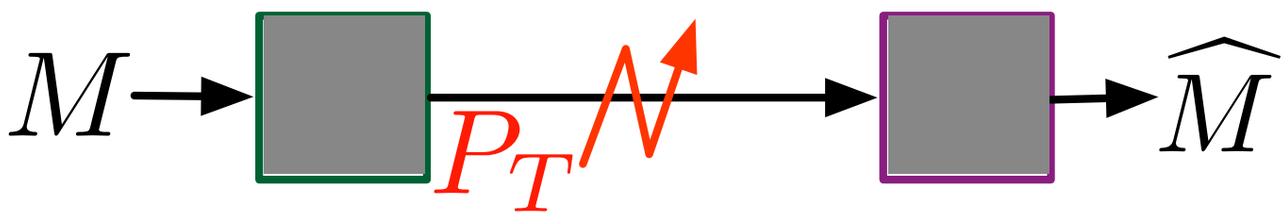


Minimizing total power: Initial theoretical approaches



Minimizing total power: Initial theoretical approaches

"Black-box" model of processing power

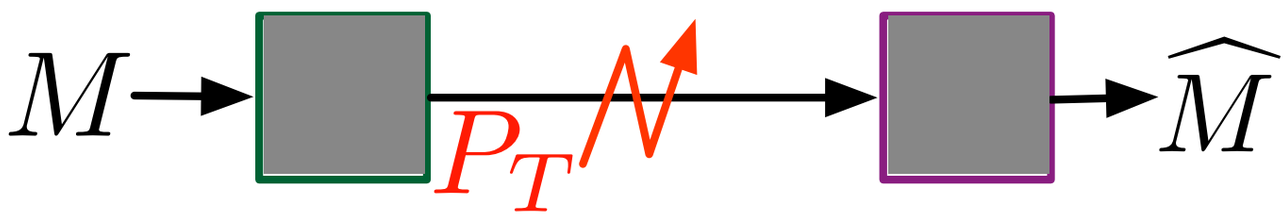


[Cui, Goldsmith, Bahai '05]

[Massaad, Medard, Zheng '08]

Minimizing total power: Initial theoretical approaches

"Black-box" model of processing power



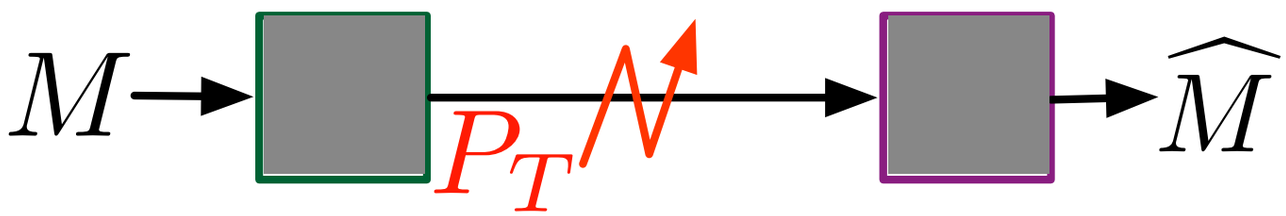
[Cui, Goldsmith, Bahai '05]

[Massaad, Medard, Zheng '08]

Optimal "burstiness"

Minimizing total power: Initial theoretical approaches

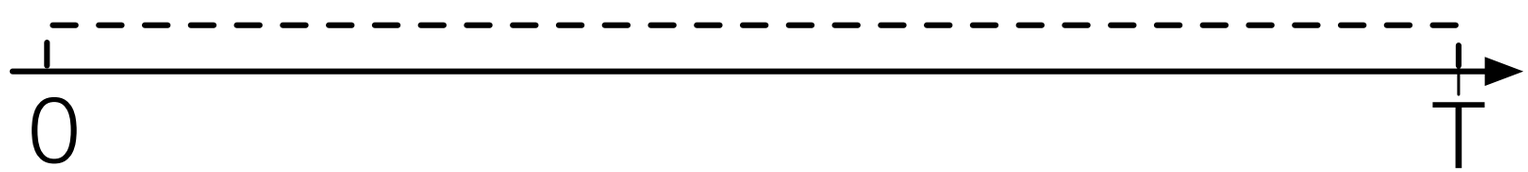
"Black-box" model of processing power



[Cui, Goldsmith, Bahai '05]

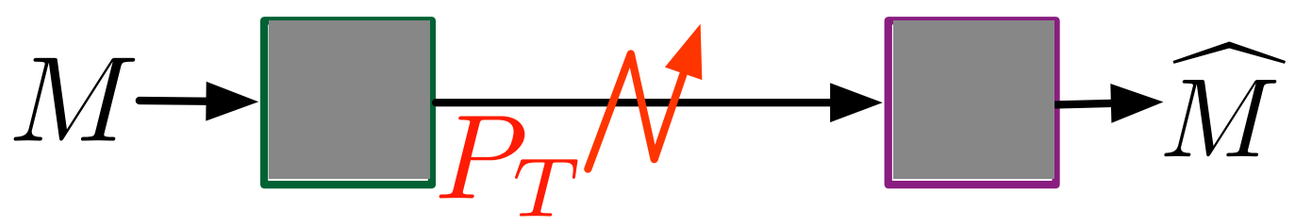
[Massaad, Medard, Zheng '08]

Optimal "burstiness"



Minimizing total power: Initial theoretical approaches

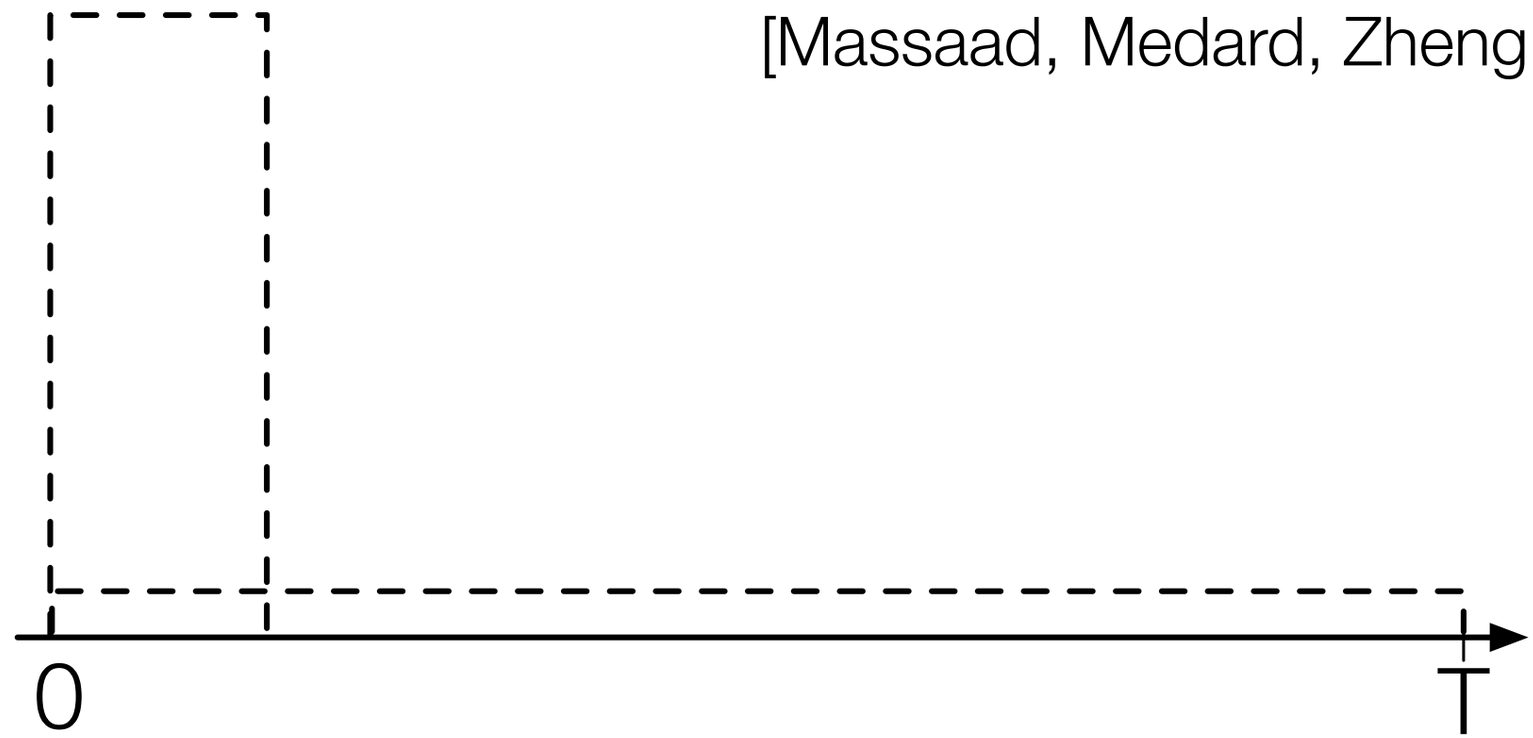
"Black-box" model of processing power



[Cui, Goldsmith, Bahai '05]

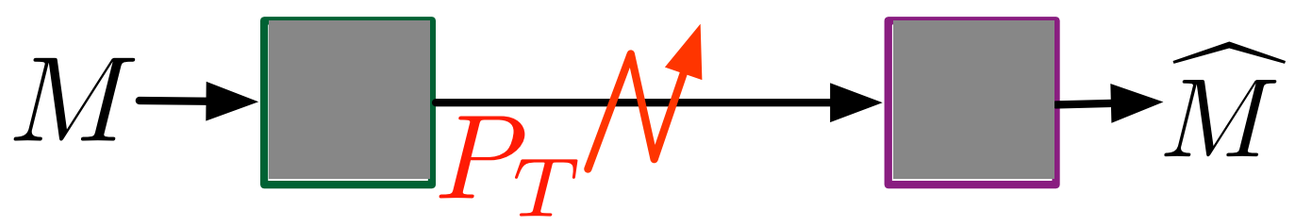
[Massaad, Medard, Zheng '08]

Optimal "burstiness"



Minimizing total power: Initial theoretical approaches

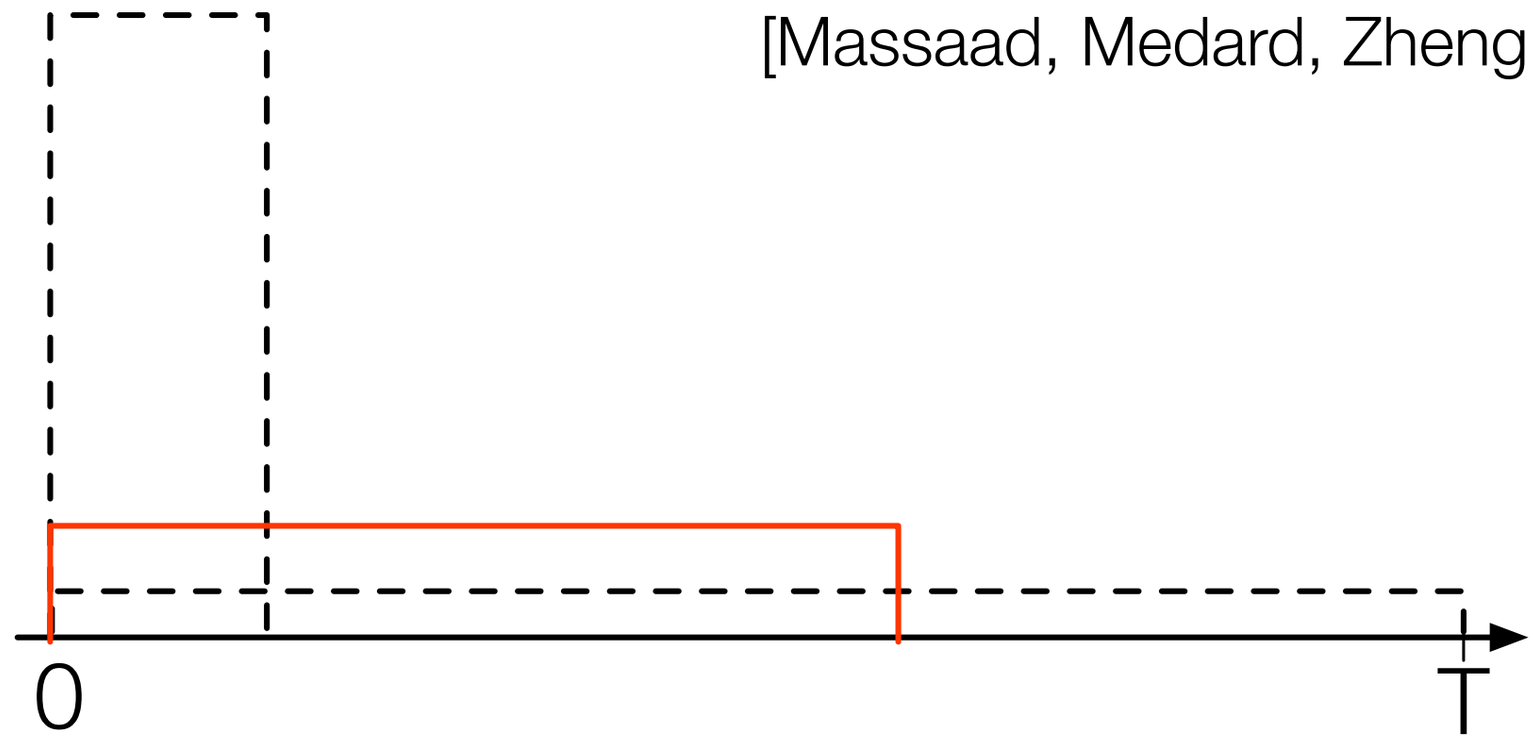
"Black-box" model of processing power



[Cui, Goldsmith, Bahai '05]

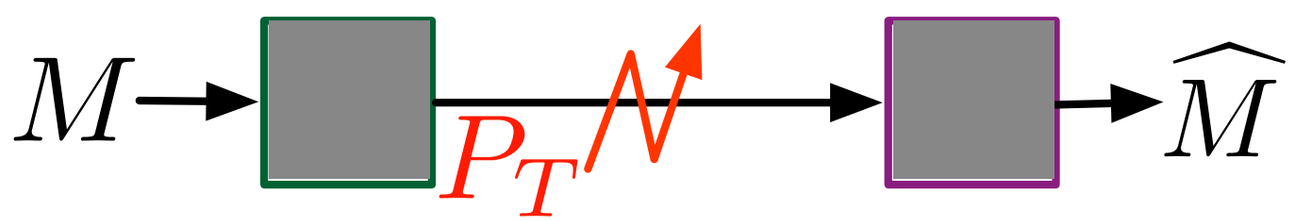
[Massaad, Medard, Zheng '08]

Optimal "burstiness"



Minimizing total power: Initial theoretical approaches

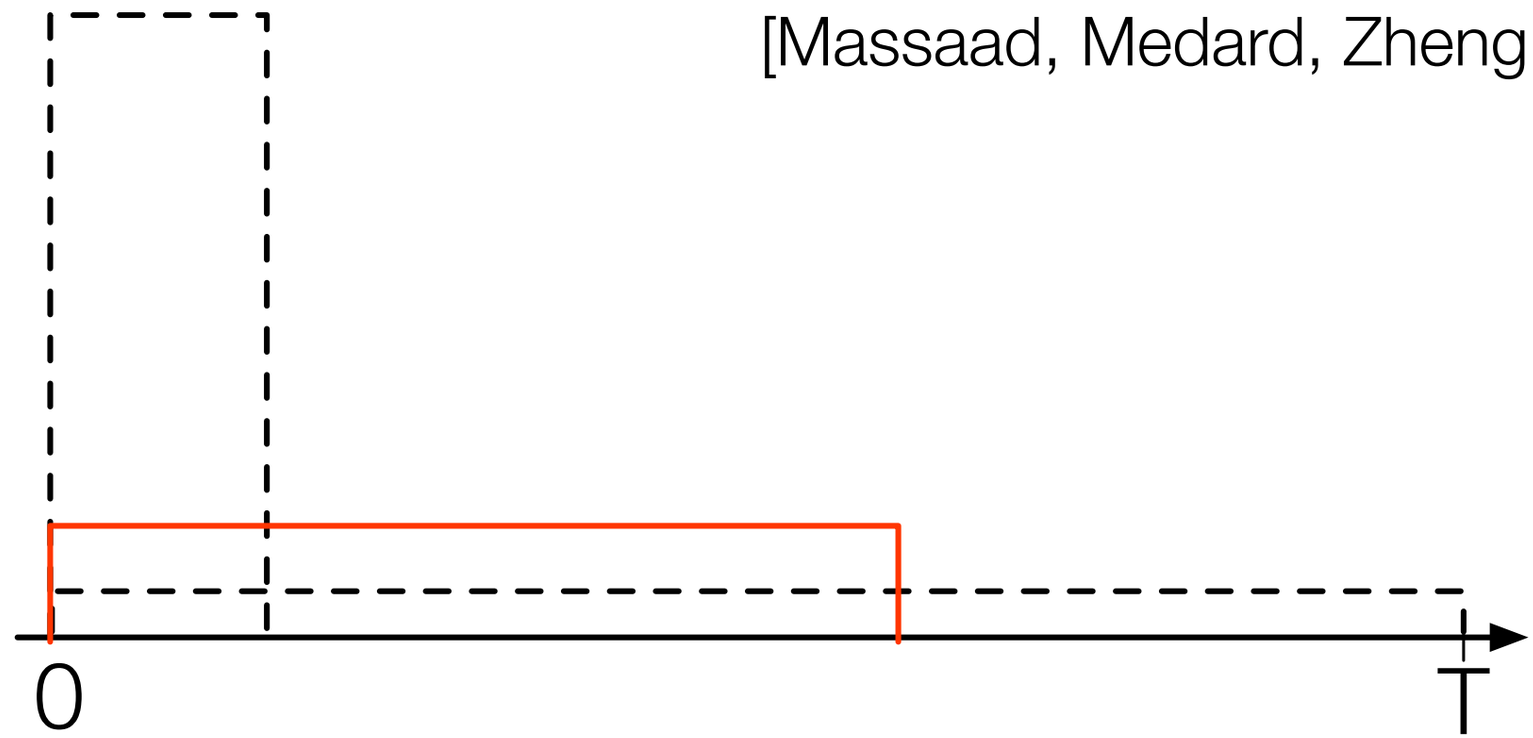
"Black-box" model of processing power



[Cui, Goldsmith, Bahai '05]

[Massaad, Medard, Zheng '08]

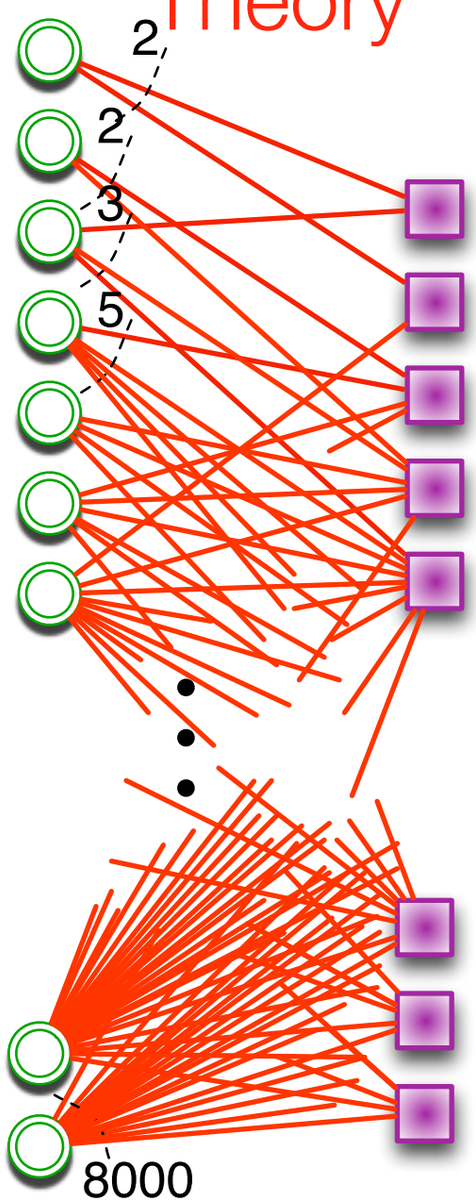
Optimal "burstiness"



What are the "greenest" codes?

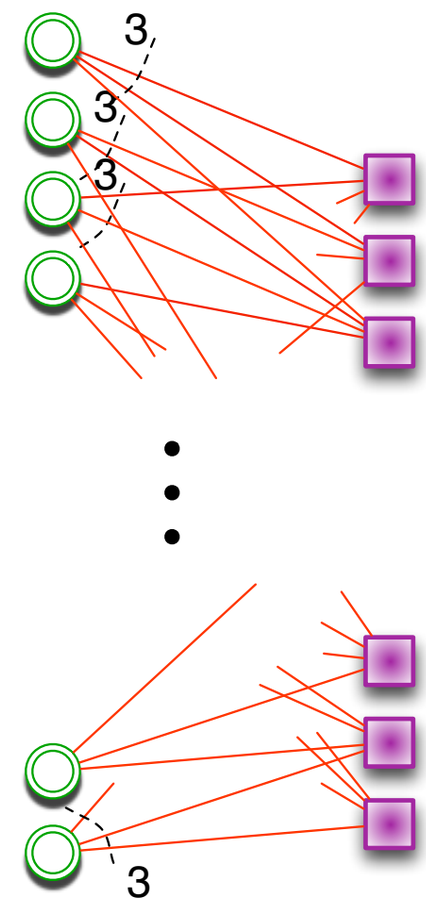
“Green” system = Capacity-code + “Green” implementation?

Theory



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

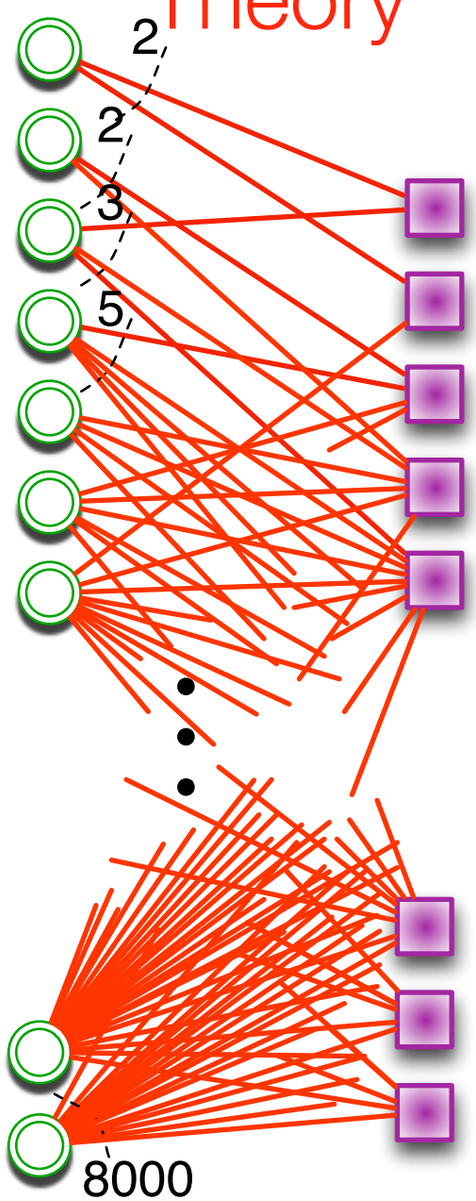
Implementation



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

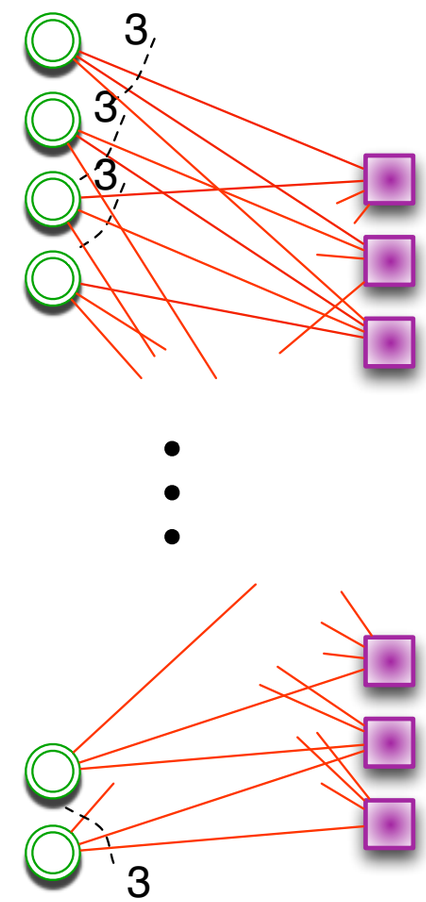
“Green” system = Capacity-code + “Green” implementation?

Theory



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

Implementation

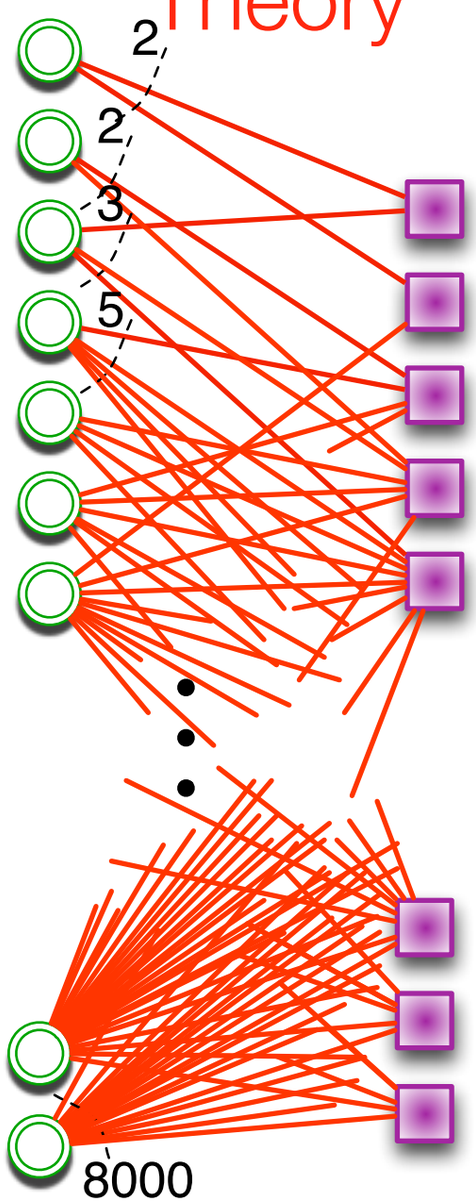


“A 47-Gbps LDPC decoder . . . ” (10 GBase-T)
[Zhang, Anantharam, Wainwright, Nikolic JSSC ’10]

Wireless LAN and 60 GHz band

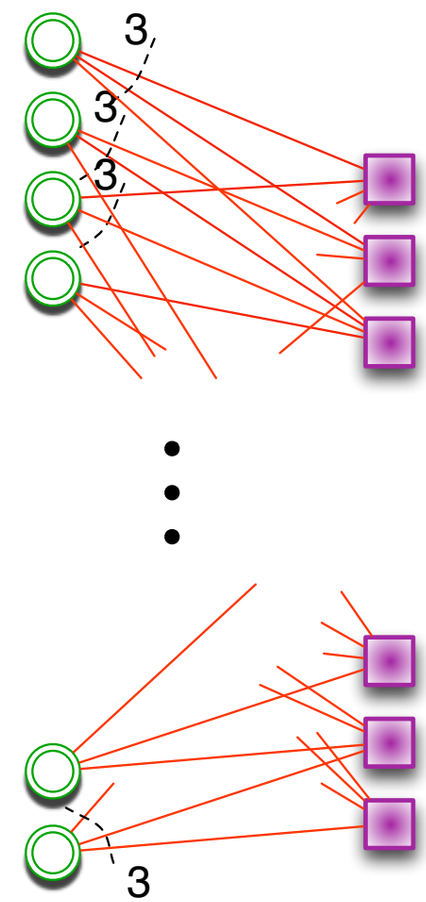
“Green” system = Capacity-code + “Green” implementation?

Theory



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

Implementation



“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

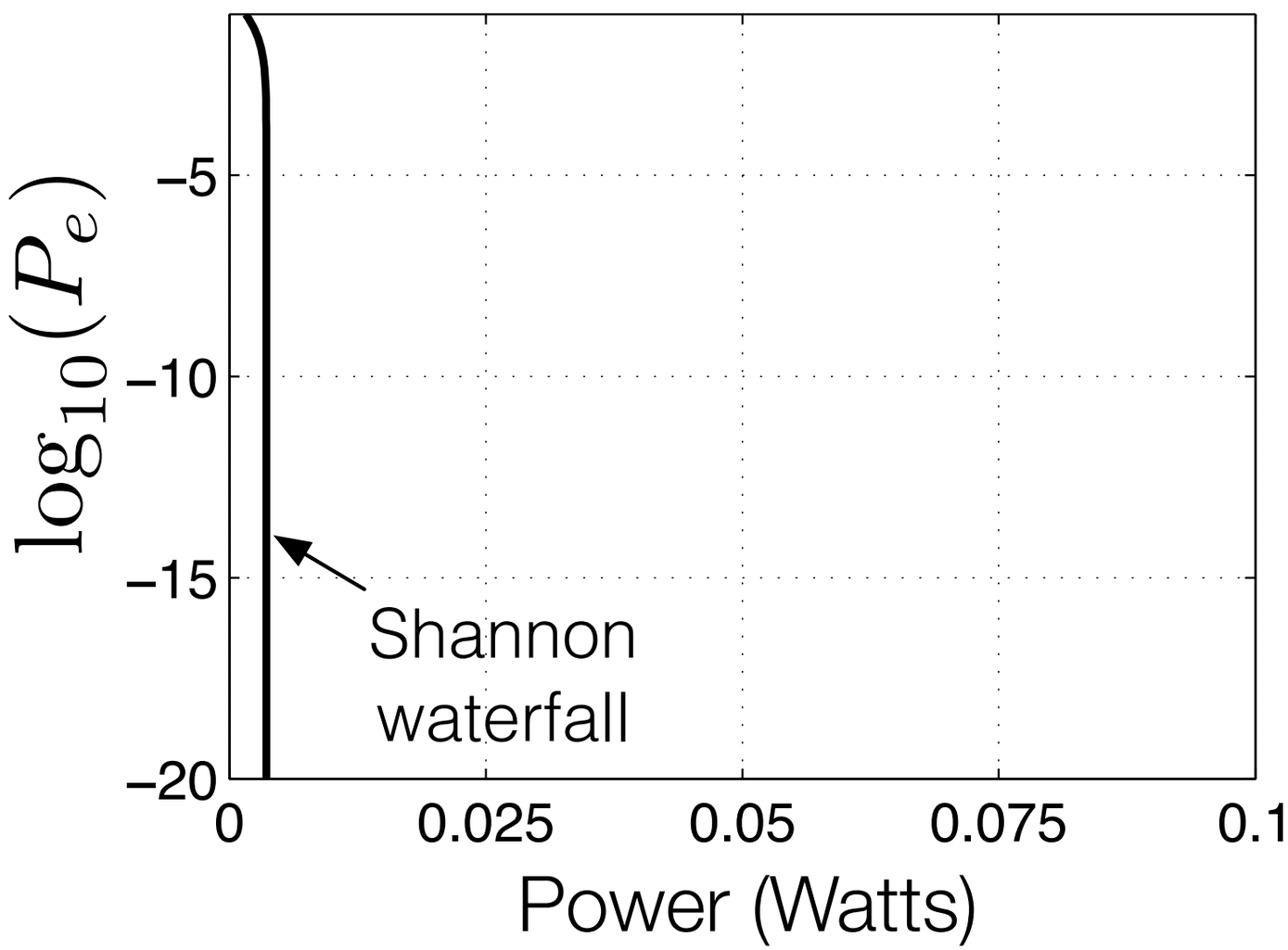
Circuit-simulation-based modeling of decoding power

[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

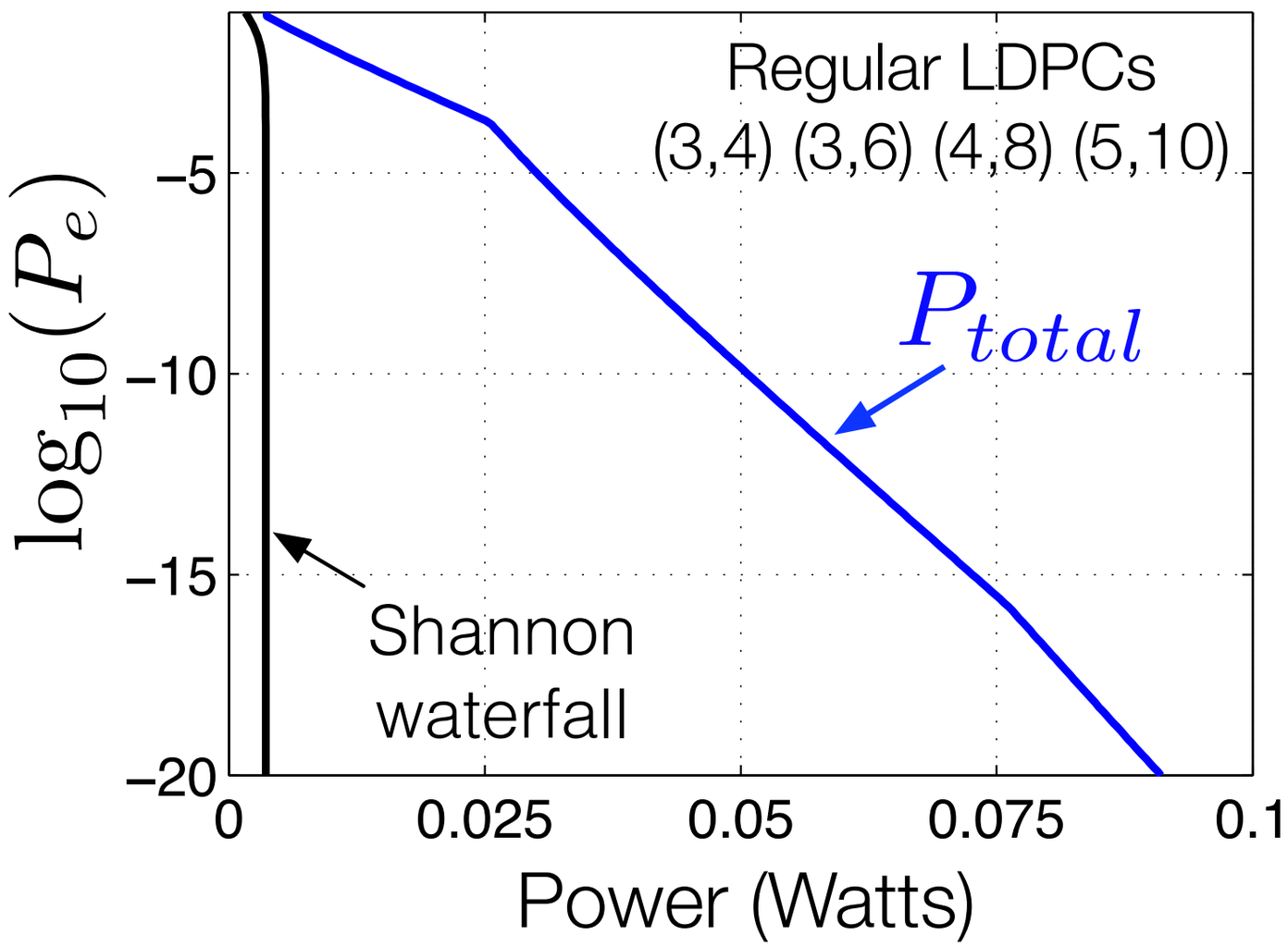
[Ganesan, Grover, Rabaey SiPS'11][Ganesan, Wen, Grover, Rabaey'12]



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]

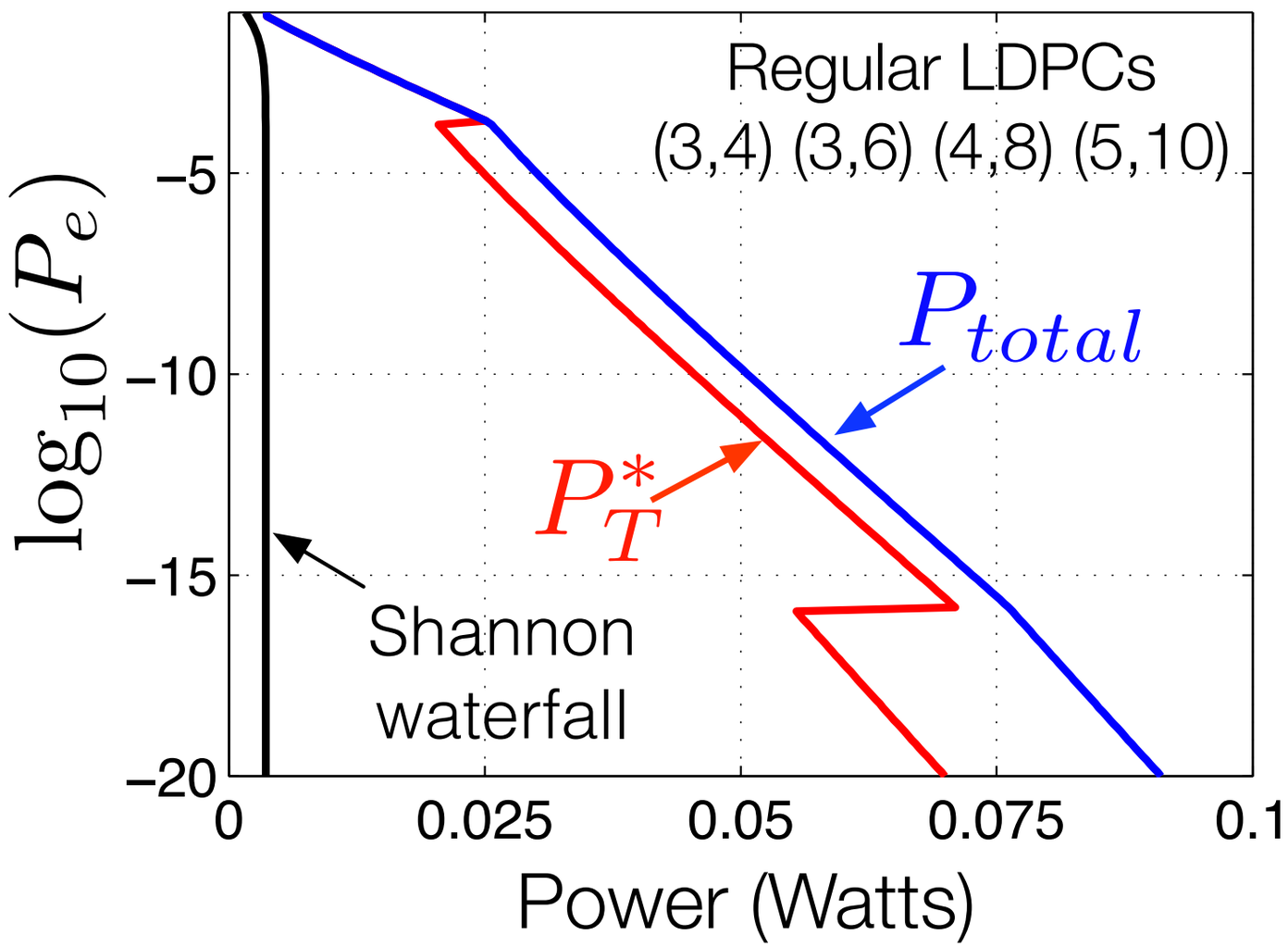


R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE

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]

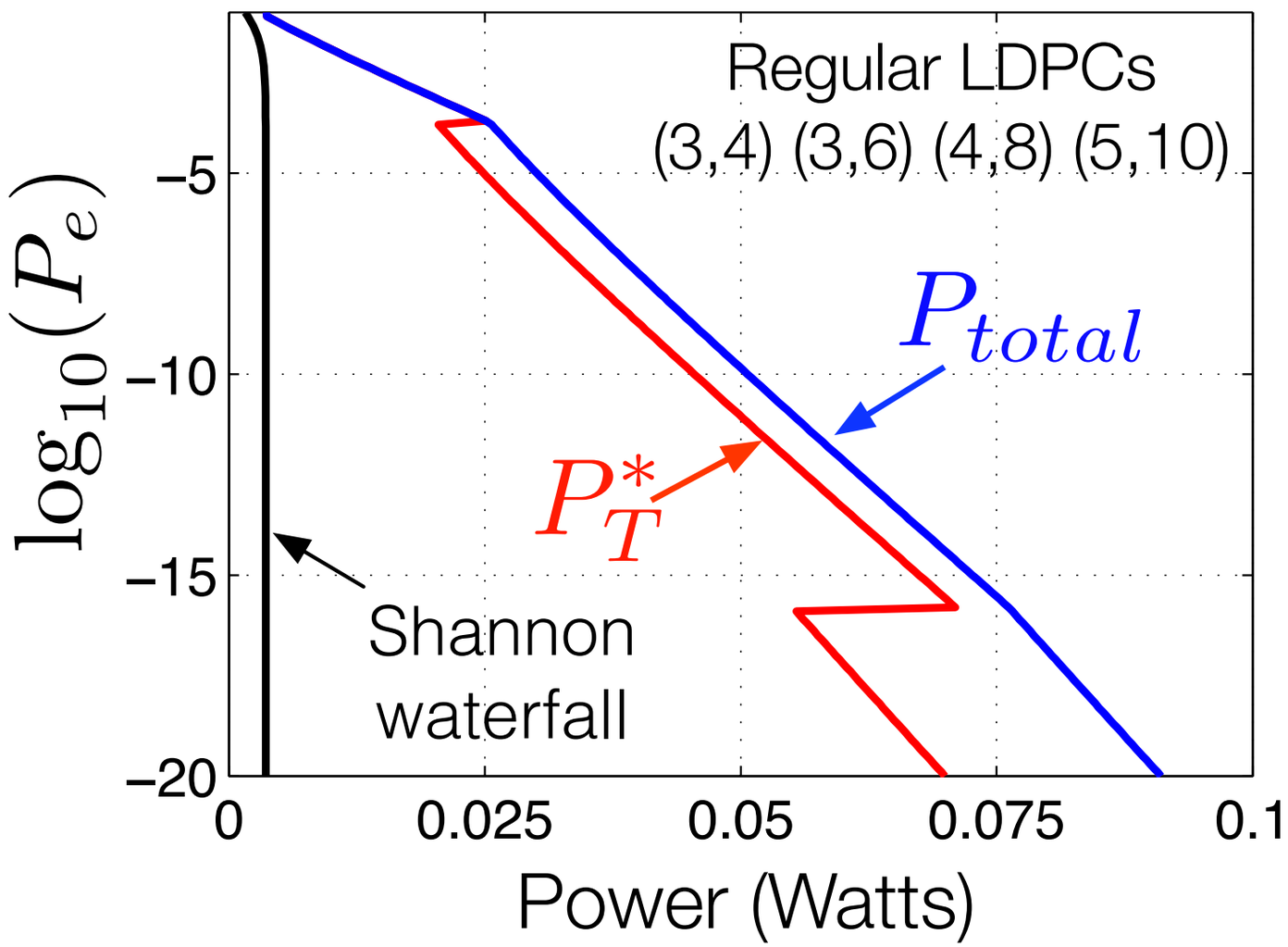


R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE

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]



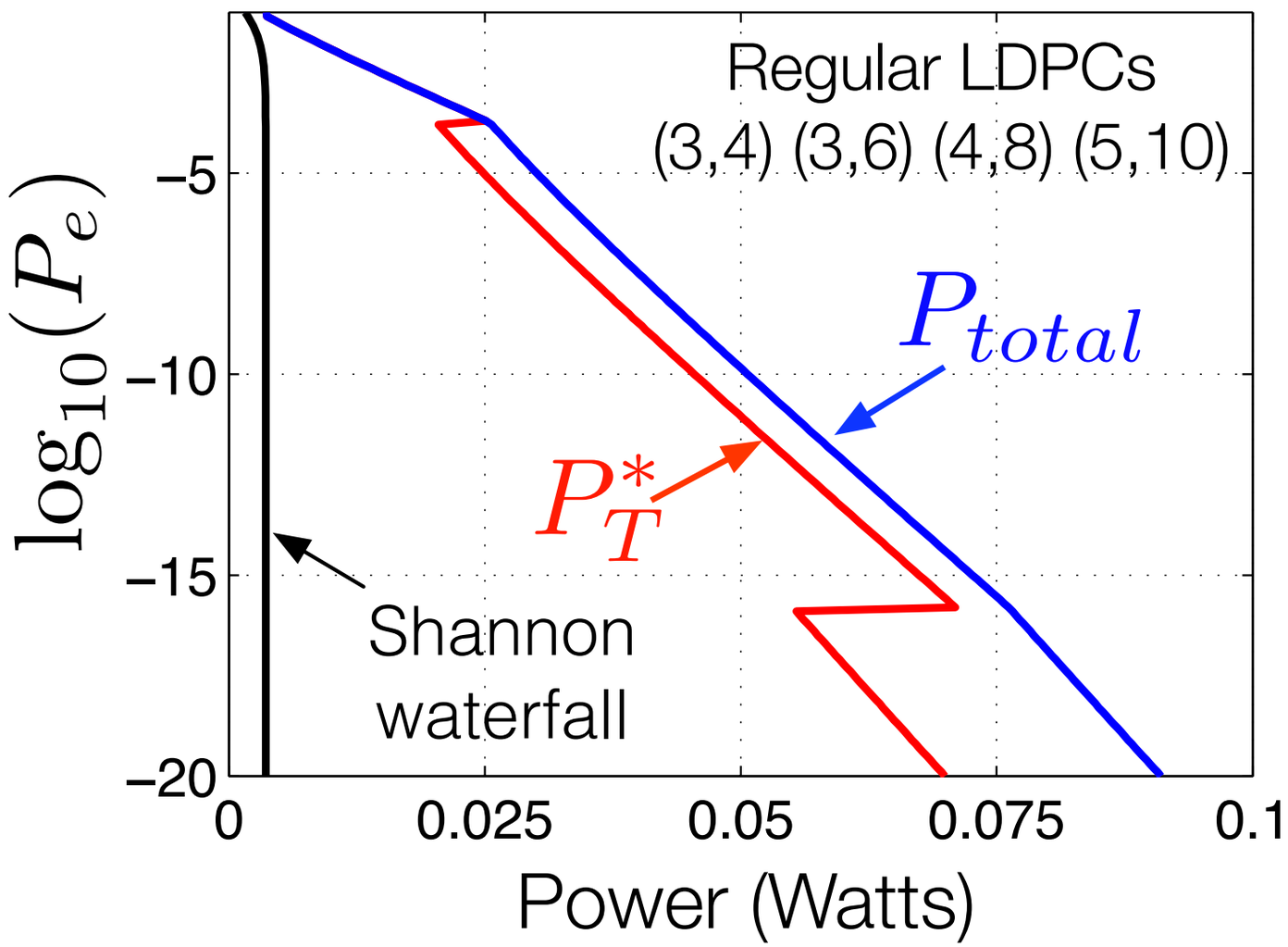
R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE

Bad code choice?

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]



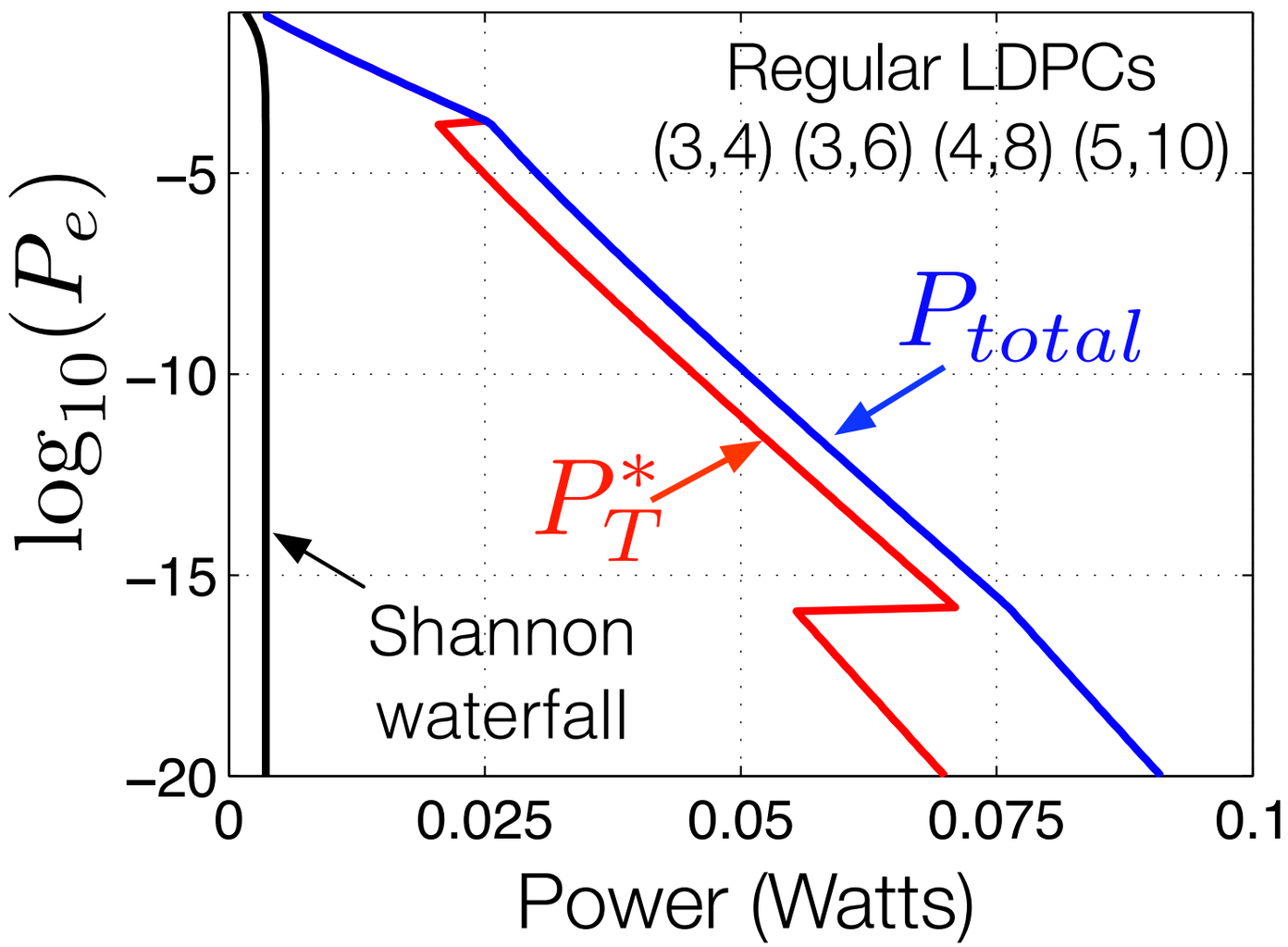
R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE

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]



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

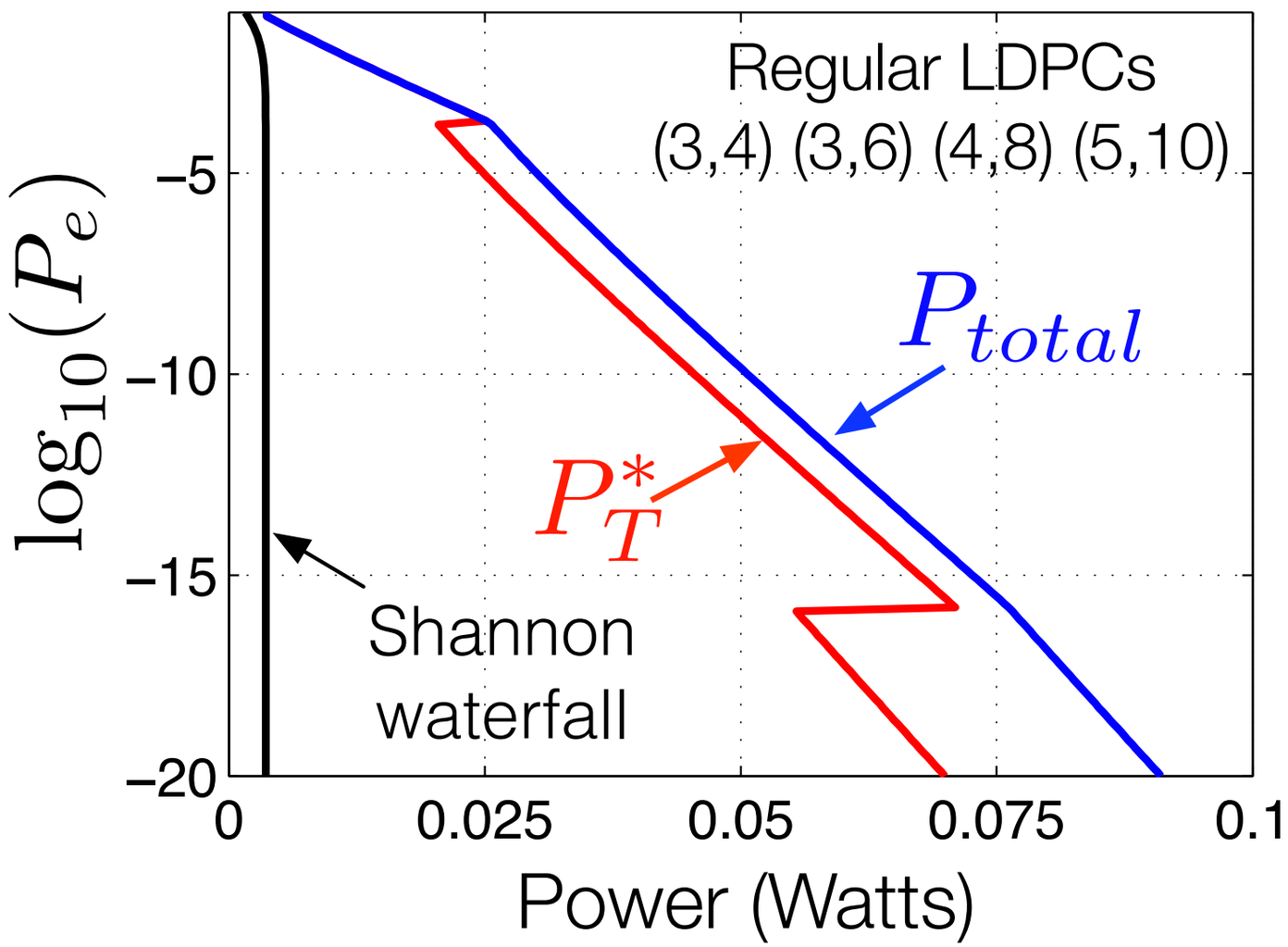
Bad code choice? Bad implementations?

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]

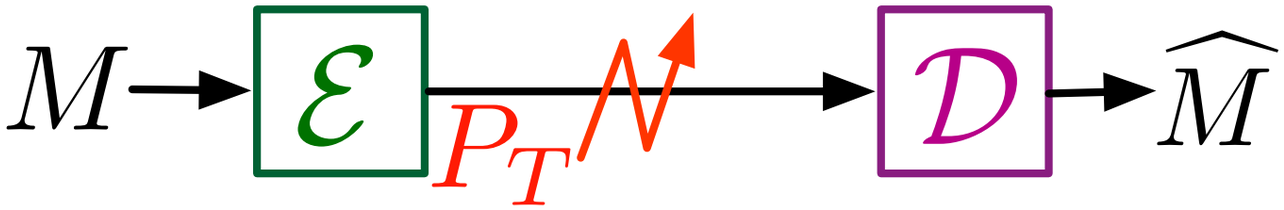


R = 7 Gbps
W = 7 GHz
path-loss coeff. = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encounter (Layout)
Synopsis HSPICE

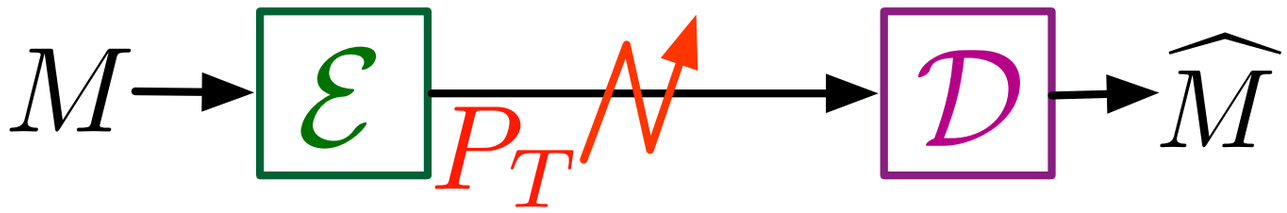
Bad code choice? Bad implementations?

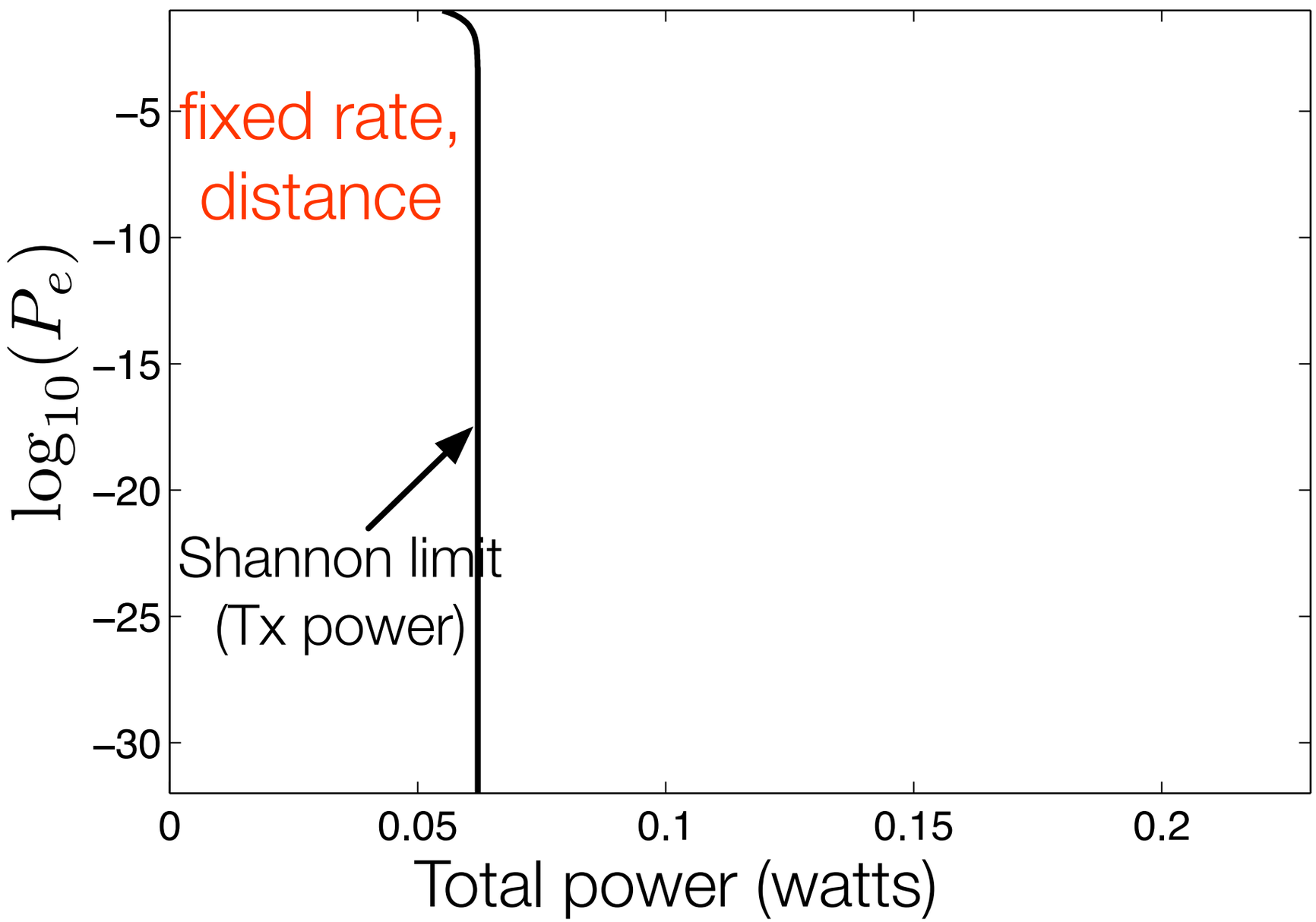
Both? Neither?

Fundamental limits on total power?

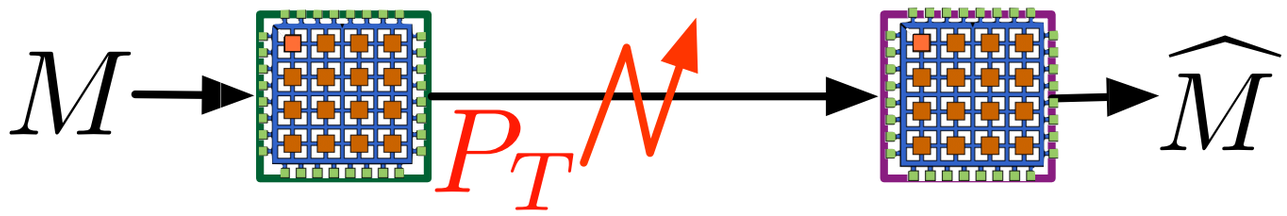


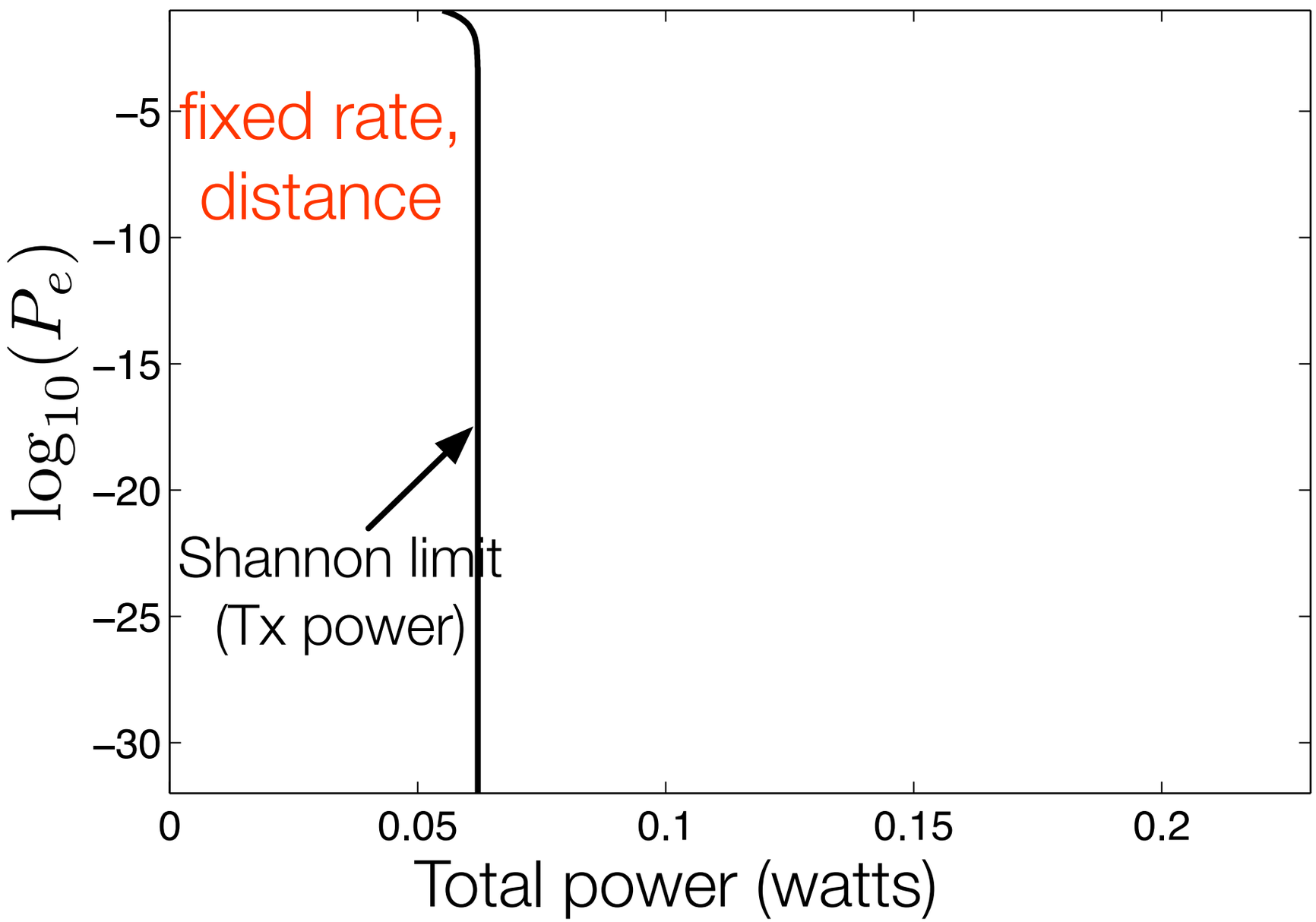
Fundamental limits on total power?



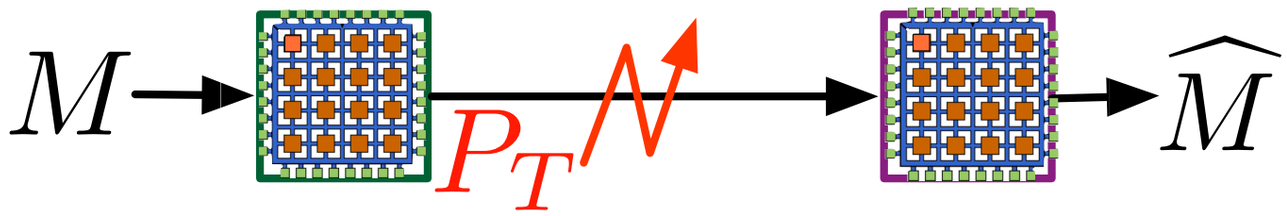


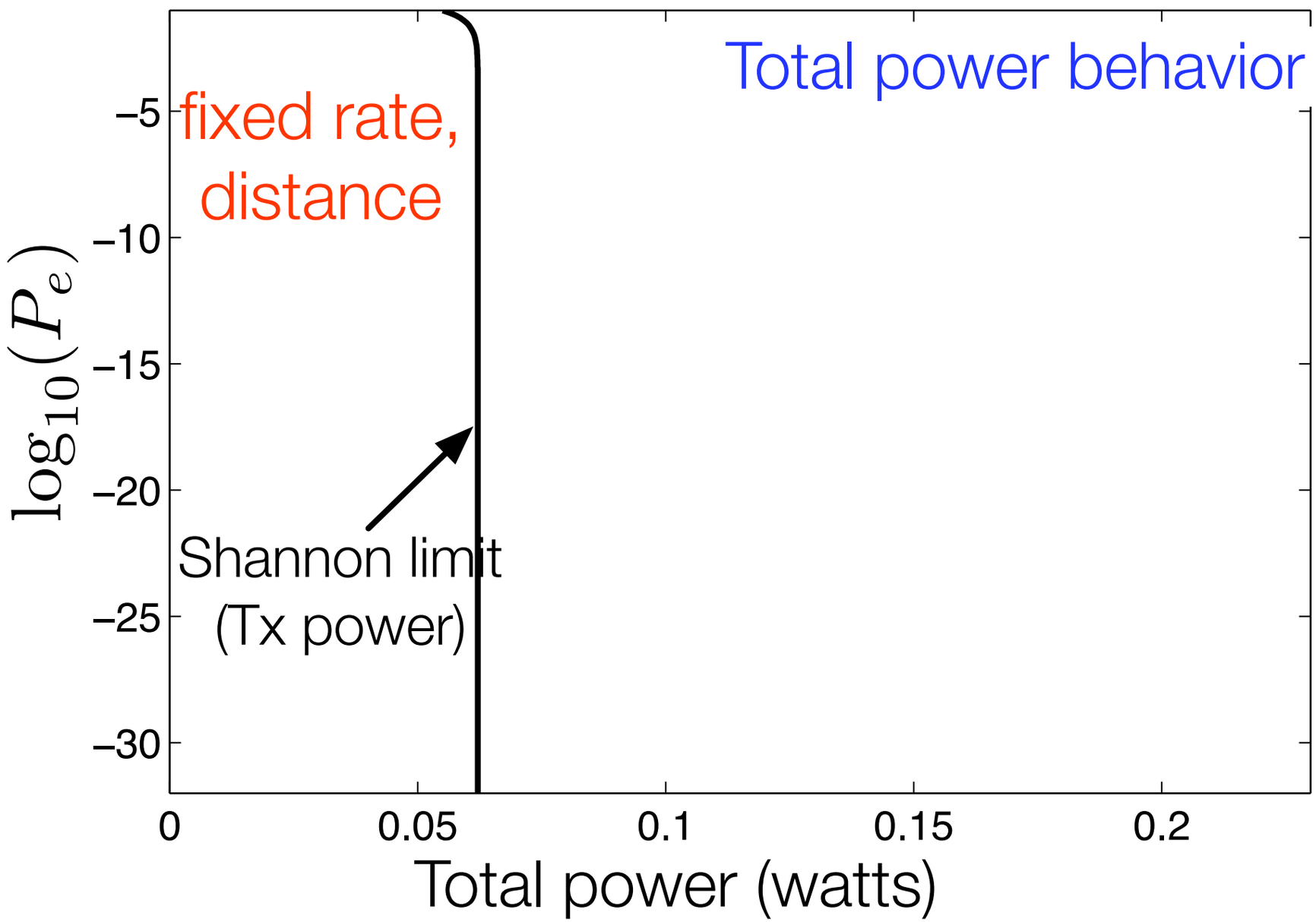
Fundamental limits on total power?



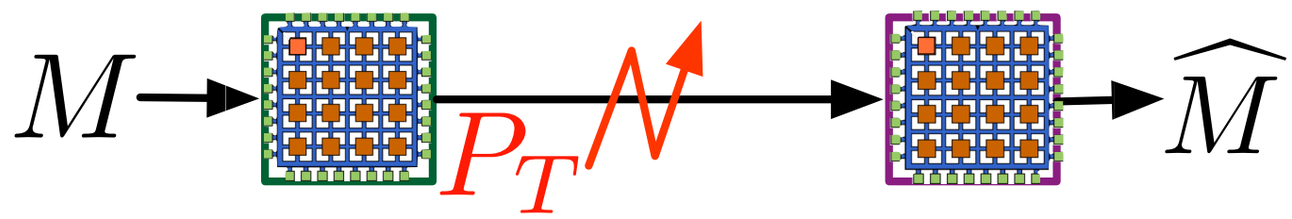


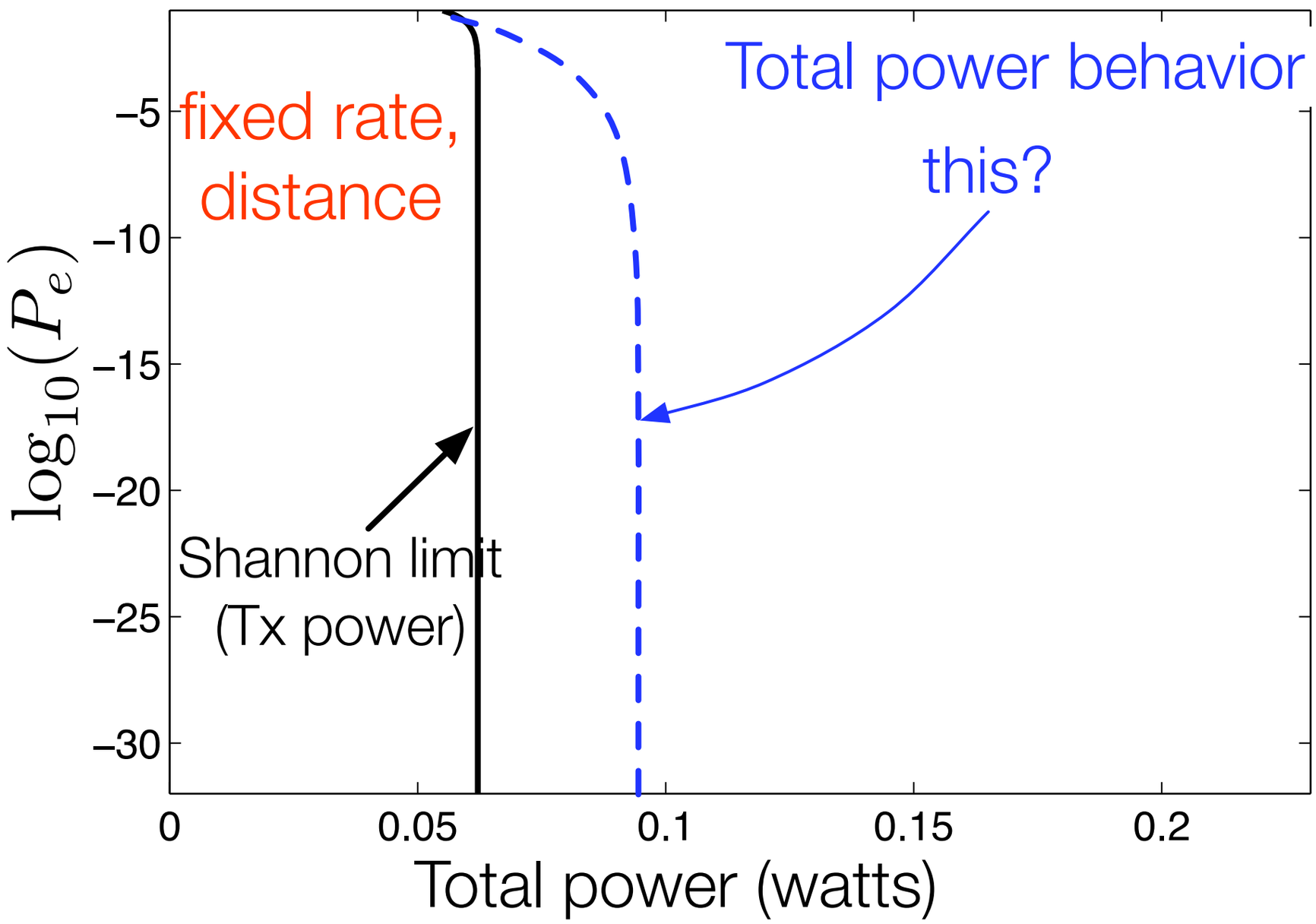
Fundamental limits on total power?



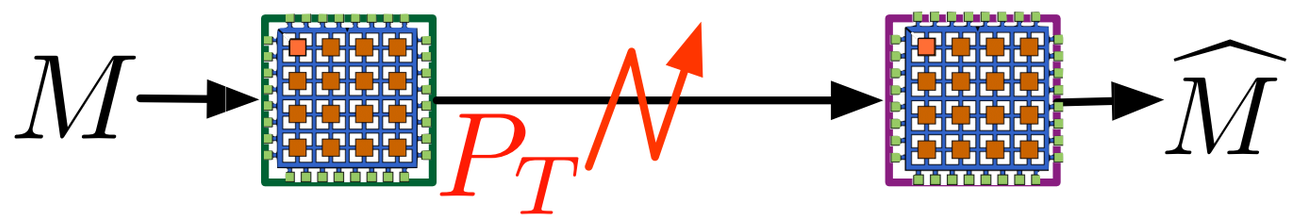


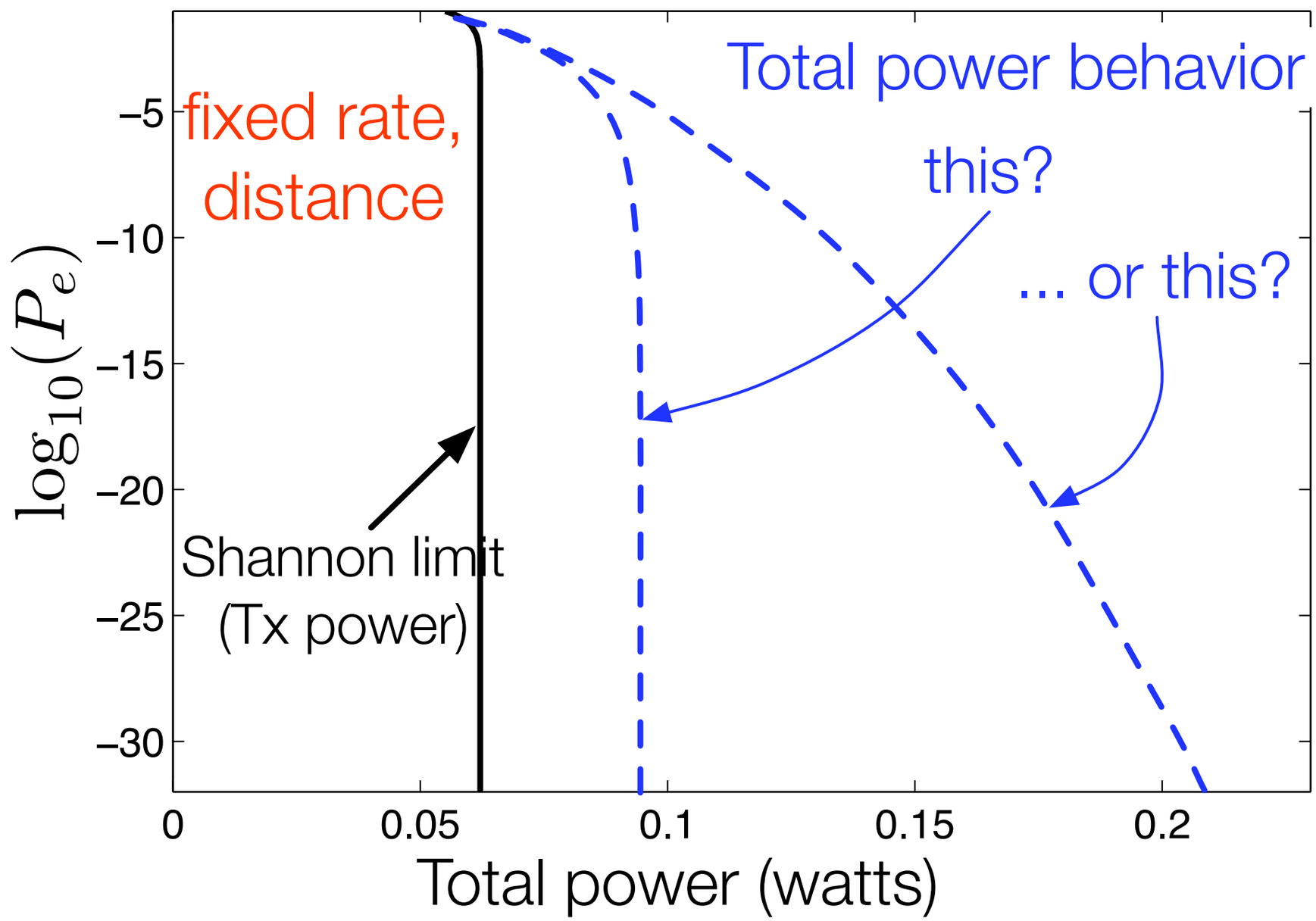
Fundamental limits on total power?



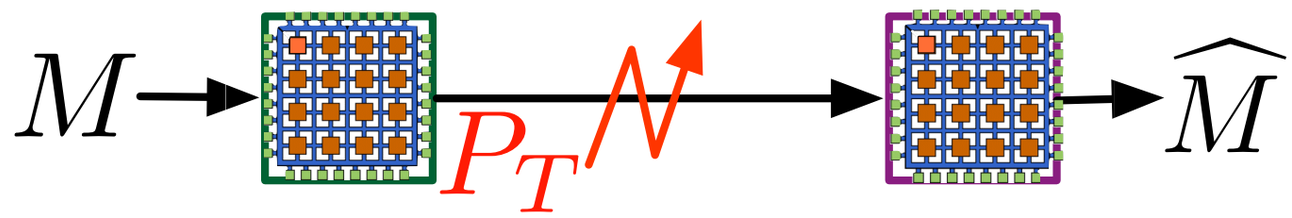


Fundamental limits on total power?





Fundamental limits on total power?



Questions for this talk:

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

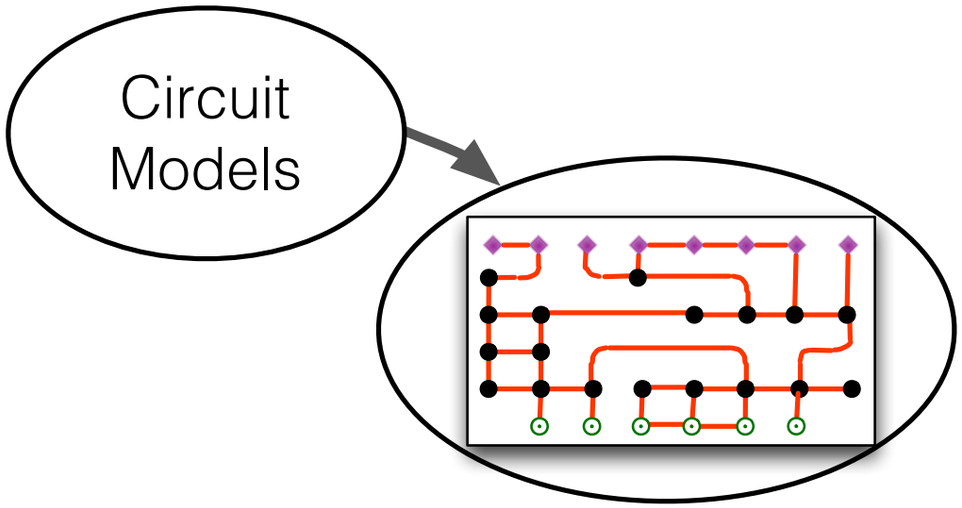
2) Should we operate close to capacity?

3) Can we use bounded P_T ?

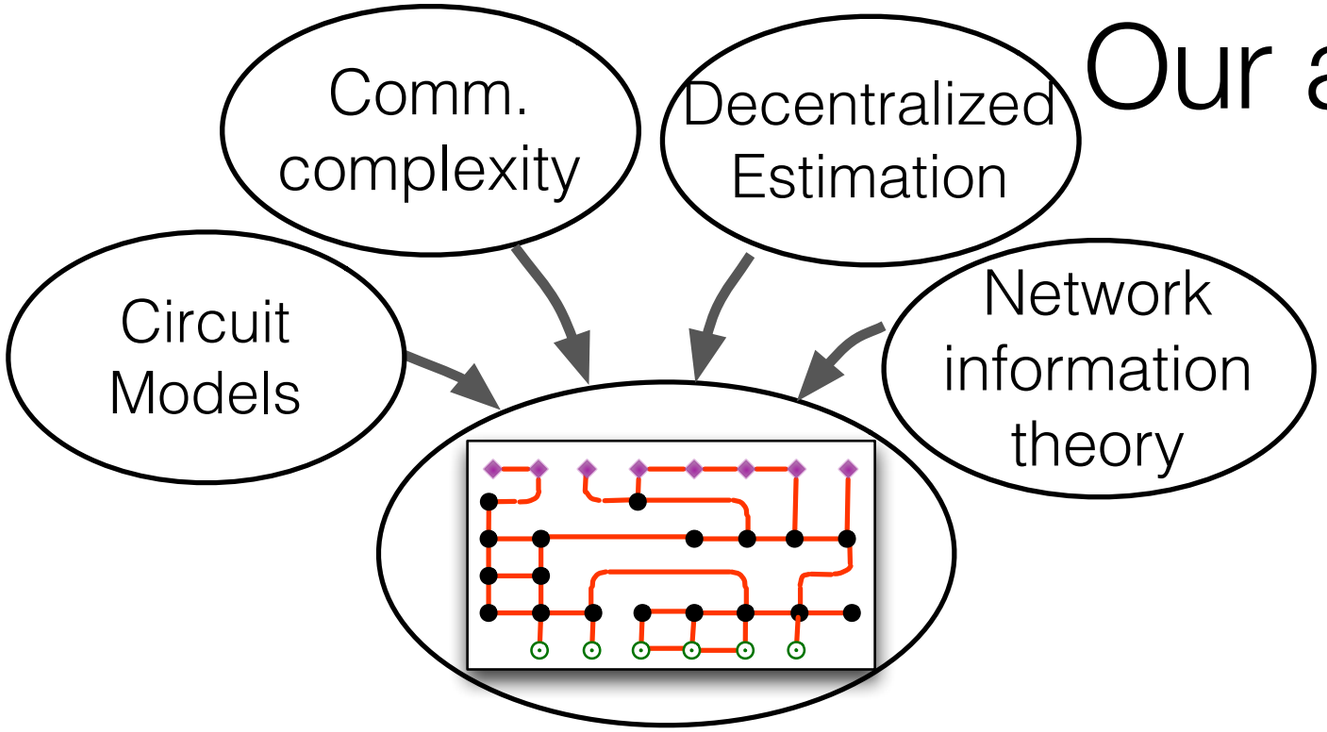
4) What codes minimize P_{total} ?

Our approach

Our approach

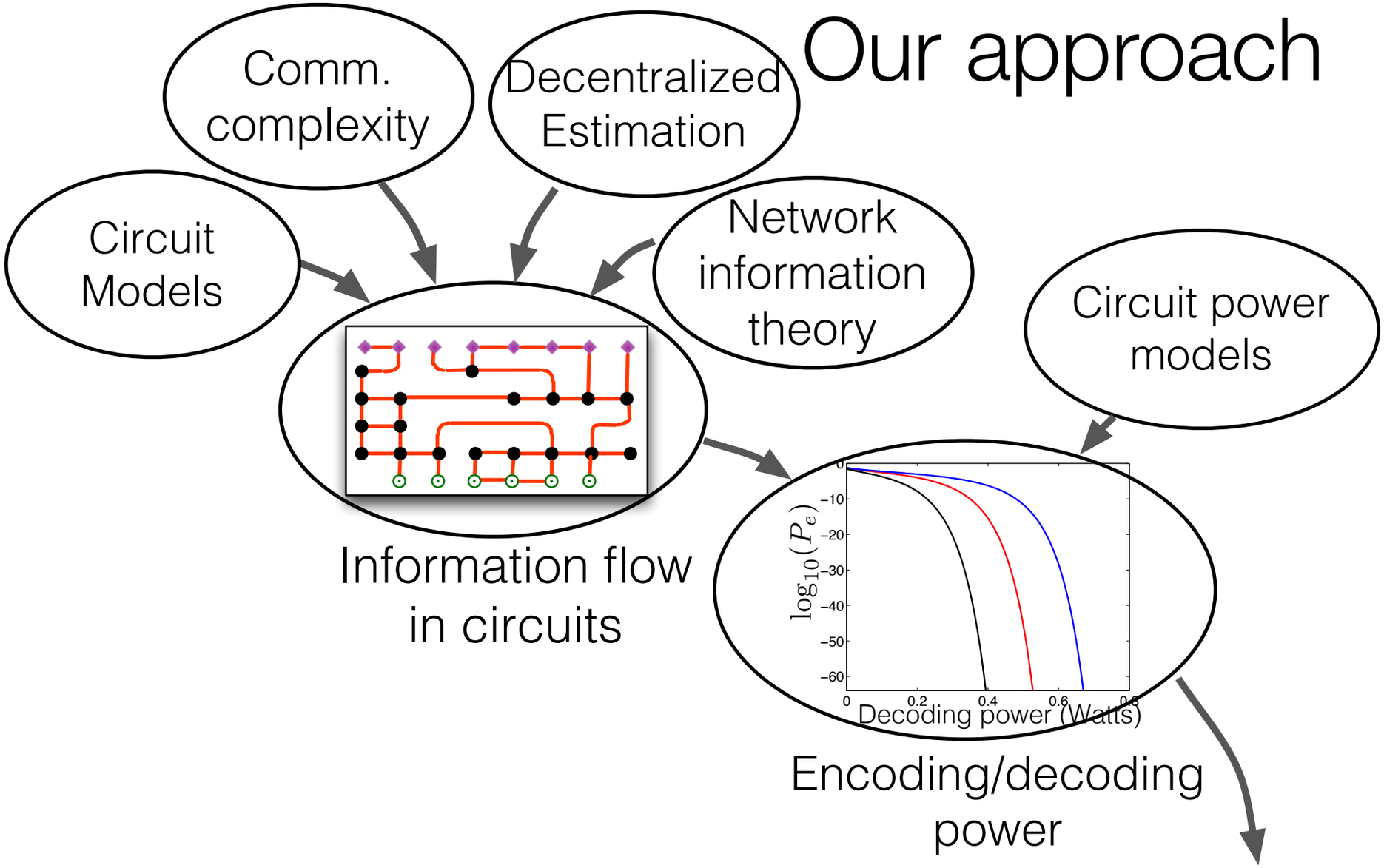


Our approach

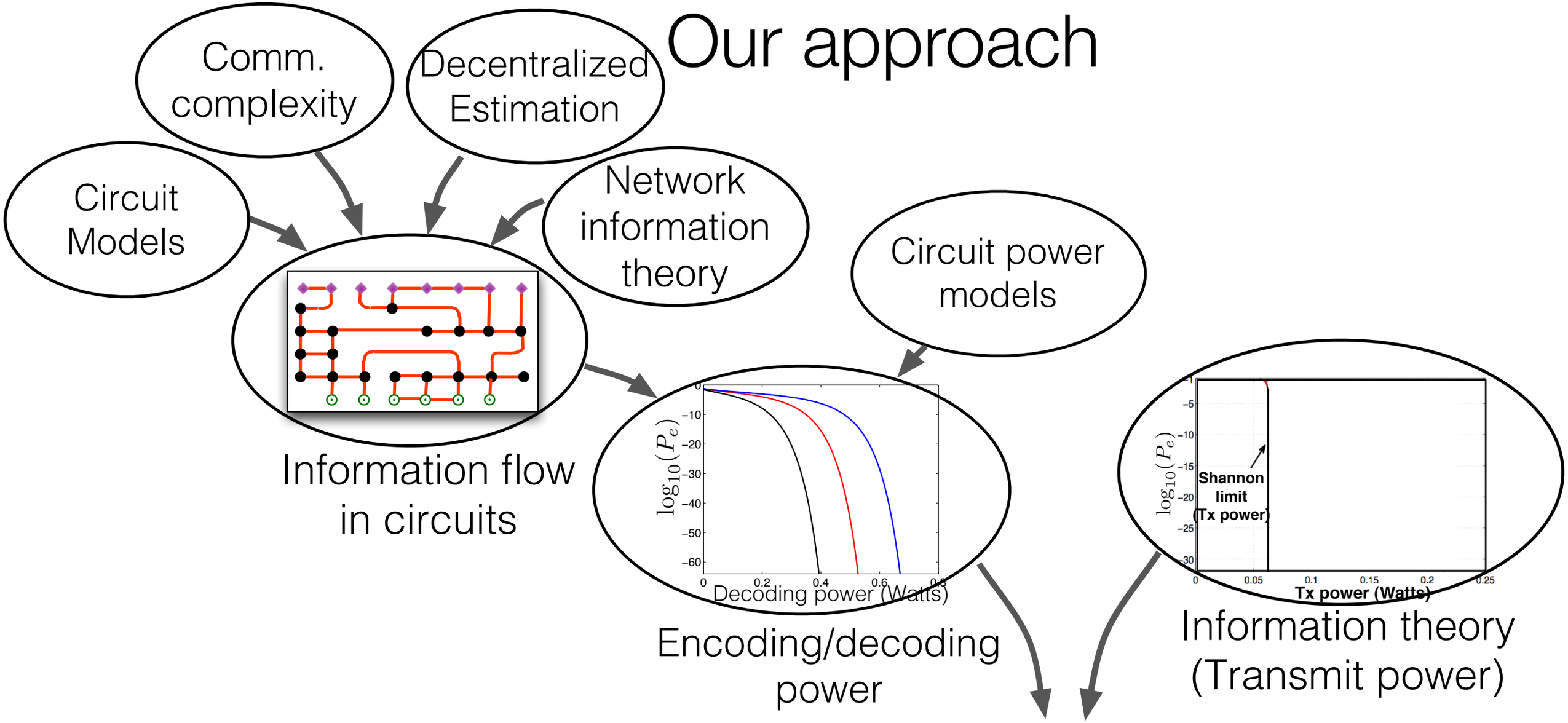


Information flow in circuits

Our approach



Our approach



Our approach

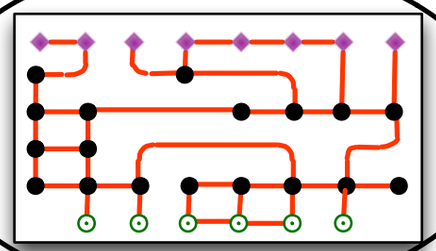
Comm. complexity

Decentralized Estimation

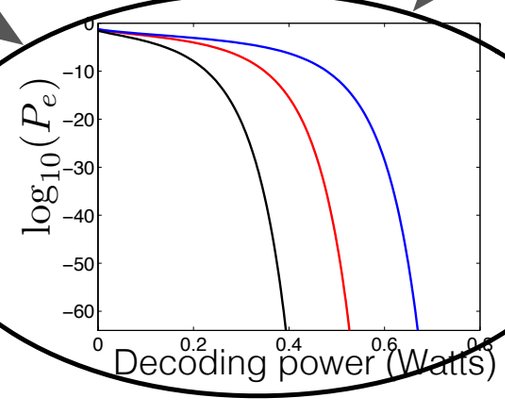
Circuit Models

Network information theory

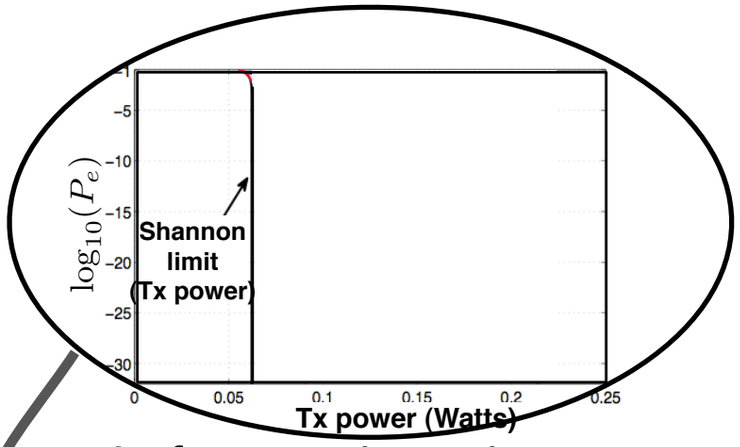
Circuit power models



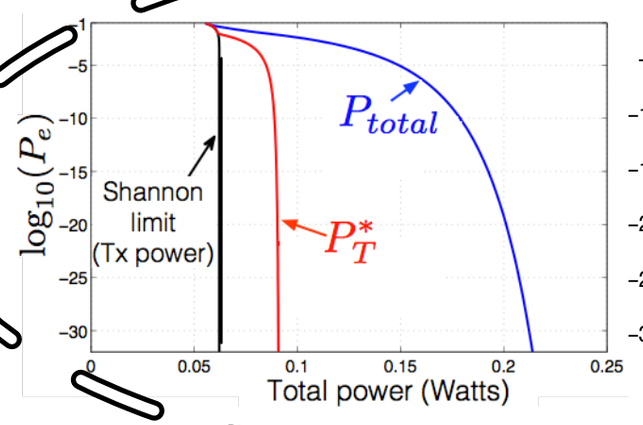
Information flow in circuits

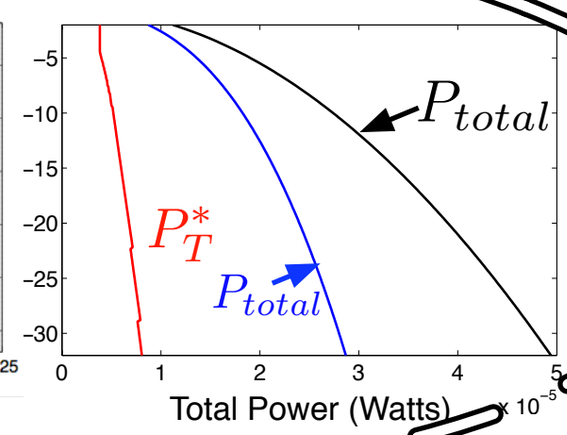


Encoding/decoding power



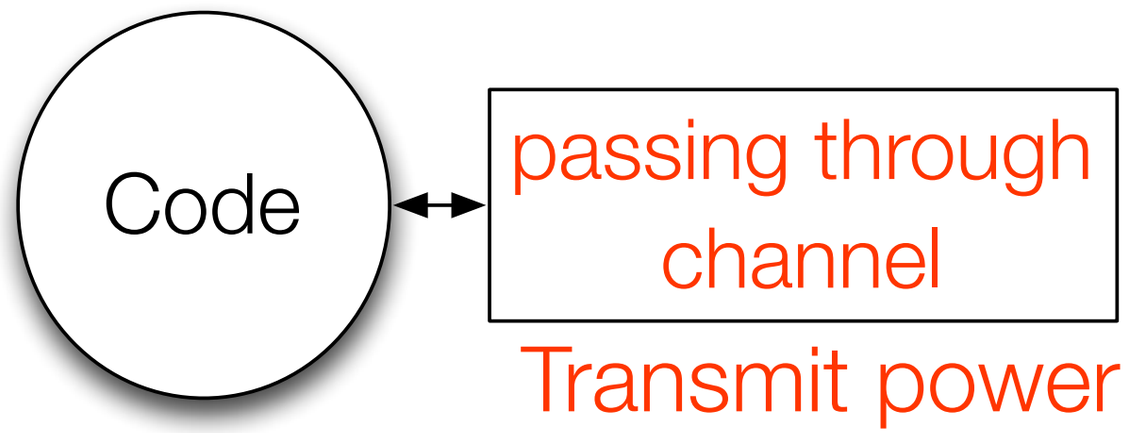
Information theory (Transmit power)





Fundamental bounds on total power

Obtaining fundamental limits: Main difficulty



Transmit power P_T

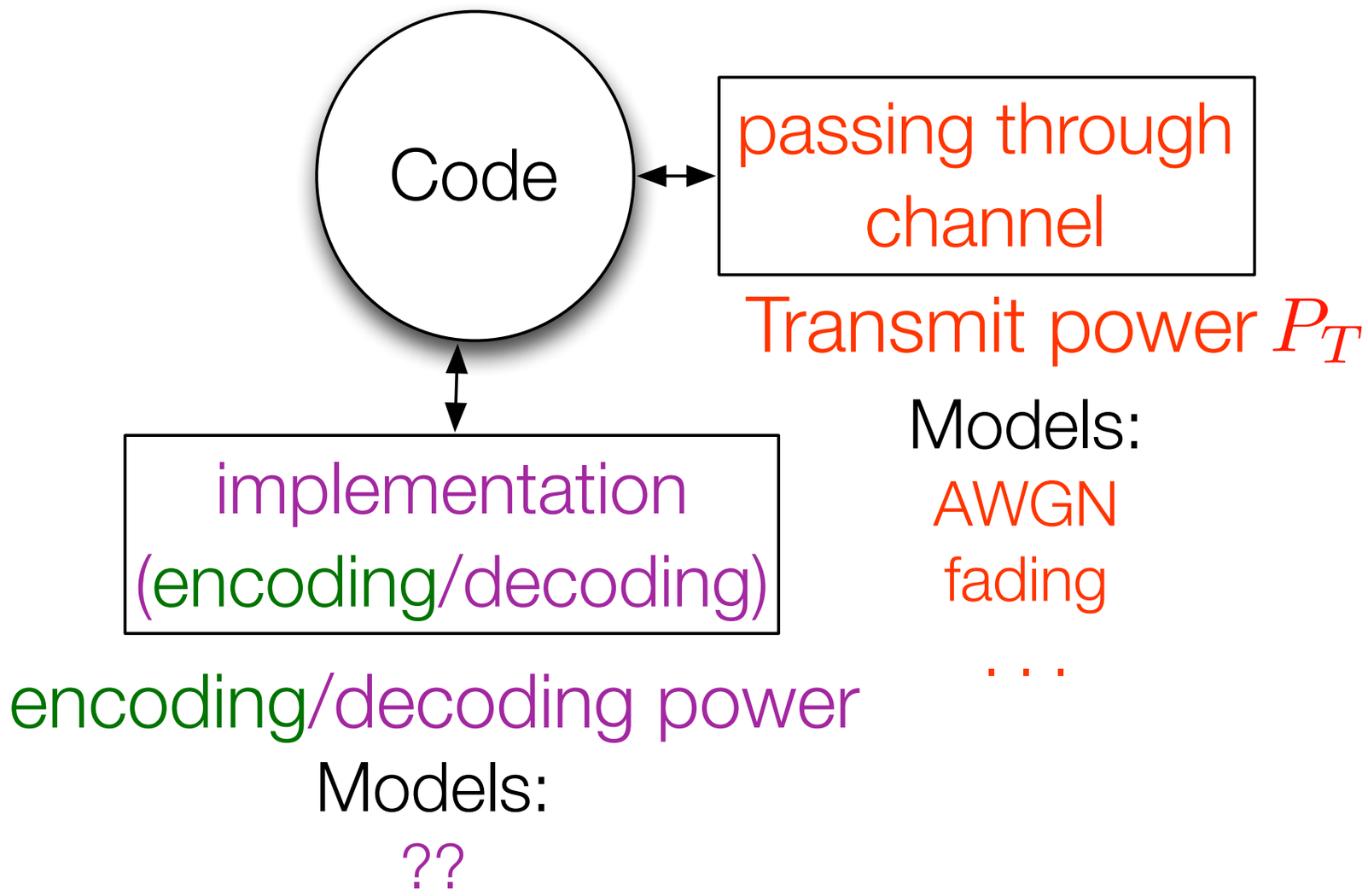
Models:

AWGN

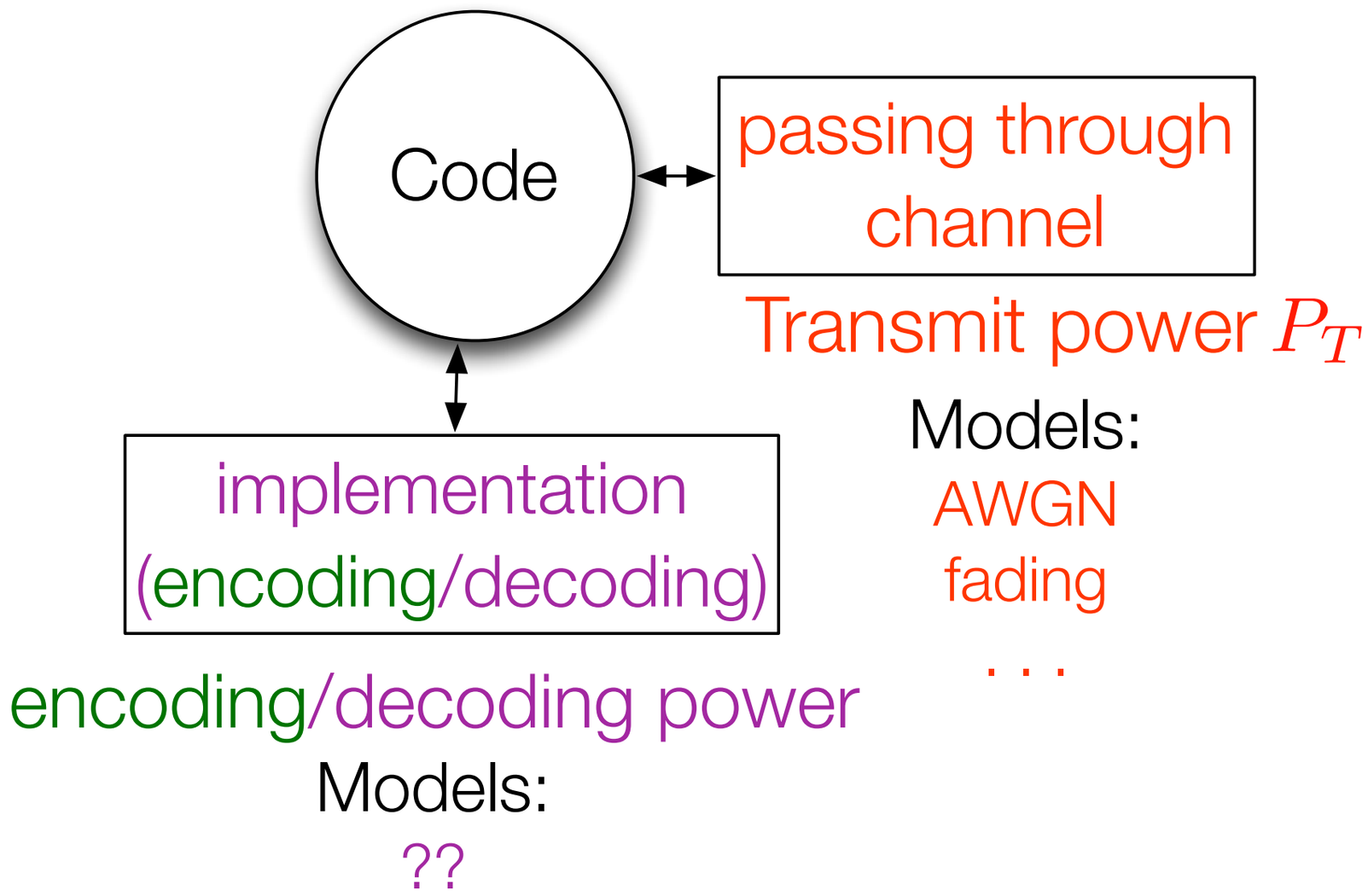
fading

. . .

Obtaining fundamental limits: Main difficulty

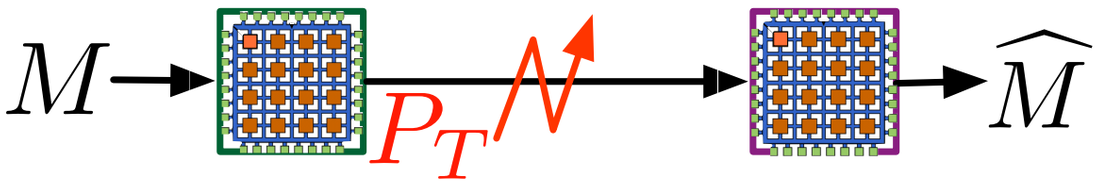


Obtaining fundamental limits: Main difficulty

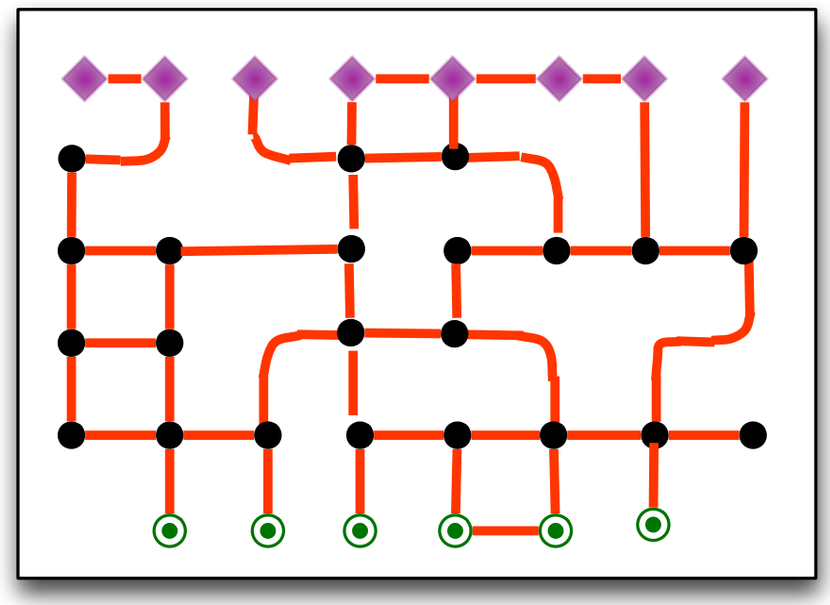


Need implementation models that are relevant and analyzable

A “graphical” model for circuit implementation



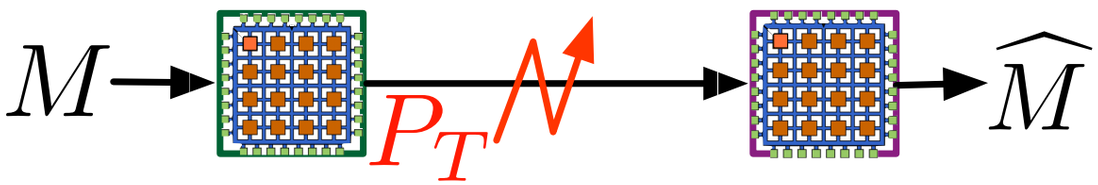
- Output nodes
- Helper nodes
- Input nodes



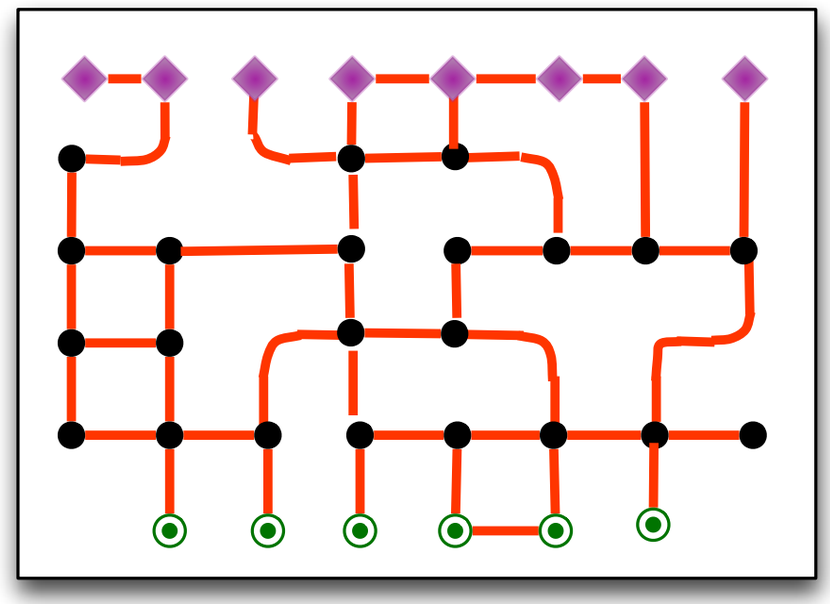
“VLSI model” of computing [Thompson ’80]

- degree of each node at most α (e.g. 4)

A “graphical” model for circuit implementation



- Output nodes
- Helper nodes
- Input nodes

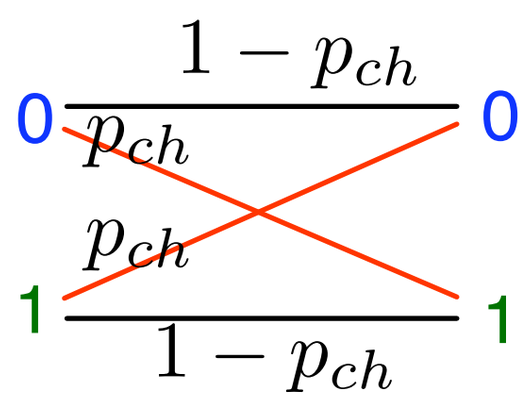


“VLSI model” of computing [Thompson ’80]

- degree of each node at most α (e.g. 4)

Channel model:

Binary Symmetric Channel



$$p_{ch} = \mathbb{Q}\left(\sqrt{\frac{\eta P_T}{\frac{N_0}{2}}}\right)$$

Decoding complexity/power: Fundamental limits

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

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

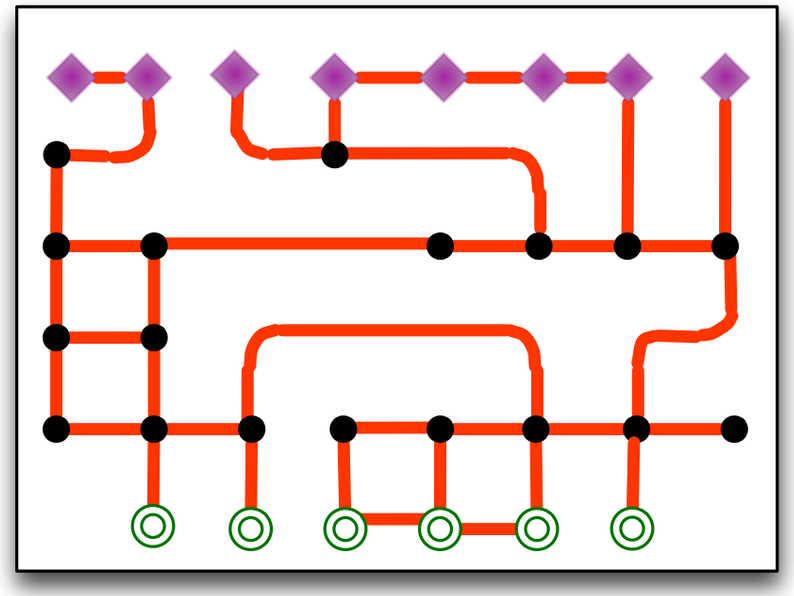
. . . for **any** code and **any** decoding algorithm.

α : Max. node degree

$V(P_T)$: "channel dispersion"

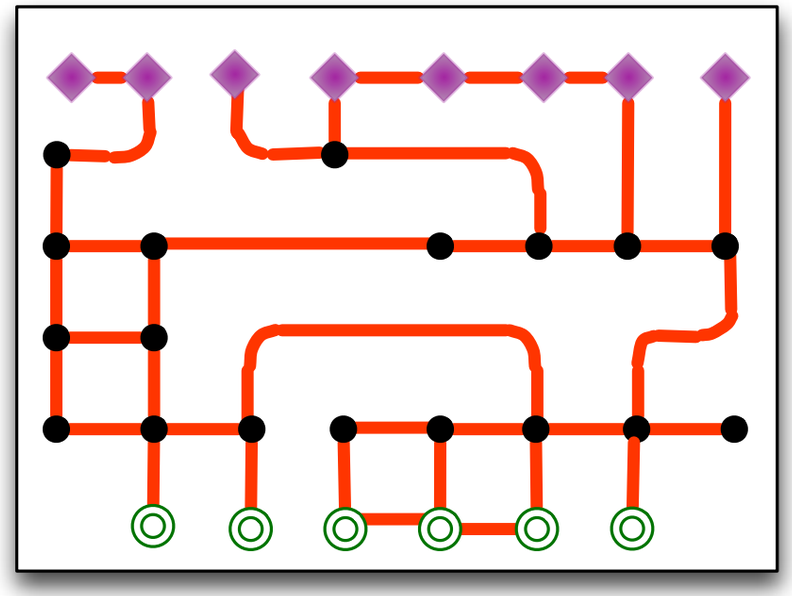
* precise result for any P_e and any gap appears in [Grover, Woyach, Sahai '11]

Decoding complexity: Proof Technique

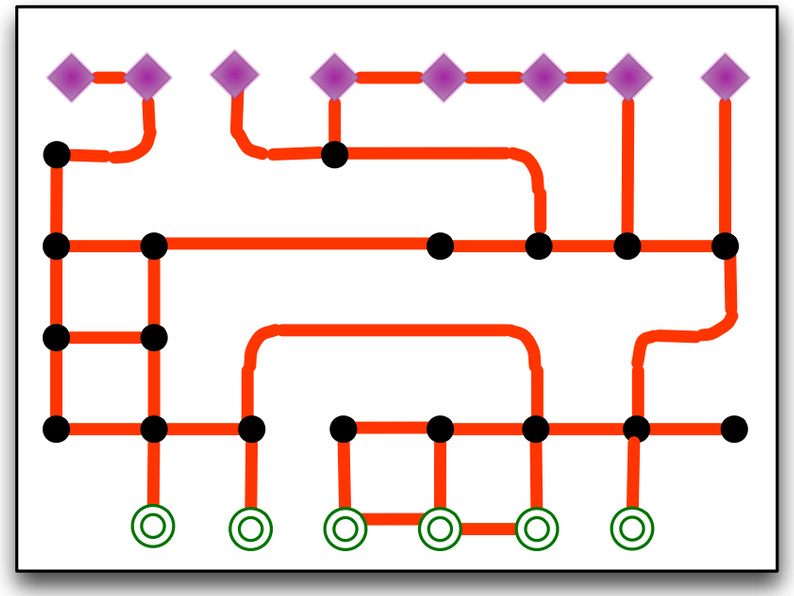


Decoding complexity: Proof Technique

Information-flow constrained by “fan-out”



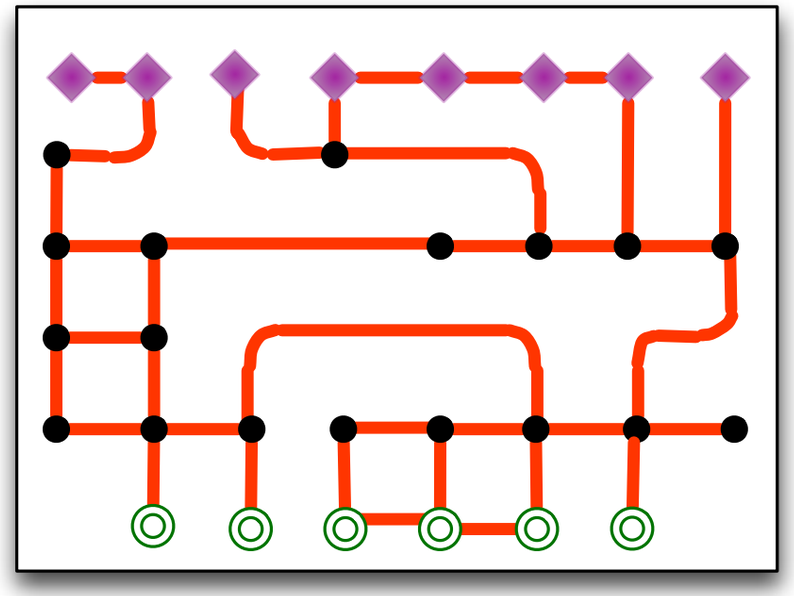
Decoding complexity: Proof Technique



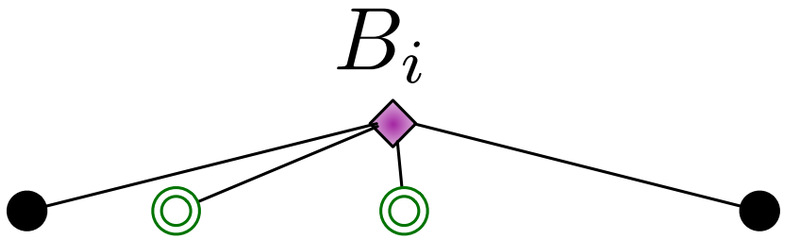
Information-flow constrained by “fan-out”

B_i

Decoding complexity: Proof Technique

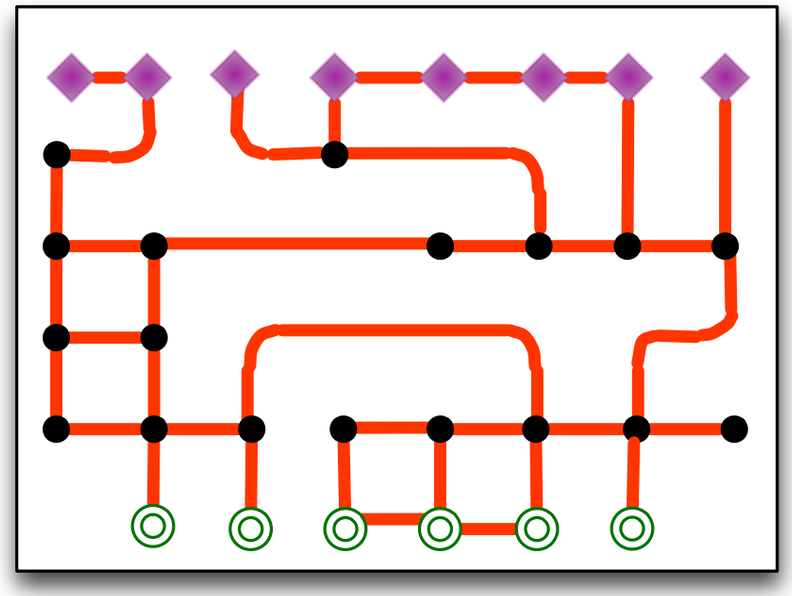


Information-flow constrained by “fan-out”

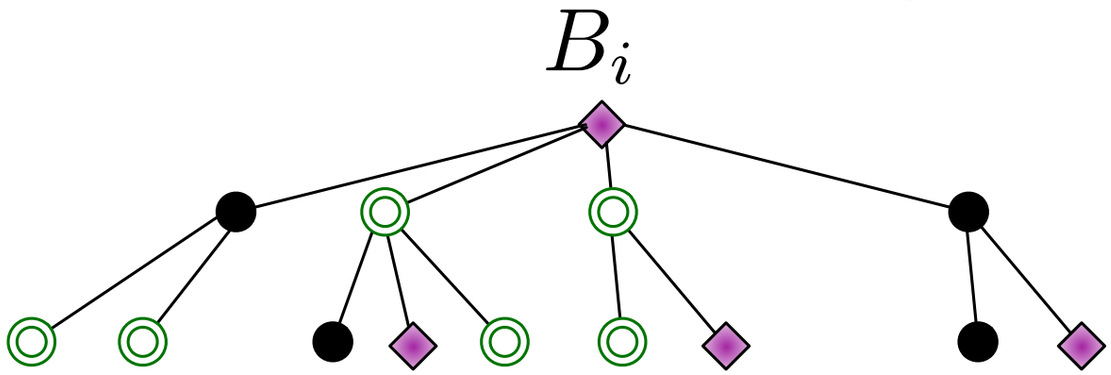


$\tau = 1$

Decoding complexity: Proof Technique



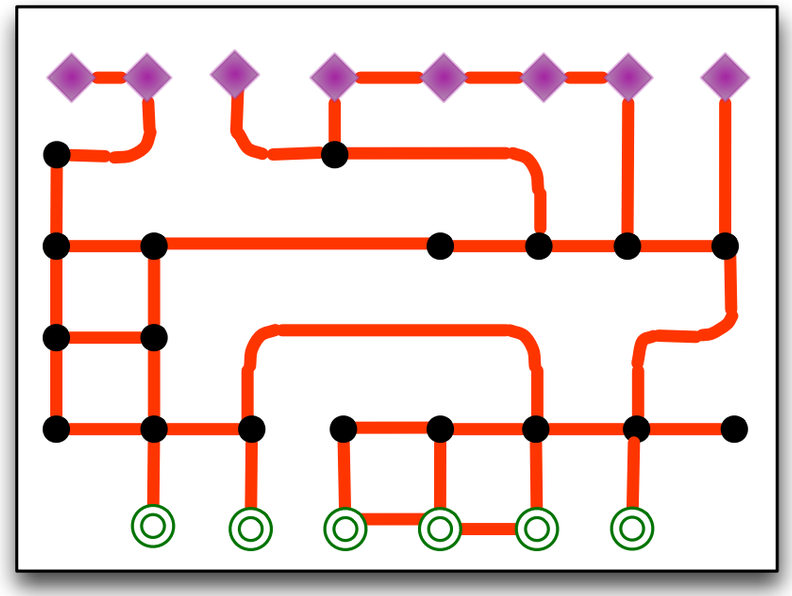
Information-flow constrained by “fan-out”



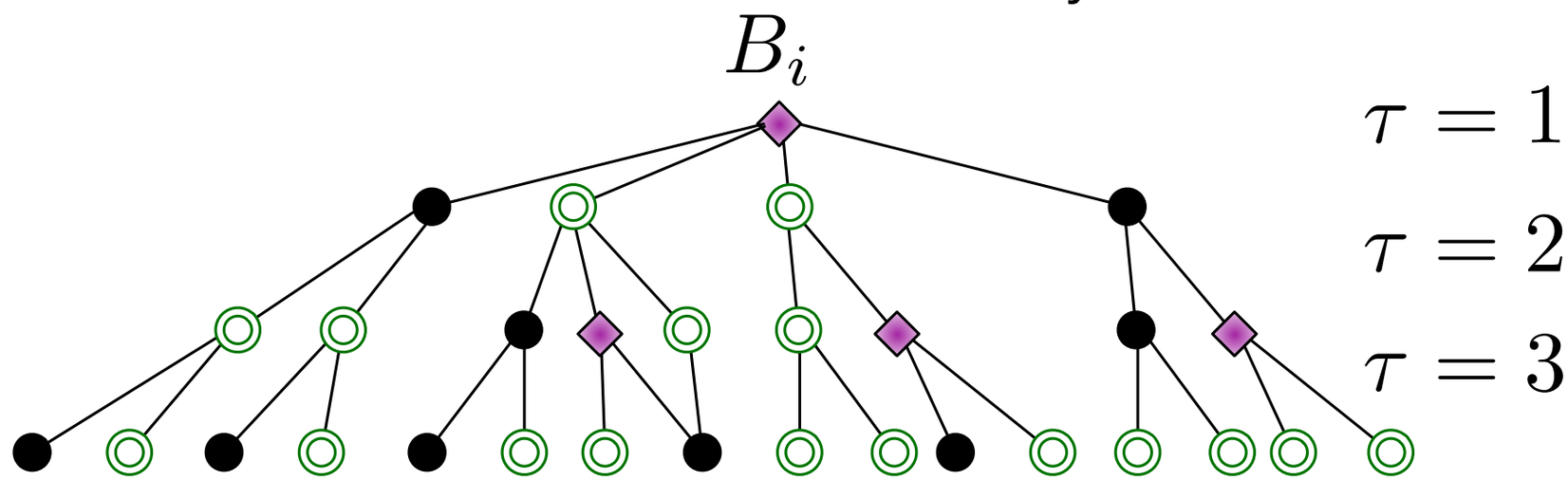
$\tau = 1$

$\tau = 2$

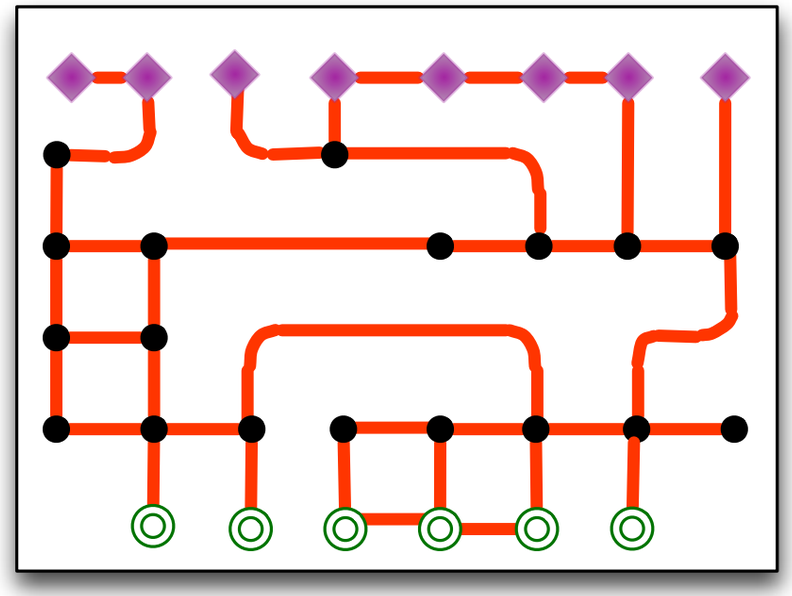
Decoding complexity: Proof Technique



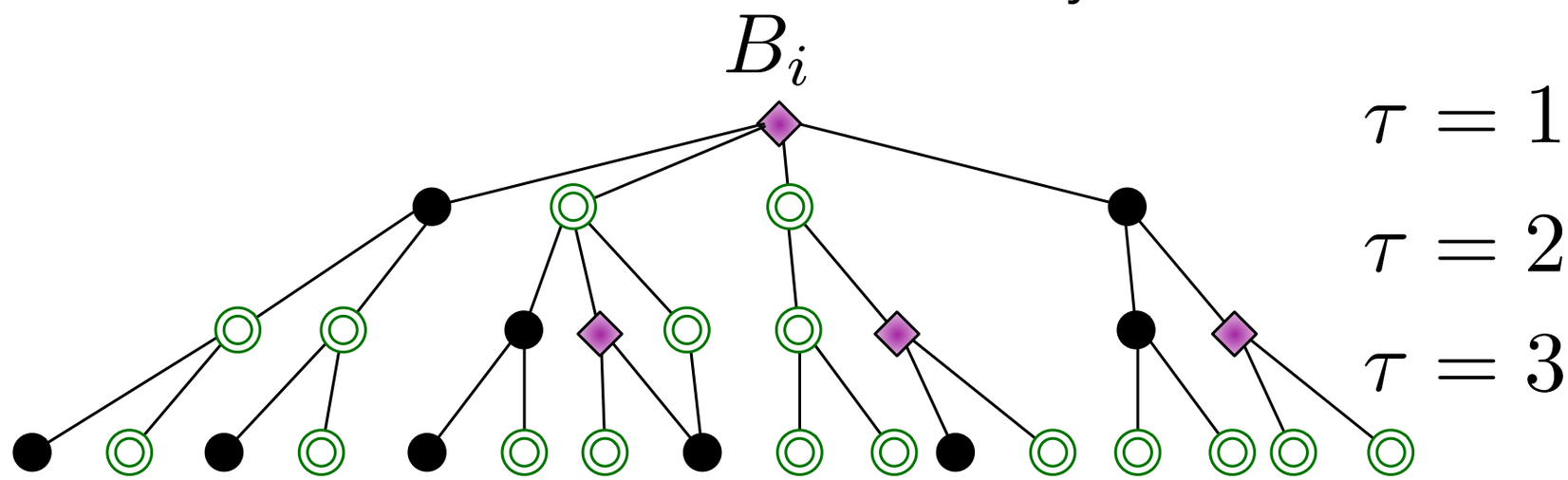
Information-flow constrained by “fan-out”



Decoding complexity: Proof Technique



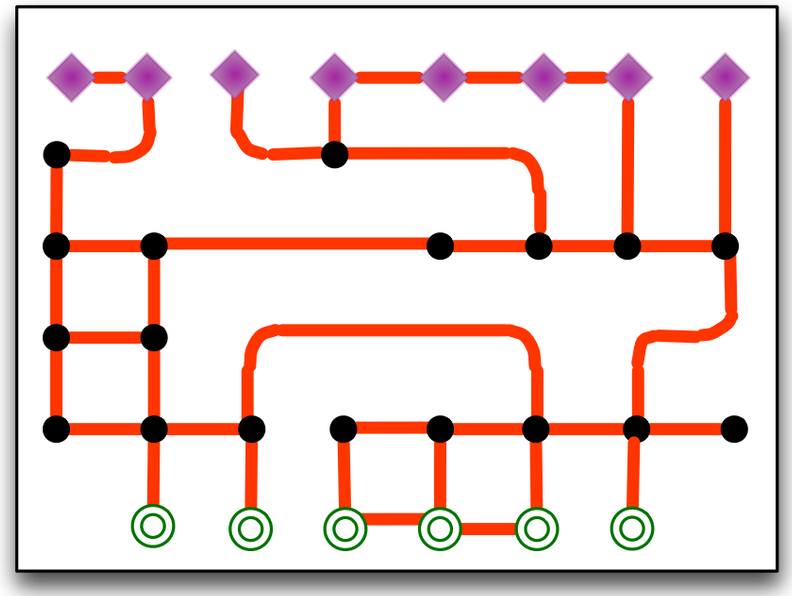
Information-flow constrained by “fan-out”



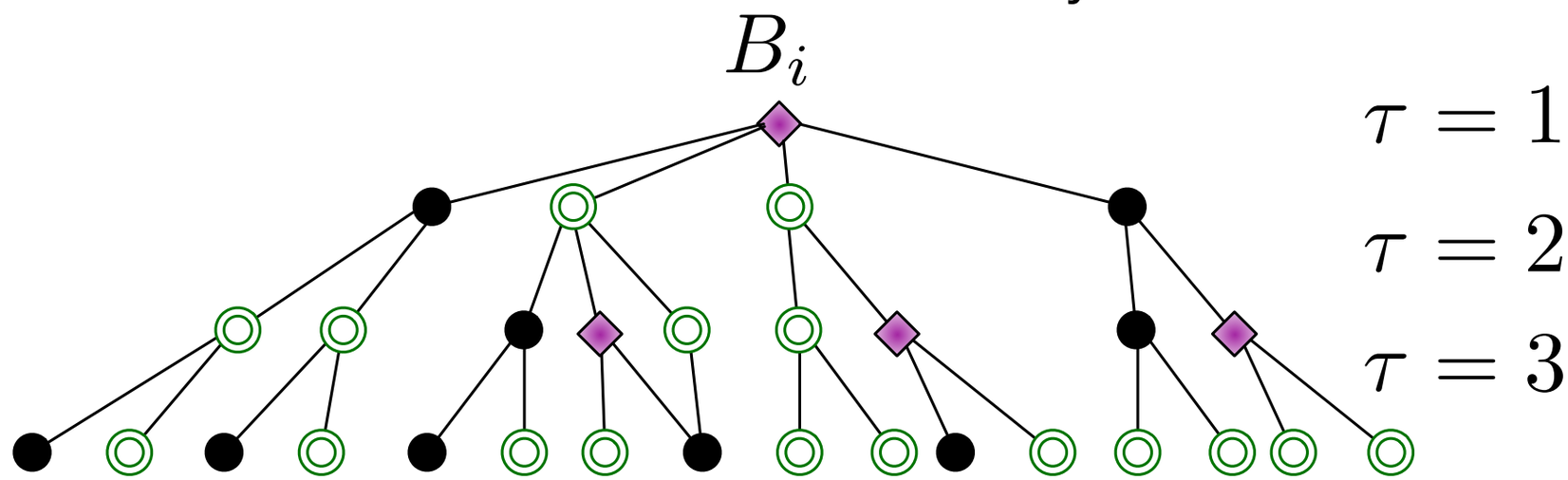
“Sphere-packing” for decoding neighborhoods

[Grover ’07][Grover, Sahai ’08][Grover, Woyach, Sahai ’11]

Decoding complexity: Proof Technique



Information-flow constrained by "fan-out"

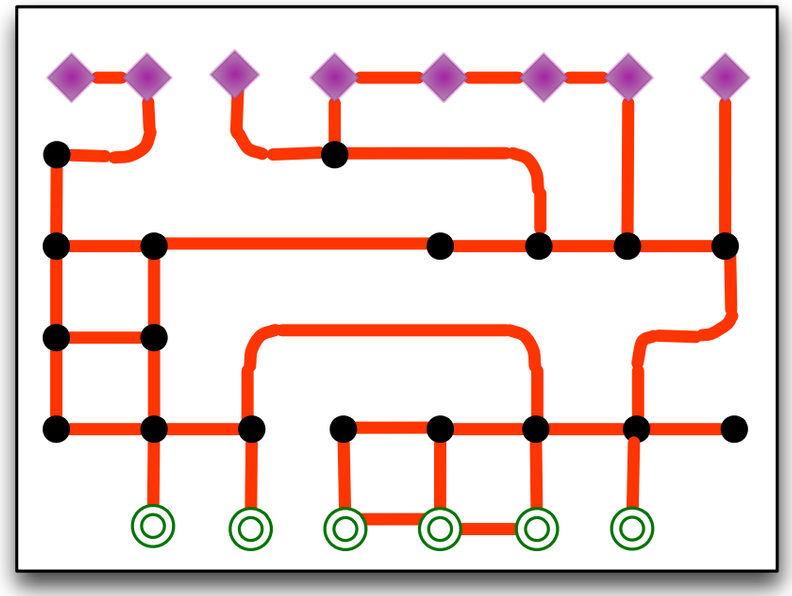


"Sphere-packing" for decoding neighborhoods

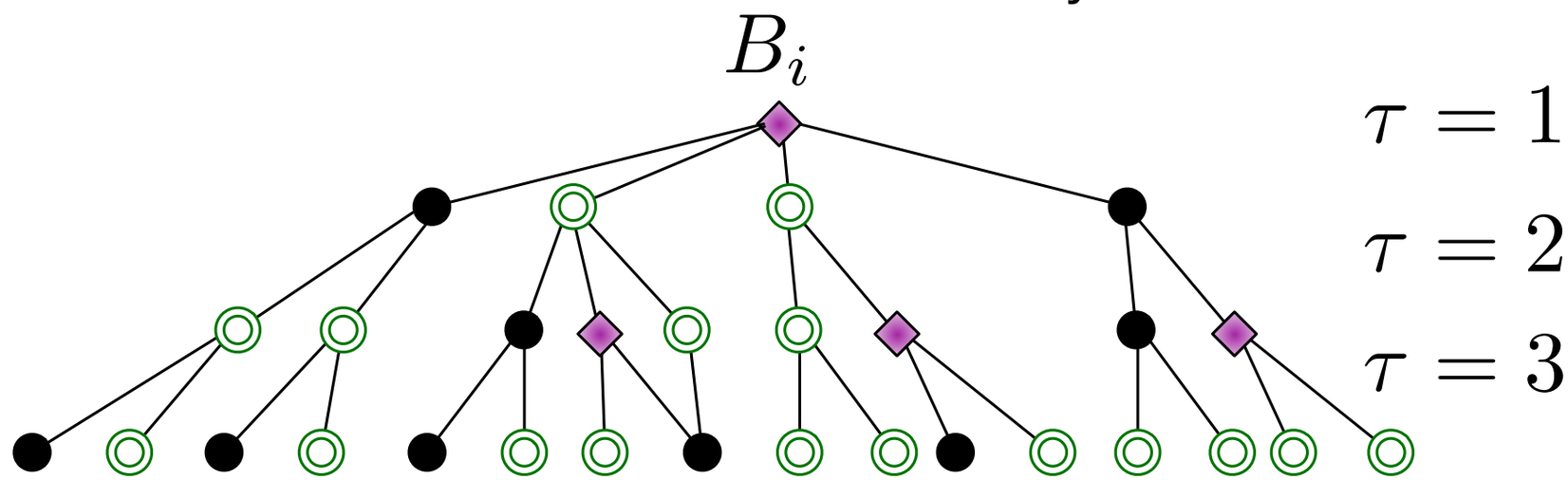
[Grover '07][Grover, Sahai '08][Grover, Woyach, Sahai '11]

n_{nbd} : number of nodes in decoding neighborhood

Decoding complexity: Proof Technique



Information-flow constrained by "fan-out"



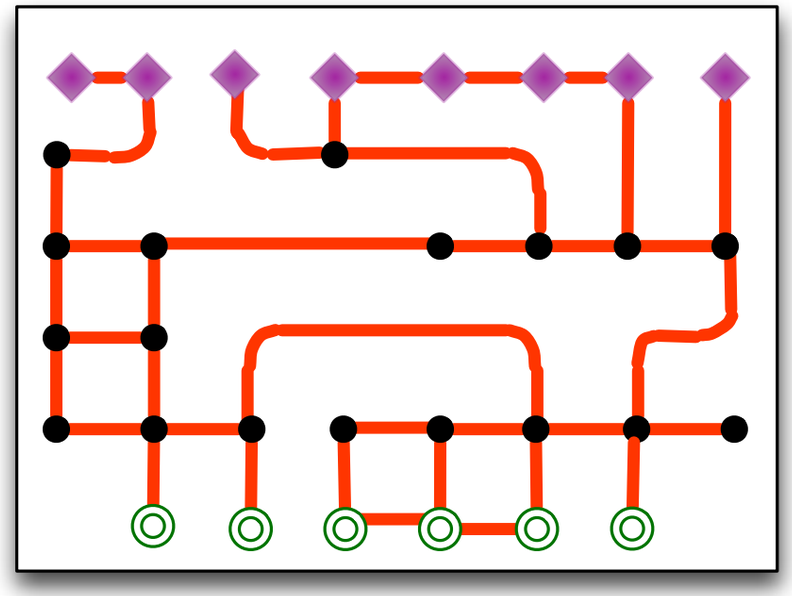
"Sphere-packing" for decoding neighborhoods

[Grover '07][Grover, Sahai '08][Grover, Woyach, Sahai '11]

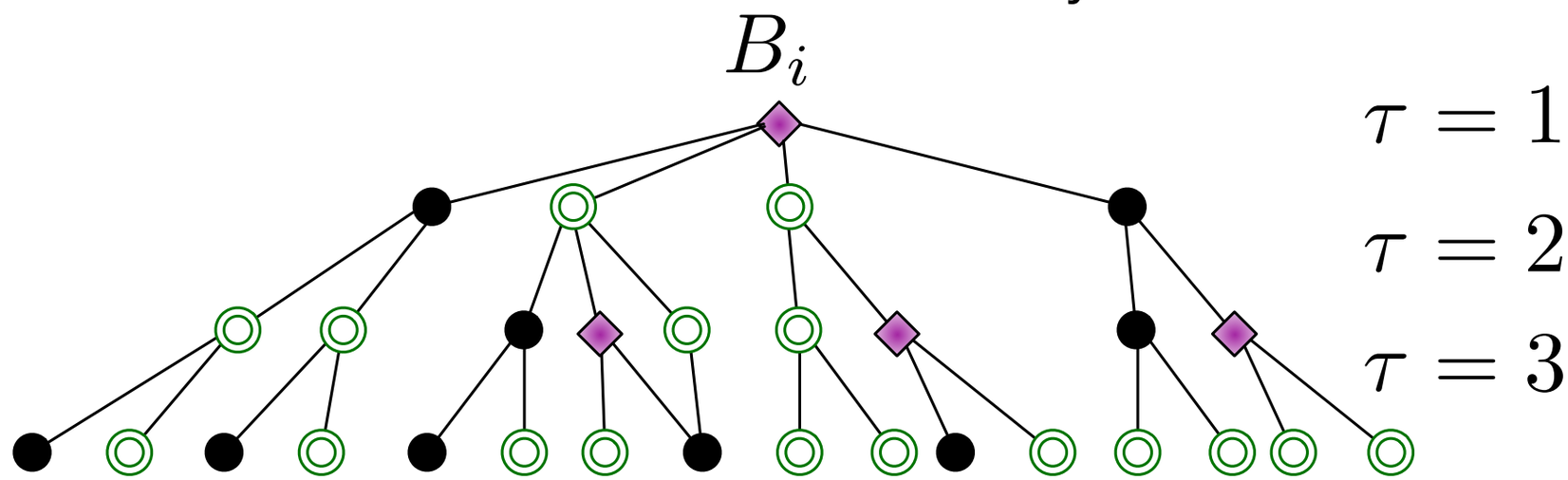
n_{nbd} : number of nodes in decoding neighborhood

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

Decoding complexity: Proof Technique



Information-flow constrained by "fan-out"



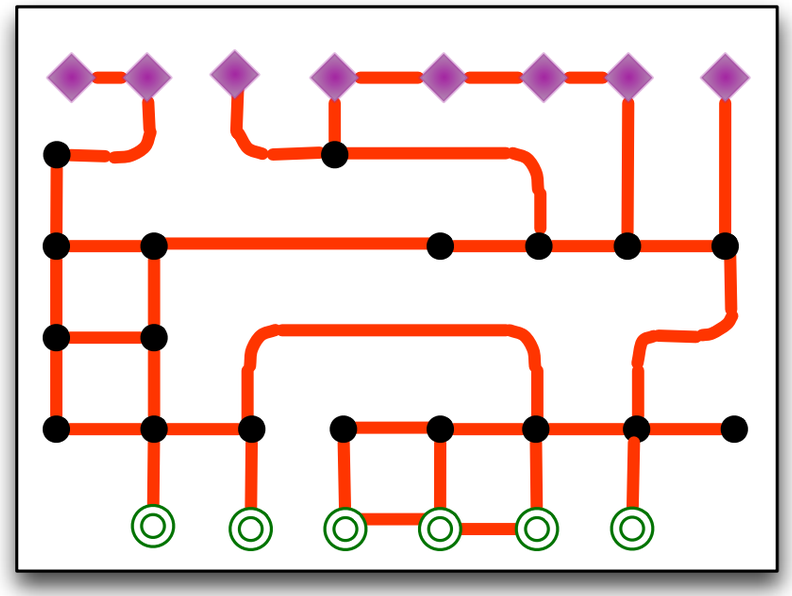
"Sphere-packing" for decoding neighborhoods

[Grover '07][Grover, Sahai '08][Grover, Woyach, Sahai '11]

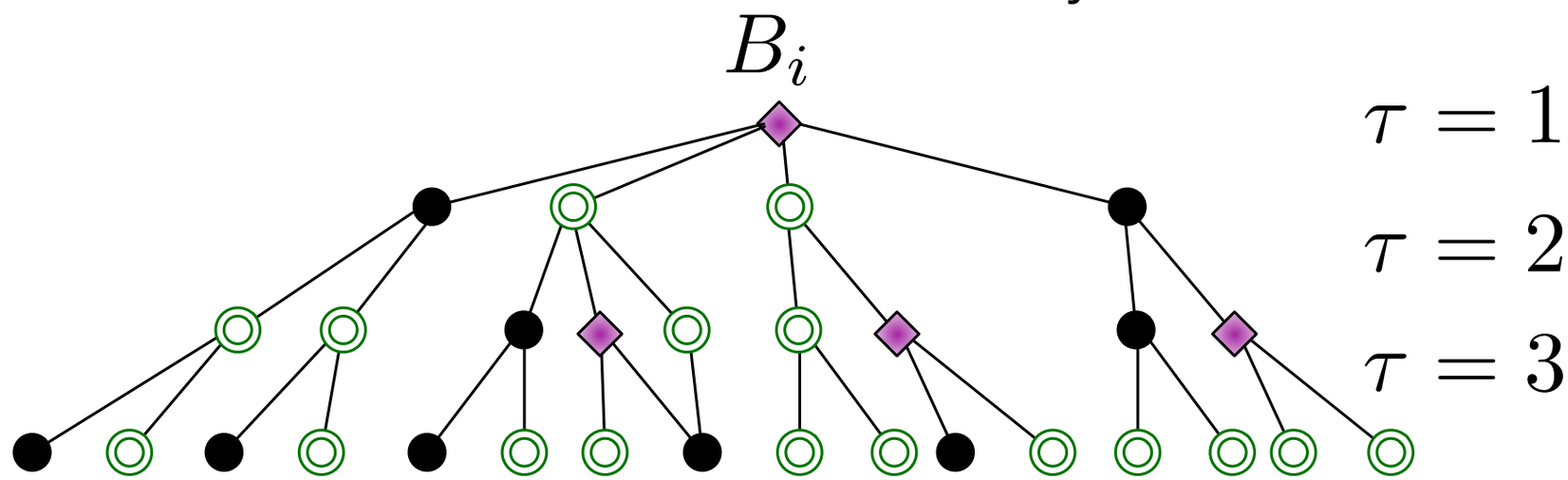
n_{nbd} : number of nodes in decoding neighborhood

$$P_e \gtrsim e^{-n_{nbd}\frac{(C(P_T)-R)^2}{V(P_T)}} \qquad \tau \gtrsim \frac{1}{\log(\alpha-1)}\log(n_{nbd})$$

Decoding complexity: Proof Technique



Information-flow constrained by "fan-out"



"Sphere-packing" for decoding neighborhoods

[Grover '07][Grover, Sahai '08][Grover, Woyach, Sahai '11]

n_{nbd} : number of nodes in decoding neighborhood

$$P_e \gtrsim e^{-n_{nbd}\frac{(C(P_T)-R)^2}{V(P_T)}} \qquad \tau \gtrsim \frac{1}{\log(\alpha-1)}\log(n_{nbd})$$

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

Node model: 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

$$P_{total} = \min_{P_T} P_T + P_{decoding}$$

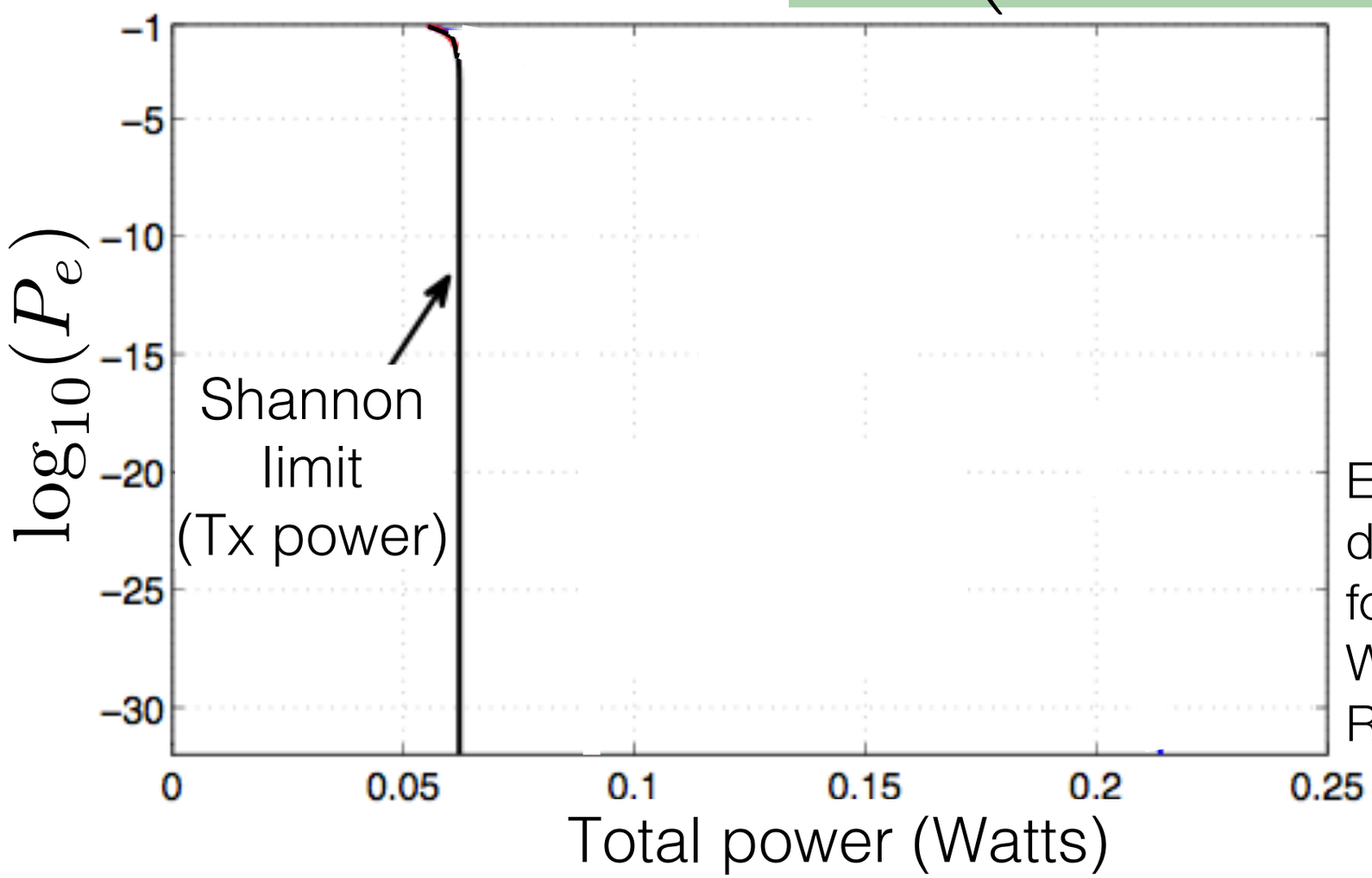
$$\geq \min_{P_T} P_T + \gamma \log \left(\frac{\log \frac{1}{P_e}}{(C(P_T) - R)^2 / V(P_T)} \right)$$

Node model:
Fundamental bounds on total power

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

$$P_{total} = \min_{P_T} P_T + P_{decoding}$$

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



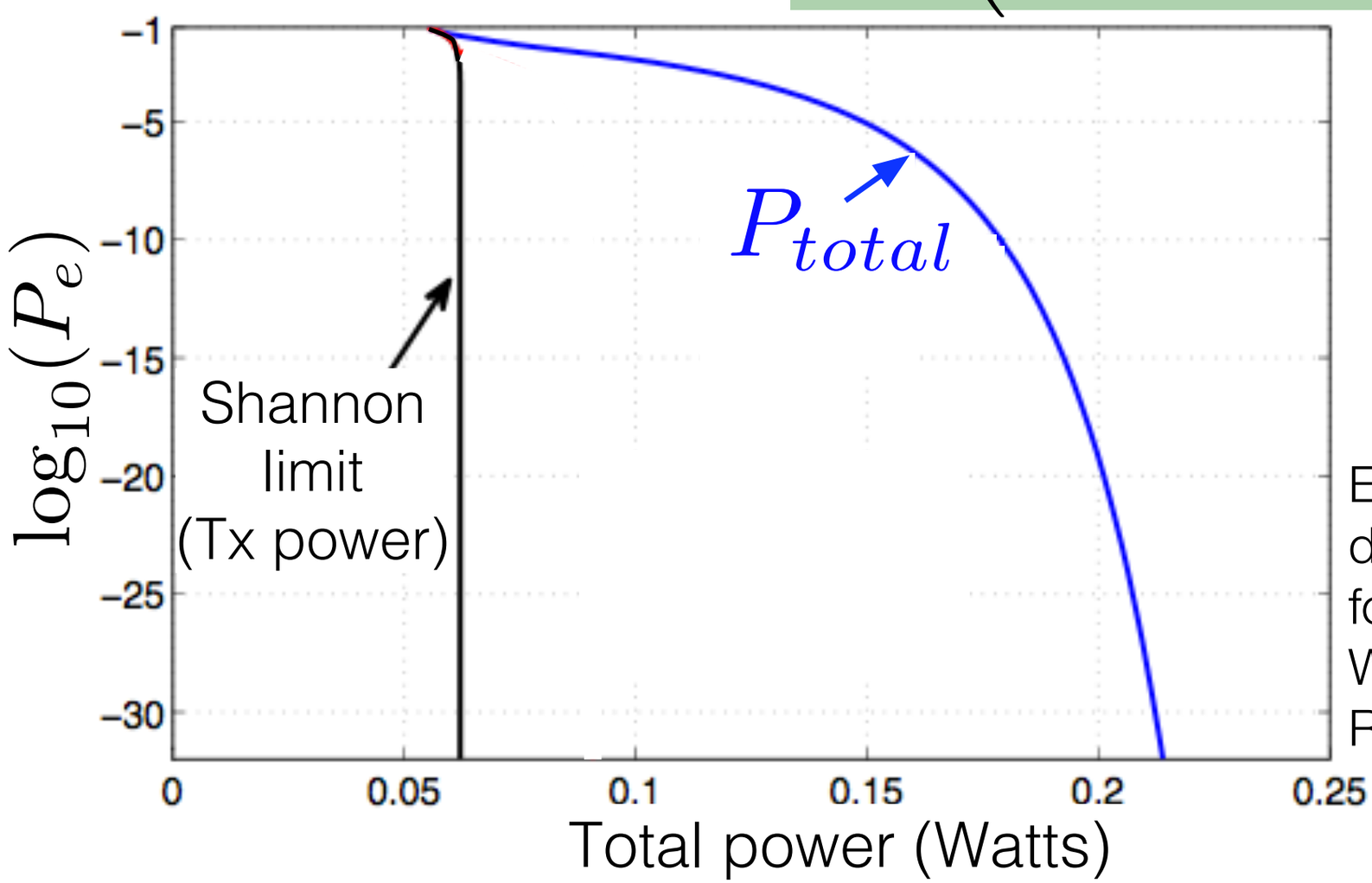
Enode = 3 pJ
distance = 10 m
fc = 60 GHz
W = 3 GHz
R = 1.5 Gbps

Node model:
Fundamental bounds on total power

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

$$P_{total} = \min_{P_T} P_T + P_{decoding}$$

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



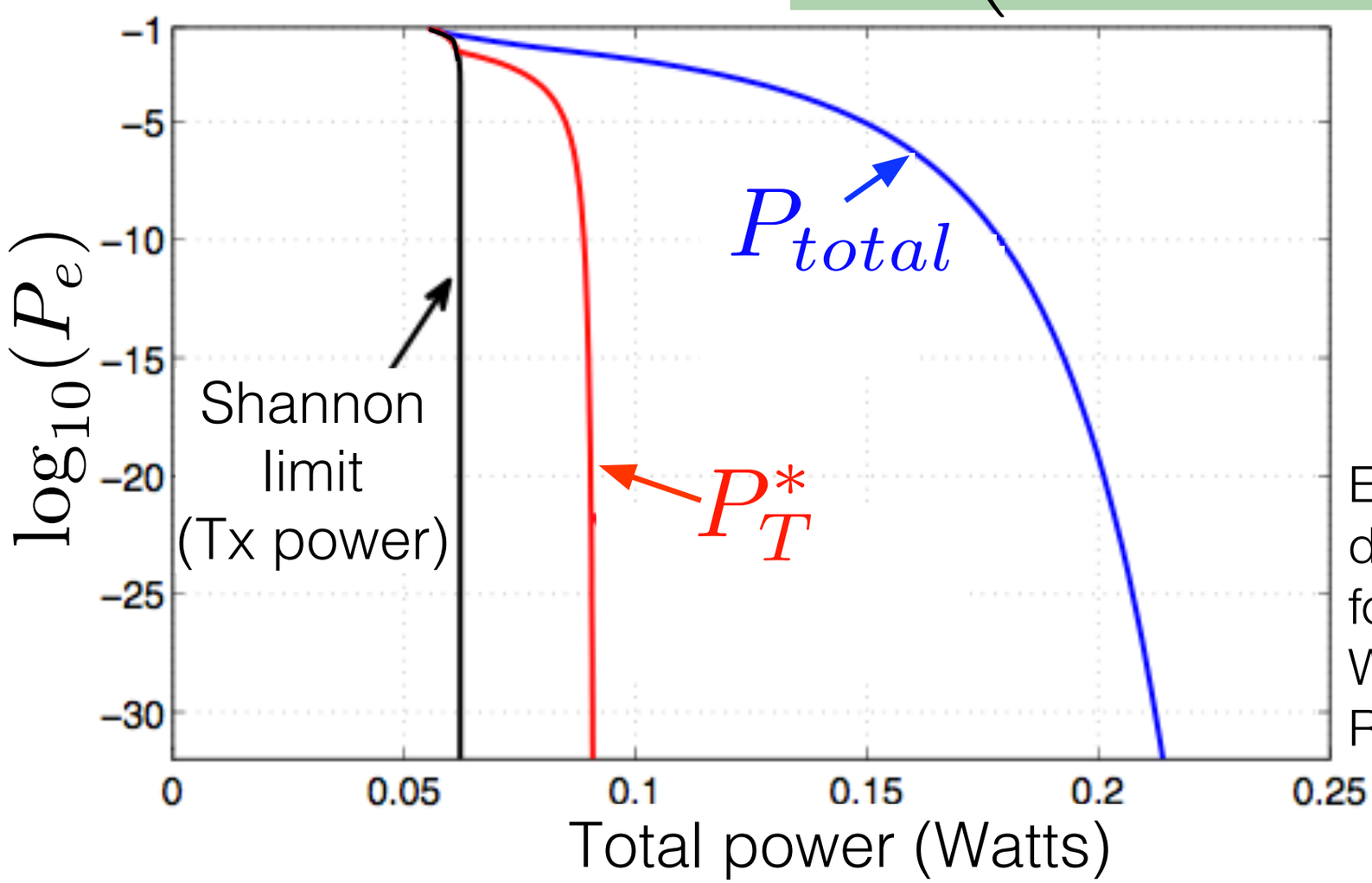
Enode = 3 pJ
distance = 10 m
fc = 60 GHz
W = 3 GHz
R = 1.5 Gbps

Node model:
Fundamental bounds on total power

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

$$P_{total} = \min_{P_T} P_T + P_{decoding}$$

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



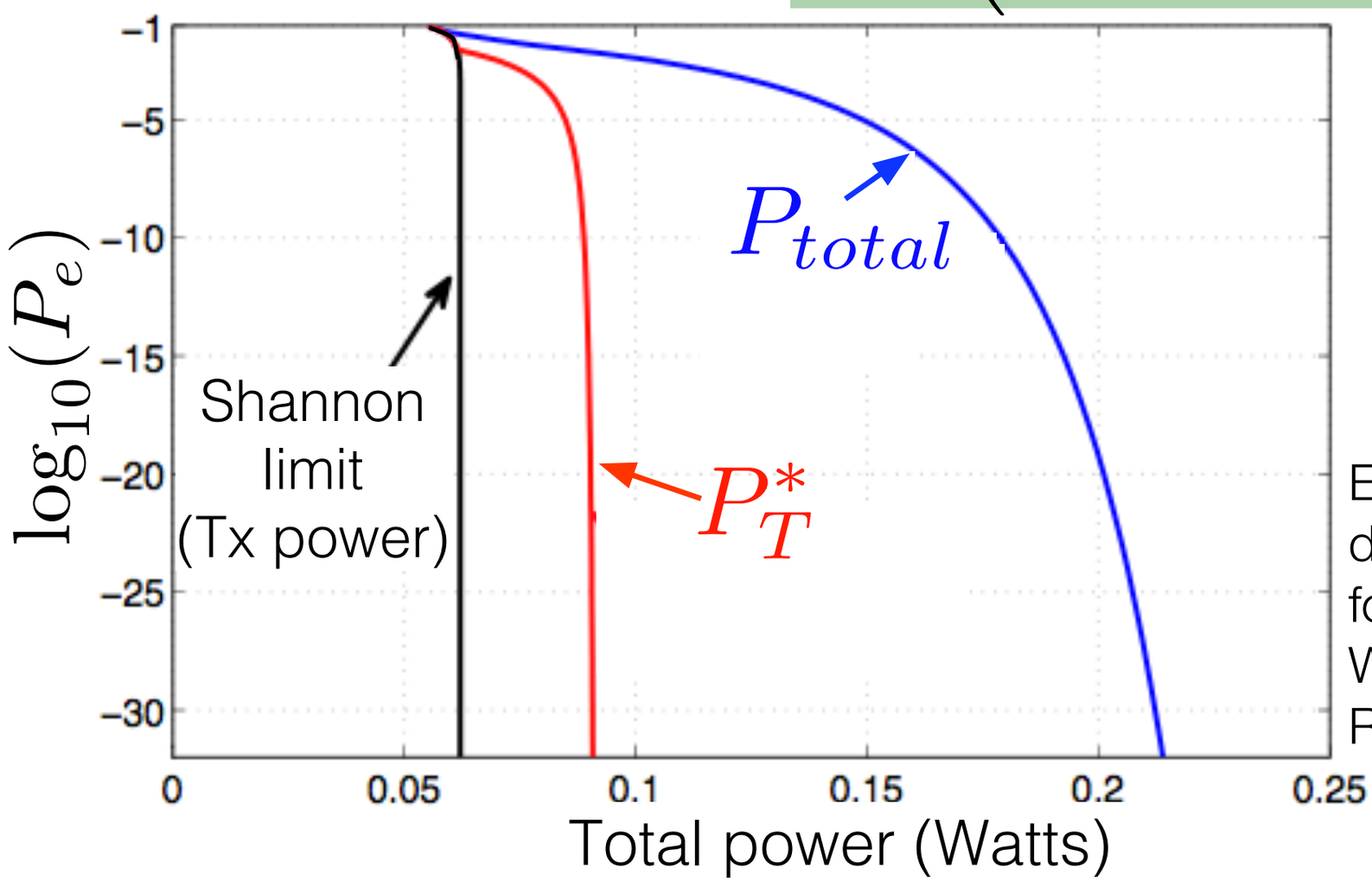
Enode = 3 pJ
distance = 10 m
fc = 60 GHz
W = 3 GHz
R = 1.5 Gbps

Node model:
Fundamental bounds on total power

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

$$P_{total} = \min_{P_T} P_T + P_{decoding}$$

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



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

Moral: do not operate too close to capacity!

Node model
What codes are good?

Node model
What codes are good?

Capacity-approaching LDPCs,
iterative decoding?

Node model
What codes are good?

Capacity-approaching LDPCs,
iterative decoding?

Random block codes,
ML decoding?

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?

Capacity-approaching LDPCs,
iterative decoding?

Convolutional codes,
Viterbi decoding

Random block codes,
ML decoding?

Polar codes,
successive cancelation?

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?

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

~~Random block codes, ML decoding?~~

~~Convolutional codes, Viterbi decoding~~

~~Polar codes, successive cancelation?~~

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

Node model
What codes are good?

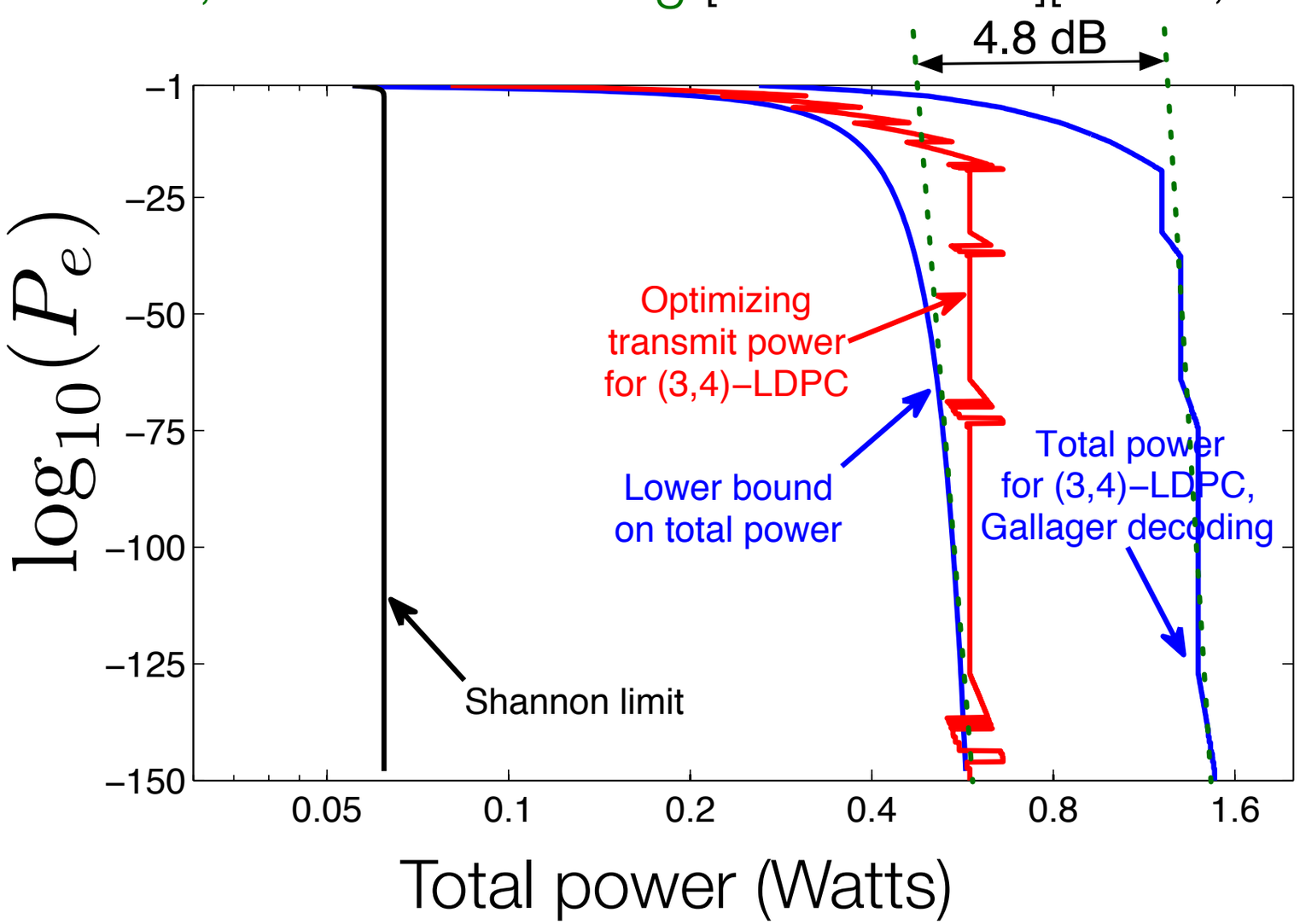
~~Capacity-approaching LDPCs, iterative decoding?~~

~~Random block codes, ML decoding?~~

~~Convolutional codes, Viterbi decoding~~

~~Polar codes, successive cancelation?~~

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



Node model
What codes are good?

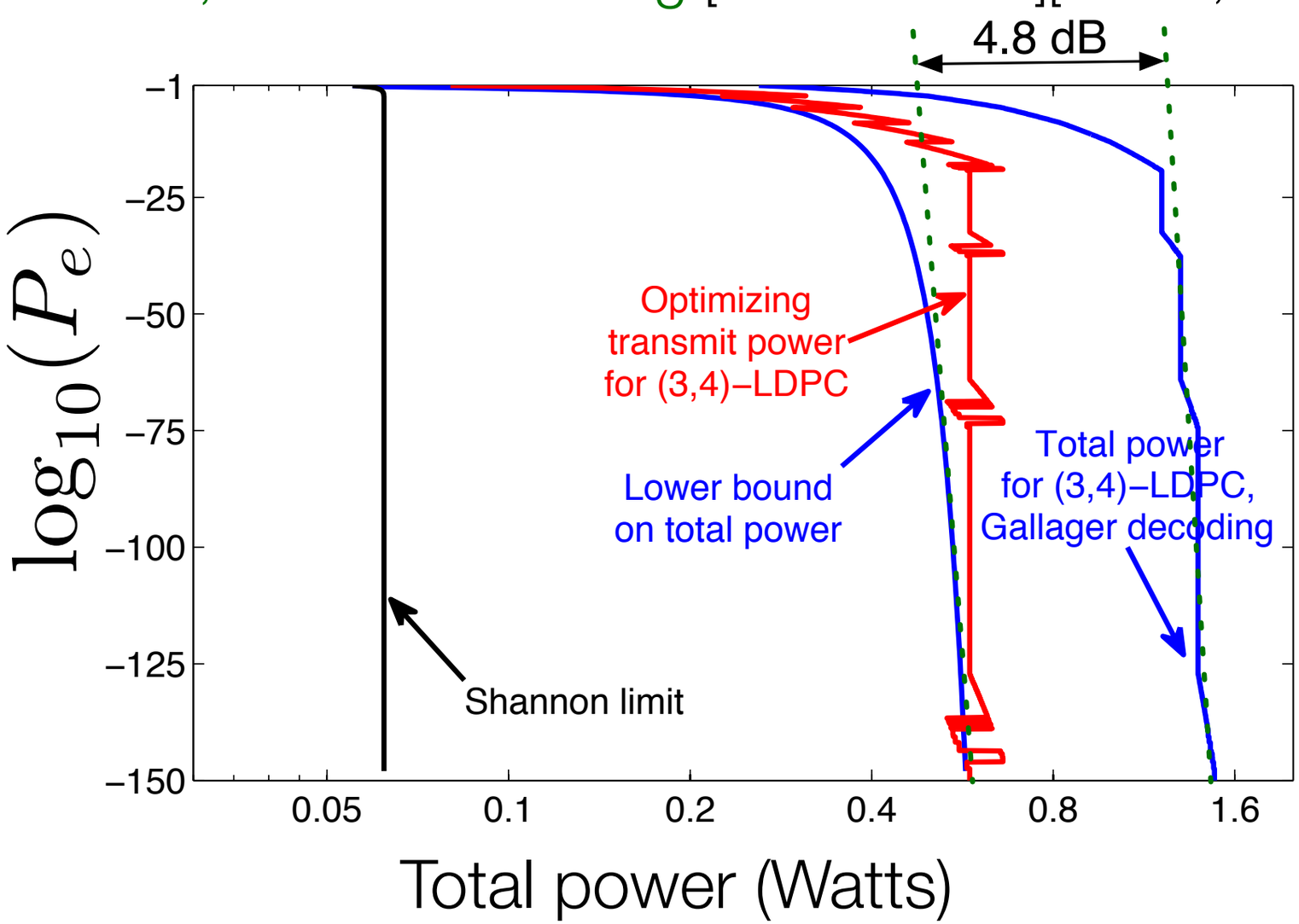
~~Capacity-approaching LDPCs, iterative decoding?~~

~~Random block codes, ML decoding?~~

~~Convolutional codes, Viterbi decoding~~

~~Polar codes, successive cancelation?~~

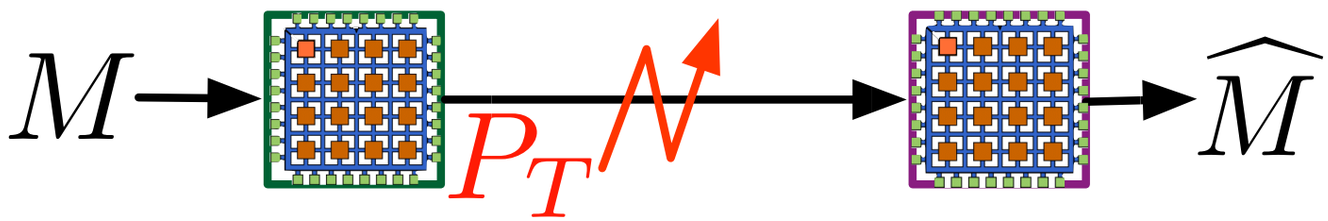
Regular LDPCs, iterative decoding [Lentmaier '05][Grover, Woyach, Sahai JSAC'11]



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

Talk outline

Information flows: visible and invisible



Traditional: Minimize P_T

Now: Minimize total (transmit + proc.) power

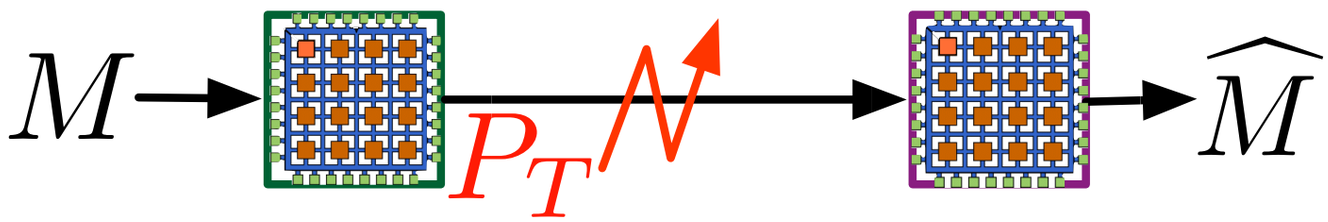
Fundamental limits

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

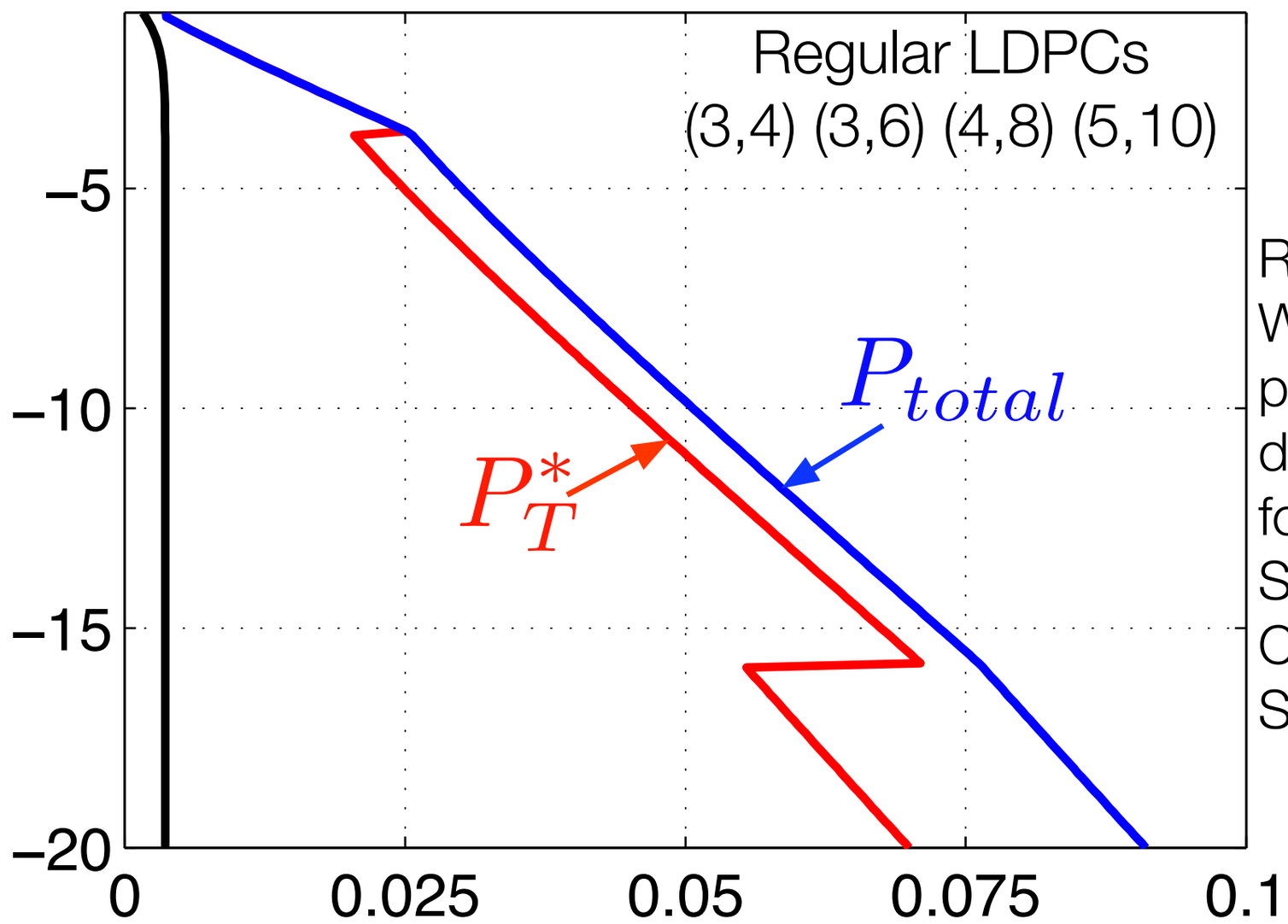
Node model: stay away from capacity

regular LDPCs are order optimal

Wire model: using bounded P_T suboptimal

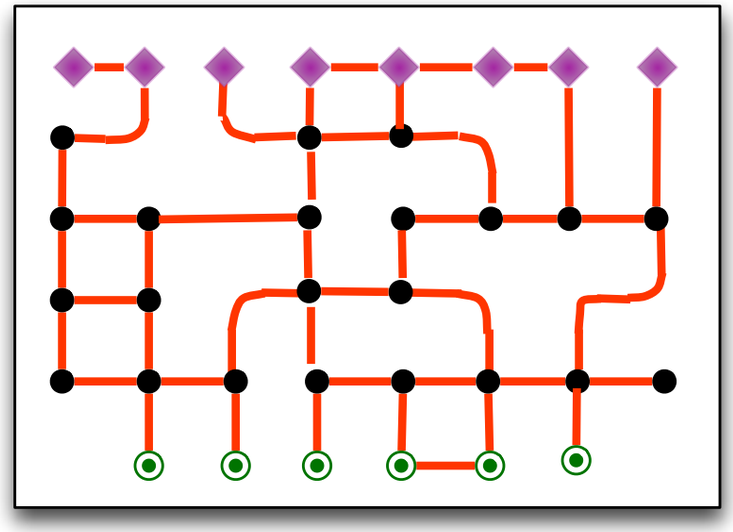
What did we expect from our experimental results?

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

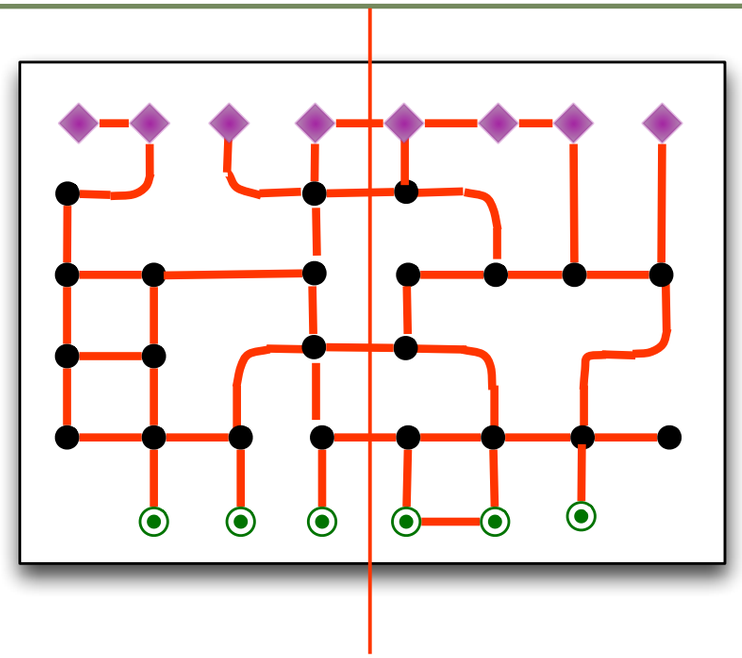


R = 7 Gbps
W = 7 GHz
path-loss coefft = 3
d = 3 meters
fc = 60 GHz
STM CMOS 90 nm
Cadence Encnter (Layout)
Synopsis HSPICE

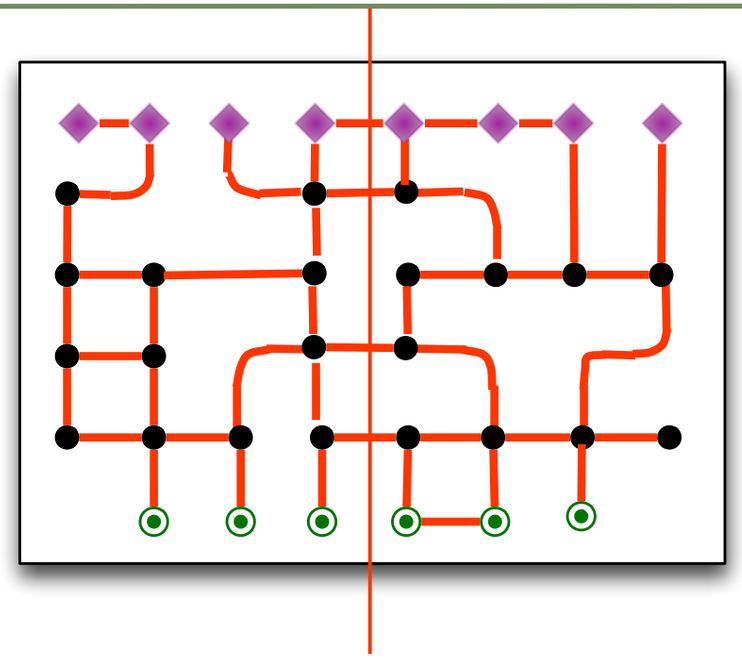
Wire model of power consumption



Wire model of power consumption



Wire model of power consumption

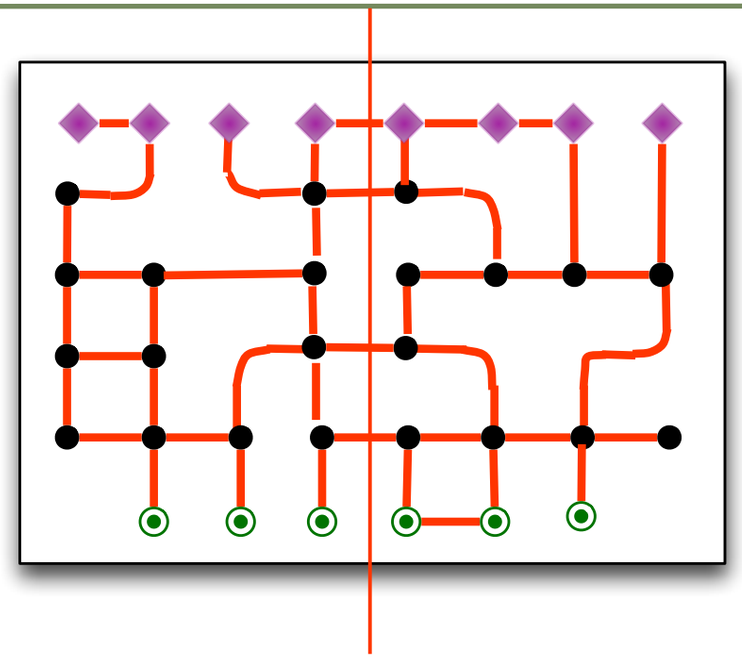


Fourier Transform [Thompson '80]

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

Similar results for sorting, boolean functions, . . .

Wire model of power consumption



Fourier Transform [Thompson '80]

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

Similar results for sorting, boolean functions, . . .

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

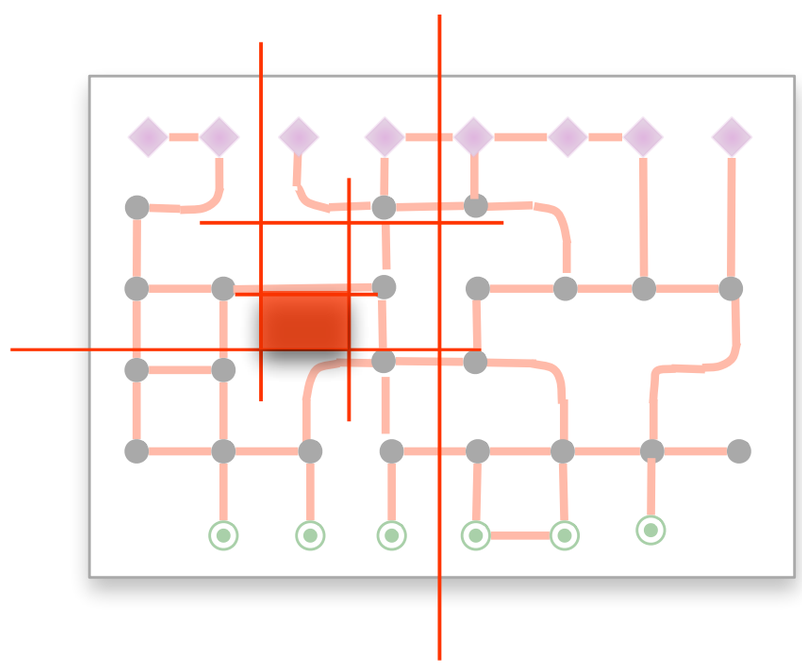
$$A_{chip}\tau^2 \geq \Omega\left(nR^2 \log \frac{n}{P_e}\right)$$

Wire model:
Fundamental limits on encoding/decoding power

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

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

. . . 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 on **total** power

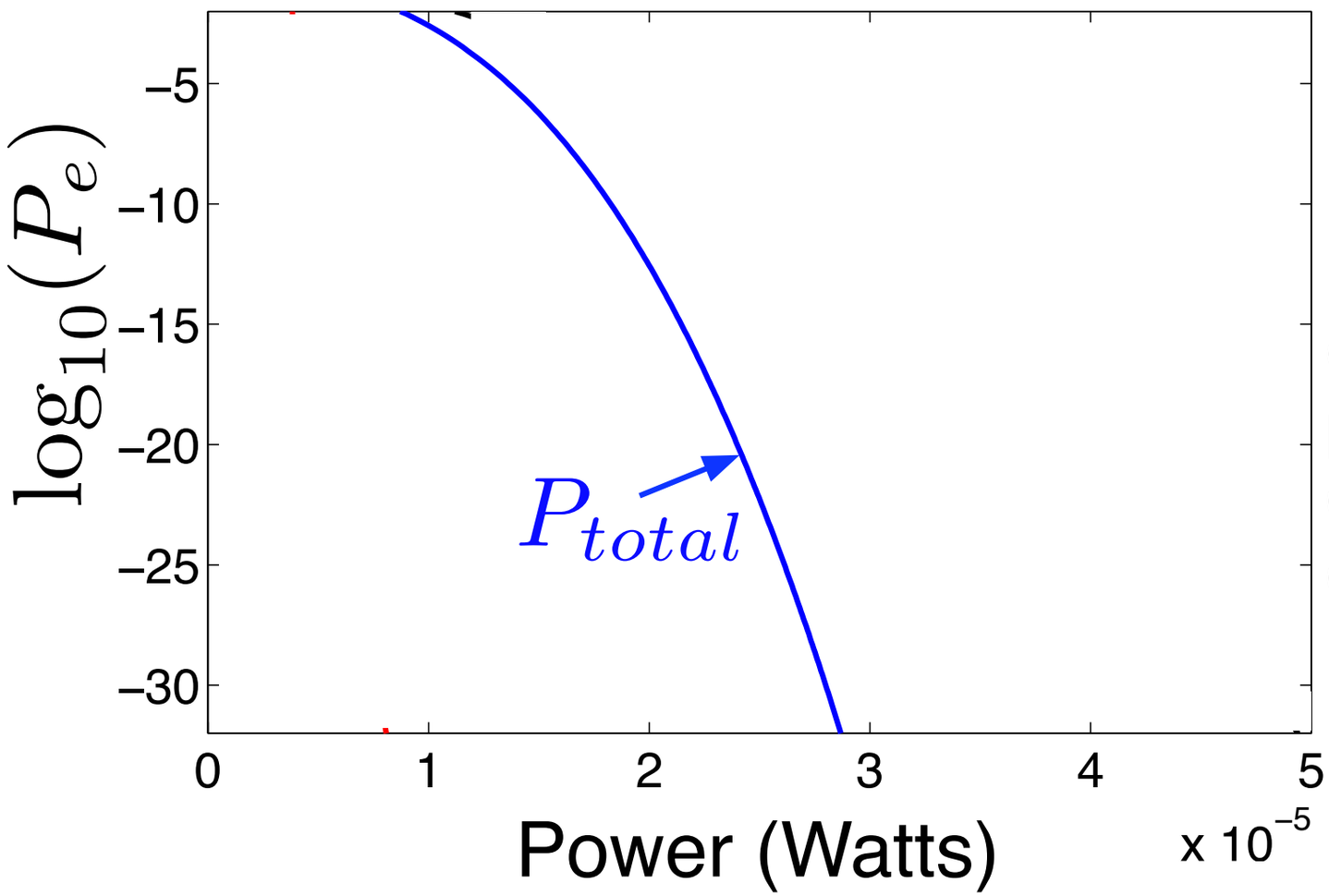
$$P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding}$$

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

Wire model:
Fundamental limits on on **total** power

$$P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding}$$

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

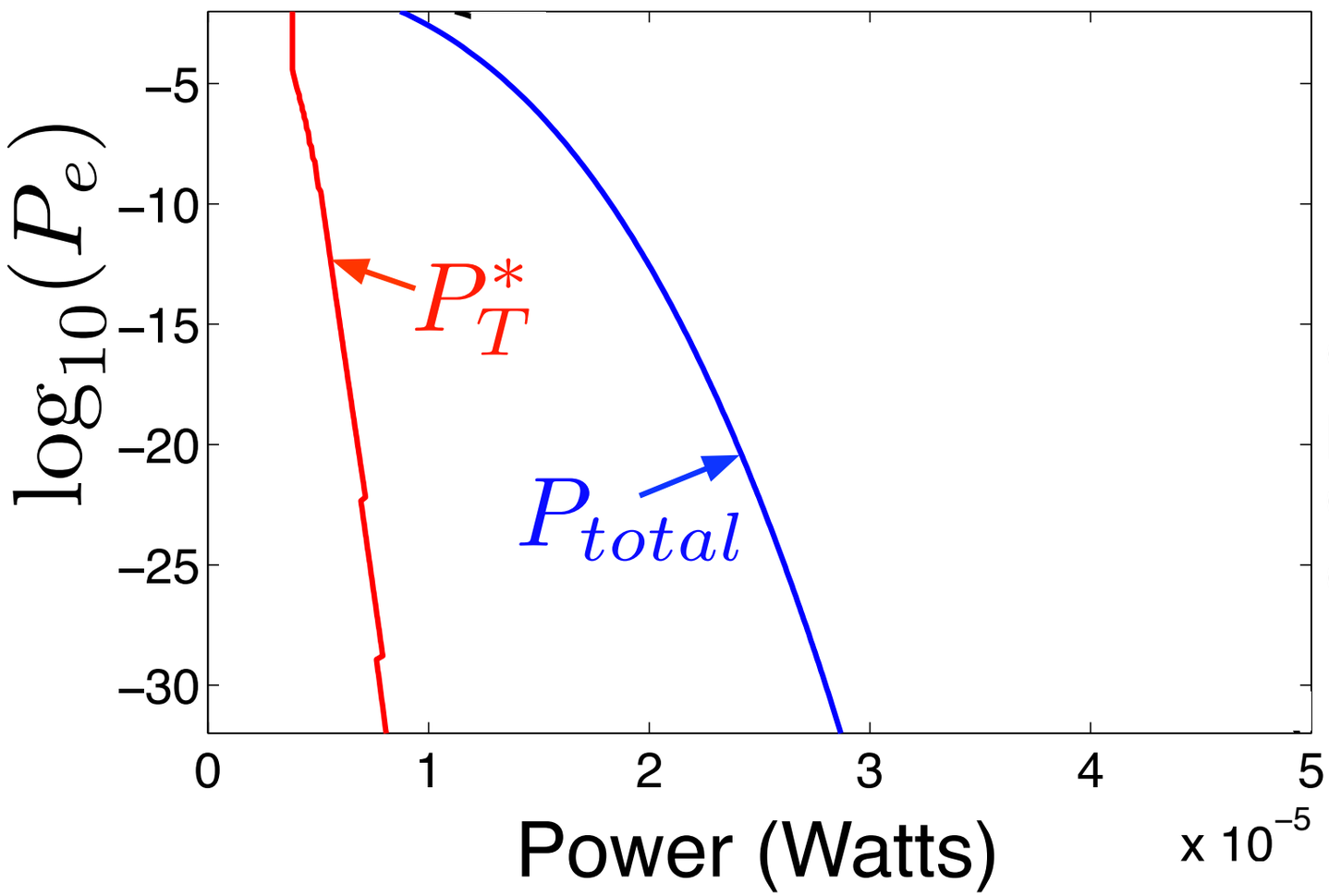


d = 0.5 m
path-loss coeff = 2
Capacitance = 1 fF/um
CMOS 90 nm tech.

Wire model:
Fundamental limits on on **total** power

$$P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding}$$

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

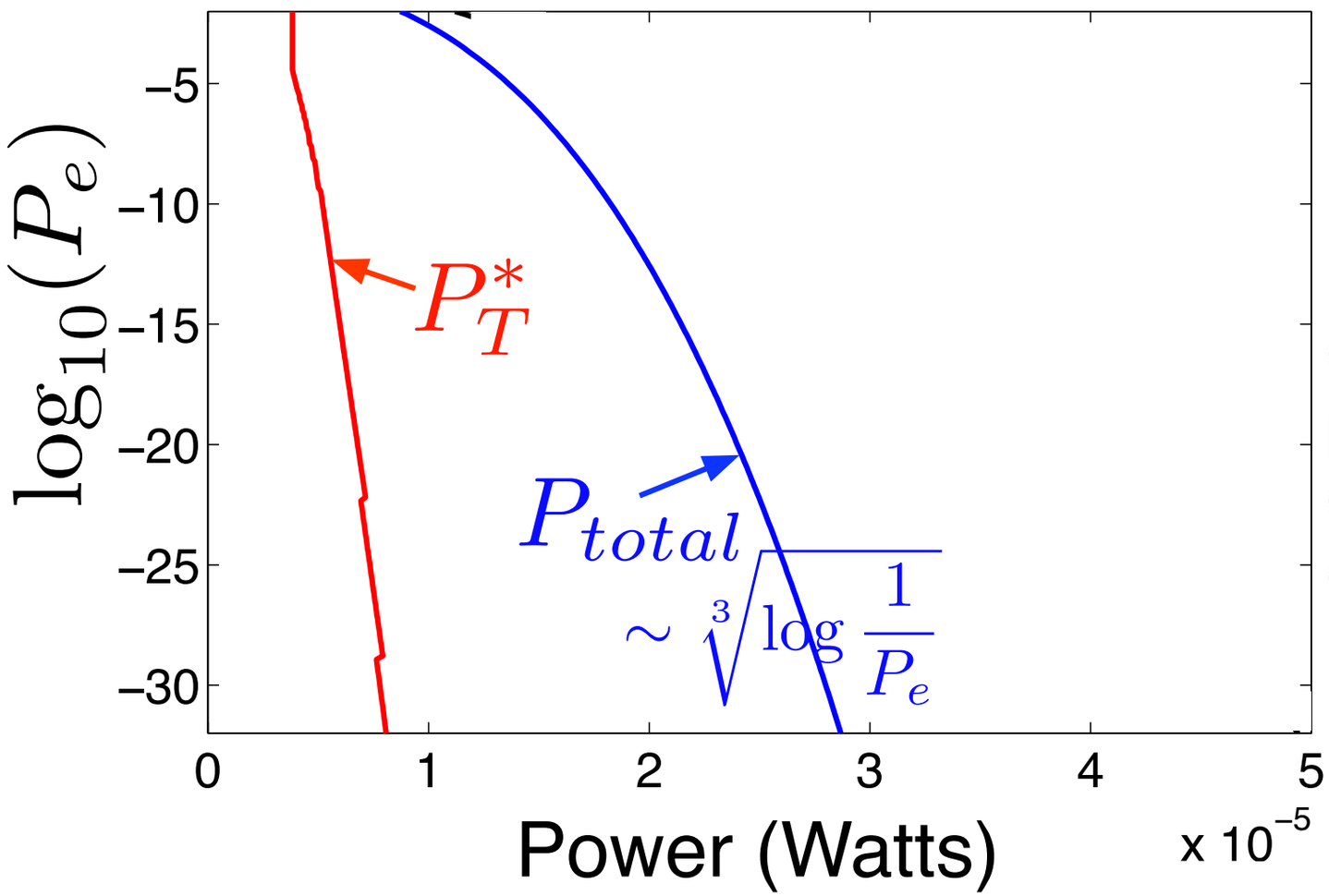


d = 0.5 m
path-loss coeff = 2
Capacitance = 1 fF/um
CMOS 90 nm tech.

Wire model:
Fundamental limits on on **total** power

$$P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding}$$

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

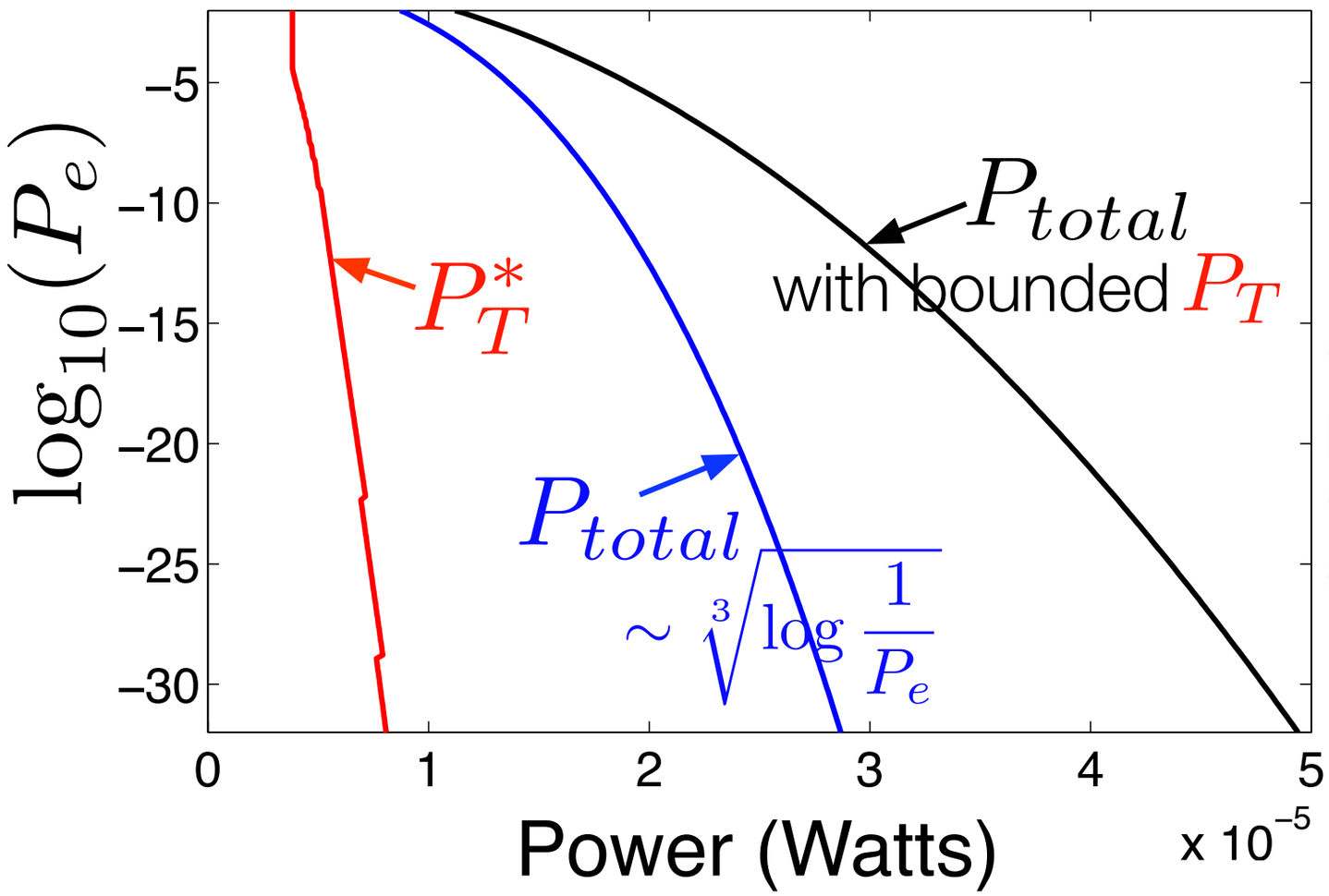


d = 0.5 m
path-loss coeff = 2
Capacitance = 1 fF/um
CMOS 90 nm tech.

Wire model:
Fundamental limits on on **total** power

$$P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding}$$

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

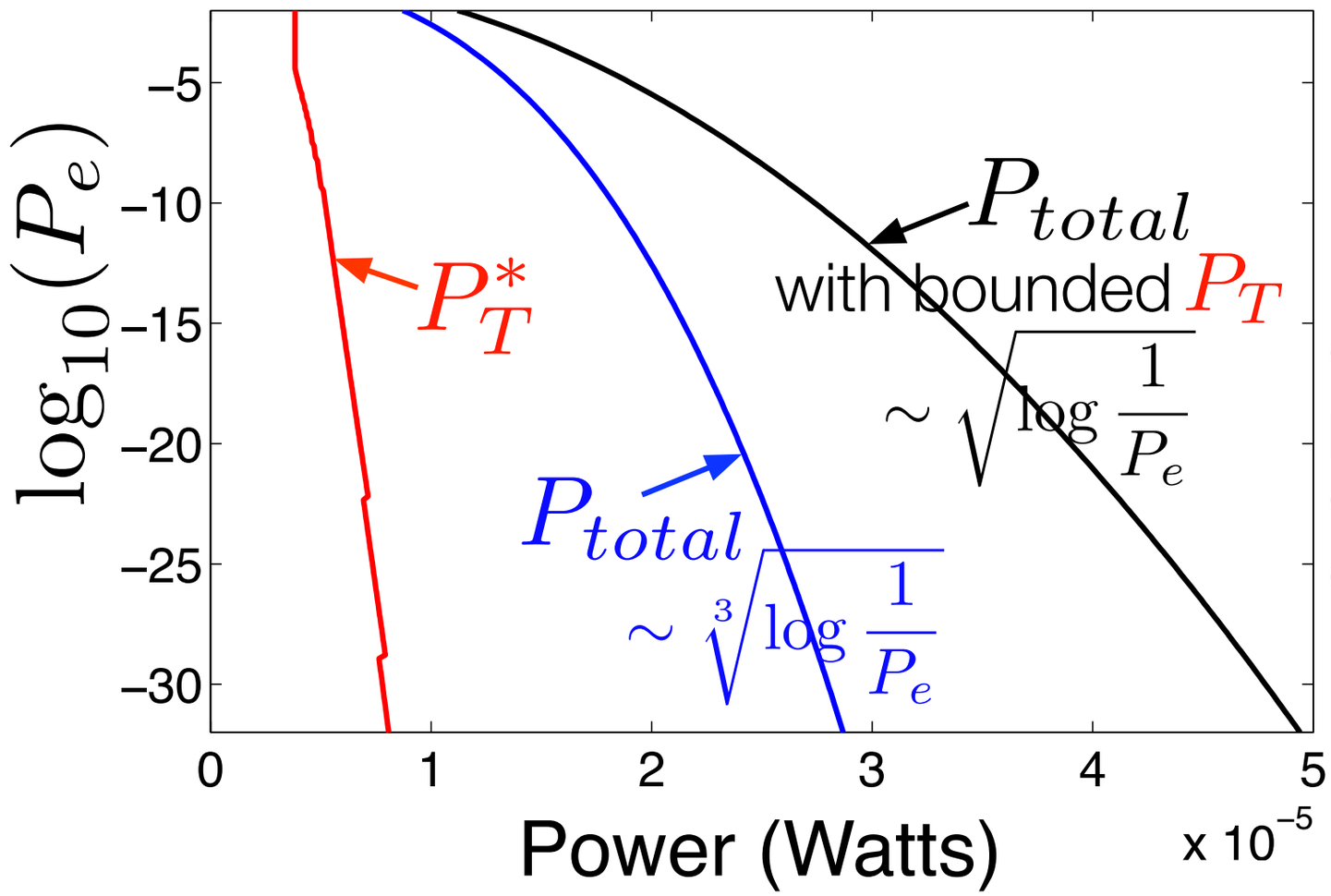


d = 0.5 m
path-loss coeff = 2
Capacitance = 1 fF/um
CMOS 90 nm tech.

Wire model:
Fundamental limits on on **total** power

$$P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding}$$

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

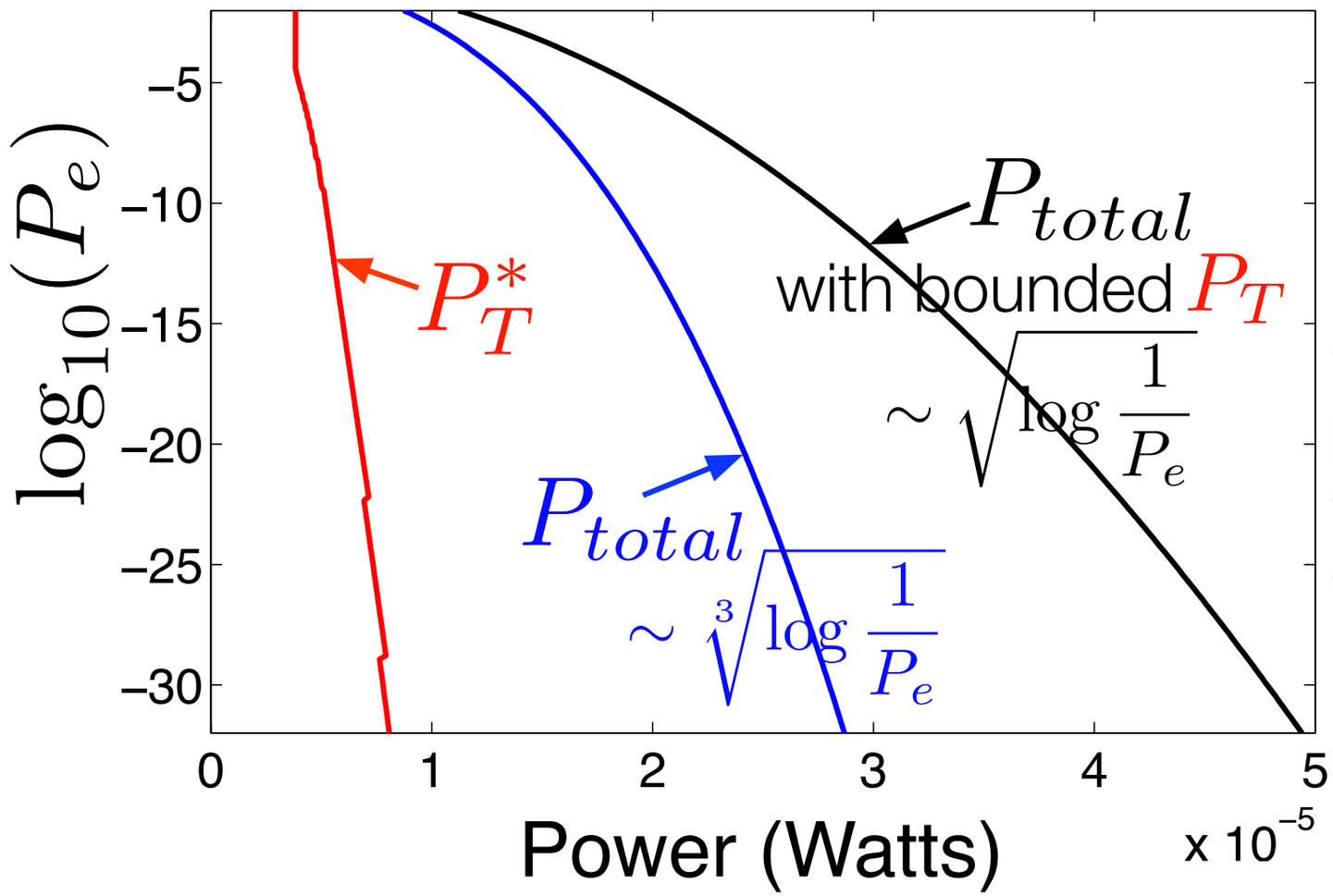


d = 0.5 m
path-loss coeff = 2
Capacitance = 1 fF/um
CMOS 90 nm tech.

Wire model:
Fundamental limits on on **total** power

$$P_{total} = \min_{P_T} P_T + P_{decoding} + P_{encoding}$$

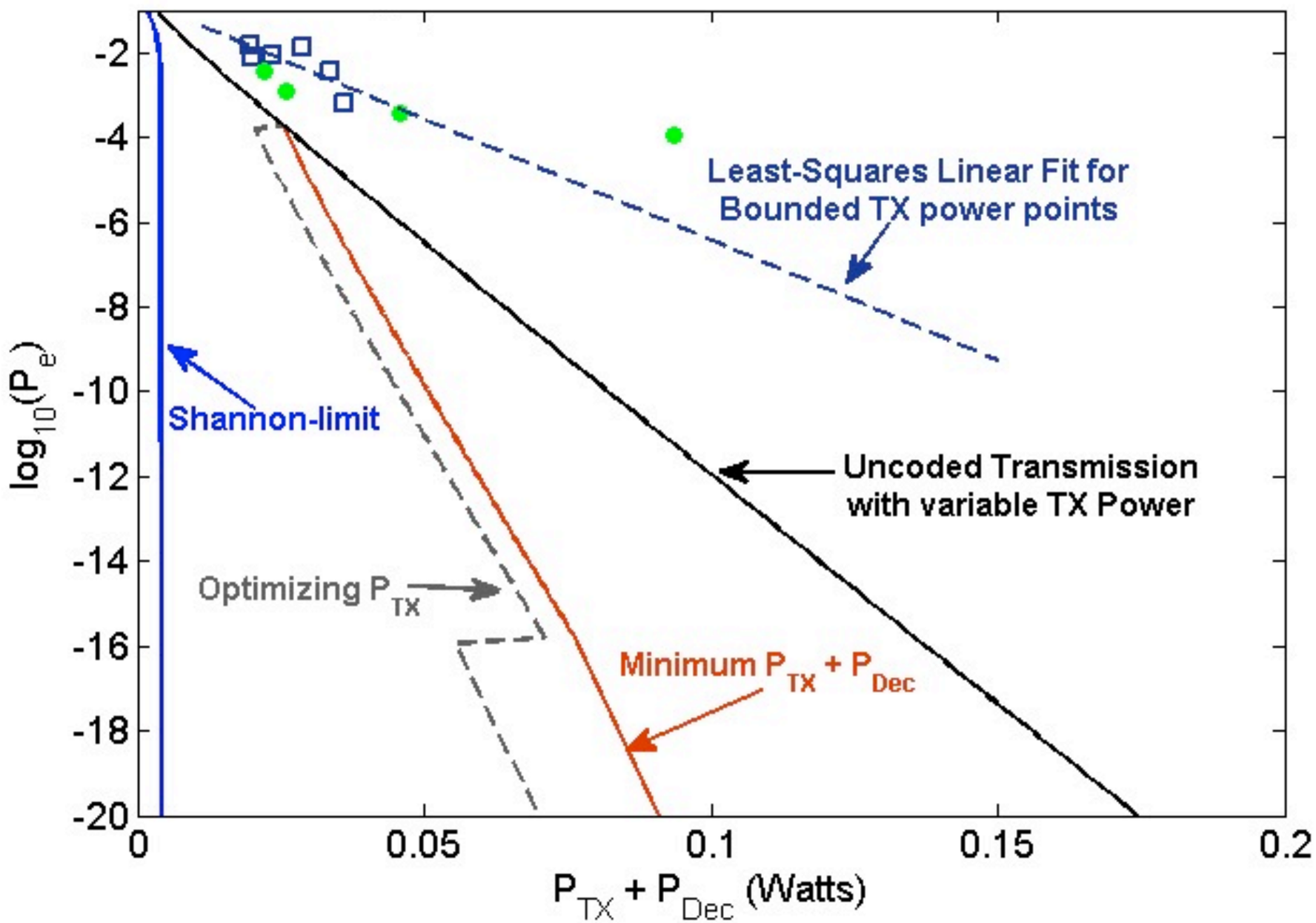
$$\gtrsim \min_{P_T} P_T + 2\beta\sqrt{\frac{\log \frac{1}{P_e}}{P_T}}$$



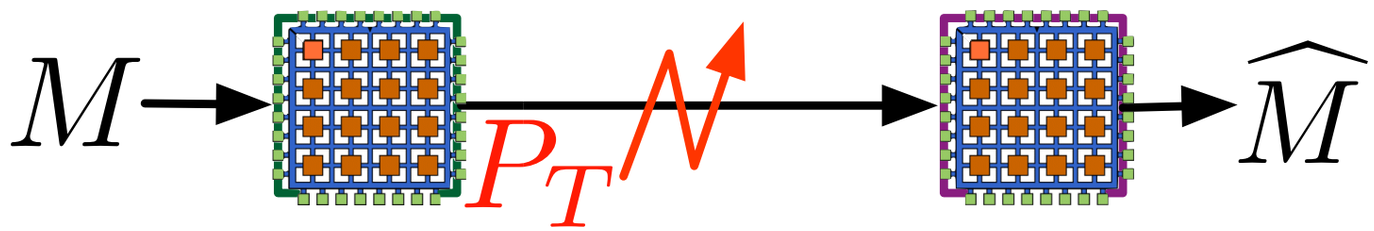
d = 0.5 m
path-loss coeff = 2
Capacitance = 1 fF/um
CMOS 90 nm tech.

Moral: Bounded P_T is suboptimal

Suboptimality of bounded transmit power: Experimental results



Information flows in circuits: Summary



Information needs to flow at the encoder and decoder
(mixing) (de-mixing)

<u>Fundamental limits:</u>

Traditional: Operate close to capacity

Total power perspective: Node model: stay away from capacity

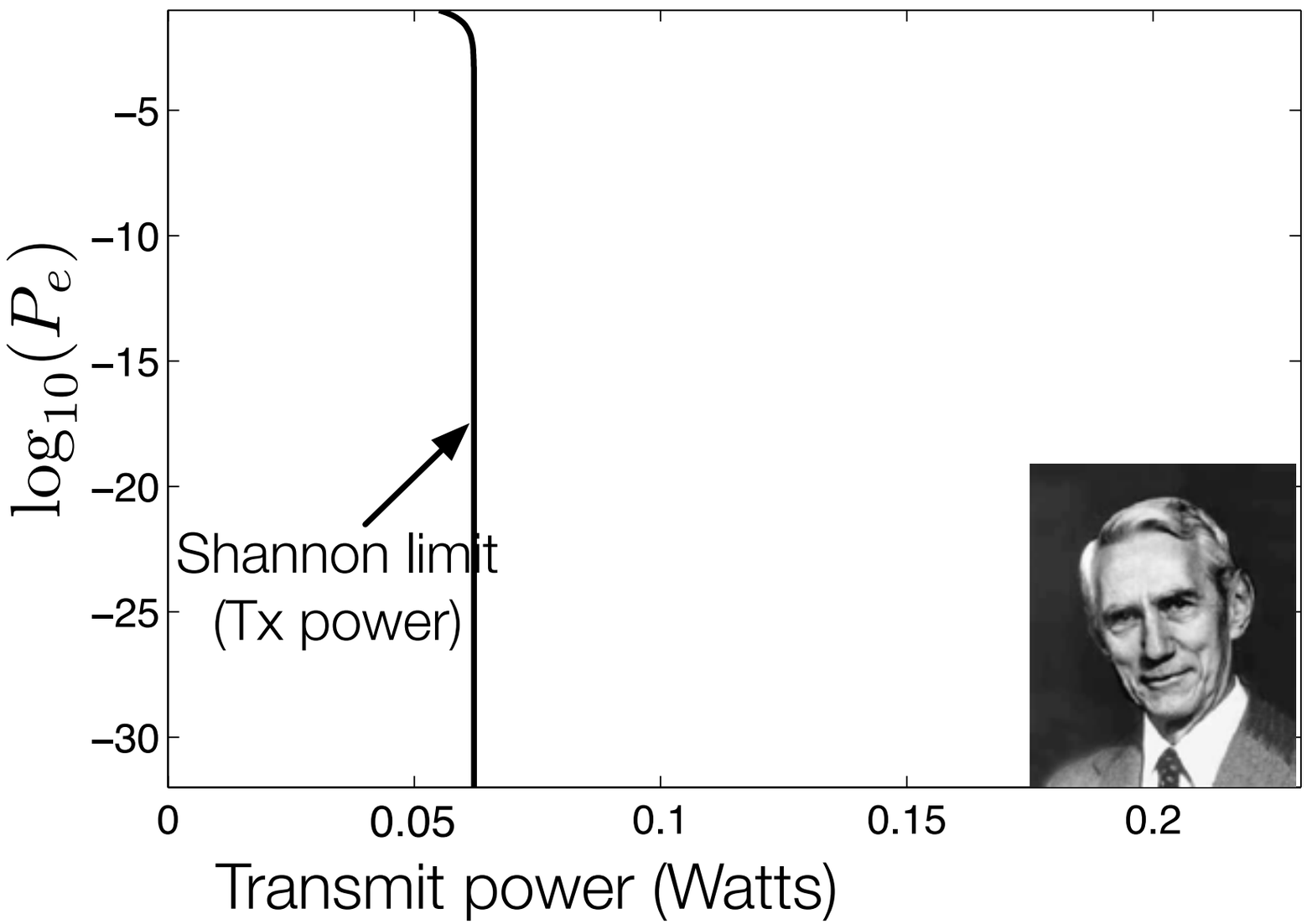
Wire model: use unbounded P_T

No assumptions on the code structure

Looking forward:
"Greenest" codes for wireless?

Grand goal:

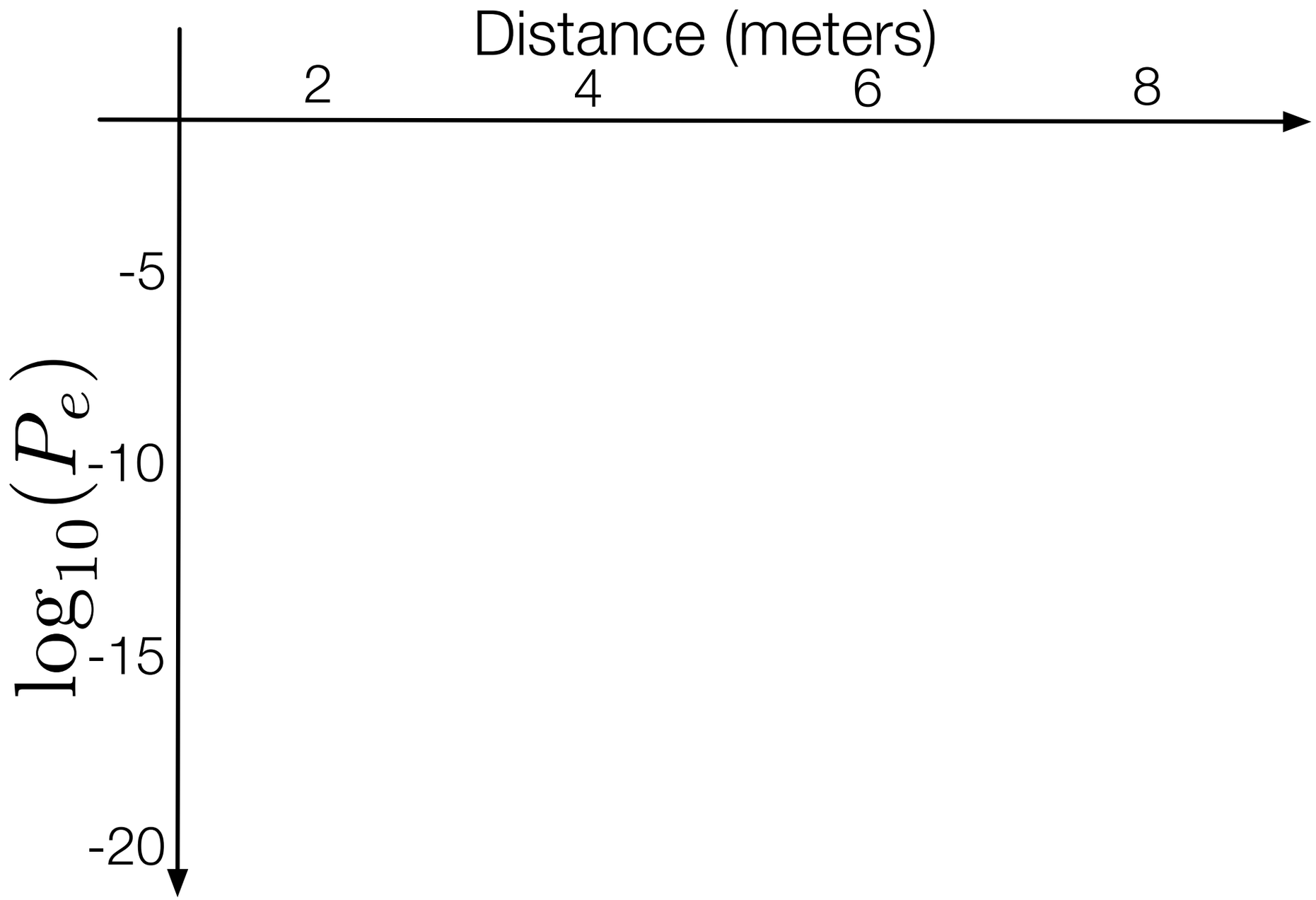
Code/encoder/decoder triples within bounded dBs of the optimal.



Looking forward:
"Greenest" codes for wireless?

Grand goal:

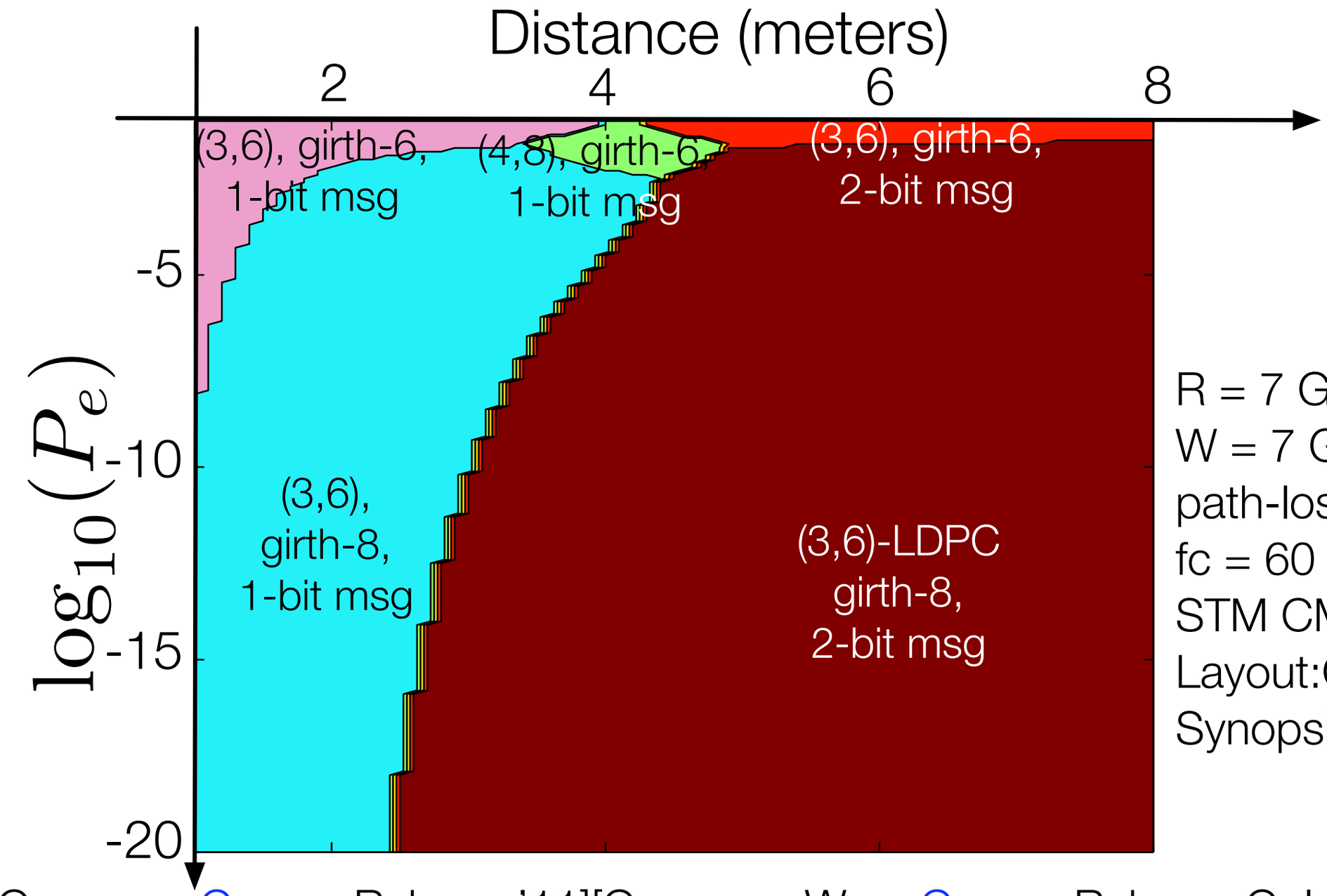
Code/encoder/decoder triples within bounded dBs of the optimal.



Looking forward:
"Greenest" codes for wireless?

Grand goal:

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



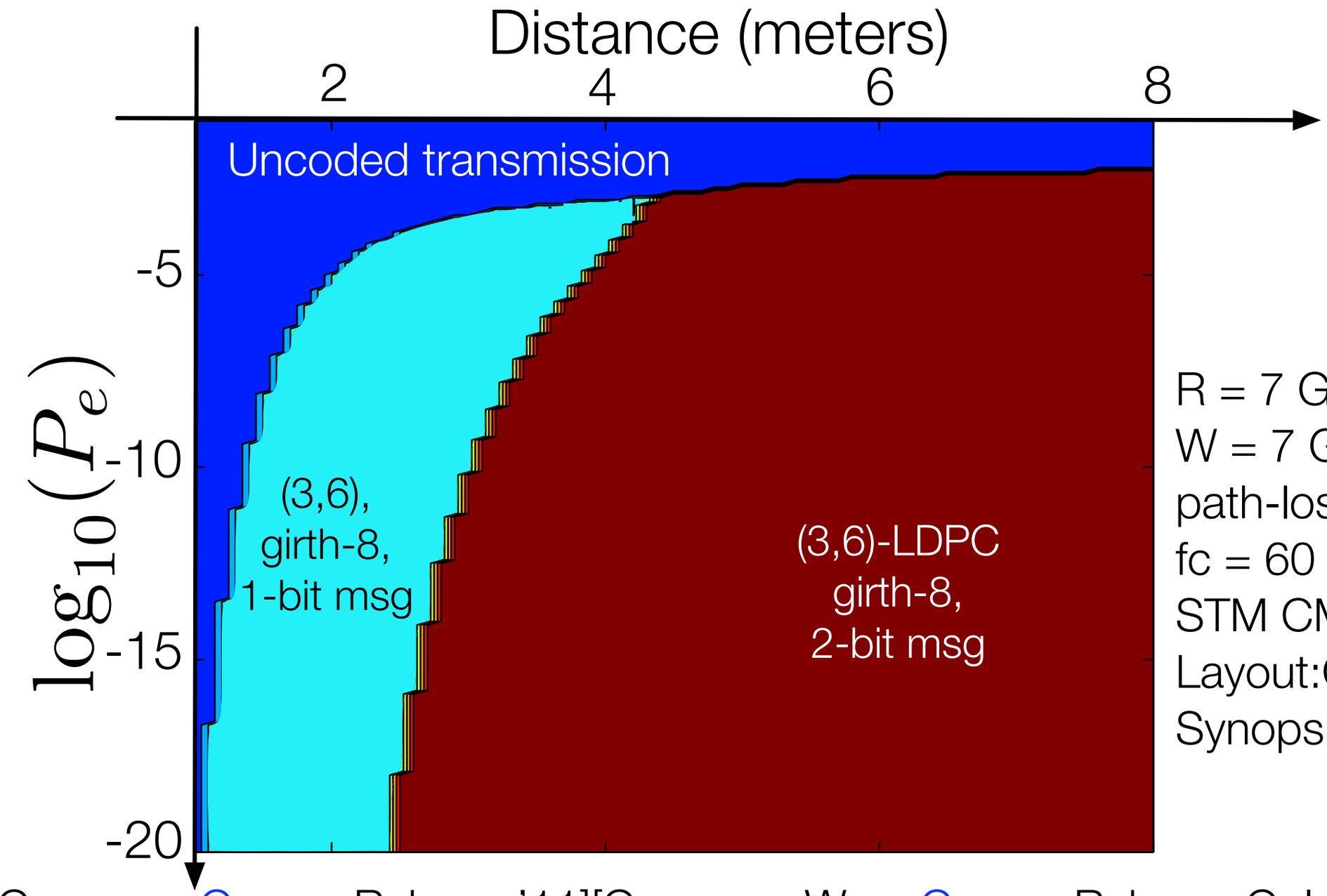
R = 7 Gbps
W = 7 GHz
path-loss coeff = 3
fc = 60 GHz
STM CMOS 90 nm
Layout:Cadence Enctr
Synopsis HSPICE

[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.



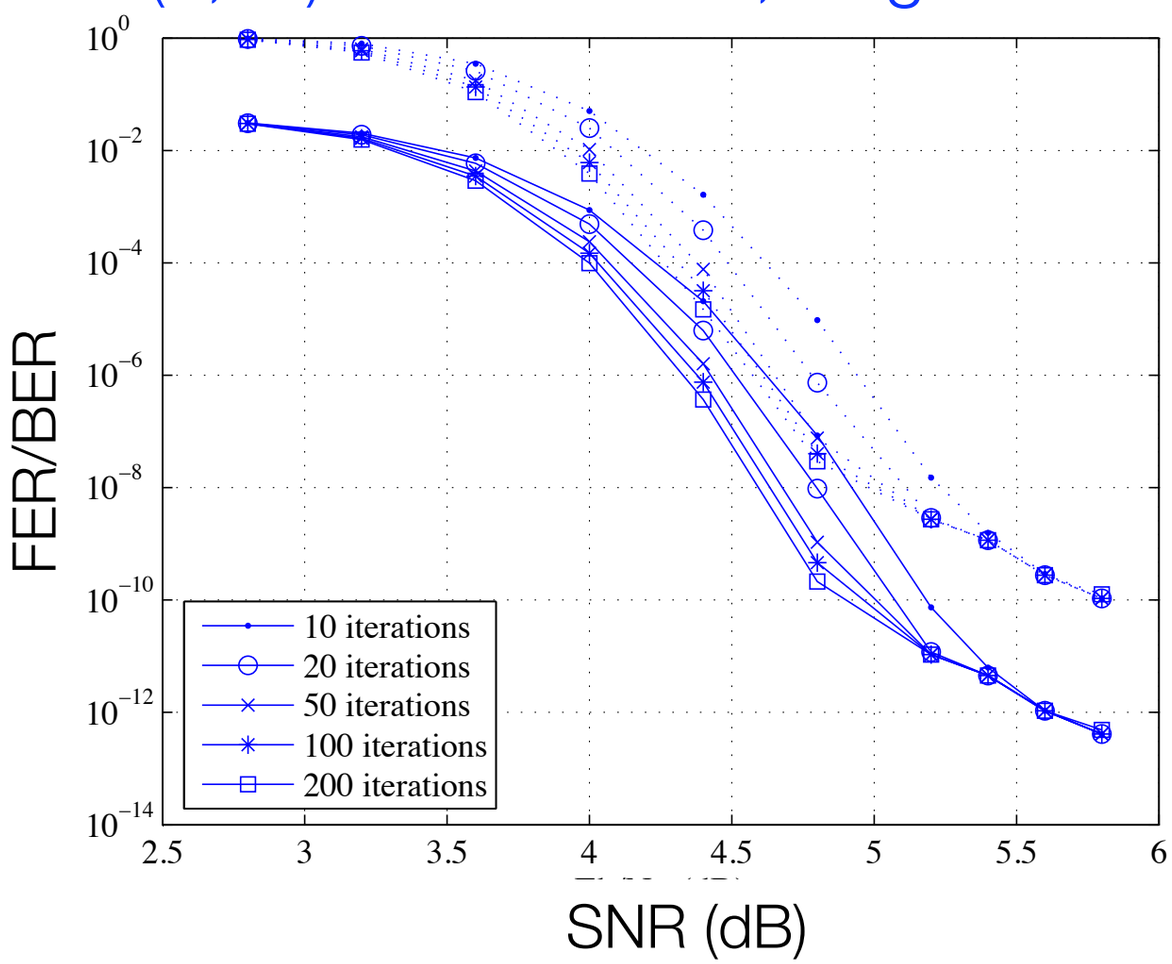
R = 7 Gbps
W = 7 GHz
path-loss coeff = 3
fc = 60 GHz
STM CMOS 90 nm
Layout:Cadence Enctr
Synopsis HSPICE

[Ganesan, Grover, Rabaey '11][Ganesan, Wen, Grover, Rabaey, Goldsmith '12]

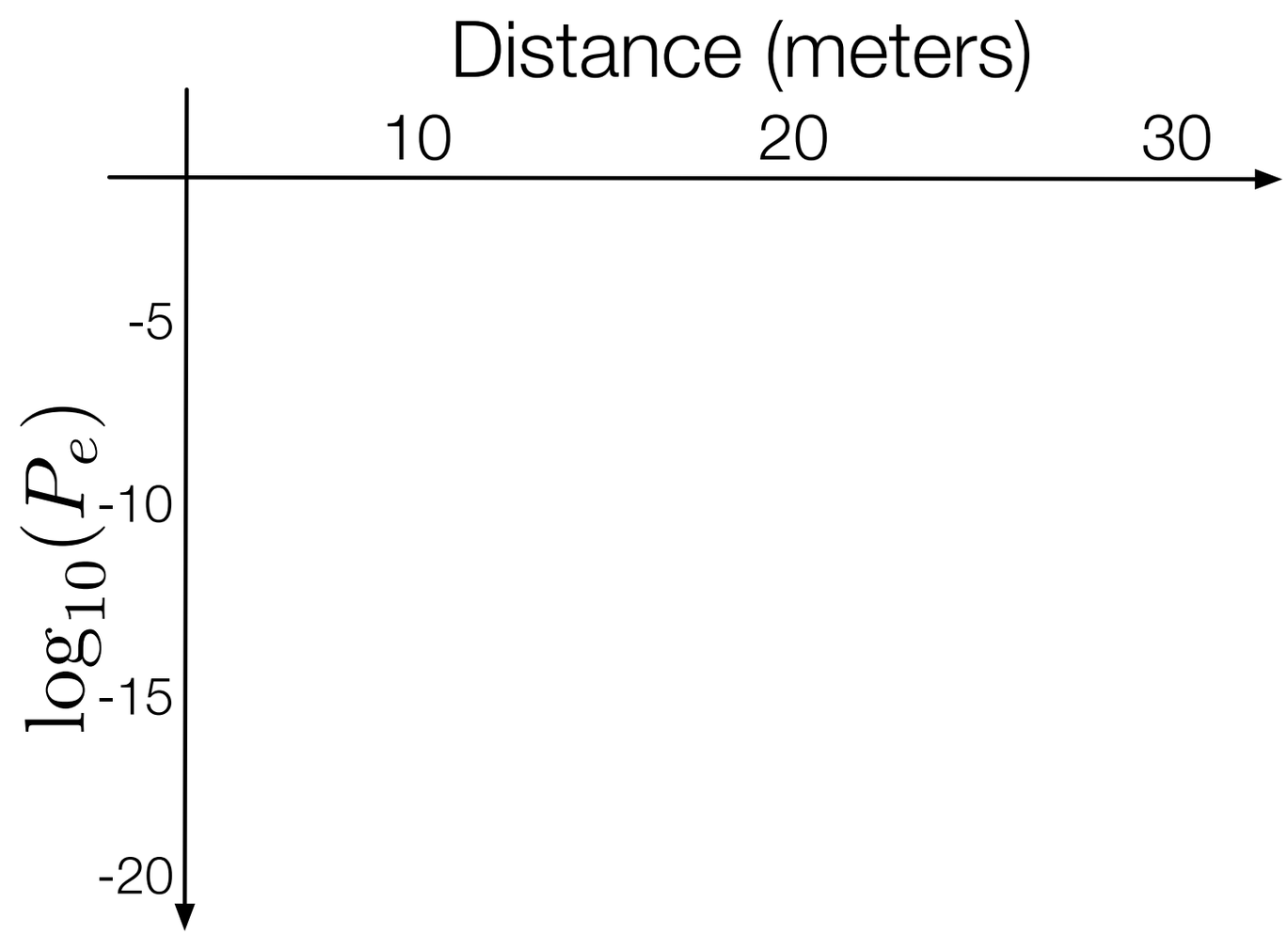
Looking forward: "Green" data-center communication

Traditional view

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



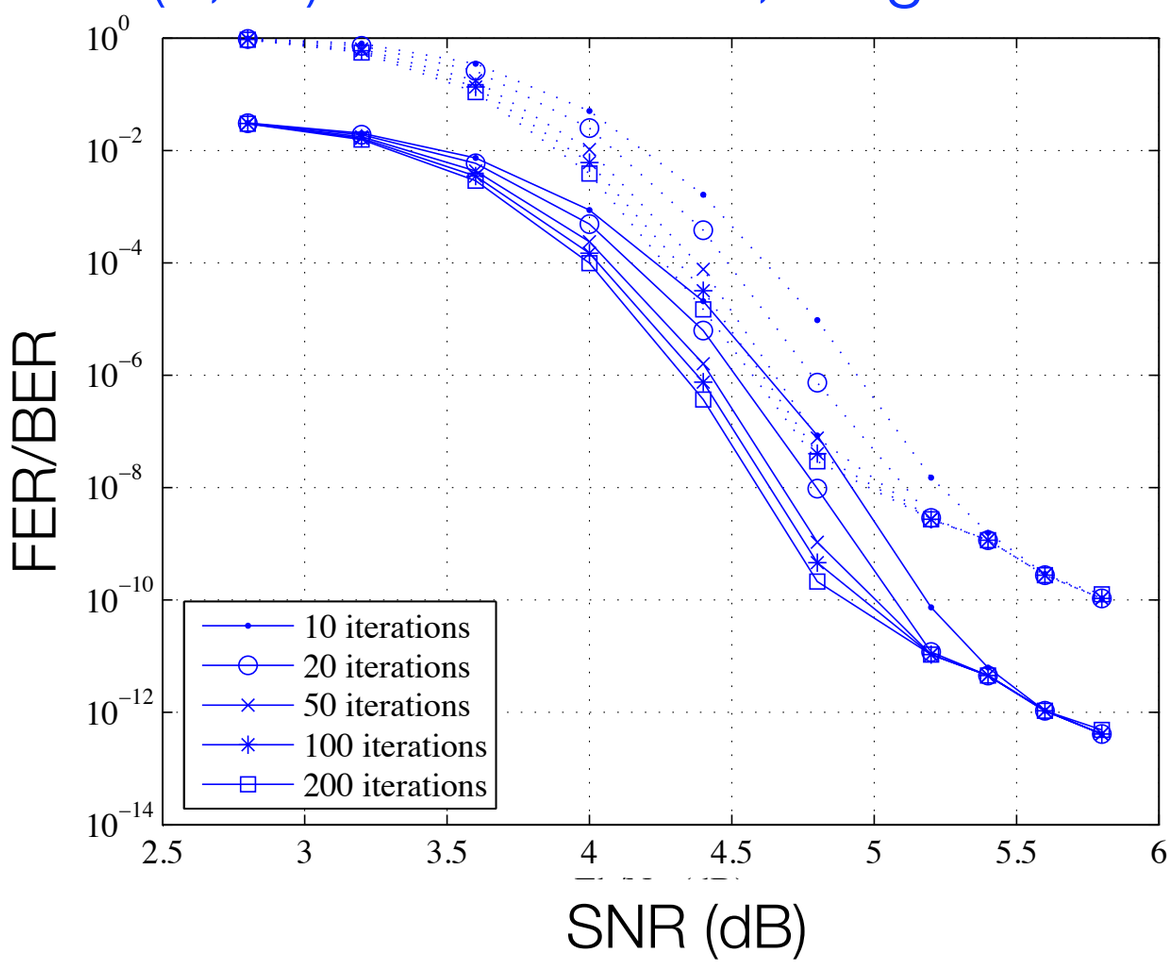
Total-power perspective



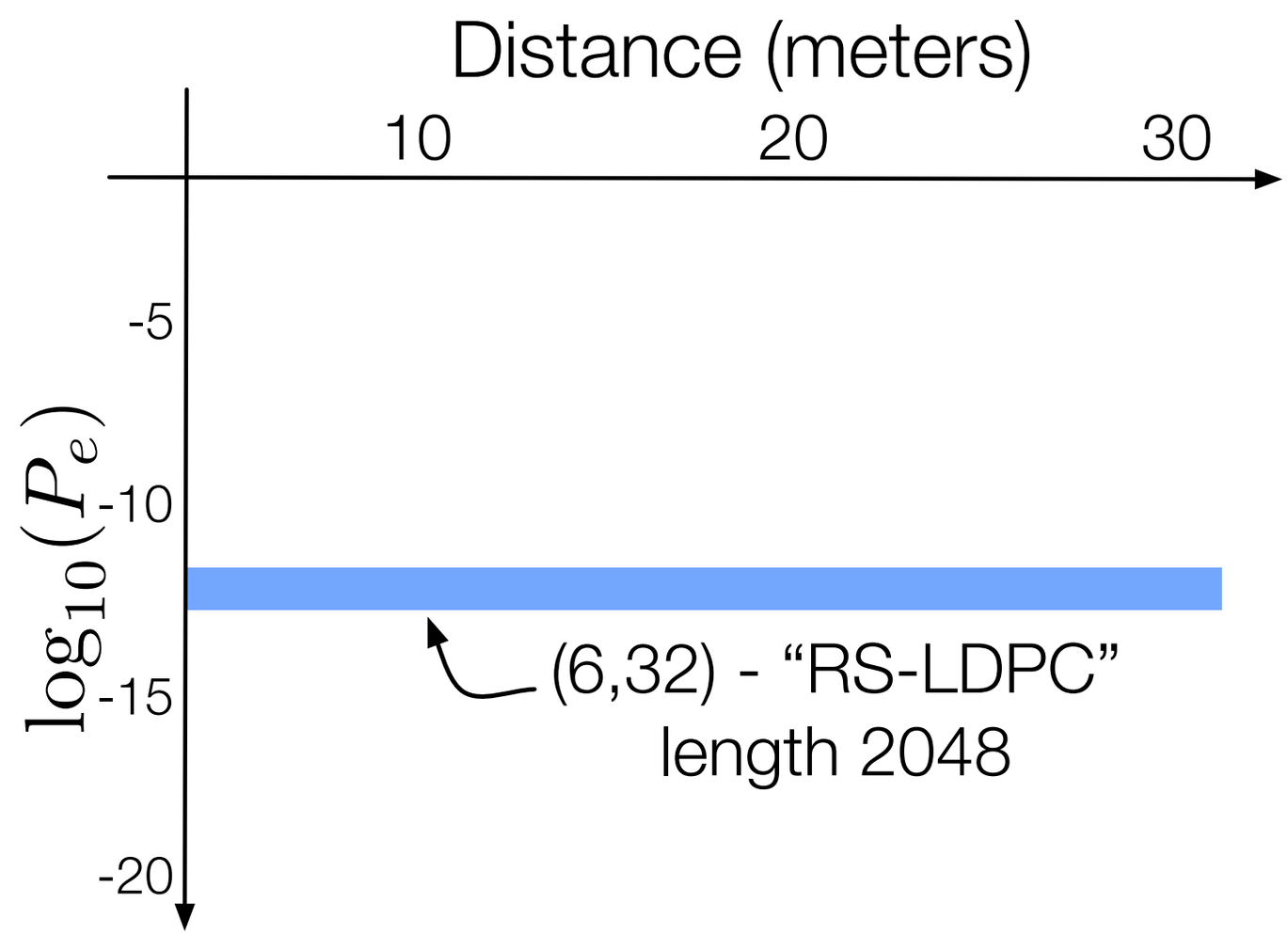
Looking forward: "Green" data-center communication

Traditional view

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



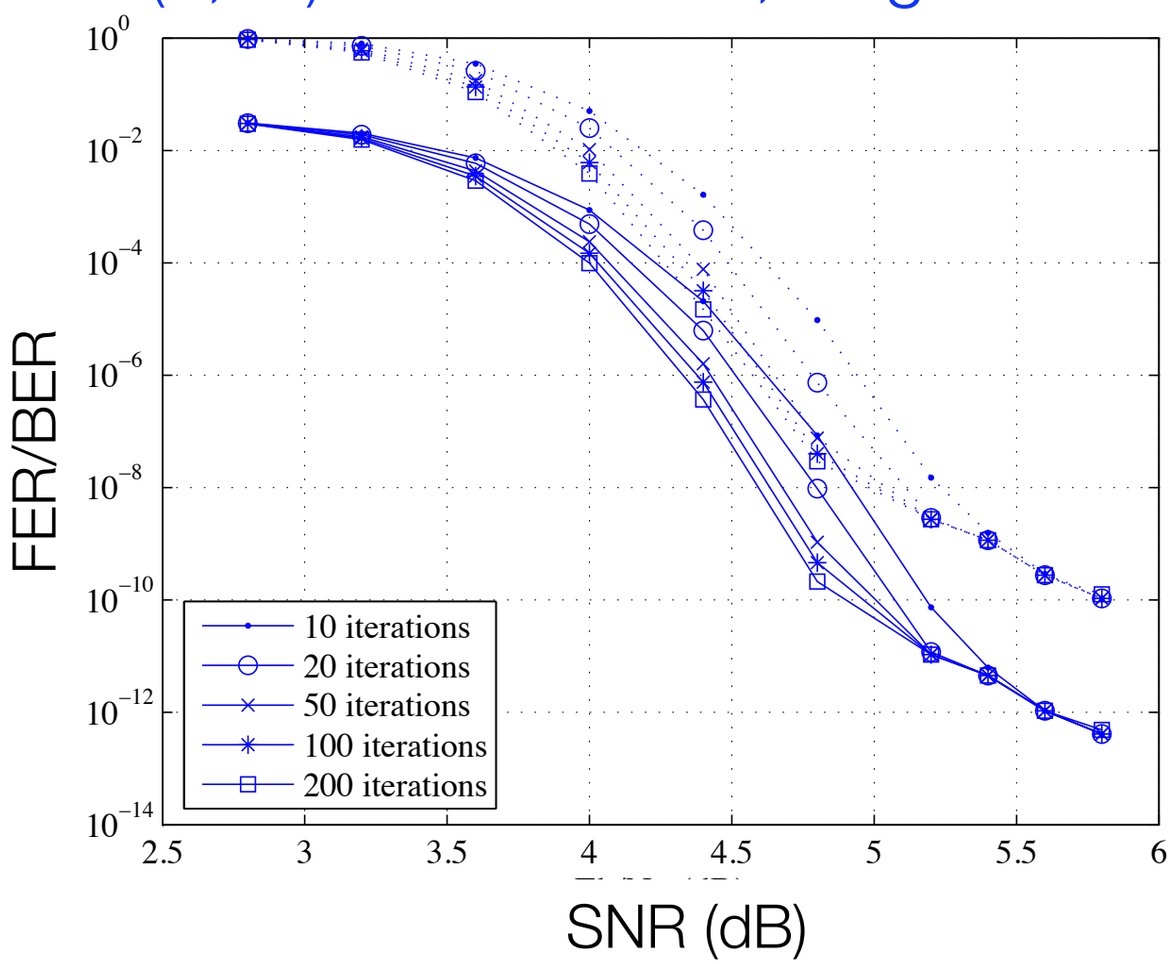
Total-power perspective



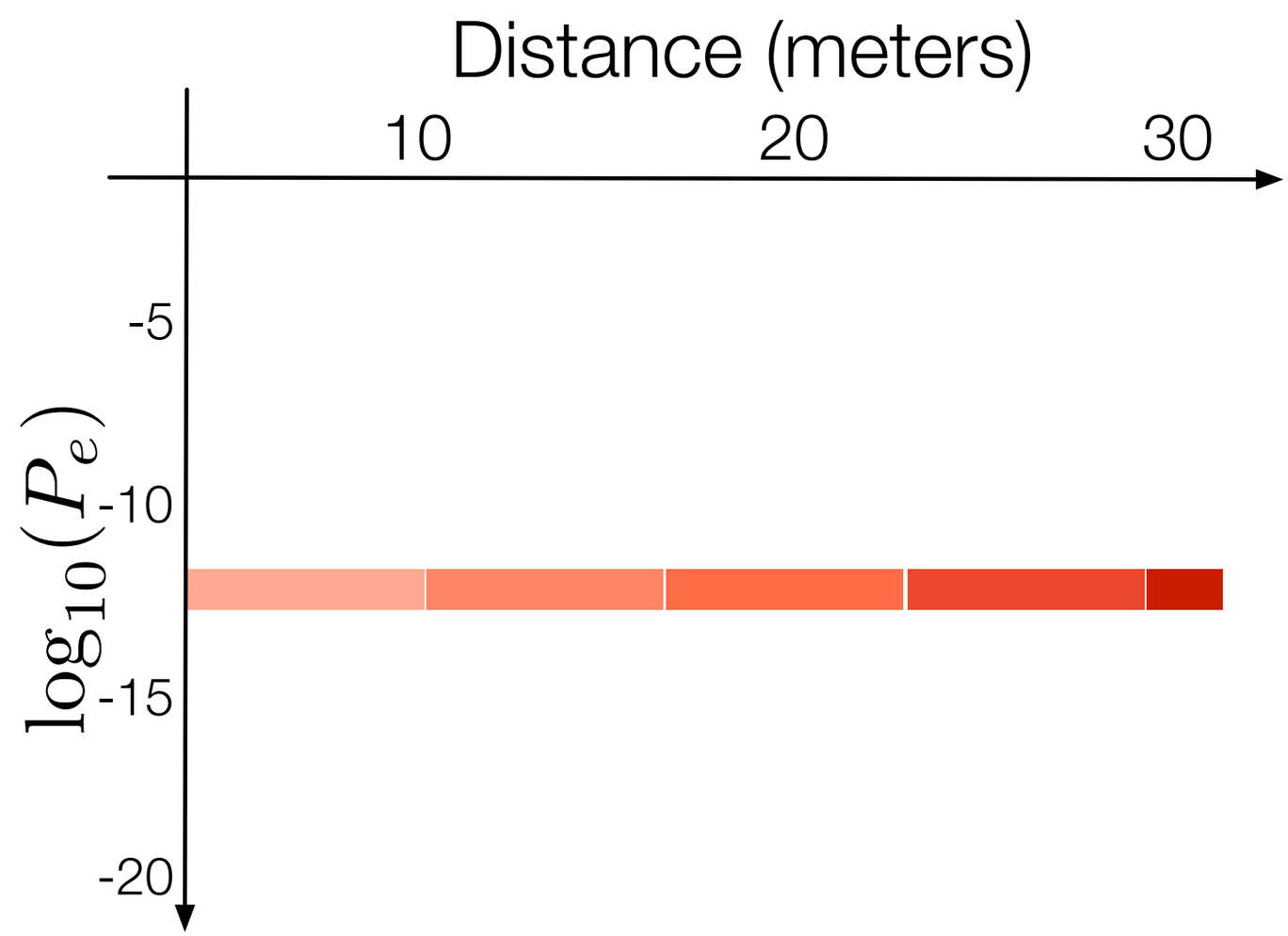
Looking forward: "Green" data-center communication

Traditional view

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



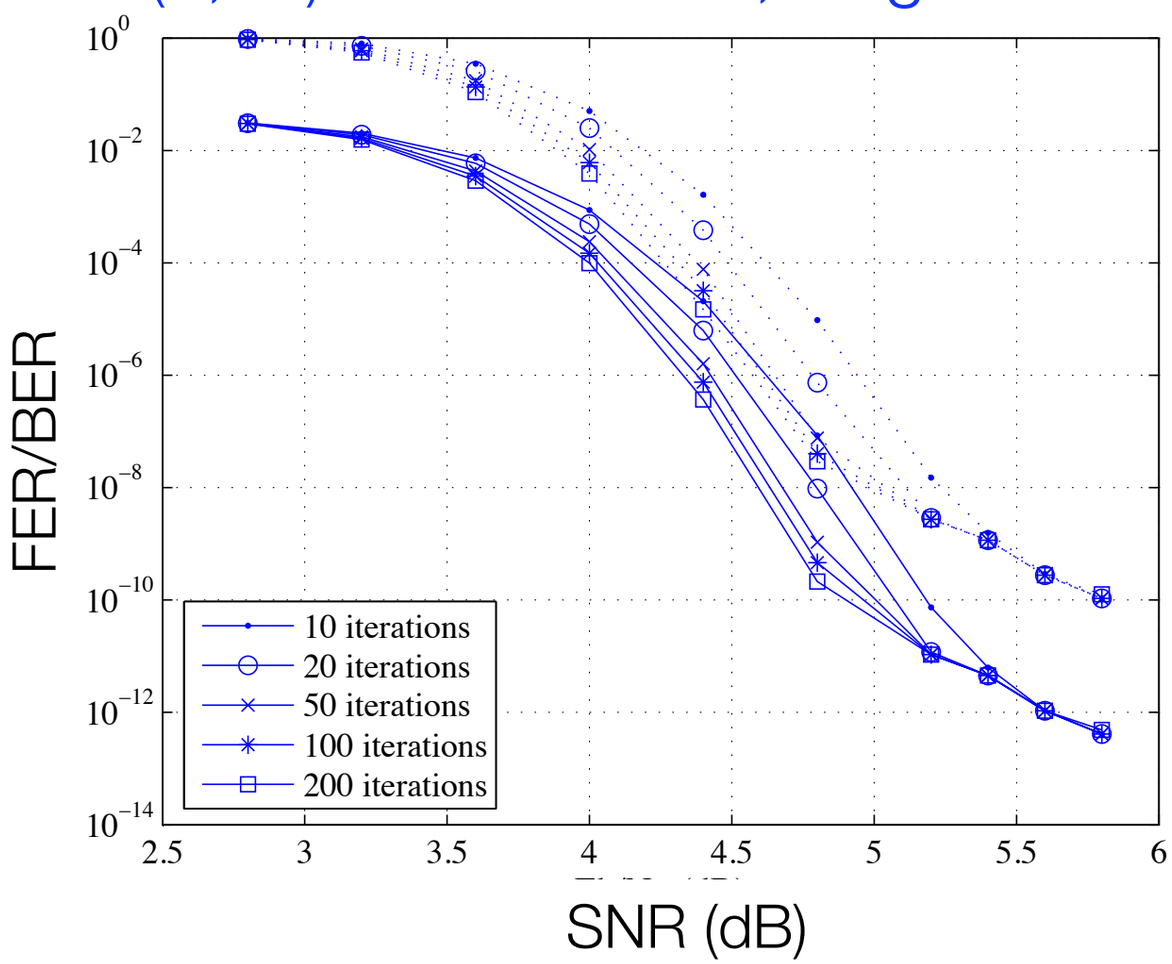
Total-power perspective



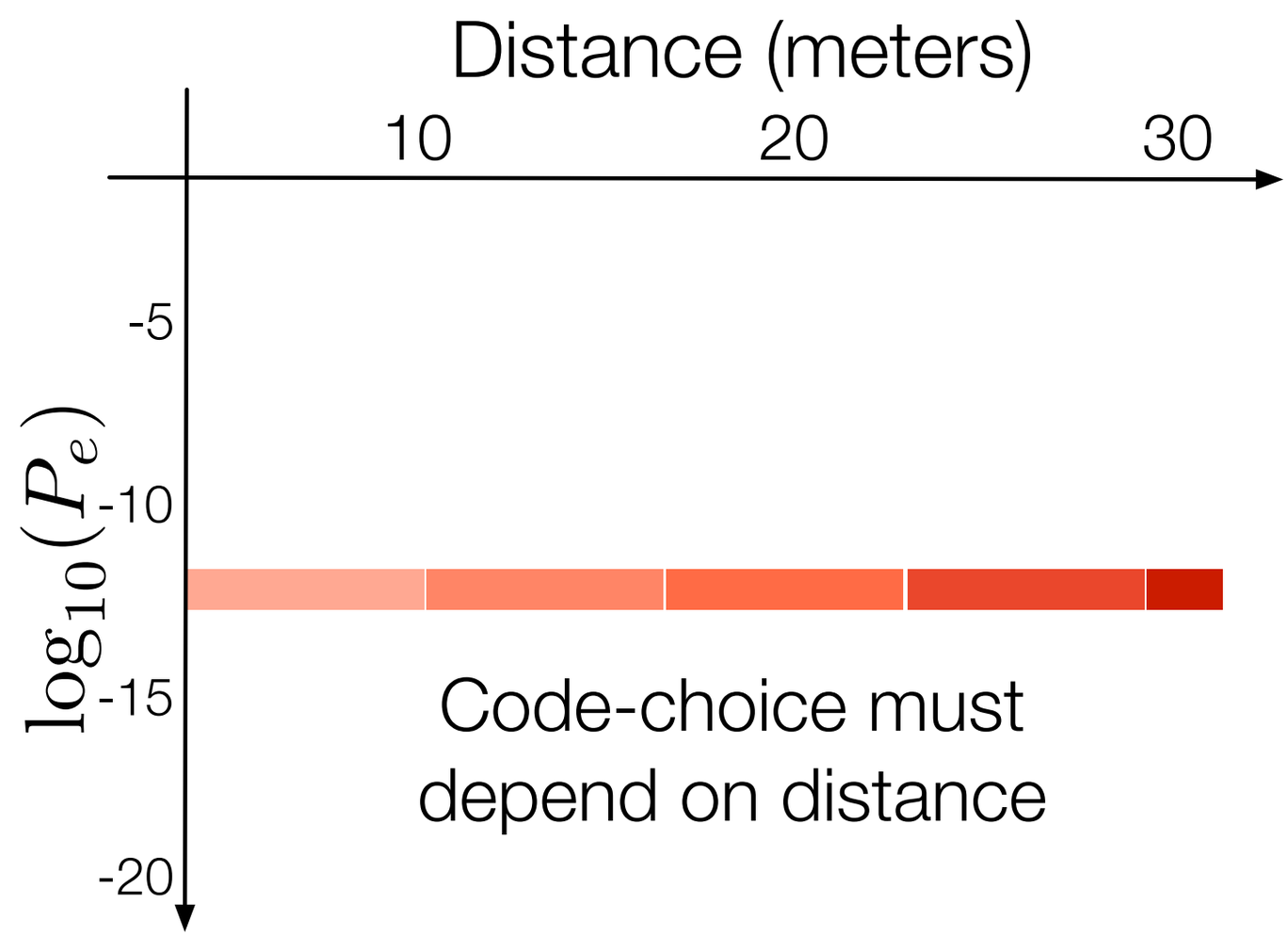
Looking forward: "Green" data-center communication

Traditional view

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



Total-power perspective



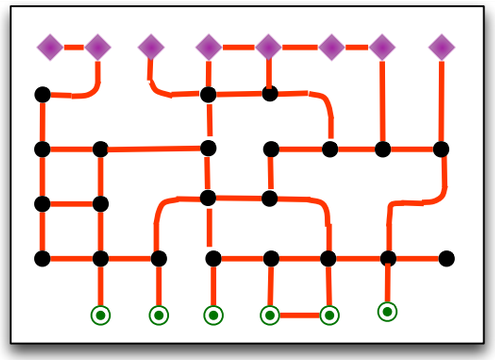
Code-choice must depend on distance

Looking forward:
Energy costs of communication

Looking forward:
Energy costs of communication

Other models of computing

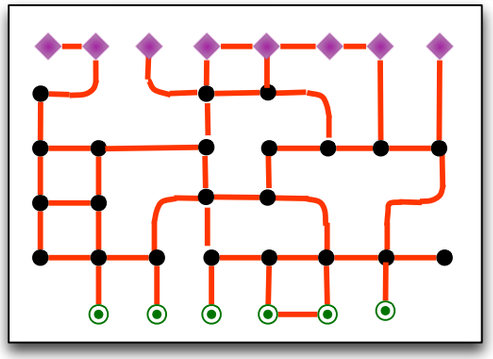
- e.g. nodes/wires can sleep



Looking forward:
Energy costs of communication

Other models of computing

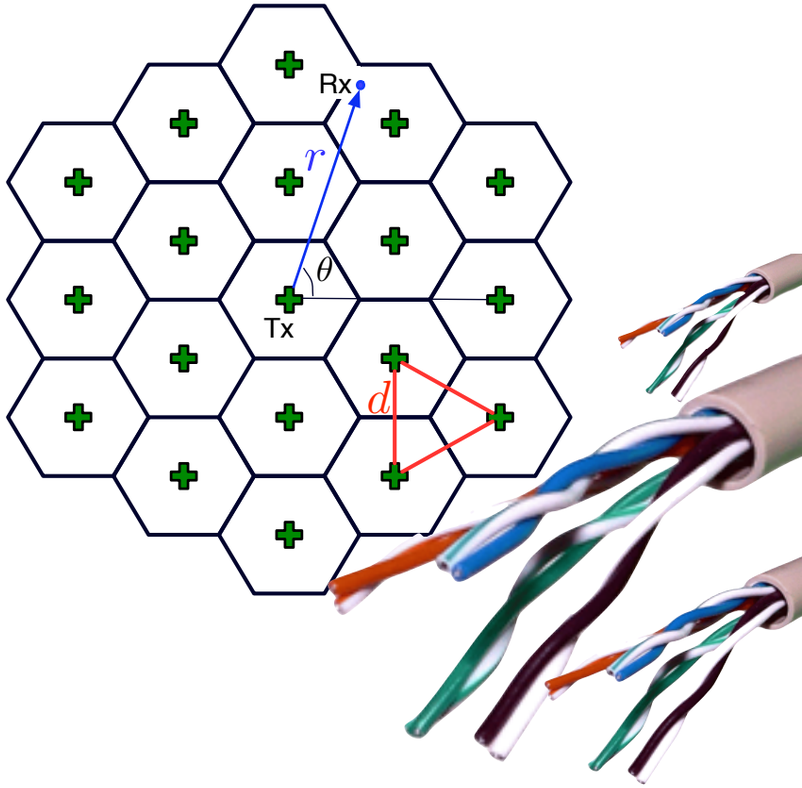
- e.g. nodes/wires can sleep



Multiple simultaneously transmitting links

- Code-choice must depend on distance, density

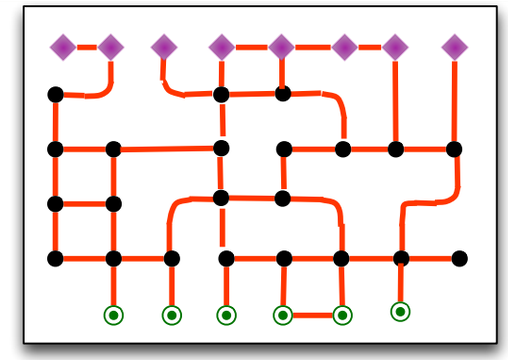
[Grover, Woyach, Sahai JSAC '11]



Looking forward:
Energy costs of communication

Other models of computing

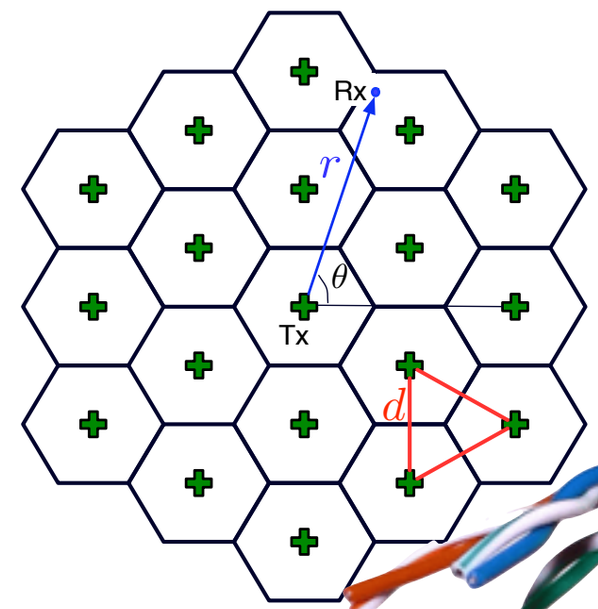
- e.g. nodes/wires can sleep

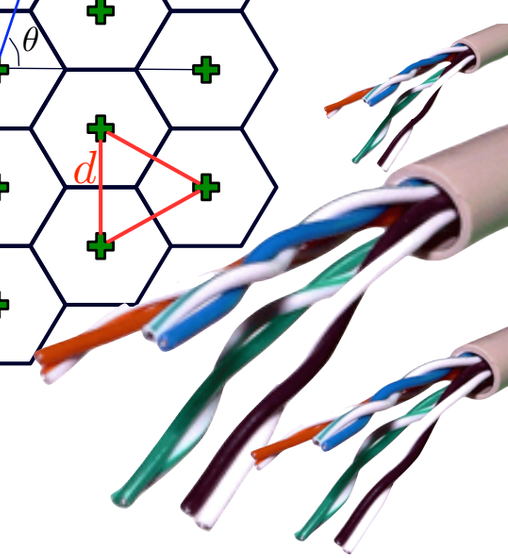


Multiple simultaneously transmitting links

- Code-choice must depend on distance, density

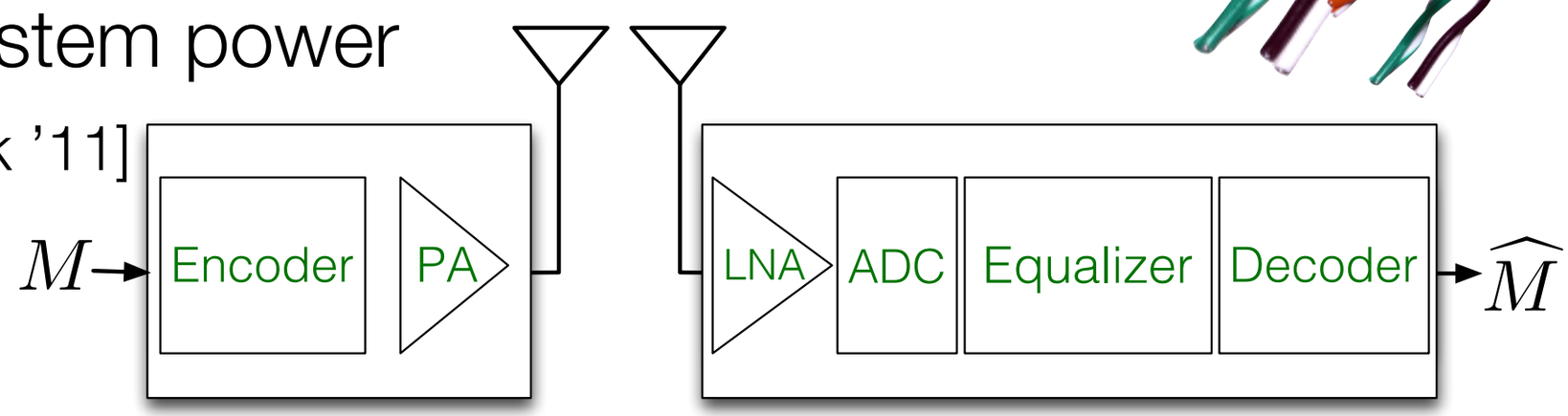
[Grover, Woyach, Sahai JSAC '11]





Other components of system power

- Equalizer [Grover, Sahai, Park '11]

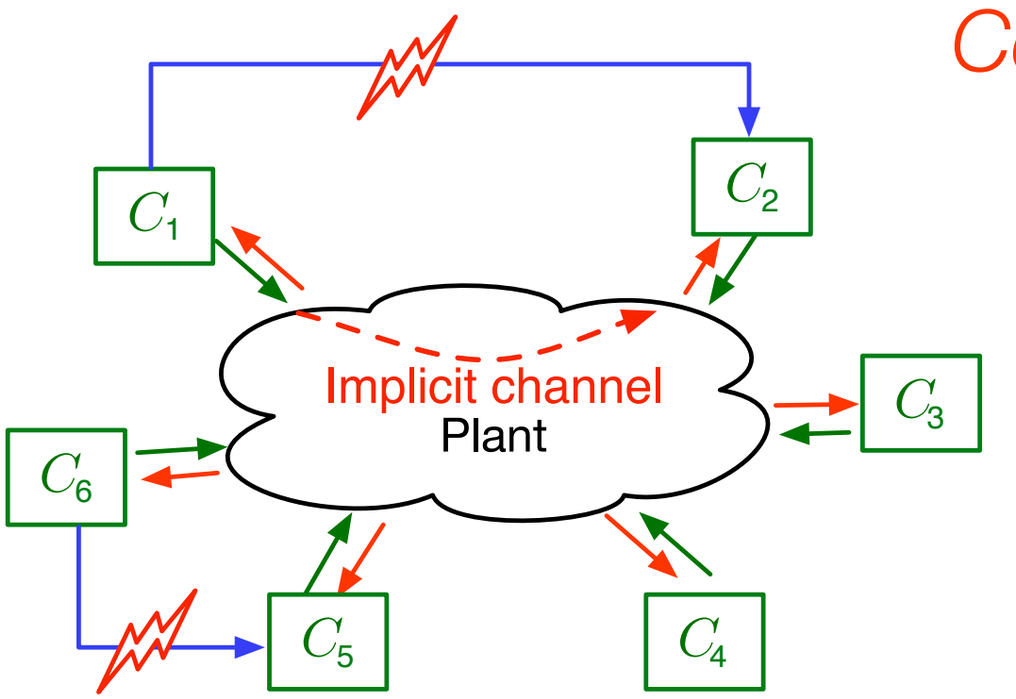


Talk outline

1) Information flows in circuits

Communicating and *processing* information

2) Information flows in physical systems



Communicating and *Using* information

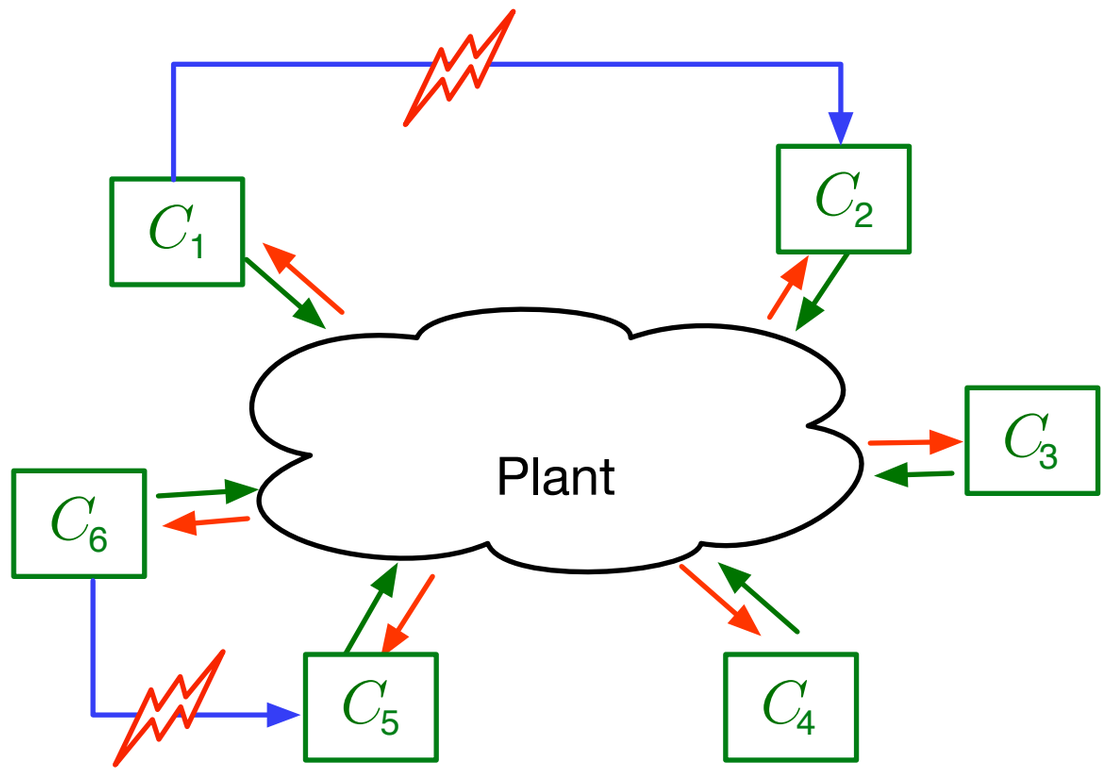
"Implicit" communication

Witsenhausen's counterexample

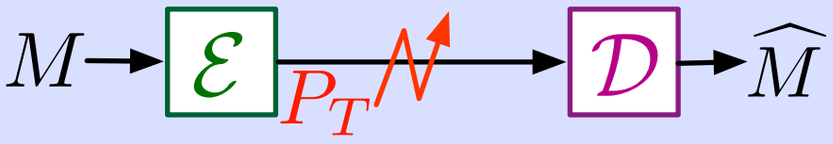
- approximately optimal solutions

Beyond the counterexample

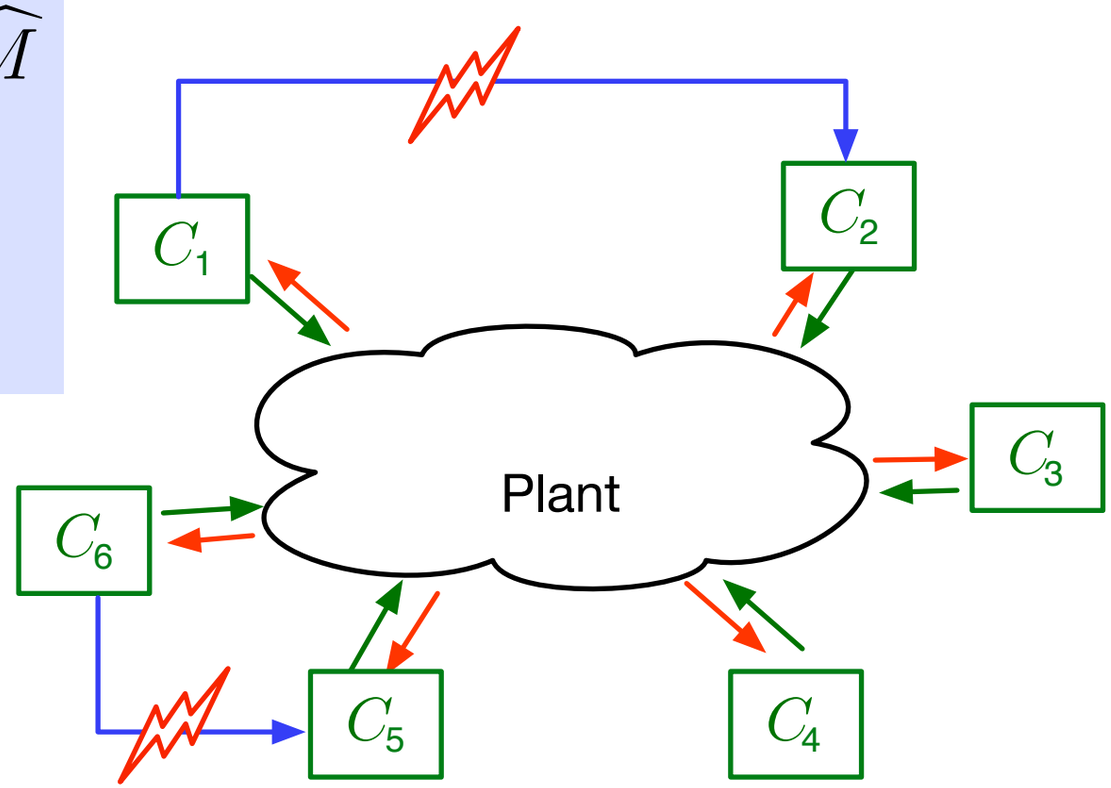
Looking forward:
Implicit information flows in control



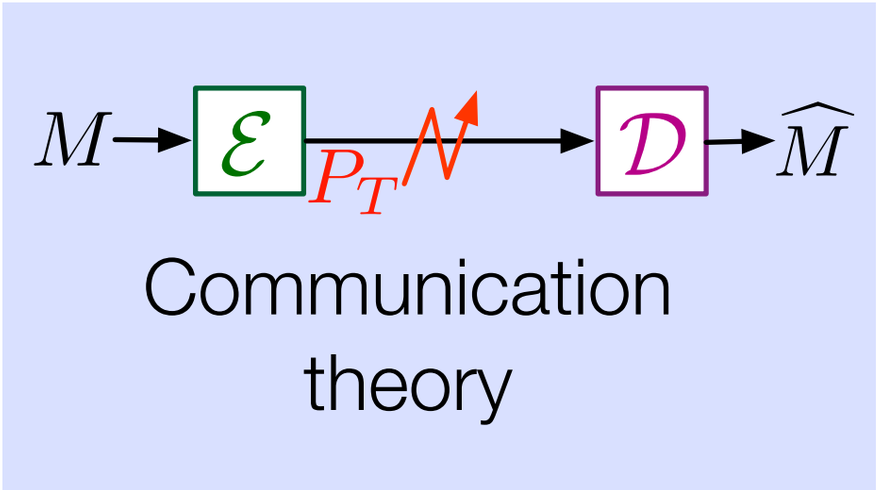
Looking forward:
Implicit information flows in control

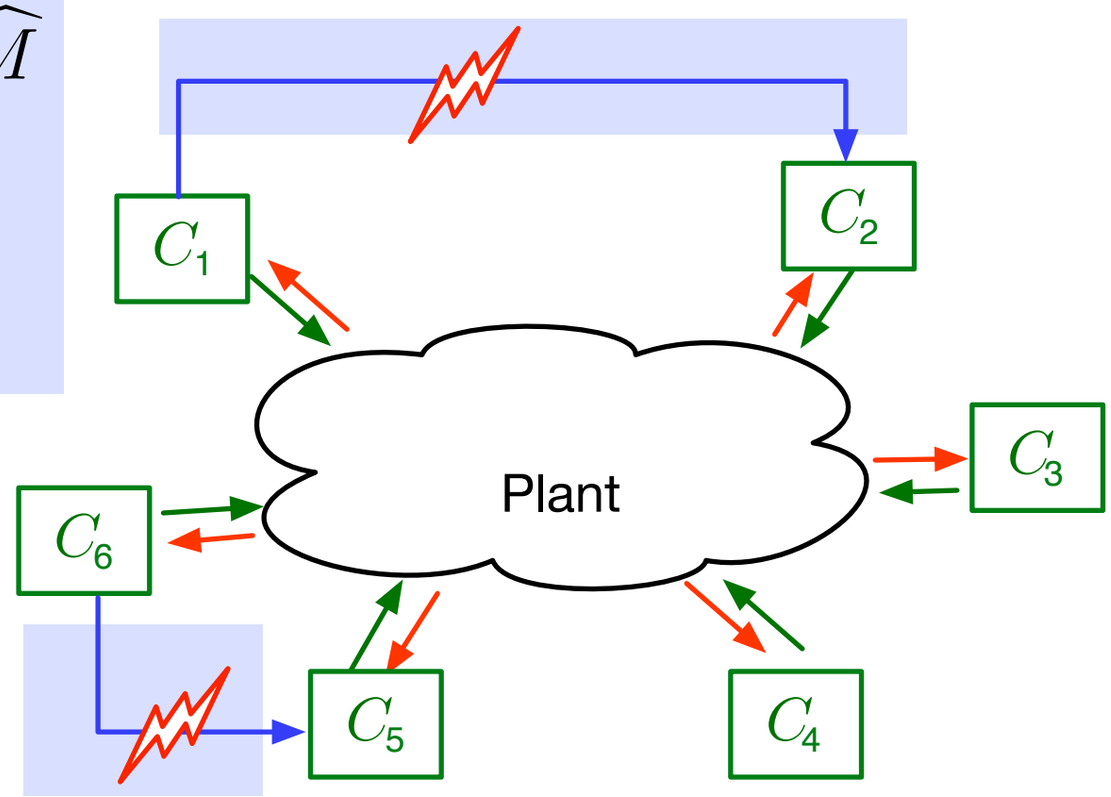


Communication theory

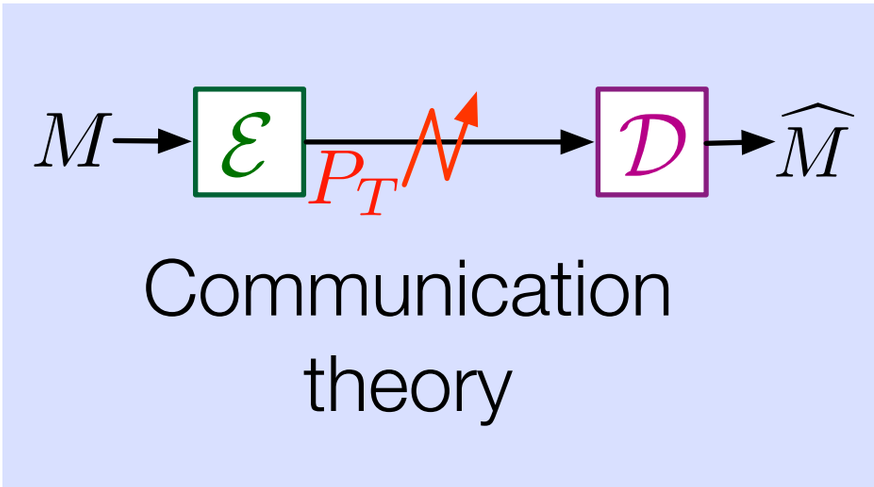


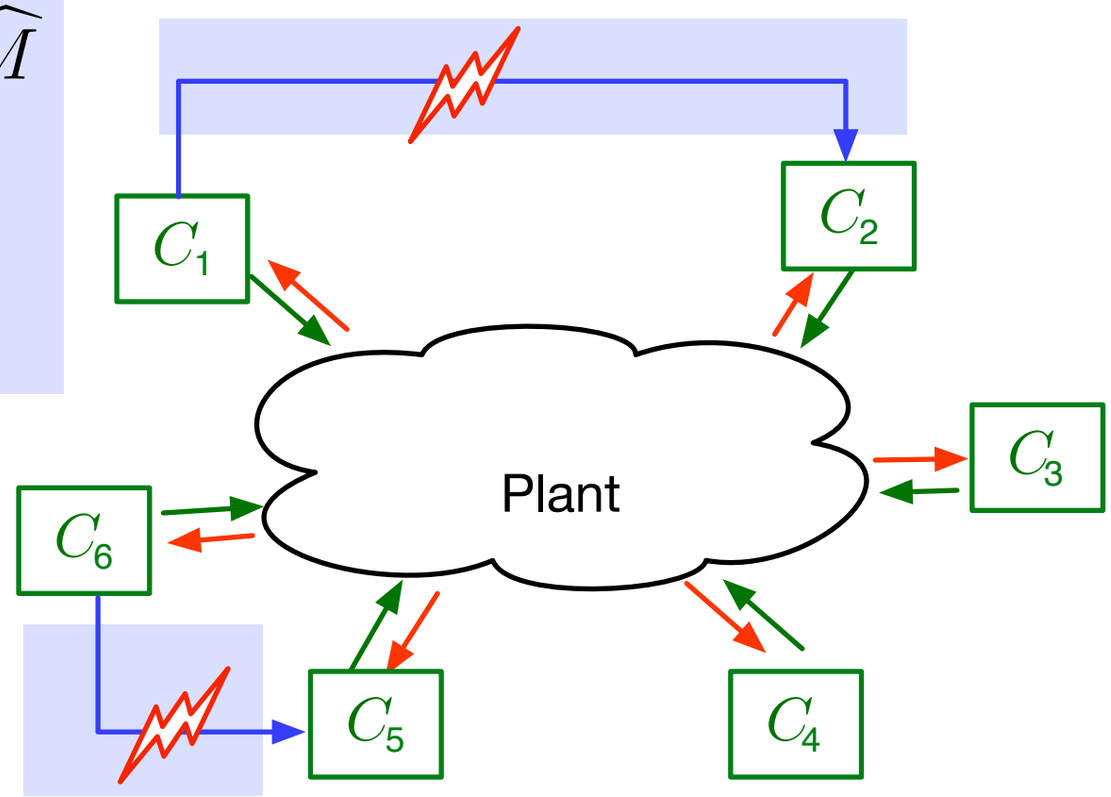
Looking forward:
Implicit information flows in control

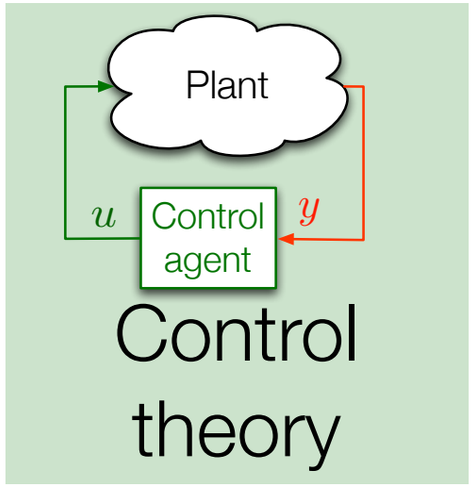




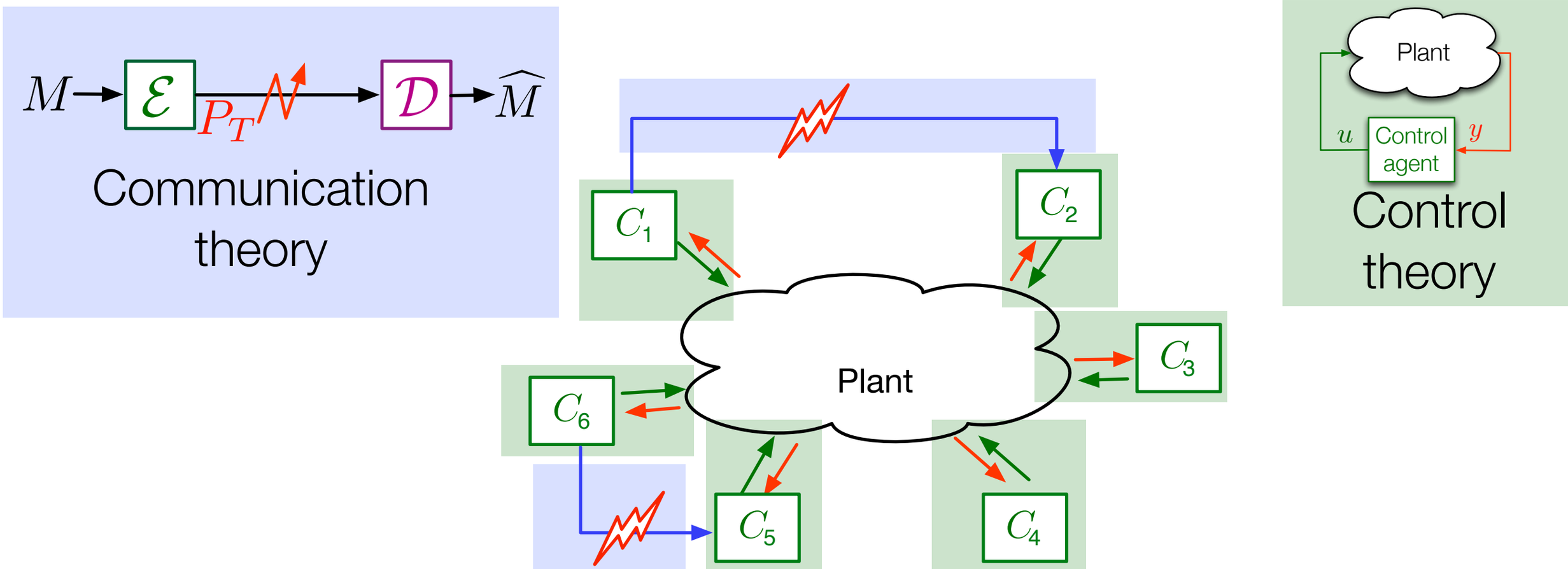
Looking forward:
Implicit information flows in control



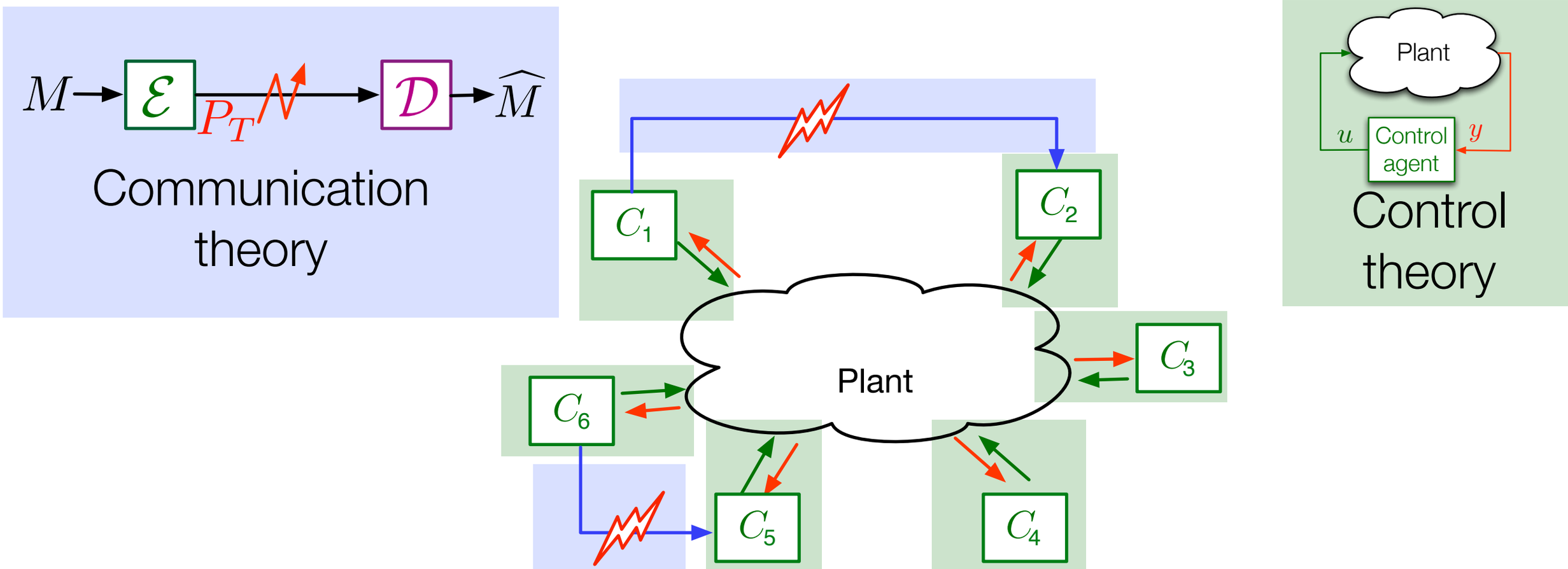




Looking forward:
Implicit information flows in control

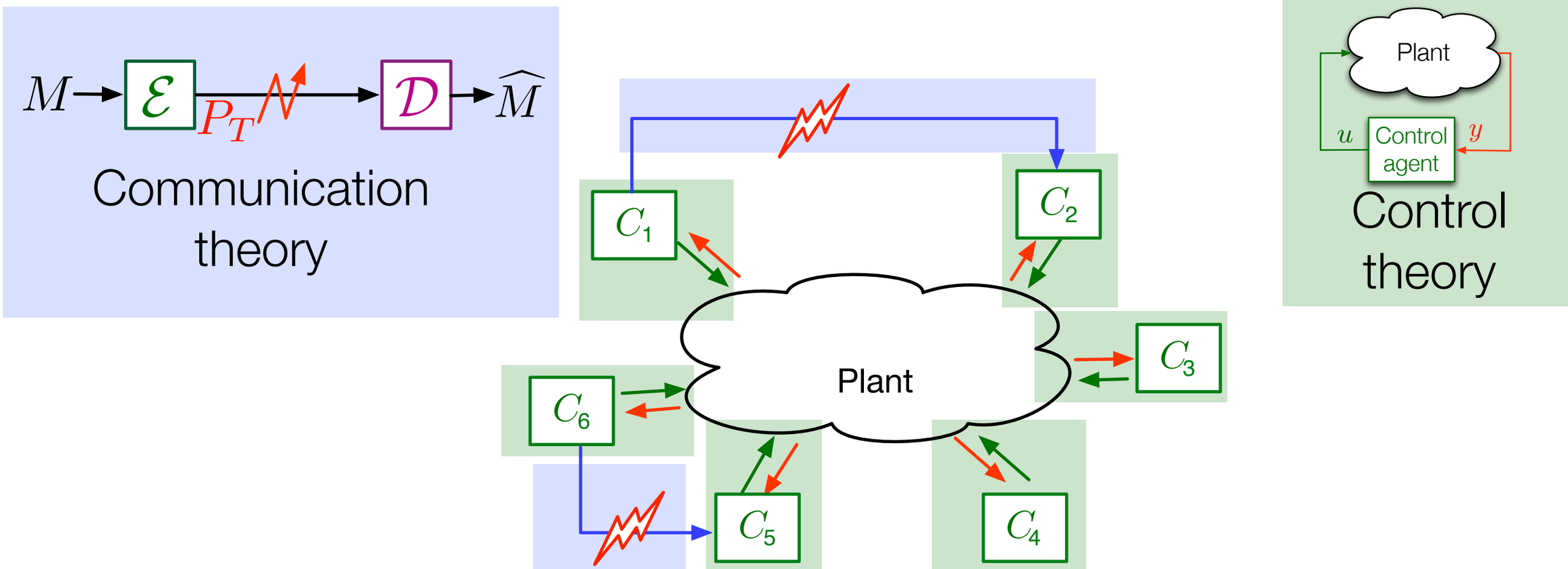


Looking forward:
Implicit information flows in control



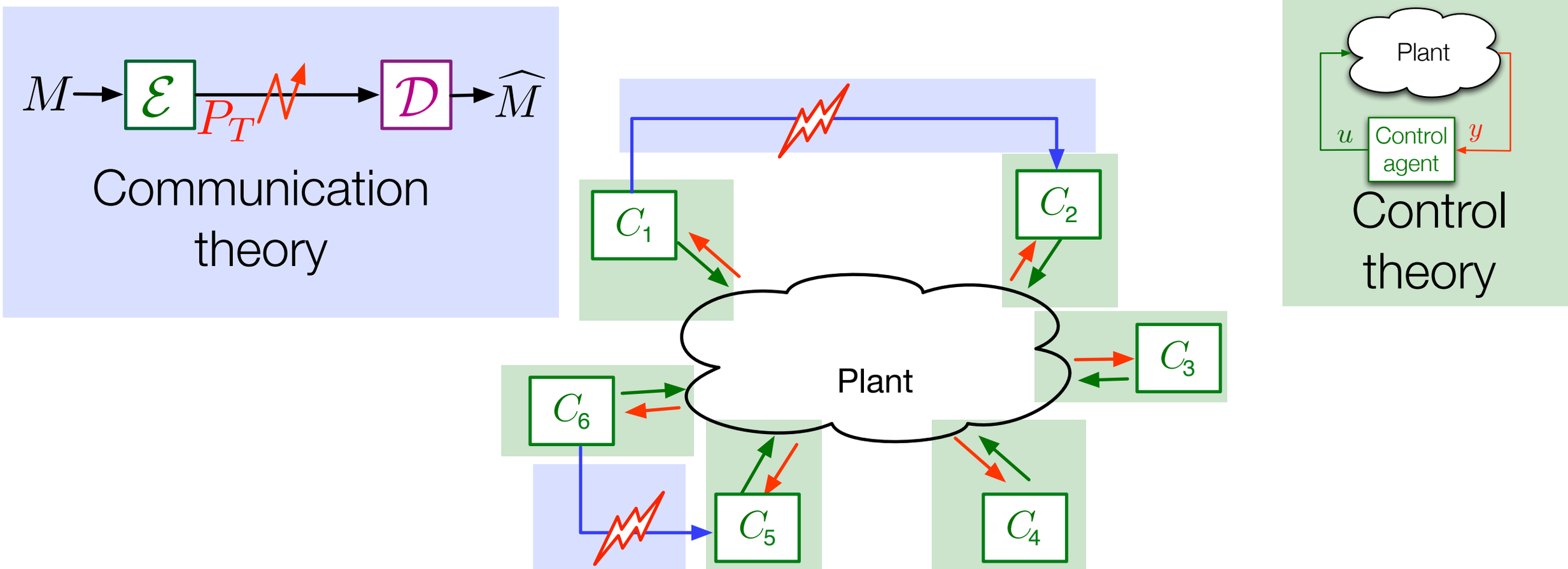
" . . . semantic aspects of communication are irrelevant to *the engineering problem.*"
[Shannon '48]

Looking forward:
Implicit information flows in control

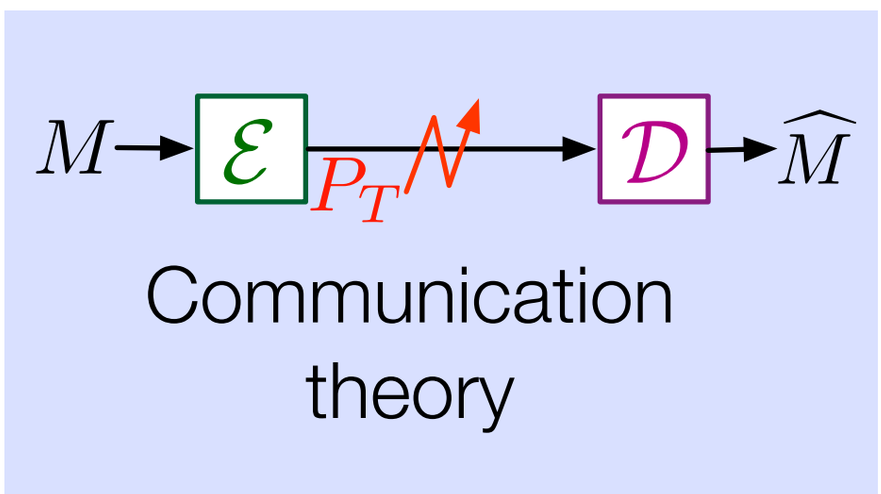


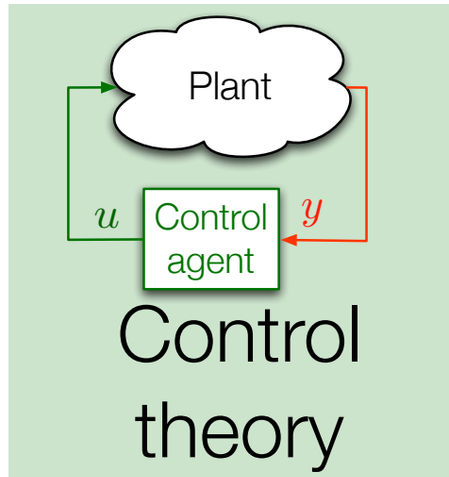
"[comm theory] deals with an essentially simple problem, because the *transmission* of information is considered independently of its *use.*" [Witsenhausen '71]

Looking forward:
Implicit information flows in control

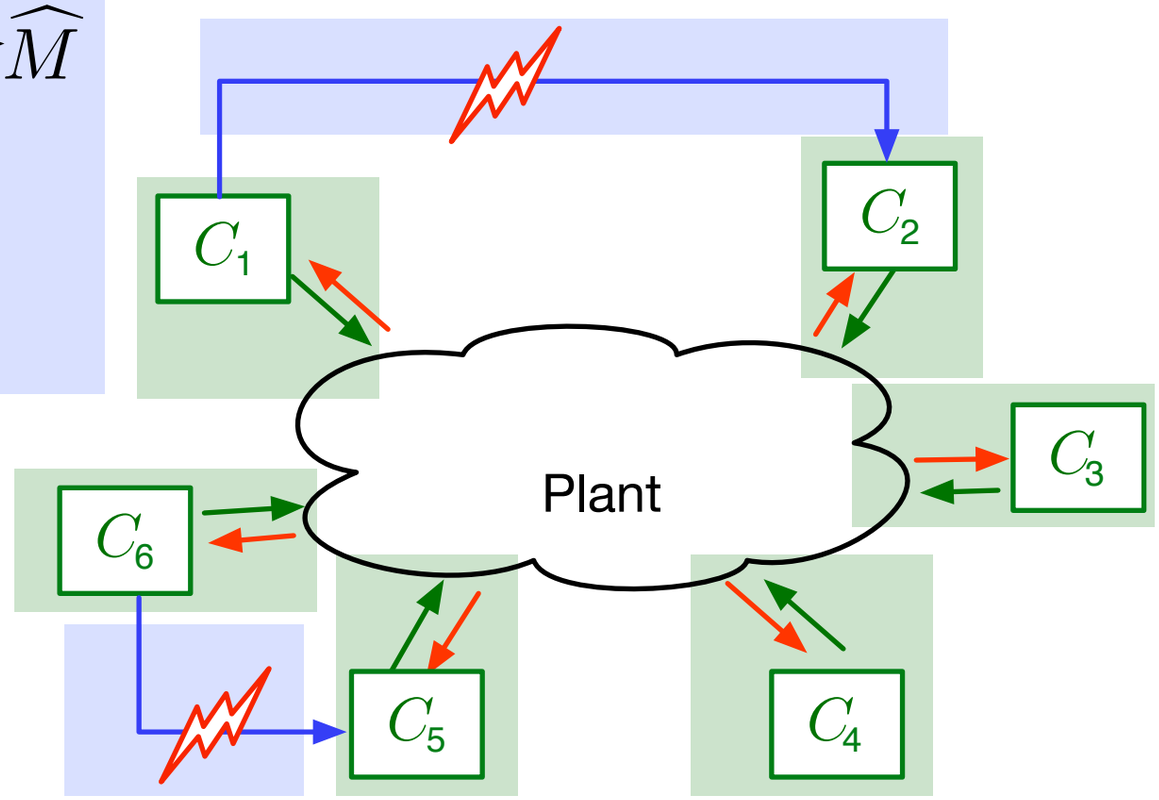


Looking forward:
Implicit information flows in control

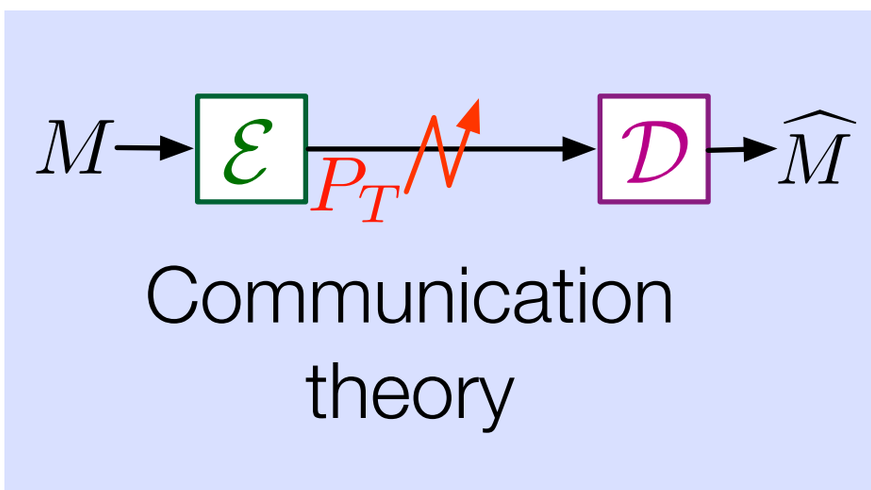


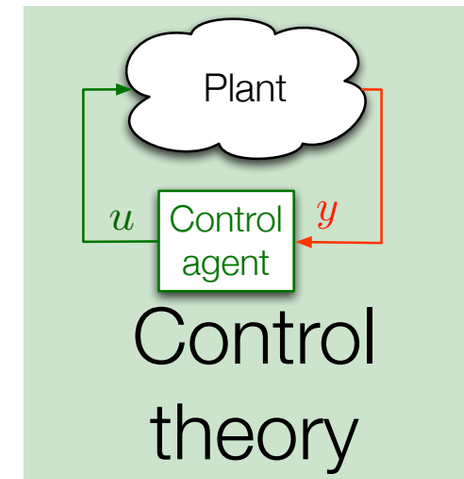


Message?



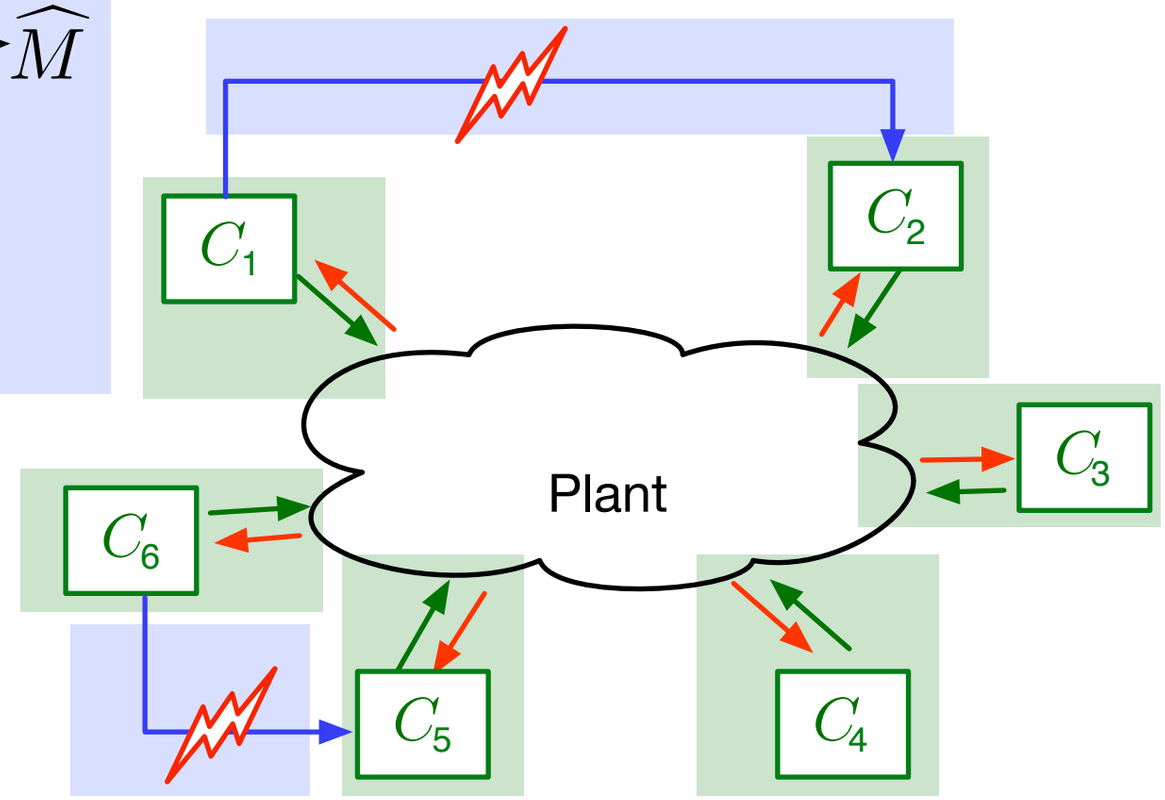
Looking forward:
Implicit information flows in control



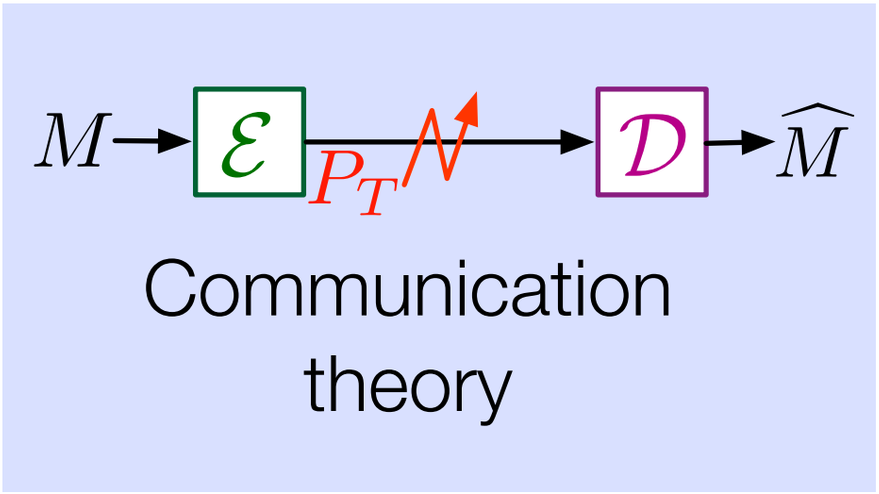


Message?

a choice

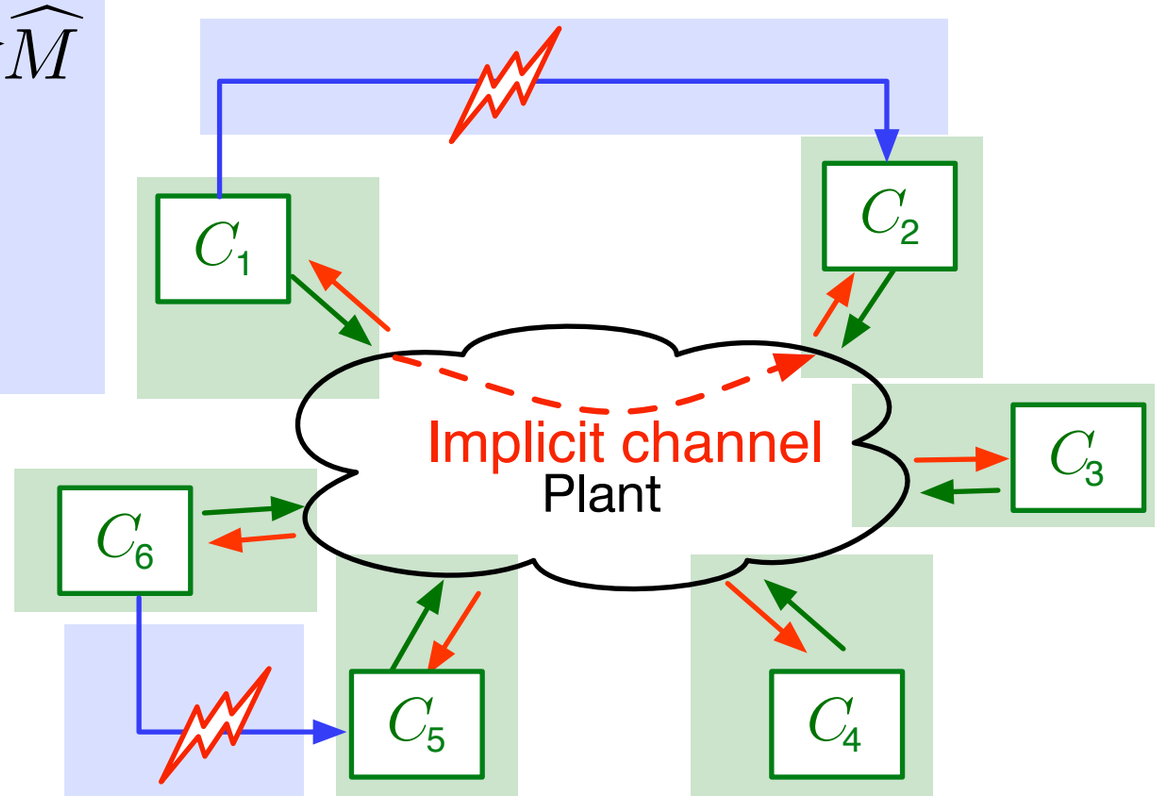


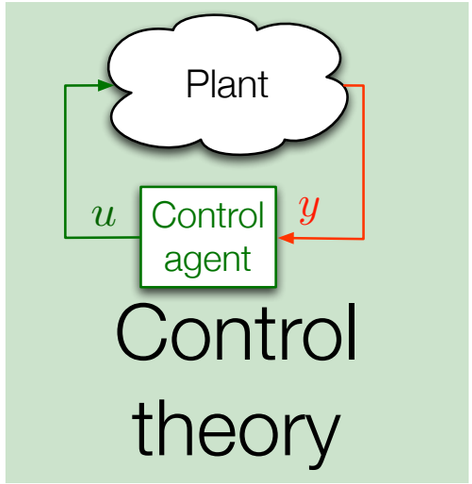
Looking forward:
Implicit information flows in control



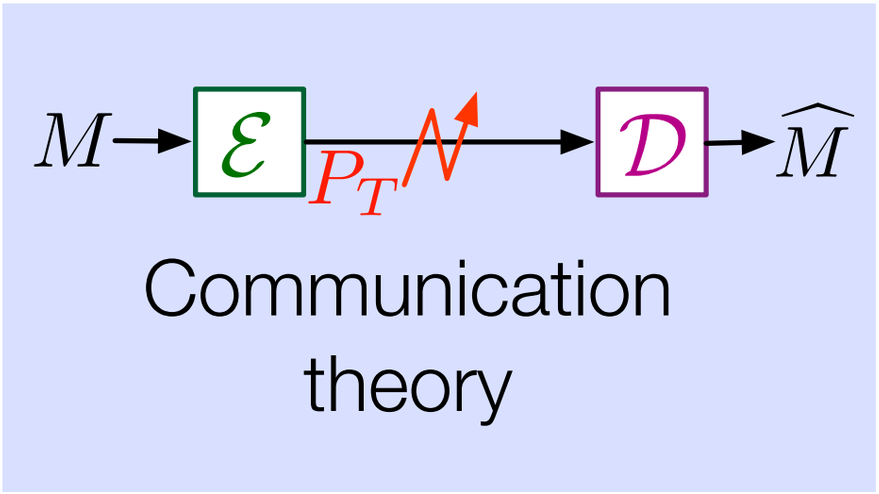
Message?

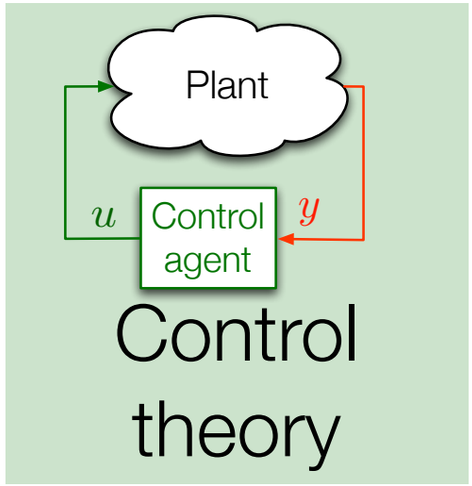
a choice





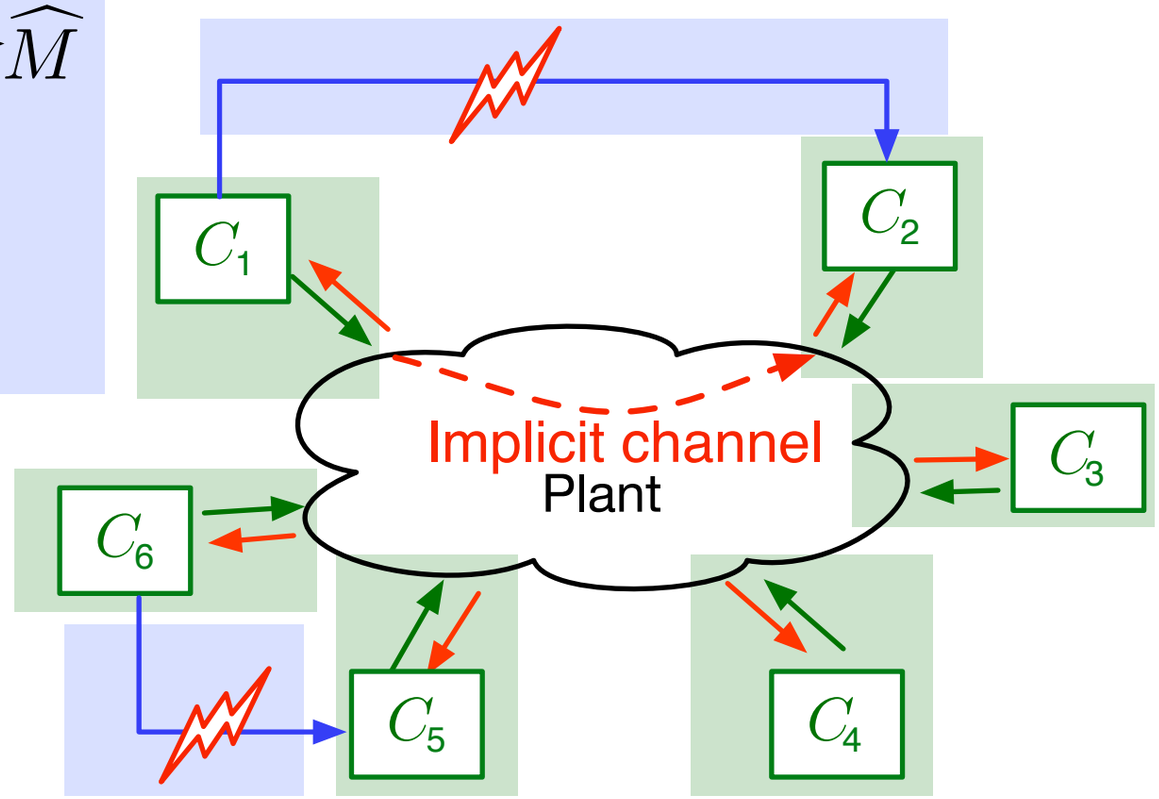
Looking forward:
Implicit information flows in control





Message?

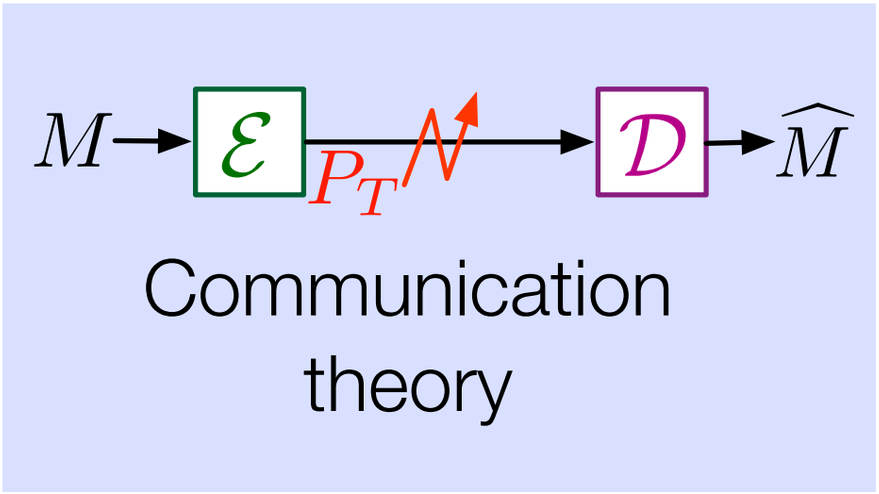
a choice

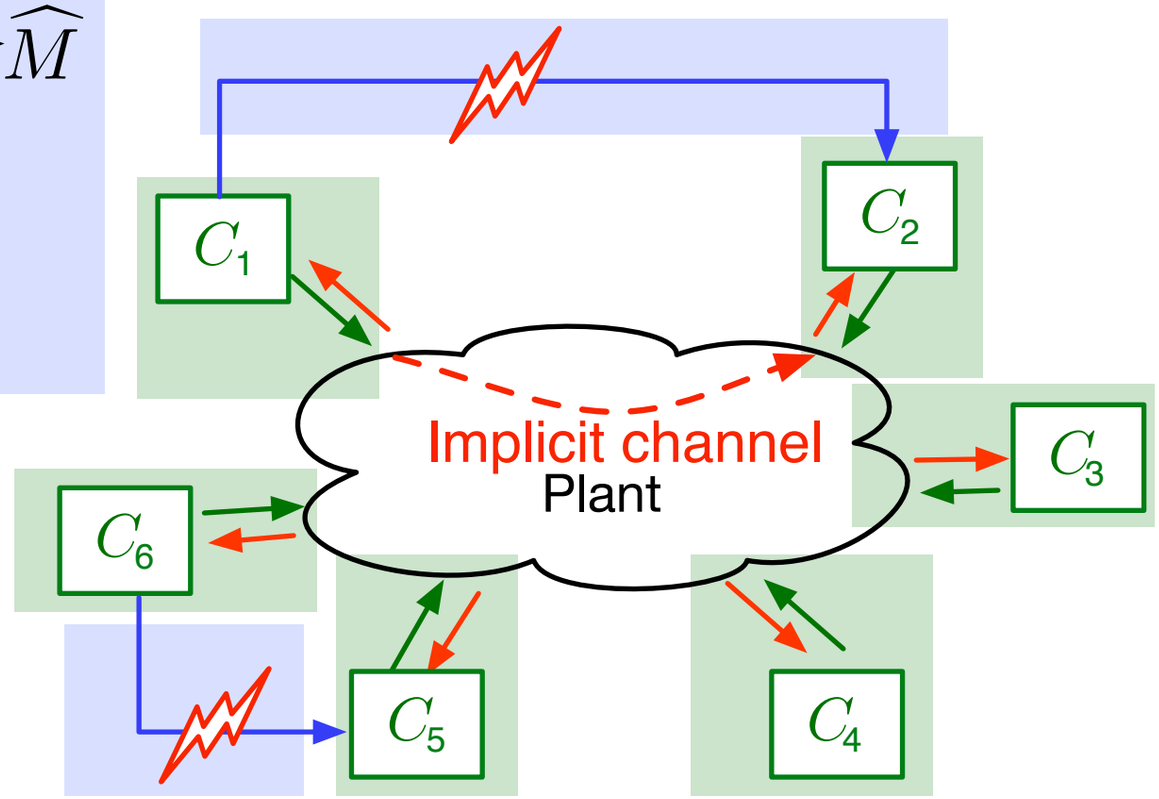


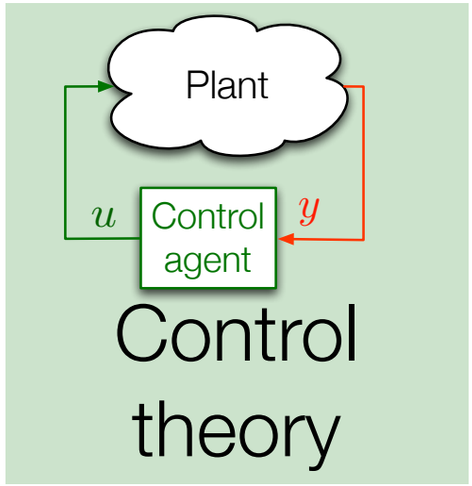
Dual role of actions

- control
- communicate

Looking forward:
Implicit information flows in control





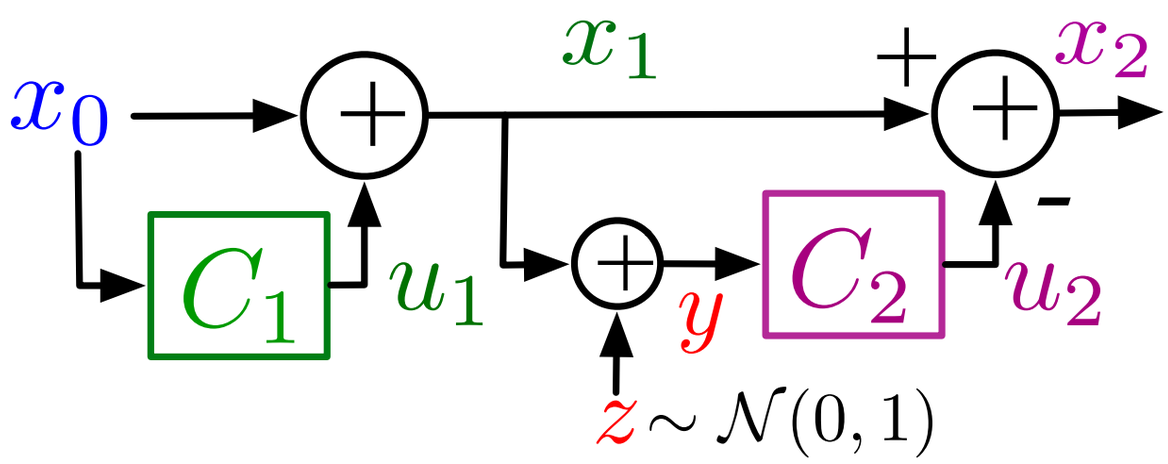


Message?

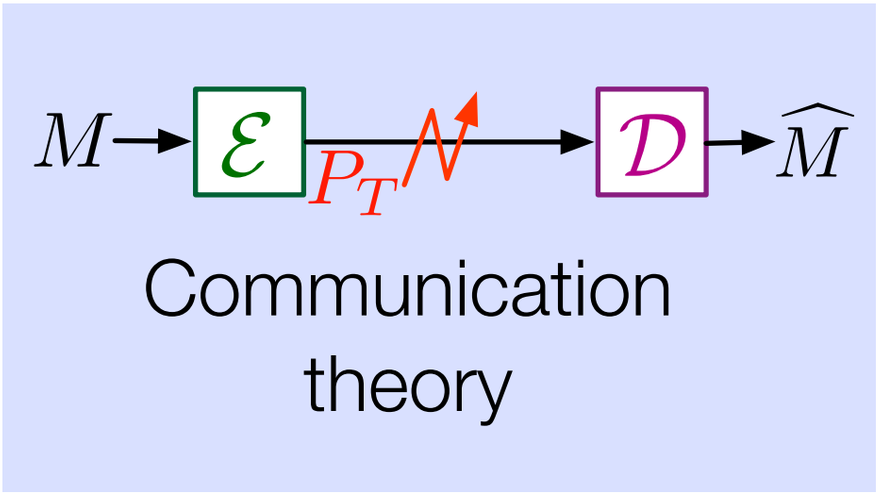
a choice

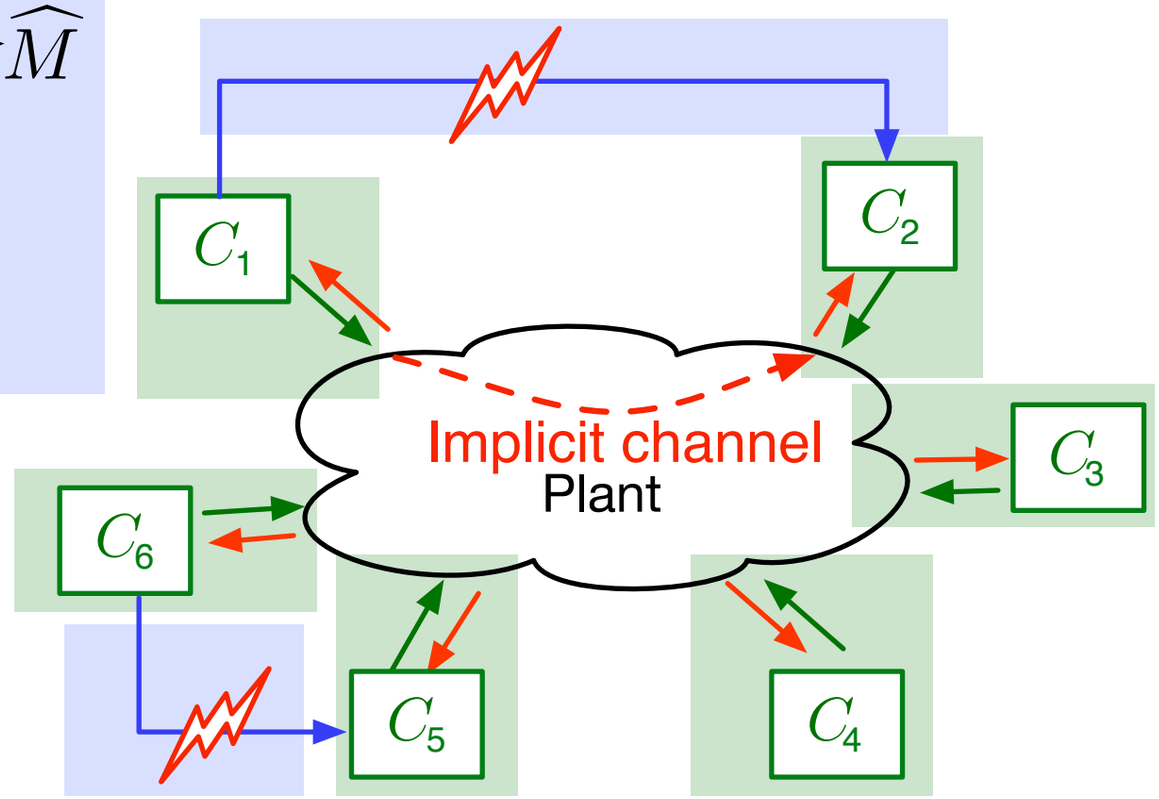
Dual role of actions

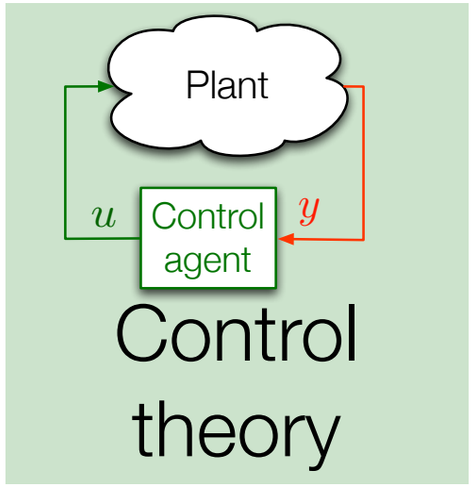
- control
- communicate



Looking forward:
Implicit information flows in control





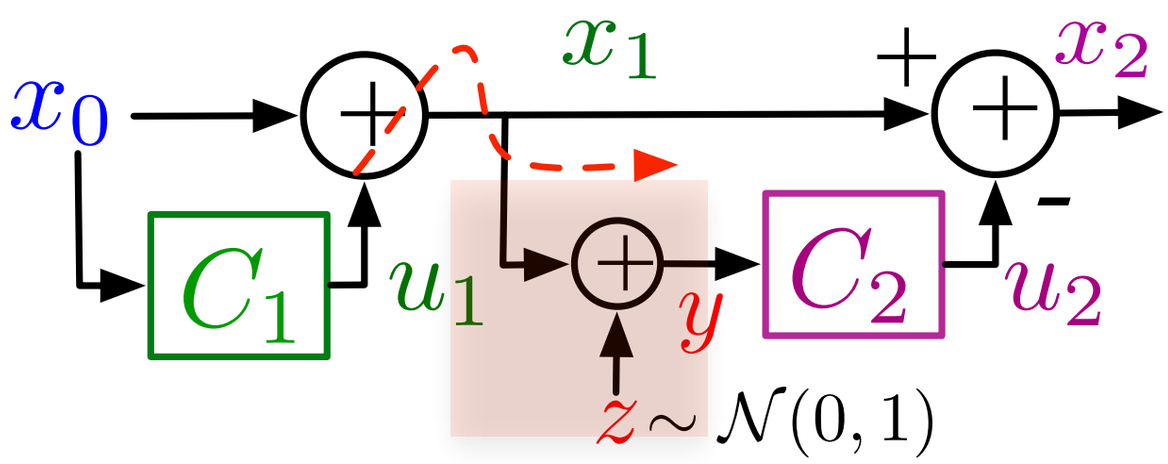


Message?

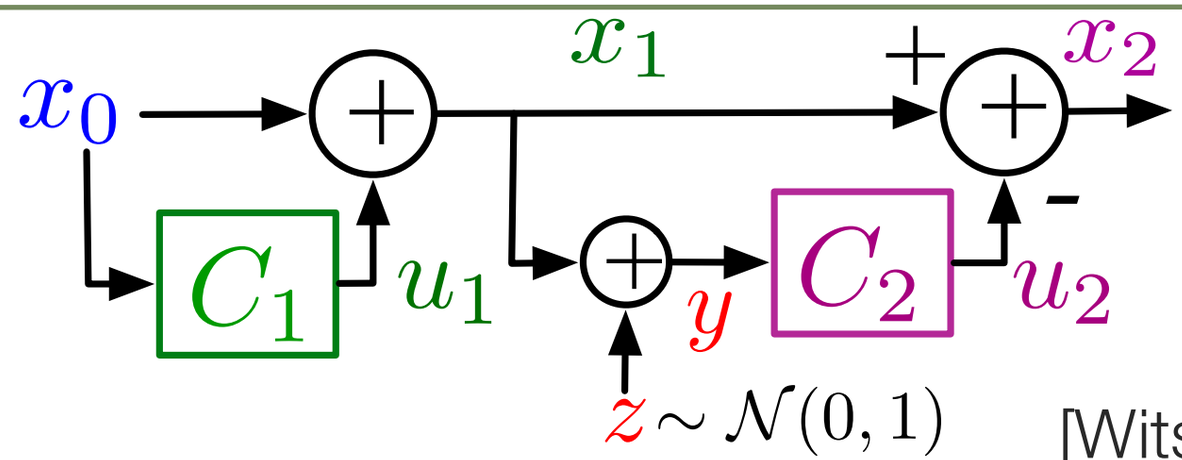
a choice

Dual role of actions

- control
- communicate



Breakdown of traditional theory: Witsenhausen's counterexample

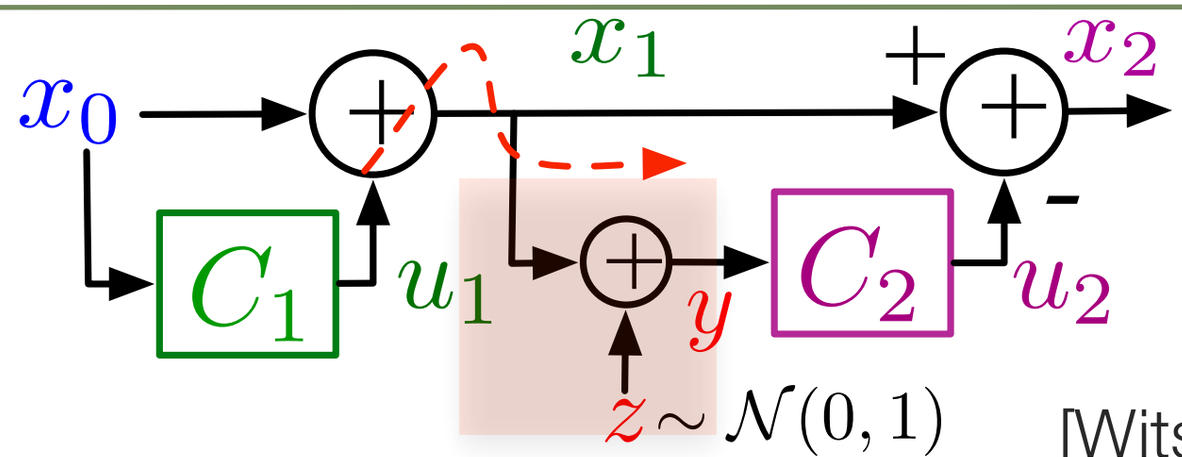


[Witsenhausen '68]

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

Linear, Quadratic, Gaussian

Breakdown of traditional theory: Witsenhausen's counterexample

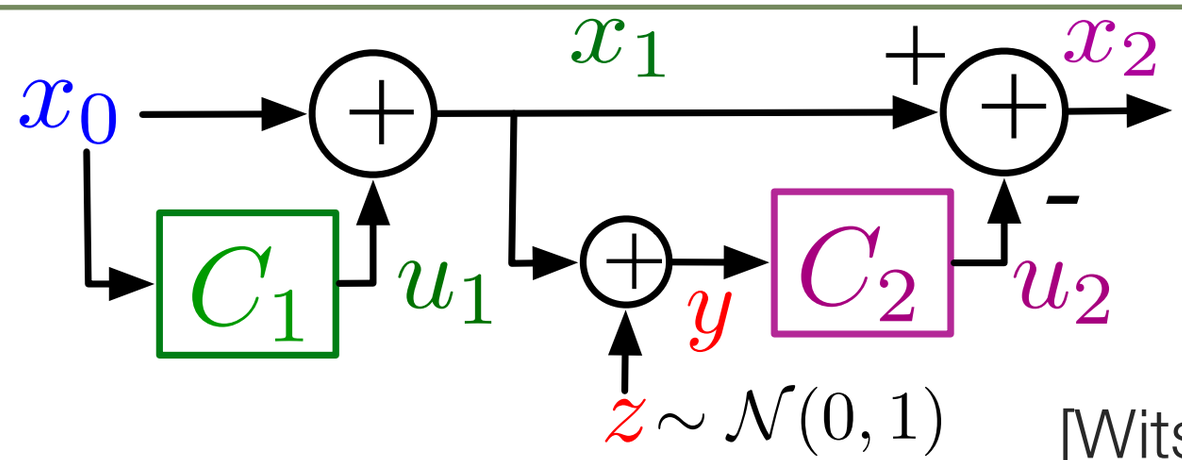


[Witsenhausen '68]

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

Linear, Quadratic, Gaussian

Breakdown of traditional theory: Witsenhausen's counterexample

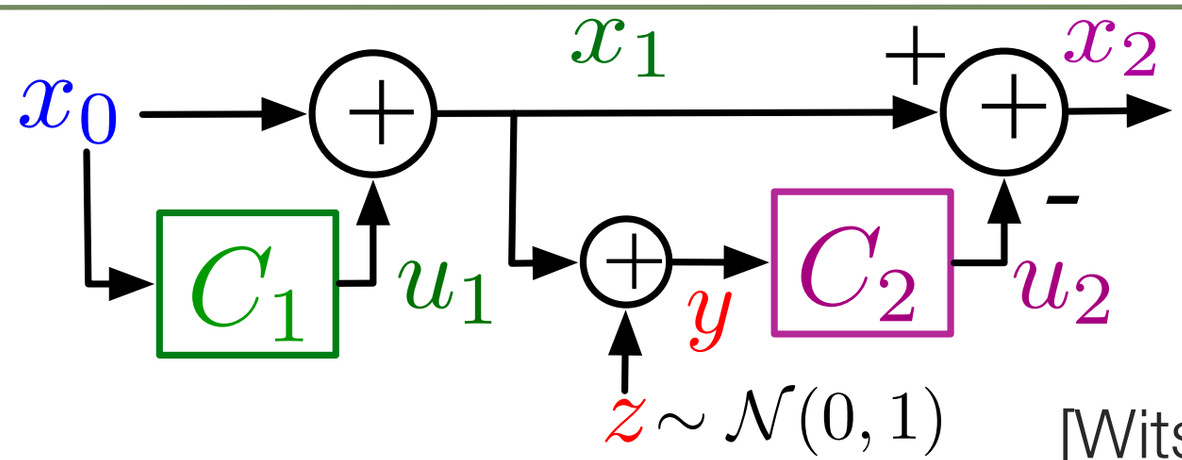


[Witsenhausen '68]

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

Linear, Quadratic, Gaussian

Breakdown of traditional theory: Witsenhausen's counterexample



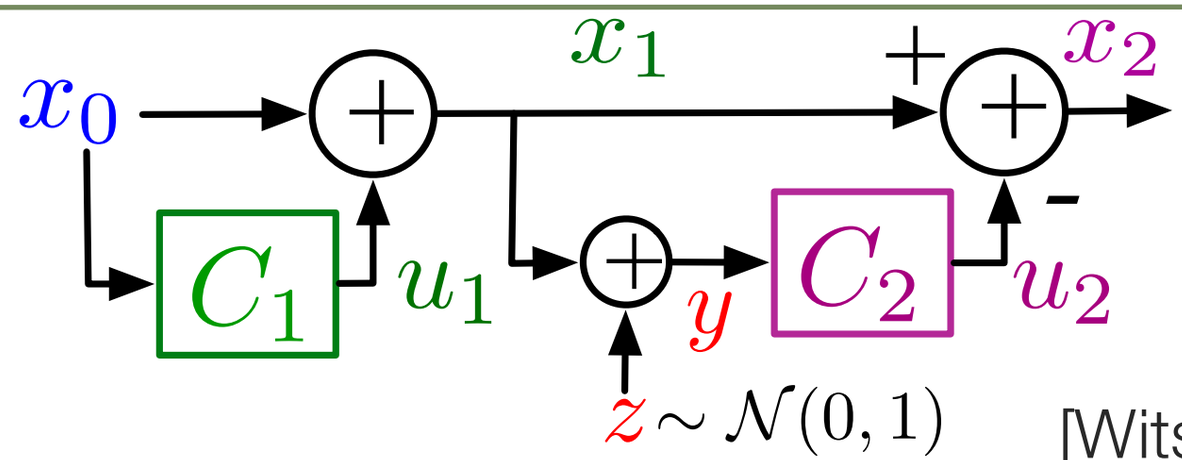
[Witsenhausen '68]

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

Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen '68]

Breakdown of traditional theory: Witsenhausen's counterexample



[Witsenhausen '68]

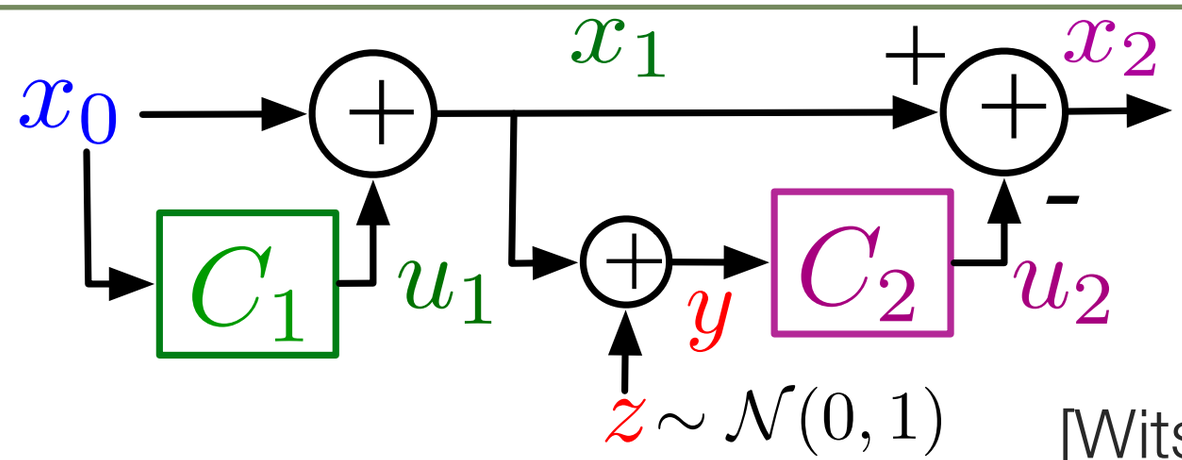
$x_0 \sim \mathcal{N}(0, \sigma_0^2)$ $\quad \min \left\{ k^2 \mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[x_2^2\right] \right\}$

Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen '68]

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

Breakdown of traditional theory: Witsenhausen's counterexample



[Witsenhausen '68]

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

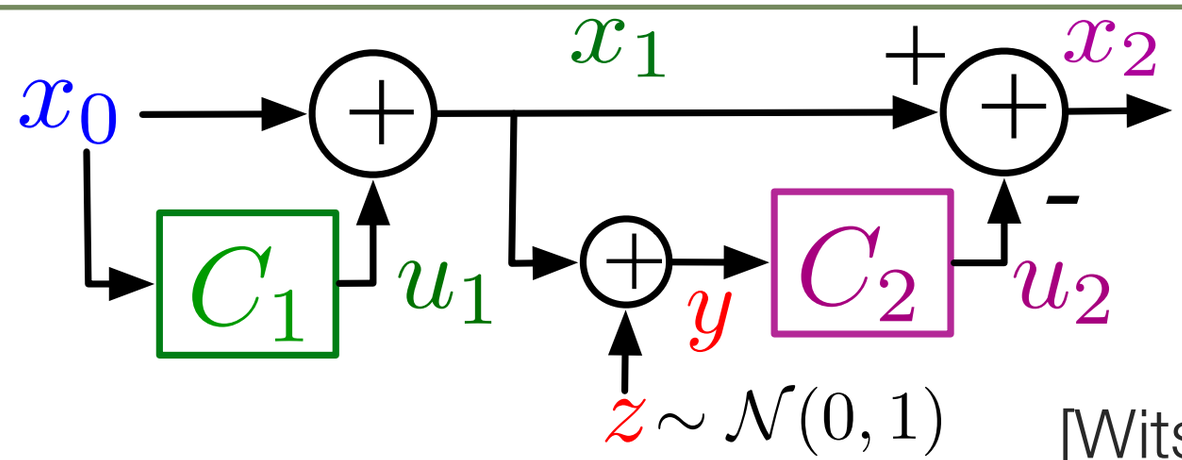
Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen '68]

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

Semi-exhaustive searches
[Deng, 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



[Witsenhausen '68]

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

Linear, Quadratic, Gaussian

Nonlinear strategies can outperform linear strategies [Witsenhausen '68]

Nonconvex [Witsenhausen '68],

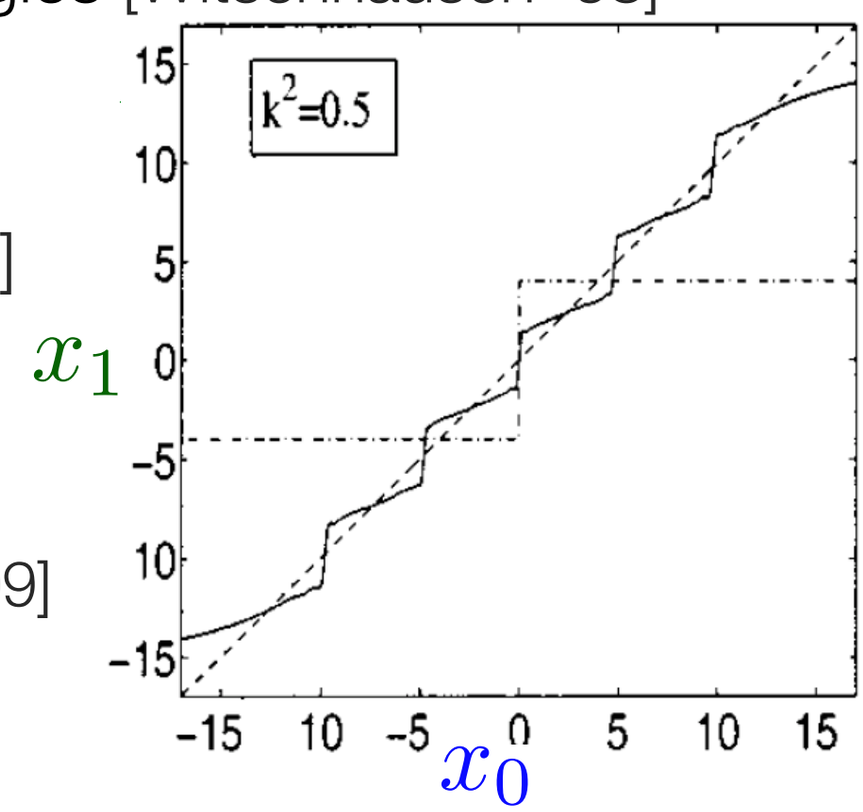
"Intractable" (NP-hard) [Papadimitriou, Tsitsiklis '86]

Semi-exhaustive searches

[Deng, Ho '99] [Baglietto *et al*. '01]

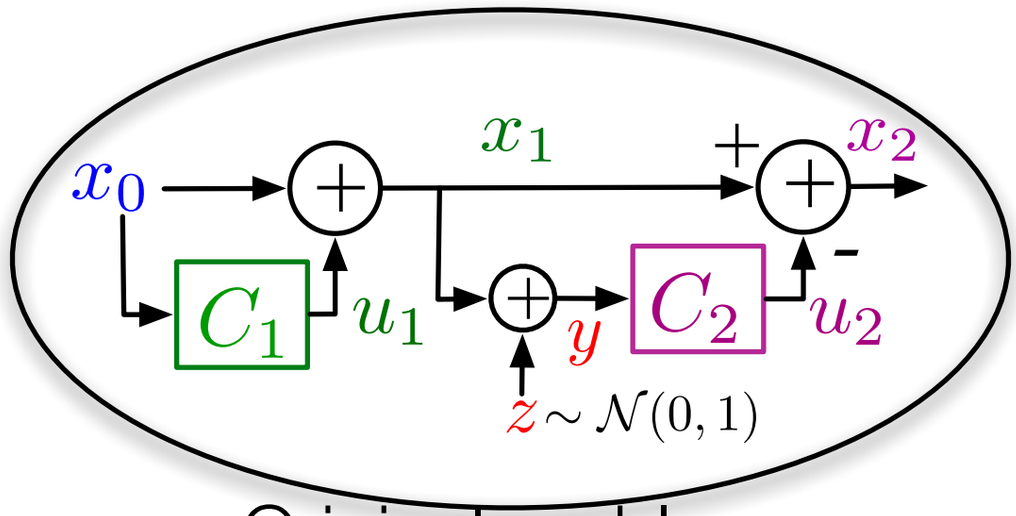
[Lee, Lau, Ho '01][Li, Marden, Shamma '09]

[Karlsson *et al*. '10]



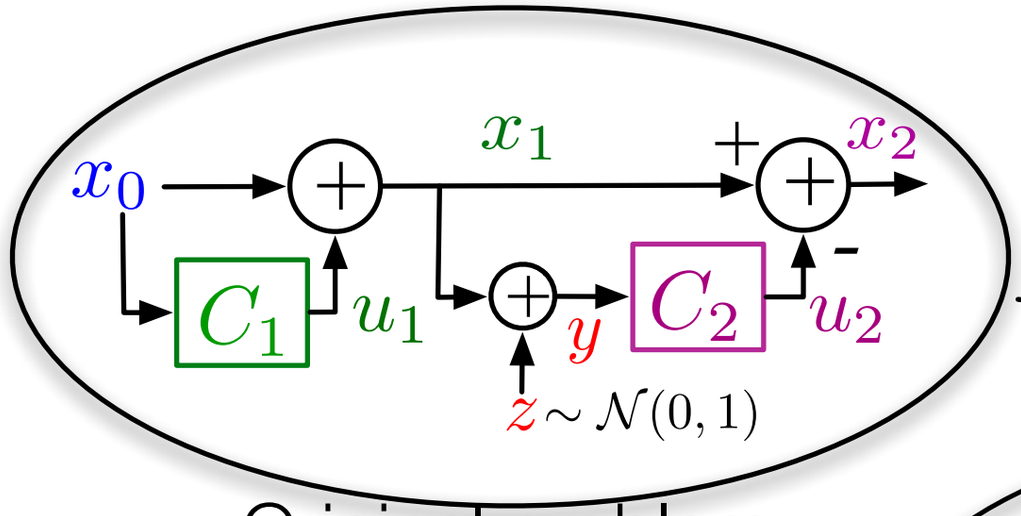
Our approach: understanding information flows to obtain fundamental limits on cost

Our approach: understanding information flows to obtain fundamental limits on cost

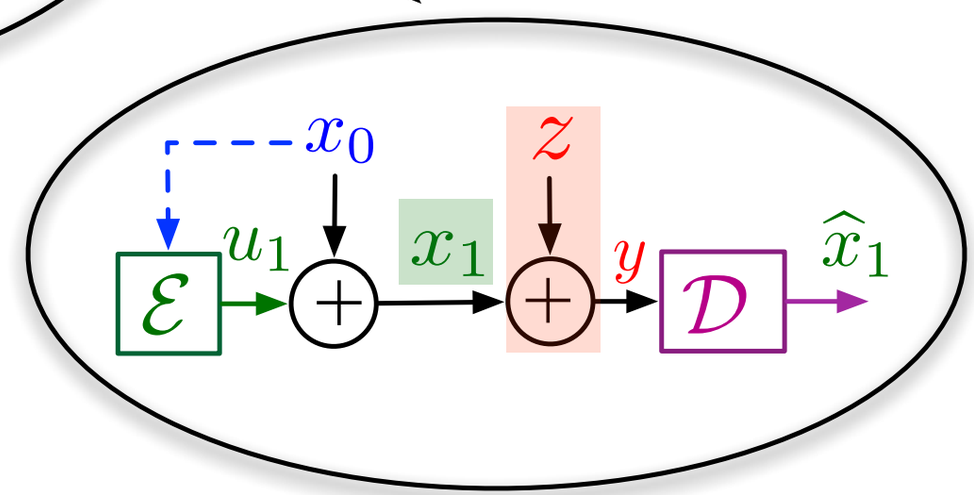


Original problem
[Witsenhausen '68]

Our approach: understanding information flows to obtain fundamental limits on cost

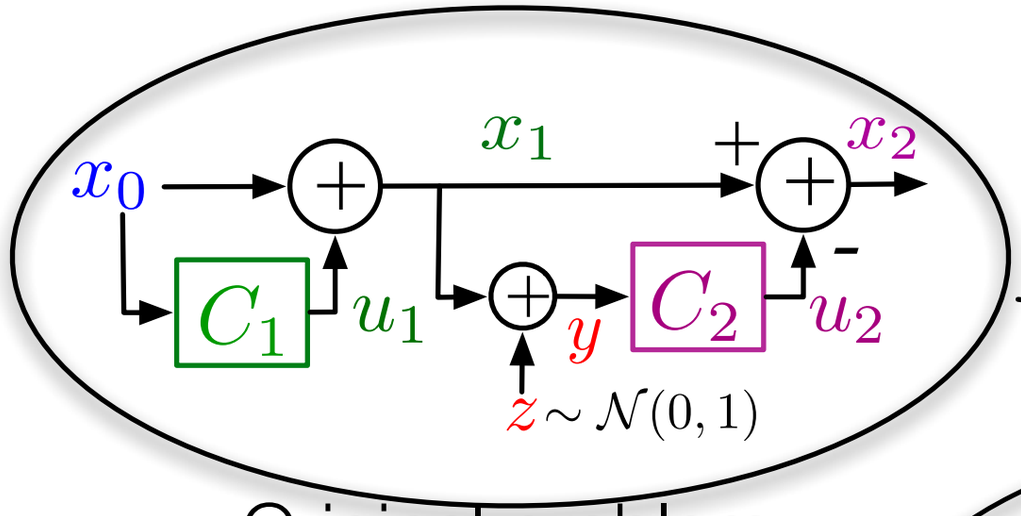


Original problem
[Witsenhausen '68]

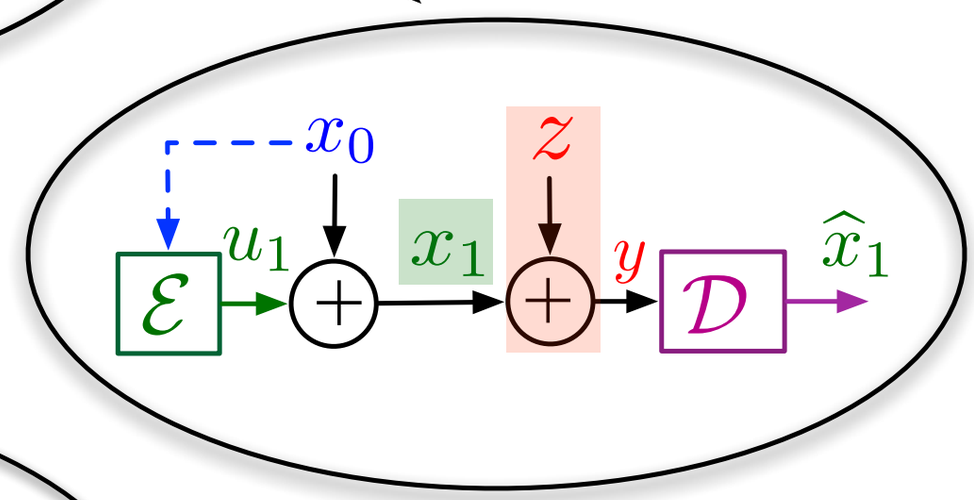


Identify
information-flow
[Grover, Sahai '08]

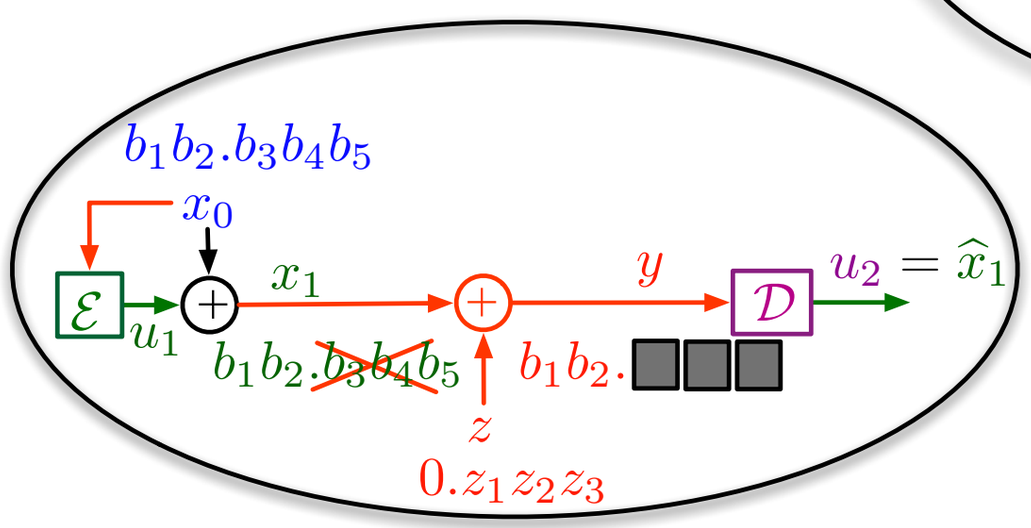
Our approach: understanding information flows to obtain fundamental limits on cost



Original problem
[Witsenhausen '68]

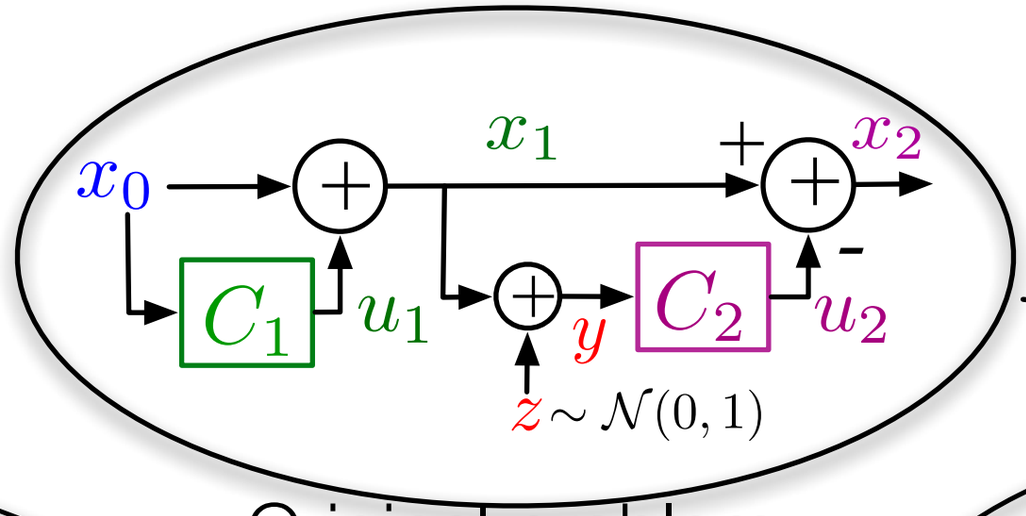


Identify information-flow
[Grover, Sahai '08]

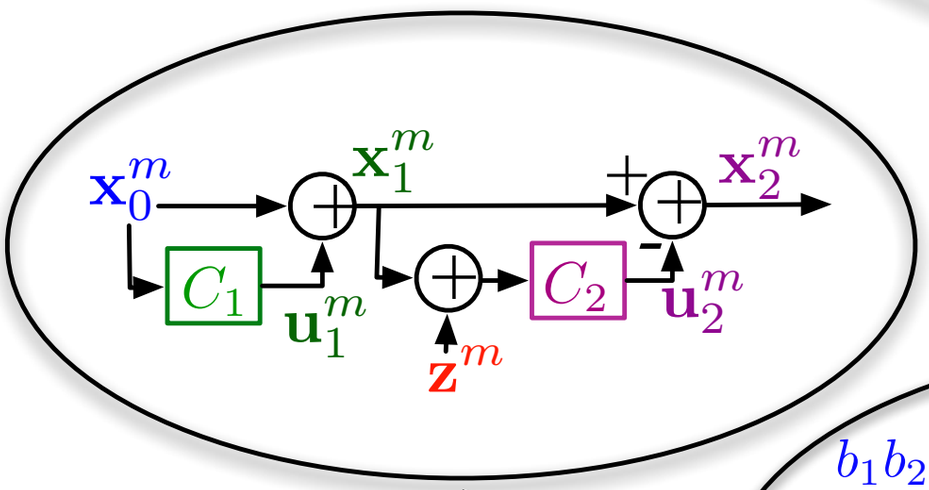


Understand information-flow:
Binary version [Avestimehr et al. '08][Grover, Sahai '09]

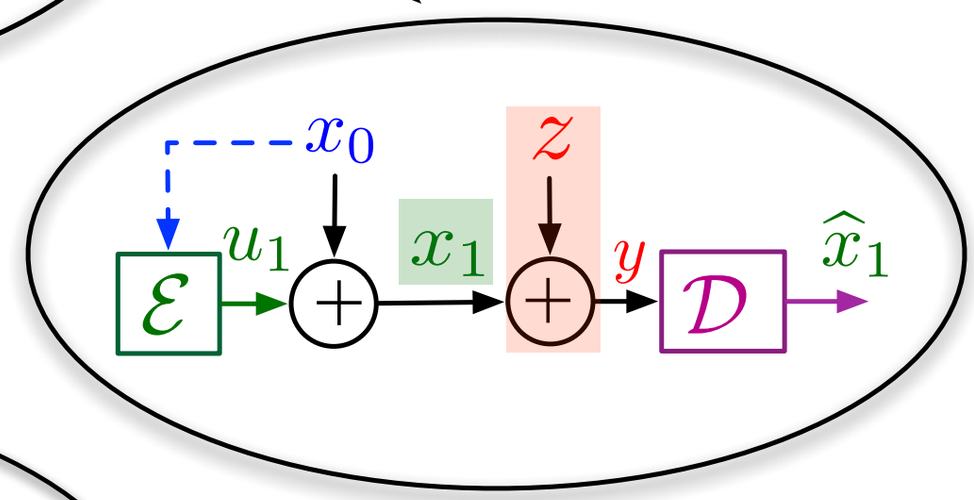
Our approach: understanding information flows to obtain fundamental limits on cost



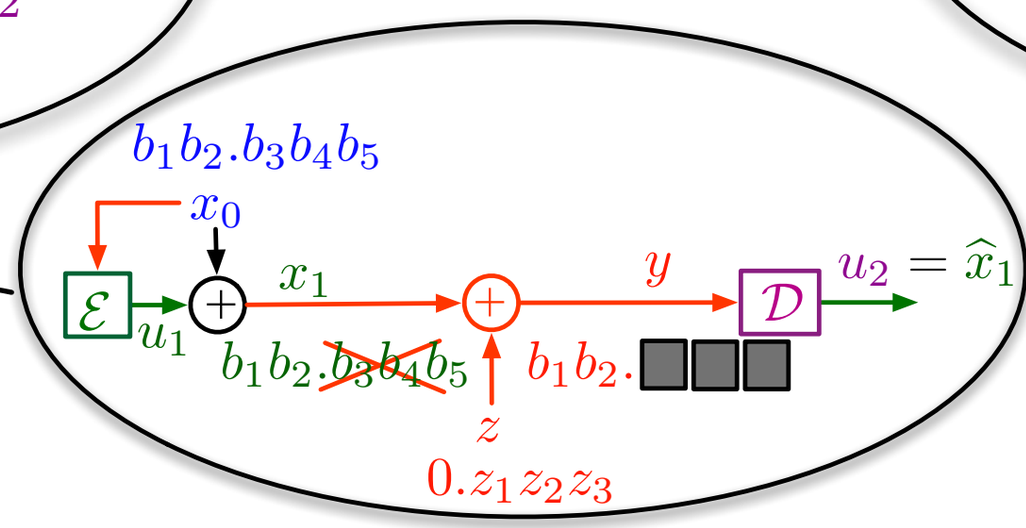
Original problem
[Witsenhausen '68]



Asymptotic vector version
[Grover, Sahai '08]



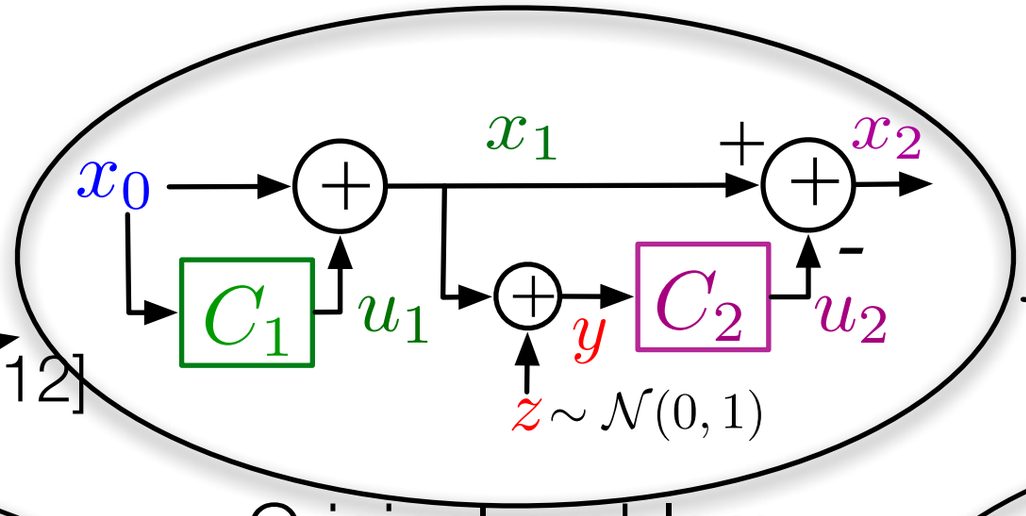
Identify information-flow
[Grover, Sahai '08]



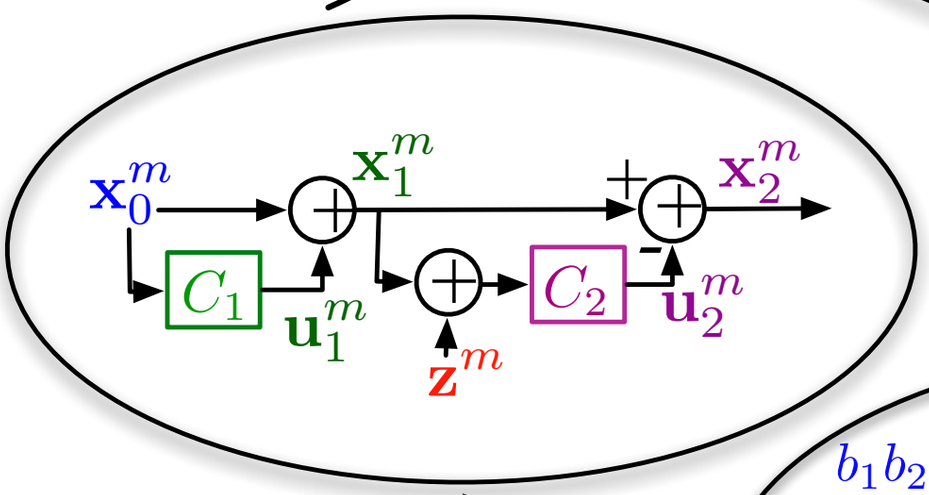
Understand information-flow:
Binary version [Avestimehr et al. '08][Grover, Sahai '09]

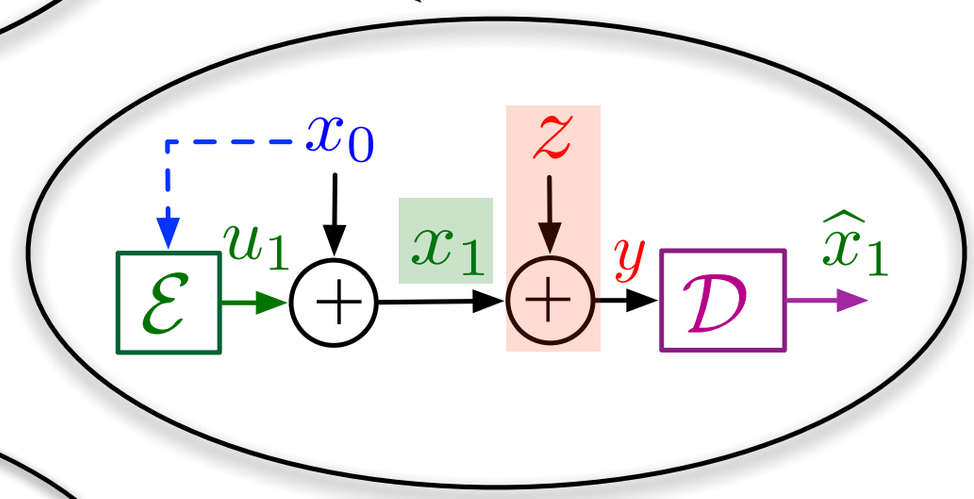
Our approach: understanding information flows to obtain fundamental limits on cost

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

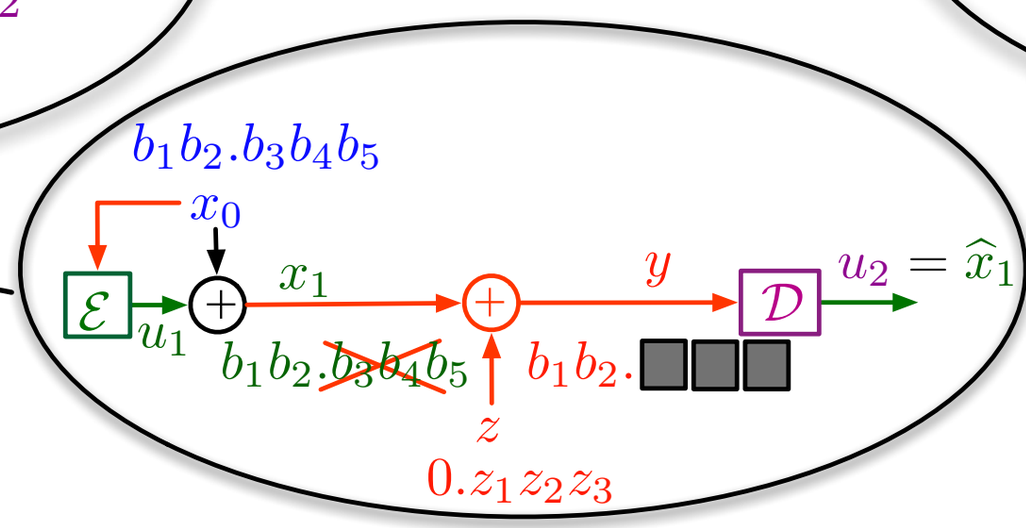


Original problem
[Witsenhausen '68]





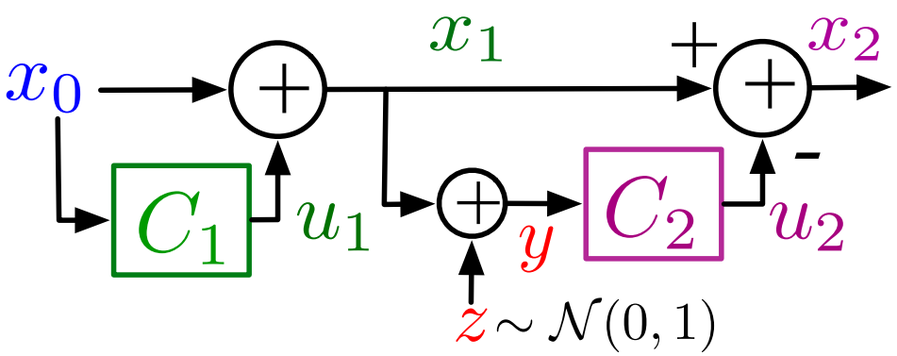
Asymptotic
vector version
[Grover, Sahai '08]



Identify
information-flow
[Grover, Sahai '08]

Understand information-flow:
Binary version [Avestimehr et al. '08][Grover, Sahai '09]

Approx. solutions to Witsenhausen's counterexample



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

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

For **all** (k, σ_0), $\dfrac{\text{Upper bound}}{\text{Lower bound}} \leq 8$

Upper bound

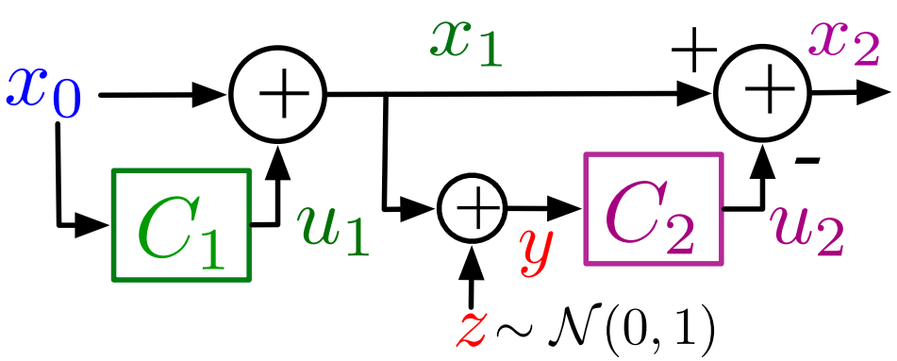
Scalar quantization

Lower bound

Info-theory + Large deviations

First approximately-optimal solutions to Witsenhausen's counterexample

Approx. solutions to Witsenhausen's counterexample



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

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

For **all** (k, σ_0), $\dfrac{\text{Upper bound}}{\text{Lower bound}} \leq 8$

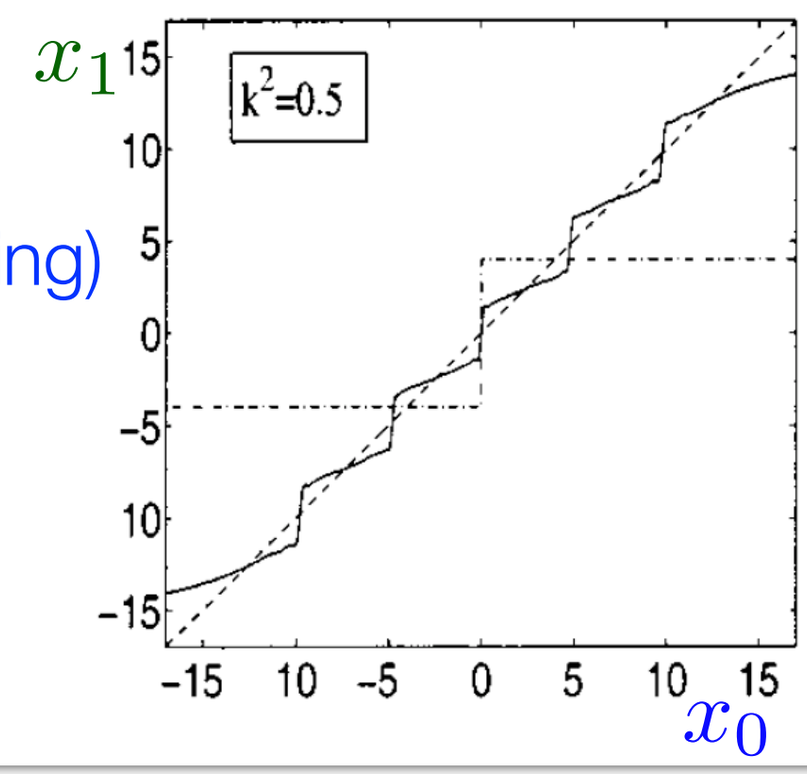
Upper bound

Scalar quantization

"Watermarking"
("dirty-paper" coding)

Lower bound

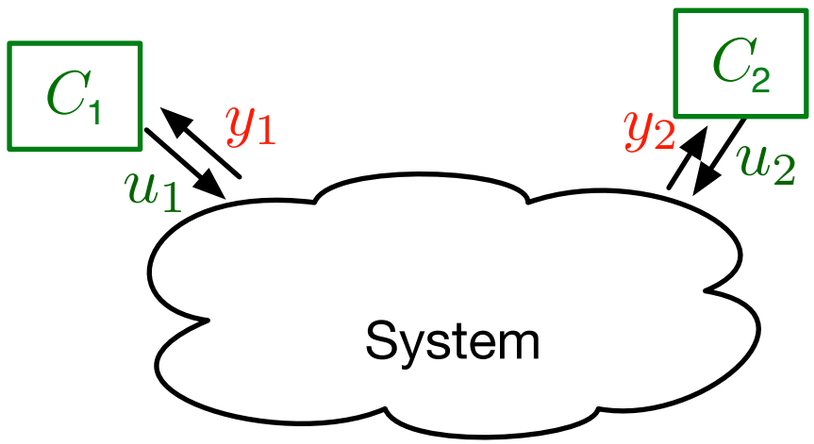
Info-theory + Large deviations



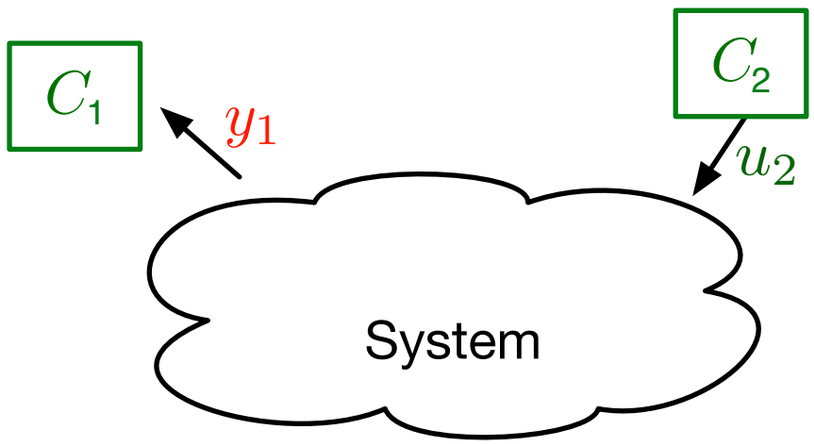
First approximately-optimal solutions to Witsenhausen's counterexample

Looking forward:
Information-flows in Cyber-Physical Systems

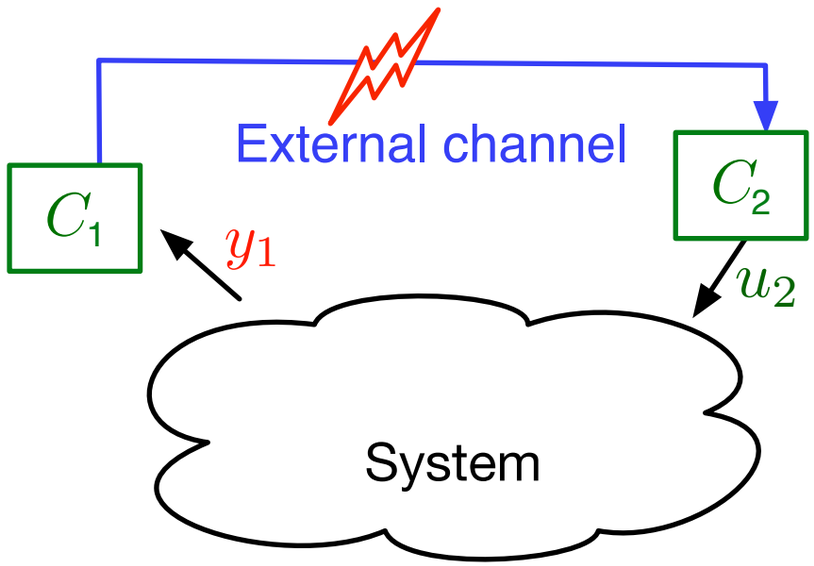
Looking forward:
Information-flows in Cyber-Physical Systems



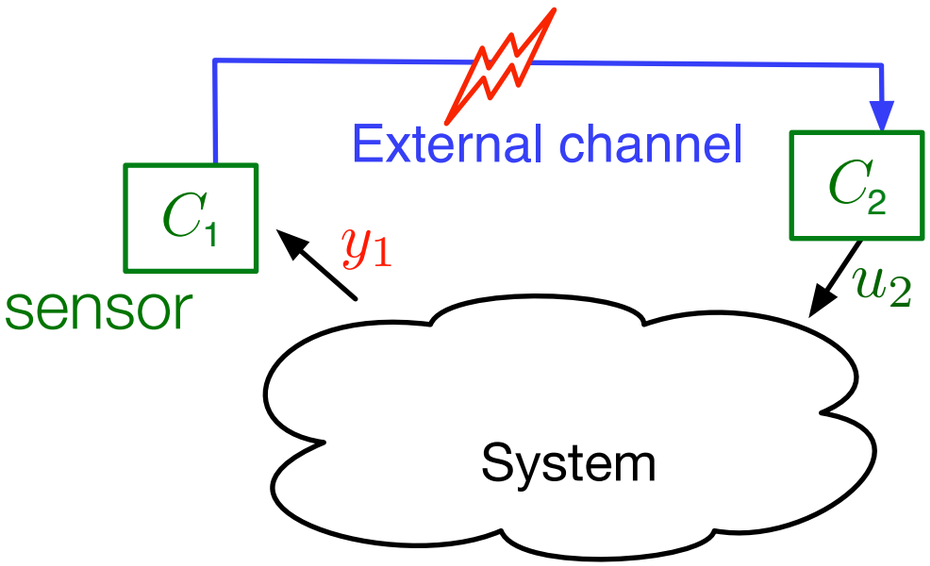
Looking forward:
Information-flows in Cyber-Physical Systems



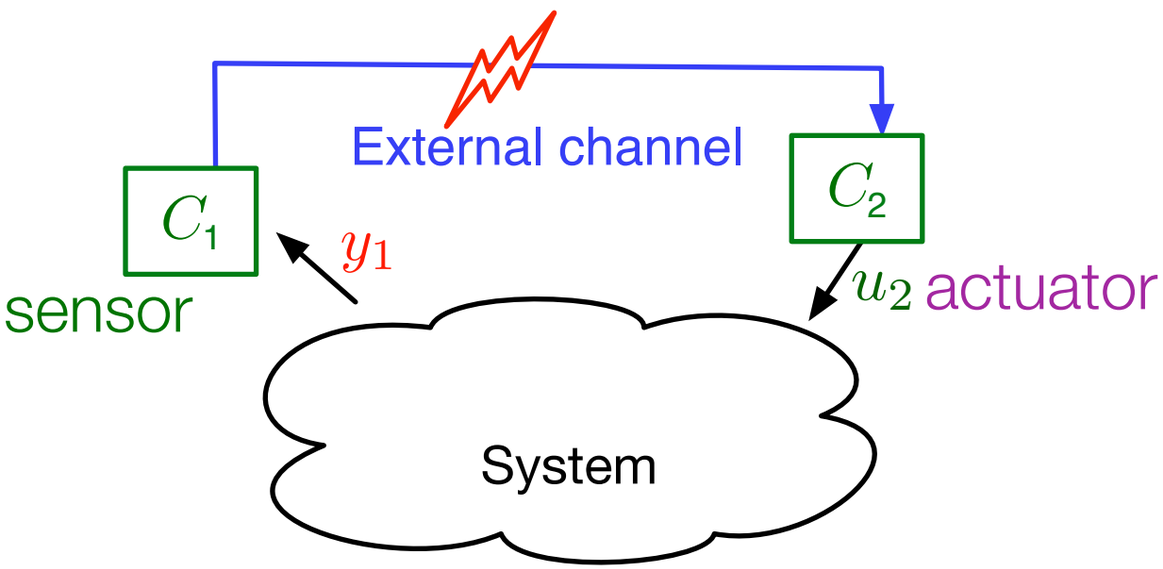
Looking forward:
Information-flows in Cyber-Physical Systems



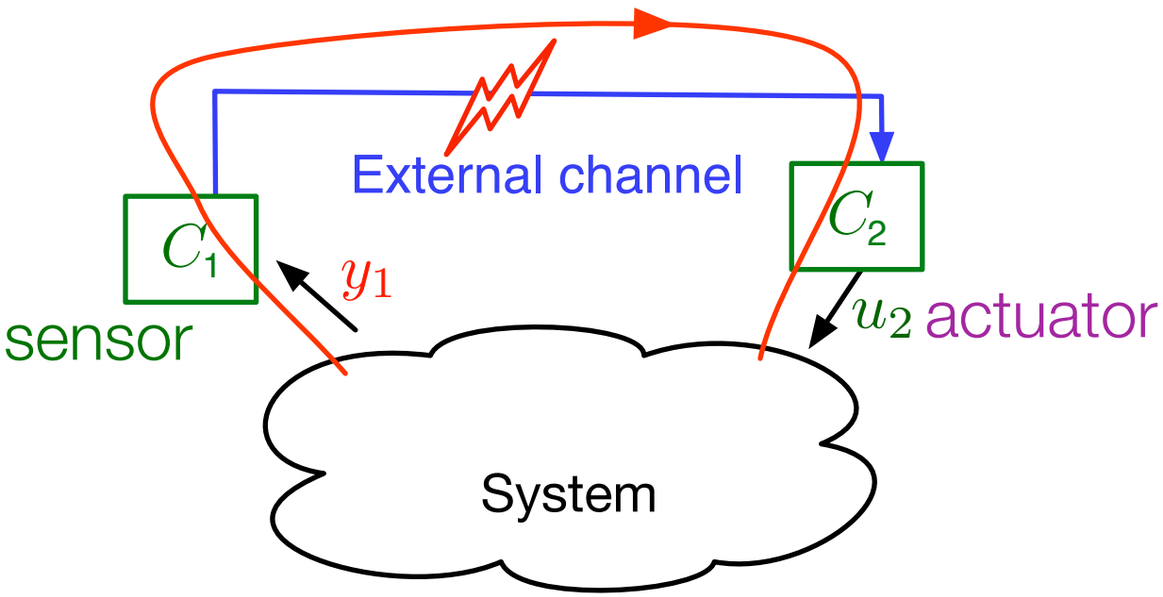
Looking forward:
Information-flows in Cyber-Physical Systems



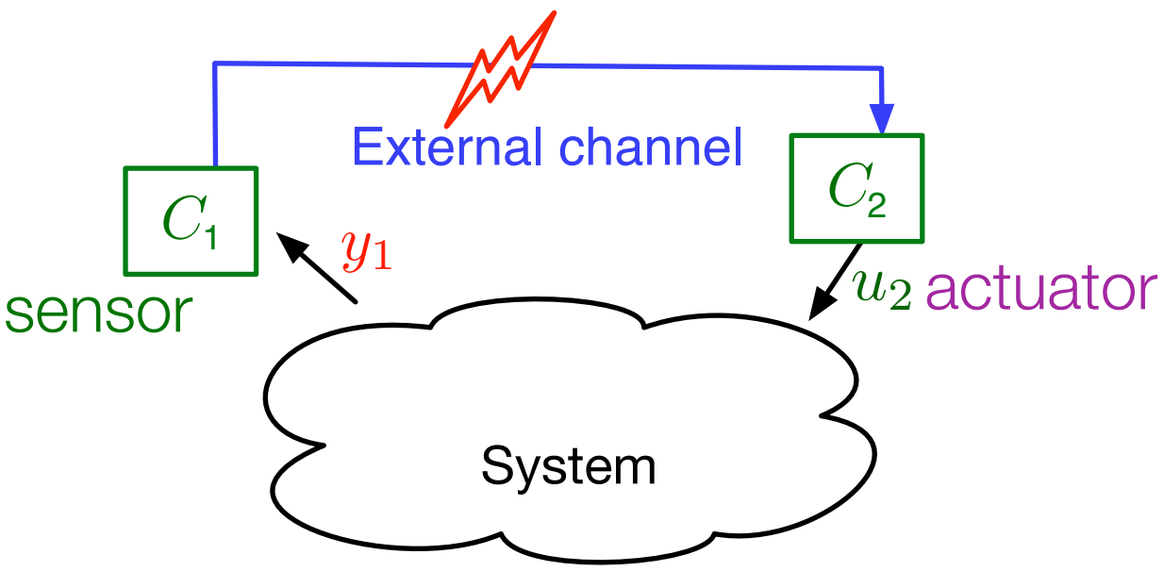
Looking forward:
Information-flows in Cyber-Physical Systems



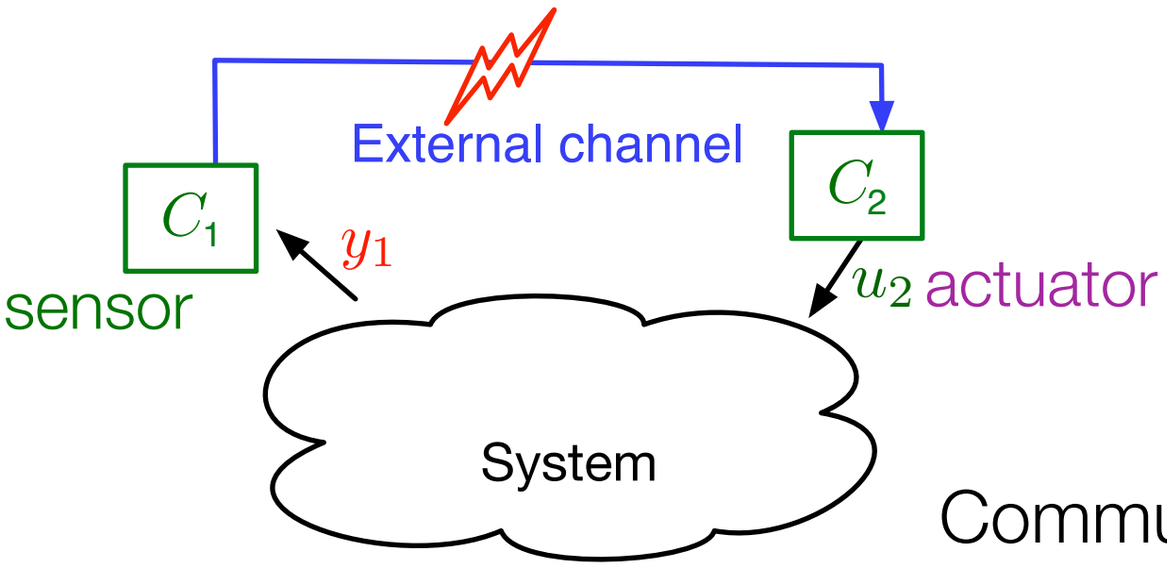
Looking forward:
Information-flows in Cyber-Physical Systems

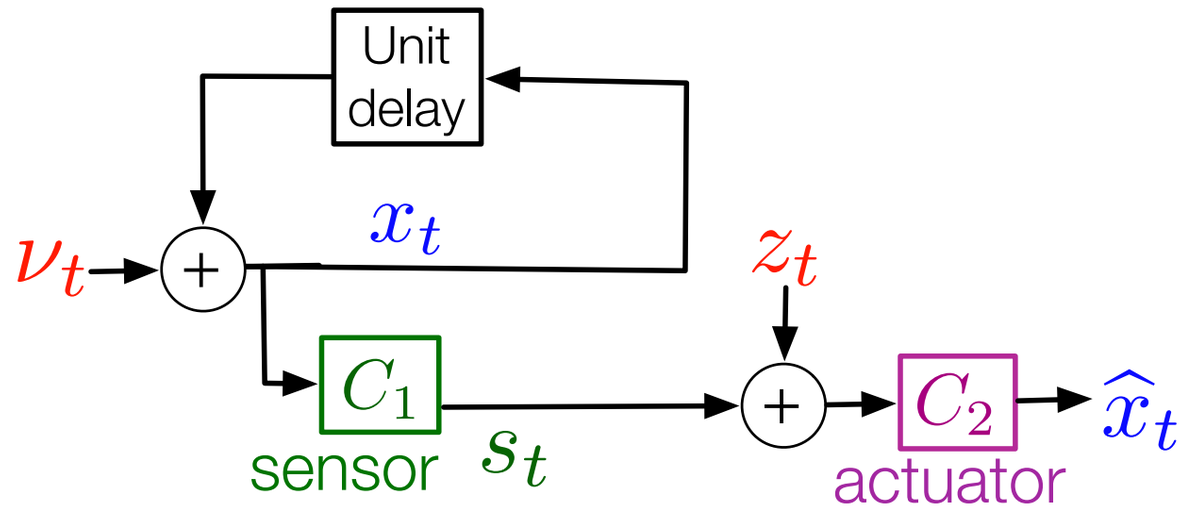


Looking forward:
Information-flows in Cyber-Physical Systems



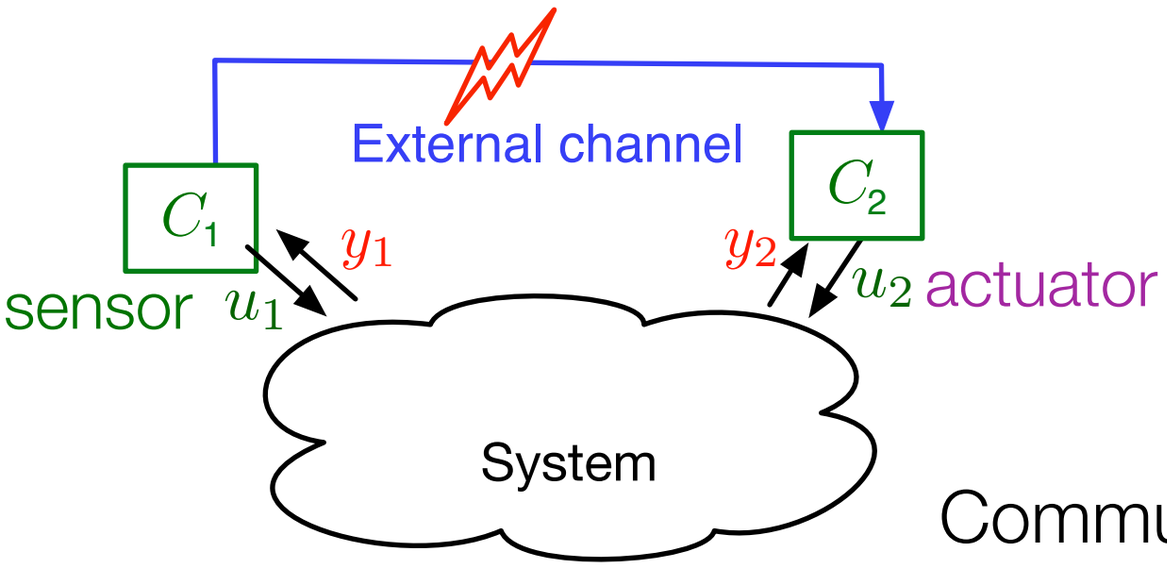
Looking forward:
Information-flows in Cyber-Physical Systems

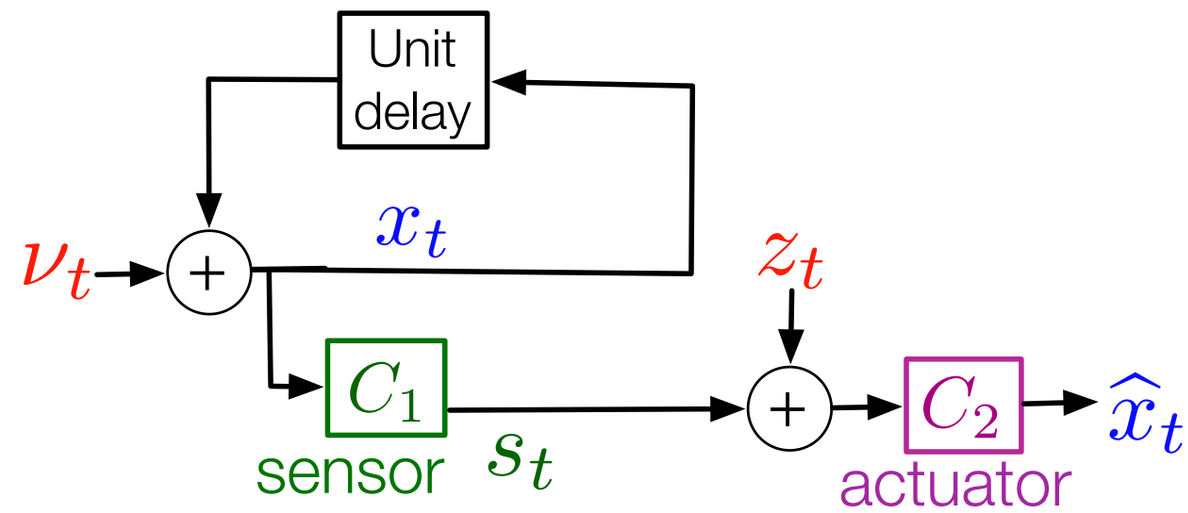




Communicating a evolving state
[Bansal, Basar '89][Wong, Brockett '90][Tatikonda, Mitter '01]
[Sahai, Mitter '06] [Sinopoli et al. '04] . . .

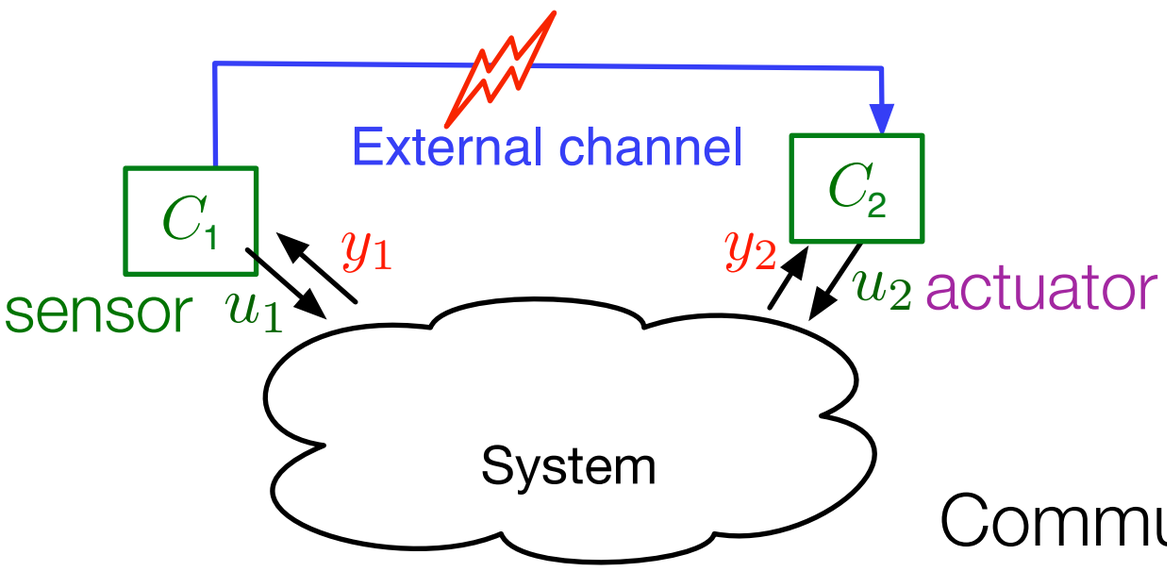
Looking forward:
Information-flows in Cyber-Physical Systems

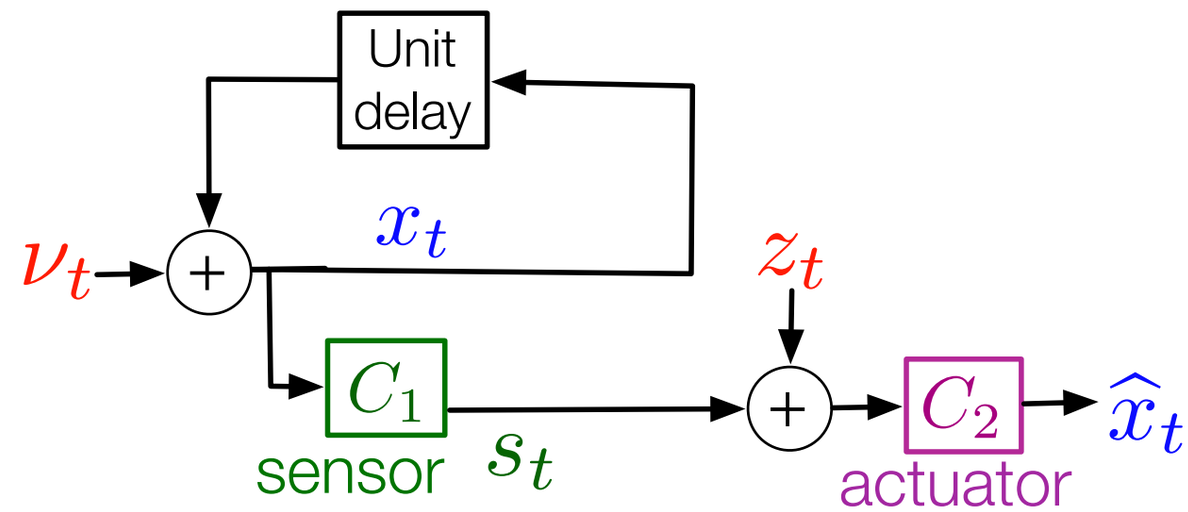




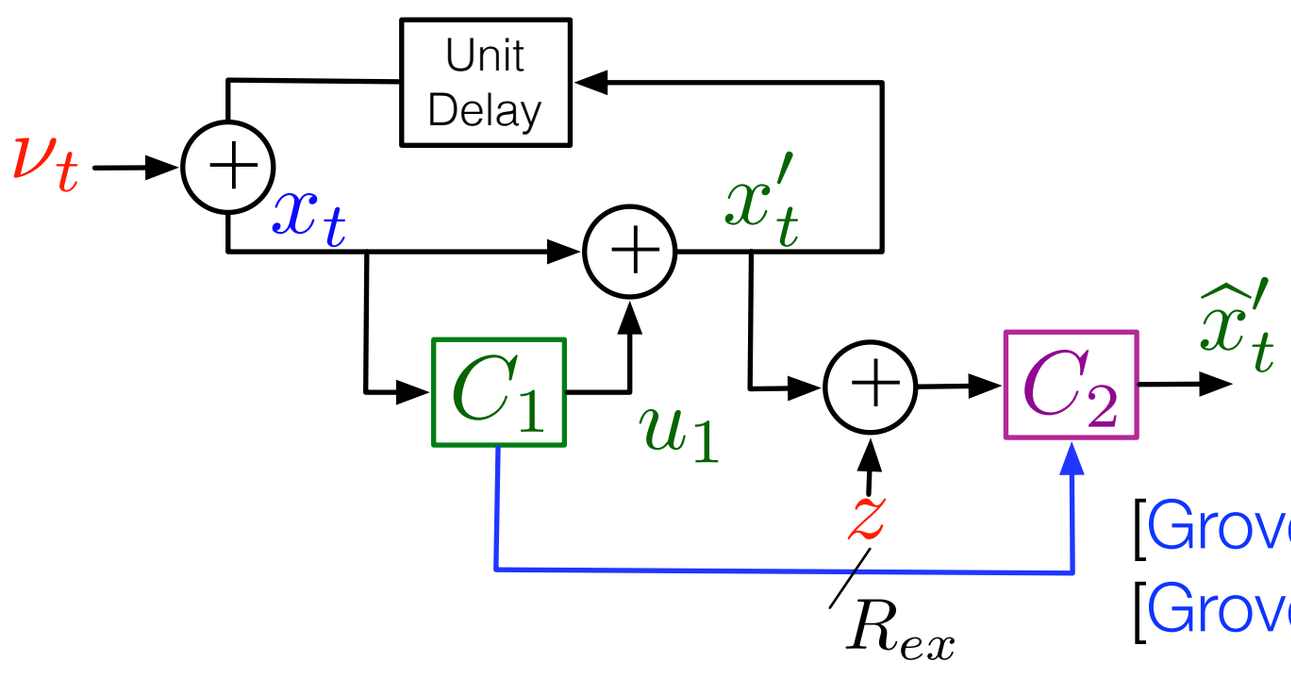
Communicating a evolving state
[Bansal, Basar '89][Wong, Brockett '90][Tatikonda, Mitter '01]
[Sahai, Mitter '06] [Sinopoli et al. '04] . . .

Looking forward:
Information-flows in Cyber-Physical Systems



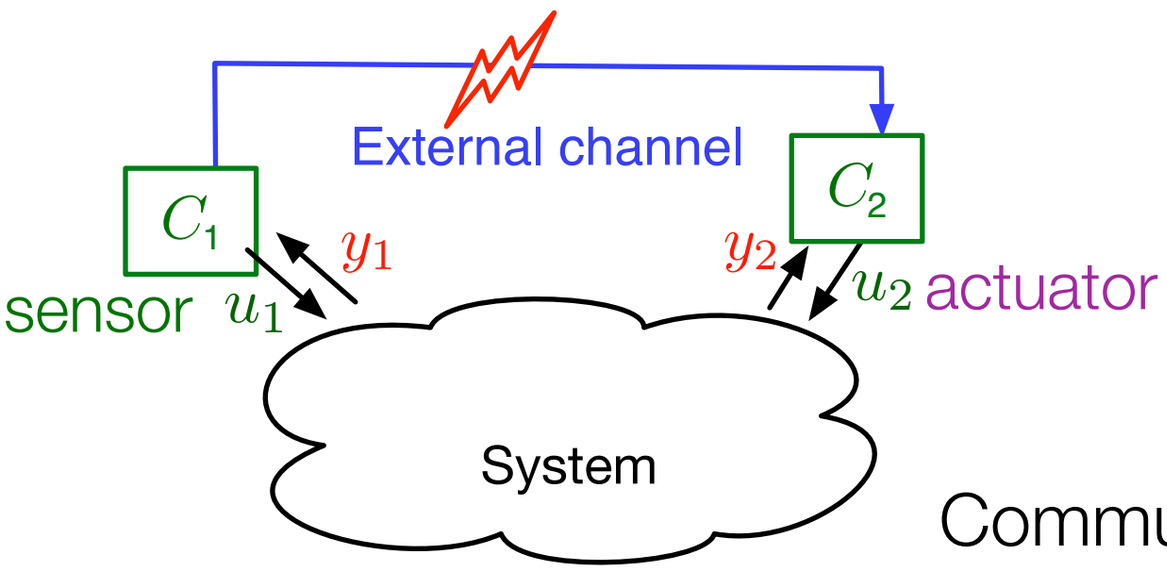


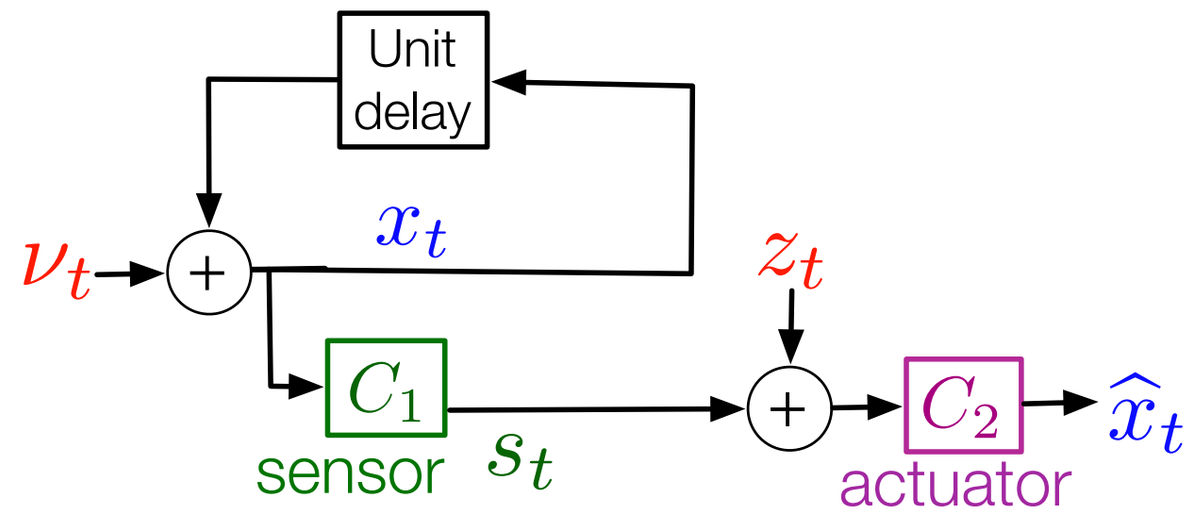
Communicating a evolving state
[Bansal, Basar '89][Wong, Brockett '90][Tatikonda, Mitter '01]
[Sahai, Mitter '06] [Sinopoli et al. '04] . . .



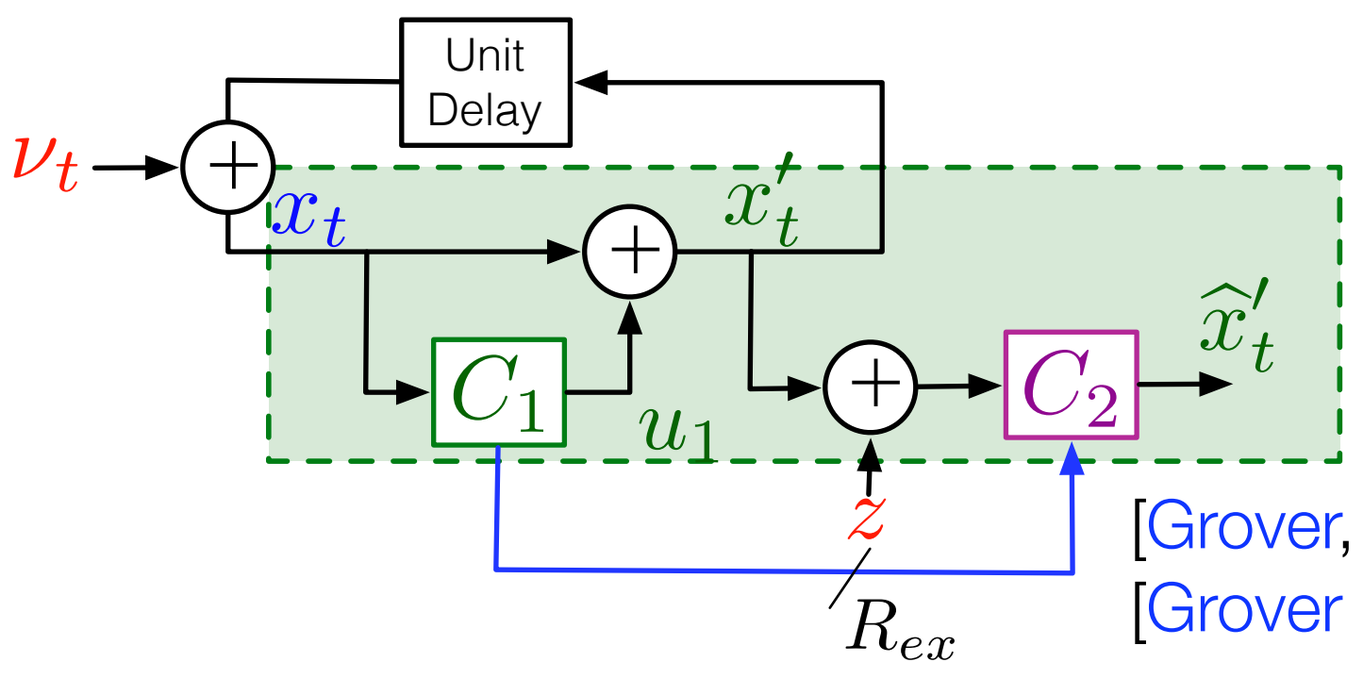
[Grover, Sahai '10]
[Grover '12]

Looking forward:
Information-flows in Cyber-Physical Systems





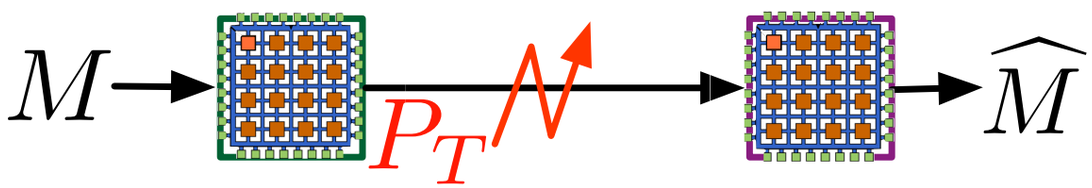
Communicating a evolving state
[Bansal, Basar '89][Wong, Brockett '90][Tatikonda, Mitter '01]
[Sahai, Mitter '06] [Sinopoli et al. '04] . . .

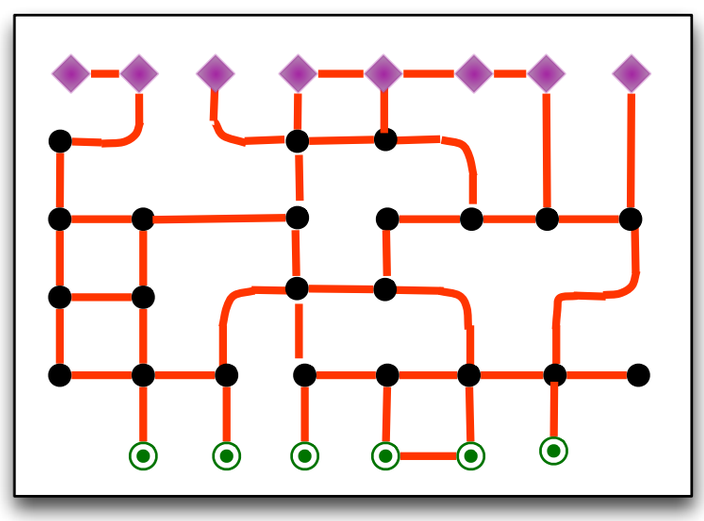


[Grover, Sahai '10]
[Grover '12]

Summary:
Understanding hidden information-flows is useful

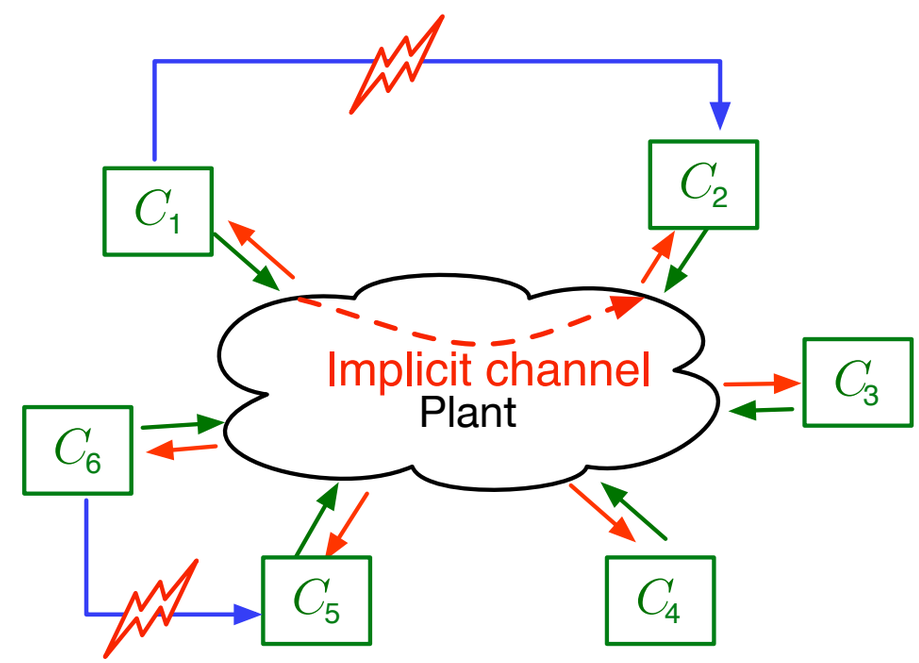
Information flows
in circuits





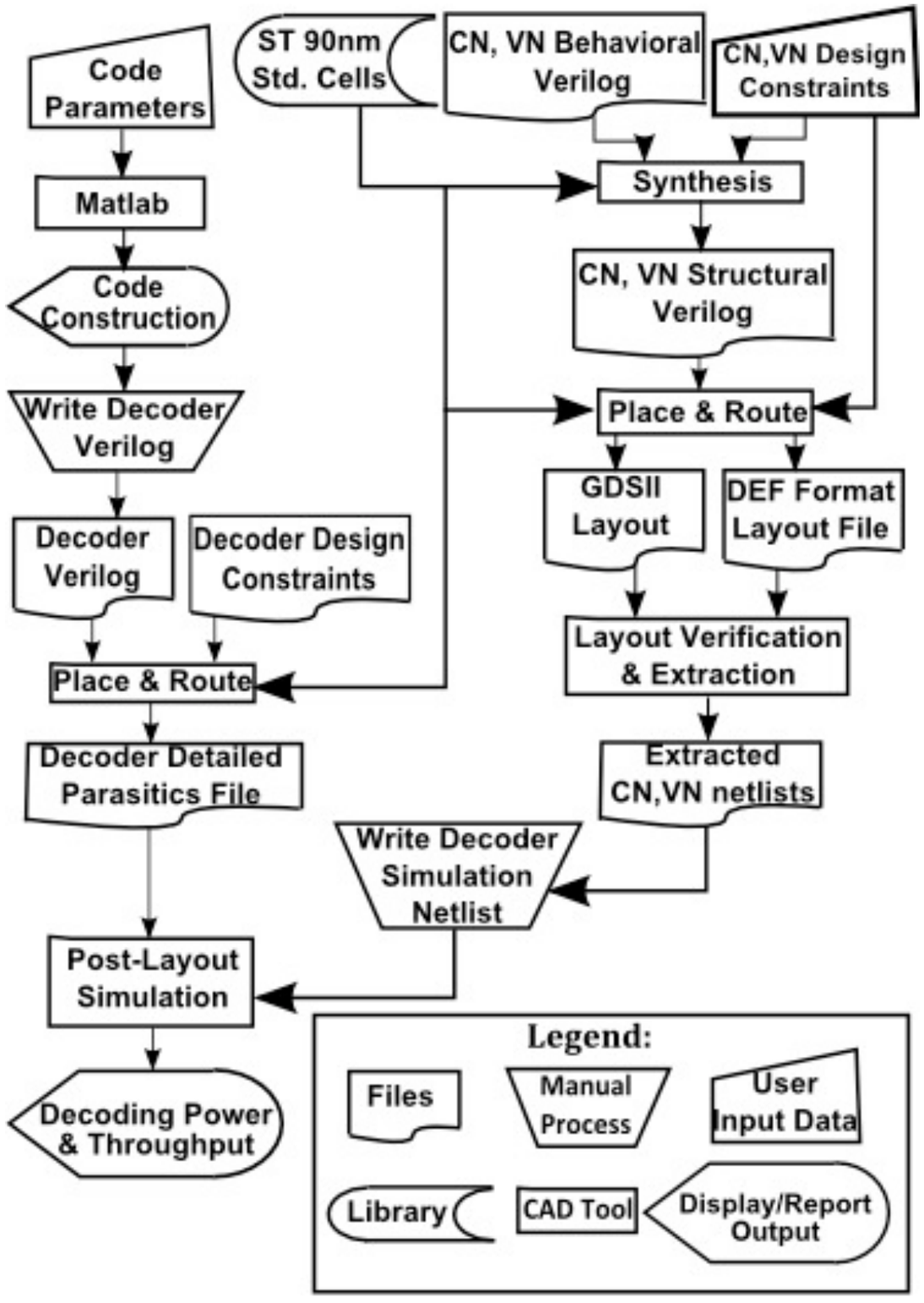
Information flow in computing:
fundamental bounds on total power

Information flows in
physical systems

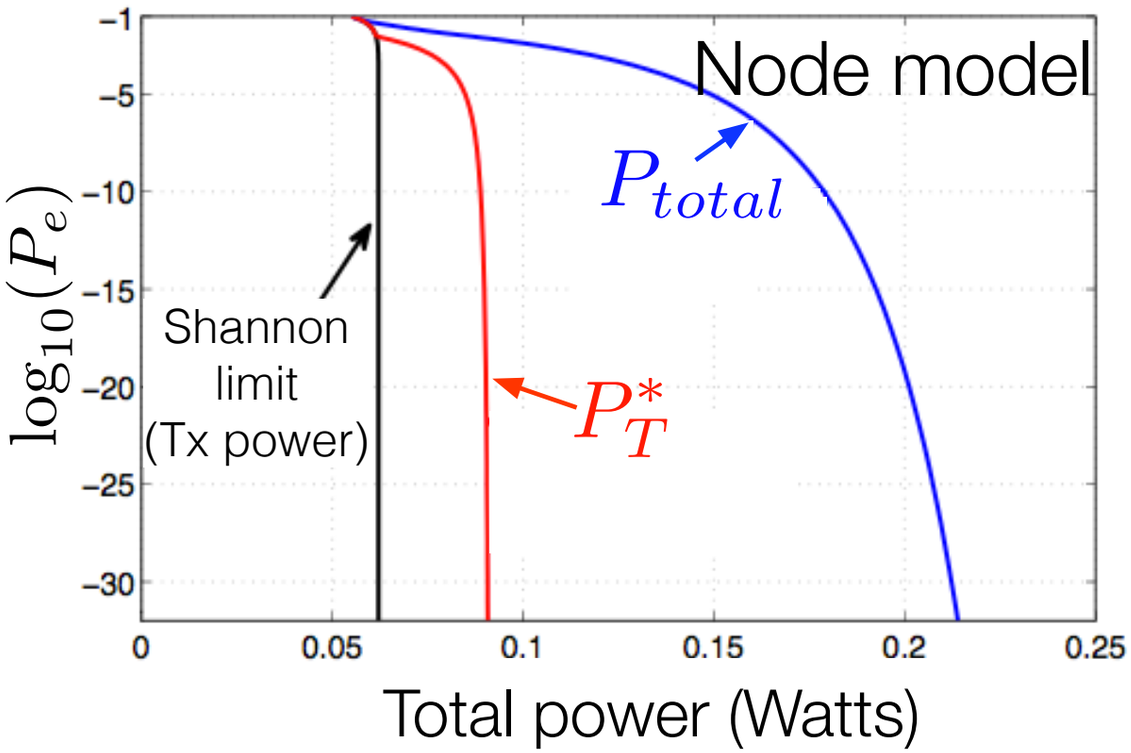


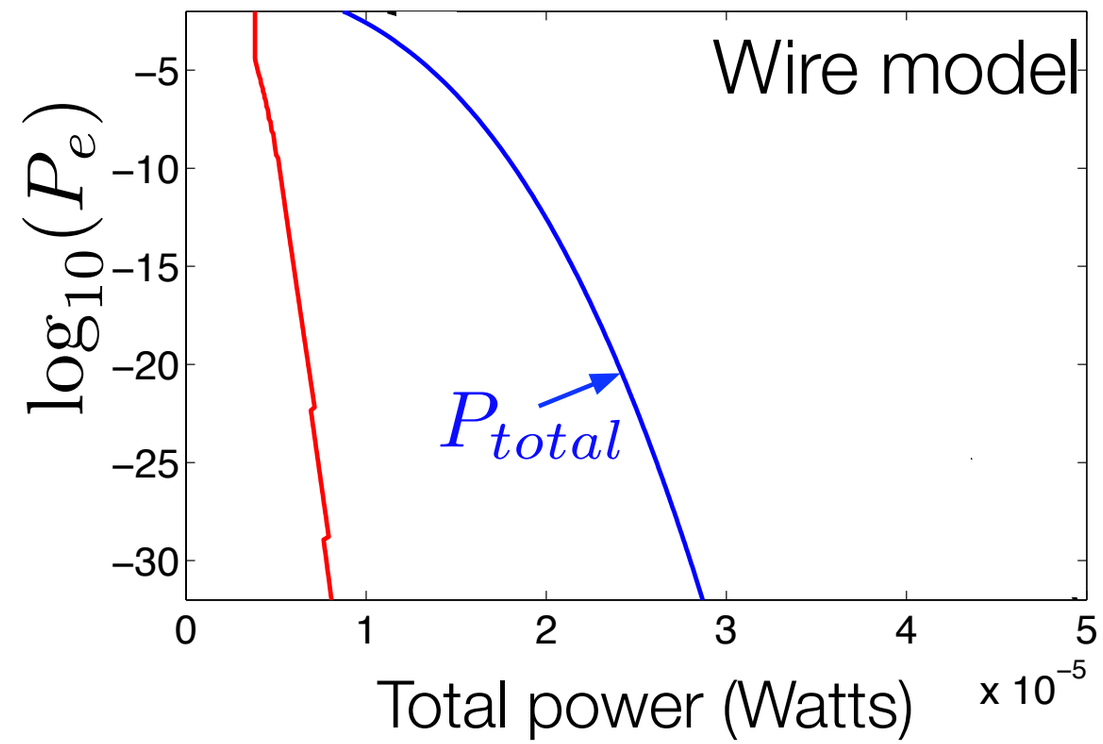
Implicit information flows:
approximately-optimal solutions

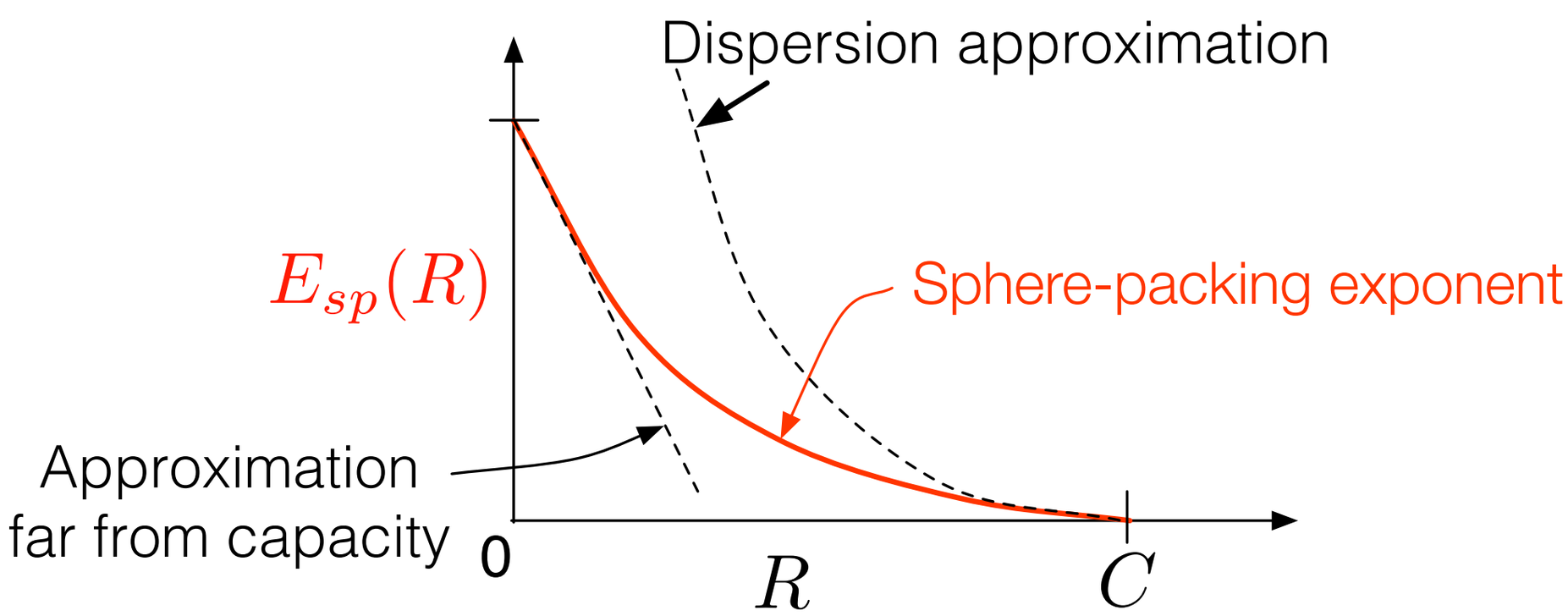
Backup slides


Code Parameters
ST 90nm Std. Cells
CN, VN Behavioral Verilog
CN,VN Design Constraints
Matlab
Synthesis
Code Construction
CN, VN Structural Verilog
Write Decoder Verilog
Place & Route
GDSII Layout
DEF Format Layout File
Decoder Verilog
Decoder Design Constraints
Layout Verification & Extraction
Place & Route
Decoder Detailed Parasitics File
Extracted CN,VN netlists
Write Decoder Simulation Netlist
Post-Layout Simulation
Decoding Power & Throughput
Legend:
Files
Manual Process
User Input Data
Library
CAD Tool
Display/Report Output

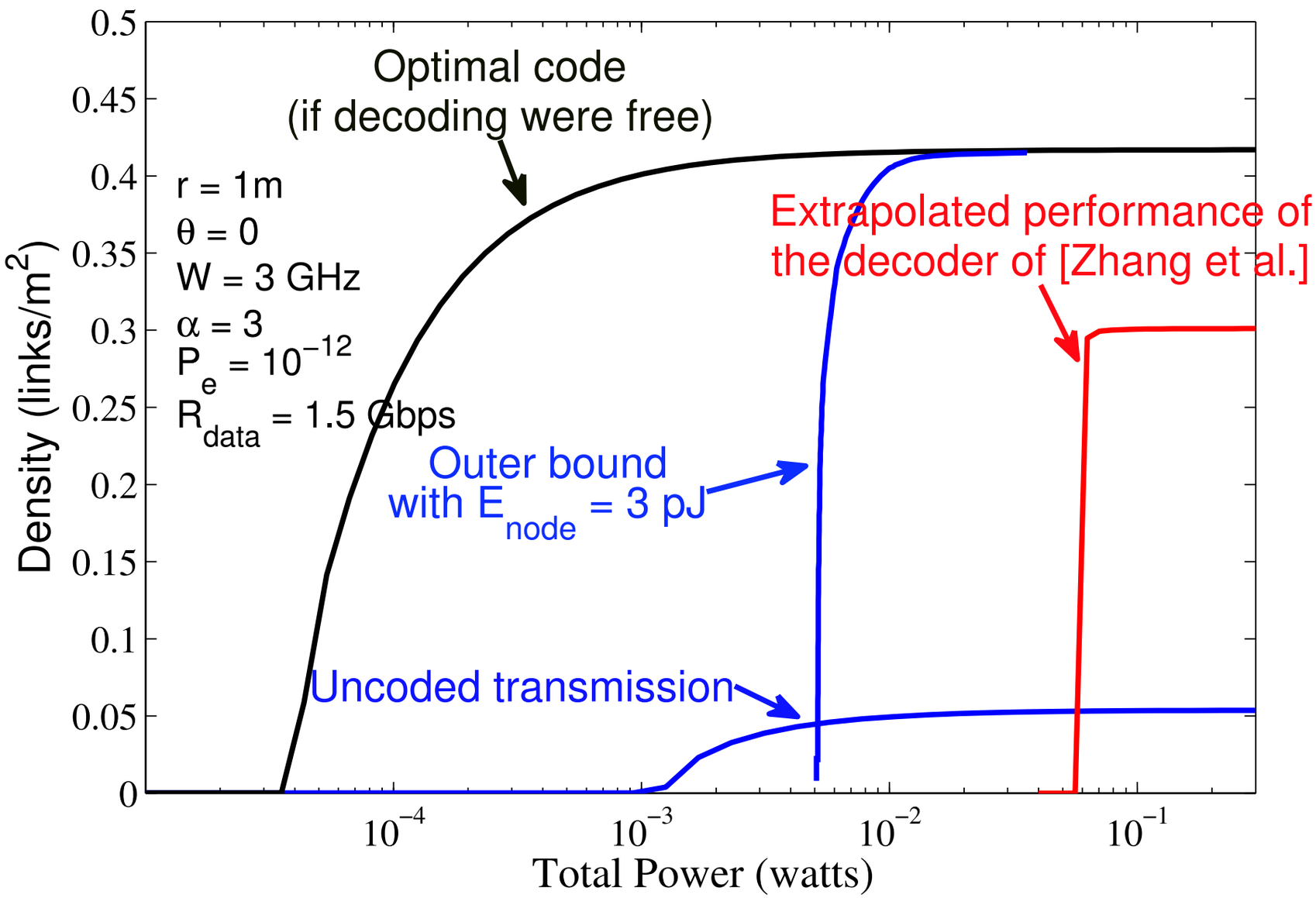
The appropriate approximations in the two bounds



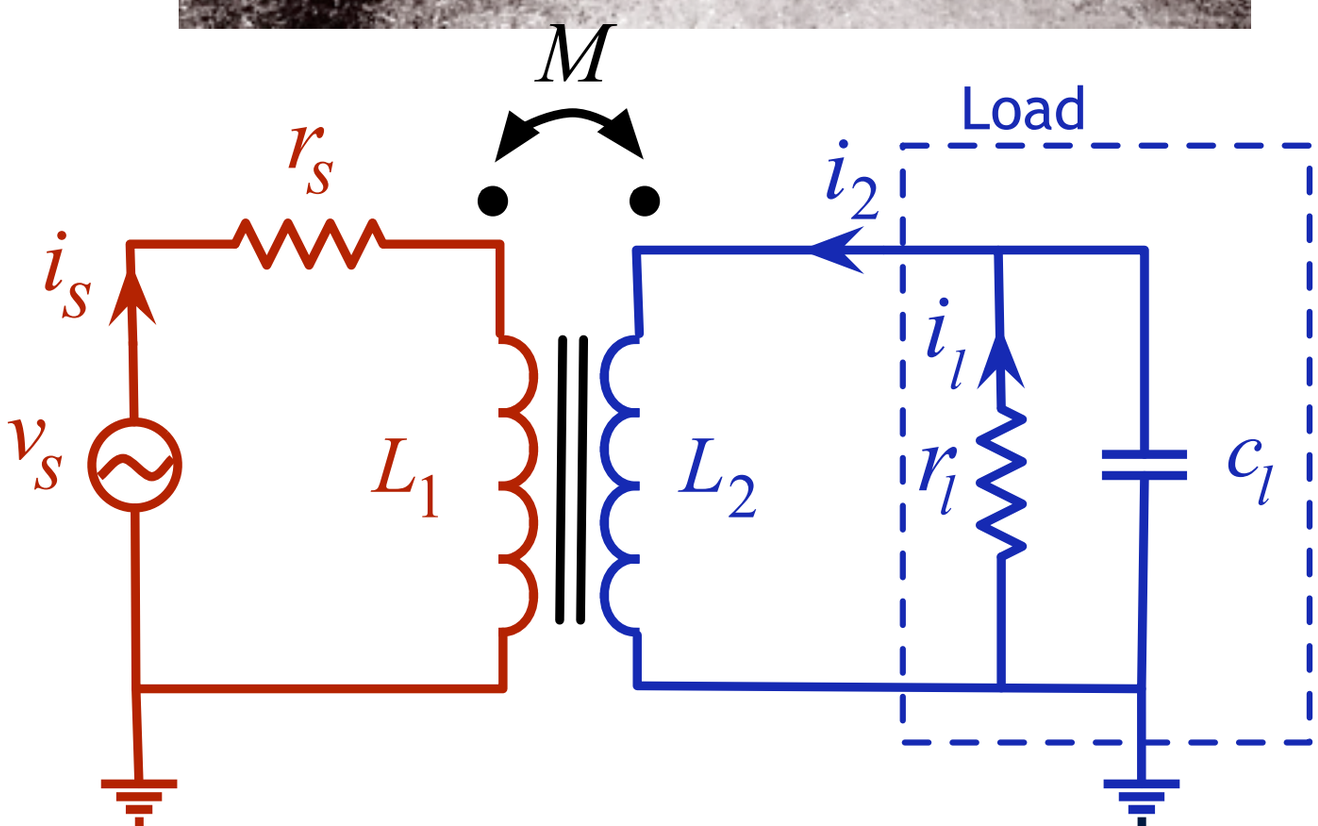




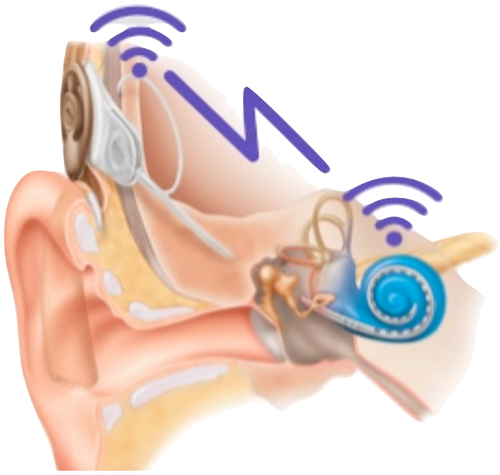
Total power minimizing density

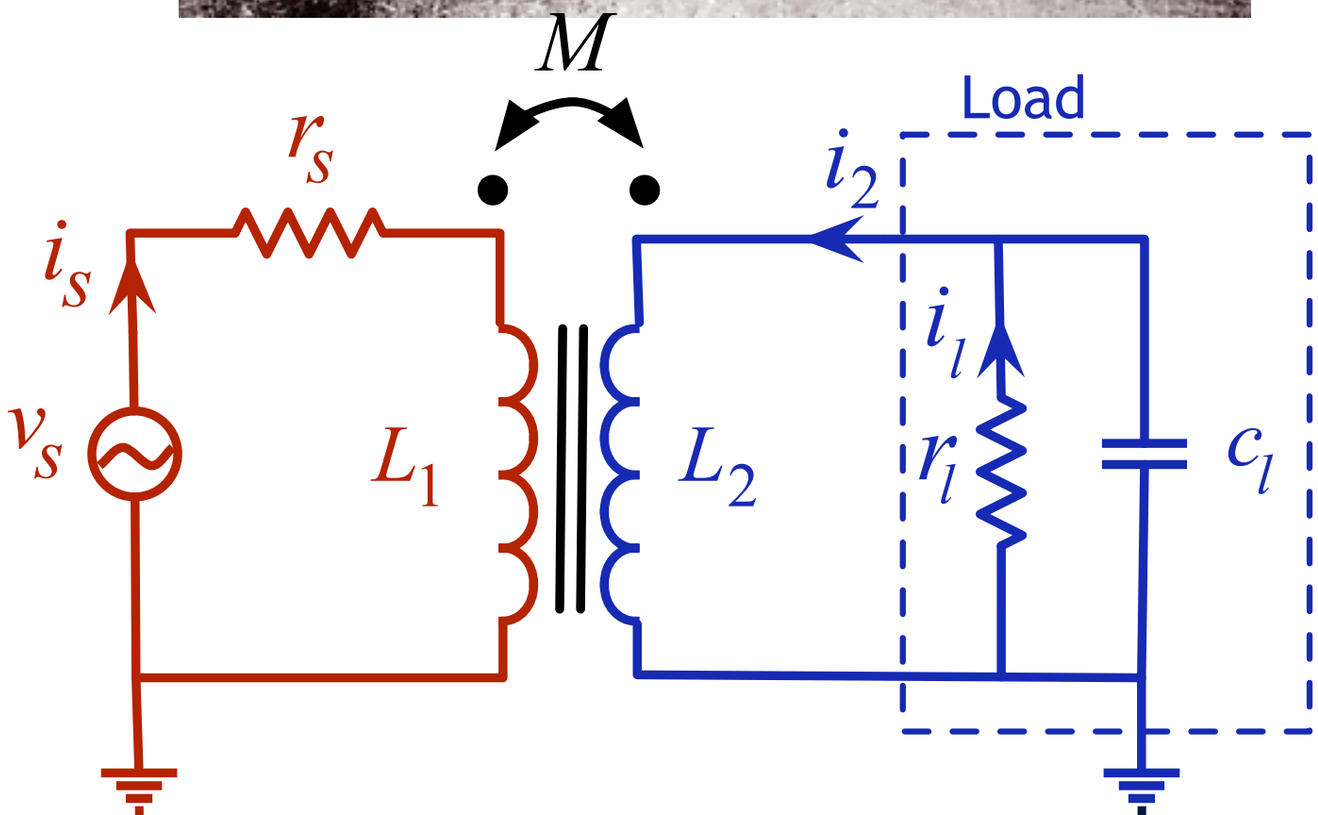


Wireless information transfer

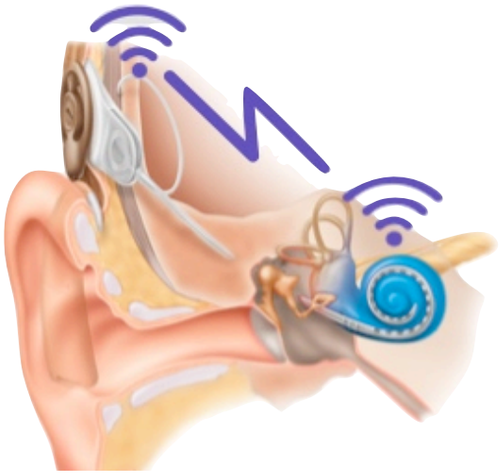


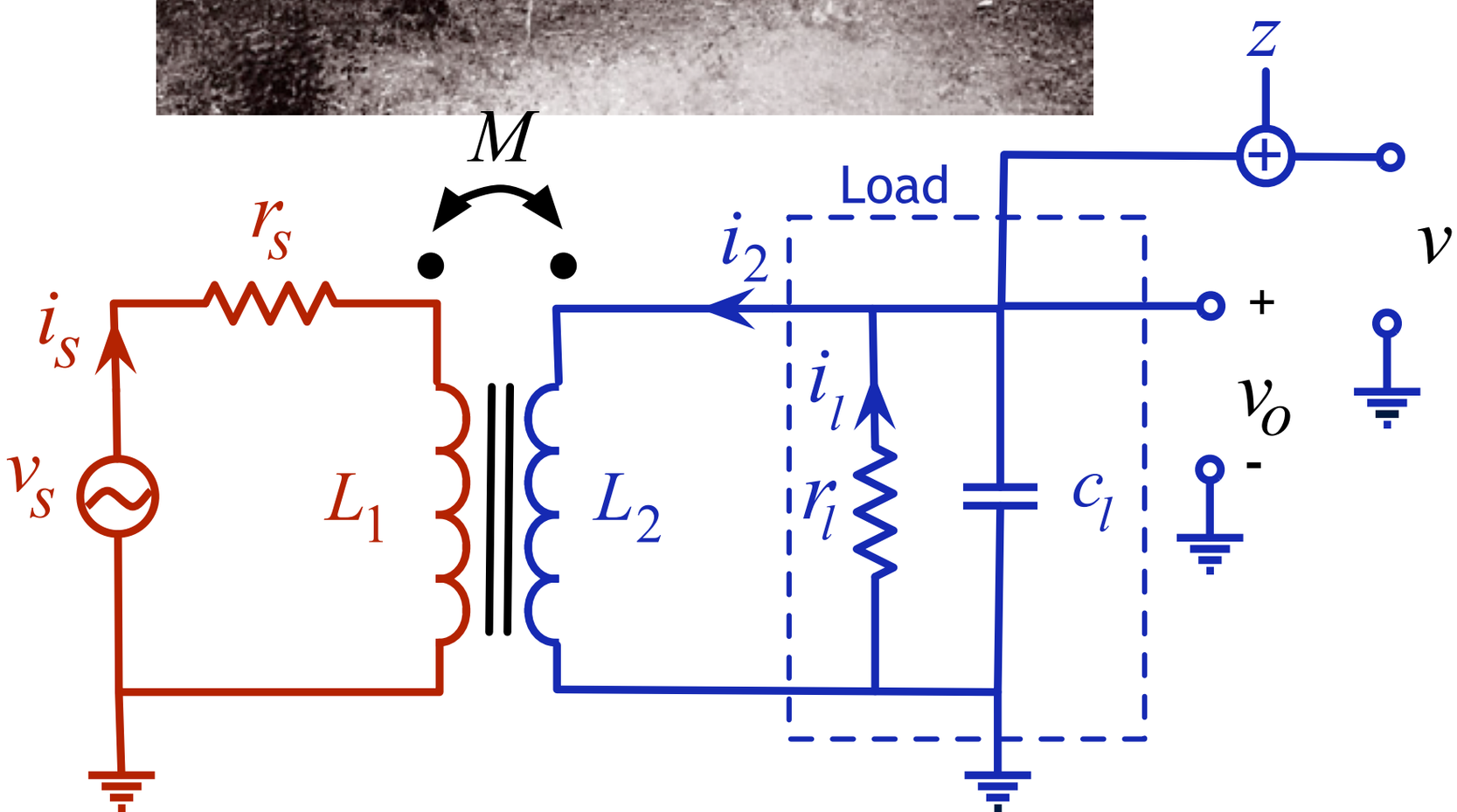
Wireless information transfer





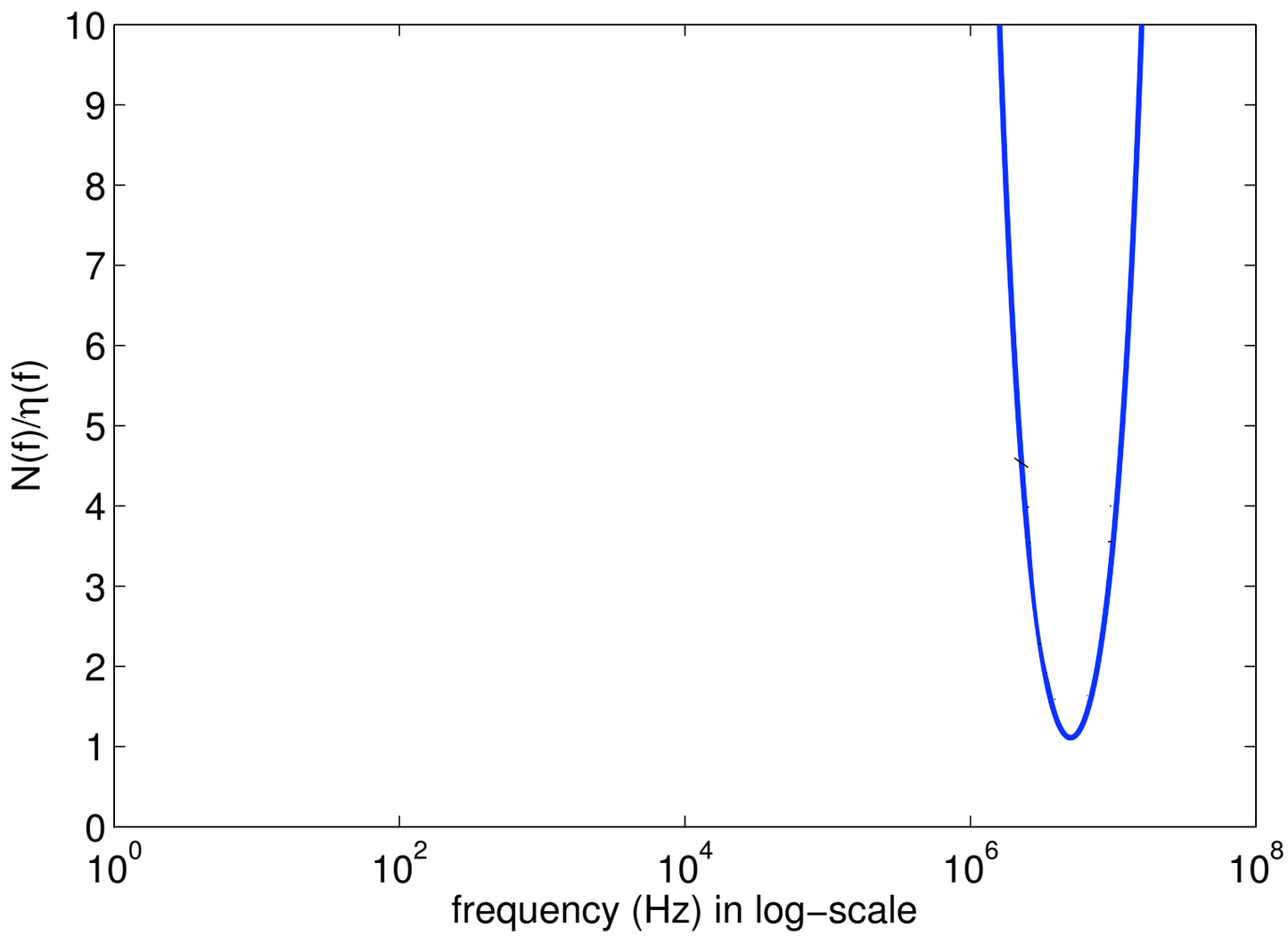
Wireless information transfer





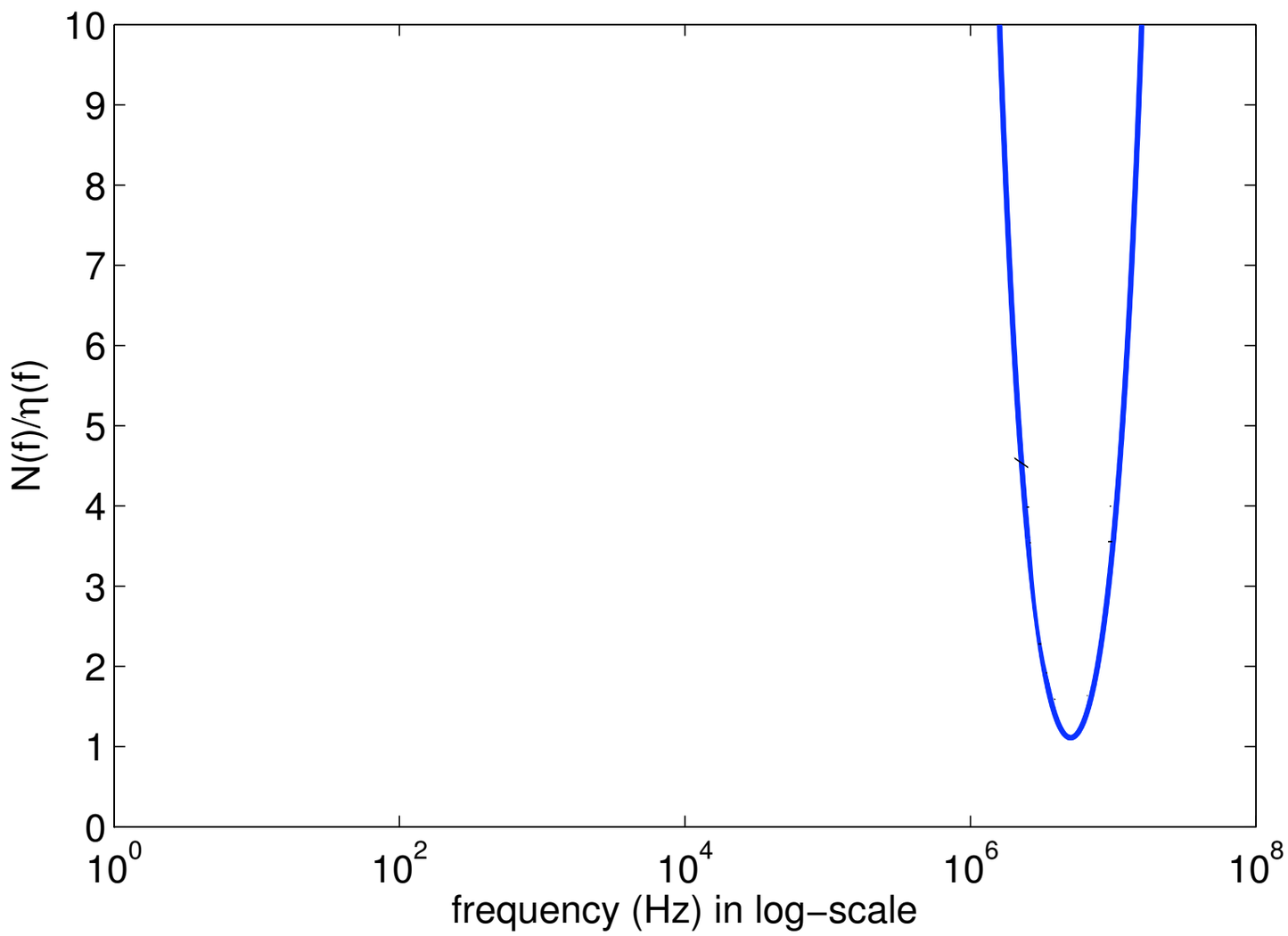
Optimal information transfer strategy?

Optimal information transfer strategy?



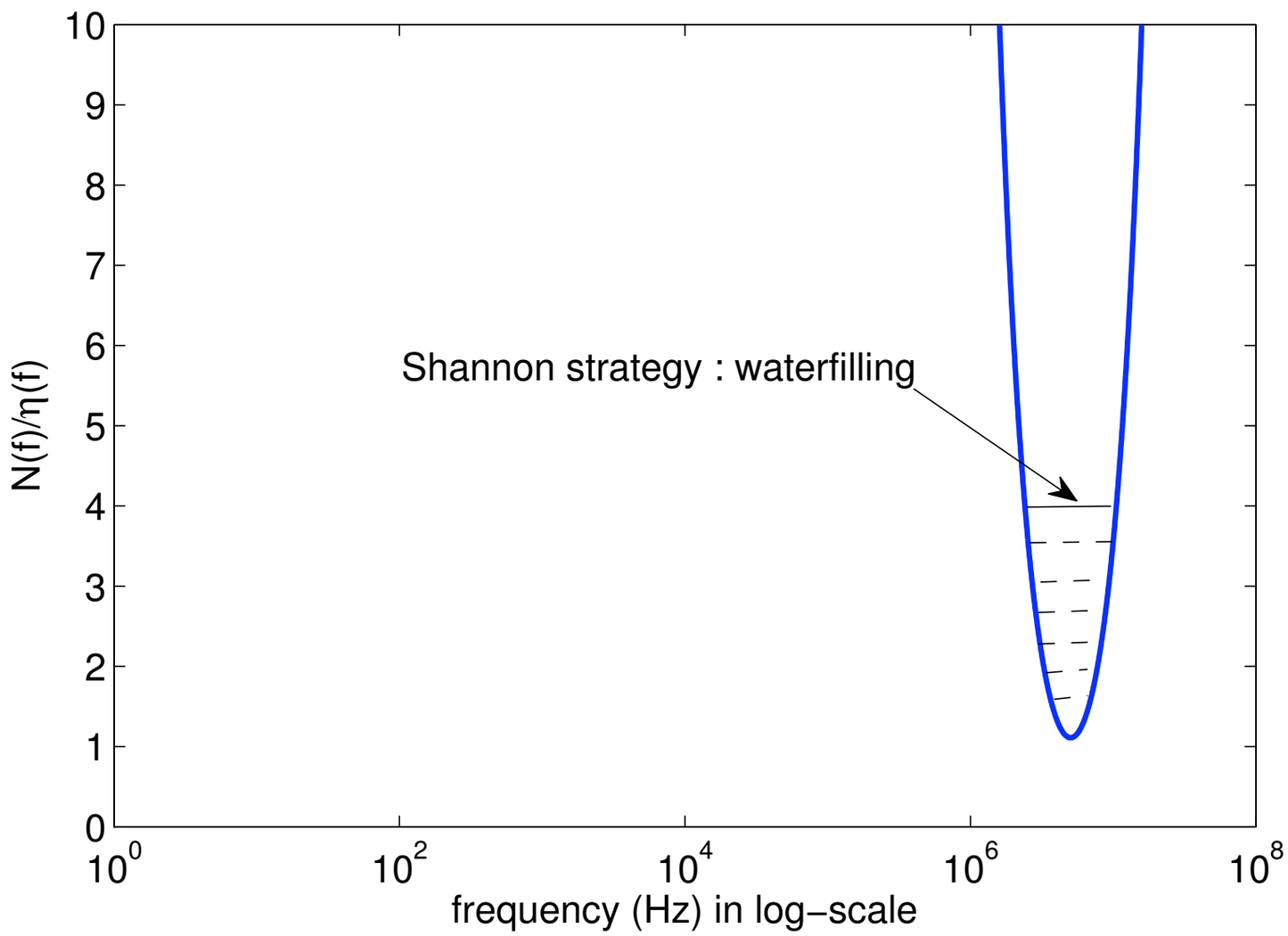
Optimal information transfer strategy?

$$C(P^{avail}) = \max_{\substack{P_1,\ldots,P_n \ s.t. \\ \sum_{i=1}^{n} P_i \leq P^{avail}}} \sum_{i=1}^{n} \log\left(1 + \frac{\eta_i P_i}{N}\right).$$



Optimal information transfer strategy?

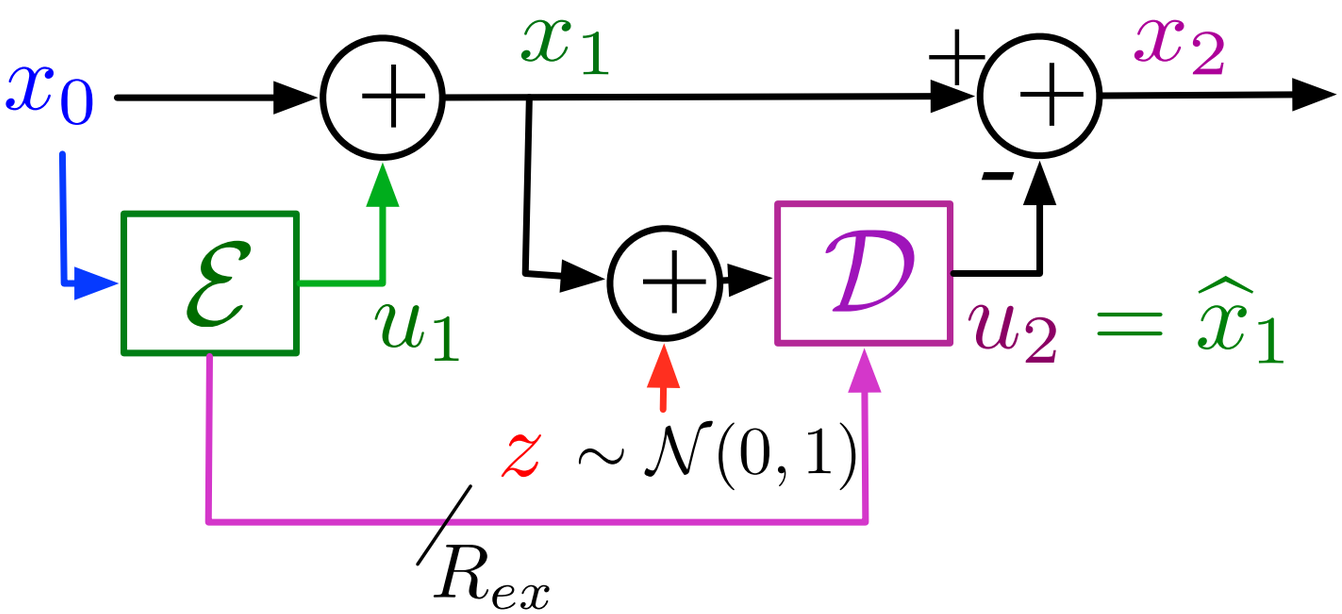
$$C(P^{avail}) = \max_{\substack{P_1,\ldots,P_n \ s.t. \\ \sum_{i=1}^{n} P_i \leq P^{avail}}} \sum_{i=1}^{n} \log\left(1 + \frac{\eta_i P_i}{N}\right).$$

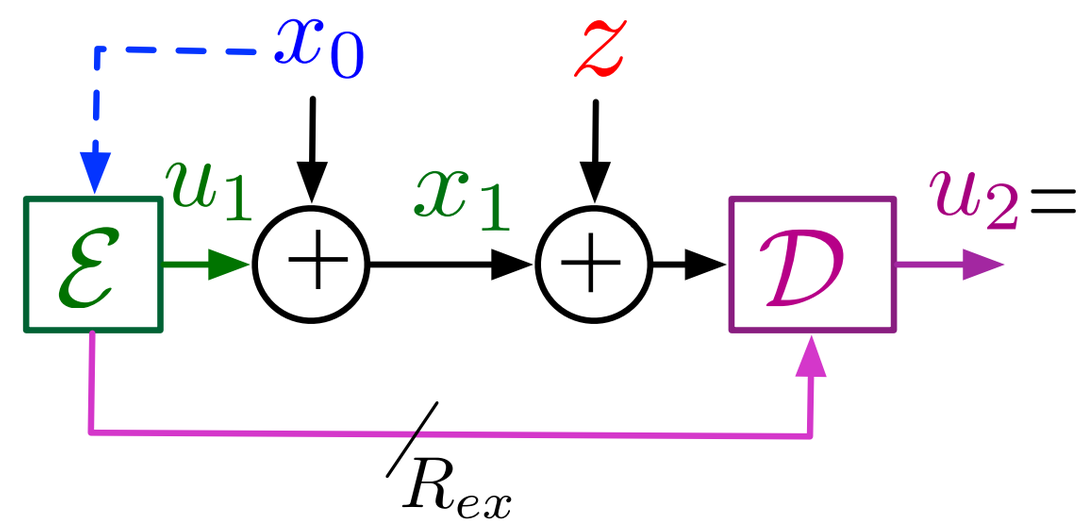


Secure Communication Using Inductive Coupling

Inductive coupling	**Quantum entanglement**
Measurement process causes detuning	Measurement process causes waveform collapse
Eavesdropper detection through spectrum sensing	Eavesdropper detection through physics
Information destruction through capacity-approaching codes and strong converse of channel coding theorem	Information destruction through physics
Inductive key distribution	Quantum key distribution

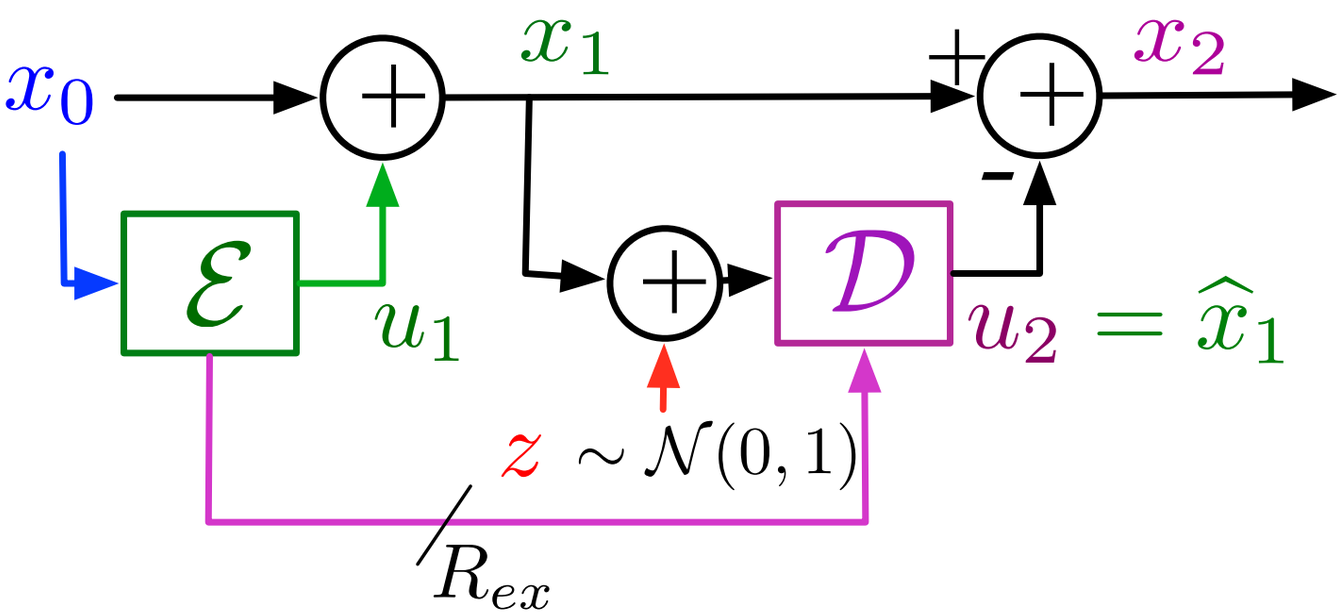
A deterministic abstraction

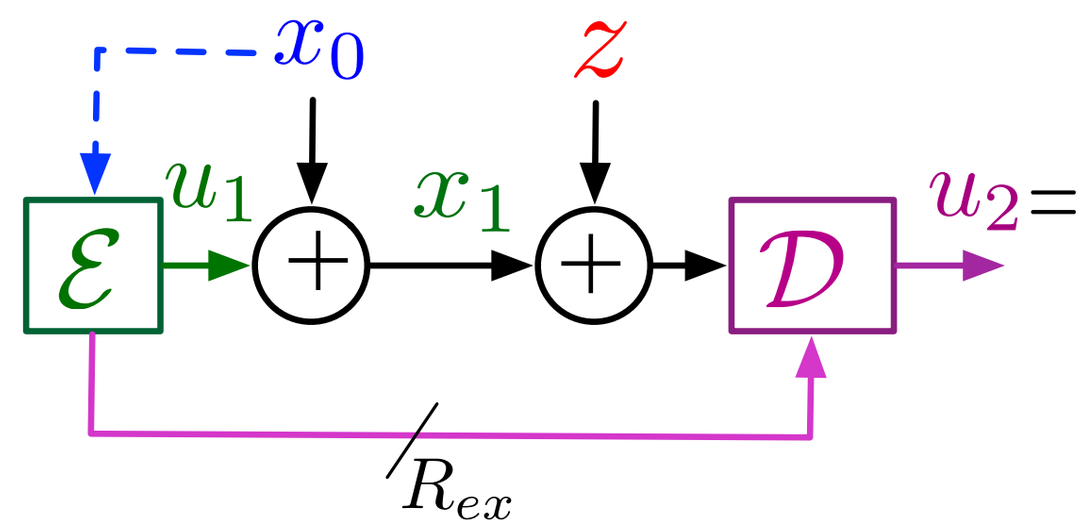




$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[(x_1 - \widehat{x}_1)^2\right]$$

A deterministic abstraction

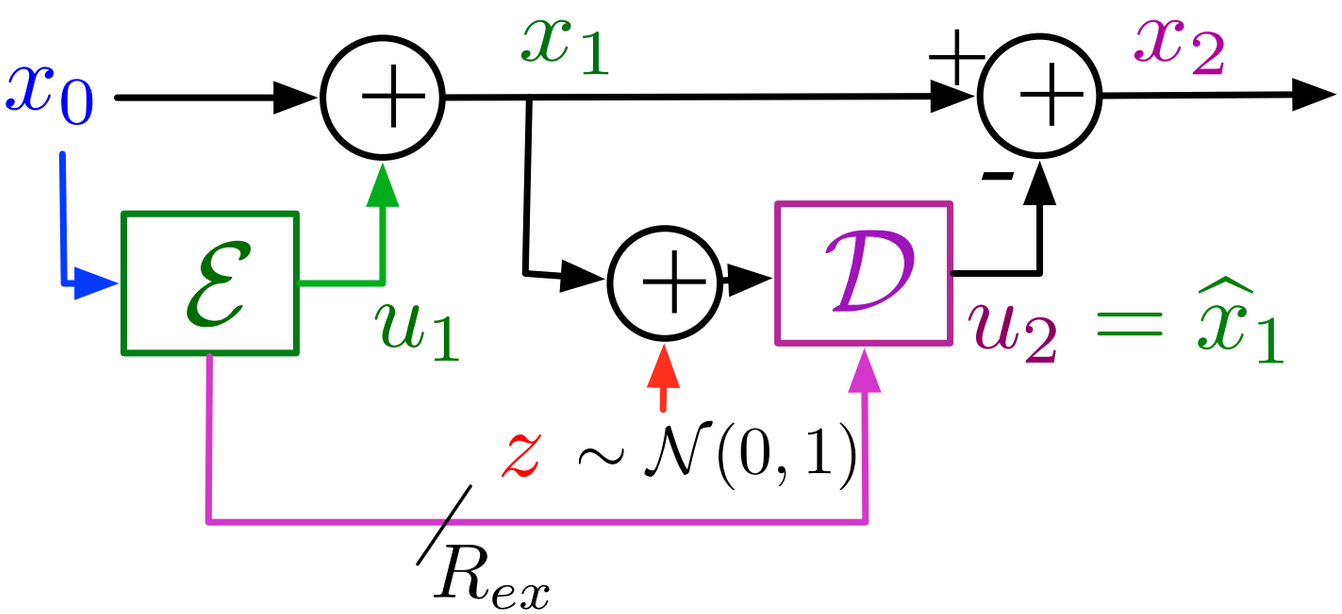


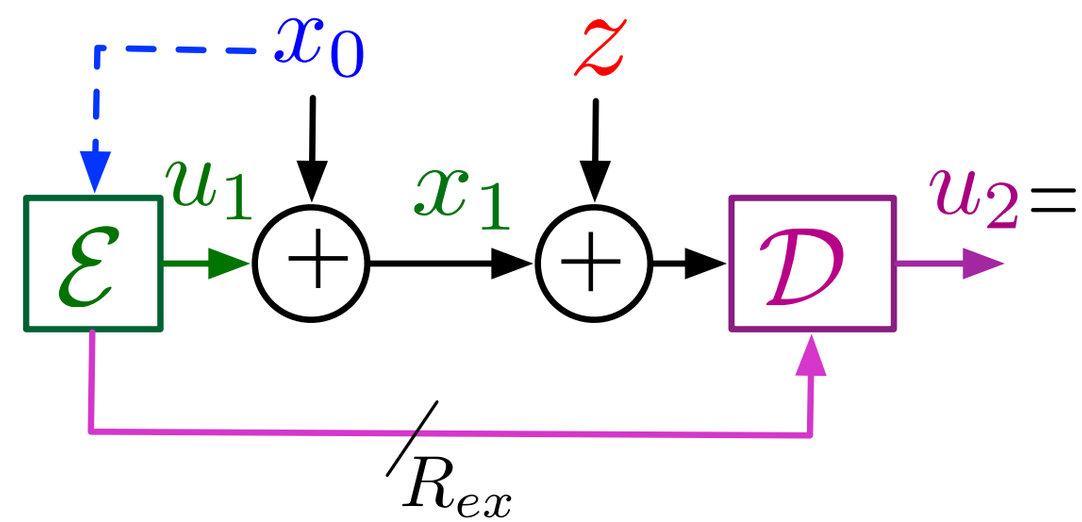


$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[(x_1 - \widehat{x}_1)^2\right]$$

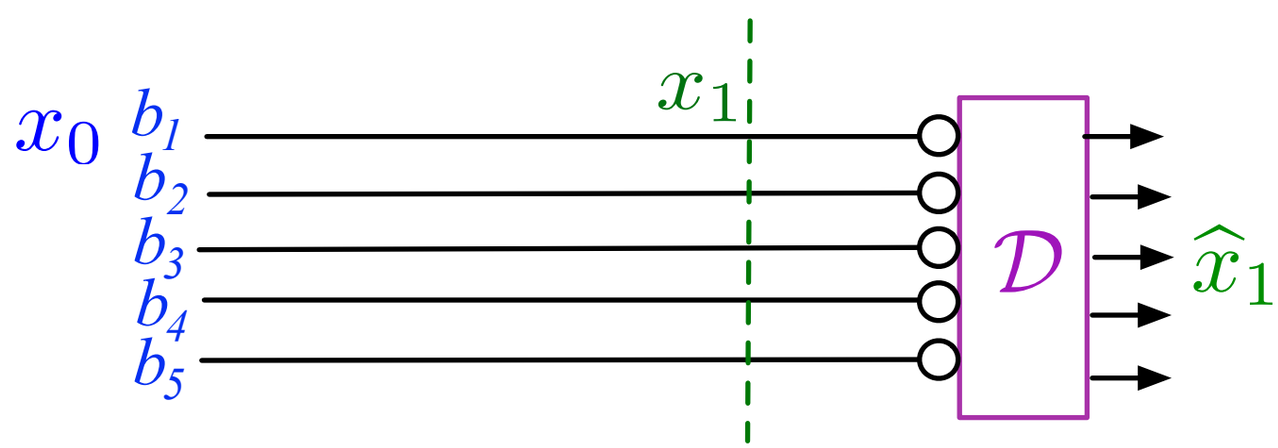
x_0 b_1
b_2
b_3
b_4
b_5

A deterministic abstraction

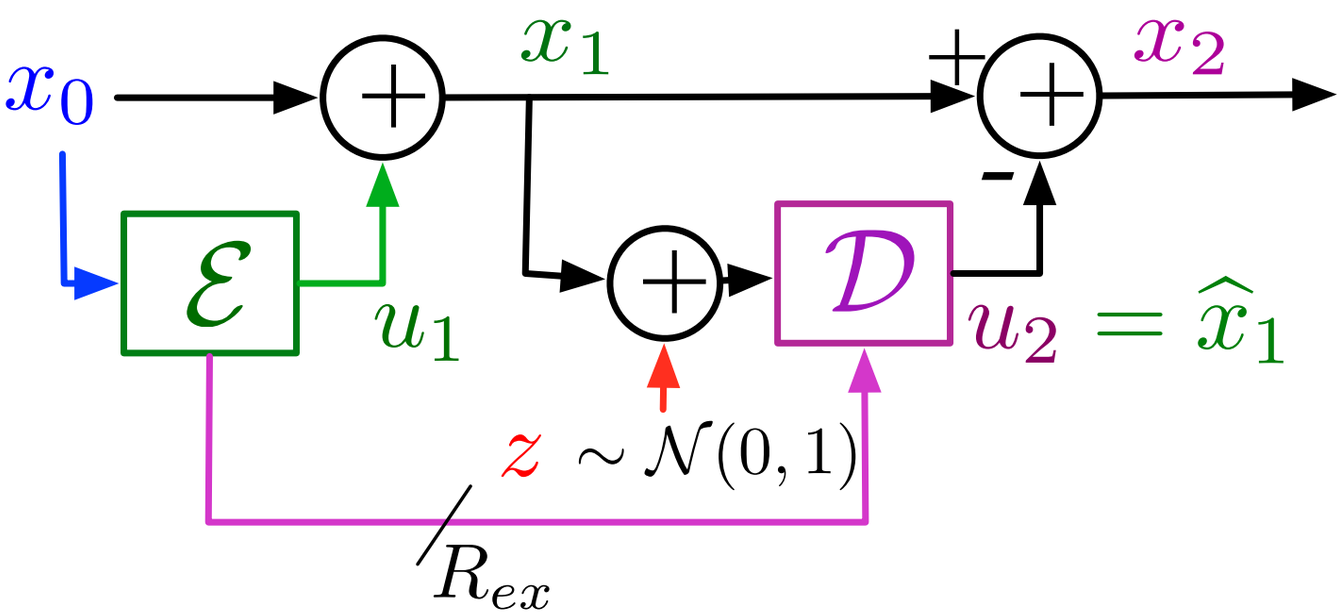


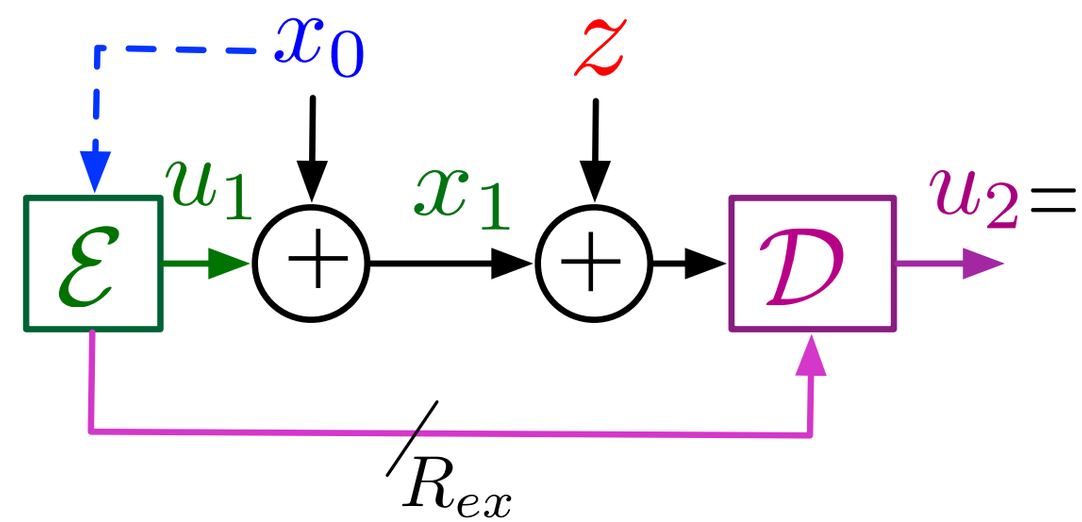


$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[(x_1 - \widehat{x}_1)^2\right]$$

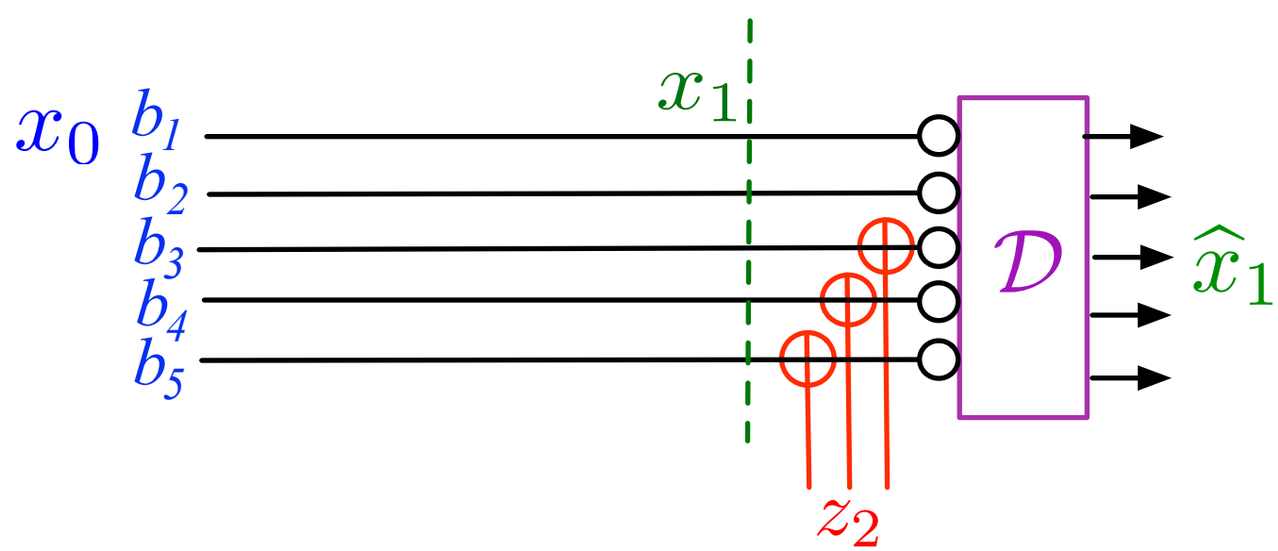


A deterministic abstraction

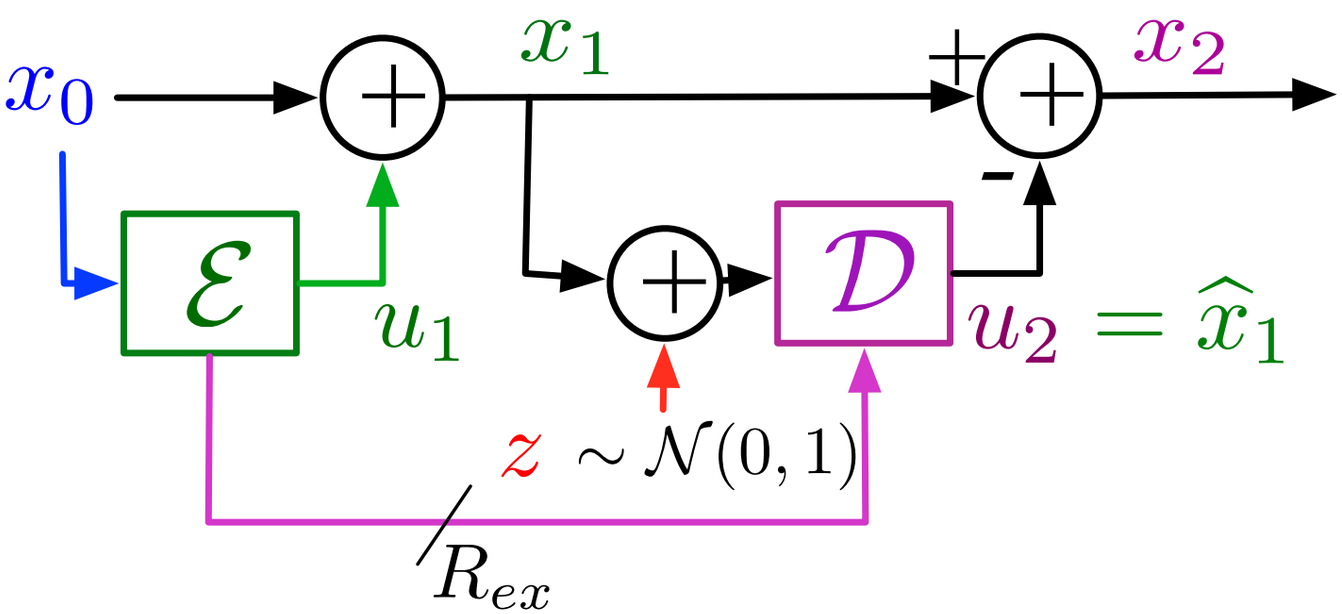


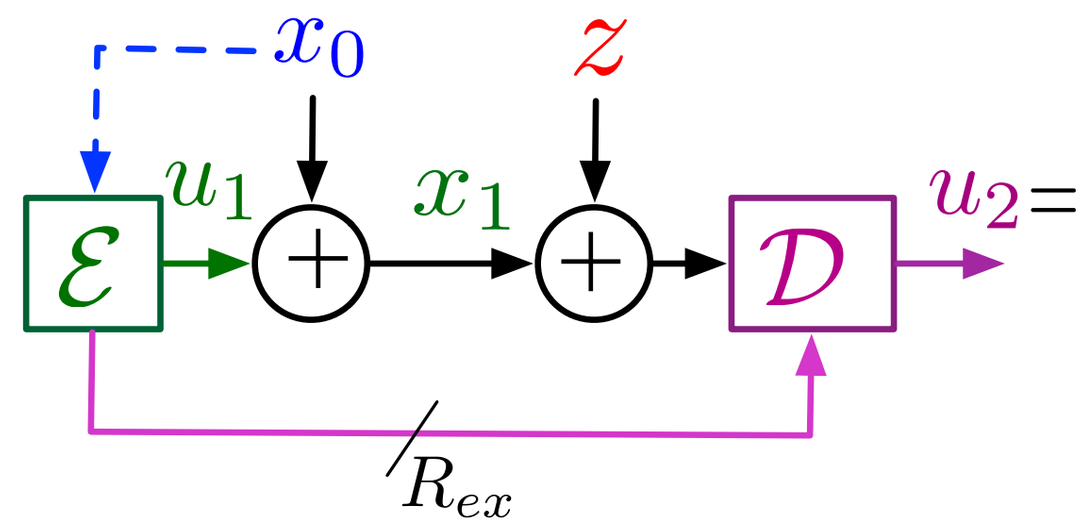


$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[(x_1 - \widehat{x}_1)^2\right]$$

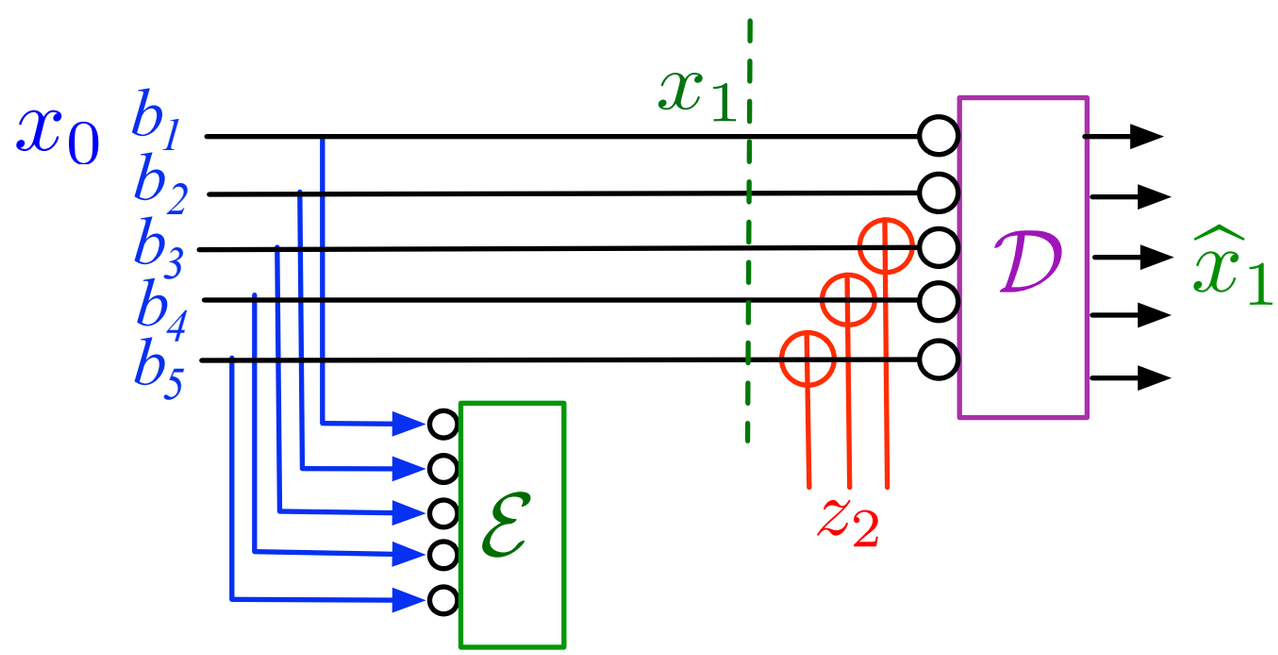


A deterministic abstraction

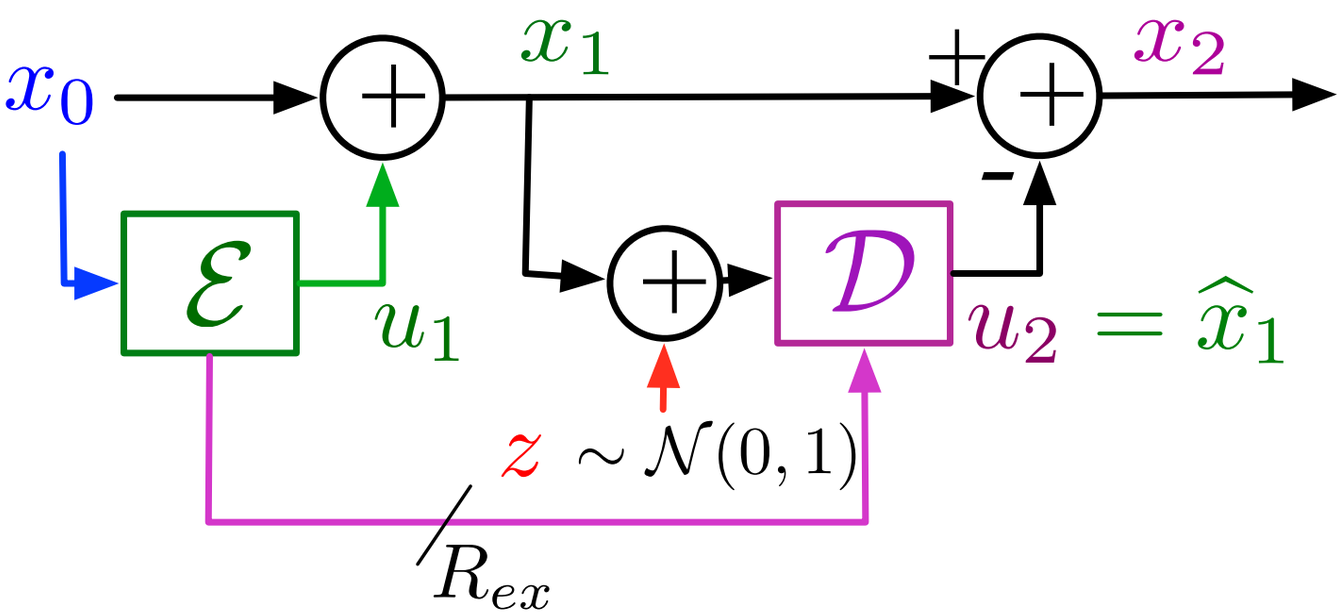


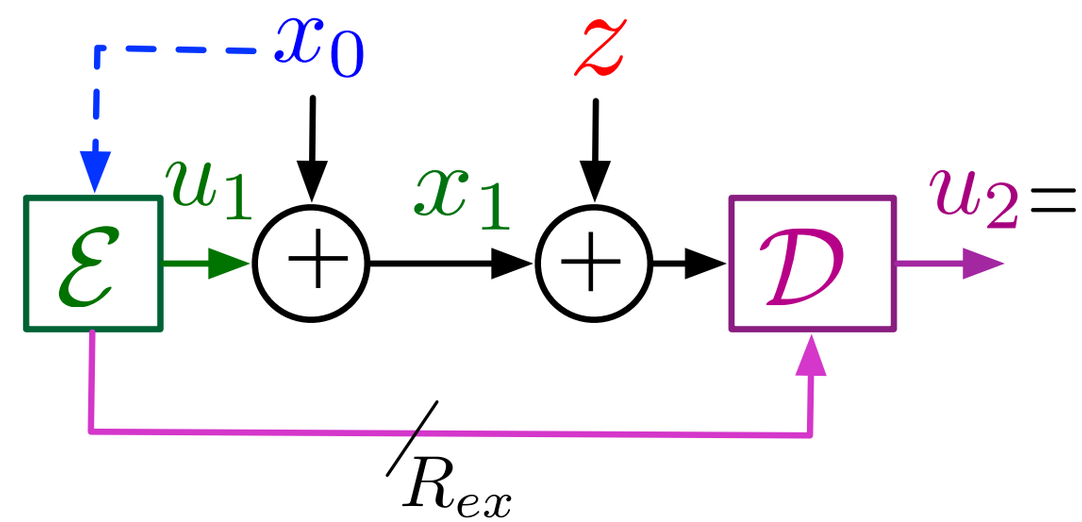


$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[(x_1 - \widehat{x}_1)^2\right]$$

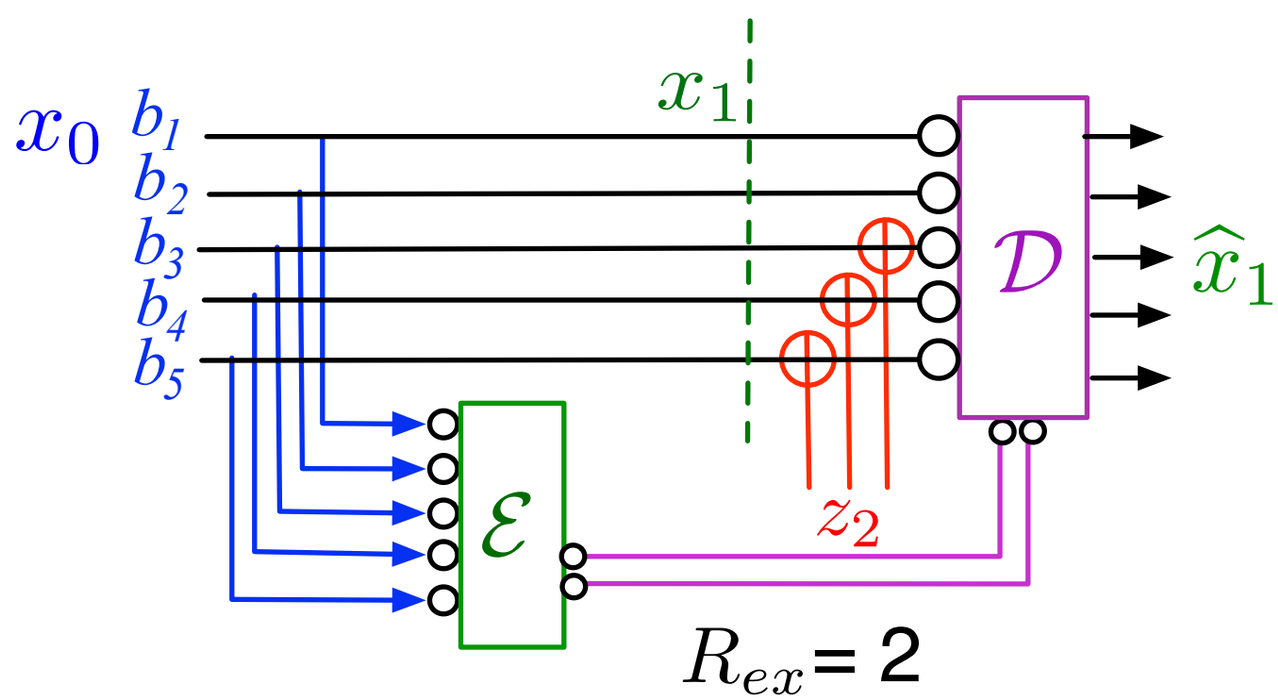


A deterministic abstraction

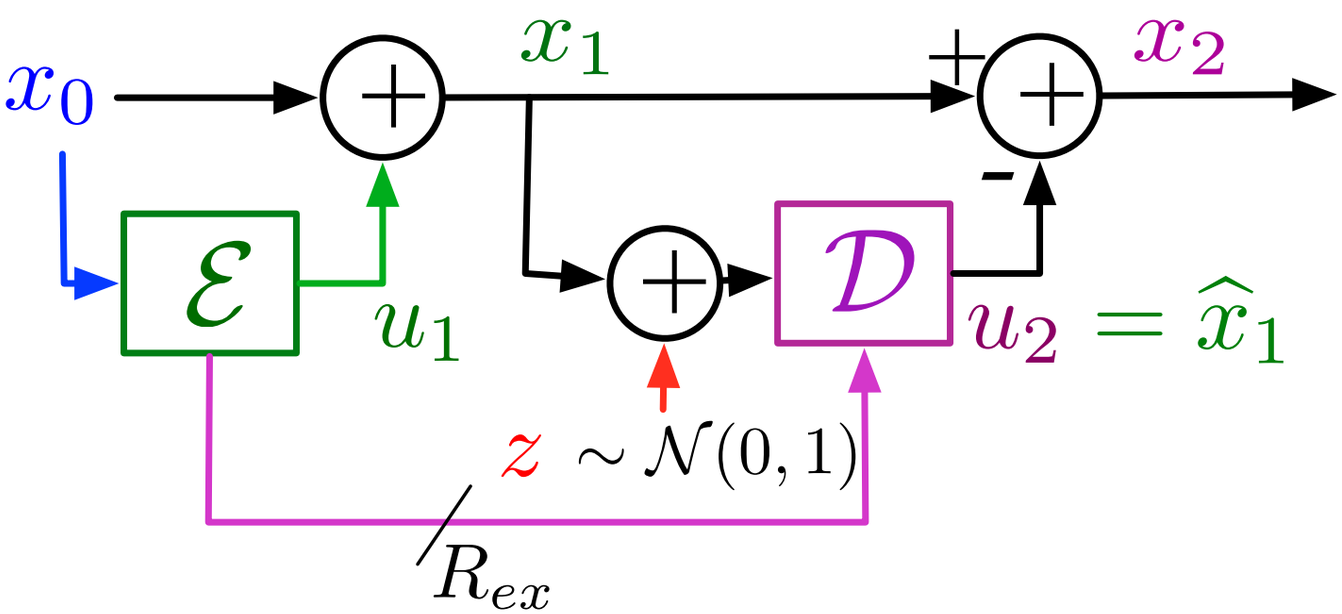


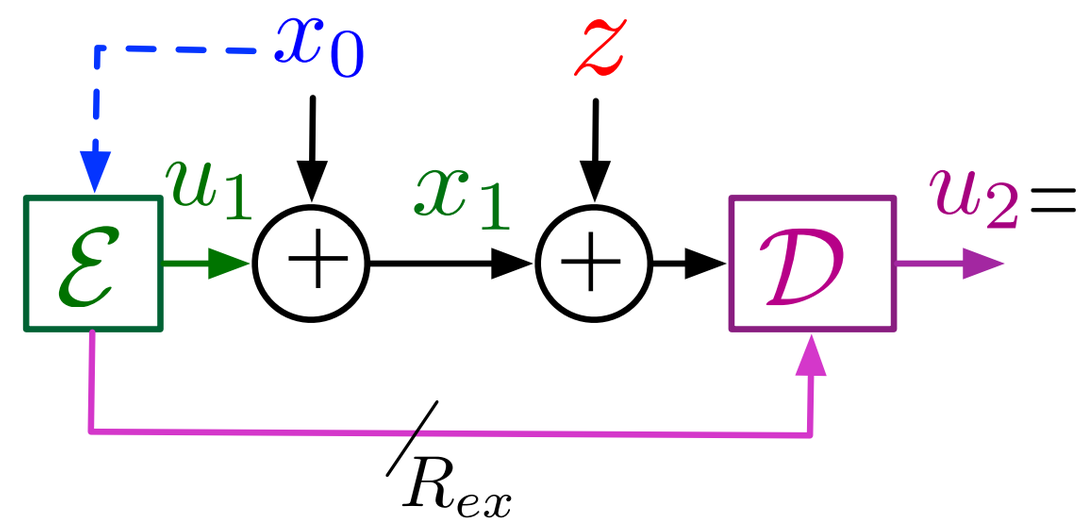


$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[(x_1 - \widehat{x}_1)^2\right]$$

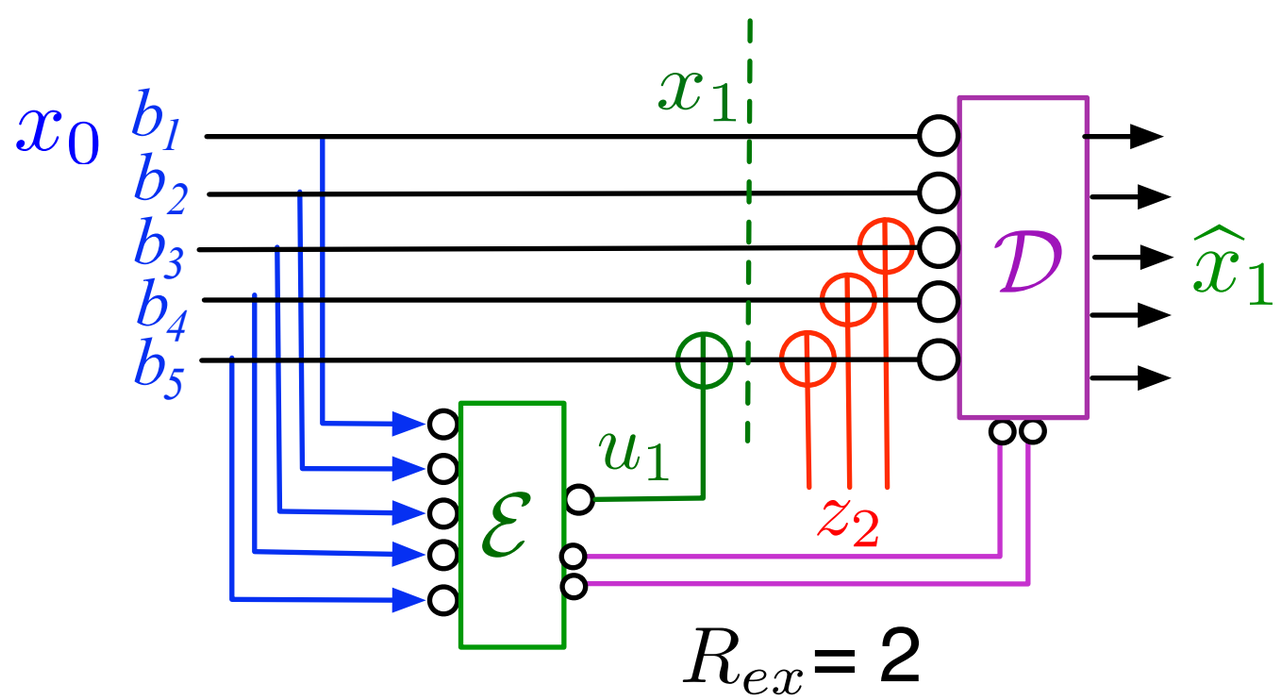


A deterministic abstraction

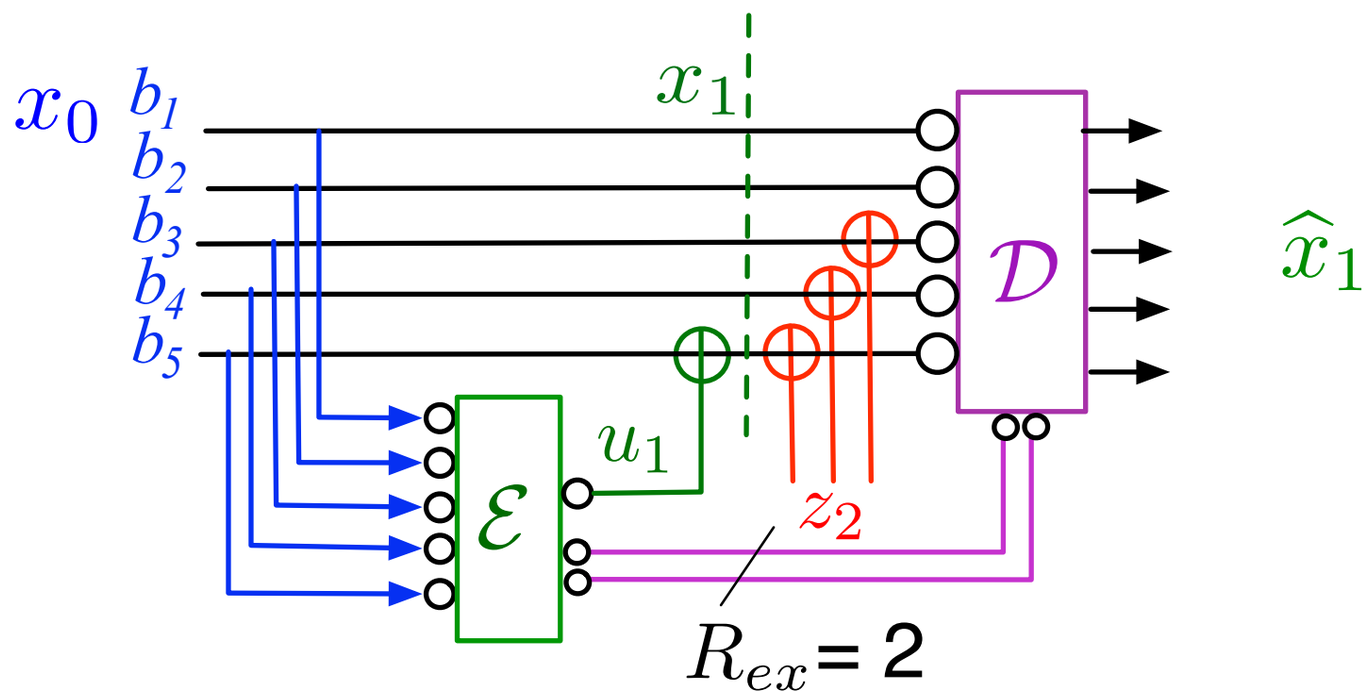




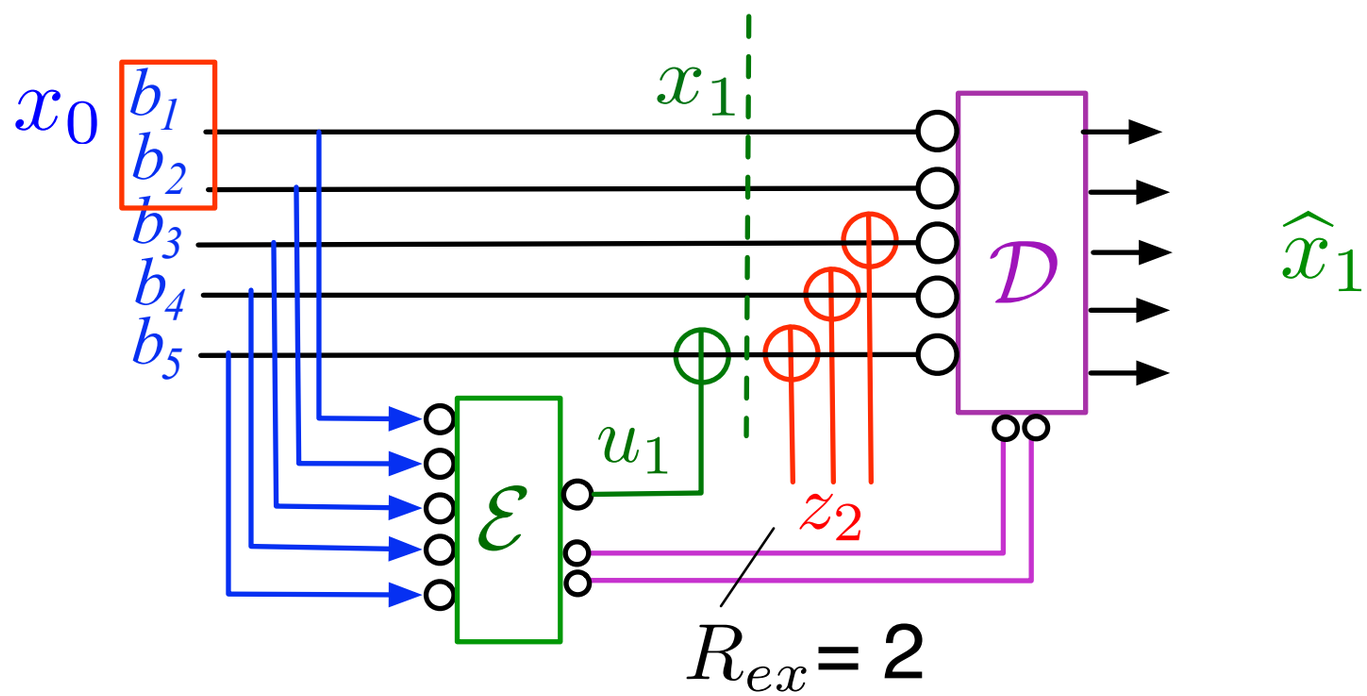
$$k^2\mathbb{E}\left[u_1^2\right] + \mathbb{E}\left[(x_1 - \widehat{x}_1)^2\right]$$



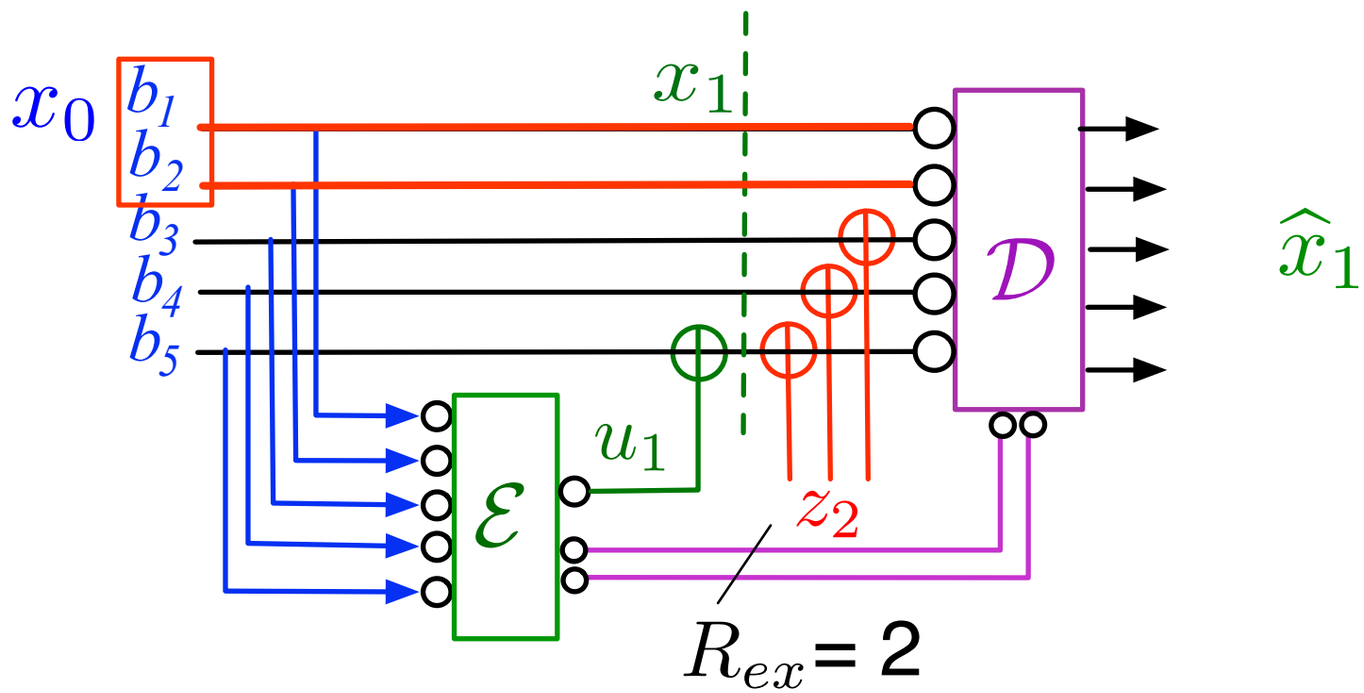
Strategy for deterministic abstraction



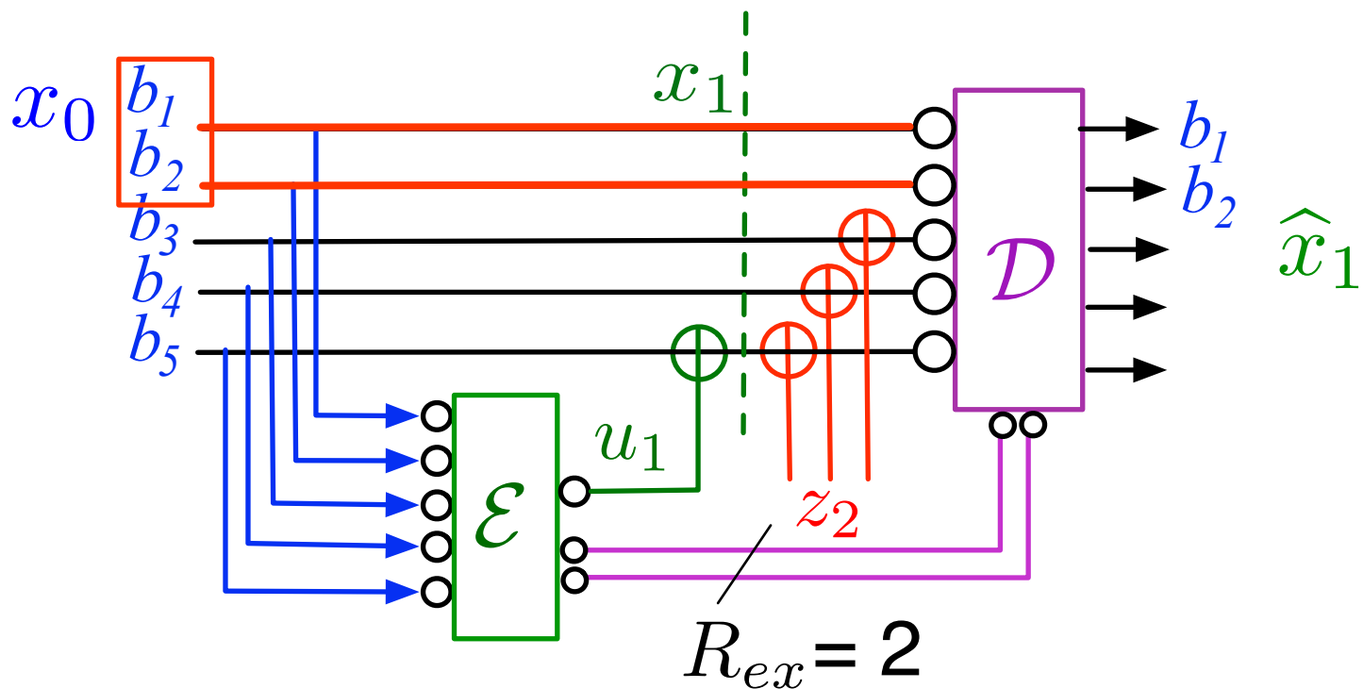
Strategy for deterministic abstraction



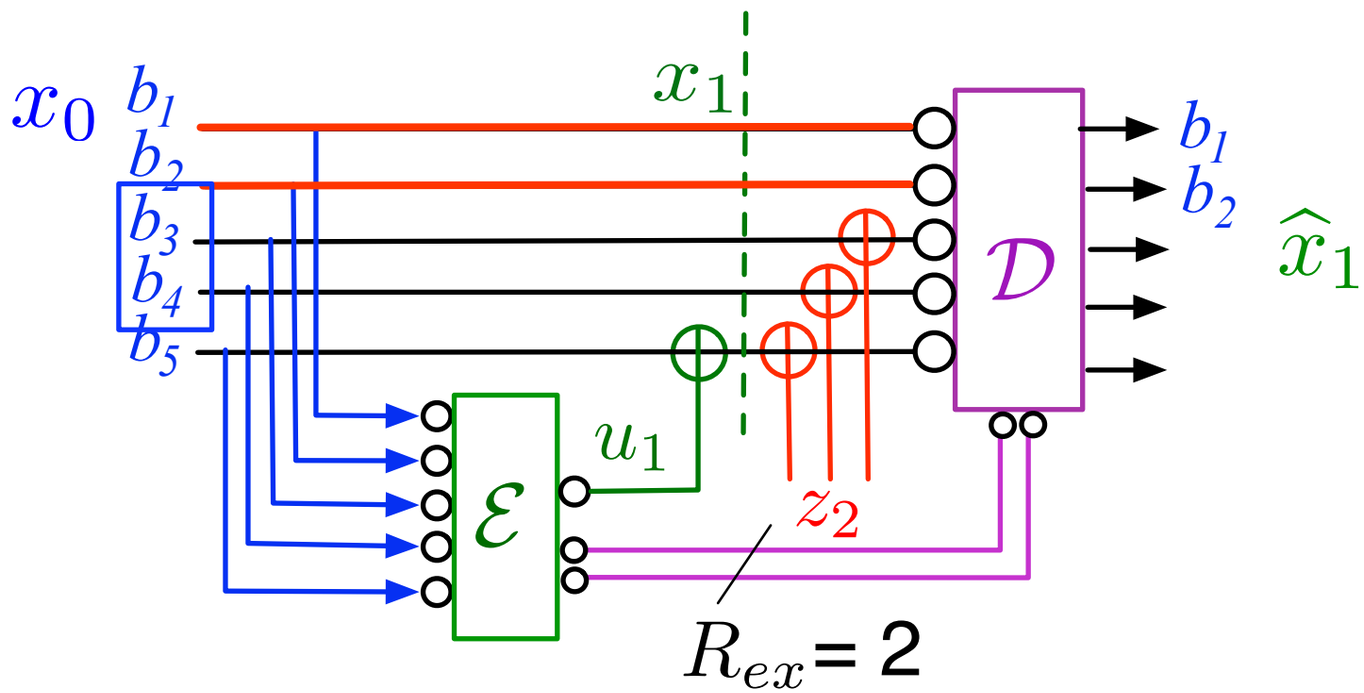
Strategy for deterministic abstraction



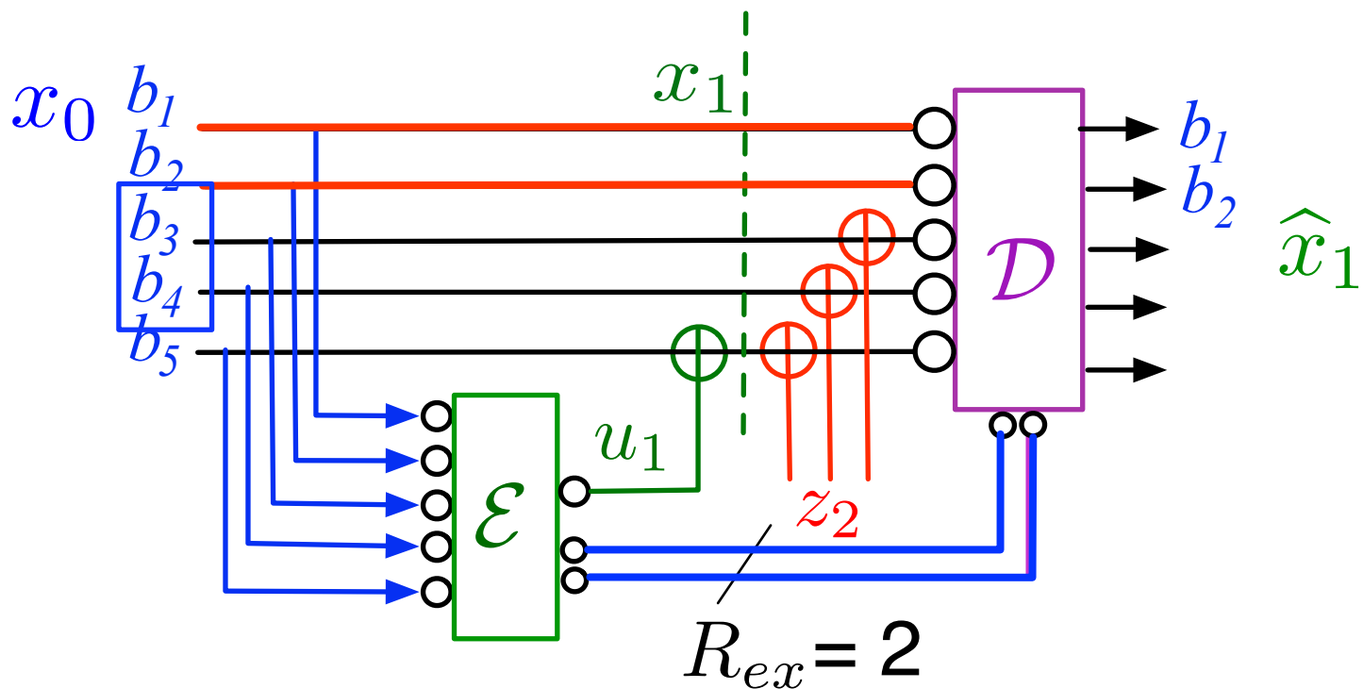
Strategy for deterministic abstraction



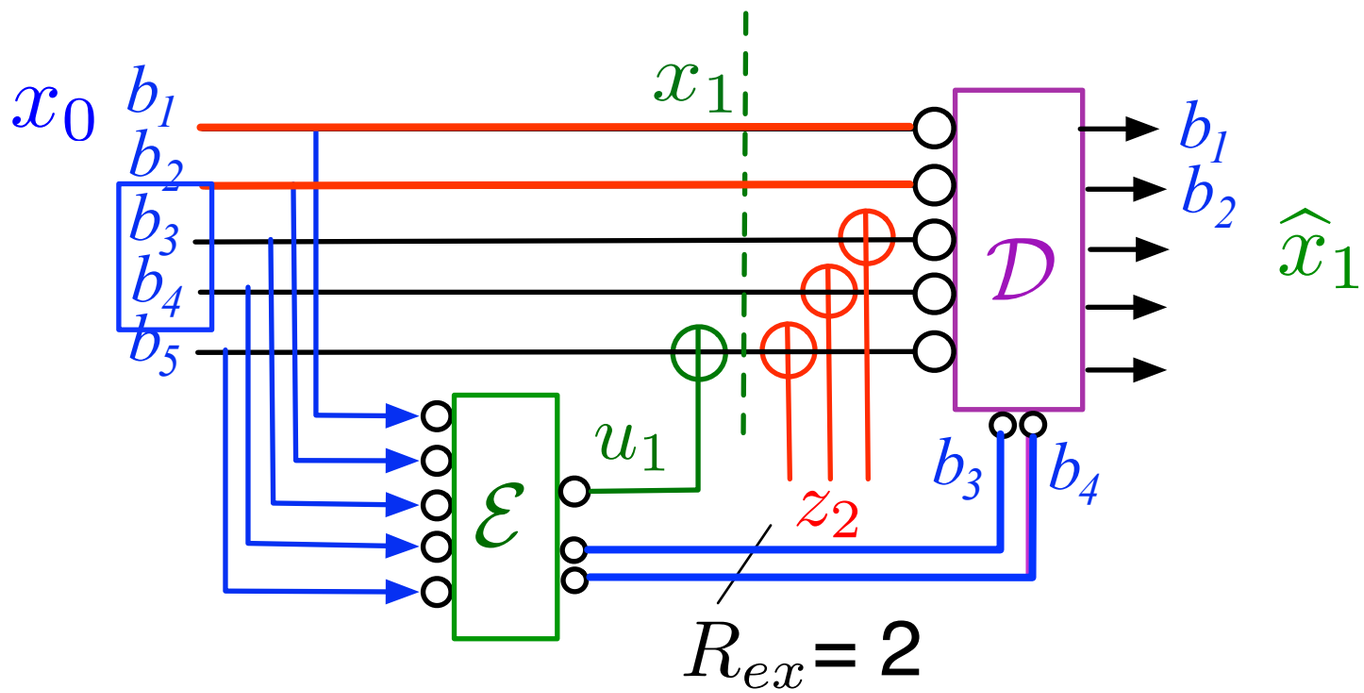
Strategy for deterministic abstraction



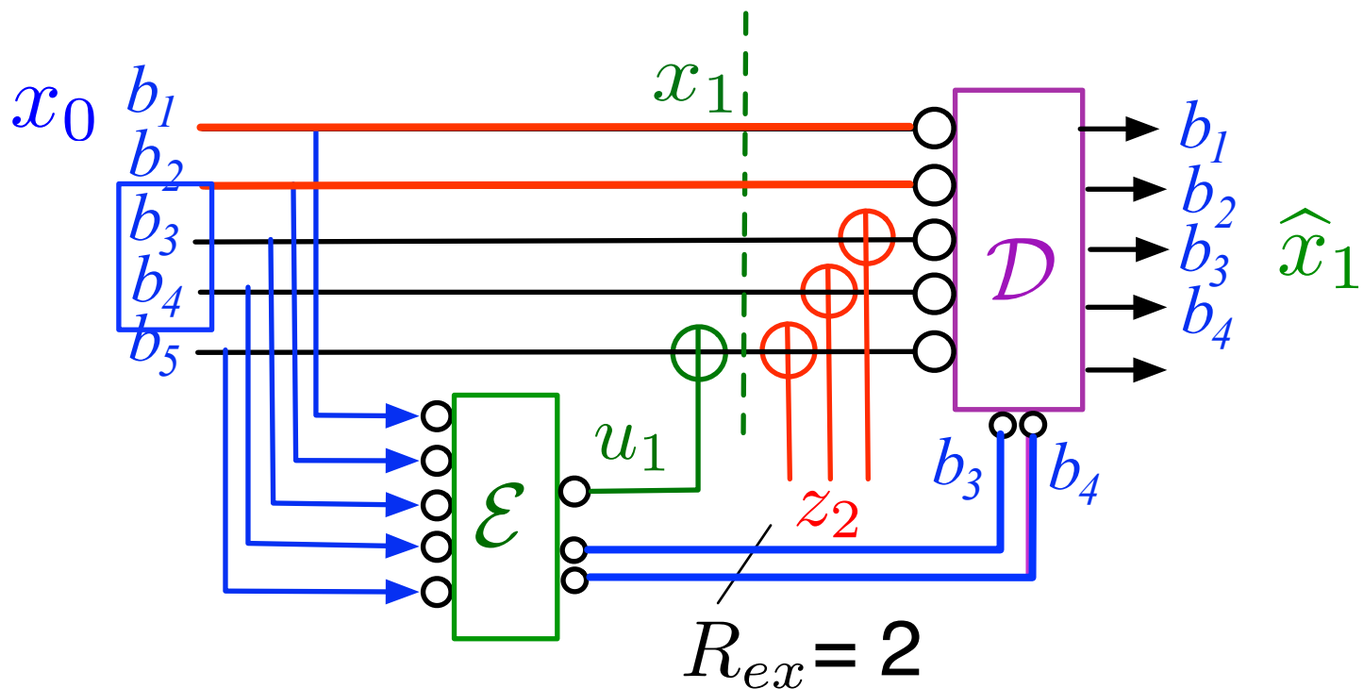
Strategy for deterministic abstraction



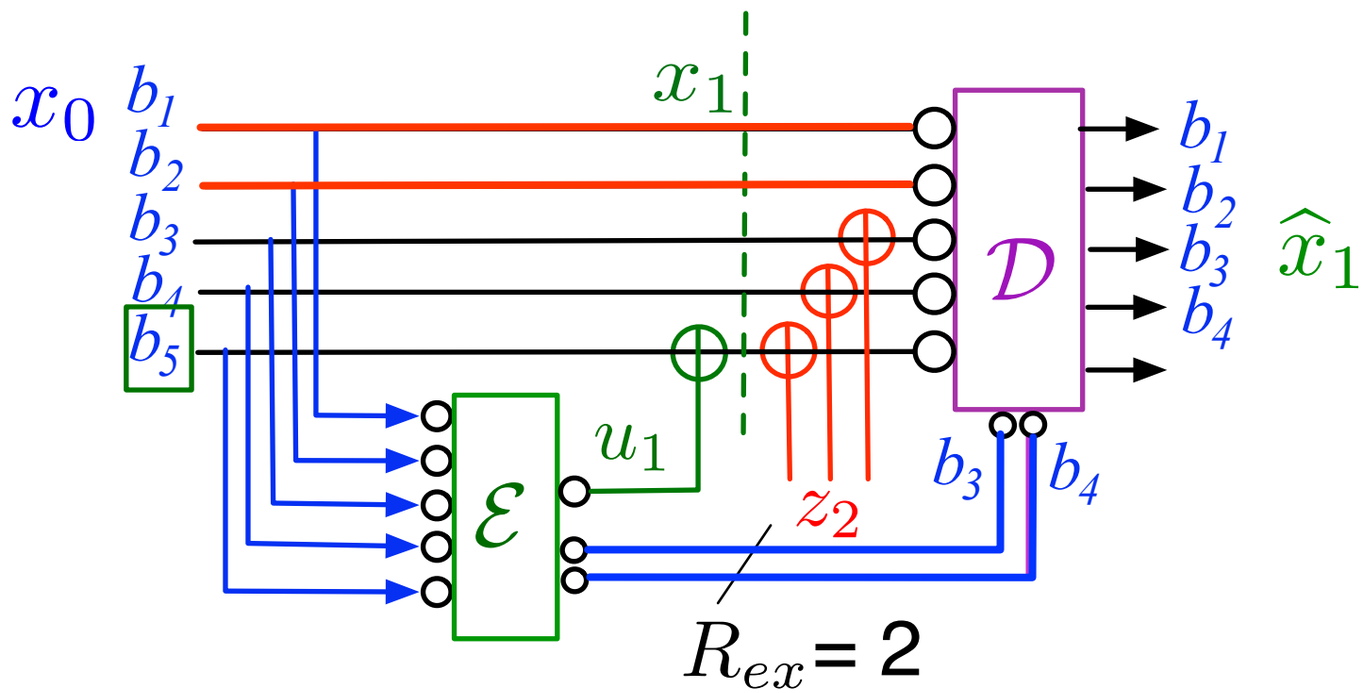
Strategy for deterministic abstraction



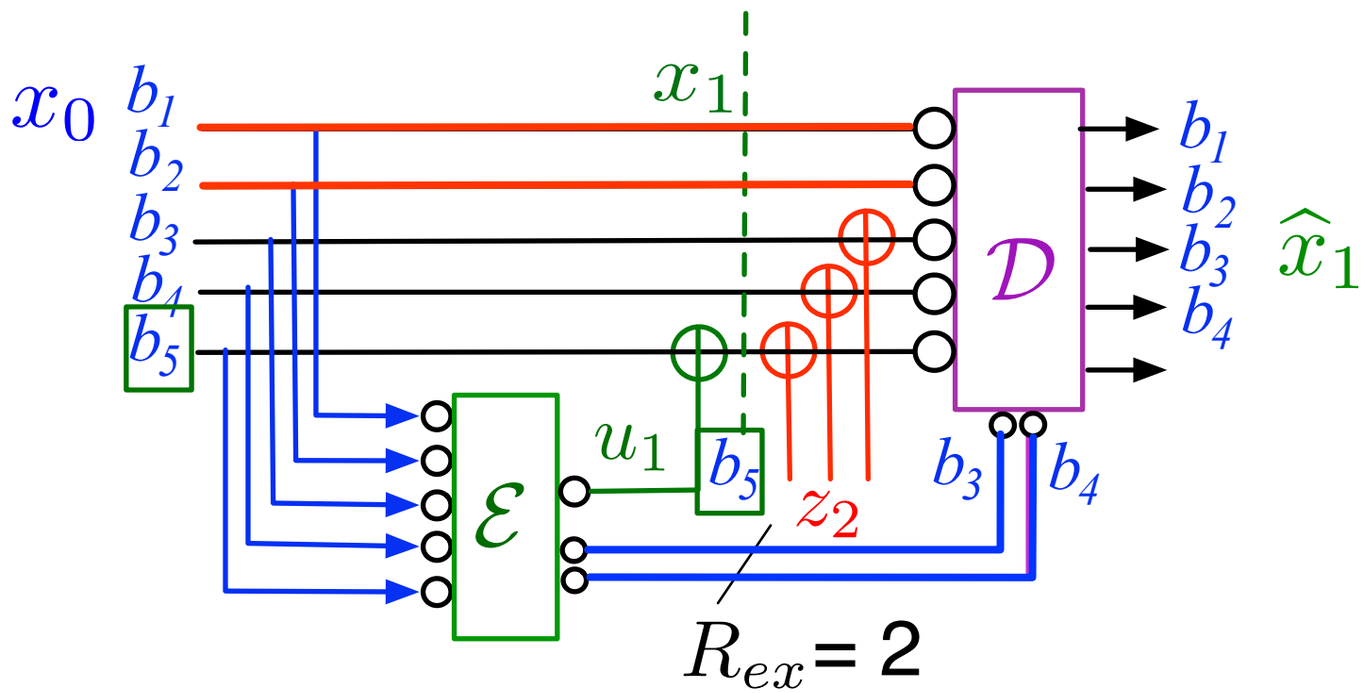
Strategy for deterministic abstraction



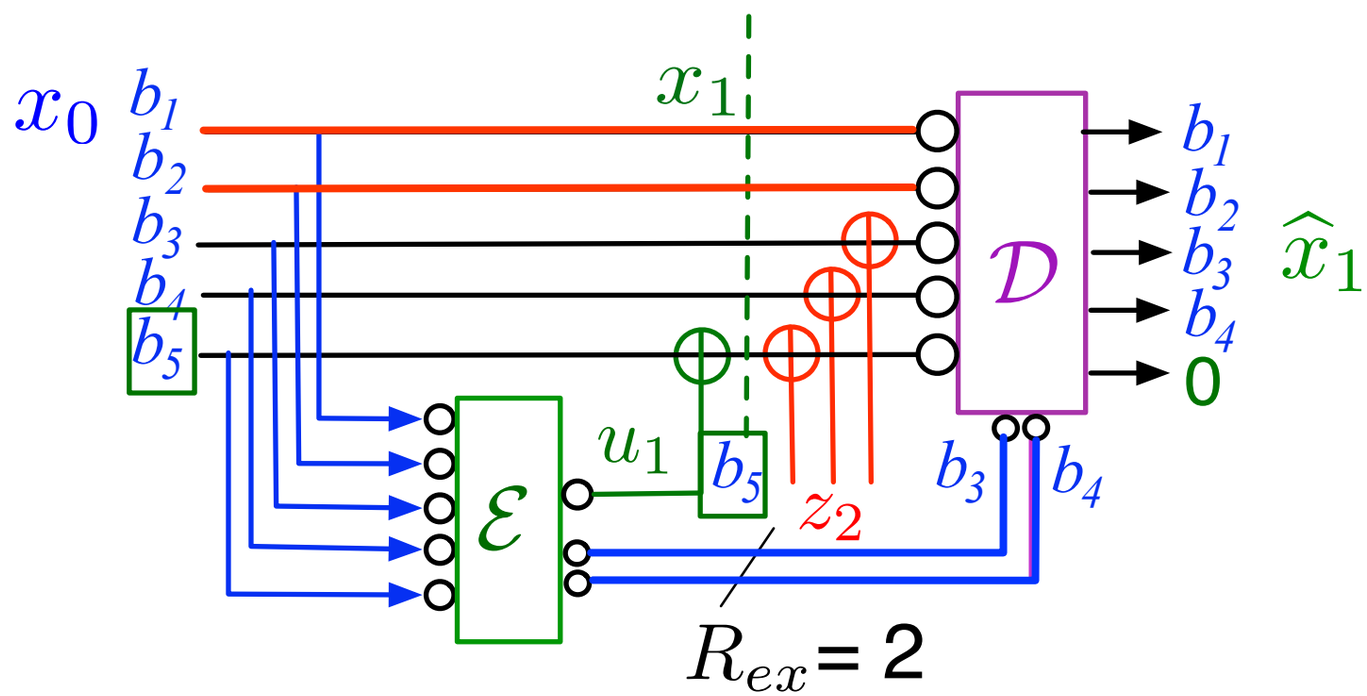
Strategy for deterministic abstraction



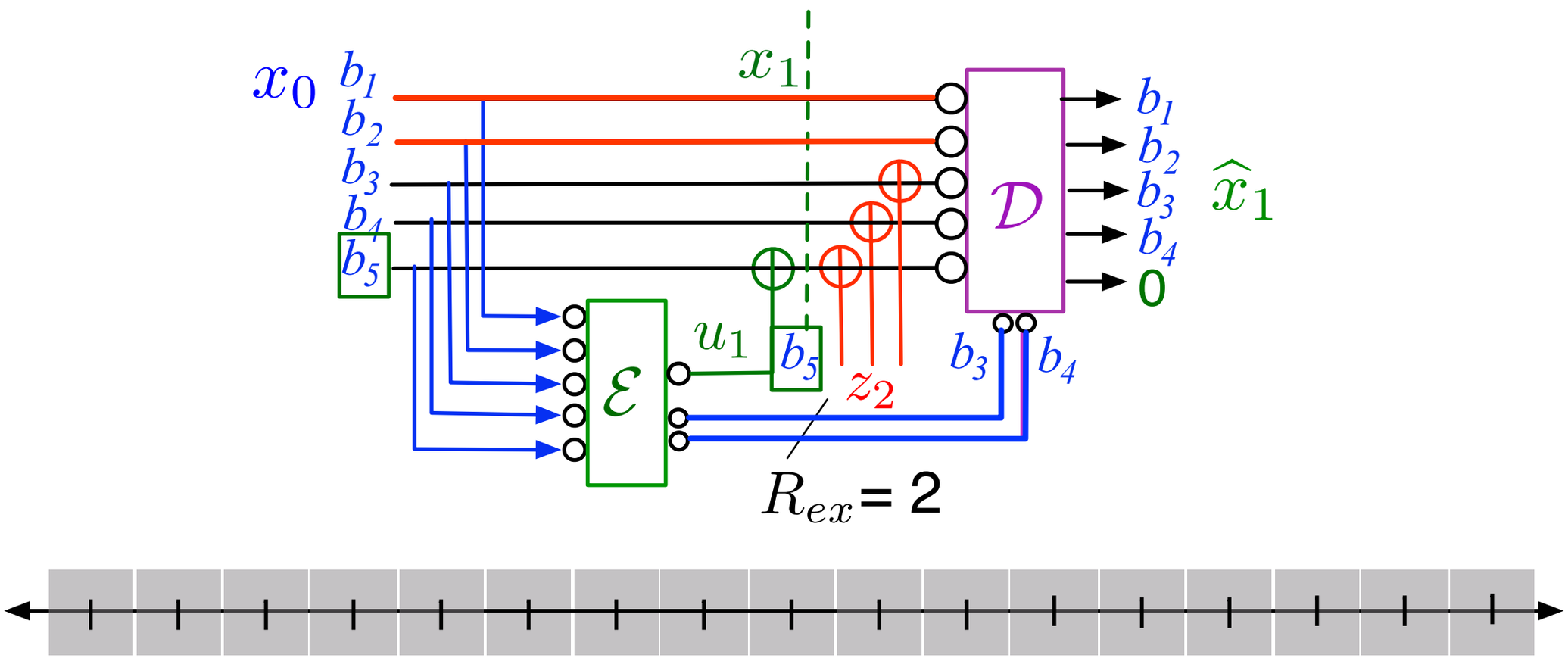
Strategy for deterministic abstraction



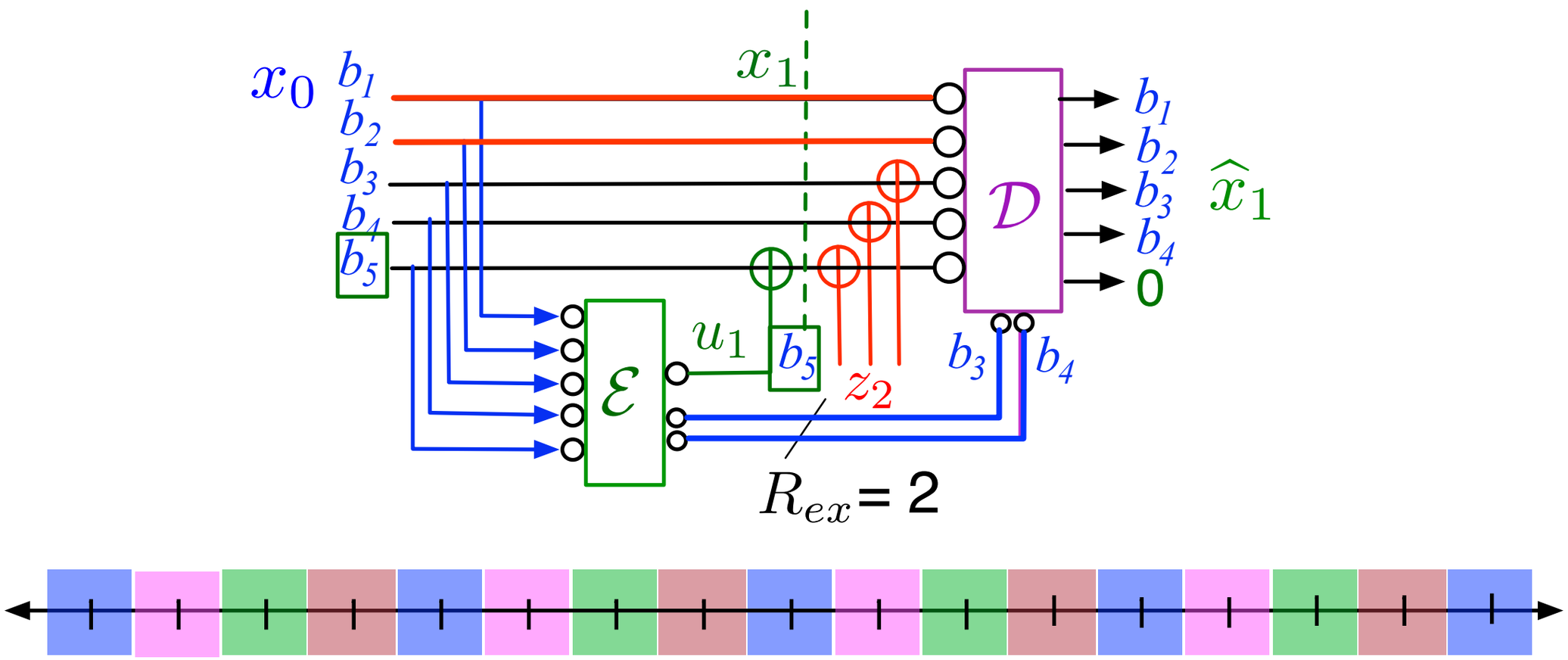
Strategy for deterministic abstraction



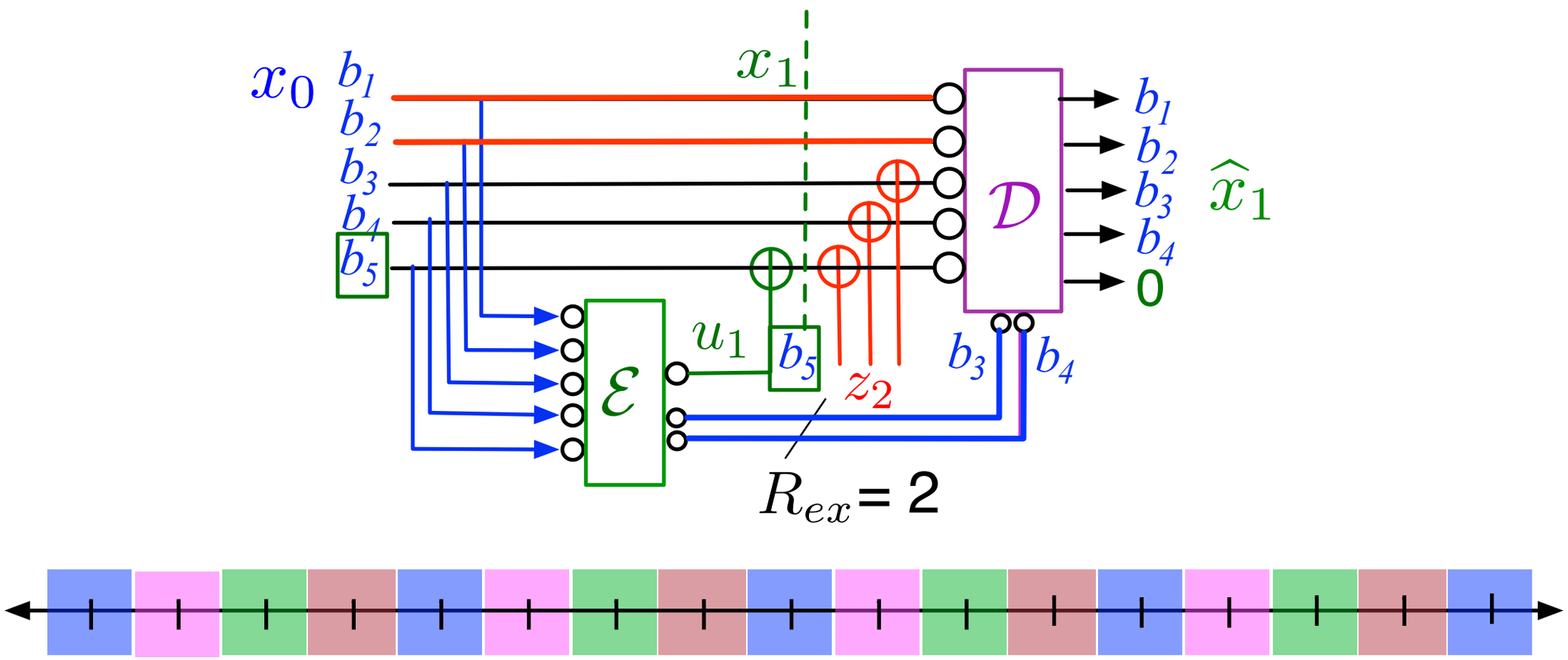
Strategy for deterministic abstraction



Strategy for deterministic abstraction

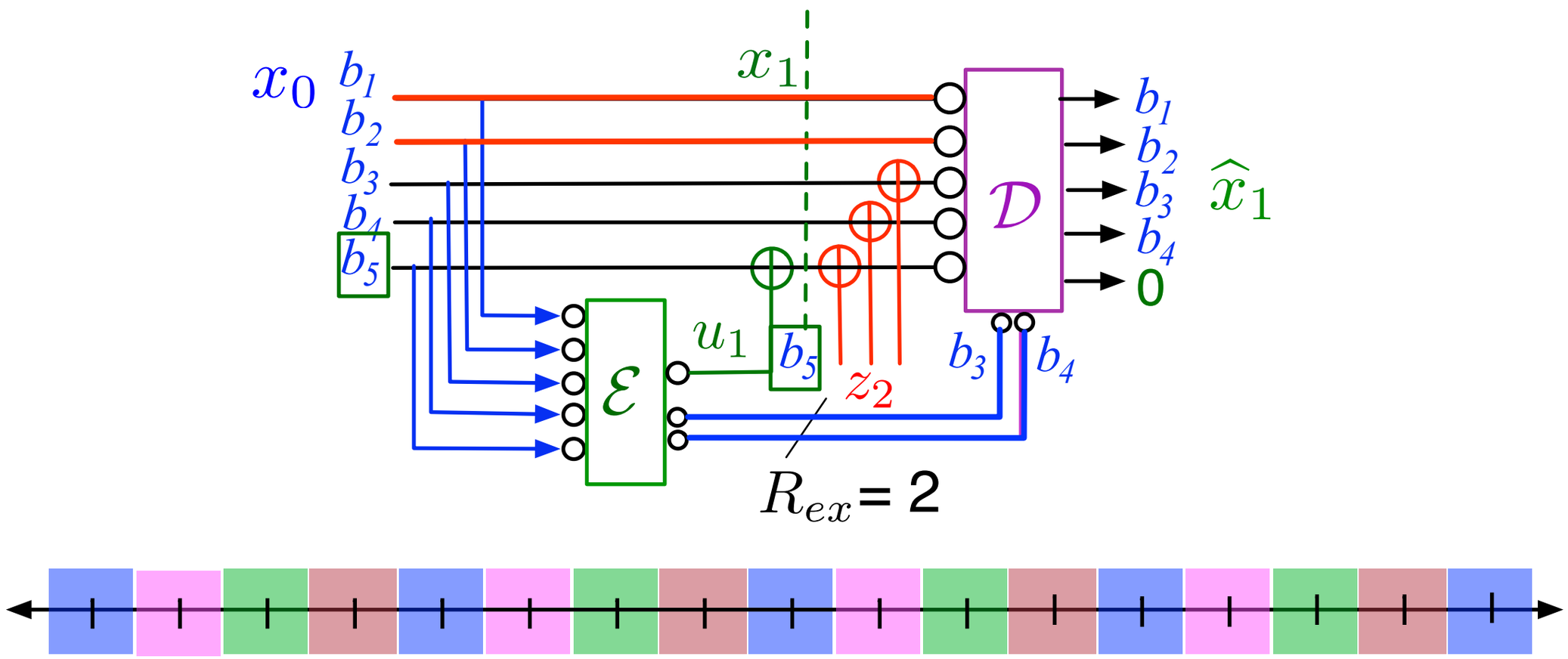


Strategy for deterministic abstraction



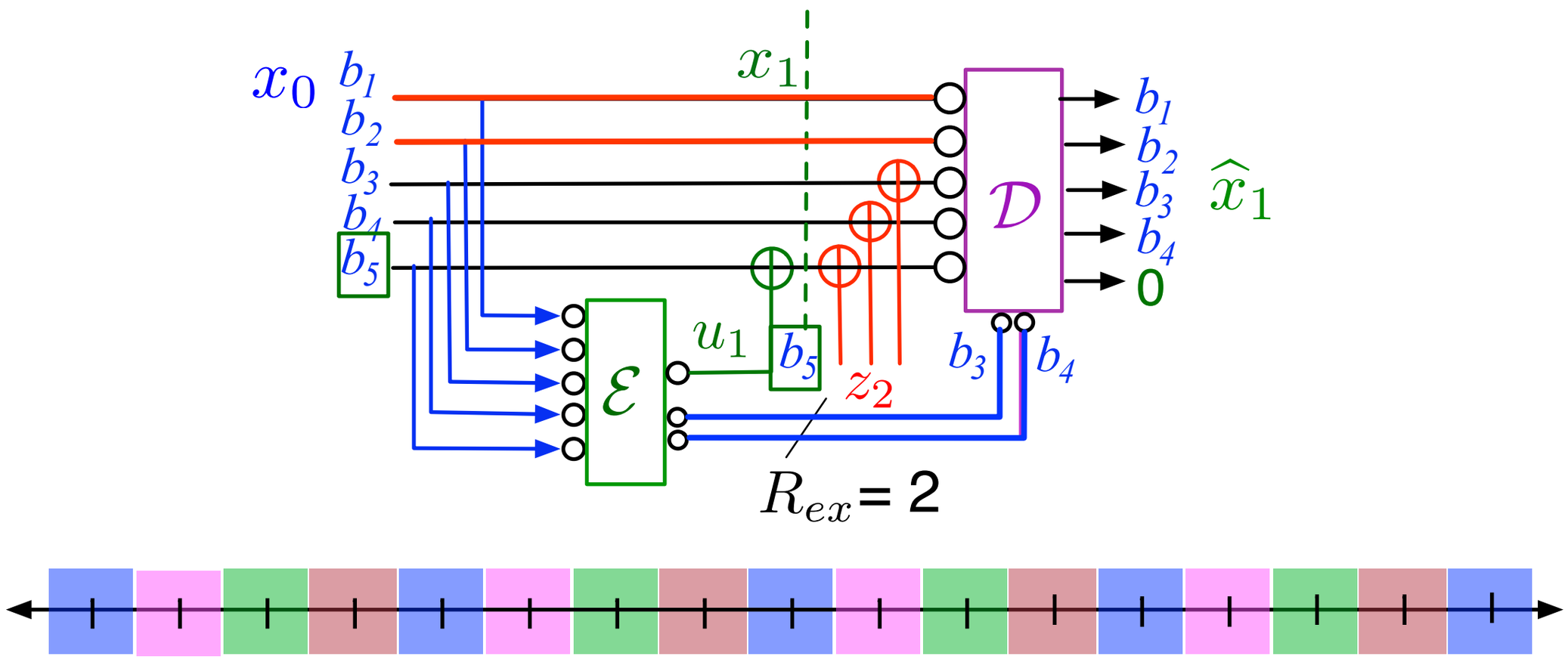
Explicit channel: fine information

Strategy for deterministic abstraction



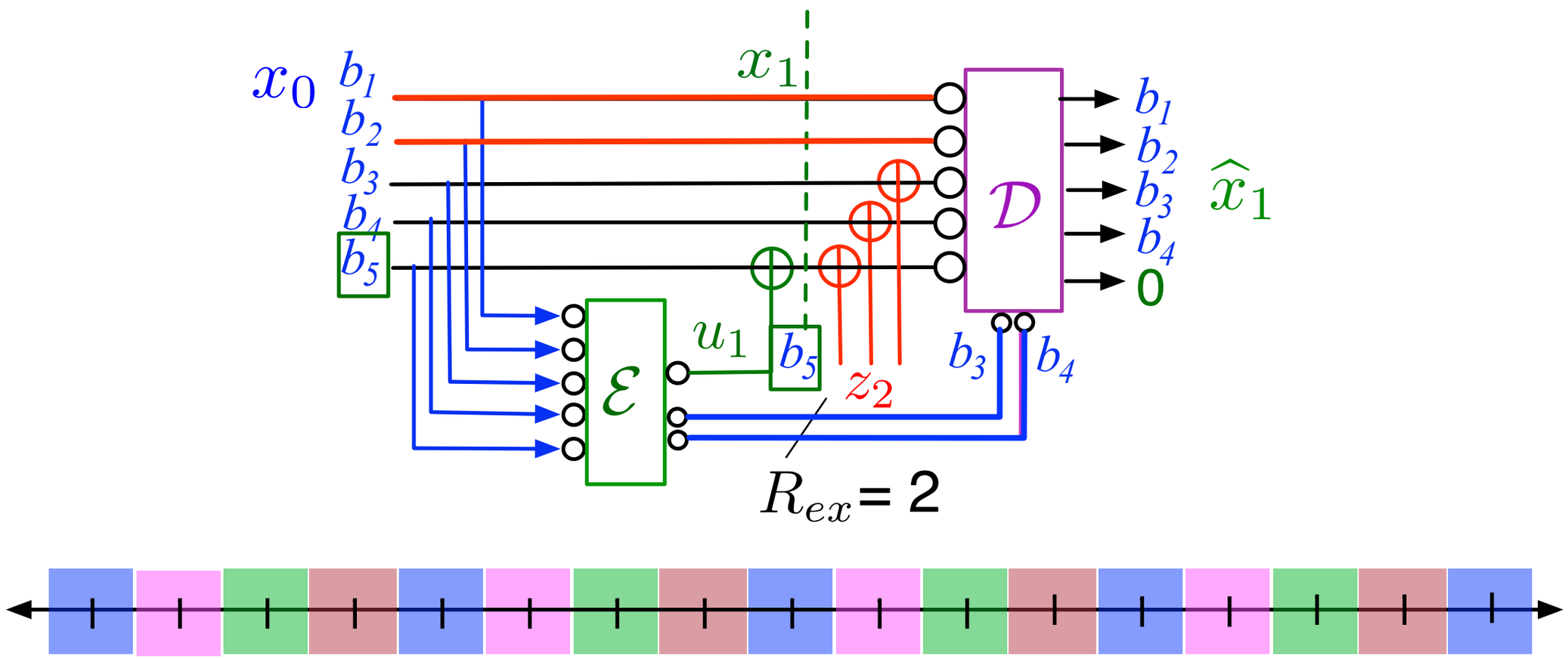
Explicit channel: fine information

Strategy for deterministic abstraction



Explicit channel: fine information

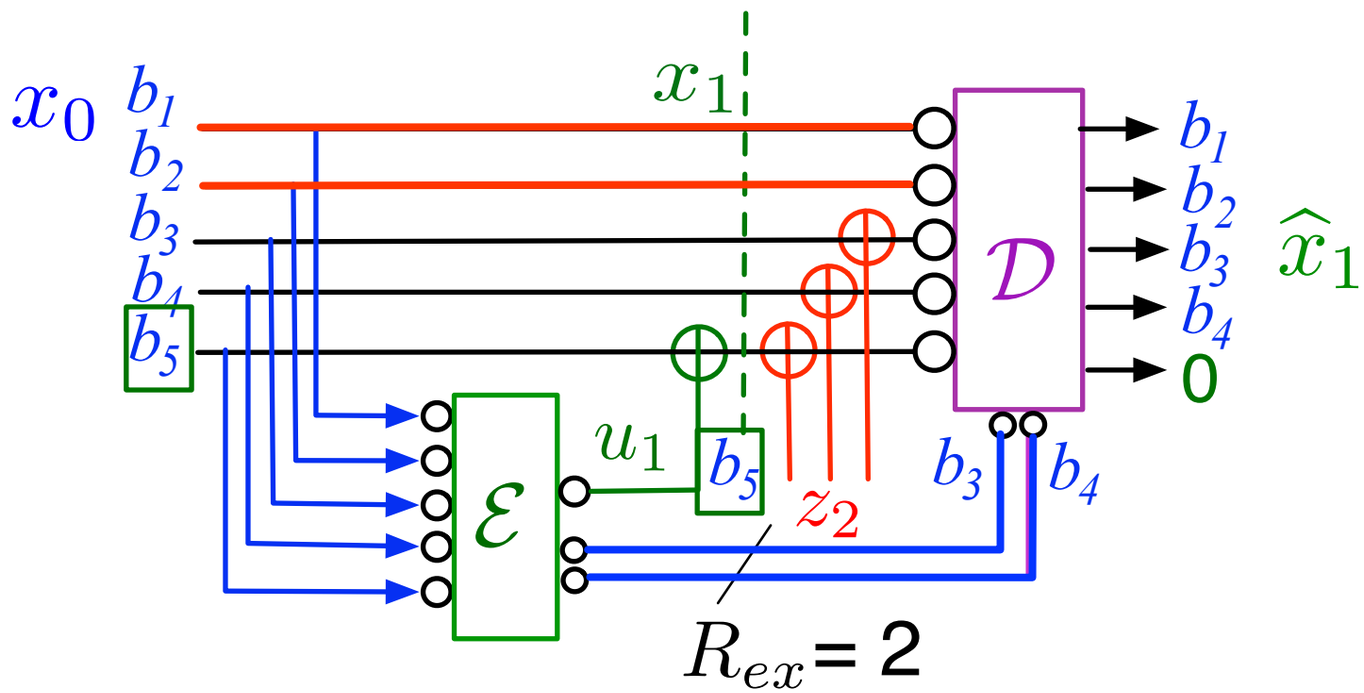
Strategy for deterministic abstraction

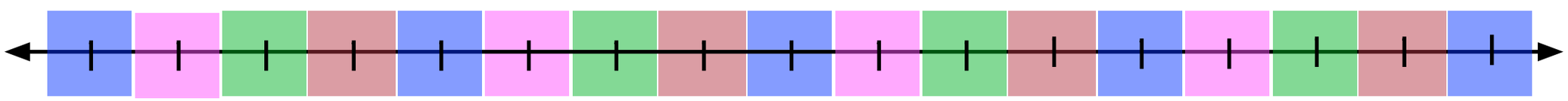


Explicit channel: fine information

Implicit channel: high order bits

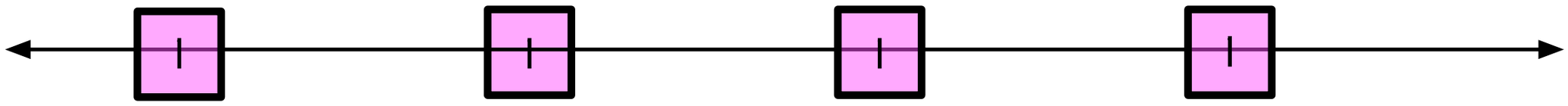
Strategy for deterministic abstraction



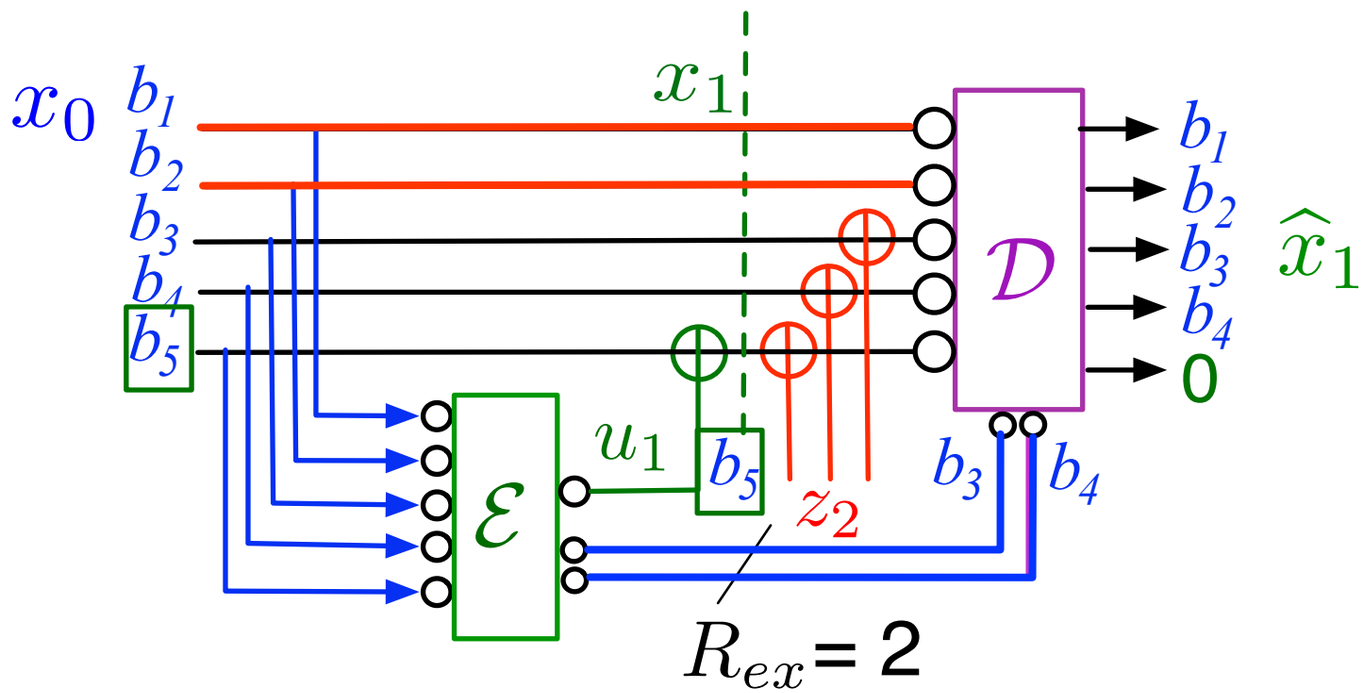


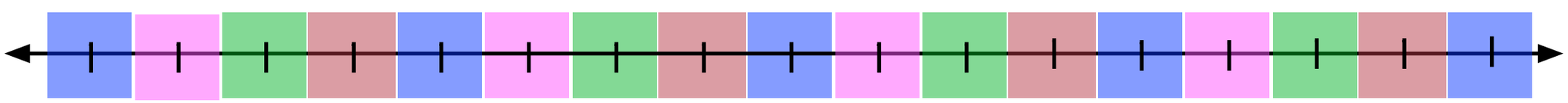
Explicit channel: fine information

Implicit channel: high order bits



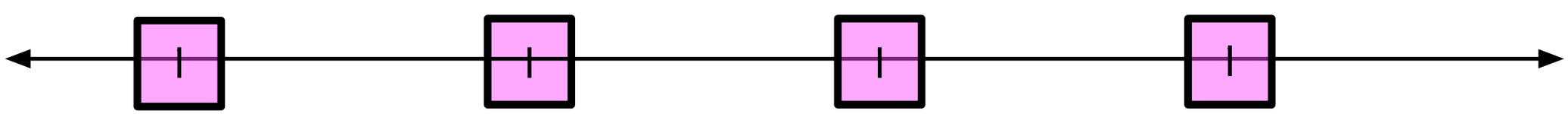
Strategy for deterministic abstraction





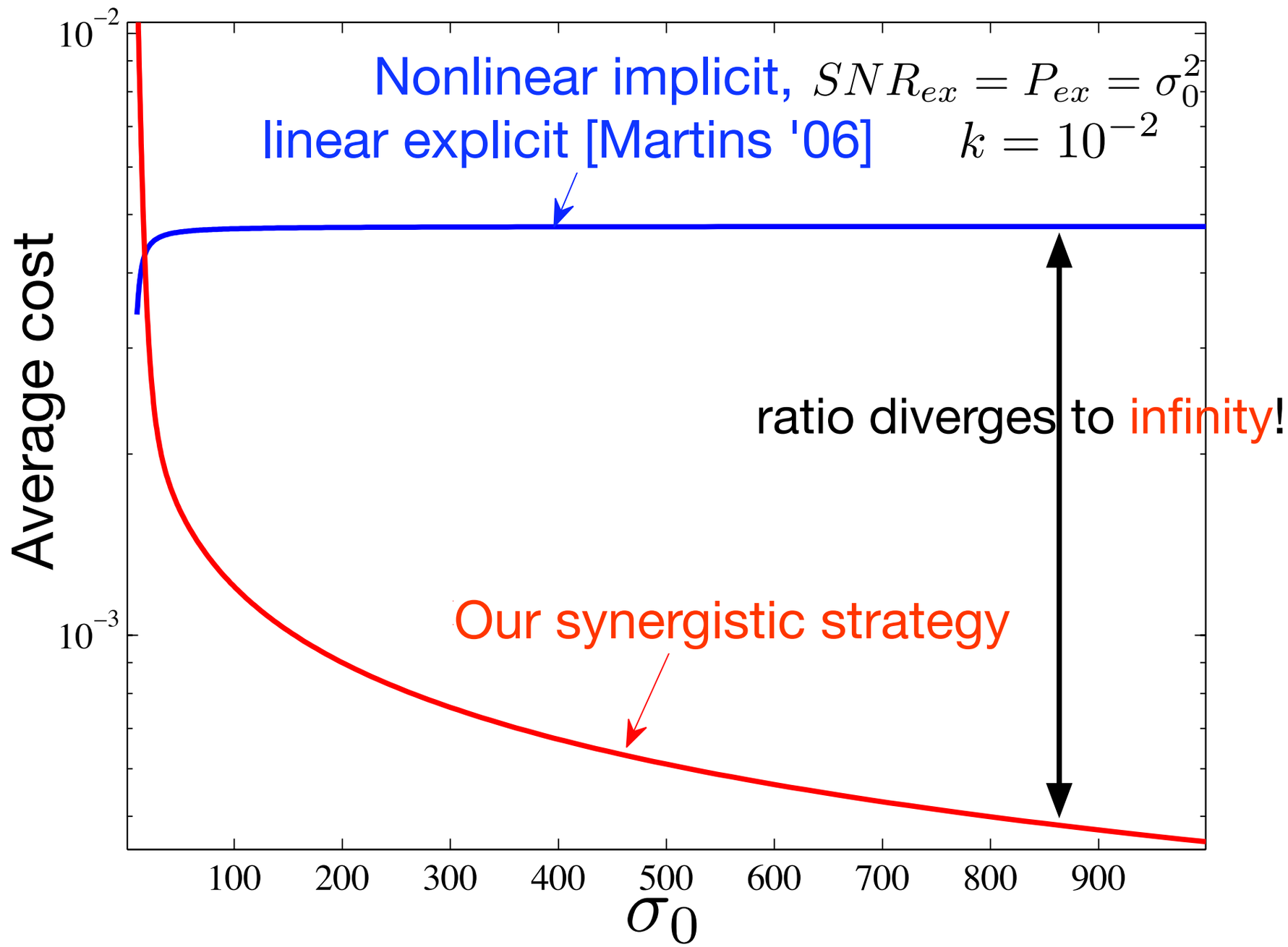
Explicit channel: fine information

Implicit channel: high order bits

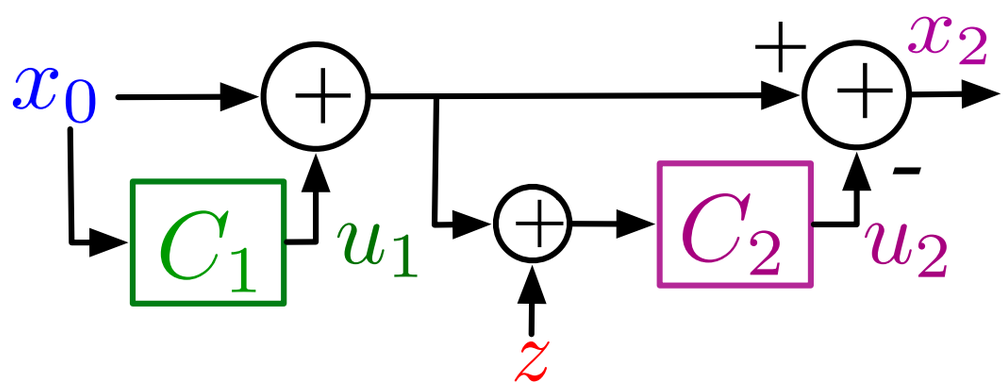


binning!

Infinitely better than the “natural” strategies!



Is the counterexample relevant?



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

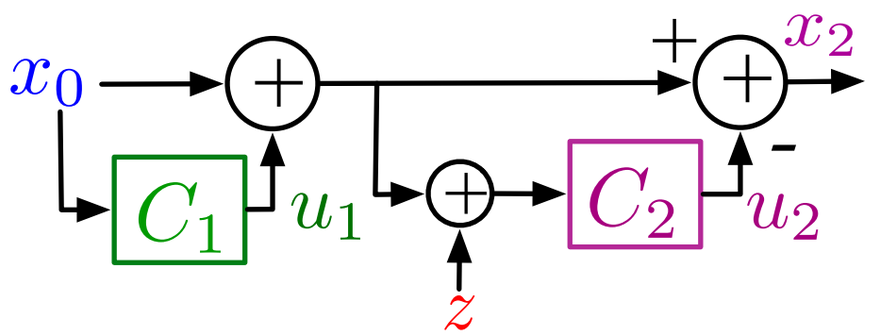
Induced norm: linear strategies optimal! [Rotkowitz '06]

Witsenhausen's LQG formulation: more assumptions harder!

Q. Why consider the quadratic norm?

[John Doyle, "Paths ahead", MIT, 2009]

Induced norm formulation too conservative?

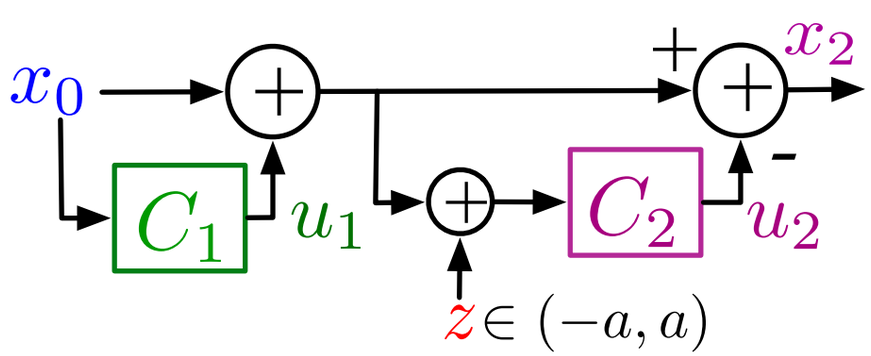


$$\min \max_{x_0, z} \frac{k^2 u_1^2 + x_2^2}{x_0^2 + z^2}$$

Makes no assumptions: noise completely arbitrary, adversarial!

Usually we know *something* about the noise, e.g. "typical" values, or bounds

say $z \in (-a, a)$



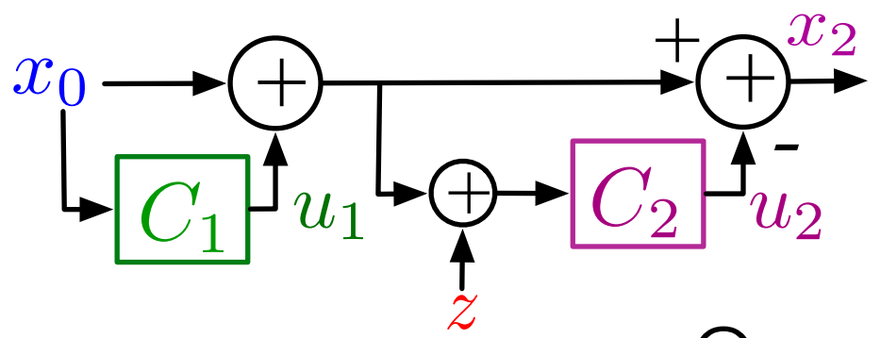
A quadratic cost

$$\min \max_{x_0, z \in (-a,a)} k^2 u_1^2 + x_2^2$$

Quantization (signaling) : can outperform linear by an unbounded factor

is approximately optimal!

Which is a “right” norm?



Induced norm

$$\min_{x_0,z} \max \frac{k^2 u_1^2 + x_2^2}{x_0^2 + z^2}$$

Quadratic norm

$$\min_{x_0} \max_{z \in (-a,a)} k^2 u_1^2 + x_2^2$$

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