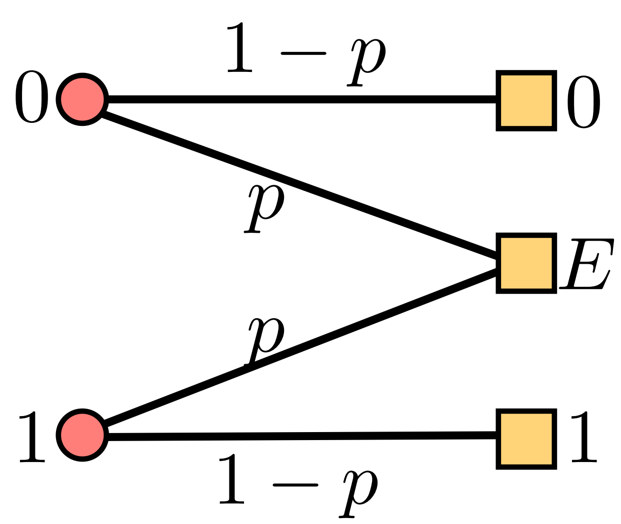
Capacity-Achieving Codes on Erasure (and other) Channels



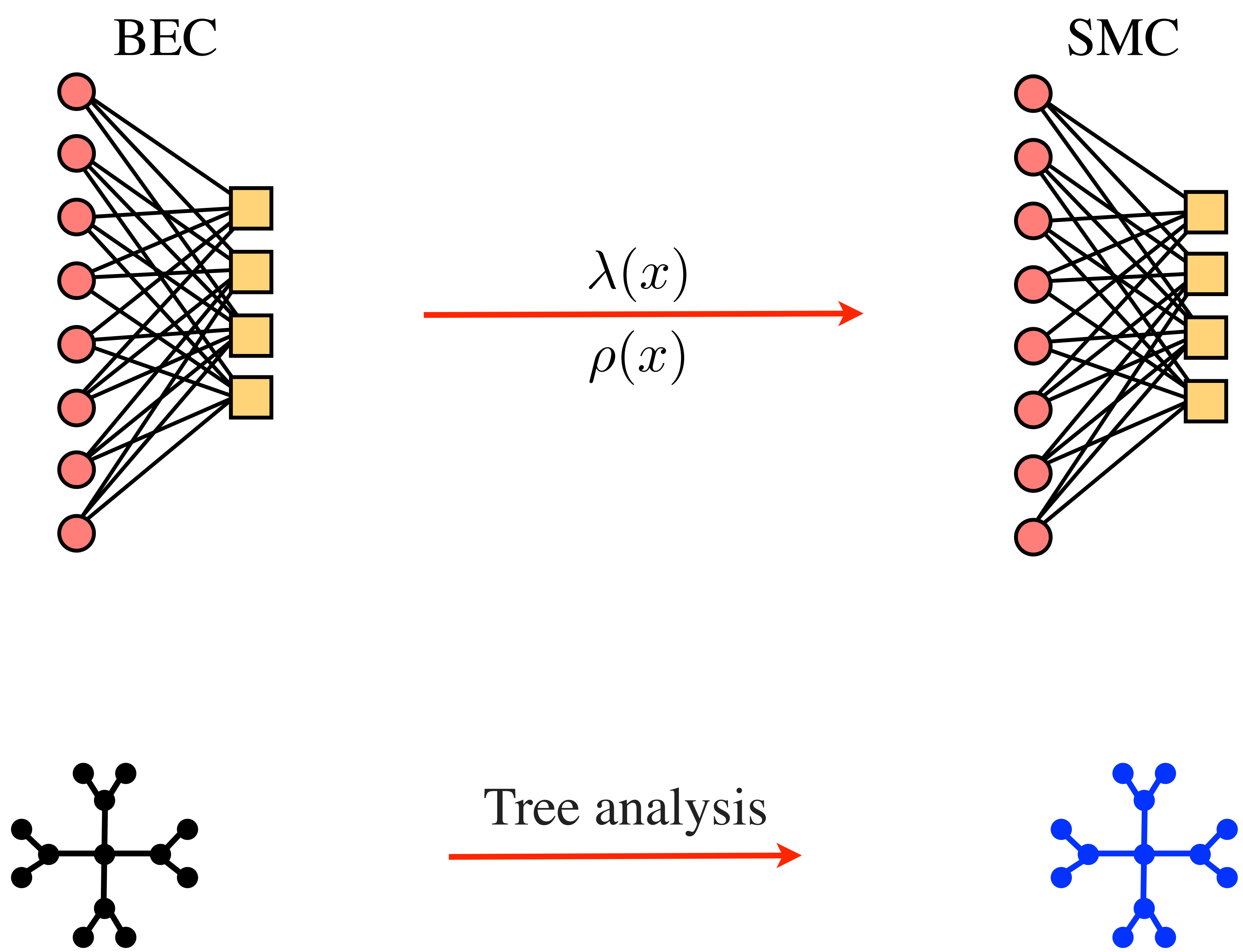
Amin Shokrollahi

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
+
DF

Why the Erasure Channel?

- Theory: Manageable example of a much richer class of channels (Symmetric Memoryless Channels)
- Transfer: Thorough understanding of codes on the BEC can help us analyze and design codes on more complicated SMC's.
- Applications: transmission of data on packet based networks.

A Lot Carries Over

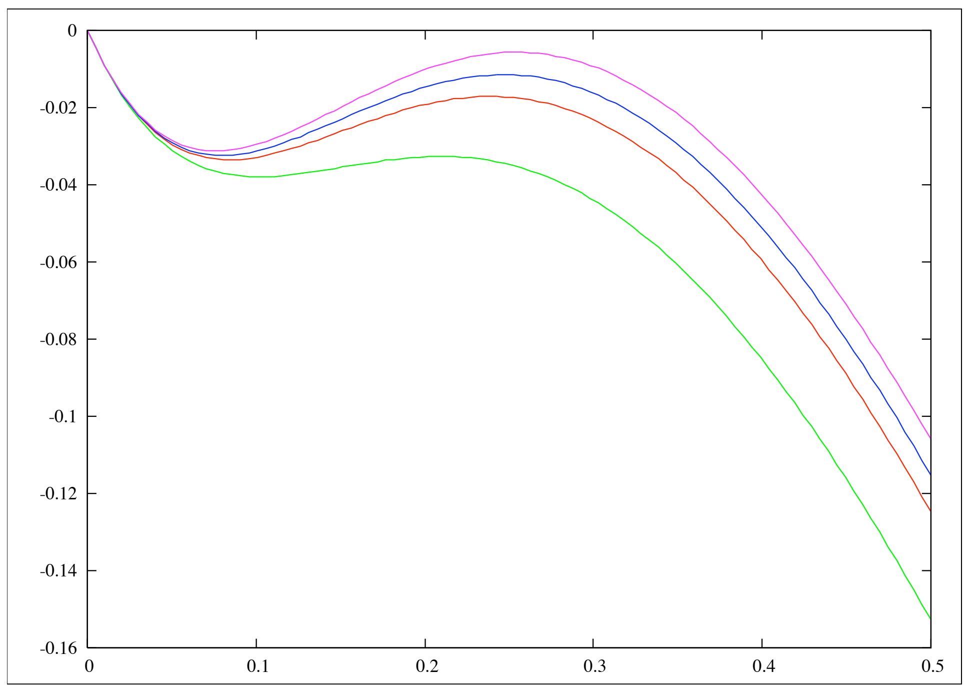


A Lot Carries Over

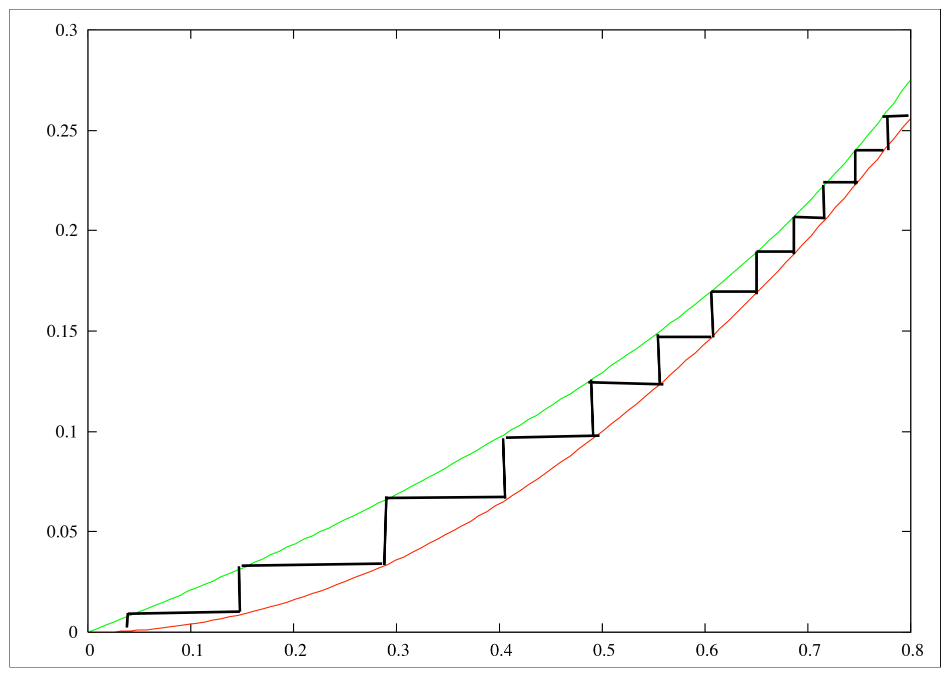
BEC   SMC

$$p_{n+1} = p_0 \lambda(1 - \rho(1 - x)) \quad \xrightarrow{\text{DE}} \quad f_{n+1} = f_0 \otimes \lambda\left(\Gamma^{-1}(\rho(\Gamma(f_n)))\right)$$

$$\lambda'(0)\rho'(1) \leq (p_0)^{-1} \quad \xrightarrow{\text{Stability}} \quad \lambda'(0)\rho'(1) \leq \left(\int_{-\infty}^{\infty} f_0(x) \mathrm{e}^{-x/2} \mathrm{d}x\right)^{-1}$$



Linear Programming →



But not Everything

BEC

$$\lambda(x) = \frac{1}{H(D)} \sum_{k=1}^{D} \frac{x^k}{k}$$

$$\rho(x) = \mathrm{e}^{\alpha(x-1)}$$

Capacity-achieving

degree distributions

SMC

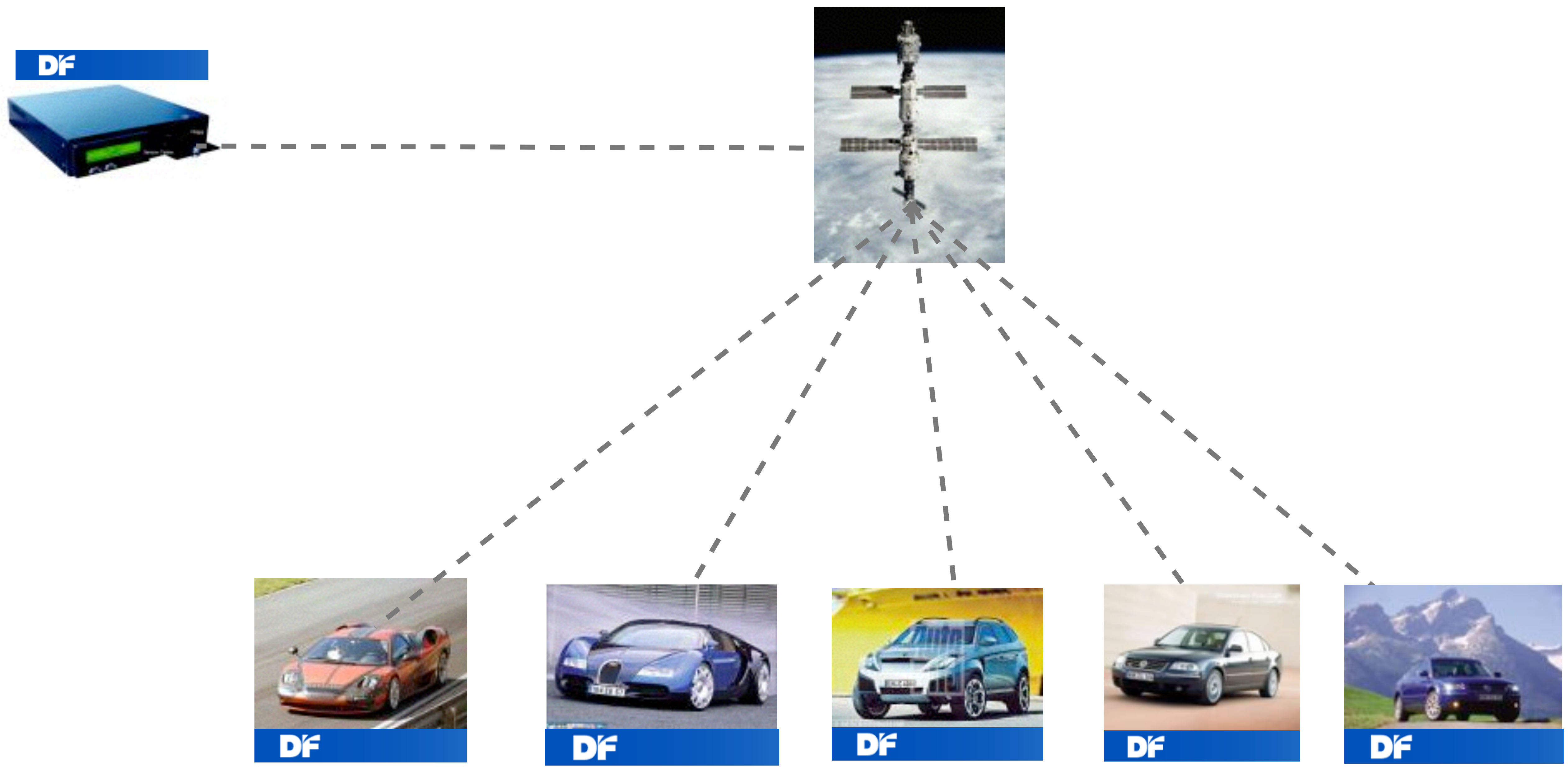
?

Can we use our knowledge of the erasure channel to design capacity-achieving codes on other channels?

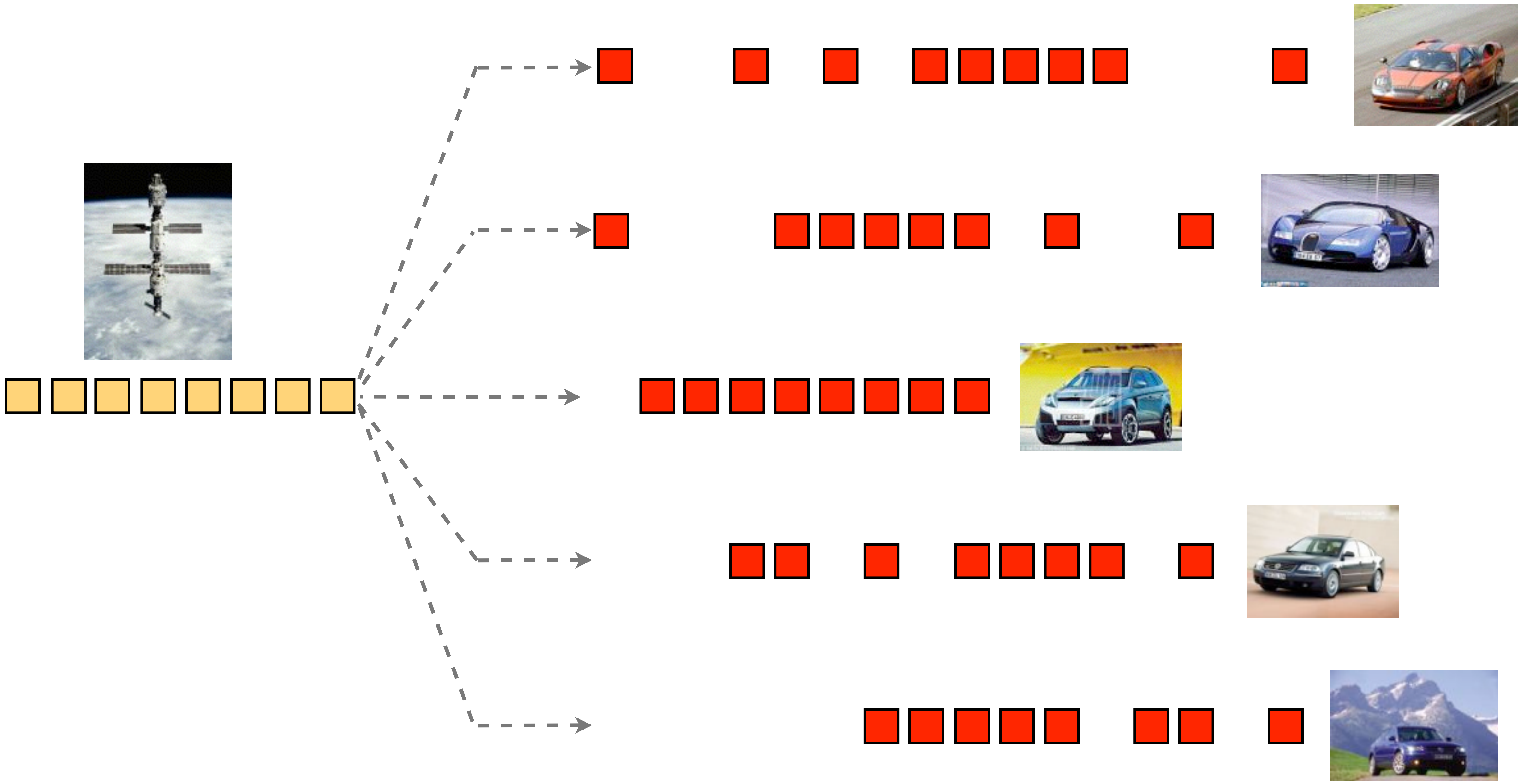
Perhaps. But we need to understand the erasure case much better!

A Real Transmission Problem

Reliable transmission over unknown channels with minimal feedback.

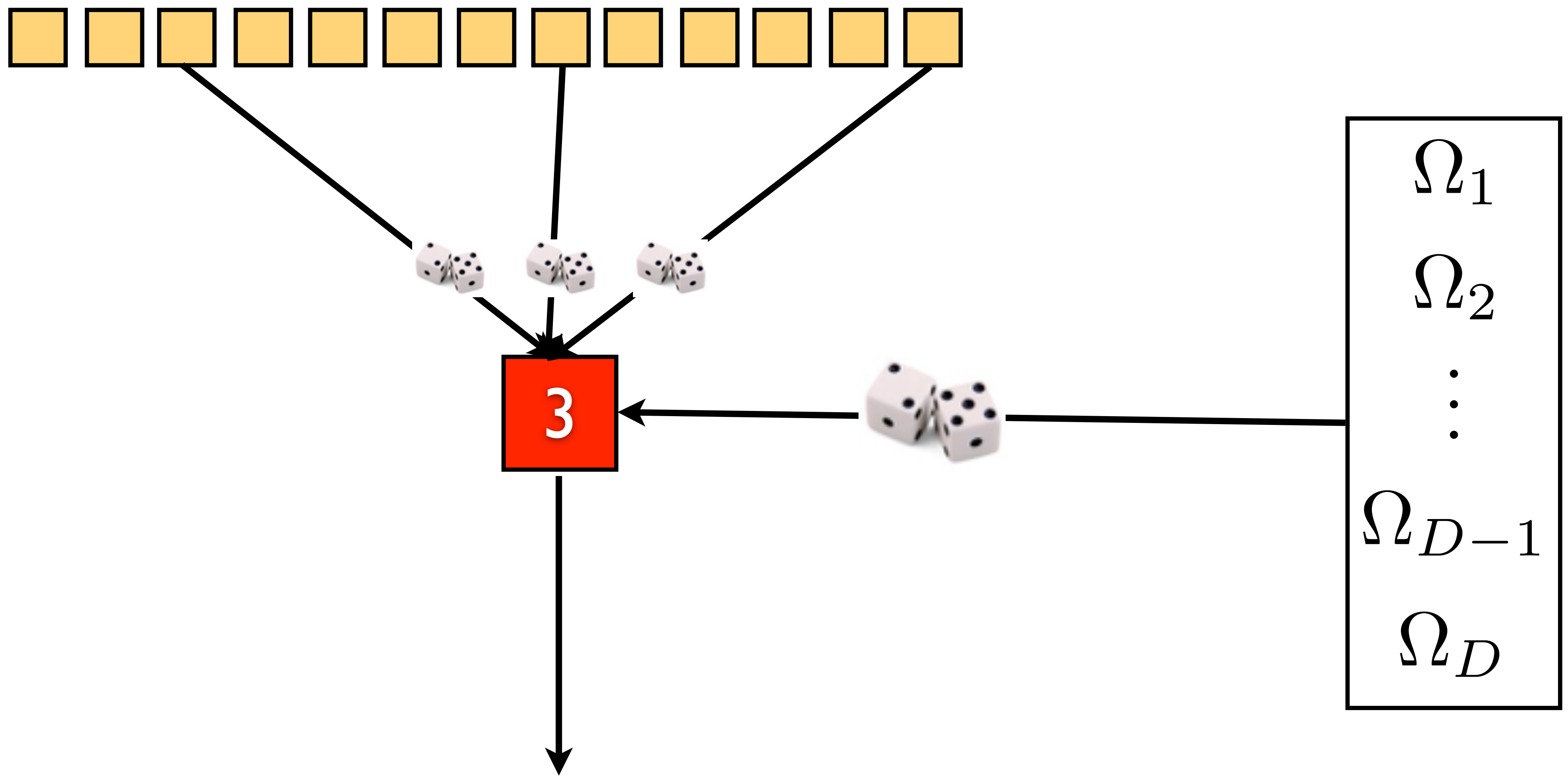


Fountain Codes

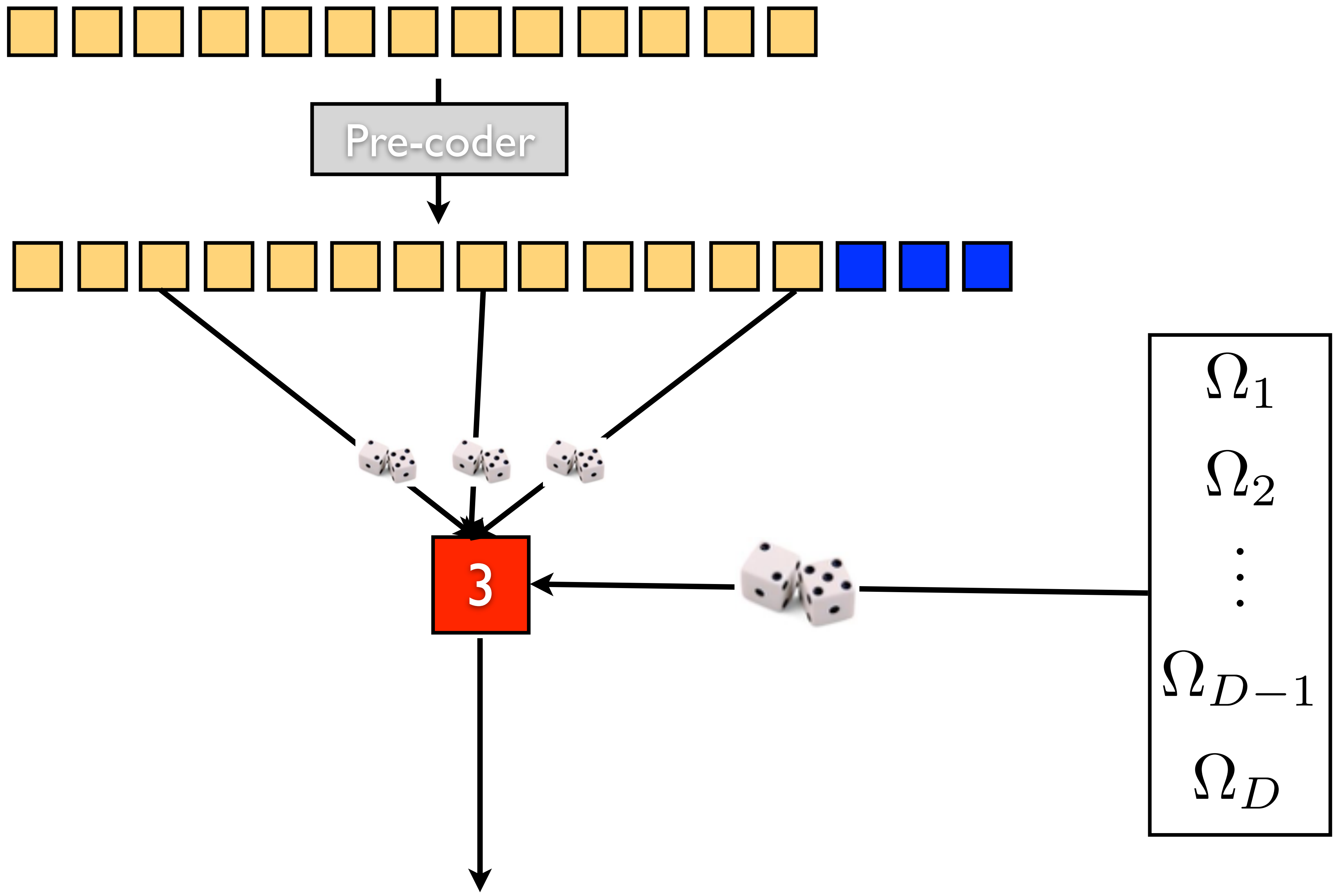


Receivers adjust to their individual loss rates

LT-Codes



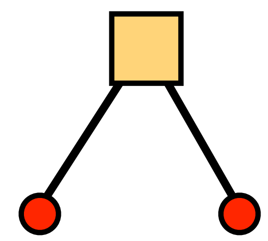
Raptor Codes



Raptor codes allow for linear time encoding/decoding

Achieving Capacity: BEC

Ω_1 has to go to zero, because information loss otherwise.

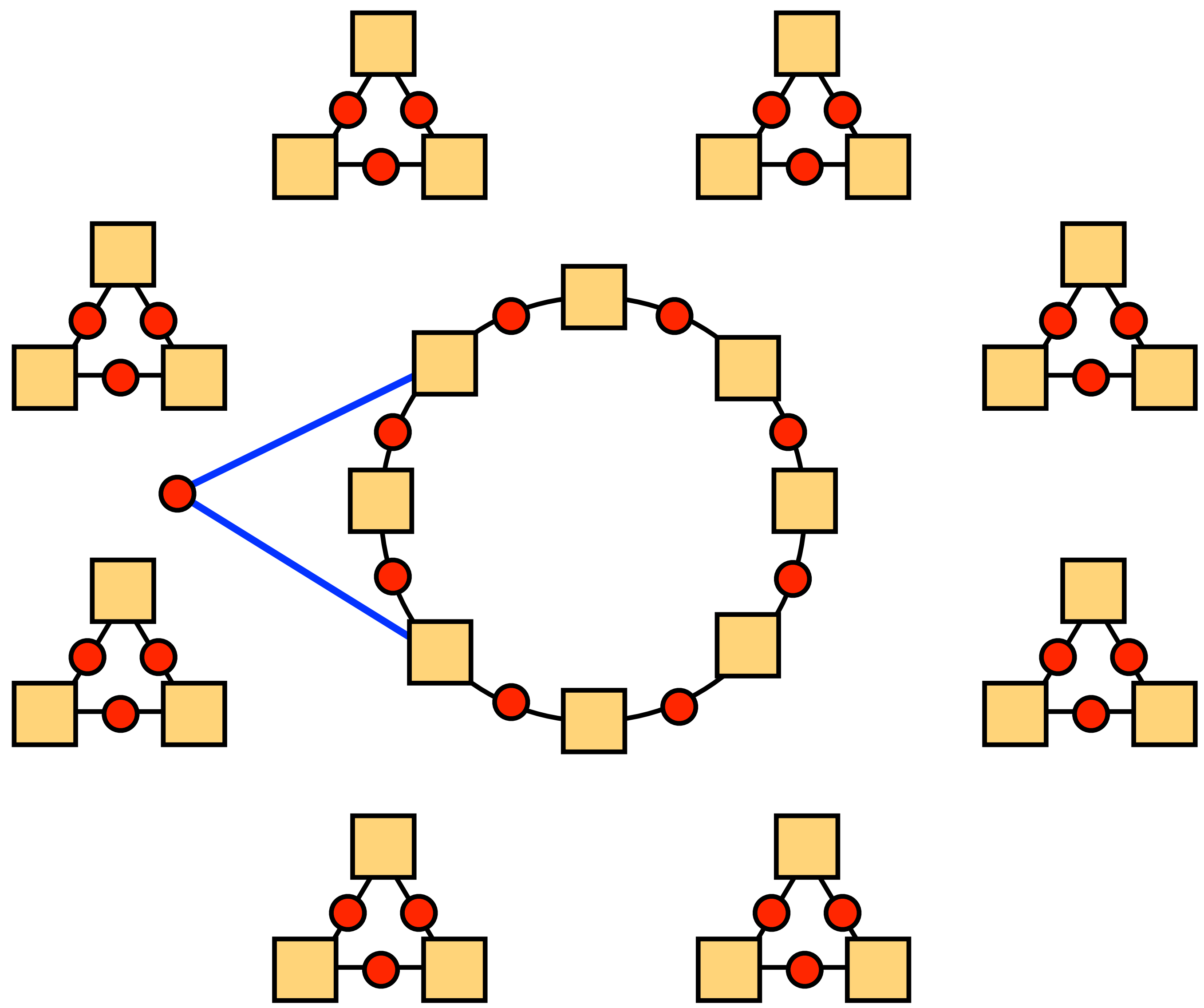


Ω_2 has to go to $\frac{1}{2}$

If graph has giant component, then new output symbol of degree 2 has both its neighbors in the component with constant probability.

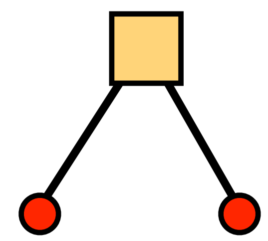
Hence, information loss.

Giant component appears iff average degree is > 1.

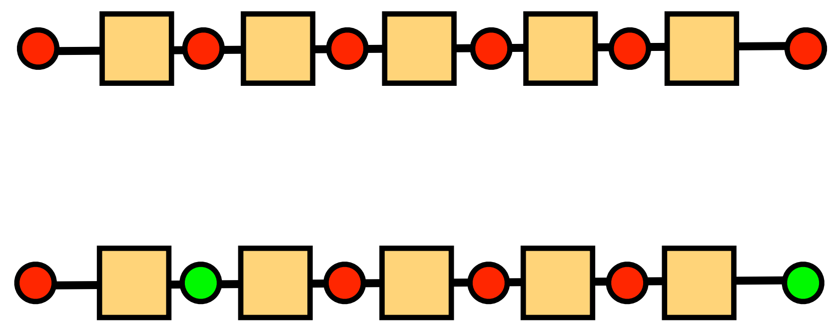


Achieving Capacity: q-SC

Ω_1 has to go to zero, because information loss otherwise.



Verified



Not verified

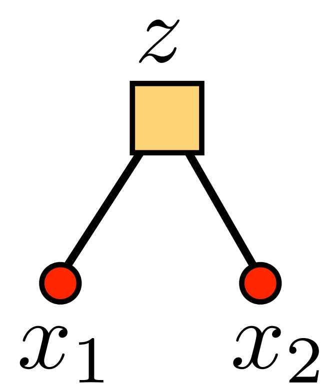
Want lots of verified paths in the induced graph.

Graph on correctly transmitted output symbols of degree 2 must have giant component.

Ω_2 has to go to $\frac{1}{2}$

Achieving Capacity: SMC

Ω_1 has to go to zero, because information loss otherwise.



$I(z; x_1, x_2) < I(z; x_1) + I(z, x_2)$

Ω_2 has to go to $\frac{1}{2}\Pi(\mathcal{C})$

$\mathcal{C}$	Binary input, memoryless, symmetric channel.
Z	LLR of the channel.

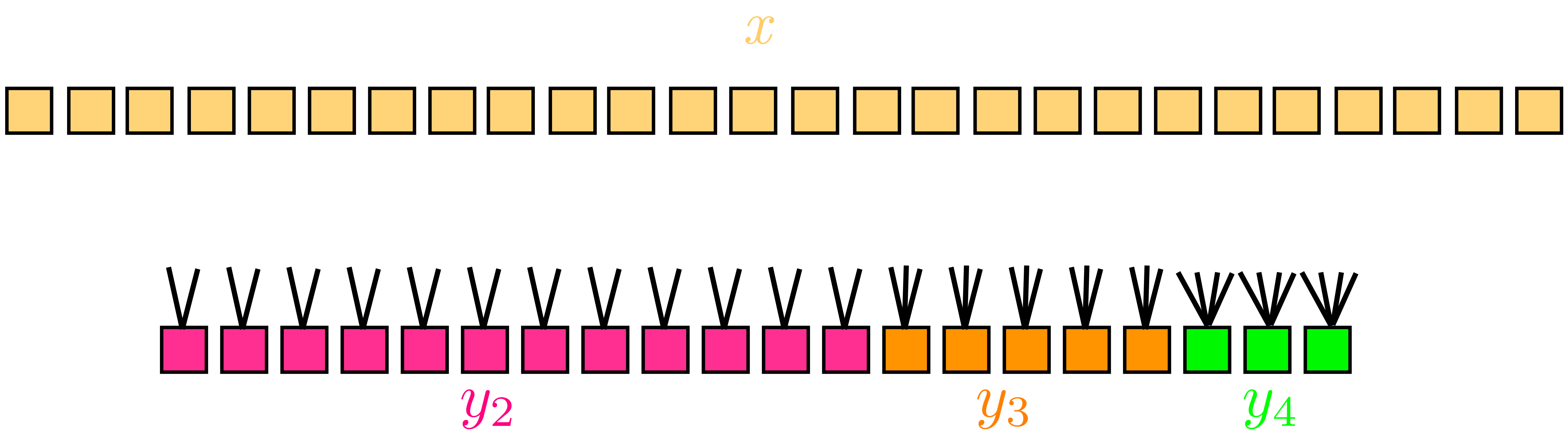
$$\Pi(\mathcal{C}) := \frac{\mathrm{Cap}(\mathcal{C})}{\mathrm{E}[\tanh(Z/2)]}$$

Proof follows same intuition as in the BEC, but uses information theoretic tools.

Examples

$\mathcal{C}$	$\Pi(\mathcal{C})$
BEC	1
$\mathrm{BSC}(p)$	$\dfrac{1-h(p)}{(1-2p)^2}$
$\mathrm{AWGN}\left(\sqrt{\frac{2}{m}}\right)$	$\dfrac{1-\frac{1}{2\sqrt{\pi m}}\int_{-\infty}^{\infty}\log_2\left(1+e^{-x}\right)e^{-\frac{(x-m)^2}{4m}}dx}{\frac{1}{2\sqrt{\pi m}}\int_{-\infty}^{\infty}\tanh\left(\frac{x}{2}\right)e^{-\frac{(x-m)^2}{4m}}dx},$

Higher Degrees?
BEC, heuristic



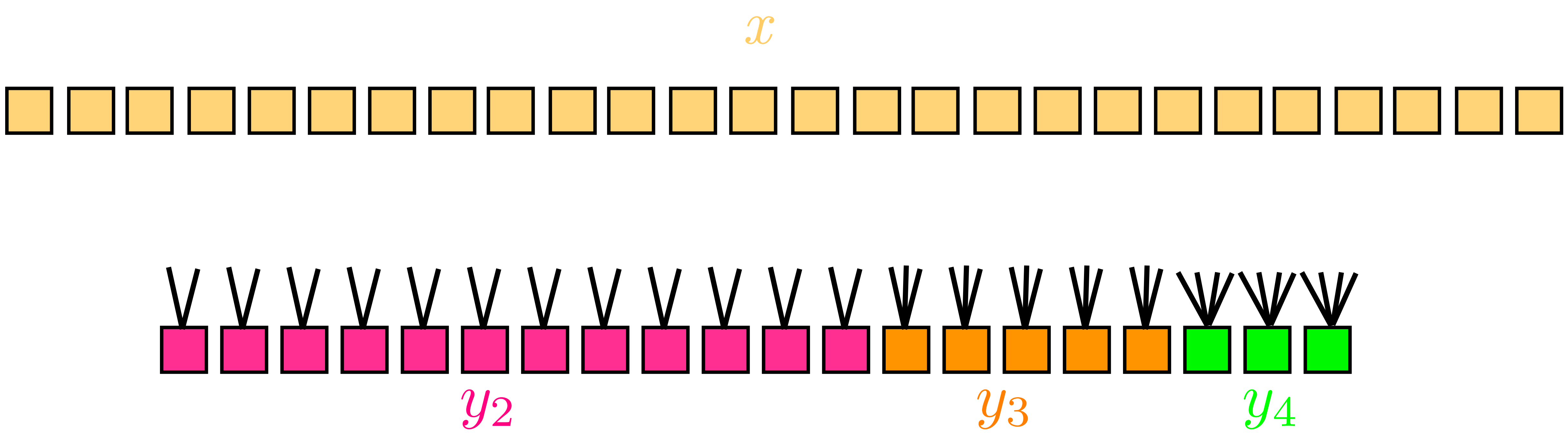
$$I(x; y_2, y_3, y_4, \ldots) = I(x; y_2) + I(x; y_3 \mid y_2) + I(x; y_4 \mid y_2, y_3) + \cdots$$

$$|y_2| = \frac{|x|}{2} \qquad |y_3| = \frac{1}{3}\left(|x| - |y_2|\right) \qquad |y_4| = \frac{1}{4}\left(|x| - |y_2| - |y_3|\right)$$

$$\Omega_2 = \frac{1}{2}, \Omega_3 = \frac{1}{6}, \ldots, \Omega_k = \frac{1}{k(k-1)}, \ldots$$

Can be made rigorous!

Higher Degrees?
SMC, heuristic
(work in progress with Luby)



$$I(x; y_2, y_3, y_4, \ldots) = I(x; y_2) + I(x; y_3 \mid y_2) + I(x; y_4 \mid y_2, y_3) + \cdots$$

$$\Omega_2 = \frac{\Pi(\mathcal{C})}{2}, \Omega_3 = \frac{\Pi(\mathcal{C})}{3}\left(1 - \Omega_2\right), \ldots, \Omega_k = \frac{\Pi(\mathcal{C})}{k}\left(1 - \sum_{j=1}^{k-1} \Omega_j\right), \ldots$$

$$\Omega(x) = \frac{1 - (1-x)^{\Pi(\mathcal{C})} - \Pi(\mathcal{C})x}{1 - \Pi(\mathcal{C})}$$

Do Capacity-Achieving Sequences carry over?

BEC		SMC
$$\Omega(x) = \sum_{k=1}^{\infty} \frac{x^k}{k(k-1)}$$	$\xrightarrow{?}$	$$\Omega(x) = \frac{1-(1-x)^{\Pi(\mathcal{C})} - \Pi(\mathcal{C})x}{1-\Pi(\mathcal{C})}$$

At least left-hand formula converges to right-hand one when the channel approaches the BEC.

Graph-theoretic Design
BEC

- The connection to the underlying graph (hyper-graph) allows a better design for the codes.

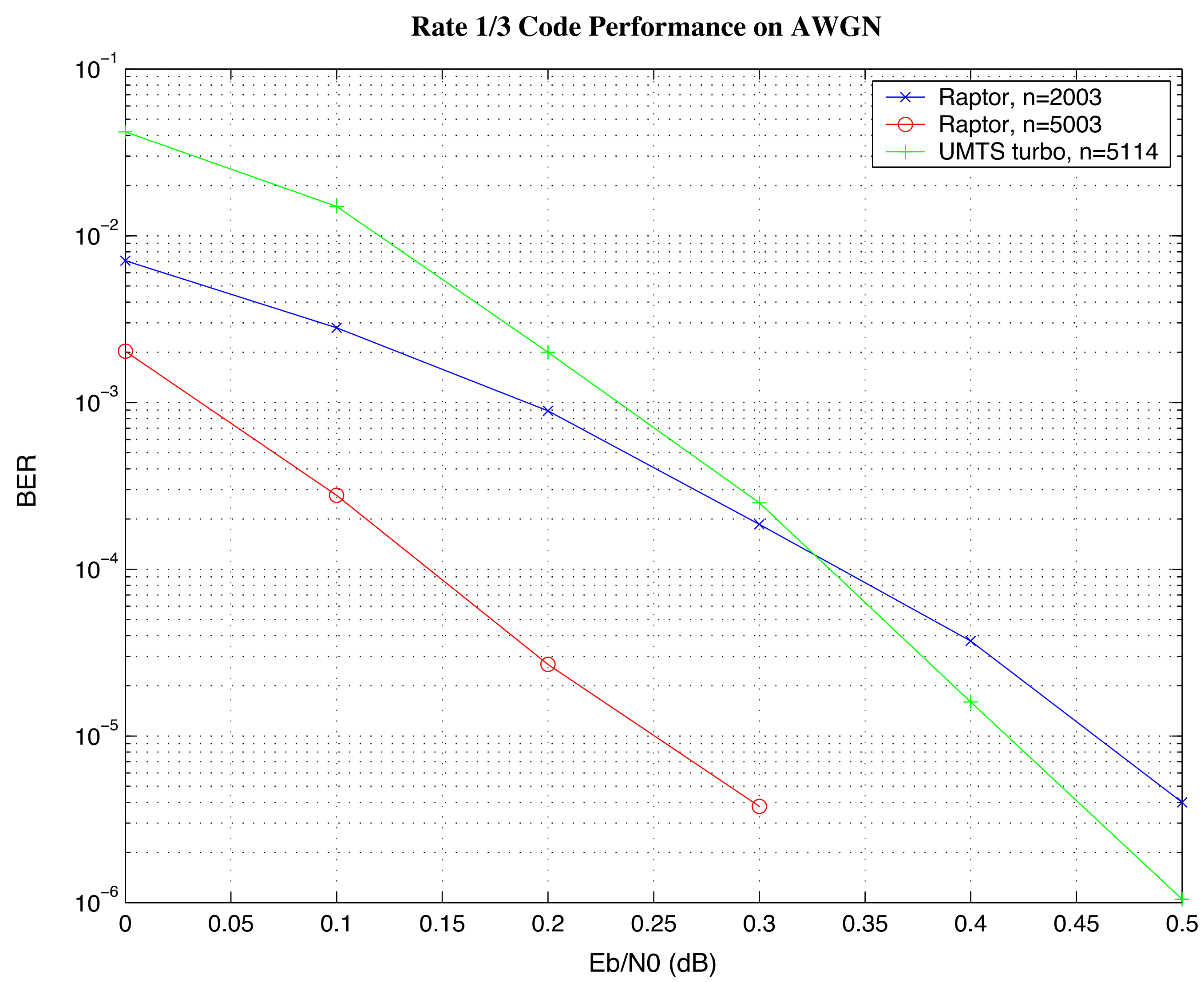
- A particular Raptor code was designed for the 3GPP MBMS Standards Body based on these principles.

- That code has been chosen as the sole mandatory FEC for all multicast/broadcast services within 3GPP MBMS. Future MBMS enabled devices will have this code embedded into their hardware.

Graph-theoretic Design
AWGN
(Work in progress with Pakzad)

- Similar design principles can be carried over to Raptor codes on other channels such as the AWGN channel.

- Corresponding design is better than, e.g., the UMTS Turbo standard code even for short lengths, and has a similar decoding complexity as the Turbo code.



Conclusions

- The connection to the BEC has proven quite fruitful.
- Many of the theoretical and practical design principles for codes on the BEC carry over to codes on more complicated channels.
- This talk concentrated on some provable and some unproved connections between the BEC and more complicated binary channels.
- The connection used results from the theory of random graphs.
- Other channels, such as the q-ary symmetric channel for large q establish more explicit provable connections to the BEC and allow the design of provably capacity-achieving codes.