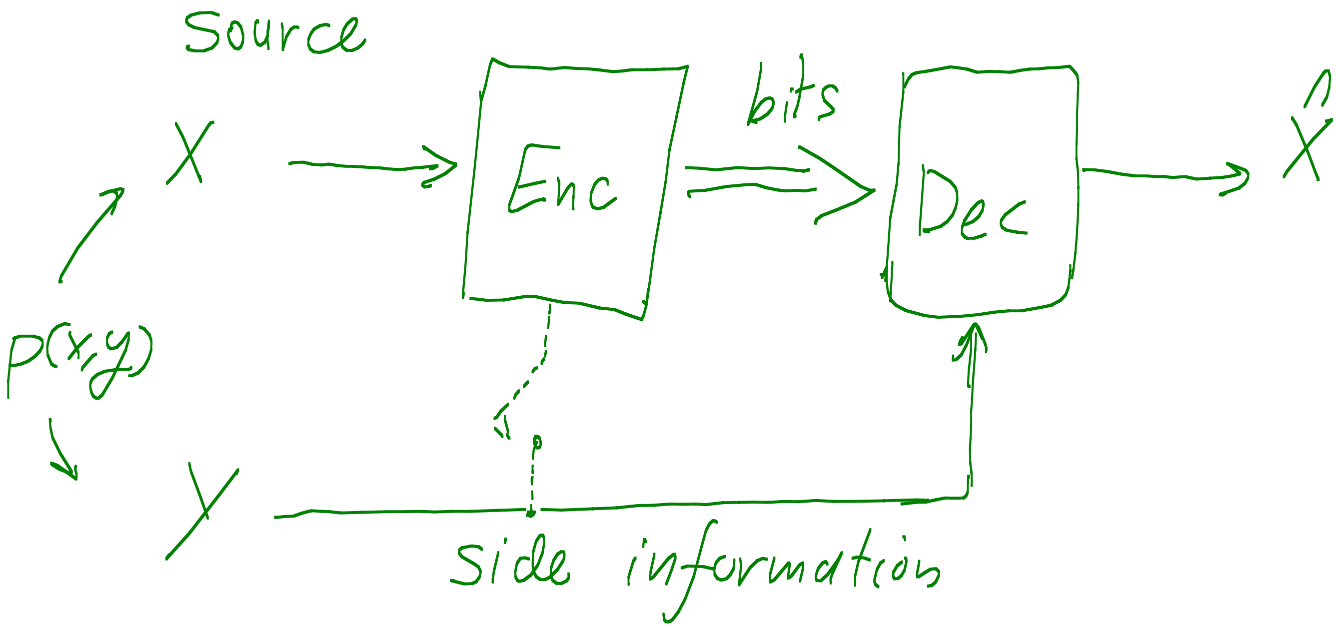
Lattice Wyner-Ziv Coding

Rami Zamir @ ETH

The Wyner-Ziv Problem
(Source coding with S.I. @ Decoder)



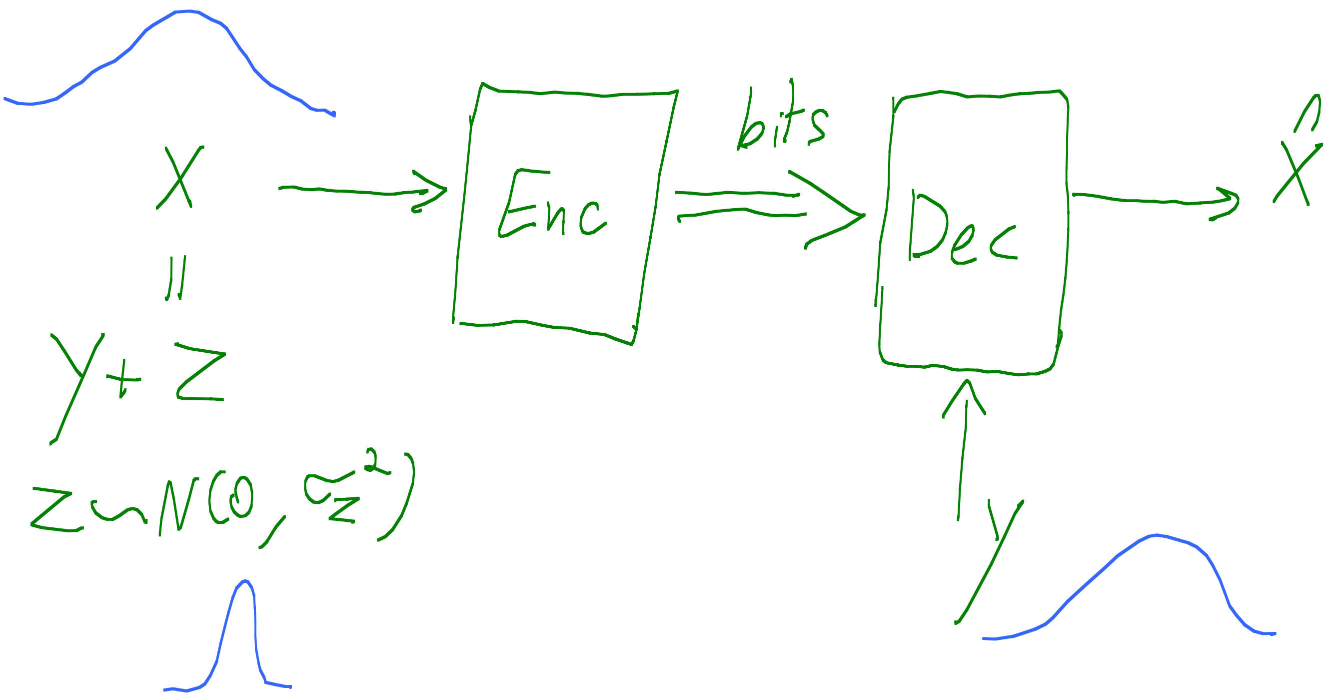
$$R^{WZ}_{X|Y}(D) = \min I(X;U|Y) \qquad \frac{\text{bit}}{\text{source sample}}$$

Wyner-Ziv 1976

minimum over U and $g(\cdot)$ such that

$$U \leftrightarrow X \leftrightarrow Y, \quad E\{d(g(U,Y)-X)\} \le D$$

The Quadratic-Gaussian WZ



$$R^{WZ}_{X|Y}(D) = R_Z(D) = \frac{1}{2}\log\left(\frac{\sigma_Z^2}{D}\right) \quad \frac{\text{bit}}{\text{source sample}}$$

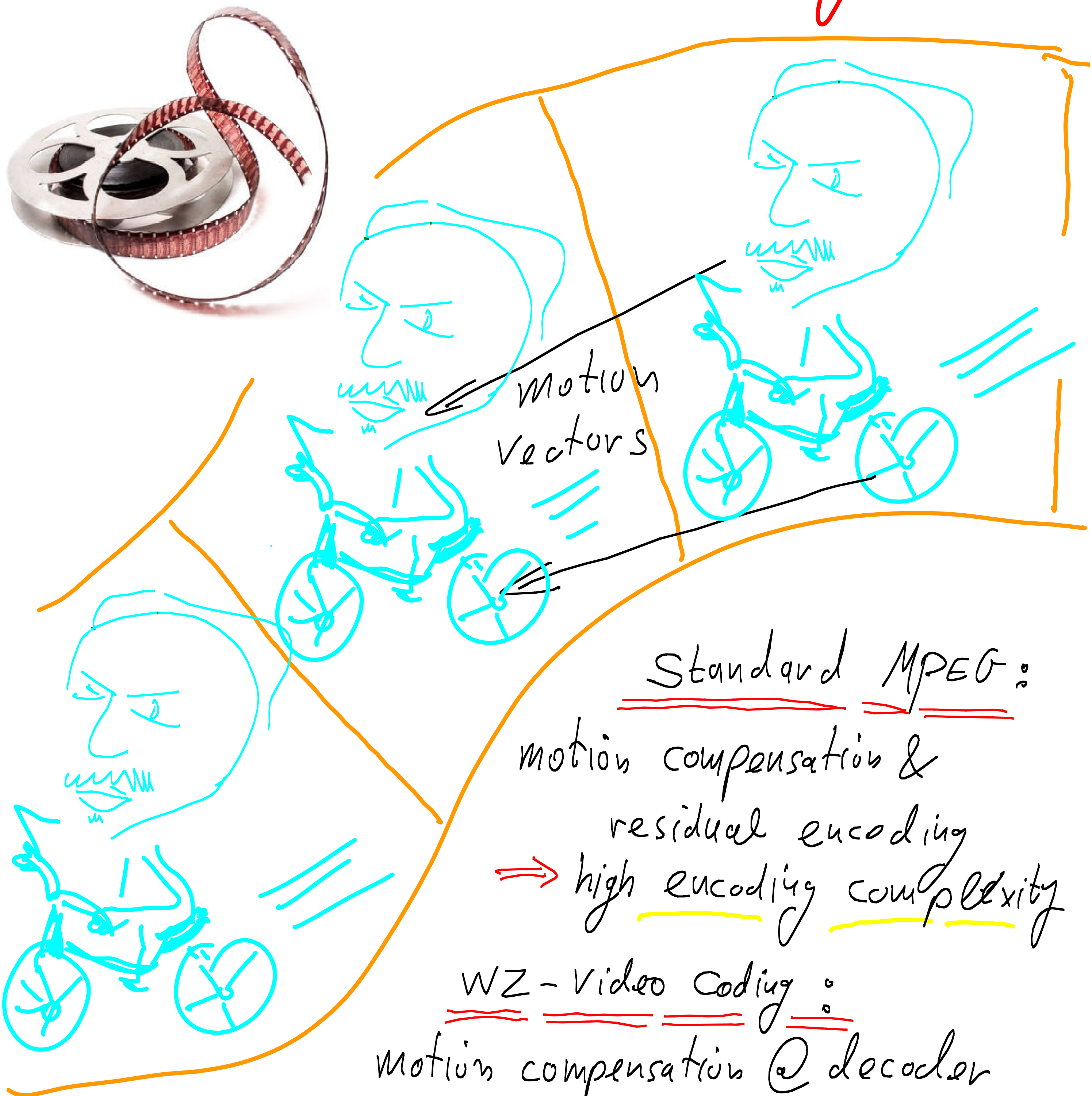
$$= R_{X|Y}(D)$$

Wyner-Ziv 1976

Wyner 1978

Note: for non-GQ $R^{WZ}_{X|Y}(D) > R_{X|Y}(D)$ (Rate loss)

Wyner-Ziv Video Coding



Standard MPEG:

motion compensation &
residual encoding
⇒ high encoding complexity

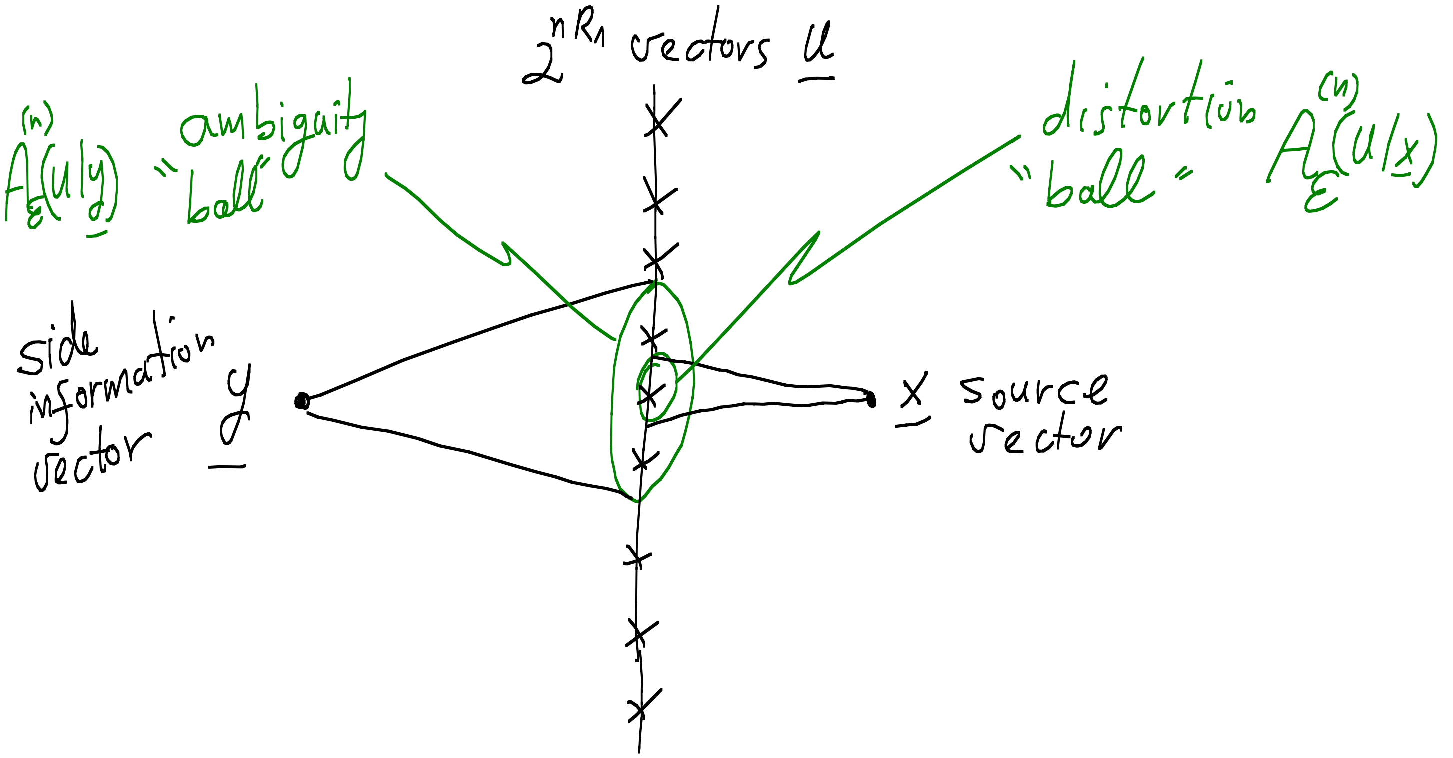
WZ-Video Coding:

motion compensation @ decoder
⇒ encoding = simple / decoding = complex

Random Binning Scheme

WZ Formula: $R^{WZ}_{X|Y}(D) = \min I(X;Y|U) = \min [I(X;U) - I(Y;U)]$

$U \leftrightarrow X \leftrightarrow Y$



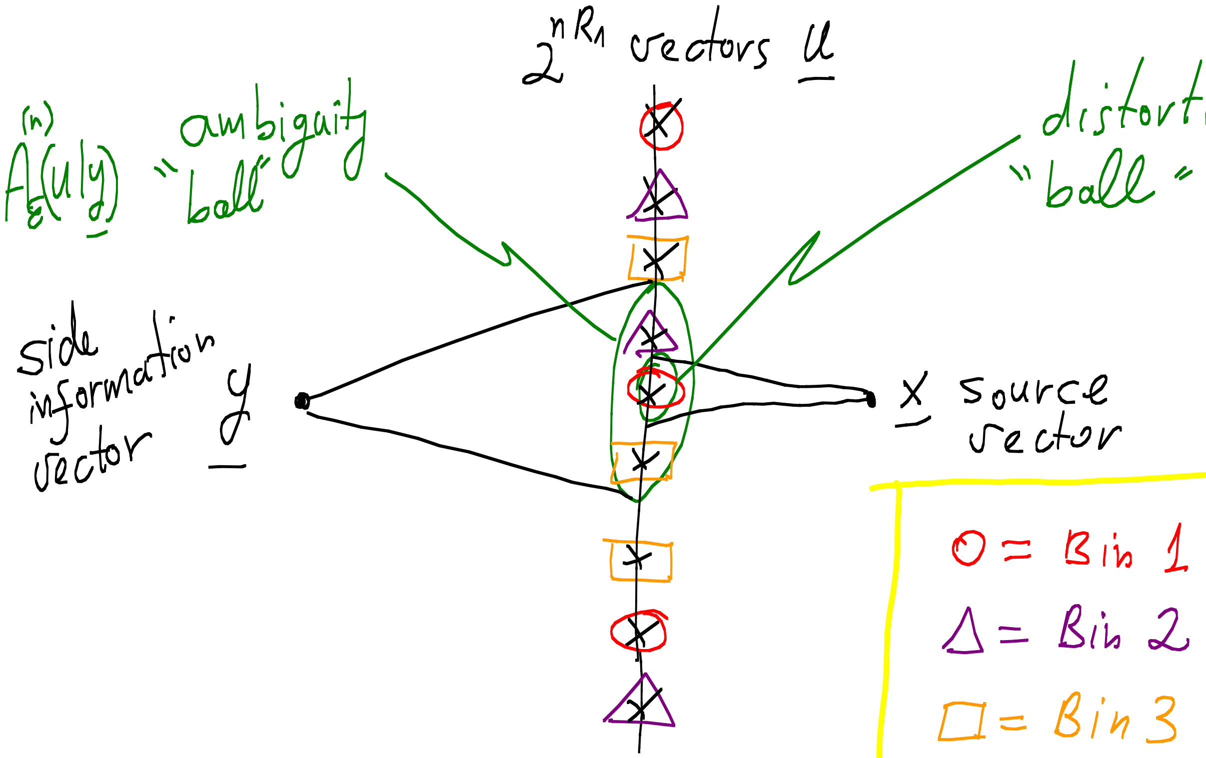
$R_1 > I(X;U) \Rightarrow$ at least one jointly-typical $\underline{u}$

$(\underline{u} \in A(U|\underline{x}))$

Random Binning Scheme

WZ Formula: $R^{WZ}_{X|Y} = \min I(X;Y|U) = \min\left[I(X;U) - I(Y;U)\right]$

$$U \leftrightarrow X \leftrightarrow Y$$



$\bigcirc$ = Bin 1
$\triangle$ = Bin 2
$\square$ = Bin 3
$+$ = ⋮

2^{nR} Bins

$R_1 > I(X;U) \Rightarrow$ at least one jointly-typical $\underline{u}$ ($\underline{u} \in A(U|\underline{x})$)

$R_1 < I(Y;U) + R \Rightarrow Pr\left\{\begin{array}{l}\text{false } \underline{u} \text{ typical} \\ \text{with } \underline{y} \text{ inside Bin}\end{array}\right\} \to 0$

$\therefore$ $R \simeq I(X;U) - I(X;Y)$ is achievable!

Nested Lattices

$$\Lambda_2 \subset \Lambda_1 \Rightarrow \underline{\underline{G}}_2 = \underline{\underline{G}}_1 \cdot \underline{\underline{J}}$$

course lattice

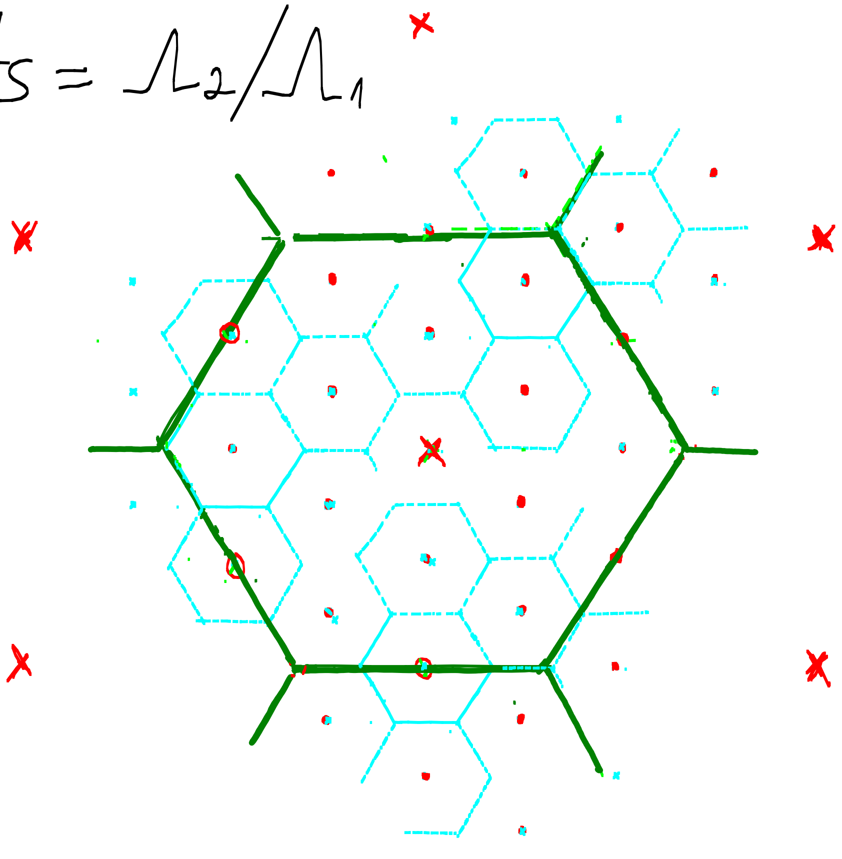
fine lattice

integer matrix

$$\text{Nesting Ratio} = \left(\frac{V_2}{V_1}\right)^{1/K} = |\det(J)|^{1/K}$$

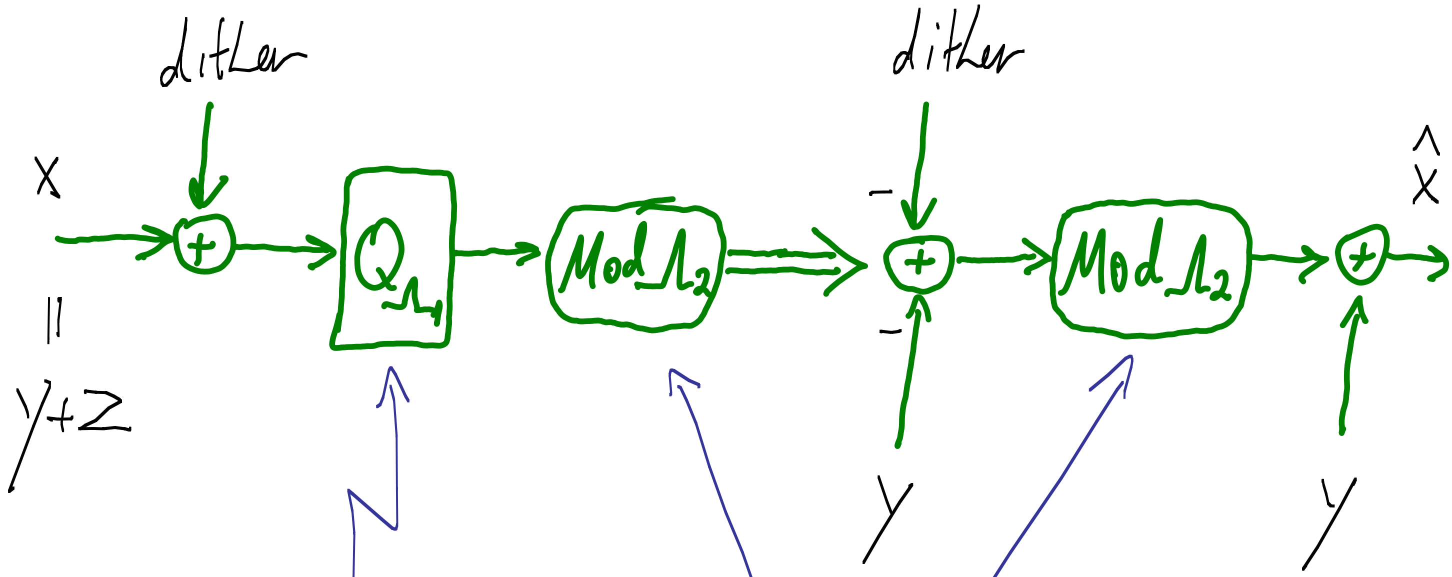
$$\text{Relative Cosets} = \Lambda_2 / \Lambda_1$$

4:1



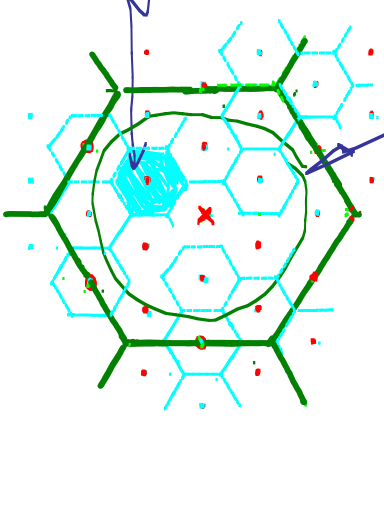
Lattice Wyner-Ziv Coding

[Z & Shamai Verdu]



Good quantizer for desired distortion: $\sigma^2(\Lambda_1) = D$

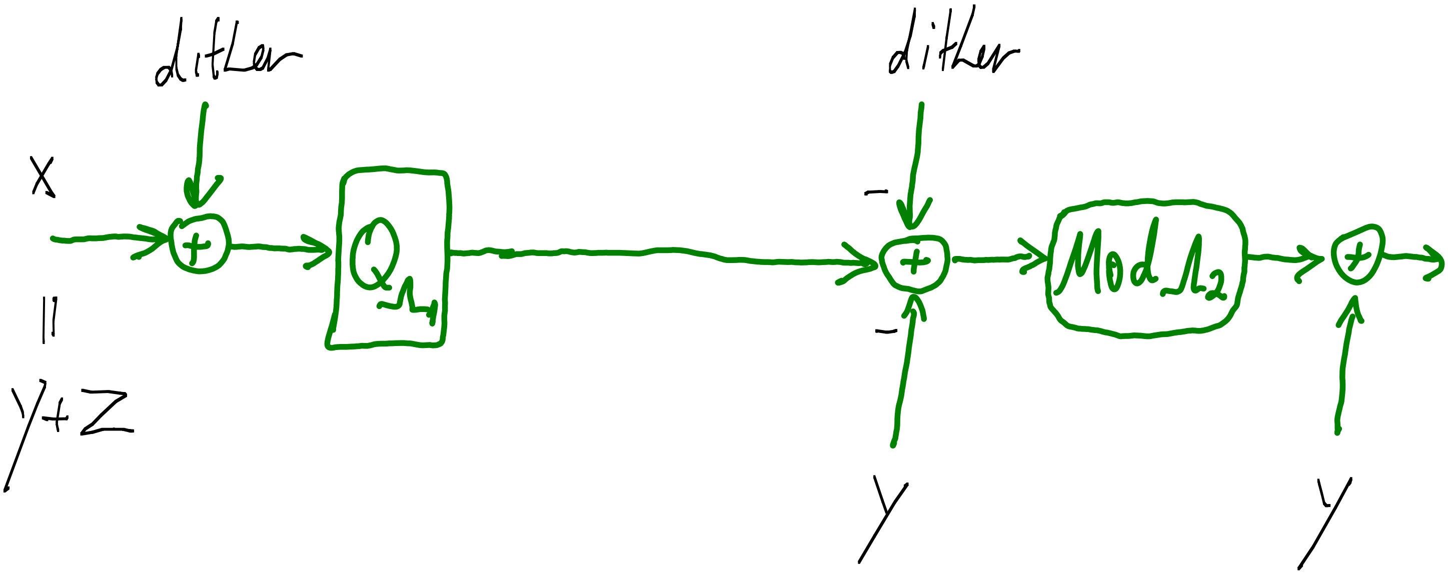
Good channel code for the noise Z: $P_e(\Lambda_2, \sigma_Z^2) < \varepsilon$



Lattice Wyner-Ziv Coding

$$(A \bmod \Lambda + B) \bmod \Lambda = (A+B) \bmod \Lambda$$

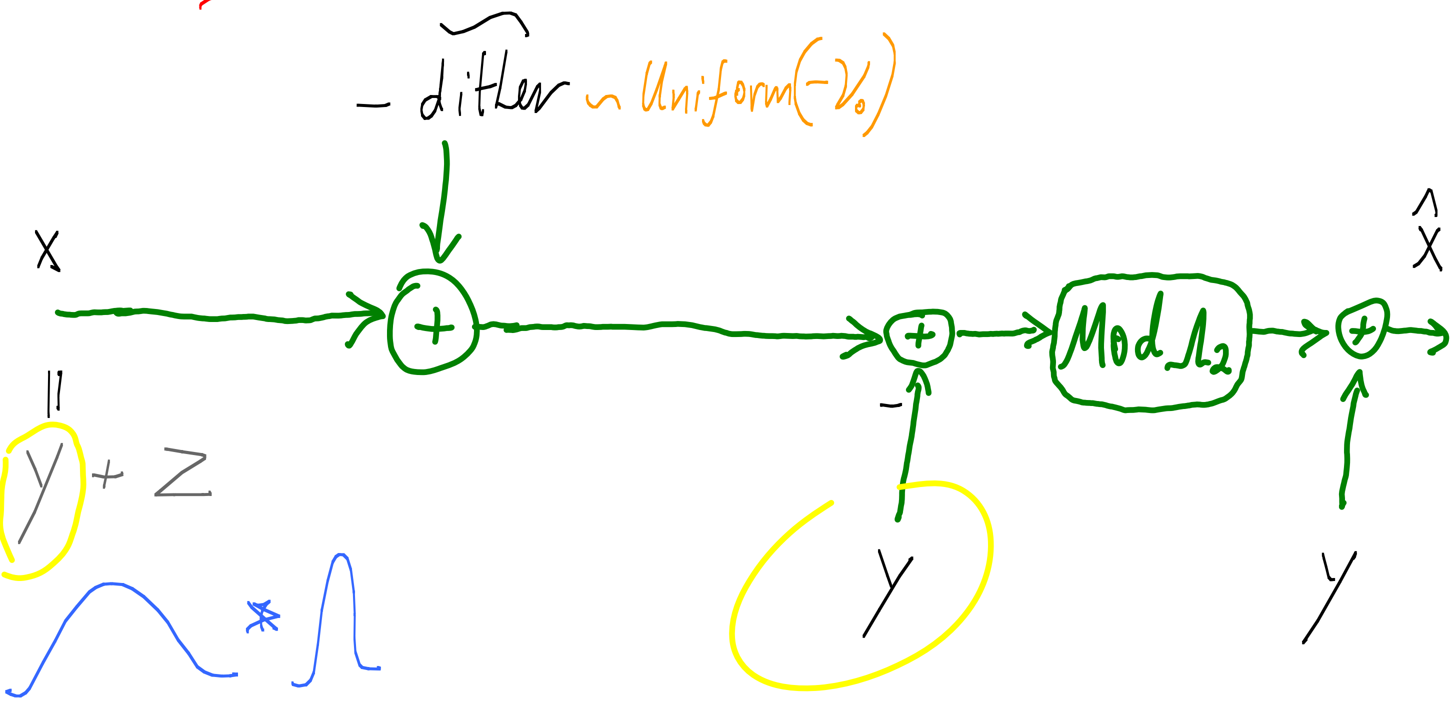
$\Longrightarrow$



Lattice Wyner-Ziv Coding

dithered quantization $\equiv$ additive noise

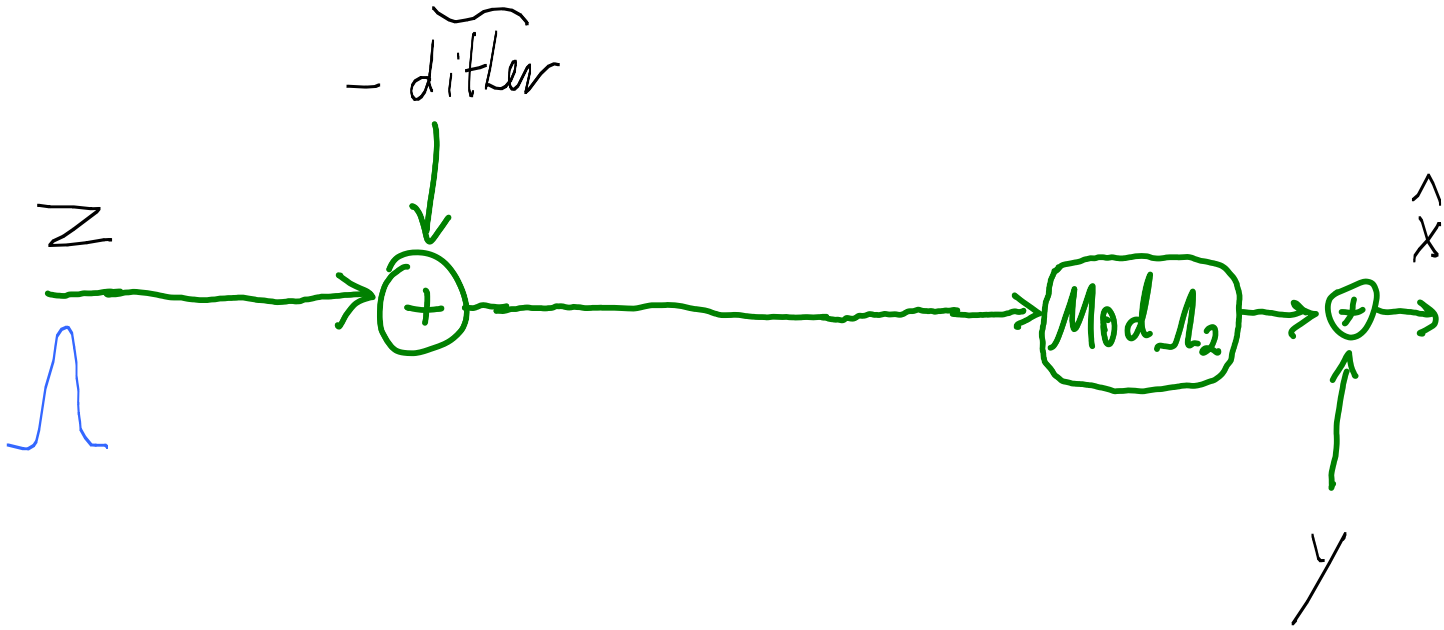
$\Longrightarrow$



Lattice Wyner-Ziv Coding

dithered quantization $\equiv$ additive noise

$\Rightarrow$

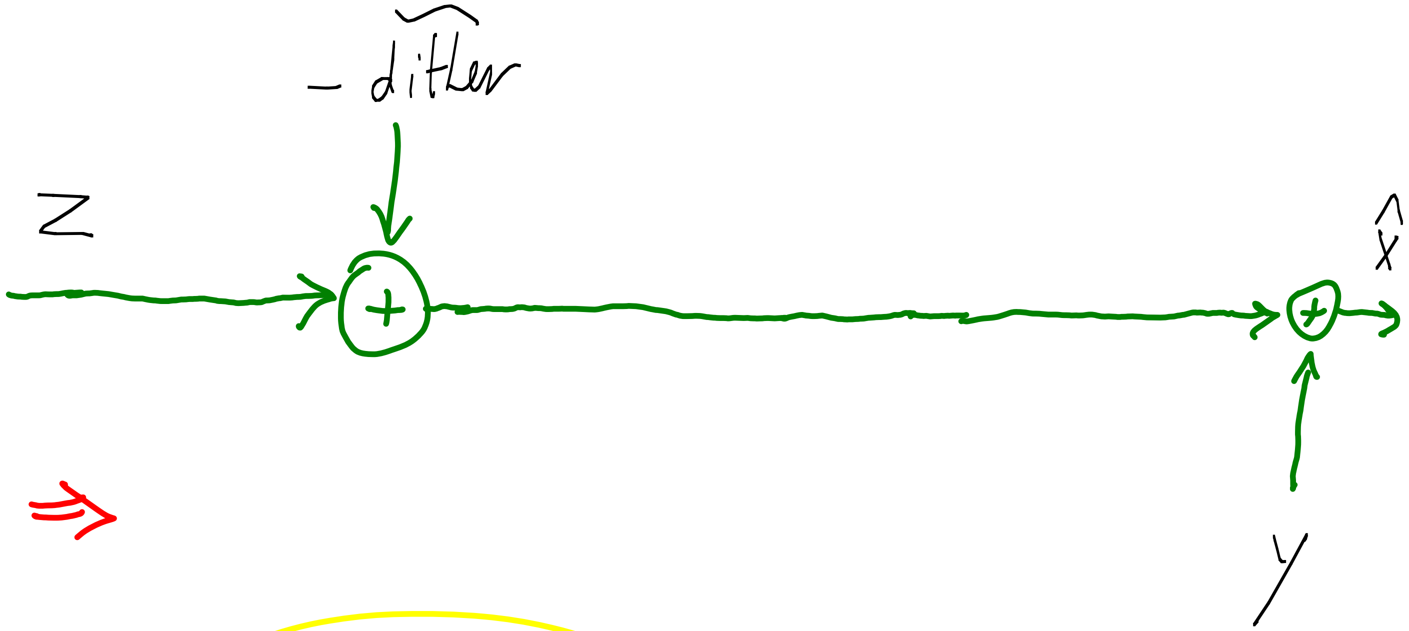


Lattice Wyner-Ziv Coding

Λ_2 = good channel code for $Z \sim \mathcal{N}(0, \sigma_z^2)$.

$D \ll \sigma_z^2$.

⇒ with prob. $> 1 - \varepsilon$,



⇒

$$\hat{X} = X - \widetilde{\text{dither}}, \quad \text{w.p.} > 1 - \varepsilon$$

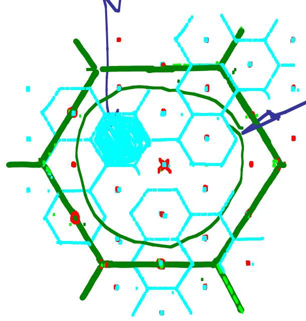
⇒ distortion $= \sigma^2(\Lambda_1) = D$

Lattice Wyner-Ziv Coding

Nesting Ratio:

$\sigma^2(\Lambda_1) = D$

$P_e(\Lambda_2, \sigma_Z^2) < \varepsilon$



$\Longrightarrow$

$$G(\Lambda_1) \triangleq \frac{\sigma^2(\Lambda_1)}{V_1^{2/K}}$$

$$\mu(\Lambda_2, P_e) \triangleq \frac{V_2^{2/K}}{\sigma_Z^2}$$

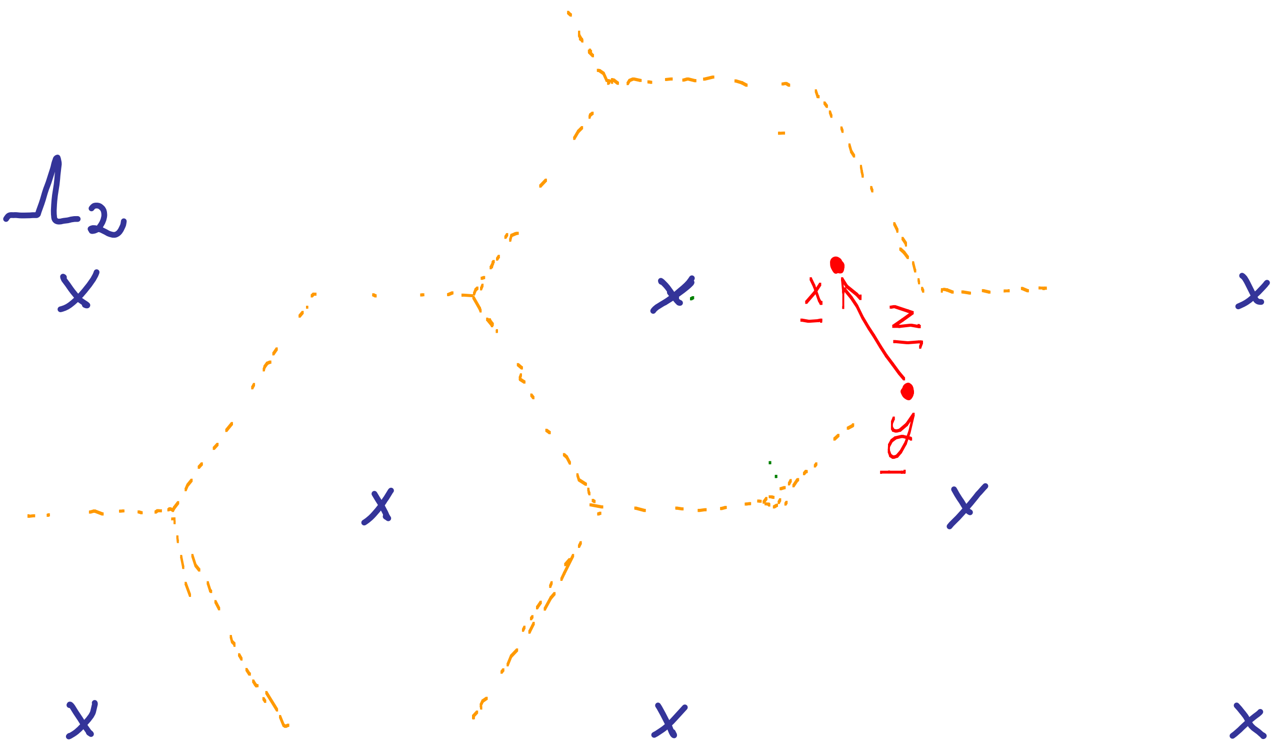
$$\text{Rate} = \frac{1}{k} \log\left(\frac{V_2}{V_1}\right) \text{ bit/sample}$$

$$= \underbrace{\frac{1}{2} \log\left(\frac{\sigma_Z^2}{D}\right)}_{R_Z(D)} + \underbrace{\frac{1}{2} \log\left(G(\Lambda_1) \cdot \mu(\Lambda_2, \varepsilon)\right)}_{\text{Redundancy} \longrightarrow 0}$$

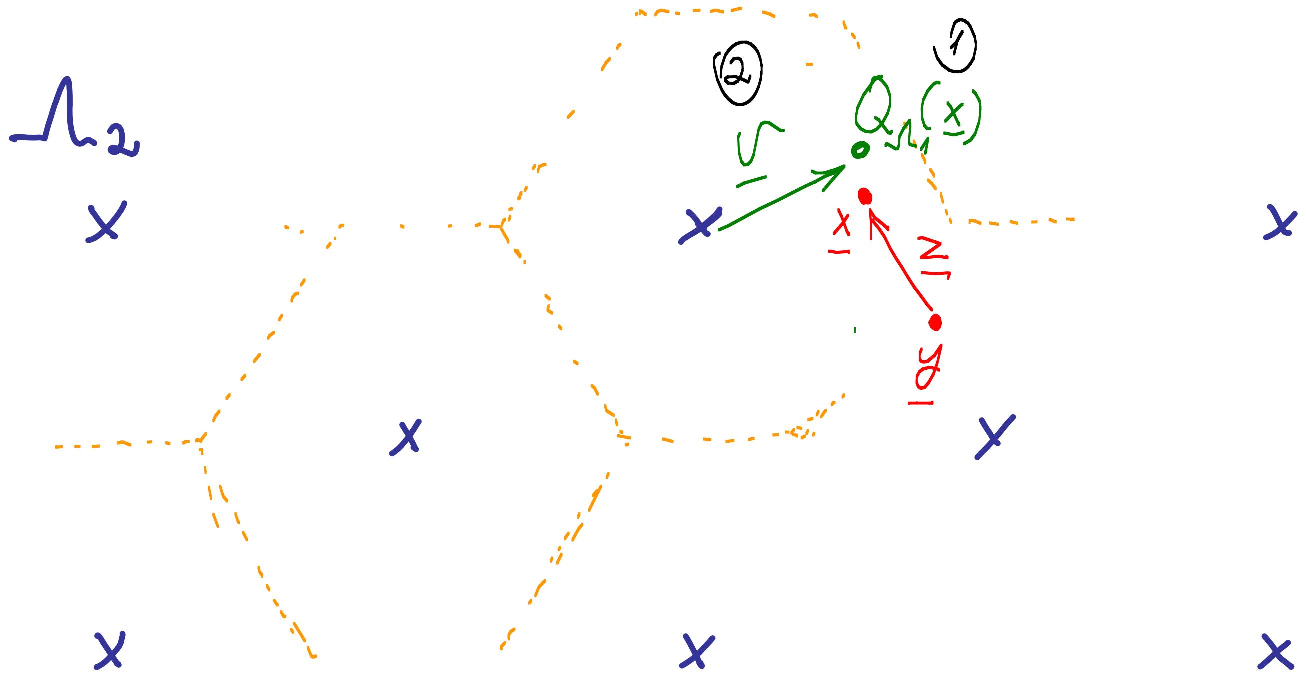
$K \to \infty$

for good lattices....

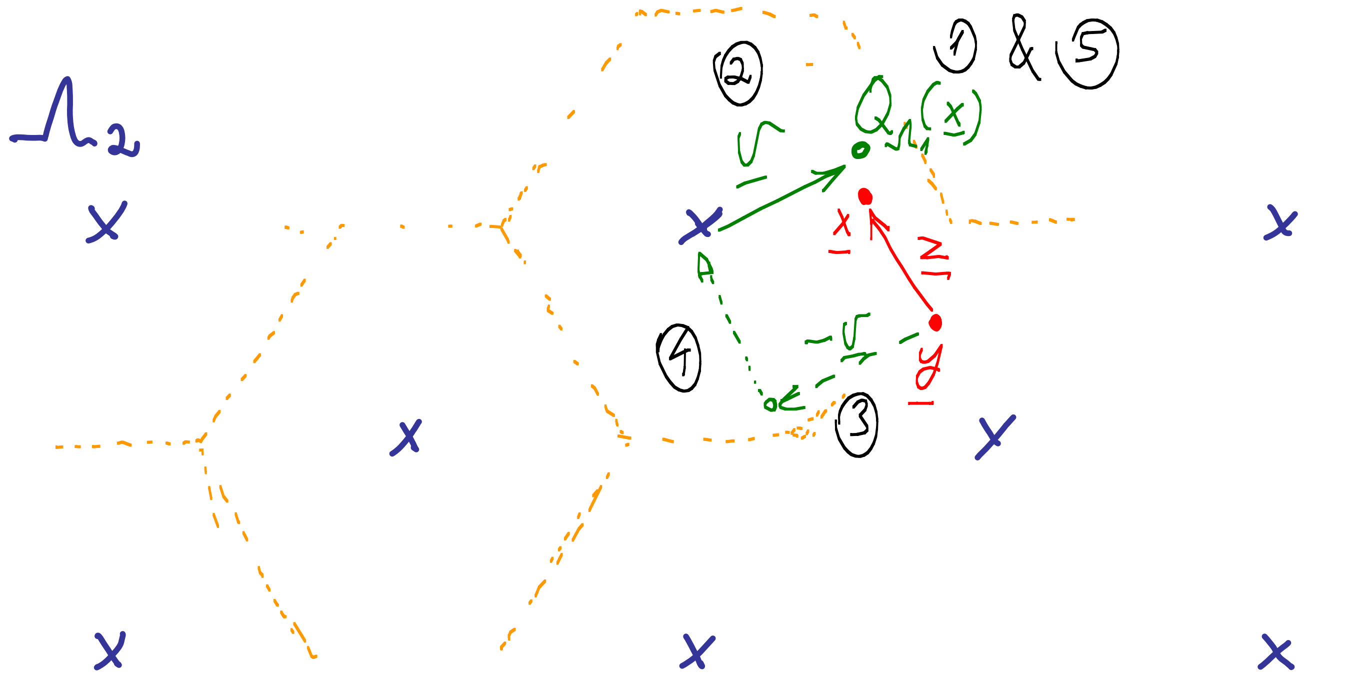
Geometric Picture of Lattice Encoding & Dec



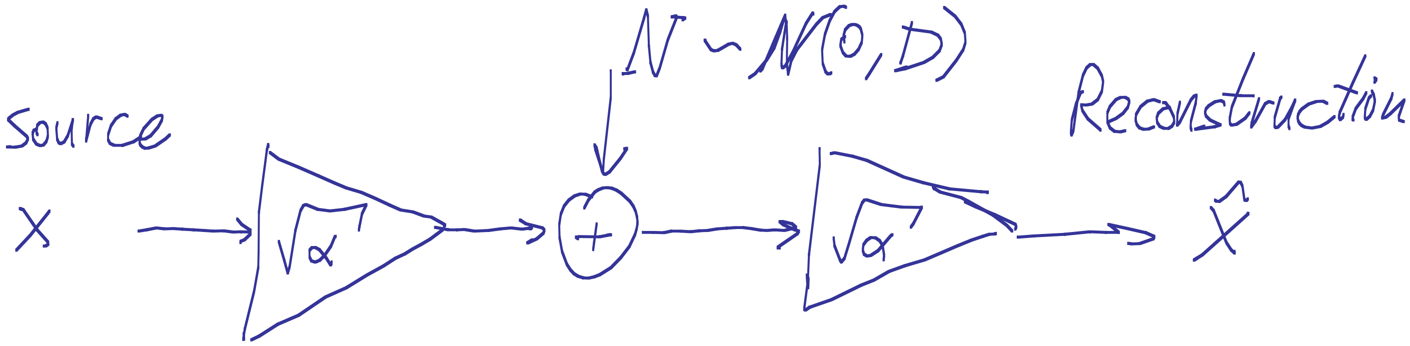
Geometric Picture of Lattice Encoding & Dec



Geometric Picture of Lattice Encoding & Dec



Forward Channel Realization of QG - R(D)



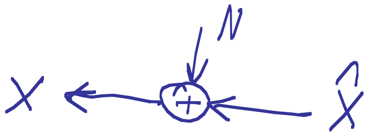
$$\alpha = 1 - D/\sigma_x^2 \qquad \hat{X} = \alpha X + \sqrt{\alpha}\, Z$$

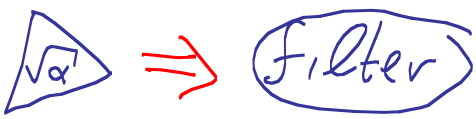
$\Rightarrow$ distortion: $E(\hat{X} - X)^2 = D$

information rate : $I(X;\hat{X}) = \frac{1}{2}\log\left(\frac{\sigma_x^2}{D}\right)$

$= I(\sqrt{\alpha}X;\ \sqrt{\alpha}X + N) = R(D)$

Remarks :

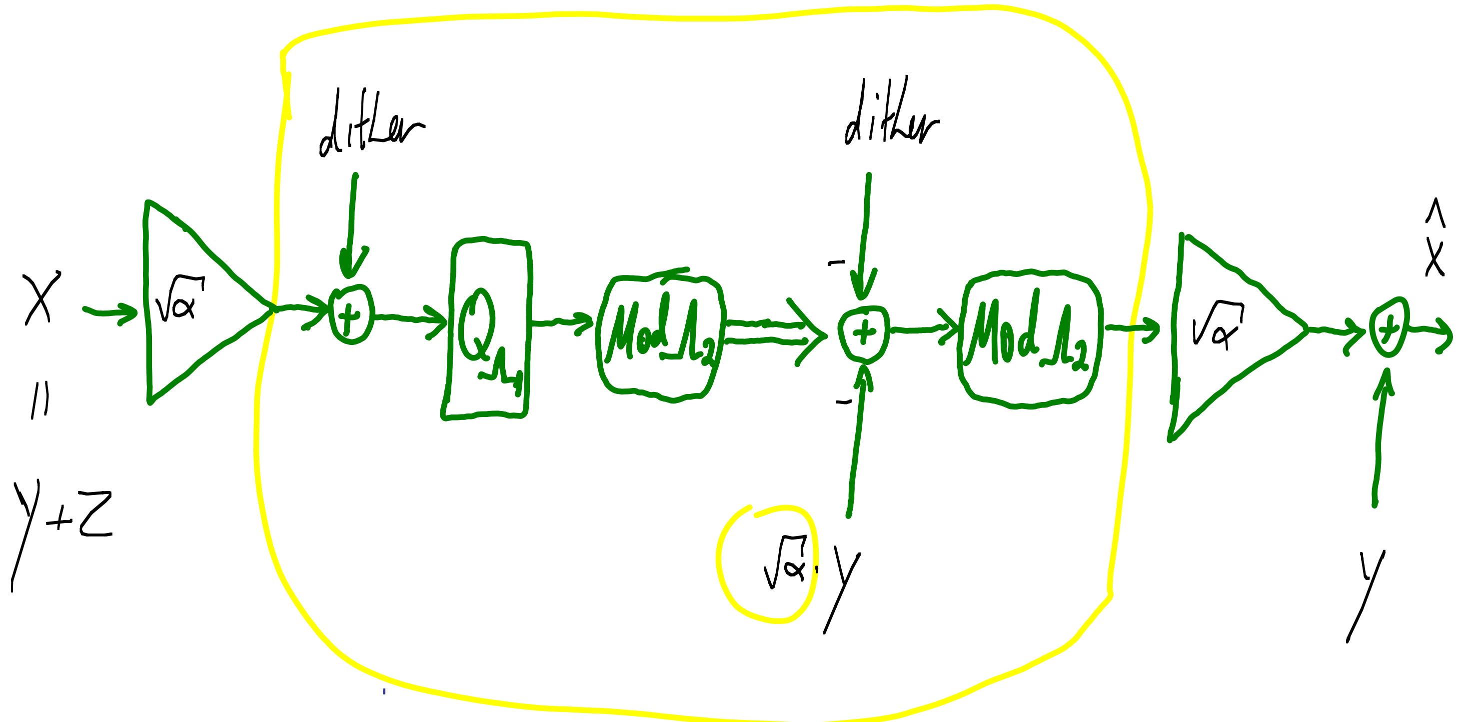
1. Backward realization 

2. Source with memory 

("water pouring")

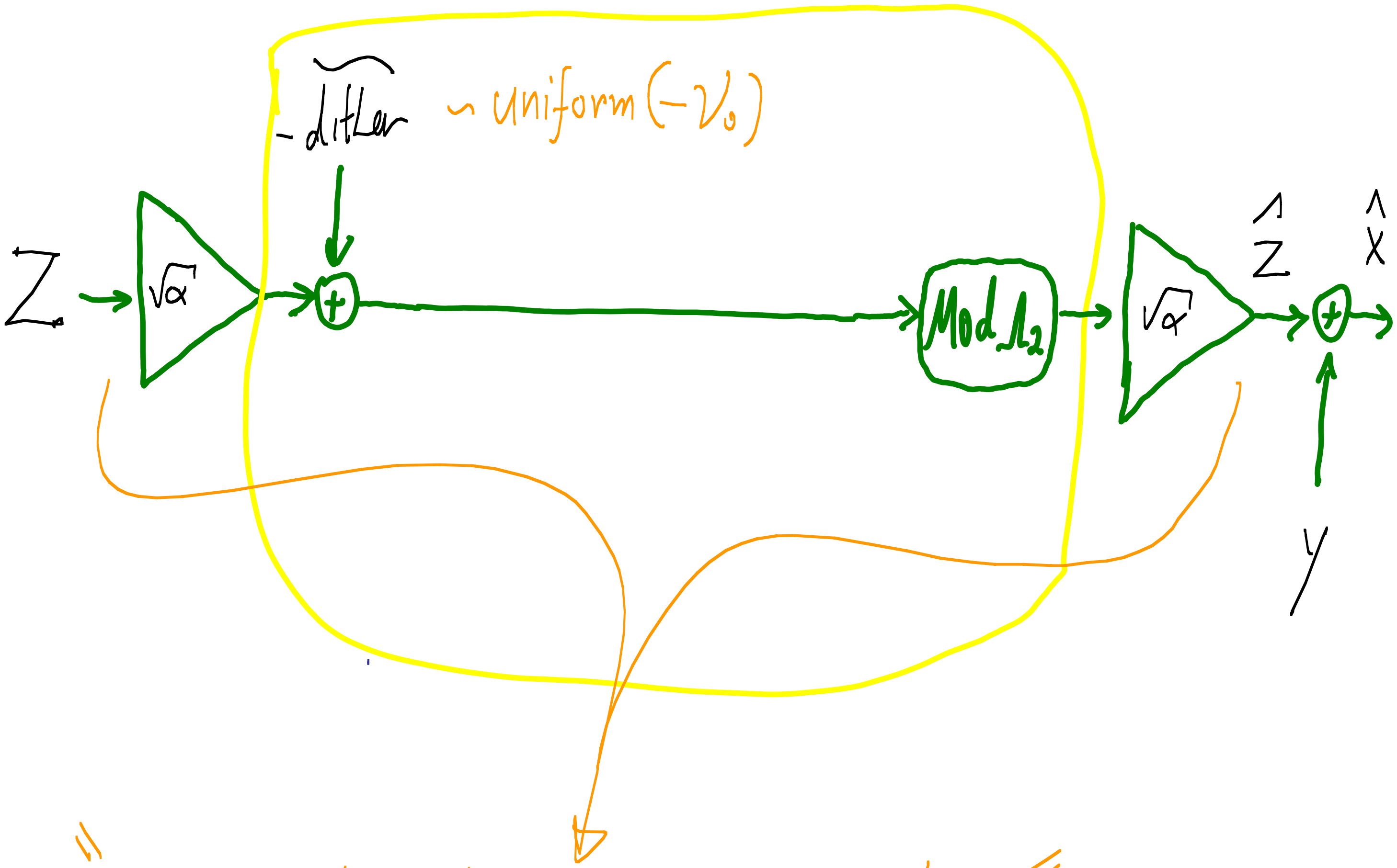
Lattice-WZ Coding: General Resolution

$\cancel{D \ll \sigma_Z^2}$



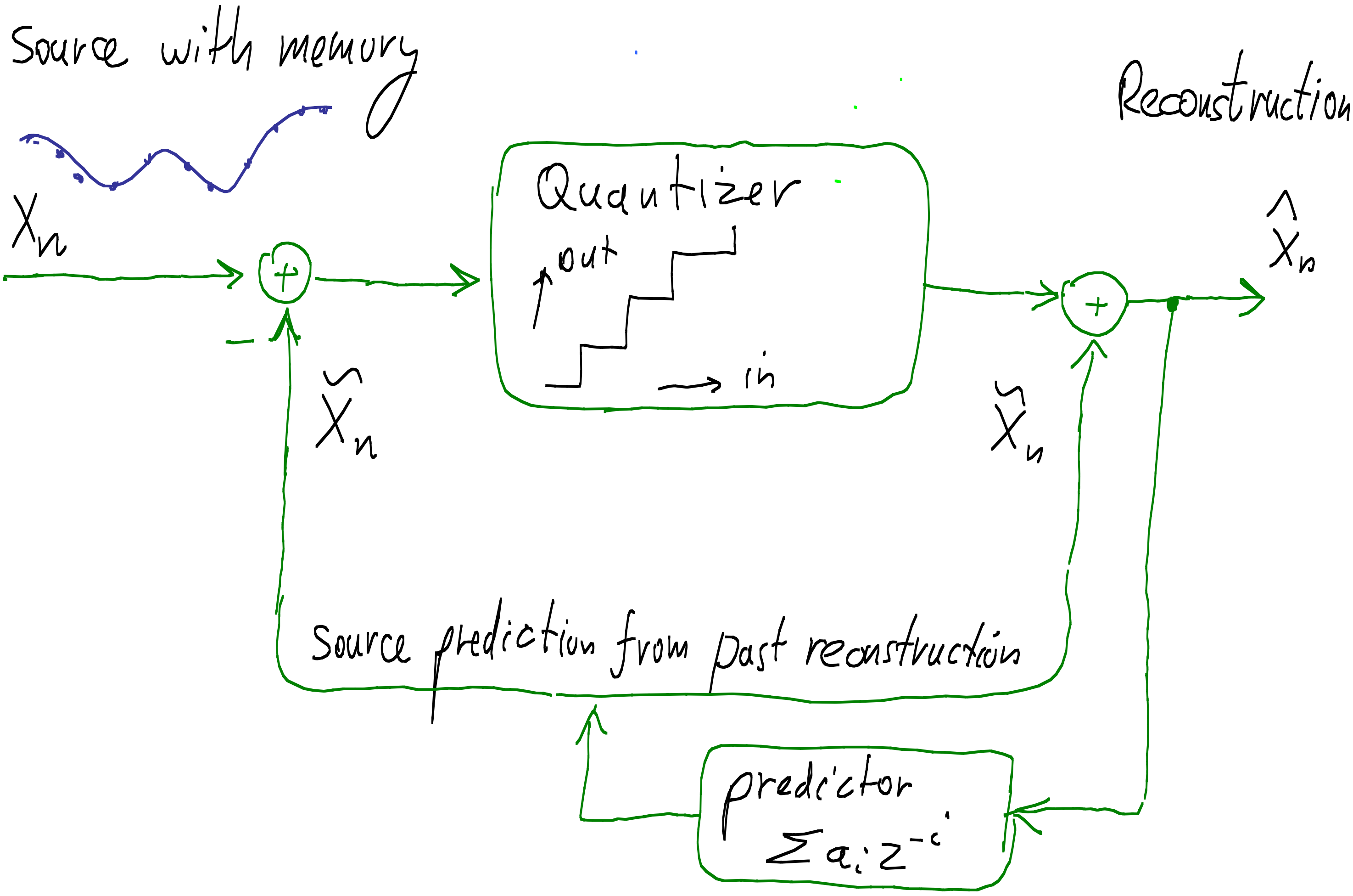
Lattice-WZ Coding: General Resolution

~~$D \ll \sigma_Z^2$~~

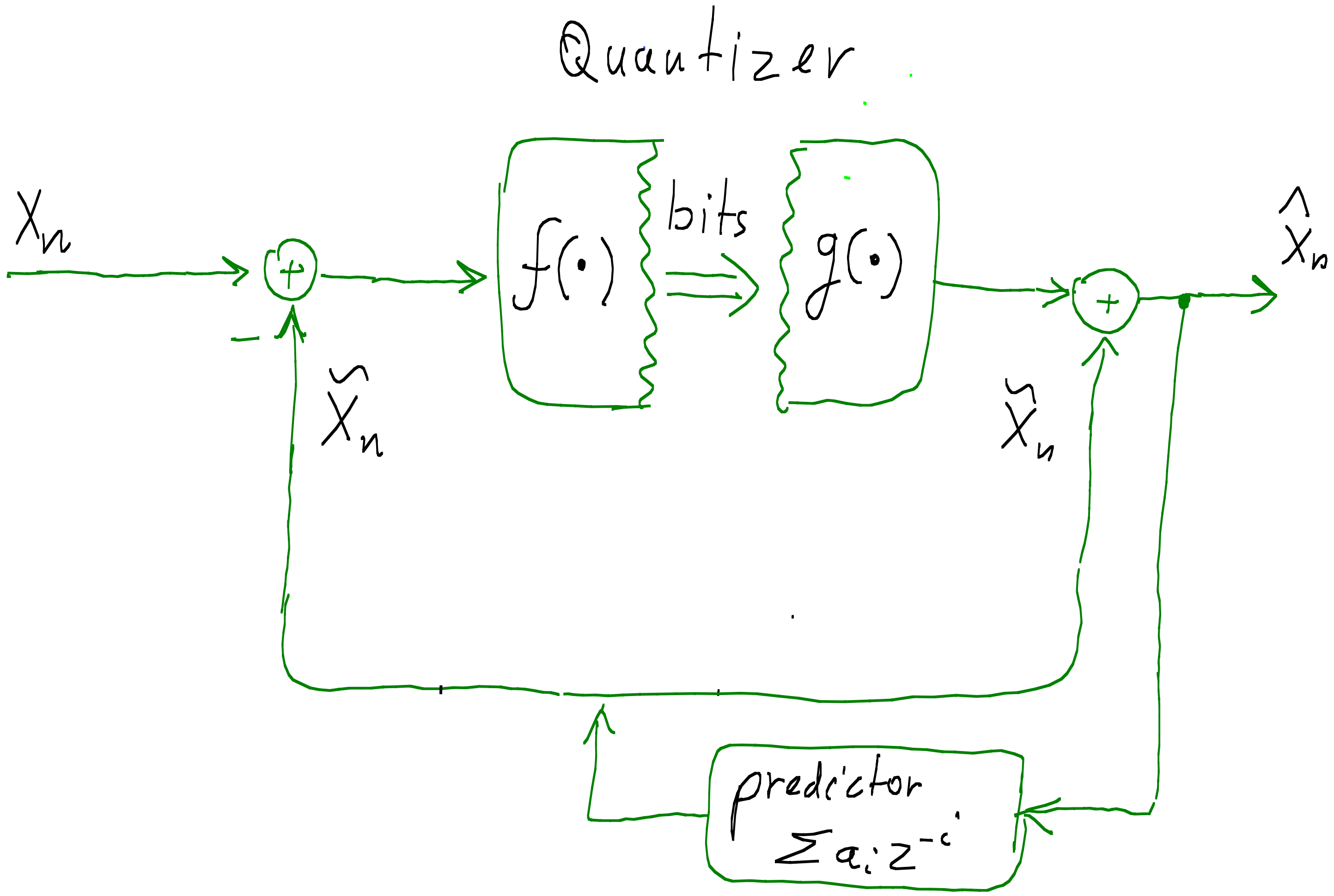


"Forward channel realization" of $R_Z(D)$ (assuming Gaussian dither)

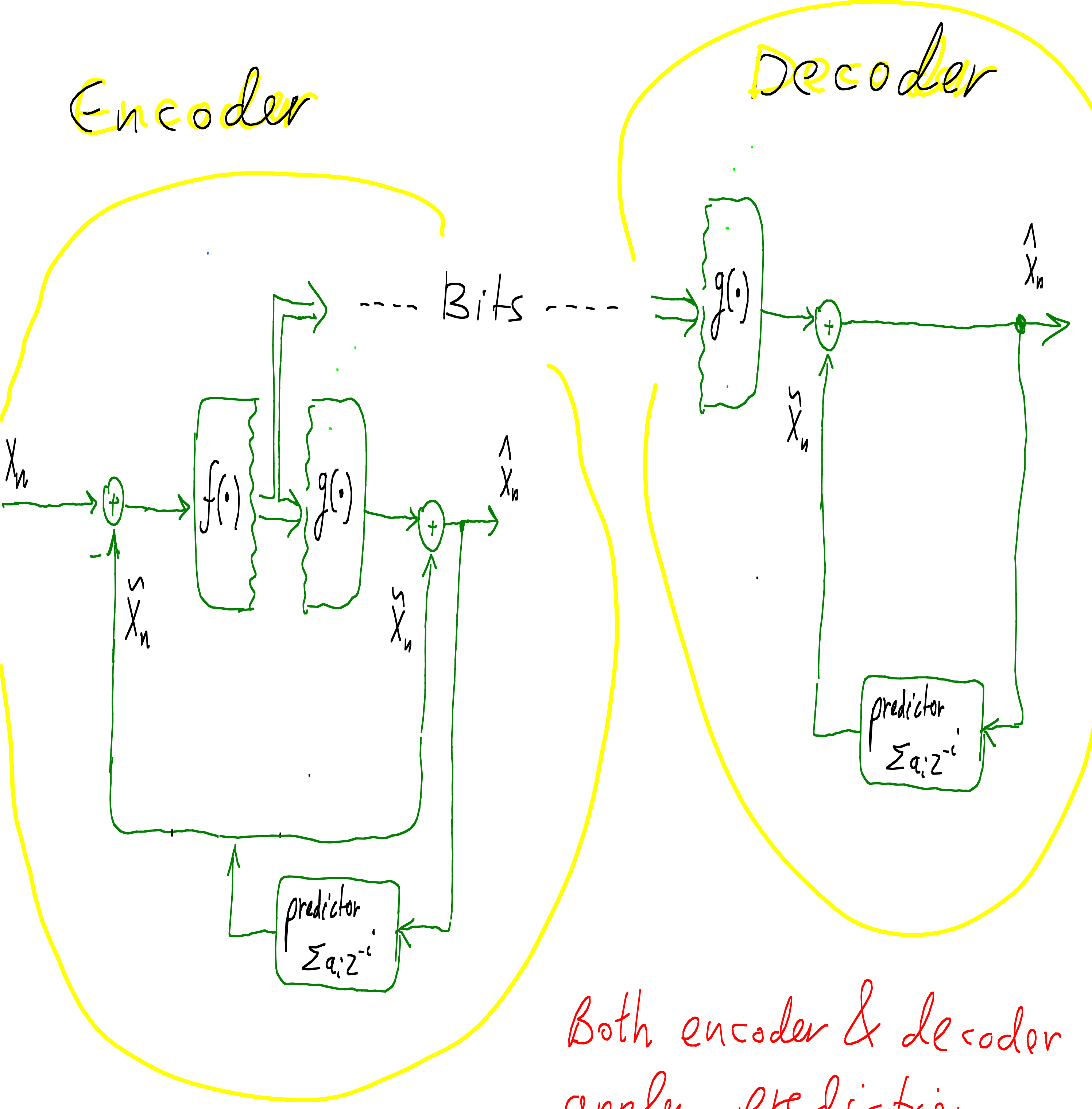
Wyner - Ziv - D.P.C.M.



Wyner-Ziv - D.P.C.M.

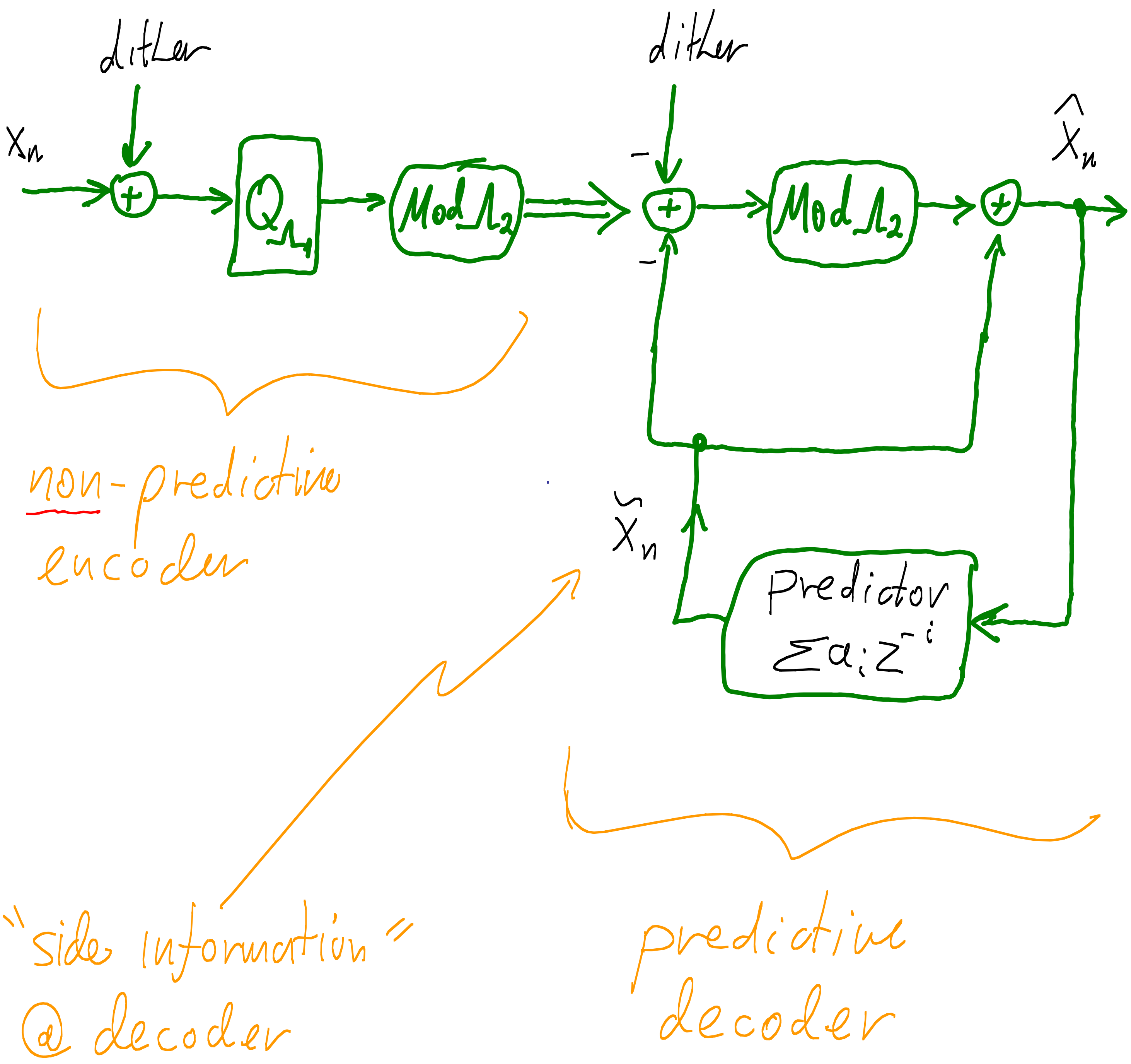


Wyner-Ziv - D.P.C.M.



Both encoder & decoder apply prediction

Wyner-Ziv - D.P.C.M.



Appendix:

Basic Information Theory

Rami Zamir

Sequences, Types ...

sequence

$\underline{x} = x_1, x_2, \ldots, x_n \qquad x_i \in \mathcal{X}$

$\mathcal{X}$ = alphabet

e.g., binary alphabet $\{0,1\}$

type (empirical distribution)

$$P_{\underline{x}}(a) = \frac{\#\, a\text{'s in } x_1 \ldots x_n}{n}, \quad a \in \mathcal{X}$$

Binary Example
$\mathcal{X} = \{0,1\}$
$n = 4$

0 0 1 0

$P_{\underline{x}}(0) = \frac{3}{4} \qquad P_{\underline{x}}(1) = \frac{1}{4}$

type class

$$T_P = \{ \underline{x} = x_1 \ldots x_n : P_{\underline{x}} = P \}$$

$$T_{(\frac{3}{4}, \frac{1}{4})} = \{ (0001), (0010), (0100), (1000) \}$$

size of type class

$$|T_P| = \binom{n}{P(1)\cdot n \quad P(2)\cdot n \quad \dots \quad P(|\mathcal{X}|)\cdot n} \simeq 2^{n \cdot H(P)}$$

stirling approximation

where $H(P) = \text{entropy}.$

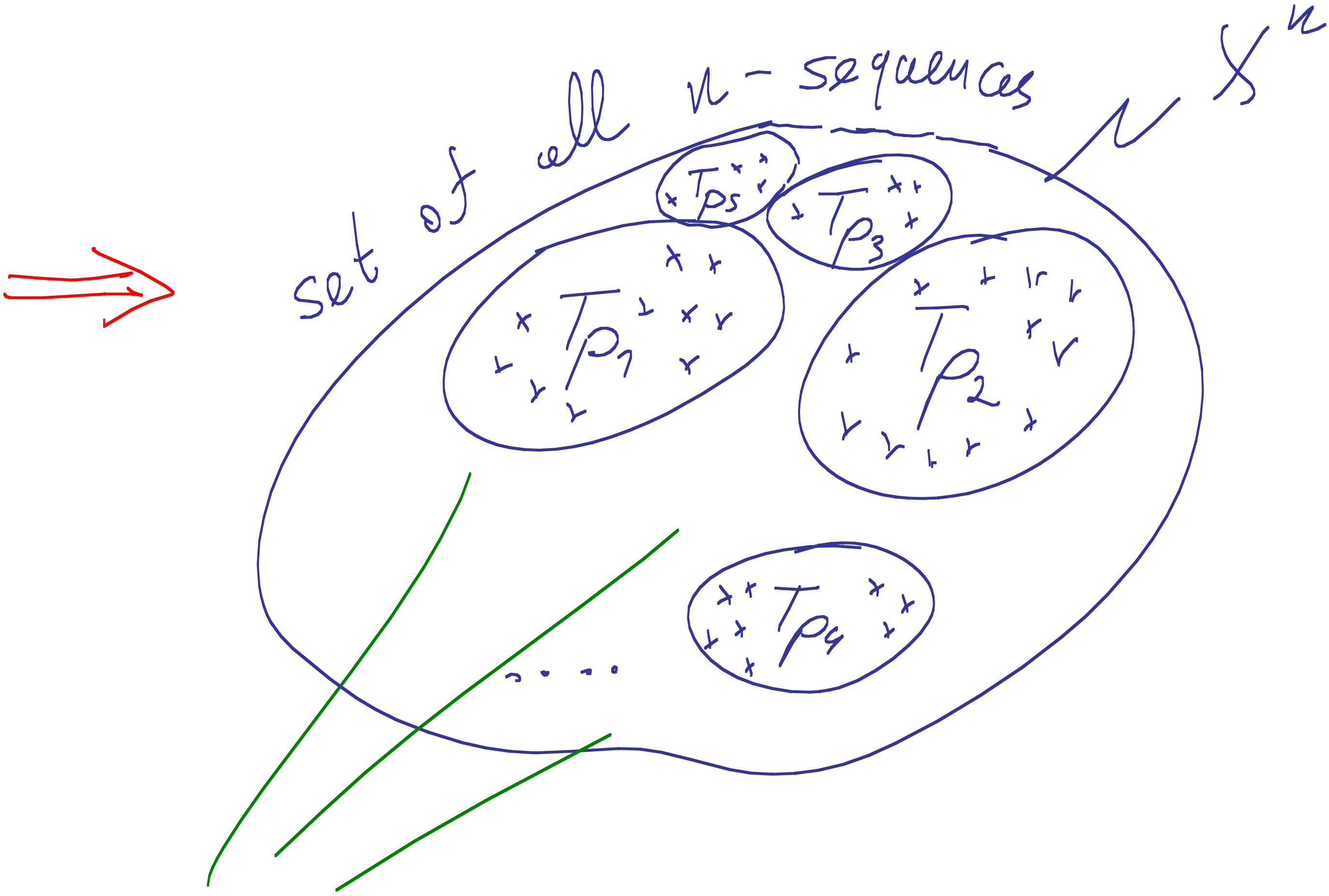
$$= -\sum_{i \in \mathcal{X}} p_i \log p_i$$

all n-types

$$\mathcal{P}_n = \{ P_{\underline{x}}(\cdot) : \underline{x} \in \mathcal{X}^n \}$$

$$\mathcal{P}_4 = \left\{ (0,1), \left(\tfrac{1}{4}, \tfrac{3}{4}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{3}{4}, \tfrac{1}{4}\right), (1,0) \right\}$$

$$|\mathcal{P}_n| \leq (n+1)^{|\mathcal{X}|}$$



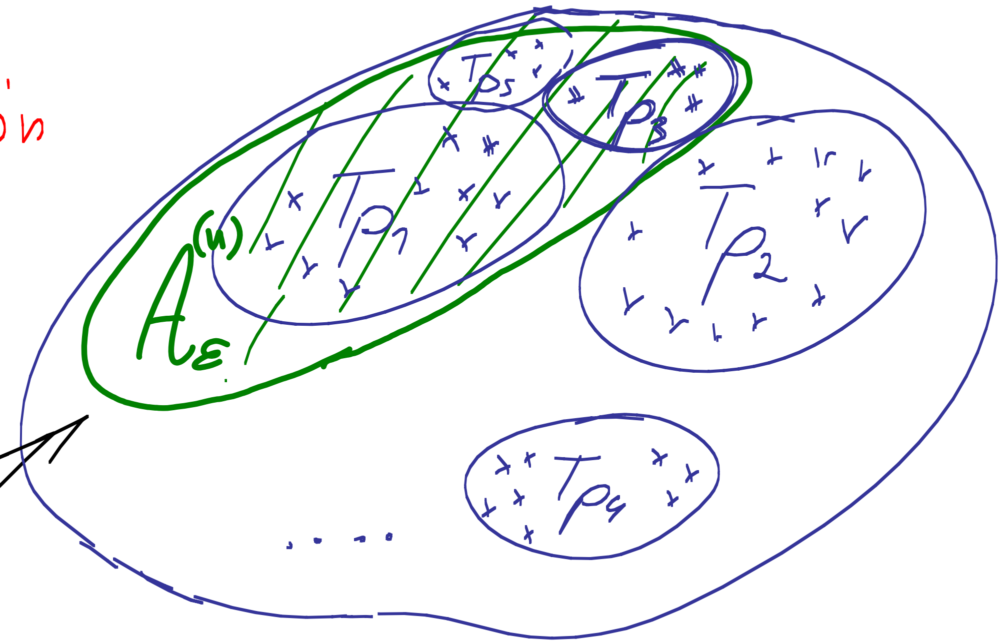
- division into subsets (=type classes)
- # of subsets polynomial in n (small)
- size of subset exponential in n (large!)

Law of Large Numbers

$X_i \sim P$ i.i.d. ("memoryless source")

L.L.N. $\Longrightarrow$ $P_{x_1 \dots x_n}(a) \xrightarrow[n \to \infty]{w.p.\,1} P(a)$

Asymptotic Equi-Partition Property



$$A_\varepsilon^{*(n)} = \bigcup T_{P \pm \varepsilon}$$

$$\Pr\{\underline{X} \in A_\varepsilon^{*(n)}\} \simeq 1 \quad , \quad \underline{X} \sim \mathrm{Unif}(A_\varepsilon^{*(n)})$$

Atypical Sequences...

Prob. of a sequence

$$P(10001100 1\ldots) = p(1)^{\#1} \cdot p(0)^{\#0}$$

$$= p(1)^{n\cdot q(1)} \cdot p(0)^{n\cdot q(0)} = 2^{n[q(1)\log p(1) + q(0)\log p(0)]}$$

$$P(x_1, \ldots, x_n) = 2^{-n\left[H(P_{\underline{x}}) + D(P_{\underline{x}} \| P)\right]}$$

Divergence (K-L)

$$D(Q\|P) \triangleq \sum_{i\in\mathcal{X}} q_i \log \frac{q_i}{p_i}$$

$\Rightarrow$ For i.i.d., sequence' type determines its probability

Atypical Sequences ...

Prob. of a sequence

$$P(10001100 1\ldots) = p(1)^{\#1} \cdot p(0)^{\#0}$$

$$= p(1)^{n \cdot q(1)} \cdot p(0)^{n \cdot q(0)} = 2^{n[q(1)\log p(1) + q(0)\log p(0)]}$$

$$P(x_1, \ldots, x_n) = 2^{-n[H(P_{\underline{x}}) + D(P_{\underline{x}} \| P)]}$$

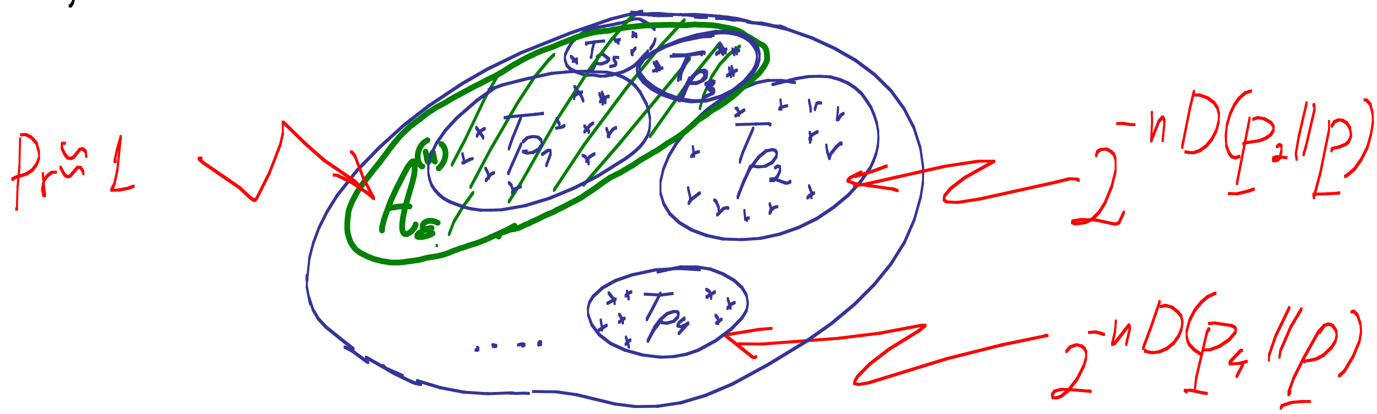
Divergence (K-L)

$$D(Q \| P) \triangleq \sum_{i \in \mathcal{X}} q_i \log \frac{q_i}{p_i}$$

$\Rightarrow$ For i.i.d., sequence' type determines its probability

Prob. of Type-Class

$$\Rightarrow P_r(\underline{X} \in T_Q) = |T_Q| \times 2^{-n[H(Q) + D(Q\|P)]} \doteq 2^{-nD(Q\|P)}$$



Joint Typicality

sequence pair

$\underline{x} = x_1, x_2, \ldots, x_n$

$\underline{y} = y_1, y_2, \ldots, y_n$

joint type

$$P_{\underline{x},\underline{y}}(a,b) = \frac{\#\binom{a}{b} \text{ in } \begin{matrix} x_1 \ldots x_n \\ y_1 \ldots y_n \end{matrix}}{n}, \quad \begin{matrix} a \in \mathcal{X} \\ b \in \mathcal{Y} \end{matrix}$$

$\underline{x} = 0\ 0\ 1\ 0$

$\underline{y} = 0\ 0\ 0\ 0$

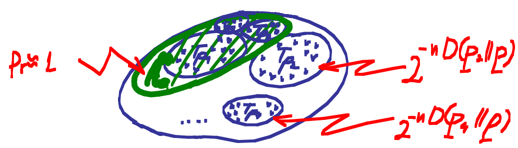
$y \backslash x$	0	1
0	3/4	1/4
1	0	0

All properties above hold. In particular,

If $(X_i, Y_i)_{i=1}^{\infty}$ i.i.d. $\sim P(x,y)$, then

L.L.N. $\Longrightarrow P_{\underline{X},\underline{Y}}(a,b) \xrightarrow[n\to\infty]{w.p.1} P(a,b)$

$\Longrightarrow$ **Joint A.E.P.**



Mutual Information & "Pretending-Correlated"

Prob. of sequence pair

$$\Pr\left\{ X \begin{array}{ccc} 0 & \xrightarrow{\;p\;} & 0 \\ 1 & \xrightarrow{\;p\;} & 1 \end{array} Y \;\middle\|\; \begin{array}{c} \text{Bernulli}(\tfrac{1}{2}) \\ \times \\ \text{Bernulli}(\tfrac{1}{2}) \end{array} \right\} = \frac{|T_{BSC\text{-type}}|}{2^n \times 2^n}$$

$$\doteq 2^{n[1 - p\log p - (1-p)\log(1-p)]}$$

$$\Pr\left\{ \underline{X} \sim P \rightarrow \boxed{W} \rightarrow \underline{Y} \sim Q \;\middle|\; \underline{X} \perp\!\!\!\perp \underline{Y},\ \begin{array}{c} \underline{X} \sim P \\ \underline{Y} \sim Q \end{array} \right\}$$

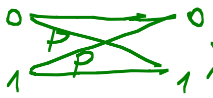
$$= 2^{-nD(P\circ W \| P \times Q)} = 2^{-nI(X;Y)}$$

mutual information

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)\cdot q(y)}$$

Mutual Information & "Pretending-Correlated"

Prob. of sequence pair



$$Pr\left\{ X \;\; Y \,\middle\|\, \begin{matrix}\text{Bernoulli}(\tfrac{1}{2}) \\ \times \\ \text{Bernoulli}(1/2)\end{matrix} \right\} = \frac{|T_{BSC\text{-type}}|}{2^n \times 2^n}$$

$$\doteq 2^{n[1 - p\log p - (1-p)\log(1-p)]}$$

$$Pr\left\{ \underline{x} \sim p \rightarrow \boxed{W} \rightarrow \underline{y} \sim q \;\middle|\; \underline{X} \perp\!\!\!\perp \underline{Y}, \; \begin{matrix} \underline{X} \sim p \\ \underline{Y} \sim q \end{matrix} \right\}$$

$$= 2^{-nD(p \circ W \,\|\, p \times q)} = 2^{-nI(X;Y)}$$

mutual information

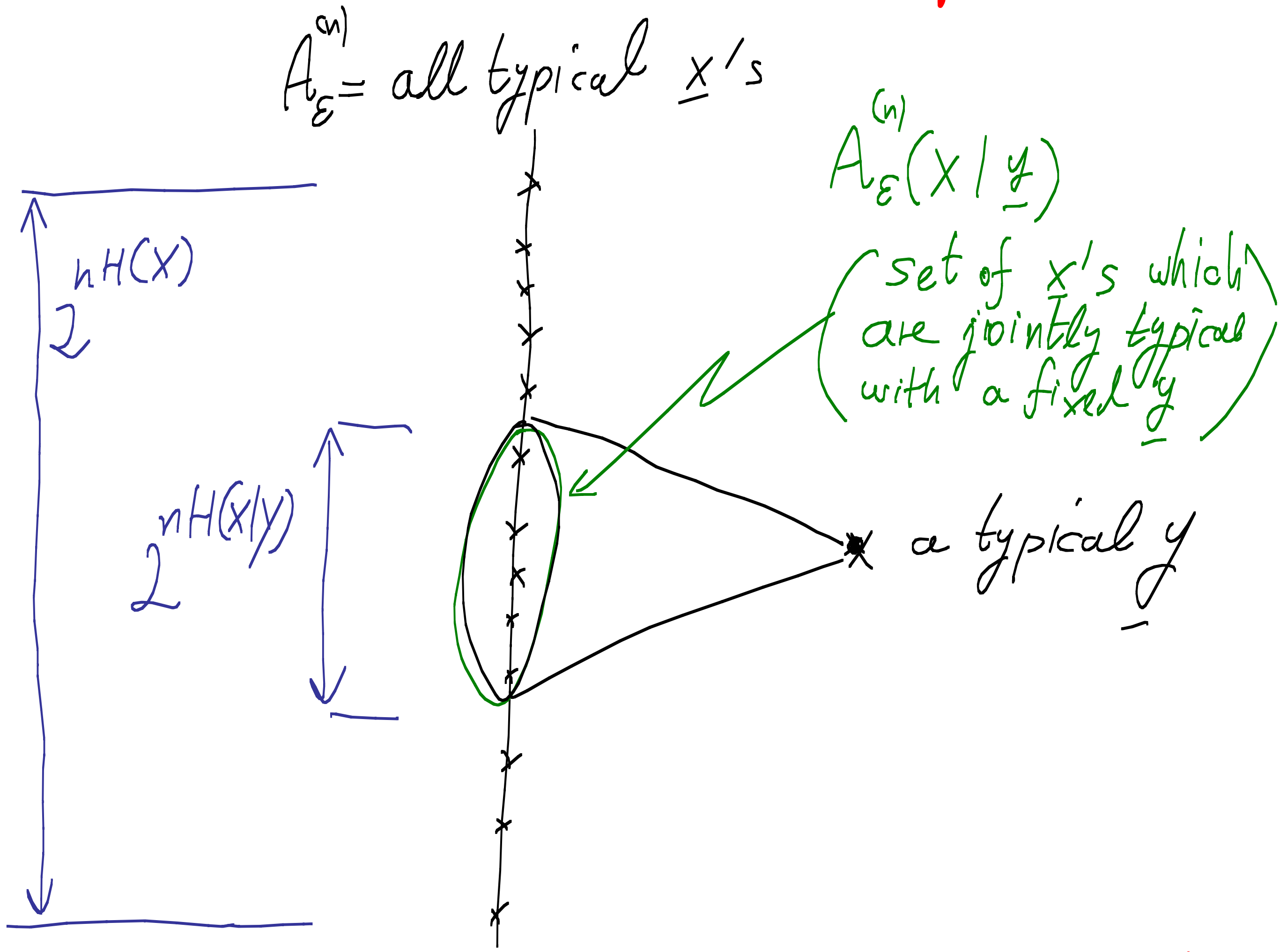
$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x) \cdot q(y)}$$

⇒ Probability that a random $\underline{x}$ is jointly typical with $\underline{y}$ (and vice versa) is $2^{-nI(X;Y)}$.

Exponential Threshold

⇒ $$Pr\begin{pmatrix} \text{one of } 2^{nR} \text{ random } \underline{X}\text{'s} \\ \text{is jointly-typical} \\ \text{a typical } \underline{y} \end{pmatrix} \xrightarrow[n \to \infty]{} \begin{cases} 1, & \text{if } R > I(X;Y) \\ 0, & \text{if } R < I(X;Y) \end{cases}$$

Alternative View: Fan Picture



⇒ Probability of "pretending correlated":

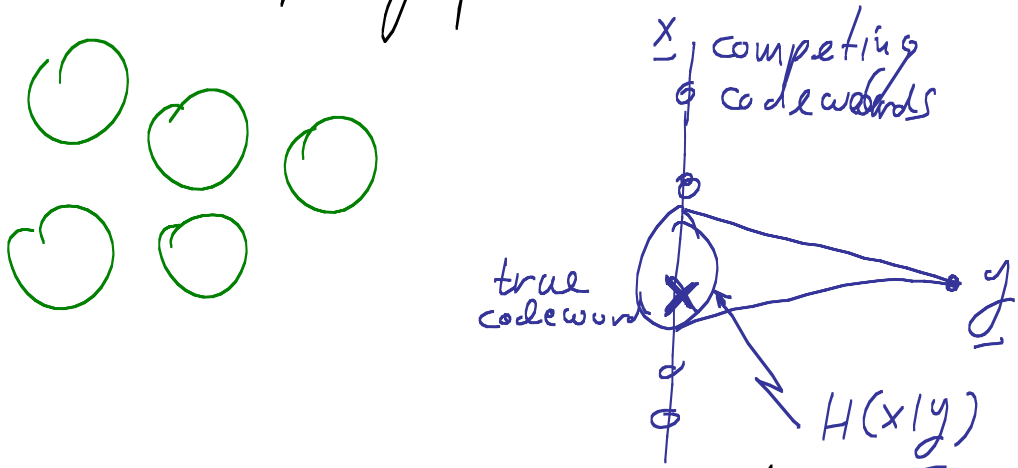
AEP

$I(X;Y) = H(X) - H(X|Y)$

$$\Pr \approx \frac{|A_\varepsilon^{(n)}(X|\underline{y})|}{|A_\varepsilon^{(n)}(X)|} \approx \frac{2^{nH(X|Y)}}{2^{nH(X)}} = 2^{-nI(X;Y)}$$

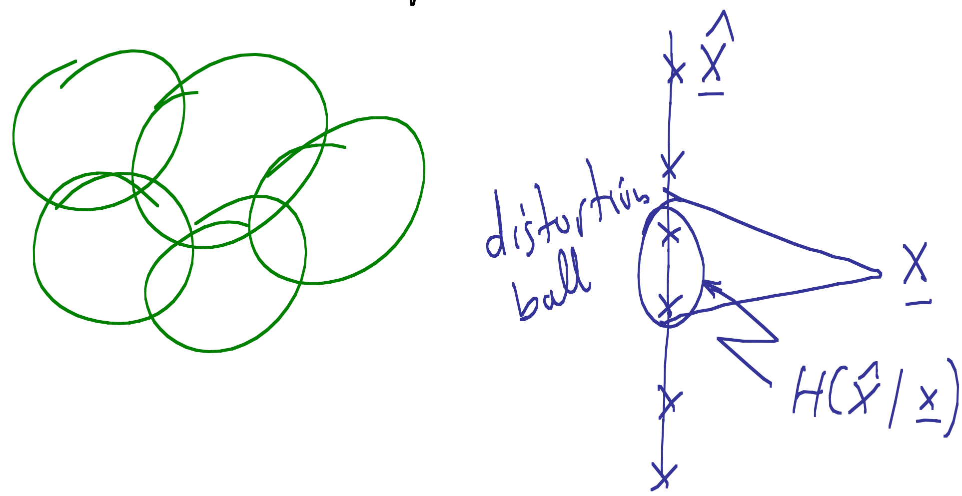
Random Coding for Channels & Sources

Channel capacity problem ≡ "random packing of Balls"



$R < I(X;Y) \Rightarrow$ no pretending $\Rightarrow C = \max_{p(x)} I(X;Y)$

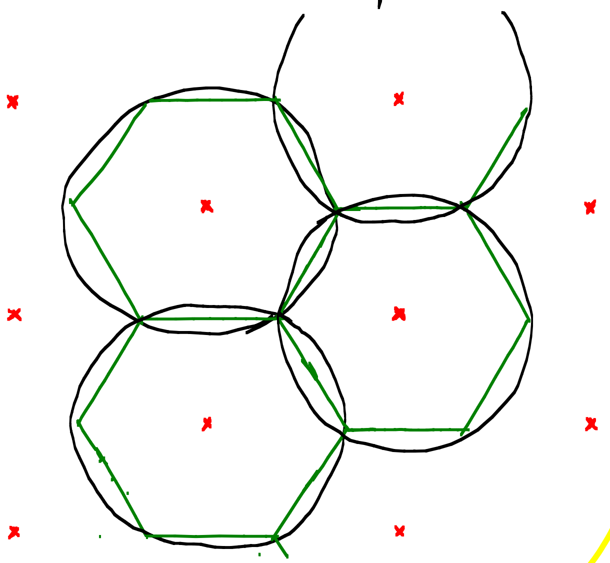
Rate-Distortion Problem ≡ "random covering by Balls"



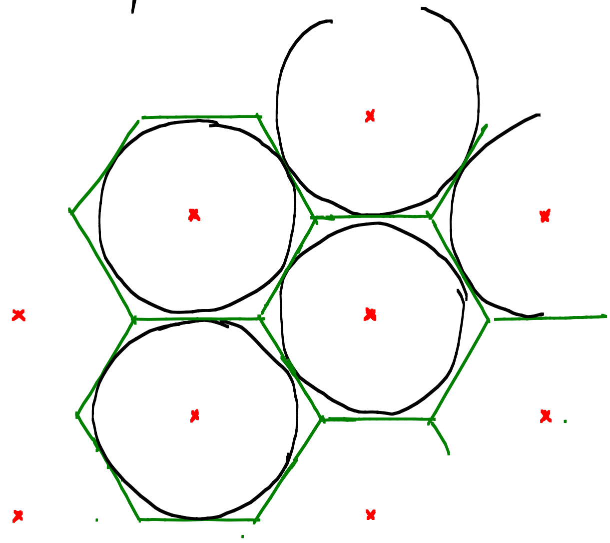
$R > I(X;\hat{X}) \Rightarrow$ guaranteed pretending $\Rightarrow R(D) = \min_{p(\hat{x}|x)} I(X;\hat{X})$

Lattices for Structured Coding

Covering $\mathbb{R}^k$ with (few) spheres



Packing (many) spheres in $\mathbb{R}^k$



Voronoi Codebooks

using

Nested Lattices