Universal Fixed-Length Coding Redundancy

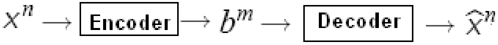
Cheng Chang   Anant Sahai

University of California, Berkeley

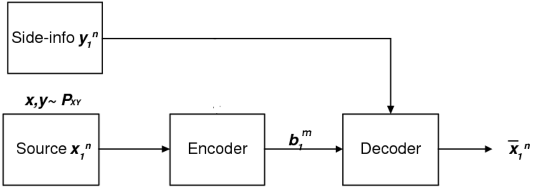
A unified approach to non-asymptotic information theory

The Non-asymptotic problems

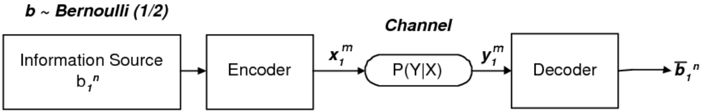
- Lossless source coding's architectural question: given ϵ and n, what's the minimum $m = m(n, \epsilon, p_X)$?



- Source coding with side-info



- Compound channel coding

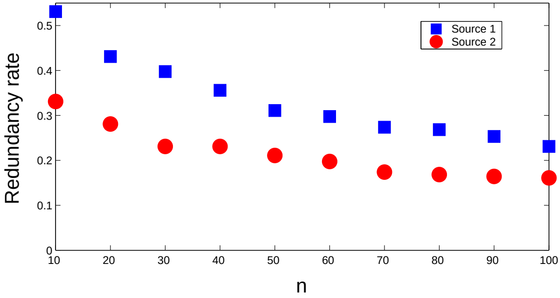


Two sources with the same entropy

- How many bits needed for $\epsilon = 0.01$? Is $nH(x)$ even close?

n	10	20	30	40	50	60	70	80	90	100
$nH(x)$	6	12	18	23	29	35	40	46	52	57
Source 1	11	20	29	37	44	52	59	67	74	80
Source 2	9	17	24	32	39	46	52	59	66	73

- *Redundancy* rate: $\frac{m(\epsilon,n,p_x)}{n} - H(x)$



Related work

- Variable length coding redundancy
  - Minimum description length (Rissanen 1978)
  - Large alphabet (Shamir 2006)
  - Lossless source coding (Kontoyiannis, 97)
  - Lossy source coding (Kontoyiannis, Dembo, 2001)
- Non-asymptotic fixed length coding
  - Non-asymptotic channel coding (Baron 2004)
  - Non-asymptotic Slepian-Wolf coding (Baron 2004)
    - CLT based results: $O(\sqrt{\frac{1}{n}})$ for fixed $\epsilon > 0$
  - Compound channel coding (Ahlswede 1967, Lapidoth 1998)
- This paper: a *unified* approach to the *universal* upper bound on the redundancy rate based on LDP

Our approach

- Upper bound on error probability
  - $\Pr(\widehat{x}^n \neq x^n) \leq 2^{-nE(\frac{m}{n})} \leq \epsilon$
- Error exponent $E(R)$ is strictly increasing with R
  - Sufficient condition: $m = m(n, \epsilon, p_x) \geq nE^{-1}(\frac{\log_2(\epsilon^-1)}{n})$
- Redundancy rate
  - $\mathcal{R}(n, \epsilon, p_x) \sim E^{-1}(\frac{\log_2(\epsilon^{-1})}{n}) - H(x)$
- Lower bound on the error exponents $\rightarrow$ *achievable* redundancy rate
- Universal penalty:

  $\Pr(\widehat{x}^n \neq x^n) \leq 2^{-n(E(\frac{m}{n} - \frac{\log_2(n+1)}{n}|\mathcal{X}|) - \frac{\log_2(n+1)}{n}|\mathcal{X}|)}$

$\epsilon \leq 2^{-n[E_r(R-Univ(n))-Univ(n)]}$ for all n

	R(limit)	Error Exponent	Univ(n)
PtoP	$H(x)$	$\max_{\rho\in[0,1]} \rho R - E_0(\rho)$	$\frac{\lvert\mathcal{X}\rvert \log_2(n)}{n}$
S-W	$H(x\|y)$	$\max_{\rho\in[0,1]} \rho R - E_1(\rho)$	$\frac{\lvert\mathcal{X}\rvert\lvert\mathcal{Y}\rvert \log_2(n)}{n}$
Chan	$\max_x I(x;y)$	$\max_{\rho\in[0,1]} -\rho R + E_2(\rho, Q)$	$\frac{\lvert\mathcal{X}\rvert\lvert\mathcal{Y}\rvert \log_2(n)}{n}$

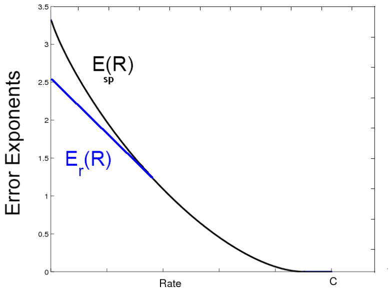
$$E_0(\rho) = (1+\rho)\ln(\sum_{x\in\mathcal{X}} p_x(x)^{\frac{1}{1+\rho}})$$

$$E_1(\rho) = \ln(\sum_y (\sum_x p_{xy}(x,y)^{\frac{1}{1+\rho}})^{1+\rho})$$

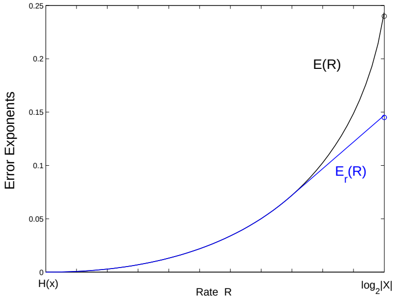
$$E_2(\rho, Q) = -\ln(\sum_y (\sum_x Q(x)P(y|x)^{\frac{1}{1+\rho}})^{1+\rho})$$

What error exponents look like

Channel coding

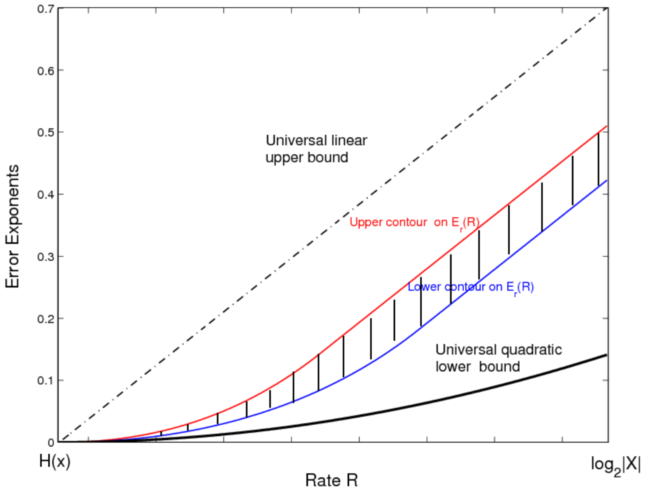


Source coding



Large Deviation result (counting and Chernoff bound)

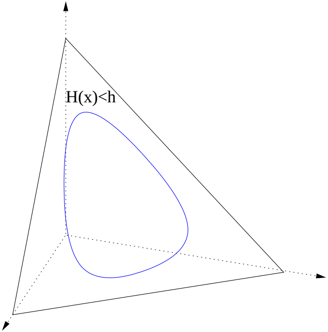
Universally lower bounding error exponents



$|\mathcal{X}| = 3, H(x) = 0.5689$

Quadratic lower bound on source coding exponent

It is a non-convex optimization problem– HARD!



We correct and apply Gallager's technique (Gallager's Exercise 5.23) in bounding the channel coding error exponent.

$$E_r(R) \geq \alpha(R - H(x))^2$$

$$\alpha = \begin{cases} \frac{\log_2 e}{2(\log_2 |\mathcal{X}|)^2} & \text{if } |\mathcal{X}| \geq 3 \\ \frac{1}{2} & \text{if } |\mathcal{X}| = 2 \end{cases}$$

Universality: depends only on the alphabet size $|\mathcal{X}|$ and the entropy rate of the source $H(x)$, but not p_x

Our achievable redundancy rate on source coding

$$\mathcal{R}(n,\epsilon,|\mathcal{X}|) = (\log_2|\mathcal{X}|)\sqrt{\frac{2\log_2(\epsilon^{-1})}{n} + \frac{4(|\mathcal{X}|-1)\log_2(n)}{n}} + \frac{1}{n}$$

- Universality penalty
- Consistent with CLT results, but $\epsilon = \epsilon(n)$ does not have to be constant
- $\mathcal{X}(n)$: large alphabet

- Compound channels

$$\mathcal{R}(n,\epsilon,|\mathcal{X}|,|\mathcal{Y}|) = \sqrt{\frac{\frac{8}{e^2}+4(\ln|\mathcal{Y}|)^2}{\ln 2}\times(\frac{\log_2(\epsilon^{-1})}{n}+\frac{2(|\mathcal{X}||\mathcal{Y}|-1)\log_2(n)}{n})}+\frac{1}{n}$$

- Source coding with decoder side-information (Slepian-Wolf)

$$\mathcal{R}(n,\epsilon,|\mathcal{X}|,|\mathcal{Y}|) = (\log_2|\mathcal{X}|)\sqrt{\frac{2\log_2(\epsilon^{-1})}{n}+\frac{4(|\mathcal{X}||\mathcal{Y}|-1)\log_2(n)}{n}}+\frac{1}{n}$$

Future work

- Tightening the quadratic lower bound
- Lower bounds of other error exponents (lossy source coding, joint source-channel coding, channel coding “focusing” bound)