# The Redundancy of Non-Singular Channel Simulation

Gergely Flamich    Sharang Sriramu

Aaron Wagner

# Channel Simulation: General setup

**Definition and goal**

- Dependent $X, Y \sim P_{X,Y}$
- Independent common randomness: $Z \sim P_Z$, with $X \perp Z$
- Want to encode **one sample** $Y \sim P_{Y|X}$ with finitely many bits:
  - **Encoder:** $f(\cdot, z) : \mathcal{X} \to \{0,1\}^*$ prefix code
  - **Decoder:** $g(\cdot, z) : \{0,1\}^* \to \mathcal{Y}$

$$Z \sim P_Z \quad \to \quad g(f(x, Z), Z) \sim P_{Y|X=x}$$

- **Rate:** $R = \mathbb{E}\left[|f_Z(X)|\right]$
- Want to characterise: $R^* = \inf_{g,f} R$

**Applications**

- $Y = X + \epsilon$, where $\epsilon \sim \mathcal{N}(0,1)$: data compression with diffusion
- $Y = X + \epsilon$, where $\epsilon \sim \mathcal{L}(0,1)$: differential privacy

# Channel Simulation: Facts

**One-shot**

Li and El Gamal [2018]:

$$\mathbb{I}[X : Y] \leq R^* \leq \mathbb{I}[X : Y] + \log(\mathbb{I}[X : Y] + 1) + 5$$

**Asymptotic**

For $P_{X,Y}$, channel $P_{Y|X}$ **singular** if $dP_{X|Y}/dP_X \propto 1$, otherwise **non-singular**.

Sriramu and Wagner [2024]: Let $X^n, Y^n \sim P_{X,Y}^{\times n}$, normalised rate $R_n^* = R^*/n$. Then:

$$\begin{cases} R_n^* \sim \mathbb{I}[X : Y] & : P_{Y|X} \text{ singular} \\ R_n^* \lesssim \mathbb{I}[X : Y] + \frac{1}{2}\frac{\log n}{n} & : P_{Y|X} \text{ non-singular} \end{cases}$$

# Our results

**One-shot**

Channel simulation divergence [Goc and Flamich, 2024]: $D_{CS}[Q\|P]$. For Polish $X, Y \sim P_{X,Y}$

$$\mathbb{I}[X:Y] \leq \mathbb{E}_{Y\sim P_Y}\left[D_{CS}[P_{X|Y}\|P_X]\right] \leq R^*.$$

**Asymptotic**

For non-singular $P_{Y|X}$:

$$\mathbb{I}[X:Y] + \frac{1}{2}\frac{\log n}{n} \lesssim R^*_n$$

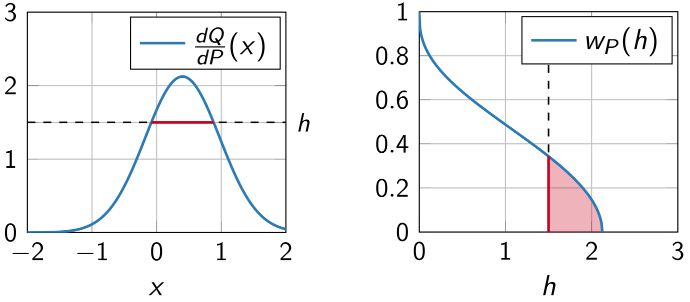
thus

$$\begin{cases} R^*_n \sim \mathbb{I}[X:Y] & : P_{Y|X} \text{ singular} \\ R^*_n \sim \mathbb{I}[X:Y] + \frac{1}{2}\frac{\log n}{n} & : P_{Y|X} \text{ non-singular} \end{cases}$$

# Part I: width function

Probability measures $Q \ll P$. Width function:

$$w_P(h) = \mathbb{P}_{Z\sim P}\left[\frac{dQ}{dP}(Z) \geq h\right]$$




$w_P$ is a PDF! Let $H$ be a RV with PDF $w_P$. Then:

$$D_{CS}[Q\|P] = h[H] = -\int_0^\infty w_P(h)\log w_P(h)\, dh$$

# Part I: Result & Proof Sketch

For Polish $X, Y \sim P_{X,Y}$, $Z \sim P_Z$, $Z \perp X$:

$$\mathbb{E}_{Y \sim P_Y}\left[D_{CS}[P_{X|Y} \| P_X]\right] \leq R^*.$$

Proof when $Y$ discrete (Prop. 1 of Li and El Gamal [2018]):

$$\begin{aligned} R^* \approx \mathbb{H}[Y \mid Z] &= -\mathbb{E}_{Y,Z \sim P_{Y,Z}}[\log p(Y \mid Z)] \\ &= -\sum_{y \in \mathcal{Y}} \mathbb{E}_{Z \sim P_Z}[p(y \mid Z) \log p(y \mid Z)]] \\ &\geq \mathbb{E}_{Y \sim P_Y}\left[D_{CS}[P_{X|Y} \| P_X]\right] \end{aligned}$$

# Part I: Fixing the troublesome step

Let $\phi(z) = P_{Y|z}$ and set $\mathcal{P} = \phi \sharp P_Z$ (Palm kernel)

$$\begin{aligned}
\mathbb{H}[Y \mid Z] &= -\mathbb{E}_{Z\sim P_Z}\left[\mathbb{E}_{Y\sim P_{Y|Z}}[\log p(Y \mid Z)]\right] \\
&= -\int_{\mathfrak{P}_{\mathcal{Y}}} \int_{\mathcal{Y}} \log \pi(y)\, d\pi(y)\, d\mathcal{P}(\pi) \\
&= -\int_{\mathcal{Y}} \int_{\mathfrak{P}_{\mathcal{Y}}} \log \pi(y)\, d\mathcal{P}_y(\pi)\, dP_Y(y)
\end{aligned}$$

$\mathcal{P}_y$ – local Palm kernel

$$\operatorname{supp} \mathcal{P}_y = \{\pi \in \mathfrak{P}_{\mathcal{Y}} \mid y \in \operatorname{supp} \pi\}$$

Adapt stochastic dominance argument for $\int_{\mathfrak{P}_{\mathcal{Y}}} \log \pi(y)\, d\mathcal{P}_y(\pi)$.

# Part II: Asymptotic result & Proof sketch

$P_{X,Y}$ with $P_{Y|X}$ non-singular. Then, for any sequence $Z_n$ of common randomness:

$$\lim_{n\to\infty} \frac{nR_n^* - \mathbb{I}[X^n : Y^n]}{\log n} \geq \frac{1}{2}$$

Proof idea:

- From part I: $nR_n^* \geq \mathbb{E}_{Y\sim P_Y}\left[D_{CS}[P_{X|Y} \| P_X]\right]$
- Show

$$\lim_{n\to\infty} \frac{\mathbb{E}_{Y\sim P_Y}\left[D_{CS}[P_{X|Y} \| P_X]\right] - \mathbb{I}[X^n : Y^n]}{\log n} = \frac{1}{2}$$

- Observe

$$\mathbb{E}_{Y^n}[D_{CS}[P_{X^n|Y^n} \| P_X^{\times n}]] - \mathbb{I}[X^n : Y^n] = \mathbb{E}_{Y^n}\left[h[\log H_{Y^n}]\right] + \log e$$

# Part II: Proof sketch: CLT

Prove a CLT for $\log H_{Y^n}$:

1. When $P_{Y|X}$ non-singular: $\mathrm{Var}[\log H_{Y^n}] = \Theta(n)$:

$$h[\log H_{Y^n}] = \frac{\log n}{2} + h\left[\frac{1}{\sqrt{n}}(\log H_{Y^n} - \mathbb{E}\left[\log H_{Y^n}\right])\right]$$

2. For large $n$

$$\log H_{Y^n} \approx \log \frac{dP_{X^n|Y^n}}{dP_{X^n}}(X^n \mid Y^n) = \sum_{k=1}^{n} \log \frac{dP_{X|Y}}{dP_X}(X_k \mid Y_k)$$

3. Show Lindeberg-Feller condition for $\log \frac{dP_{X|Y}}{dP_X}(X_k \mid Y_k)$, hence

$$\frac{1}{\sqrt{n}}(\log H_{Y^n} - \mathbb{E}\left[\log H_{Y^n}\right]) \to \mathcal{N}(0, \sigma^2)$$

# Our results

## One-shot

Channel simulation divergence [Goc and Flamich, 2024]: $D_{CS}[Q \| P]$. For Polish $X, Y \sim P_{X,Y}$

$$\mathbb{I}[X : Y] \leq \mathbb{E}_{Y \sim P_Y}\left[D_{CS}[P_{X|Y} \| P_X]\right] \leq R^*.$$

## Asymptotic

For non-singular $P_{Y|X}$:

$$R_n^* \gtrsim \mathbb{I}[X : Y] + \frac{1}{2}\frac{\log n}{n}$$

thus

$$\begin{cases} R_n^* \sim \mathbb{I}[X : Y] & : P_{Y|X} \text{ singular} \\ R_n^* \sim \mathbb{I}[X : Y] + \frac{1}{2}\frac{\log n}{n} & : P_{Y|X} \text{ non-singular} \end{cases}$$

# Part III: Operational proof of the asymptotic result

Let $(f, g, Z)$ be a valid scheme simulating the channel with rate $R$

$$
\begin{aligned}
nR &\geq \mathbb{H}[f(X^n, Z)|Z] \\
&\geq \mathbb{I}[X^n : f(X^n, Z)|Z] \\
&\geq \mathbb{I}[X^n : Y^n|Z] \\
&= \mathbb{E}\left[\log \frac{dP_{X^n|Y^n,Z}}{dP_{X^n}}\right]
\end{aligned}
$$

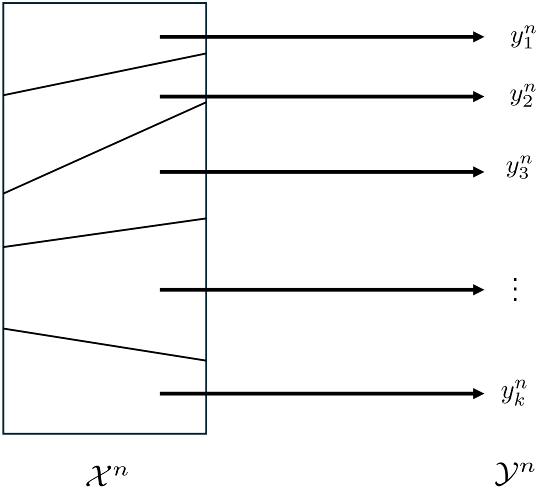
# Part III: Operational proof of the asymptotic result

Let $(f, g, Z)$ be a valid scheme simulating the channel with rate $R$

$$\begin{aligned} nR &\geq \mathbb{H}[f(X^n, Z) | Z] \\ &\geq \mathbb{I}[X^n : f(X^n, Z) | Z] \\ &\geq \mathbb{I}[X^n : Y^n | Z] \\ &= \mathbb{E}\left[\log \frac{dP_{X^n | Y^n, Z}}{dP_{X^n}}\right] \end{aligned}$$

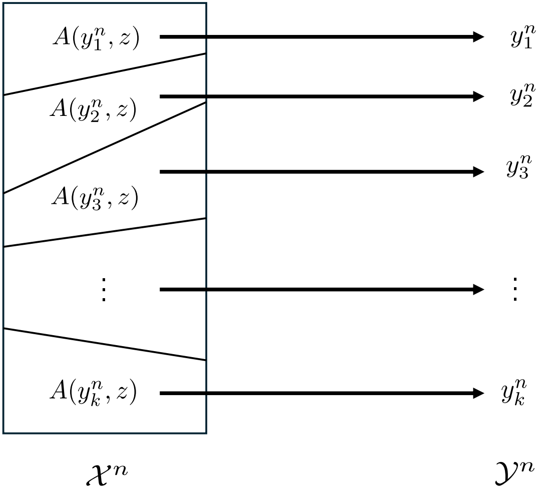
# Deterministic mapping induced by the CR

$$g(f(\cdot, z), z) : \mathcal{X}^n \mapsto \mathcal{Y}^n$$



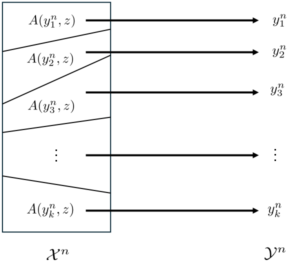
# Deterministic mapping induced by the CR

$$A(y^n, z) = \{x^n \in \mathcal{X}^n : g(f(x^n, z), z) = y^n\}$$



# Simplified density ratio

$$
\begin{aligned}
nR &\geq \mathbb{H}[f(X^n, Z) \mid Z] \\
&\geq \mathbb{I}[X^n : f(X^n, Z) \mid Z] \\
&\geq \mathbb{I}[X^n : Y^n \mid Z] \\
&= \mathbb{E}\left[\log \frac{dP_{X^n|Y^n,Z}}{dP_{X^n}}\right]
\end{aligned}
$$

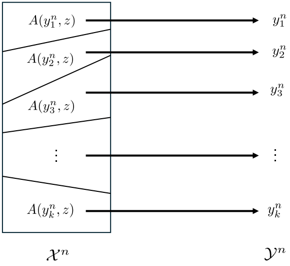



$$
\frac{dP_{X^n|Y^n,Z}}{dP_{X^n}}(x^n|y^n, z) = \frac{\mathbf{1}\left[x^n \in A(y^n, z)\right]}{P_{X^n}(A(y^n, z))}
$$

$$
\therefore nR \geq -\mathbb{E}_{Y^n,Z}\left[\log P_{X^n}(A(Y^n, Z))\right]
$$

# Simplified density ratio

$$
\begin{aligned}
nR &\geq \mathbb{H}[f(X^n, Z) \mid Z] \\
&\geq \mathbb{I}[X^n : f(X^n, Z) \mid Z] \\
&\geq \mathbb{I}[X^n : Y^n \mid Z] \\
&= \mathbb{E}\left[\log \frac{dP_{X^n|Y^n,Z}}{dP_{X^n}}\right]
\end{aligned}
$$




$$\frac{dP_{X^n|Y^n,Z}}{dP_{X^n}}(x^n|y^n, z) = \frac{\mathbf{1}\left[x^n \in A(y^n, z)\right]}{P_{X^n}(A(y^n, z))}$$

$$\therefore\ nR \geq -\mathbb{E}_{Y^n,Z}\left[\log P_{X^n}(A(Y^n, Z))\right]$$

Coincidence probability

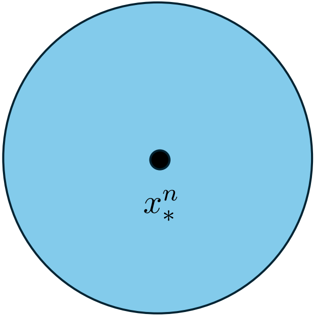
# Bounding the coincidence probability

Replace $A(y^n, z)$ with a more structured set

# Likelihood balls

We define likelihood balls with radius $\alpha$ as

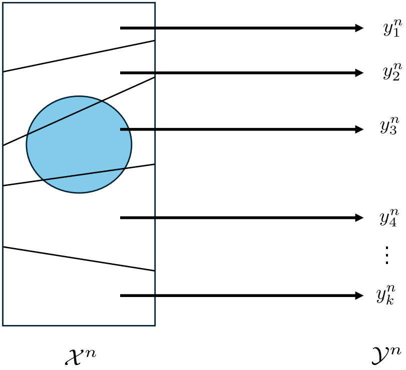
$$B_\alpha = \left\{ x^n \in \mathcal{X}^n : \frac{1}{n} \sum_{i=1}^{n} \log \frac{dP_{X|Y}}{dP_X}(x_i|y_i) \geq \alpha \right\}$$




$x_*^n$ maximizes $\frac{dP_{X^n|Y^n}}{dP_{X^n}}(\cdot|y^n)$

# Mean likelihood

$$\mathrm{LLR}(S) = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \log \frac{dP_{X|Y}}{dP_X}(X_i|y_i) \,\middle|\, X^n \in S\right]$$




Choose the radius $\imath_{y^n,z}$ s.t.

$$\mathrm{LLR}(B_{\imath_{y^n,z}}(y^n)) \approx \mathrm{LLR}(A(y^n,z))$$

Then,

$$P_{X^n}(B_{\imath_{y^n,z}}(y^n)) \geq P_{X^n}(A(y^n,z))$$

Note that

$$\imath_{y^n,z} \approx \mathrm{LLR}(B_{\imath_{y^n,z}}(y^n))$$

# Completing the proof

$$\begin{aligned} nR &\geq -\mathbb{E}_{Y^n,Z}\left[\log P_{X^n}(A(Y^n,Z))\right] \\ &\geq -\mathbb{E}_{Y^n,Z}\left[\log P_{X^n}(B_{\imath_{Y^n,Z}})\right] \\ &= -\mathbb{E}_{Y^n,Z}\left[\log P_{X^n}\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{dP_{X|Y}}{dP_X}(X_i|Y_i) \geq \imath_{Y^n,Z}\right)\right] \end{aligned}$$

# Warmup: Coarse large deviations result

**Theorem**[Dembo and Zeitouni, 2011]. Given independent random variables $Z_1, Z_2, \cdots, Z_n$ and $\gamma > \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Z_i]$, for sufficiently large $n$,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i > \gamma\right) \leq \exp(-n\Lambda_n^*(\gamma)),$$

$\Lambda_n^*(\gamma)$: Large deviations rate function

# Warmup: First order result

In our setting,

$$P_{X^n}\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{dP_{X|Y}}{dP_X}(X|y_i) > \imath_{y^n,z}\right) \leq \exp(-n\Lambda_n^*(\imath_{y^n,z})),$$

where

$$\begin{aligned}\Lambda_n^*(\imath_{y^n,z}) &\approx \imath_{y^n,z}\\ &\approx \mathrm{LLR}(B_{\imath_{y^n,z}}(y^n))\\ &\approx \mathrm{LLR}(A(y^n,z)).\end{aligned}$$

Then,

$$\begin{aligned}nR &\geq -\mathbb{E}_{Y^n,Z}\left[\log P_{X^n}\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{dP_{X|Y}}{dP_X}(X_i|Y_i) \geq \imath_{Y^n,Z}\right)\right]\\ &\geq -\mathbb{E}_{Y^n,Z}\left[\log\exp\left(-n\mathrm{LLR}(A(Y^n,Z))\right)\right]\\ &= I(X;Y)\end{aligned}$$

# Refined large deviations result

**Theorem**[e.g. Altuğ and Wagner [2021]]. Given independent random variables $Z_1, Z_2, \cdots, Z_n$ and $\gamma > \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[Z_i]$,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i > \gamma\right) \leq \frac{(\cdots)}{\sqrt{\sum_{i=1}^{n} \mathrm{Var}(\tilde{Z}_i)}} \exp(-n\Lambda_n^*(\gamma)),$$

$\Lambda_n^*(\gamma)$: Large deviations rate function,

$\tilde{Z}_k$: Exponentially tilted version of $Z_k$

# Refined large deviations result

In our setting,

$$P_{X^n}\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{dP_{X|Y}}{dP_X}(X|y_i) > \imath_{y^n,z}\right) \leq \frac{(\cdots)\exp(-n\Lambda_n^*(\imath_{y^n,z}))}{\sqrt{\sum_{i=1}^{n}\mathrm{Var}_{\tilde{X}_i}\left(\log\frac{dP_{X|Y}}{dP_X}(\tilde{X}_i|y_i)\right)}},$$

$$P_{\tilde{X}} \approx P_{X|Y}(\cdot|y_k)$$

For non-singular channels, $\mathrm{Var}\left(\log\frac{dP_{X|Y}}{dP_X}|Y\right) > 0$

Hence, $\sum_{i=1}^{n}\mathrm{Var}_{\tilde{X}_i}\left(\log\frac{dP_{X|Y}}{dP_X}(\tilde{X}_i|y_i)\right) = \Theta(n)$

# Refined large deviations result

In our setting,

$$P_{X^n}\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{dP_{X|Y}}{dP_X}(X|y_i) > \imath_{y^n,z}\right) \leq \frac{(\cdots)\exp(-n\Lambda_n^*(\imath_{y^n,z}))}{\sqrt{n}},$$

$P_{\tilde{X}} \approx P_{X|Y}(\cdot|y_k)$

For non-singular channels, $\mathrm{Var}\left(\log\frac{dP_{X|Y}}{dP_X}|Y\right) > 0$

Hence, $\sum_{i=1}^{n}\mathrm{Var}_{\tilde{X}_i}\left(\log\frac{dP_{X|Y}}{dP_X}(\tilde{X}_i|y_i)\right) = \Theta(n)$

# Contributions

**One-shot**

For Polish $X, Y \sim P_{X,Y}$

$$\mathbb{I}[X : Y] \leq \mathbb{E}_{Y \sim P_Y}\left[D_{CS}[P_{X|Y} \| P_X]\right] \leq R^*.$$

**Asymptotic**

For non-singular $P_{Y|X}$:

$$\mathbb{I}[X : Y] + \frac{1}{2}\frac{\log n}{n} \lesssim R_n^*$$

thus

$$\begin{cases} R_n^* \sim \mathbb{I}[X : Y] & : P_{Y|X} \text{ singular} \\ R_n^* \sim \mathbb{I}[X : Y] + \frac{1}{2}\frac{\log n}{n} & : P_{Y|X} \text{ non-singular} \end{cases}$$

# References I

Y. Altuğ and A. B. Wagner. On exact asymptotics of the error probability in channel coding: symmetric channels. *IEEE Transactions on Information Theory*, 67(2):844–868, 2021.

D. Goc and G. Flamich. On channel simulation with causal rejection samplers. In *2024 IEEE International Symposium on Information Theory (ISIT)*, 2024. doi: 10.1109/ISIT57864.2024.10619339.

C. T. Li and A. El Gamal. Strong functional representation lemma and applications to coding theorems. *IEEE Transactions on Information Theory*, 64(11):6967–6978, 2018.

S. M. Sriramu and A. B. Wagner. Optimal redundancy in exact channel synthesis. In *2024 IEEE International Symposium on Information Theory (ISIT)*, pages 1913–1918, 2024. doi: 10.1109/ISIT57864.2024.10619703.