Asymptotic Capacity of a Random Channel

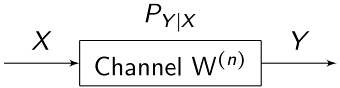
Tobias Sutter, David Sutter, and John Lygeros

Automatic Control Laboratory & Institute for Theoretical Physics

Allerton Conference, 2014

ETH zürich

Channel Capacity



- Discrete memoryless channel (DMC) $W^{(n)} : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X} = \{1, 2, \cdots, n\}$, $\mathcal{Y} = \{1, 2, \cdots, n\}$
- Capacity of $W^{(n)}$ [Shannon'48]

$$C(W^{(n)}) = \begin{cases} \max_{p} & I\big(p, W^{(n)}\big) \\ & p \in \Delta_n \end{cases}$$

- What if the channel matrix $W^{(n)}$ is chosen at random ?
  - e.g. Dirichlet distributed row vectors of $W^{(n)}$

Outline

- Random channel construction
- Capacity of a random channel
- Proof (sketch)
- Numerical example
- Rate of convergence
- Summary & Outlook

Random channel construction

- Consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$
- Construction
  1. Let $(V_{x,y})_{x,y\in[n]}$ be i.i.d. nonnegative random variables on Ω such that $\mathbb{E}\big[(V_{x,y}\log V_{x,y})^2\big] < \infty$
  2. Channel matrix $\mathrm{W}^{(n)}$ with components $\mathrm{W}^{(n)}_{x,y} := \frac{V_{x,y}}{\sum_{y\in[n]} V_{x,y}}$
- Special case: i.i.d. Dirichlet distributed row vectors of $\mathrm{W}^{(n)}$

Lemma (Measurability)

The capacity of such a channel $C(\mathrm{W}^{(n)}) := \max_{p\in\Delta_n} I\big(p, \mathrm{W}^{(n)}\big)$ and the optimal input distributions $p^\star(\mathrm{W}^{(n)}) \in \arg\max_{p\in\Delta_n} I\big(p, \mathrm{W}^{(n)}\big)$ are random variables.

- Questions: $C(\mathrm{W}^{(n)}) \underset{n\to\infty}{\longrightarrow} ?, \qquad p^\star(\mathrm{W}^{(n)}) \underset{n\to\infty}{\longrightarrow} ?$

Capacity of a random channel

- Define $\mu_1 := \mathbb{E}[V_{x,y}]$ and $\mu_2 := \mathbb{E}[V_{x,y} \log V_{x,y}]$

Theorem (Asymptotic capacity)

For $n \to \infty$ the capacity $C(\mathrm{W}^{(n)})$ of the DMC defined above converges to $\frac{\mu_2}{\mu_1} - \log \mu_1$ almost surely and in L^2.

Capacity of a random channel

- Define $\mu_1 := \mathbb{E}[V_{x,y}]$ and $\mu_2 := \mathbb{E}[V_{x,y} \log V_{x,y}]$

Theorem (Asymptotic capacity)

For $n \to \infty$ the capacity $C(\mathrm{W}^{(n)})$ of the DMC defined above converges to $\frac{\mu_2}{\mu_1} - \log \mu_1$ almost surely and in L^2.

The asymptotic capacity

- is nonnegative (direct consequence of Jensen's inequality)
- can be zero (e.g., consider $V_{x,y}$ concentrated at a point)
- can be arbitrarily large (e.g., consider random variables $V_{x,y}$ such that $\mathbb{P}[V_{x,y} = 0] = 1 - \varepsilon$ and $\mathbb{P}[V_{x,y} = 1] = \varepsilon$ leading to $\frac{\mu_2}{\mu_1} - \log \mu_1 = \log \frac{1}{\varepsilon}$)

Capacity of a random channel (cont'd)

- Define $\mu_1 := \mathbb{E}[V_{x,y}]$ and $\mu_2 := \mathbb{E}[V_{x,y} \log V_{x,y}]$

Theorem (Asymptotic capacity)

For $n \to \infty$ the capacity $C(\mathrm{W}^{(n)})$ of the DMC defined above converges to $\frac{\mu_2}{\mu_1} - \log \mu_1$ almost surely and in L^2.

Example 1 (Uniform distribution)

Consider $V_{x,y} \overset{\text{i.i.d.}}{\sim} \mathcal{U}([0, A])$ for $A > 0$, then $\lim_{n\to\infty} C(\mathrm{W}^{(n)}) = 1 - \frac{1}{2 \ln 2}$

Capacity of a random channel (cont'd)

- Define $\mu_1 := \mathbb{E}[V_{x,y}]$ and $\mu_2 := \mathbb{E}[V_{x,y} \log V_{x,y}]$

Theorem (Asymptotic capacity)

For $n \to \infty$ the capacity $C(\mathrm{W}^{(n)})$ of the DMC defined above converges to $\frac{\mu_2}{\mu_1} - \log \mu_1$ almost surely and in L^2.

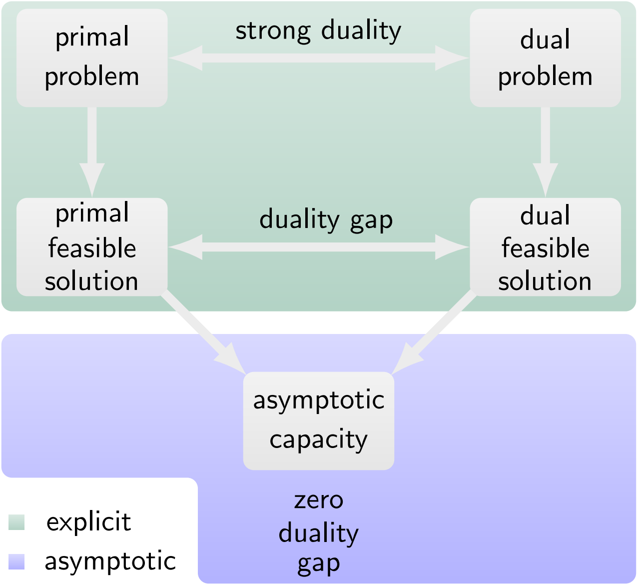
Example 1 (Uniform distribution)

Consider $V_{x,y} \overset{\text{i.i.d.}}{\sim} \mathcal{U}([0, A])$ for $A > 0$, then $\lim_{n\to\infty} C(\mathrm{W}^{(n)}) = 1 - \frac{1}{2 \ln 2}$

Example 2 (Symmetric Dirichlet distribution)

Consider $V_{x,y} \sim \mathcal{E}(\lambda)$ for $\lambda > 0$, then the channel rows are i.i.d. Dirichlet distributed, i.e., ${\mathrm{W}^{(n)}}_{x,\cdot} \overset{\text{i.i.d.}}{\sim} \mathrm{Dir}(\lambda, \cdots, \lambda)$ and $\lim_{n\to\infty} C(\mathrm{W}^{(n)}) = \frac{1-\gamma}{\ln 2}$, where γ (≈ 0.5772) denotes Euler's constant

Proof sketch



Proof sketch (lower bound)

- Primal program $C(\mathrm{W}^{(n)}) = \begin{cases} \max_p & I\big(p, \mathrm{W}^{(n)}\big) \\ & p \in \Delta_n \end{cases}$
- Lower bound $C_{\mathrm{LB}}^{(p\sim\mathcal{U})}(\mathrm{W}^{(n)}) := I\big(p, \mathrm{W}^{(n)}\big)$, where p is the uniform distribution on Δ_n

Lemma (Lower bound)

For $n \to \infty$, the random variable $C_{\mathrm{LB}}^{(p\sim\mathcal{U})}(\mathrm{W}^{(n)})$ converges to $\frac{\mu_2}{\mu_1} - \log \mu_1$ almost surely and in L^2.

Proof sketch (lower bound)

- Primal program $C(\mathrm{W}^{(n)}) = \begin{cases} \max\limits_{p} & I\big(p, \mathrm{W}^{(n)}\big) \\ & p \in \Delta_n \end{cases}$
- Lower bound $C_{\mathrm{LB}}^{(p\sim\mathcal{U})}(\mathrm{W}^{(n)}) := I\big(p, \mathrm{W}^{(n)}\big)$, where p is the uniform distribution on Δ_n

Lemma (Lower bound)

For $n \to \infty$, the random variable $C_{\mathrm{LB}}^{(p\sim\mathcal{U})}(\mathrm{W}^{(n)})$ converges to $\frac{\mu_2}{\mu_1} - \log \mu_1$ almost surely and in L^2.

- Introduce $q := {\mathrm{W}^{(n)}}^\top p$ [Chiang & Boyd'04]
- Equivalent formulation (prerequisite for the upper bound)

$$C(\mathrm{W}^{(n)}) = \begin{cases} \max\limits_{p,q} & -r^\top p + H(q) \\ \text{s. t.} & {\mathrm{W}^{(n)}}^\top p = q \\ & p \in \Delta_n, \;\; q \in \Delta_n \end{cases}$$

with $r_i := -\sum_{j=1}^N \mathrm{W}_{i,j}^{(n)} \log \mathrm{W}_{i,j}^{(n)}$

Proof sketch (upper bound)

Dual program (with strong duality)

$$C(\mathrm{W}^{(n)}) = \min_{\lambda} \{ G(\lambda) + F(\lambda) \ : \ \lambda \in \mathbb{R}^n \},$$

parametric LP ↓

$$G(\lambda) = \begin{cases} \max_{p} & -r^\top p + \lambda^\top \mathrm{W}^{(n)\top} p \\ \text{s.t.} & p \in \Delta_n, \end{cases}$$

$$= \max_{i \in [n]} \left(\mathrm{W}^{(n)} \lambda - r \right)_i$$

Entropy maximization ↓

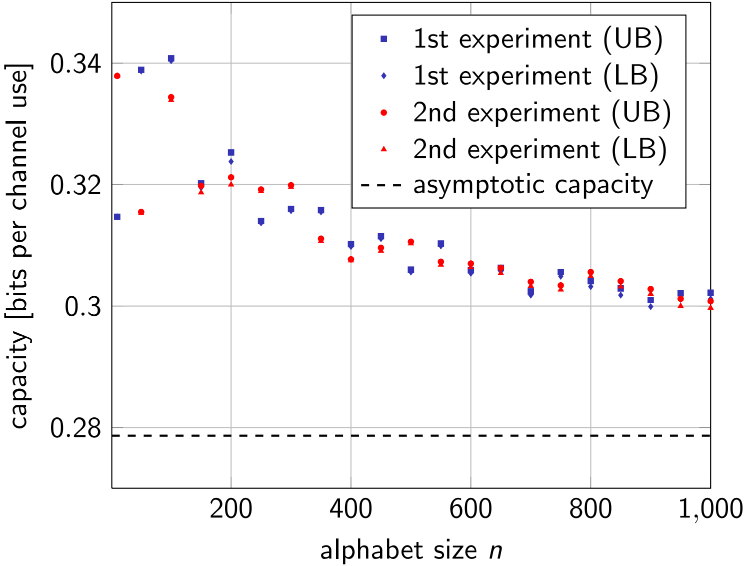
$$F(\lambda) = \begin{cases} \max_{q} & H(q) - \lambda^\top q \\ \text{s.t.} & q \in \Delta_n \end{cases}$$

$$= \log \left(\sum_{i=1}^{n} 2^{-\lambda_i} \right)$$

Lemma (Upper bound)

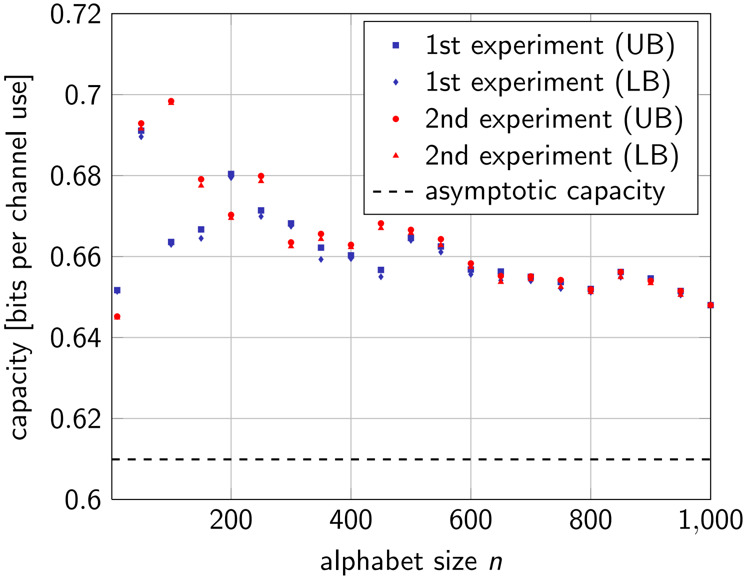
For $n \to \infty$, the random variable $C_{\mathrm{UB}}^{(\lambda=0)}(\mathrm{W}^{(n)}) := G(0) + F(0)$ converges to $\frac{\mu_2}{\mu_1} - \log \mu_1$ almost surely and in L^2.

Example 1 (Uniform distribution)



[TS *et al.*'14]

Example 2 (Symmetric Dirichlet distribution)



[TS *et al.*'14]

Rate of convergence

- Assume $V_{x,y} \in [a, b]$ almost surely for $0 \leq a < b < \infty$
- $c := \min_{x \in [a,b]} x \log x, \; d := \max_{x \in [a,b]} x \log x$
- $f_1(t, n) := 2 \exp\left(-\frac{2nt^2}{(b-a)^2}\right), \; f_2(t, n) := 2 \exp\left(-\frac{2nt^2}{(d-c)^2}\right)$

Theorem (Rate of convergence)

The capacity of the DMC defined above satisfies for any $t \in \mathbb{R}_{>0}$

$$
\begin{aligned}
&\mathbb{P}\left[\left|C(\mathrm{W}^{(n)}) - \left(\frac{\mu_2}{\mu_1} - \log \mu_1\right)\right| \geq t\right] \\
&\leq \left(f_1(\alpha_{t/2}, n) + f_2(\alpha_{t/2}, n) + f_1(\tfrac{t}{2L}, n)\right) \vee \\
&\qquad \left(f_1(\alpha_{t/4}, n) + f_2(\alpha_{t/4}, n) + f_1(\tfrac{t}{4L}, n) + 2f_1(\beta_{t/(2L)}, n)\right),
\end{aligned}
$$

$$
\beta_t = \frac{t\mu_1}{2+t}, \quad L = \frac{1}{a \ln 2}, \quad \alpha_t = \begin{cases} \frac{t\mu_1^2}{\mu_1(1+t)+\mu_2} & \text{if } \mu_1 + \mu_2 \geq 0 \\ \frac{t\mu_1^2}{\mu_1(1-t)+\mu_2} & \text{otherwise} \end{cases}
$$

Proof (sketch)

Key steps

- Main idea is to derive concentration bounds for the random variables $C_{\mathrm{UB}}^{(\lambda=0)}(\mathrm{W}^{(n)})$ and $C_{\mathrm{LB}}^{(p\sim\mathcal{U})}(\mathrm{W}^{(n)})$ around $\frac{\mu_2}{\mu_1} - \log \mu_1$
- Main tool is Hoeffding's inequality

Remarks

- Boundedness assumption of the random variables can be relaxed (e.g., by using a general Bernstein inequality)
- The rate is by no means optimal

Summary & Outlook

Summary

- Asymptotic capacity of DMCs, whose channel entries are chosen at random
- Exponential rate of convergence

Outlook

- Application and interpretation of the limiting capacity of a random channel in the context of Bayesian estimation and optimal experiment design.
- Decay rate of the variance
- Remove the i.i.d. assumption