Round Complexity of Common Randomness Generation: The Amortized Setting

Noah Golowich

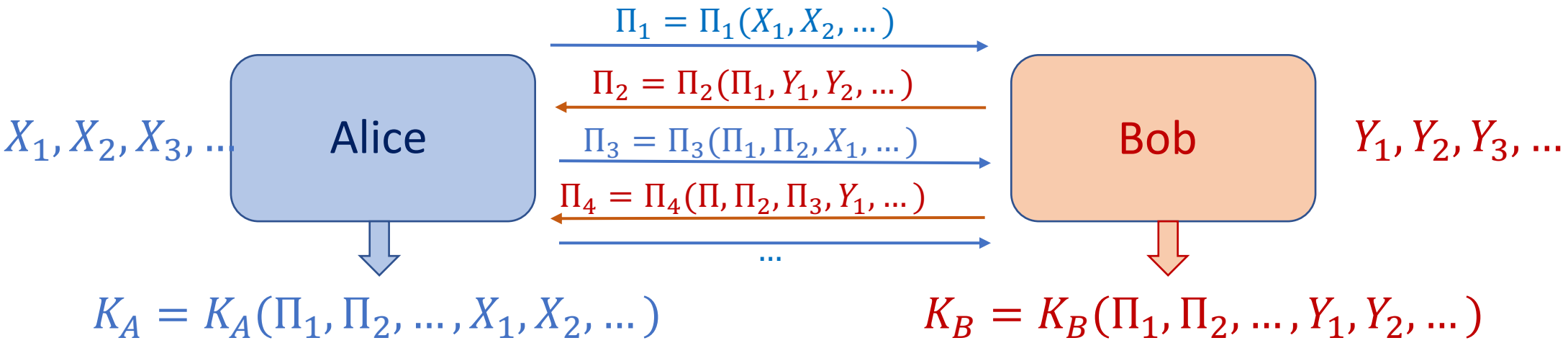
MIT

Madhu Sudan

Harvard University

Common Randomness Generation (CRG)

- Given **source distribution** μ & iid samples $(X_i, Y_i) \sim \mu$, $X_i, Y_i \in \{0,1\}^*$.
- **r rounds** of communication via messages $\Pi_1, \Pi_2, \dots, \Pi_r \in \{0,1\}^*$:

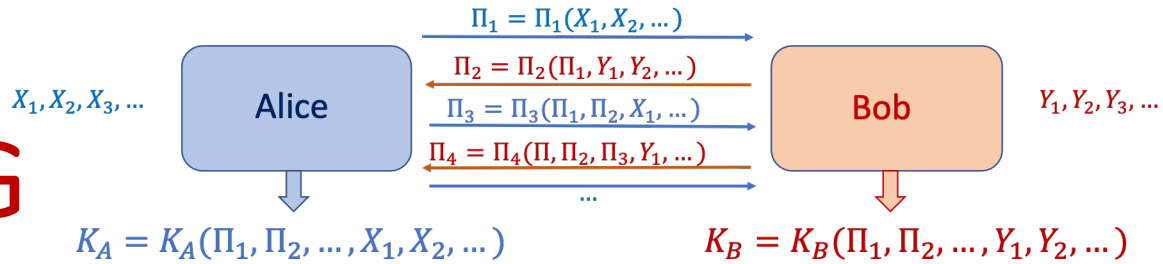


- Goal: output keys $K_A, K_B \in \{0,1\}^*$ of Alice, Bob satisfy:
  - $K_A = K_B$ with high probability.
  - K_A, K_B have large entropy.

Motivation for CRG

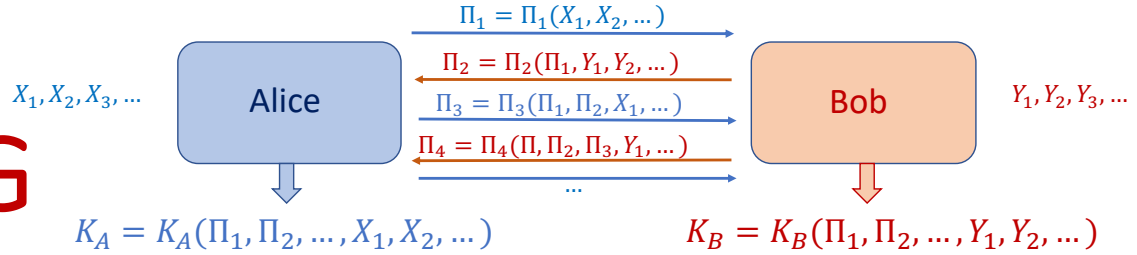
- Information-theoretically secure **Secret Key Generation (SKG):**
  - Same as CRG, but also require key $K_A = K_B$ to be secure against eavesdropper watching communication of $\Pi_1, \Pi_2, \Pi_3, \ldots$.
  - Once parties agree on a key, can use **private-key crypto** to communicate securely, so avoid needing computational assumptions of **public-key crypto**.
  - (Our results for CRG give analogous immediate corollaries for SKG.)
- **Communication complexity** w/ imperfectly shared randomness:
  - Generating shared key is one way to proceed.
- Others: **Coding theory**, **locality-sensitive hashing**.

Round separations in CRG



- Two measures of efficiency in CRG:
  - Total number of communicated bits, C.
  - Number of rounds of interaction (i.e., number of messages), r.
- **Main question: are there sources μ such that having more rounds can lead to lower communication protocols for CRG?**
- [Bafna et al., SODA'19]: answer is "yes" in so-called *non-amortized setting*.
- [This work]: answer is "yes" in the *amortized setting*.
  - Also: we show stronger tradeoff than [Bafna et al.] for both settings.

Amortized setting of CRG



- **r-round protocol** $\Pi = (\Pi_1, \ldots, \Pi_r)$ consists of sequence of r message functions.
- **C…n complexity** of Π, $CC(\Pi)$, is max number of bits c…l (taken over all inputs $(X_1, Y_1), (X_2, Y_2), \ldots$).

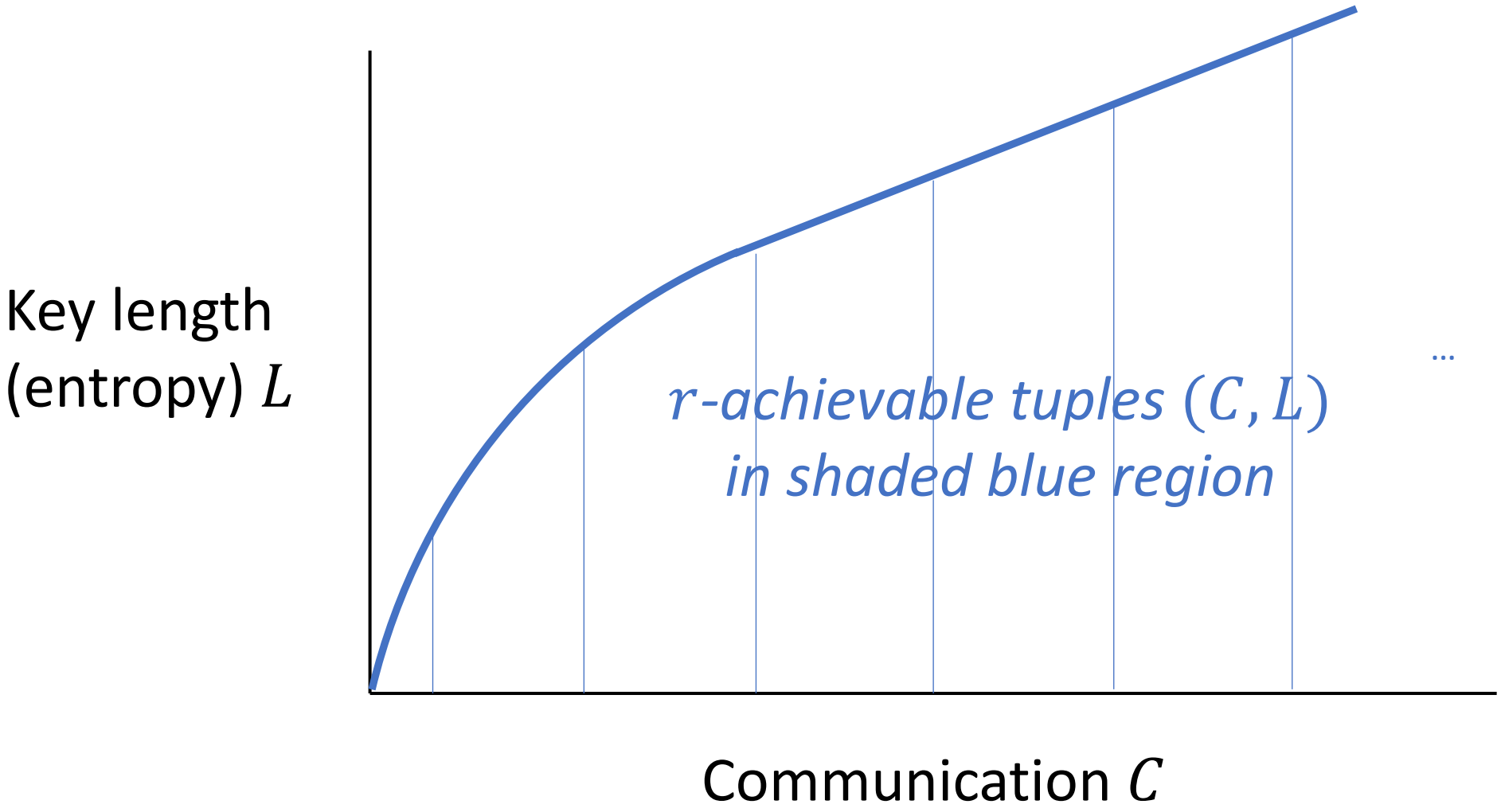
$C > 0$: comm. complexity
$L > 0$: key length (entropy)

Defin… (Amortized CRG rates; eliminating some technical details):

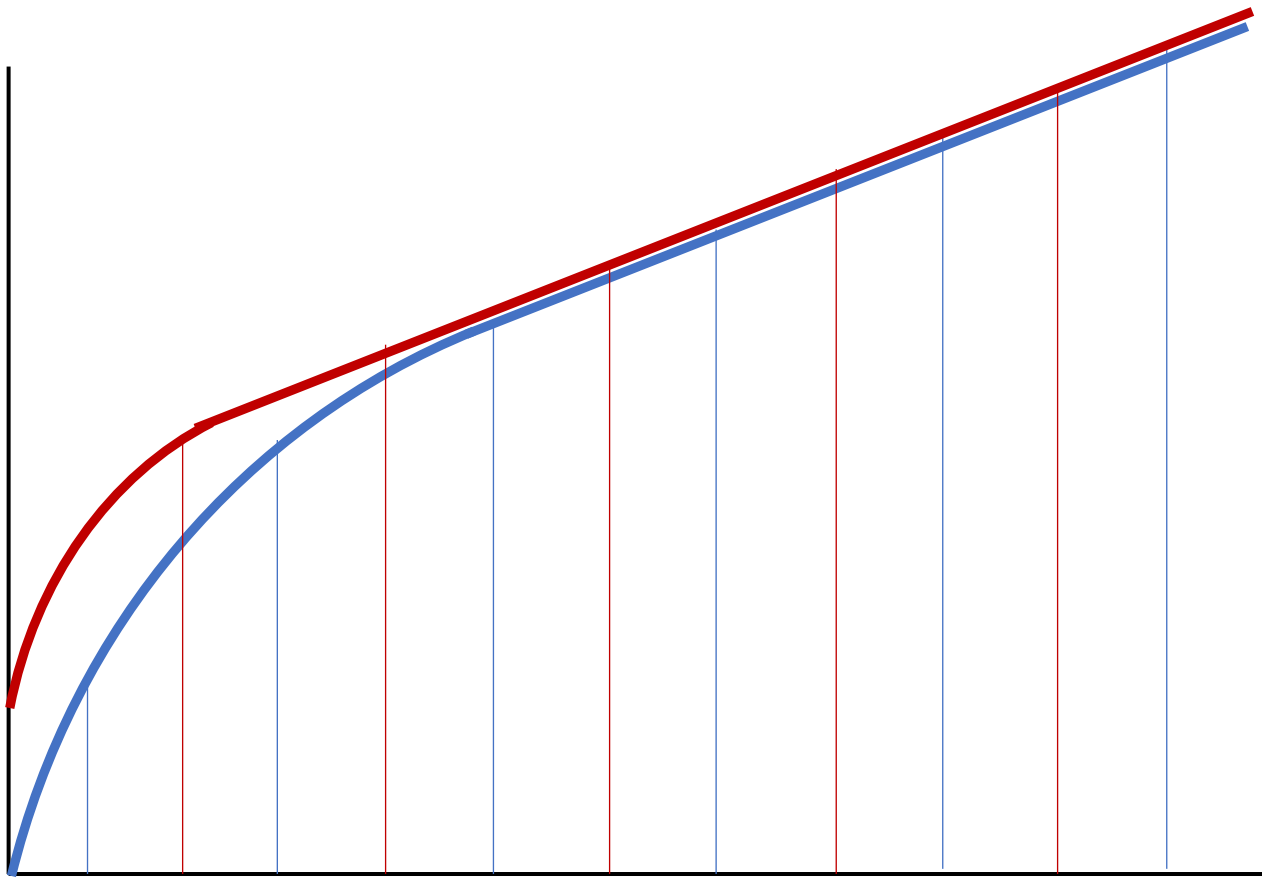
Tuple (C, L) is **r-achievable for CRG** under μ if there are protocols $\Pi^N, N \in \mathbb{N}$ with input $(X_1, Y_1), \ldots, (X_N, Y_N) \sim \mu$ iid, producing output keys (K_A^N, K_B^N), so that:

1. "Amortized communication at most C": $\limsup_{N\to\infty} \frac{CC(\Pi^N)}{N} \leq C$.
2. "Amortized key entropy at least L": $\liminf_{N\to\infty} \frac{H(K_A^N)}{N} \geq L$.
3. "Key disagreement probability close to 0": $\lim_N \Pr[K_A^N \neq K_B^N] = 0$.

Representative shape of r-achievable region for some μ



Our question: exists μ so that increasing r increases size of achievable region?



Key length (entropy) L

Communication C

r'-achievable tuples (C, L) for μ in red region, $r' > r$?

...

r-achievable tuples (C, L) for μ in blue region

Prior work: rounds-communication tradeoffs

- [Ahlswede & Csiszár, '98], [Tyagi, '13], [Liu et al., '16]: for all r, μ, proved characterization of r-achievable rates (C, L) for μ:
  - **Single-letter characterization**: algorithm, given μ, deciding whether some rate (C, L) is arbitrarily close to r-achievability for μ.
- **Using above results**: if μ is **binary** (i.e., $X, Y \in \{0, 1\}$) and **symmetric**, additional rounds (probably) does not help:
  - [Liu et al., '16] conjecture: set of 1-achievable (C, L) = set of r-achievable (C, L) $\forall r > 1$ (+ prove a weak version of this).
- For **ternary** sources (i.e., $X, Y \in \{0, 1, 2\}$):
  - [Tyagi, '13]: showed there exists a ternary source μ s.t.:

    2-achievable region for μ is strictly larger than 1-achievable region

    But: 1. Only show a constant factor gap;

    2. What about $r > 2$?

Our theorem: exponential rounds-communication tradeoffs for all r

Theorem [This work]:

For all $r \in \mathbb{N}$, $L > 0$, there is a source $\mu_{r,L}$ so that:

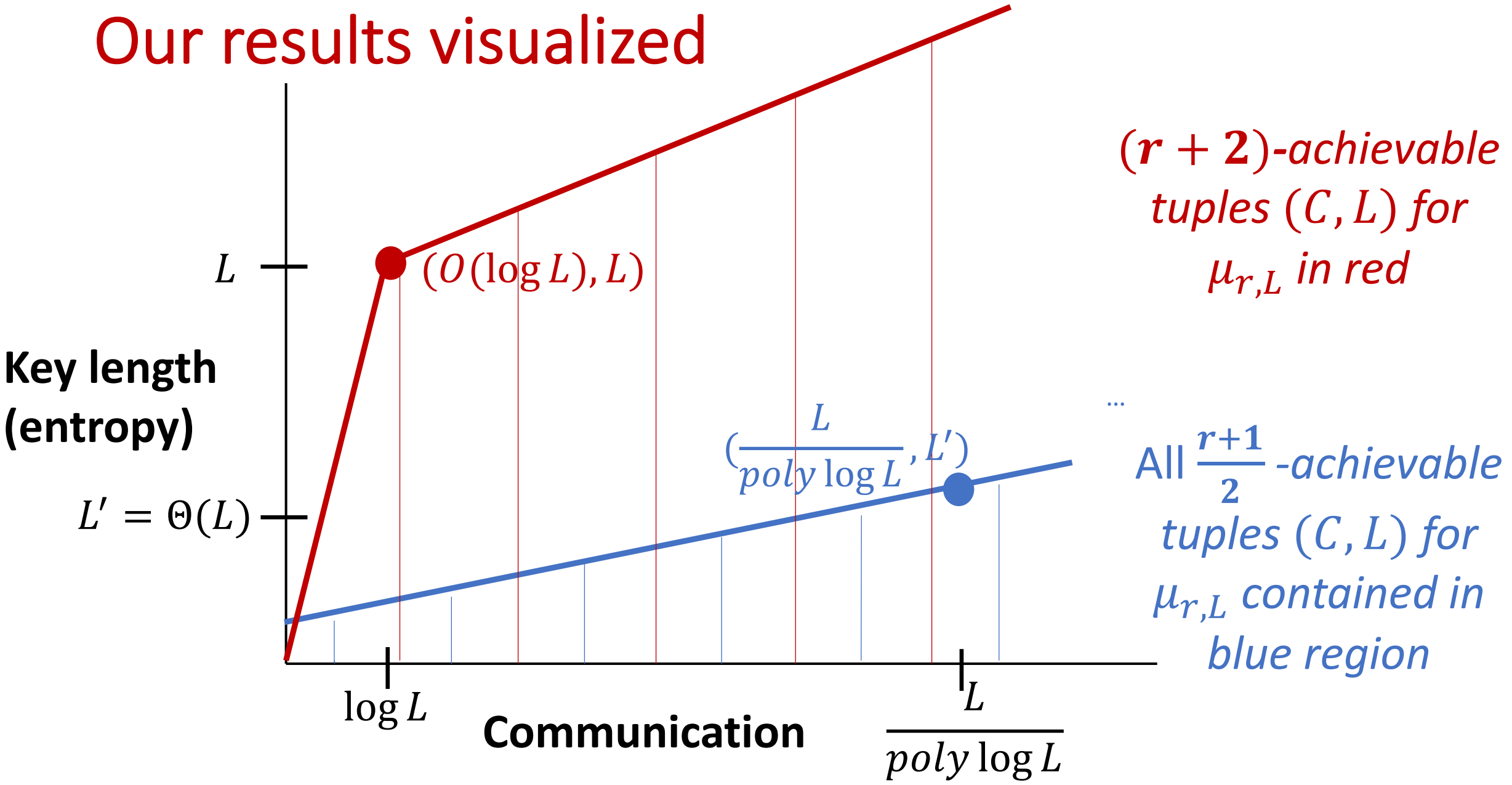
- $\exists C \leq O(\log L)$ s.t. (C, L) is $(r+2)$-achievable from $\mu_{r,L}$;

But; for all $L' \geq \Omega(L)$:

1. If (C, L') is $\frac{r+1}{2}$-achievable from $\mu_{r,L}$, then $C \geq \frac{L}{poly \log L}$.
2. If (C, L') is r-achievable from $\mu_{r,L}$, then $C \geq \frac{\sqrt{L}}{poly \log L}$.

- Note: (L, L) is 1-achievable from any source (so $(\frac{L}{poly \log L}, L)$ in 1. is nearly tight).

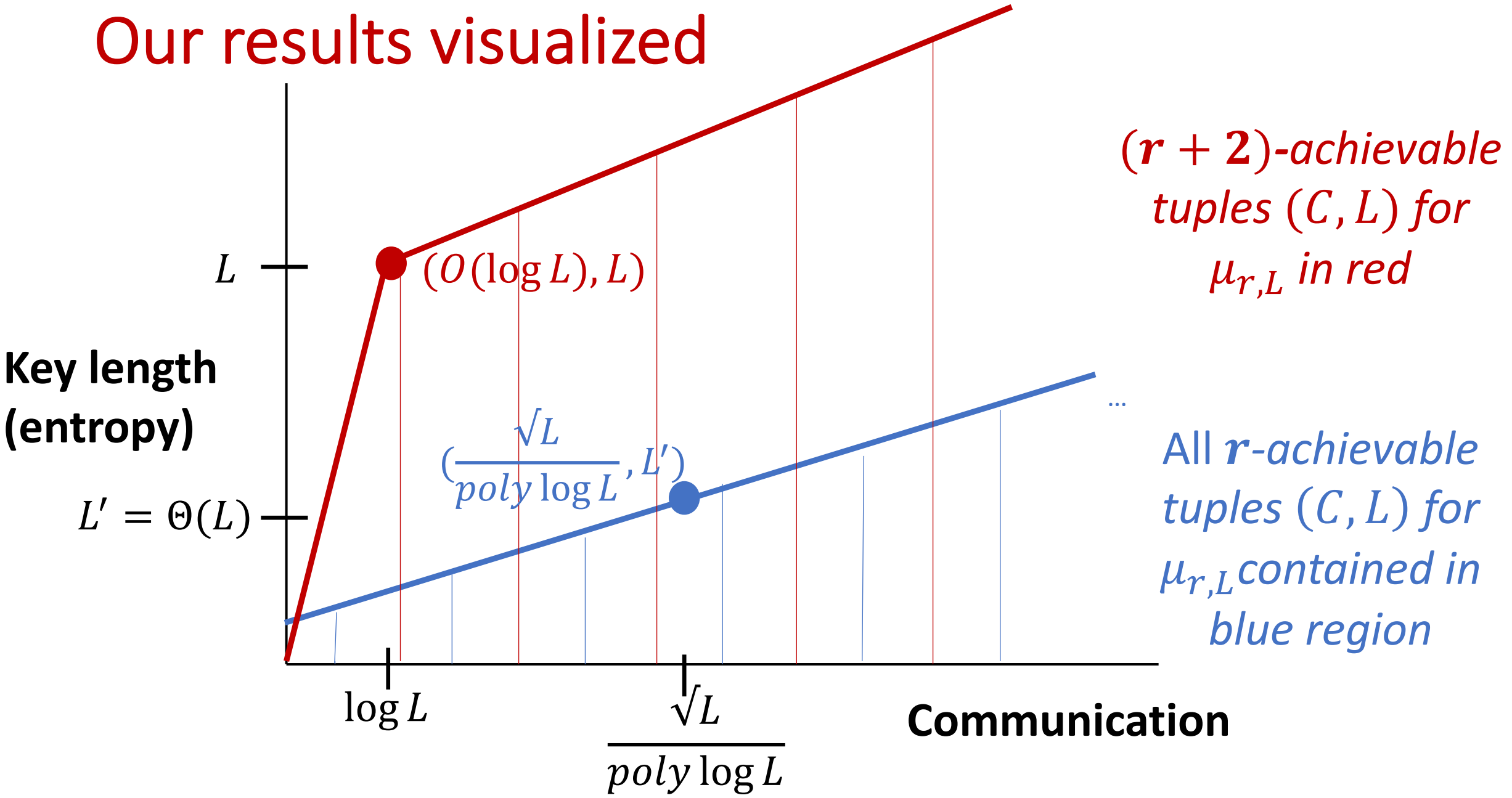
Our results visualized



$(\boldsymbol{r+2})$-*achievable tuples* (C, L) *for* $\mu_{r,L}$ *in red*

All $\frac{\boldsymbol{r+1}}{\boldsymbol{2}}$ *-achievable tuples* (C, L) *for* $\mu_{r,L}$ *contained in blue region*

Our results visualized



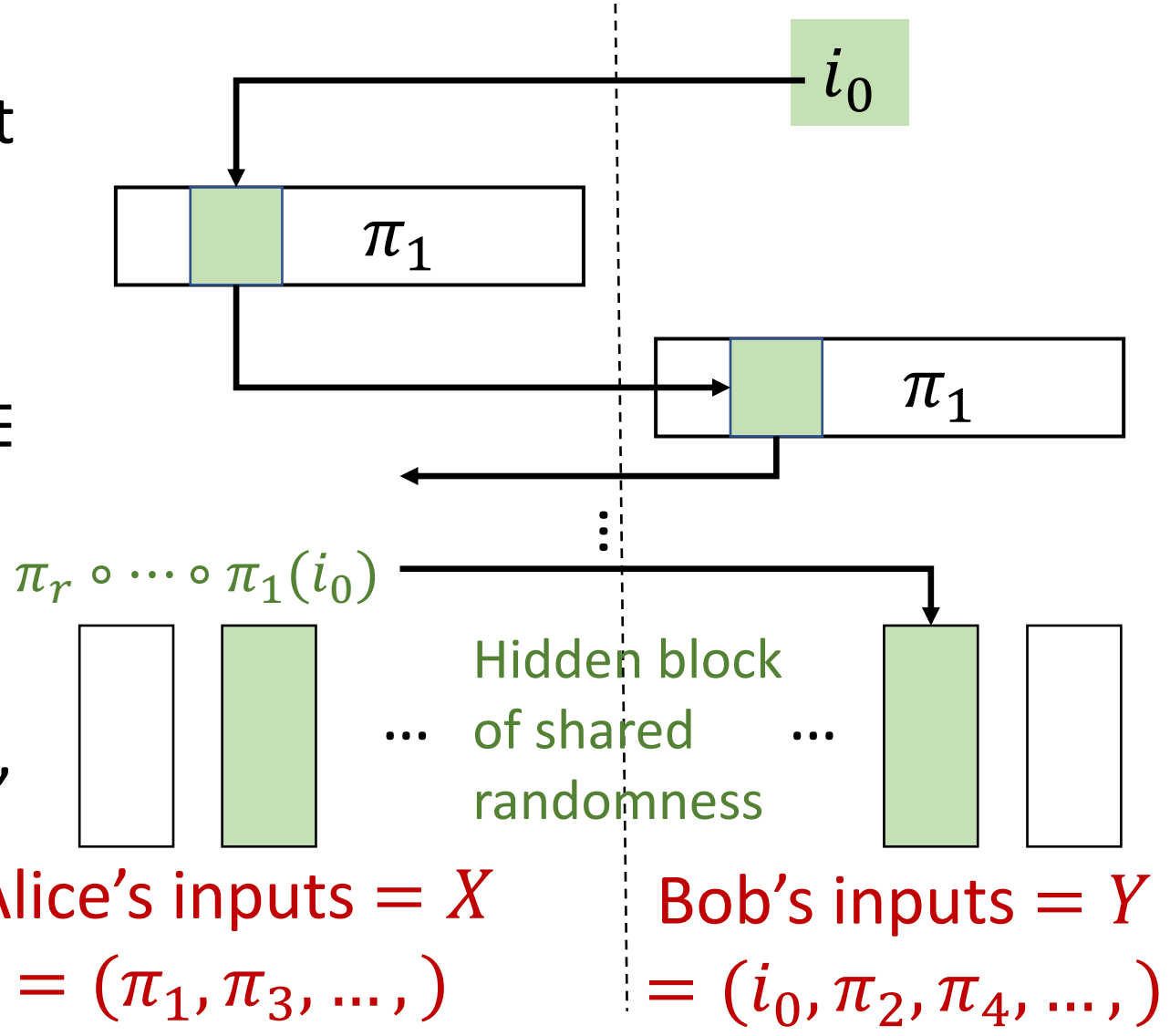
$(r+2)$-achievable tuples (C, L) for $\mu_{r,L}$ in red

All r-achievable tuples (C, L) for $\mu_{r,L}$ contained in blue region

What is the source $\mu_{r,L}$ in main theorem?

Pointer chasing source [Bafna et al, '19], [Nisan & Wigderson, '93]:

- Idea: random permutations $\pi_1, \ldots, \pi_r : [L] \to [L]$, index $i_0 \in [L]$.
- Alice + Bob get alternating $\pi_1, \pi_2, \ldots$
- $\Rightarrow$ To compute $\pi_r \circ \cdots \circ \pi_1(i_0)$, need $r + 2$ rounds.



Proof idea of main theorem

Bulk of proof: if (C, L) is $\frac{r+1}{2}$-achievable from $\mu_{r,L}$, then $C \geq \frac{L}{poly \log L}$.

- **Two main ingredients:**

1. Result of [Bafna et al., '19] showing analogous rounds-communication tradeoff for *non-amortized* setting.
   - In particular: our source $\mu_{r,L}$ is the same as that of [Bafna et al.].
2. Tools of *information complexity* [Barak et al., '10]:
   - [Braverman & Rao, '11], [Jain et al., '15]: **direct sum** results for *computing functions*:
   - I.e., relate *amortized communication complexity* to *(non-amortized) communication complexity*.

- **Main difficulty**: CRG is "easier" task than computing a function
  - So harder to prove lower bounds…

Quick detour: amortized vs. non-amortized CRG

Definition (Amortized CRG rates; eliminating some technical details):

Tuple (C, L) is **r-achievable for CRG** under μ if there are protocols $\Pi^N, N \in \mathbb{N}$ with input $(X_1, Y_1), \ldots, (X_N, Y_N) \sim \mu$ iid, producing output keys (K_A^N, K_B^N), so that:

1. "Amortized communication at most C": $\limsup_{N\to\infty} \frac{CC(\Pi^N)}{N} \leq C$.
2. "Amortized key entropy at least L": $\liminf_{N\to\infty} \frac{H(K_A^N)}{N} \geq L$.
3. "Key disagreement probability close to 0": $\lim_N \Pr[K_A^N \neq K_B^N] = 0$.

Definition (Non-Amortized CRG rates):

Tuple (C, L, ϵ) is **r-achievable for CRG** under μ if there is single $N \in \mathbb{N}$, protocol Π^N, with input $(X_1, Y_1), \ldots, (X_N, Y_N) \sim \mu$ iid, with output keys (K_A^N, K_B^N), so that:

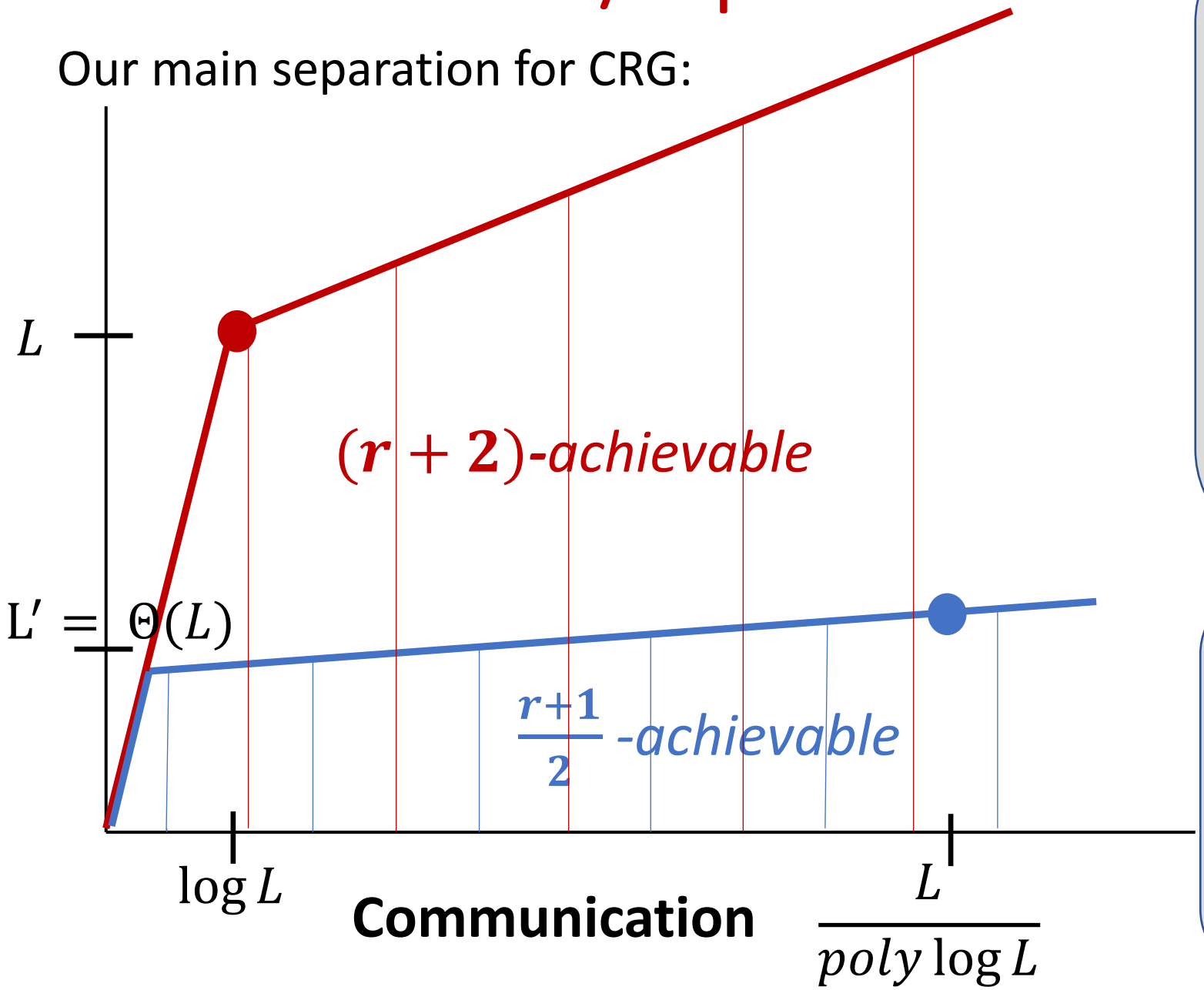
1. "*Communication* at most C": $\mathrm{CC}(\Pi^N) \leq C$.
2. "*Key min-entropy* at least L": $H_\infty(K_A^N), H_\infty(K_B^N) \geq L$.
3. "Key disagreement probability *close to 0*": $\Pr[K_A^N \neq K_B^N] \leq \epsilon$.

Lower bound proof: amortized vs non-amortized

- Why is reduction to non-amortized case [Bafna et al., ’19] nontrivial?
- May be “*easier to agree on a key*” in amortized setting (i.e., proving lower bounds is harder):
  - May be easier to extract keys for multiple samples by communicating “*in bulk*”; amortizing allows communication to be “*spread out*”.
- Our proof: shows *it is not much easier* in amortized setting:
  - Given a protocol Π with inputs from $\mu^{\otimes N}$, communication $\approx C \cdot N$, and key length $\approx L \cdot N$:
  - Can construct protocol Π' with inputs from μ, communication $\approx C$, and key length $\approx L$.

Conclusion / Open Problem

Our main separation for CRG:



Open Problem:

Can we relax the requirement that $L' = \Theta(L)$? I.e., "move the blue line down"?

Formally: show that (C, L') is not $\frac{r+1}{2}$-achievable for some $L' \ll \Theta(L), C < \frac{L}{poly \log L}$.

Bolder Open Problem:

Prove rounds-communication tradeoff for the quantity:

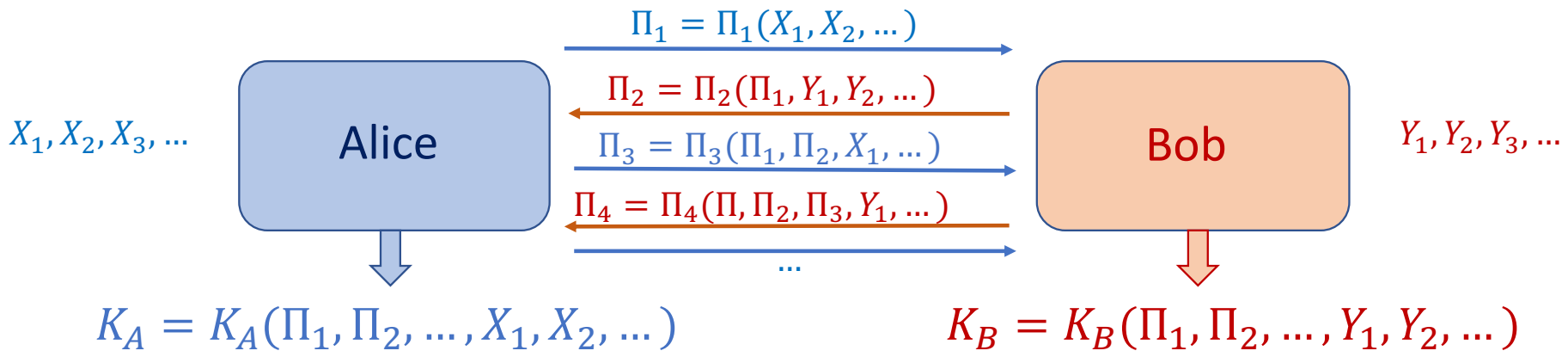
$\sup \{L/C : (C, L) \text{ is } r\text{-achievable}\}$

(known as **CBIB**, or **r-round SDPC**).

Thank you!

I am grateful to the NSF for providing a student travel grant.

Quick detour: amortized vs non-amortized CRG



Definition (Amortized CRG rates; eliminating some technical details):

Tuple (C, L) is ***r*-achievable for CRG** under μ if there are protocols $\Pi^N, N \in \mathbb{N}$ with input $(X_1, Y_1), \dots, (X_N, Y_N) \sim \mu$ iid, producing output keys (K_A^N, K_B^N), so that:

1. "Amortized communication at most C": $\limsup_{N\to\infty} \frac{CC(\Pi^N)}{N} \leq C$.
2. "Amortized key entropy at least L": $\liminf_{N\to\infty} \frac{H(K_A^N)}{N} \geq L$.
3. "Key agreement probability close to 1": $\lim_N \Pr[K_A^N \neq K_B^N] \to 0$.