

Generalized quantum asymptotic equipartition and applications

Kun FANG

Joint work with Hamza Fawzi and Omar Fawzi



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen



UNIVERSITY OF
CAMBRIDGE



arXiv: 2411.04035 & 2502.15659 & 2506.03060

BIID 2025 @ Munich, July 2025

What is “asymptotic equipartition”?

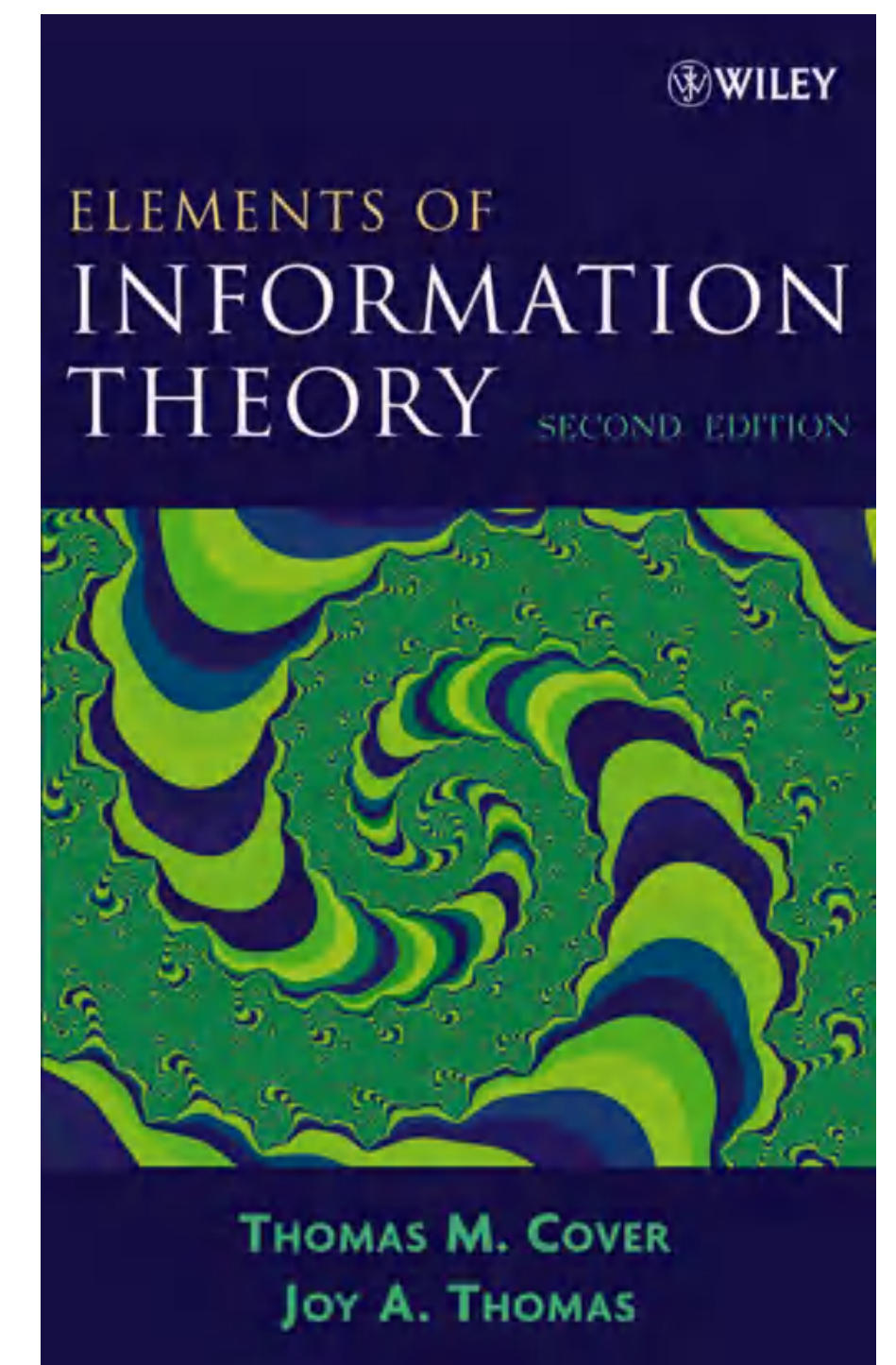
Asymptotic equipartition property (AEP)

A form of the law of large numbers in information theory

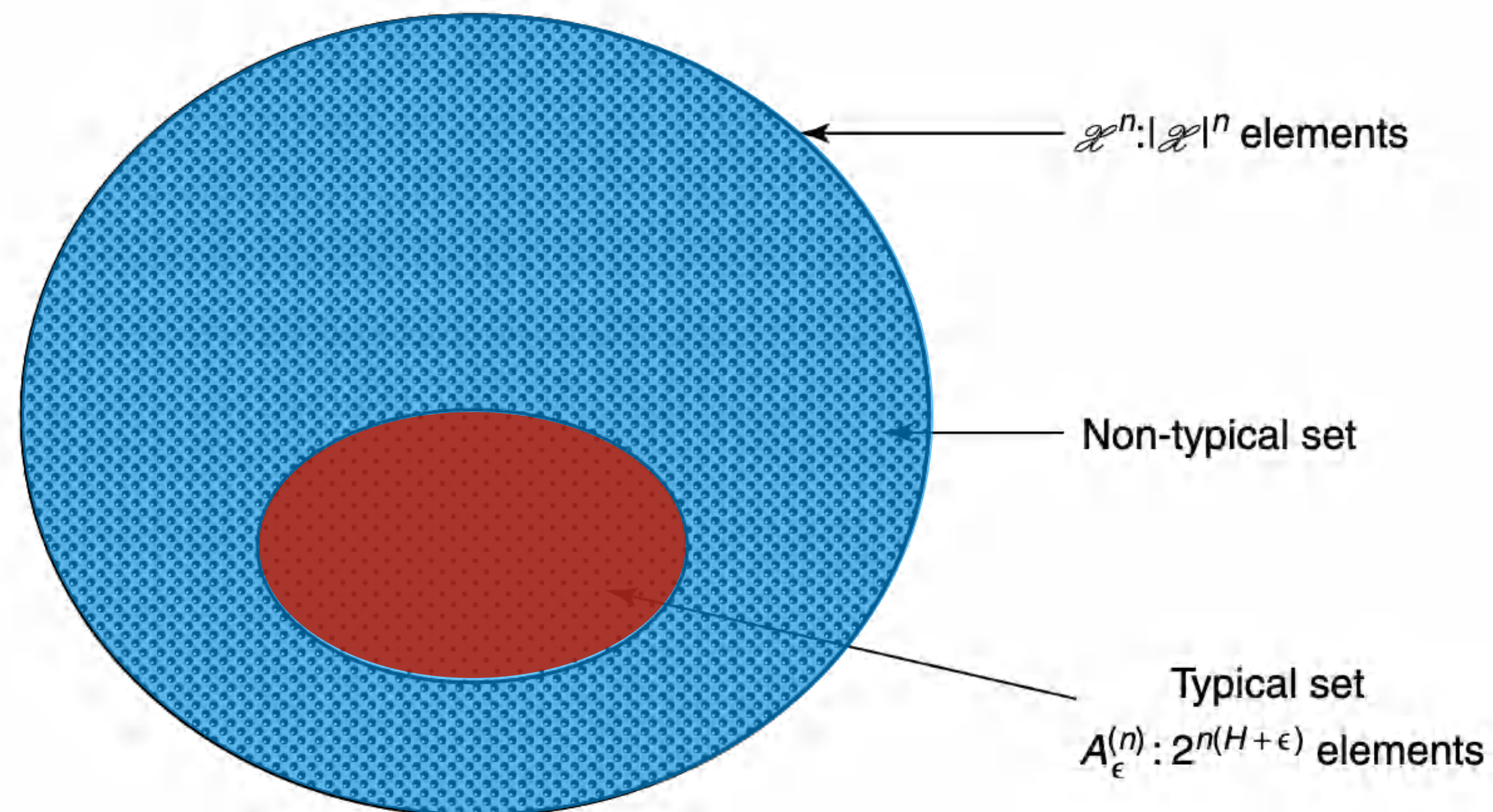
AEP or Shannon-MacMillan-Breiman theorem

Given i.i.d. random variables X_1, X_2, \dots, X_n , the probability $p(X_1, X_2, \dots, X_n)$ satisfies

$$-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) \longrightarrow H(X) \quad \text{in probability}$$



What is “asymptotic equipartition”?



Bit strings of length n

Typical set v.s. Non-typical set

Size of the typical set is nearly $2^{nH(X)}$

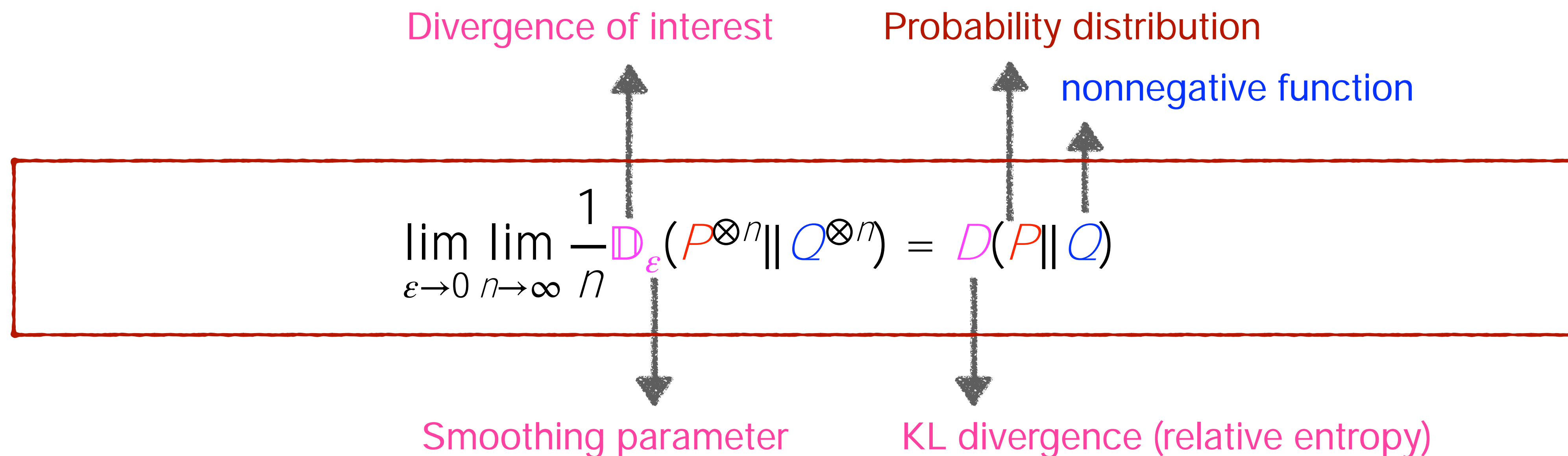
The typical set has probability nearly 1

Elements in the typical set are nearly **equiprobable**

Lie in the heart of information theory:

data compression, channel coding, cryptography...

More generic form of **AEP** in divergences



More generic form of **AEP** in divergences

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{P}^{\otimes n} \parallel \mathcal{Q}^{\otimes n}) = D(\mathcal{P} \parallel \mathcal{Q})$$

Shannon-McMillan-Breiman theorem:

$\mathbb{D} = H_{\min}$ or H_{\max} , $\mathcal{Q} = 1$ constant function e.g. [Tomamichel, Colbeck, Renner 2009]

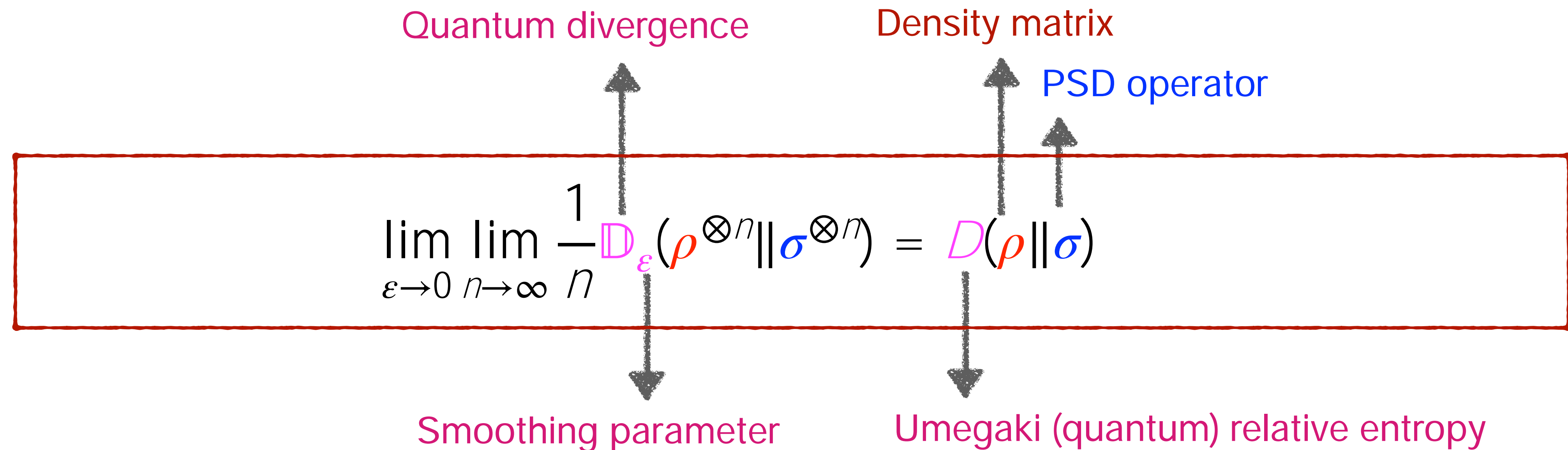
H_{\max} : the size of the typical set & H_{\min} : the distribution is uniform on the typical set

Chernoff-Stein Lemma:

$\mathbb{D} = D_H$ hypothesis testing relative entropy

Generalization to quantum AEP?

Generalization to quantum AEP?



Generalization to quantum AEP?

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho \| \sigma)$$

- Hiai and Petz 1991: $\mathbb{D} = D_H$
 - Ogawa and Nagaoka 2000: remove ε -dependence in the outer limit
 - Tomamichel, Colbeck, Renner 2009: $\sigma_{AB} = I_A \otimes \rho_B$, $H_{\min}(A|B)$ and $H_{\max}(A|B)$
 - Tomamichel, Hayashi 2013: $\mathbb{D} = D_{\max}$
- Quantum
Stein's lemma

Many applications: quantum data compression, quantum state merging, quantum channel coding, quantum cryptography, and quantum resource theory...

Generalization to quantum AEP?

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho \| \sigma)$$

Limited to singleton and i.i.d. structure

Generalization to quantum AEP?

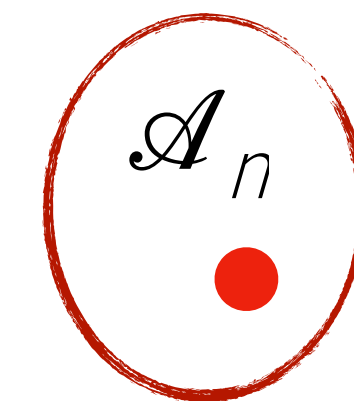
$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho \| \sigma)$$

Limited to singleton and i.i.d. structure

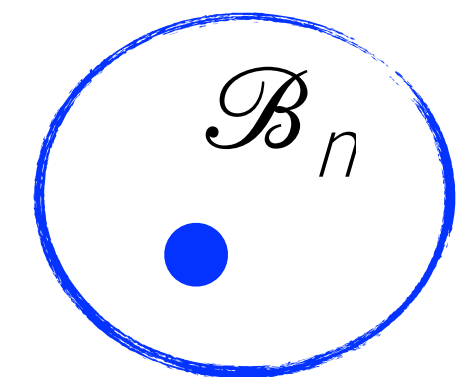
What if?

Correlation: **beyond i.i.d.** source $\rho_n \neq \rho^{\otimes n}, \sigma_n \neq \sigma^{\otimes n}$

Uncertainty: **not singleton** $\rho_n \in \mathcal{A}_n$ and $\sigma_n \in \mathcal{B}_n$



v.s.



e.g. composite hypothesis

Generalization to quantum AEP?

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho \| \sigma)$$

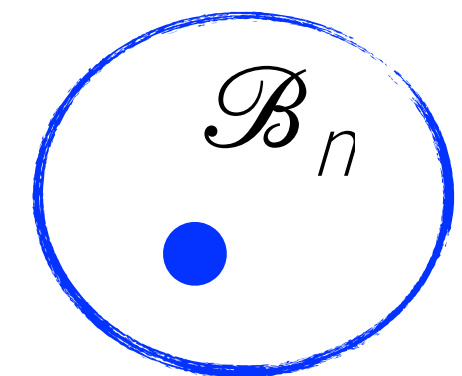
What if?

Correlation: **beyond i.i.d.** source $\rho_n \neq \rho^{\otimes n}, \sigma_n \neq \sigma^{\otimes n}$

Uncertainty: **not singleton** $\rho_n \in \mathcal{A}_n$ and $\sigma_n \in \mathcal{B}_n$



v.s.



e.g. composite hypothesis

Practical motivations in the classical setting e.g. [Leviton and Nerhav 2002, TIT]

Classification with training sequences (e.g. speech recognition, signal detection)

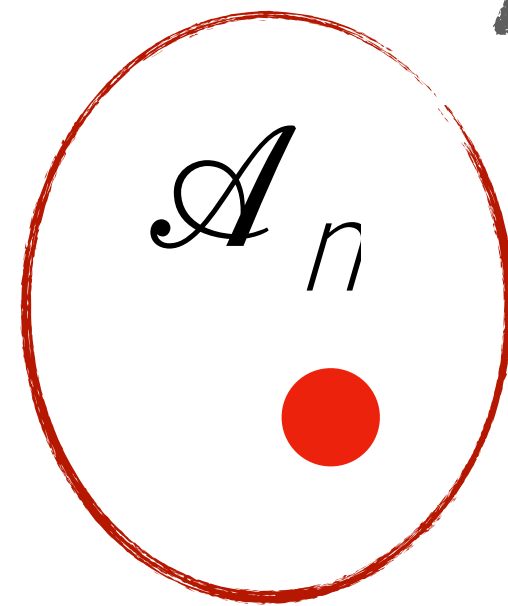
Detection of messages via unknown channels (e.g. radar target detection, watermark detection)

Generalization to quantum AEP beyond i.i.d. and singleton

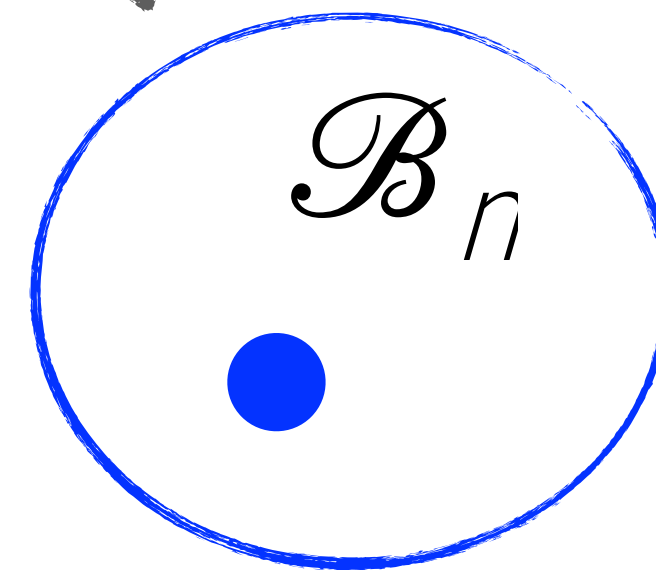
Generalization to quantum AEP beyond i.i.d. and singleton

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = ?$$

A set of quantum states



v.s.



A set of PSD operators

$$\mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) := \inf_{\rho_n \in \mathcal{A}_n, \sigma_n \in \mathcal{B}_n} \mathbb{D}_{\varepsilon}(\rho_n \| \sigma_n)$$

Generalization to quantum AEP beyond i.i.d. and singleton

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = ?$$

A very **general framework** that contains almost all existing quantum AEP in the literature

Generalized quantum Stein's lemma

$$\mathcal{A}_n = \{\rho^{\otimes n}\}$$

\mathcal{B}_n a set of quantum states

Generalized quantum Sanov setting

\mathcal{A}_n a set of quantum states

$$\mathcal{B}_n = \{\rho^{\otimes n}\}$$

Talk by Yamasaki and Talk by Lami on Monday

Generalization to quantum AEP beyond i.i.d. and singleton

Our answer

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = D^{\infty}(\mathcal{A} \| \mathcal{B})$$

$$\mathbb{D} \in \{D_H, D_{\max}\}$$

$$D^{\infty}(\mathcal{A} \| \mathcal{B}) := \lim_{n \rightarrow \infty} \frac{1}{n} D(\mathcal{A}_n \| \mathcal{B}_n)$$

Generalization to quantum AEP beyond i.i.d. and singleton

Our answer

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = D^{\infty}(\mathcal{A} \| \mathcal{B})$$

Generality (divergence):

two extreme cases $\mathbb{D} \in \{D_H, D_{\max}\}$

any divergence in between or equivalent, yield the same result

Our answer

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B})$$

Generality (sets):

Polar set $\mathcal{C}^\circ := \{X : \langle X, Y \rangle \leq 1, \forall Y \in \mathcal{C}\}$

(A.1) Each \mathcal{A}_n is convex and compact;

(A.3) $\mathcal{A}_m \otimes \mathcal{A}_k \subseteq \mathcal{A}_{m+k}$, for all $m, k \in \mathbb{N}$;

(A.2) Each \mathcal{A}_n is permutation-invariant;

(A.4) $(\mathcal{A}_m)_+^\circ \otimes (\mathcal{A}_k)_+^\circ \subseteq (\mathcal{A}_{m+k})_+^\circ$, for all $m, k \in \mathbb{N}$;

Our answer

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B})$$

Generality (sets):

Polar set $\mathcal{C}^\circ := \{X : \langle X, Y \rangle \leq 1, \forall Y \in \mathcal{C}\}$

(A.1) Each \mathcal{A}_n is convex and compact;

(A.3) $\mathcal{A}_m \otimes \mathcal{A}_k \subseteq \mathcal{A}_{m+k}$, for all $m, k \in \mathbb{N}$;

(A.2) Each \mathcal{A}_n is permutation-invariant;

(A.4) $(\mathcal{A}_m)_+^\circ \otimes (\mathcal{A}_k)_+^\circ \subseteq (\mathcal{A}_{m+k})_+^\circ$, for all $m, k \in \mathbb{N}$;

Sets	Mathematical descriptions
Singleton	$\{\rho^{\otimes n}\}$ with $\rho \in \mathcal{D}(\mathcal{H})$
Conditional states	$\{I_n \otimes \rho_n : \rho_n \in \mathcal{D}(\mathcal{H}^{\otimes n})\}$
Channel image	$\{\mathcal{N}^{\otimes n}(\rho_n) : \rho_n \in \mathcal{D}(\mathcal{H}^{\otimes n})\}$ with a quantum channel \mathcal{N}
Recovery set	$\{\mathcal{N}_{B^n \rightarrow C^n}(\rho_{AB}^{\otimes n}) : \mathcal{N} \in \text{CPTP}(B^n : C^n)\}$ with $\rho \in \mathcal{D}(AB)$
Extensions set	$\{\omega_n \in \mathcal{D}(A^n B^n) : \text{Tr}_{B^n} \omega_n = \rho_A^{\otimes n}\}$ with $\rho_A \in \mathcal{D}(A)$
Incoherent states	$\{\rho_n \in \mathcal{D}(\mathcal{H}^{\otimes n}) : \rho_n = \Delta(\rho_n)\}$ with the completely dephasing channel Δ
Rains set	$\{\rho_n \in \mathcal{H}_+(A^n B^n) : \ \rho_n^{\top_{B_1 \cdots B_n}}\ _1 \leq 1\}$ with the partial transpose \top_{B_i}
Nonpositive mana	$\{\rho_n \in \mathcal{H}_+(\mathcal{H}^{\otimes n}) : \ \rho_n\ _{W,1} \leq 1\}$ with the Wigner trace norm $\ \cdot\ _{W,1}$

Our answer

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B})$$

Generality (sets):

Polar set $\mathcal{C}^\circ := \{X : \langle X, Y \rangle \leq 1, \forall Y \in \mathcal{C}\}$

(A.1) Each \mathcal{A}_n is convex and compact;

(A.3) $\mathcal{A}_m \otimes \mathcal{A}_k \subseteq \mathcal{A}_{m+k}$, for all $m, k \in \mathbb{N}$;

(A.2) Each \mathcal{A}_n is permutation-invariant;

(A.4) $(\mathcal{A}_m)_+^\circ \otimes (\mathcal{A}_k)_+^\circ \subseteq (\mathcal{A}_{m+k})_+^\circ$, for all $m, k \in \mathbb{N}$;

More importantly, **without (A.4)**, the AEP does not hold in general.

Counterexamples e.g.

arXiv: 2501.09303v2 by Hayashi & arXiv: 2408.07067 by Lami, Berta, Regula

Our answer

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = D^{\infty}(\mathcal{A} \| \mathcal{B})$$

Efficiency: ★★★★★

Regularization instead of single-letter formula. But it can be estimated by

$$\frac{1}{m} D_M(\mathcal{A}_m \| \mathcal{B}_m) \leq D^{\infty}(\mathcal{A} \| \mathcal{B}) \leq \frac{1}{m} D(\mathcal{A}_m \| \mathcal{B}_m)$$

with explicit convergence guarantees,

$$\frac{1}{m} D(\mathcal{A}_m \| \mathcal{B}_m) - \frac{1}{m} D_M(\mathcal{A}_m \| \mathcal{B}_m) \leq \frac{1}{m} 2(d^2 + d) \log(m + d)$$

Efficiently approximate $D^{\infty}(\mathcal{A} \| \mathcal{B})$ within an additive error by a quantum relative entropy program of polynomial size. [arXiv: 2502.15659]

Our answer

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = D^{\infty}(\mathcal{A} \| \mathcal{B})$$

★ **Explicit finite n estimate:**

$$n D^{\infty}(\mathcal{A} \| \mathcal{B}) - \underbrace{O(n^{2/3} \log n)} \leq \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) \leq n D^{\infty}(\mathcal{A} \| \mathcal{B}) + \underbrace{O(n^{2/3} \log n)}$$

Leading term is regularized, but **still provide an explicit estimate** for finite n , making its **convergence controllable**; a rare case in QIT

Leading term independent of ε (strong converse property)

The second order in $O(n^{2/3} \log n)$ instead of $O(\sqrt{n})$, potential improvement exists

Key technical tools

Measured relative entropy $D_M(\rho\|\sigma) := \sup_M D(P_{\rho,M}\|P_{\sigma,M})$

Superadditivity $D_M(\rho_1 \otimes \rho_2\|\sigma_1 \otimes \sigma_2) \geq D_M(\rho_1\|\sigma_1) + D_M(\rho_2\|\sigma_2)$

$$D(\rho\|\sigma) = \lim_{n \rightarrow \infty} \frac{1}{n} D_M(\rho^{\otimes n}\|\sigma^{\otimes n})$$

Subadditivity

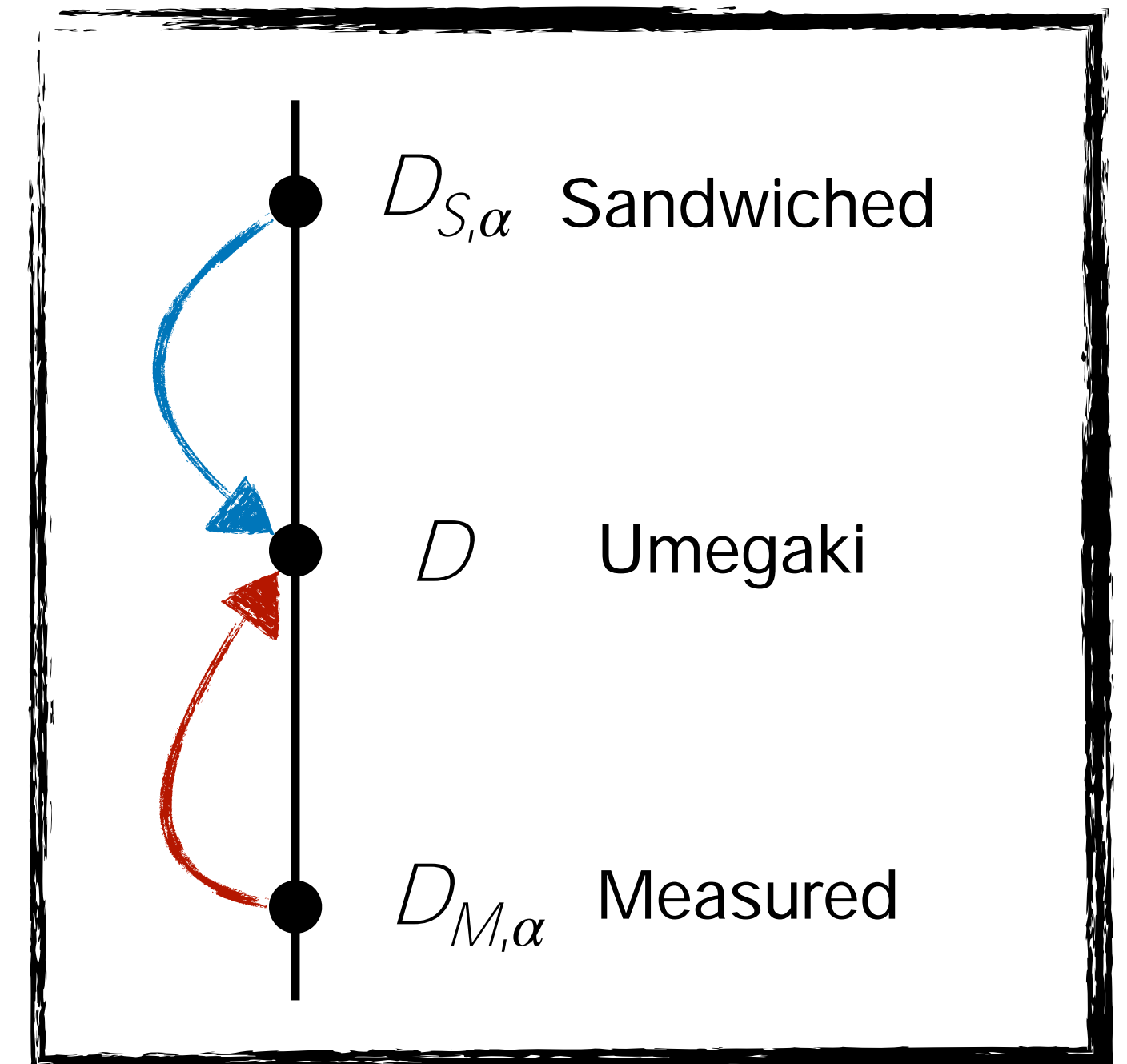
Suppose $\mathcal{A}_1 \otimes \mathcal{A}_2 \subseteq \mathcal{A}_{12}$ and $\mathcal{B}_1 \otimes \mathcal{B}_2 \subseteq \mathcal{B}_{12}$

$$D_{S,\alpha}(\mathcal{A}_{12}\|\mathcal{B}_{12}) \leq D_{S,\alpha}(\mathcal{A}_1\|\mathcal{B}_1) + D_{S,\alpha}(\mathcal{A}_2\|\mathcal{B}_2) \quad \forall \alpha > 1$$

Superadditivity

Suppose $(\mathcal{A}_1)_+^\circ \otimes (\mathcal{A}_2)_+^\circ \subseteq (\mathcal{A}_{12})_+^\circ$ and $(\mathcal{B}_1)_+^\circ \otimes (\mathcal{B}_2)_+^\circ \subseteq (\mathcal{B}_{12})_+^\circ$

$$D_{M,\alpha}(\mathcal{A}_{12}\|\mathcal{B}_{12}) \geq D_{M,\alpha}(\mathcal{A}_1\|\mathcal{B}_1) + D_{M,\alpha}(\mathcal{A}_2\|\mathcal{B}_2) \quad \forall 0 < \alpha < 1$$



A few more words on the polar assumption

Variational expression [Berta-Fawzi-Tomamichel-17]: $D_M(\rho\|\sigma) = \sup_{\omega \in \mathcal{H}_{++}} \text{Tr}[\rho \log \omega] + 1 - \text{Tr}[\sigma \omega]$

$$D_M(\mathcal{A}\|\mathcal{B}) = \sup_{W \in \mathcal{B}_{++}^\circ} -h_{\mathcal{A}}(-\log W) \quad \text{Support function} \quad h_{\mathcal{A}}(X) := \sup_{Y \in \mathcal{A}} \text{Tr}[XY]$$

Superadditivity

Suppose $(\mathcal{A}_1)_+^\circ \otimes (\mathcal{A}_2)_+^\circ \subseteq (\mathcal{A}_{12})_+^\circ$ and $(\mathcal{B}_1)_+^\circ \otimes (\mathcal{B}_2)_+^\circ \subseteq (\mathcal{B}_{12})_+^\circ$

$$D_{M,\alpha}(\mathcal{A}_{12}\|\mathcal{B}_{12}) \geq D_{M,\alpha}(\mathcal{A}_1\|\mathcal{B}_1) + D_{M,\alpha}(\mathcal{A}_2\|\mathcal{B}_2) \quad \forall 0 < \alpha < 1$$

$$(\mathcal{A}_1)_+^\circ \otimes (\mathcal{A}_2)_+^\circ \subseteq (\mathcal{A}_{12})_+^\circ \iff h_{\mathcal{A}_{12}}(X_1 \otimes X_2) \leq h_{\mathcal{A}_1}(X_1) \cdot h_{\mathcal{A}_2}(X_2), \forall X_i \in \mathcal{H}_+.$$

$$\iff \beta(\mathcal{A}_{12}, X_1 \otimes X_2) \leq \beta(\mathcal{A}_1, X_1) \cdot \beta(\mathcal{A}_2, X_2), \forall 0 \leq X_i \leq I.$$

Worst-case Type-II error prob. (@ Milan Mosonyi)

Recap: from AEP to generalized quantum AEP

$$-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) \longrightarrow H(X) \quad \text{in probability}$$

AEP

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(P^{\otimes n} \| Q^{\otimes n}) = D(P \| Q)$$

Quantum

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho \| \sigma)$$

Generalized

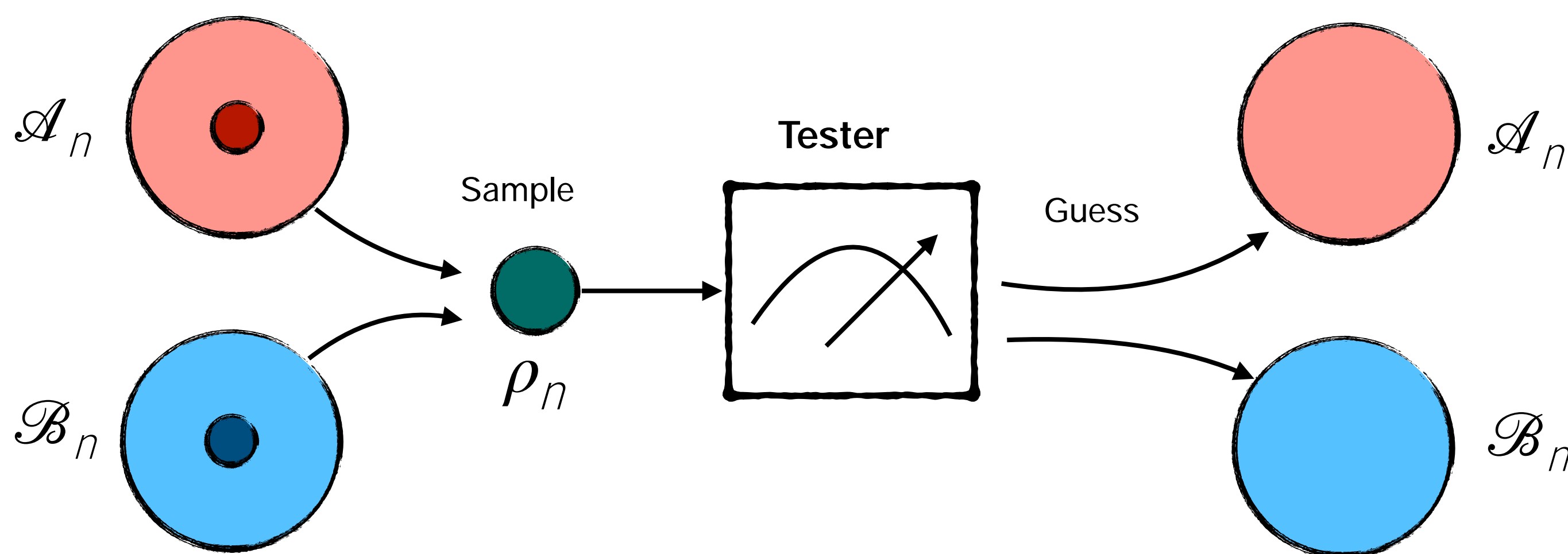
$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_{\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = D^{\infty}(\mathcal{A} \| \mathcal{B})$$

Applications

1. Quantum hypothesis testing between two sets of states
2. Adversarial quantum channel discrimination
3. A relative entropy accumulation theorem
4. Efficient bounds for quantum resource theory

Application 1: Quantum hypothesis testing between two sets of states

A tester draws samples from *two sets of quantum states*, and performs measurements to determine which set the sample belongs to.



Type-I error

$$\alpha(\mathcal{A}_n, M_n) := \sup_{\rho_n \in \mathcal{A}_n} \text{Tr} [\rho_n (I - M_n)]$$

Type-II error

$$\beta(\mathcal{B}_n, M_n) := \sup_{\sigma_n \in \mathcal{B}_n} \text{Tr} [\sigma_n M_n]$$

Worst-case

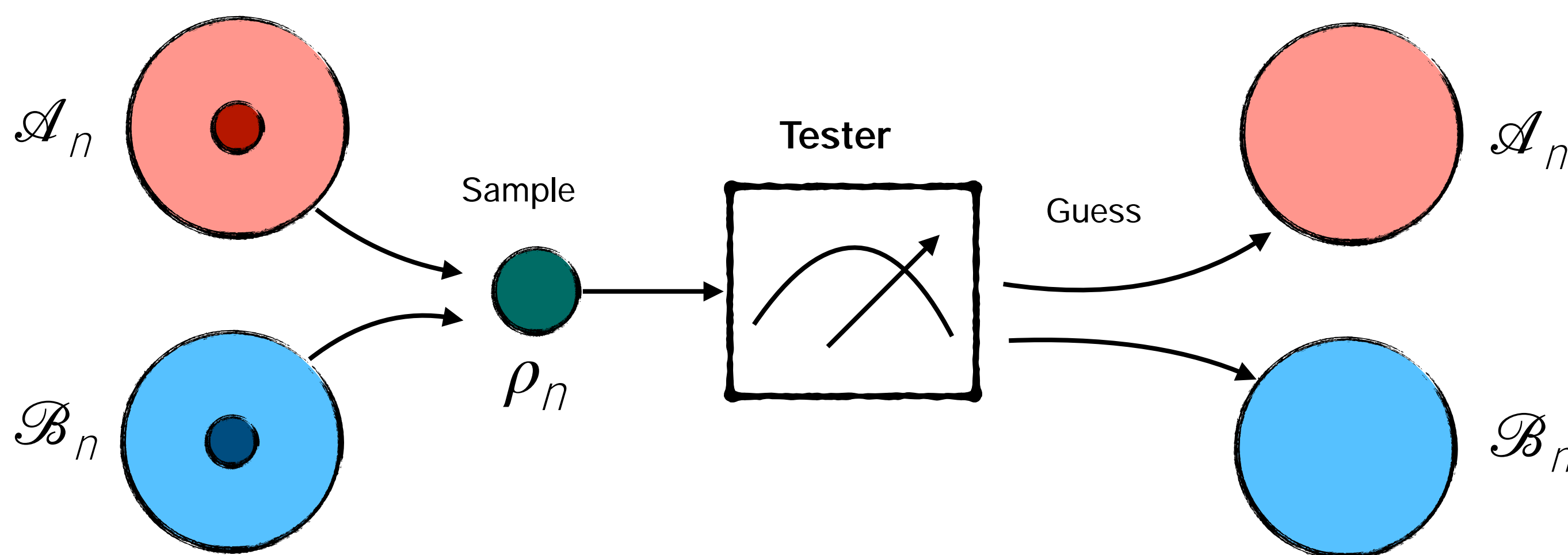
As in standard hypothesis testing, the tester will make two types of errors:

Type-I error: sample from \mathcal{A}_n , but classified as from \mathcal{B}_n .

Type-II error: sample from \mathcal{B}_n , but classified as from \mathcal{A}_n .

Application 1: Quantum hypothesis testing between two sets of states

A tester draws samples from *two sets of quantum states*, and performs measurements to determine which set the sample belongs to.



Type-I error

$$\alpha(\mathcal{A}_n, M_n) := \sup_{\rho_n \in \mathcal{A}_n} \text{Tr} [\rho_n (I - M_n)]$$

Type-II error

$$\beta(\mathcal{B}_n, M_n) := \sup_{\sigma_n \in \mathcal{B}_n} \text{Tr} [\sigma_n M_n]$$

Worst-case

Goal: Determine the optimal exponent at which the type-II error probability decays, while keeping the type-I error within a fixed threshold ε (to control over false positives)

e.g. COVID-19: healthy people get a positive test

$$\beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) := \inf_{0 \leq M_n \leq I} \{ \beta(\mathcal{B}_n, M_n) : \alpha(\mathcal{A}_n, M_n) \leq \varepsilon \}$$

$$\beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) \approx ?$$

Application 1: Quantum hypothesis testing between two sets of states

Our answer (new generalized quantum Stein's lemma)

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B}) \quad \forall \varepsilon \in (0, 1)$$

✓ Classical Chernoff-Stein Lemma

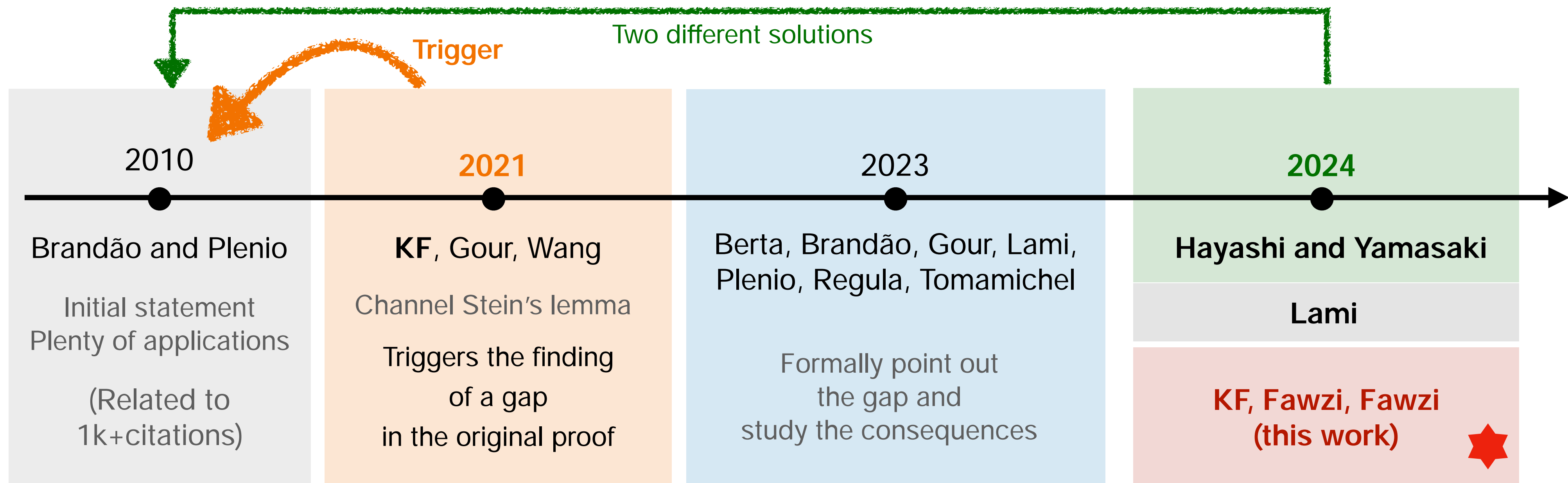
✓ Quantum Stein's Lemma [Hiai, Petz 1991; Ogawa, Nagaoka 2000]

Let $\mathcal{A}_n = \{\rho^{\otimes n}\}$ and $\mathcal{B}_n = \{\sigma^{\otimes n}\}$ be two singletons.

Generalized Quantum Stein's Lemma ($\mathcal{A}_n = \{\rho^{\otimes n}\}$)

Talk by Yamasaki and Talk by Lami on Monday

Story: Generalized Quantum Stein's Lemma ($\mathcal{A}_n = \{\rho^{\otimes n}\}$)



A.1		A.3
A.1	A.2	A.3
A.1	A.2	A.3

However, an issue has recently been found in the claimed proof of the generalised quantum Stein's lemma in [BP10a]. Specifically, after the appearance of the first version of the preprint [FGW21] that studied a related setting using the methods of [BP10a], one of us identified a mistake in [FGW21, Lemma 16], which then led to the discovery that the original result [BP10a, Lemma III.9] is incorrect. This means that the main claims of [BP10a], and in particular the generalised quantum Stein's lemma introduced therein, are not known to be correct, and the validity of a number of results that build on those findings is thus directly put into question.

Story:

Generalized Quantum Stein's Lemma ($\mathcal{A}_n = \{\rho^{\otimes n}\}$)

2024

Hayashi and Yamasaki

Lami

KF, Fawzi, Fawzi
(this work) ★

A.1		A.3		A.5	
A.1	A.2	A.3		A.5	A.6
A.1	A.2	A.3	A.4		

(A.1) Each \mathcal{A}_n is convex and compact;

(A.2) Each \mathcal{A}_n is permutation-invariant;

(A.3) $\mathcal{A}_m \otimes \mathcal{A}_k \subseteq \mathcal{A}_{m+k}$, for all $m, k \in \mathbb{N}$;

(A.5) \mathcal{A}_1 contains a full-rank state

(A.6) Each \mathcal{A}_n is closed under partial traces

(A.4) $(\mathcal{A}_m)_+^\circ \otimes (\mathcal{A}_k)_+^\circ \subseteq (\mathcal{A}_{m+k})_+^\circ$, for all $m, k \in \mathbb{N}$;

GQSL by BP

$$\lim_{n \rightarrow \infty} \frac{1}{n} D_{H,\varepsilon}(\rho^{\otimes n} \| \mathcal{B}_n) = D^\infty(\rho \| \mathcal{B})$$

New GQSL

$$\lim_{n \rightarrow \infty} \frac{1}{n} D_{H,\varepsilon}(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B})$$

Proof of the generalised quantum Stein's lemma of the first version of the paper [BP10a], one of us identified a gap in the original result [BP10a], and in particular the generalisation to be correct, and the validity of the result is into question.

Story: Generalized Quantum Stein's Lemma ($\mathcal{A}_n = \{\rho^{\otimes n}\}$)

2024

Hayashi and Yamasaki

Lami

KF, Fawzi, Fawzi
(this work)



A.1		A.3		A.5	
A.1	A.2	A.3		A.5	A.6
A.1	A.2	A.3	A.4		

(A.1) Each \mathcal{A}_n is convex and compact;

(A.2) Each \mathcal{A}_n is permutation-invariant;

(A.3) $\mathcal{A}_m \otimes \mathcal{A}_k \subseteq \mathcal{A}_{m+k}$, for all $m, k \in \mathbb{N}$;

(A.5) \mathcal{A}_1 contains a full-rank state

(A.6) Each \mathcal{A}_n is closed under partial traces

(A.4) $(\mathcal{A}_m)_+^\circ \otimes (\mathcal{A}_k)_+^\circ \subseteq (\mathcal{A}_{m+k})_+^\circ$, for all $m, k \in \mathbb{N}$;

Our result is incomparable to the previous generalized quantum Stein lemma.

Weaker: assume (A.4) for \mathcal{B}_n (not satisfied by SEP)

Stronger: 1. composite null hypothesis \mathcal{A}_n instead of $\rho^{\otimes n}$

2. efficient and controlled approximations of the Stein's exponent $D^\infty(\mathcal{A}||\mathcal{B})$

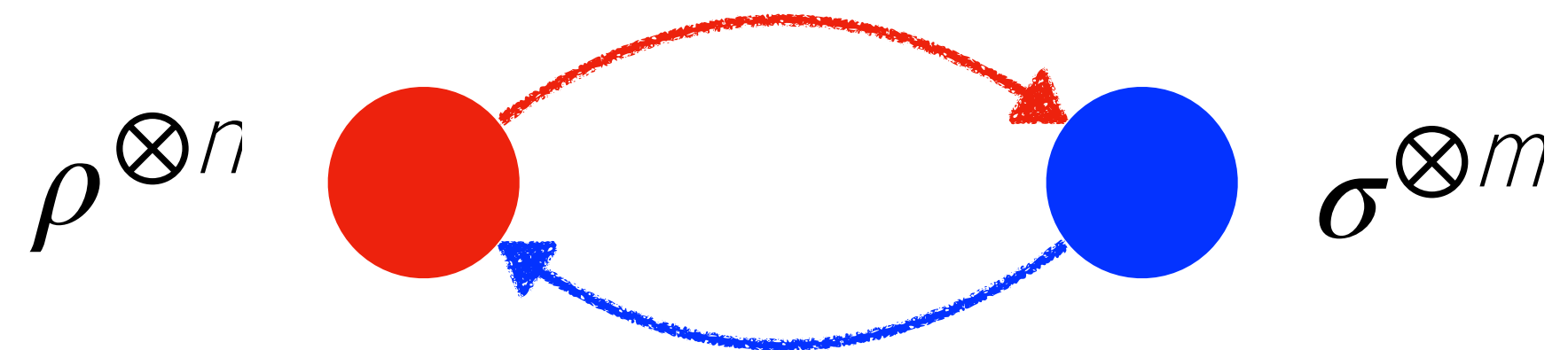


solves open problems

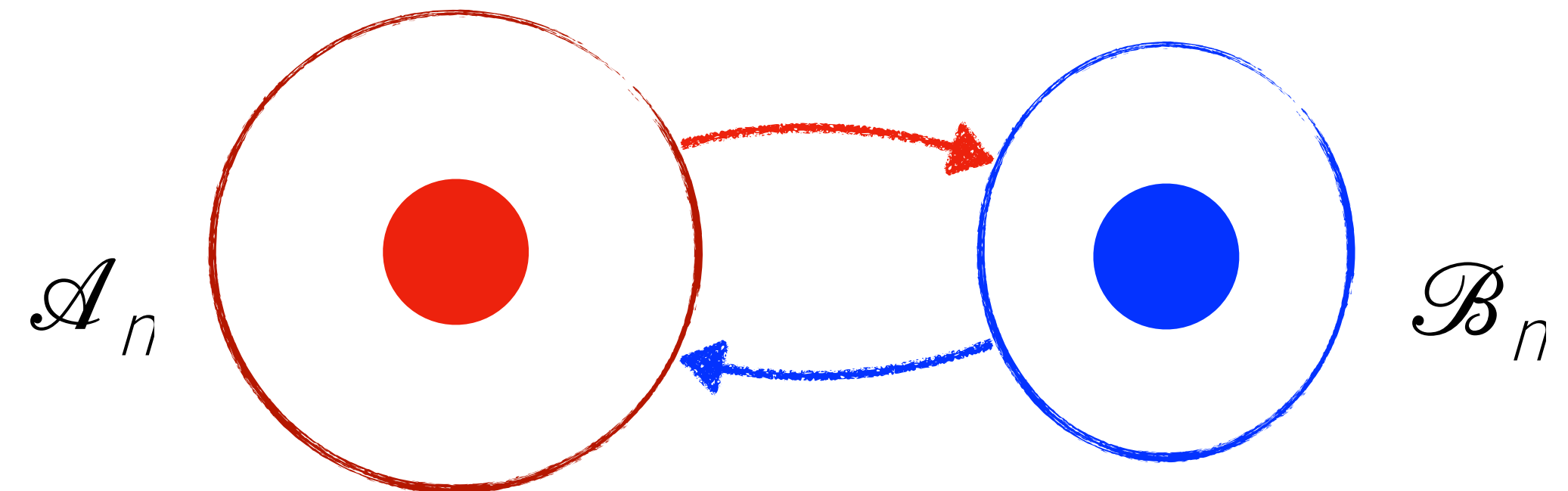
in [Brandão, Harrow, Lee, Peres, 2020, TIT] and [Mosonyi, Szilagyi, Weiner, 2022, TIT]

Application 1': Quantum resource theory and its reversibility

a.k.a, second law



Standard resource manipulation



Resource manipulation with partial information

Asymptotic resource nongenerating operations
[Brandão and Plenio, 2010]

Lack of knowledge of the states
Different copies of the sources
can exhibit correlation in nature

Talk by Watanabe on Monday

Our answer

Optimal transformation rate

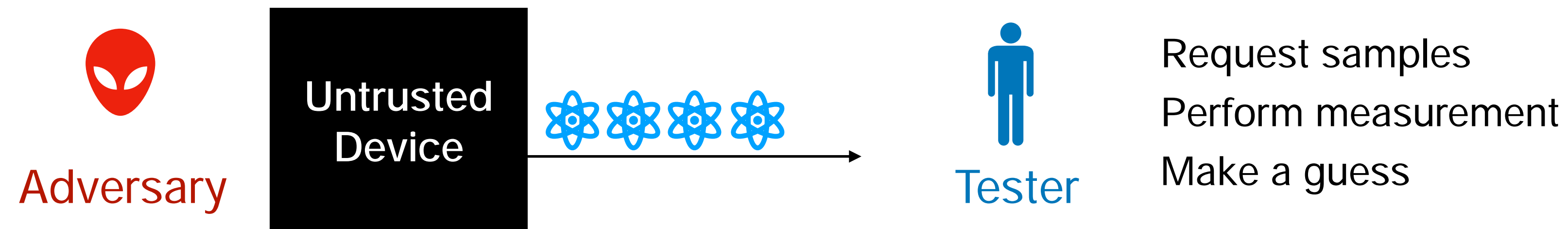
$$r(\mathcal{A} \xrightarrow{\text{ARNG}} \mathcal{B}) = \frac{D^\infty(\mathcal{A} \parallel \mathcal{F})}{D^\infty(\mathcal{B} \parallel \mathcal{F})} \quad \mathcal{F} \text{ is the set of free states}$$

Application 2: adversarial quantum channel discrimination

Operational setting:

A tester is working with an **untrusted** quantum device that generates a quantum state upon request

Guarantee: either \mathcal{N} (the bad case) or \mathcal{M} (the good case)

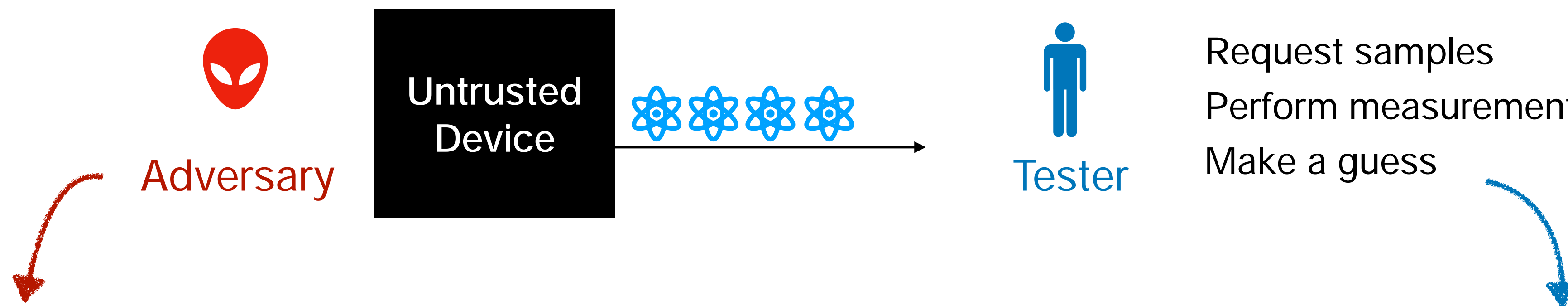


Application 2: adversarial quantum channel discrimination

Operational setting:

A tester is working with an **untrusted** quantum device that generates a quantum state upon request

Guarantee: either \mathcal{N} (the bad case) or \mathcal{M} (the good case)



Environmental system of the channel

Internal memory correlates with the generated samples

Actively misleading the tester to correctly identify the channel

How effectively can the tester distinguish between the two cases while playing against the adversary?

Classical setting refers to [Brandão, Harrow, Lee, Peres, 2020, TIT]

Application 2: adversarial quantum channel discrimination

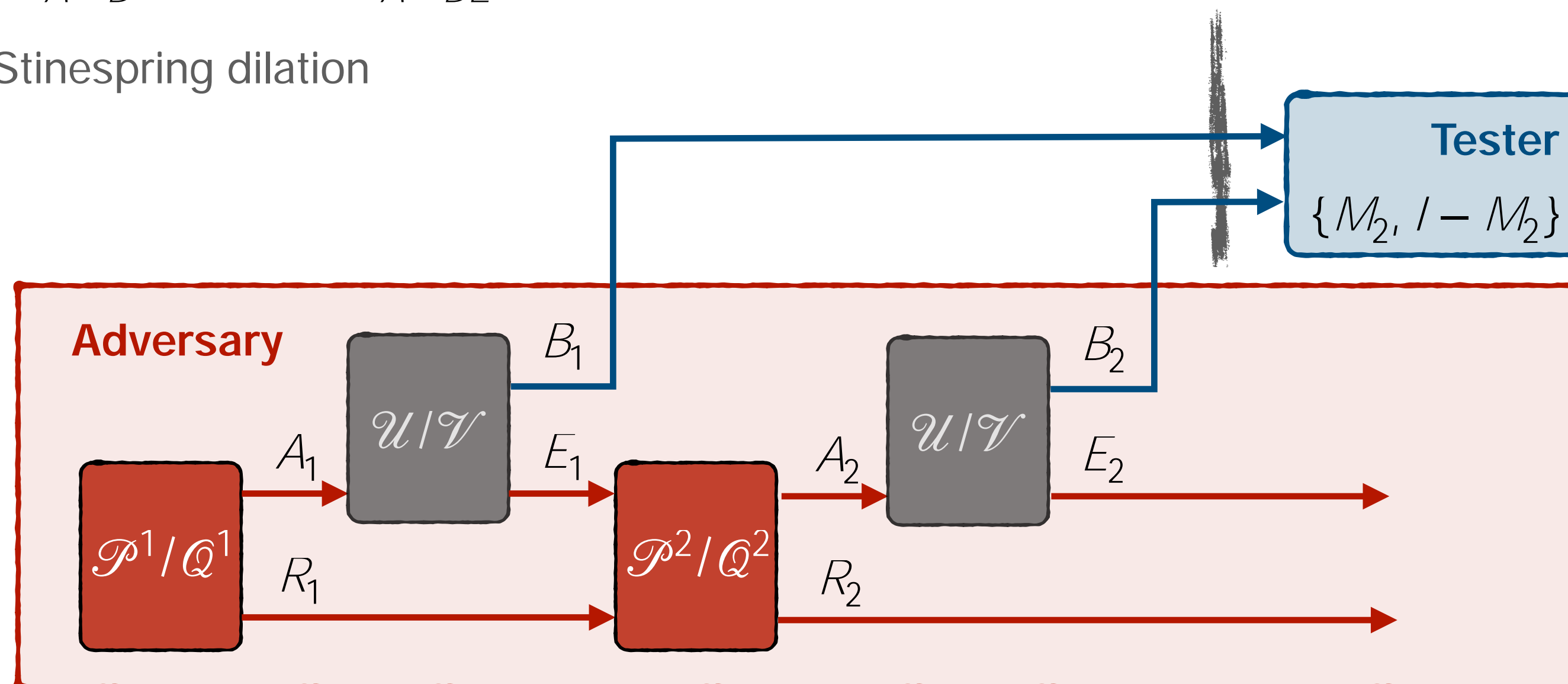
Operational setting:

A tester is working with an **untrusted** quantum device that generates a quantum state upon request

$$\mathcal{N}_{A \rightarrow B} = \text{Tr}_E \circ \mathcal{U}_{A \rightarrow BE}$$

$$\mathcal{M}_{A \rightarrow B} = \text{Tr}_E \circ \mathcal{V}_{A \rightarrow BE}$$

Stinespring dilation



Due to the lack of knowledge of what the adversary do:

- \mathcal{A}_r if device is \mathcal{N} ;
- \mathcal{B}_r if device is \mathcal{M}

Adaptive strategies by adversary

E_i environmental systems, R_i internal memories, $\mathcal{P}_i/\mathcal{Q}_i$ internal operations by adversary

Application 2: adversarial quantum channel discrimination

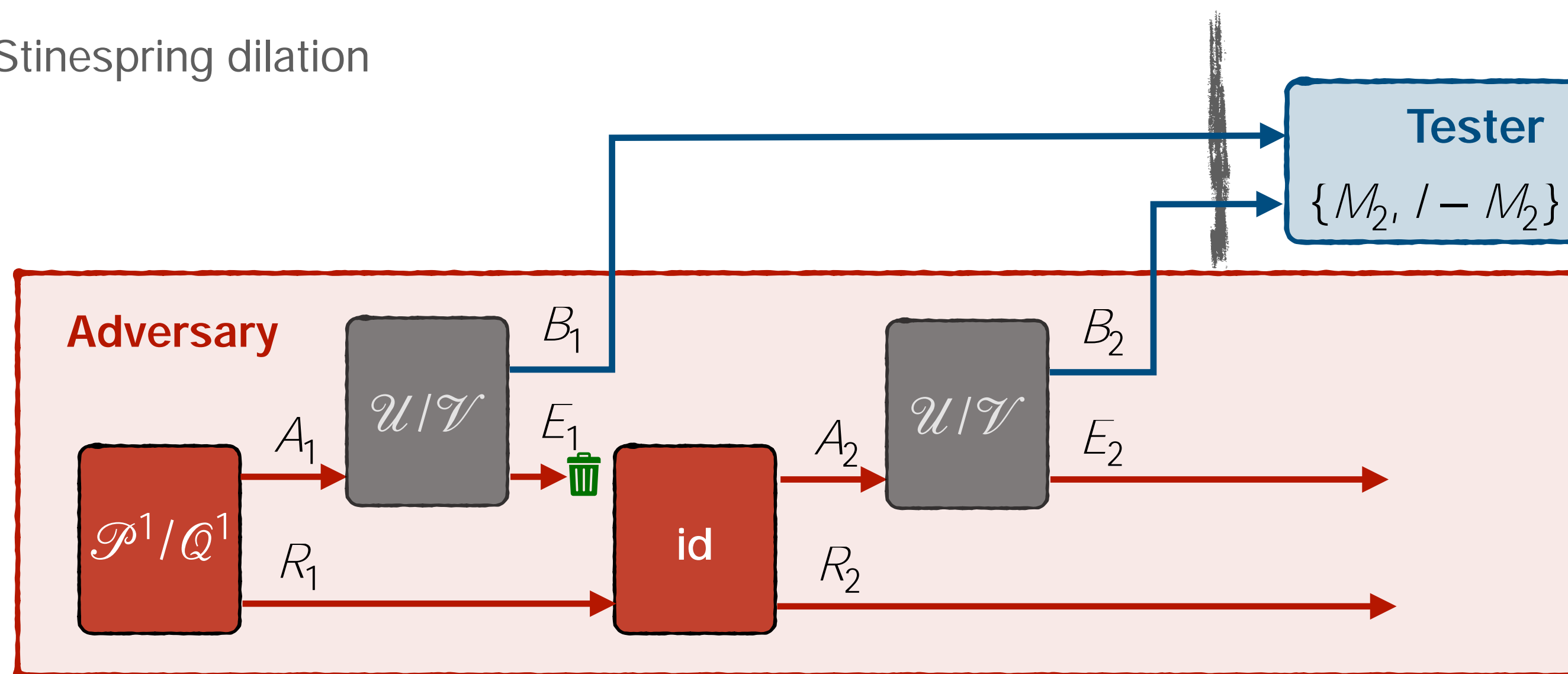
Operational setting:

A tester is working with an **untrusted** quantum device that generates a quantum state upon request

$$\mathcal{N}_{A \rightarrow B} = \text{Tr}_E \circ \mathcal{U}_{A \rightarrow BE}$$

$$\mathcal{M}_{A \rightarrow B} = \text{Tr}_E \circ \mathcal{V}_{A \rightarrow BE}$$

Stinespring dilation



Due to the lack of knowledge of what the adversary do:

- \mathcal{A}'_n if device is \mathcal{N} ;
- \mathcal{B}'_n if device is \mathcal{M}

Non-adaptive strategies by adversary

E_i environmental systems, R_i internal memories, P_i/Q_i internal operations by adversary

Application 2: adversarial quantum channel discrimination

The best performance of the tester playing against the adversary is given by:

Our answer

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = \lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\mathcal{A}'_n \| \mathcal{B}'_n) = D^{\text{inf}, \infty}(\mathcal{N} \| \mathcal{M})$$

Adaptive strategies
by adversary

Non-adaptive strategies
by adversary

Minimum output
quantum channel divergence

$$D^{\text{inf}}(\mathcal{N} \| \mathcal{M}) := \inf_{\rho, \sigma \in \mathcal{D}} D(\mathcal{N}(\rho) \| \mathcal{M}(\sigma)) \quad D^{\text{inf}, \infty}(\mathcal{N} \| \mathcal{M}) := \lim_{n \rightarrow \infty} \frac{1}{n} D^{\text{inf}}(\mathcal{N}^{\otimes n} \| \mathcal{M}^{\otimes n})$$

Adaptive strategies offer **no advantage** over non-adaptive ones
in adversarial quantum channel discrimination.

Good news for the tester!

Application 2: adversarial quantum channel discrimination

The best performance of the tester playing against the adversary is given by:

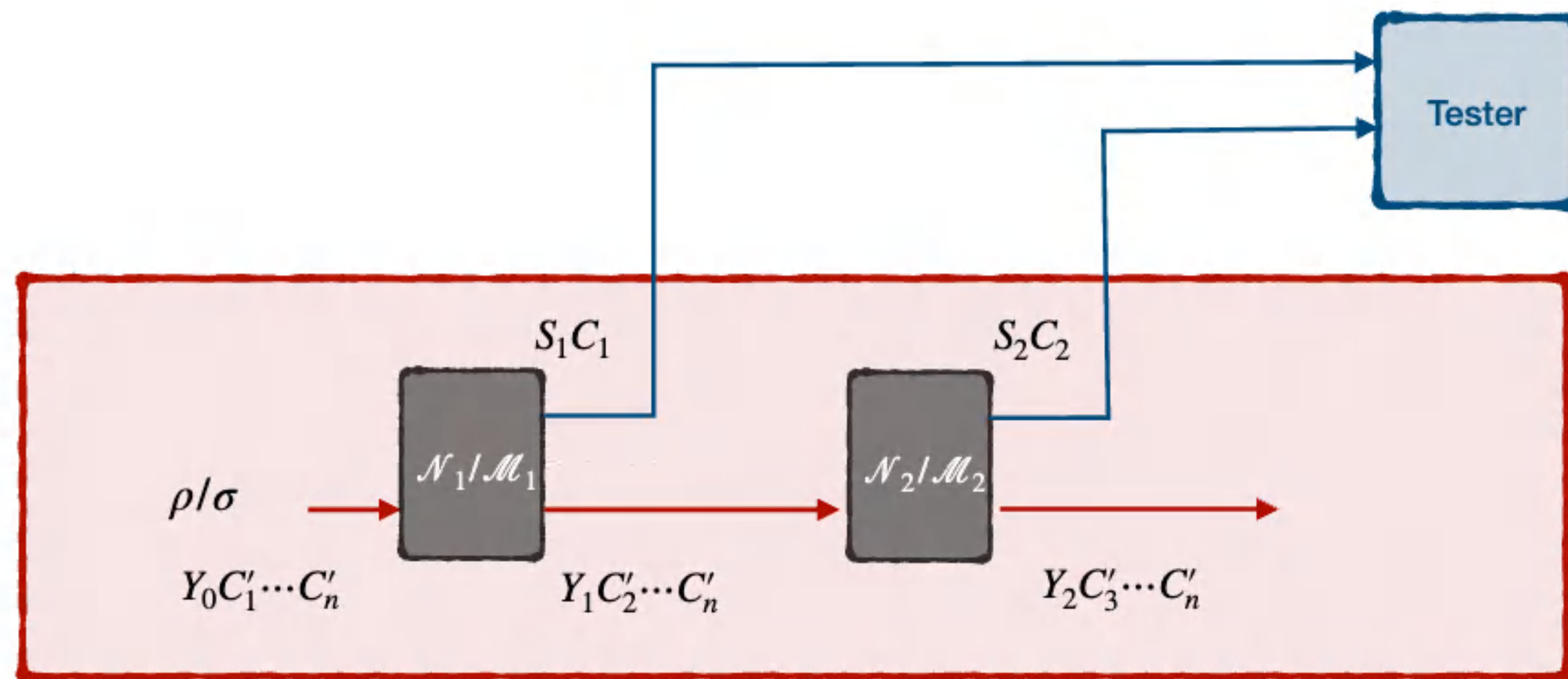
$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = \lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\mathcal{A}'_n \| \mathcal{B}'_n) = D^{\text{inf}, \infty}(\mathcal{N} \| \mathcal{M})$$

Key technical tool (chain rule):

$$D_{M, \alpha}(\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \| \mathcal{M}_{A \rightarrow B}(\sigma_{RA})) \geq D_{M, \alpha}(\rho_R \| \sigma_R) + D_{M, \alpha}^{\text{inf}}(\mathcal{N}_{A \rightarrow B} \| \mathcal{M}_{A \rightarrow B}), \quad \forall \alpha \in [0, +\infty]$$

$$D_{S, \alpha}(\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \| \mathcal{M}_{A \rightarrow B}(\sigma_{RA})) \geq D_{S, \alpha}(\rho_R \| \sigma_R) + D_{S, \alpha}^{\text{inf}, \infty}(\mathcal{N}_{A \rightarrow B} \| \mathcal{M}_{A \rightarrow B}) \quad \forall \alpha \in [1/2, +\infty]$$

Application 3: a relative entropy accumulation theorem



How entropy accumulate for sequential operations on a state?

[Dupuis, Fawzi, Renner, 2020, CMP] Find plenty of applications in quantum cryptography

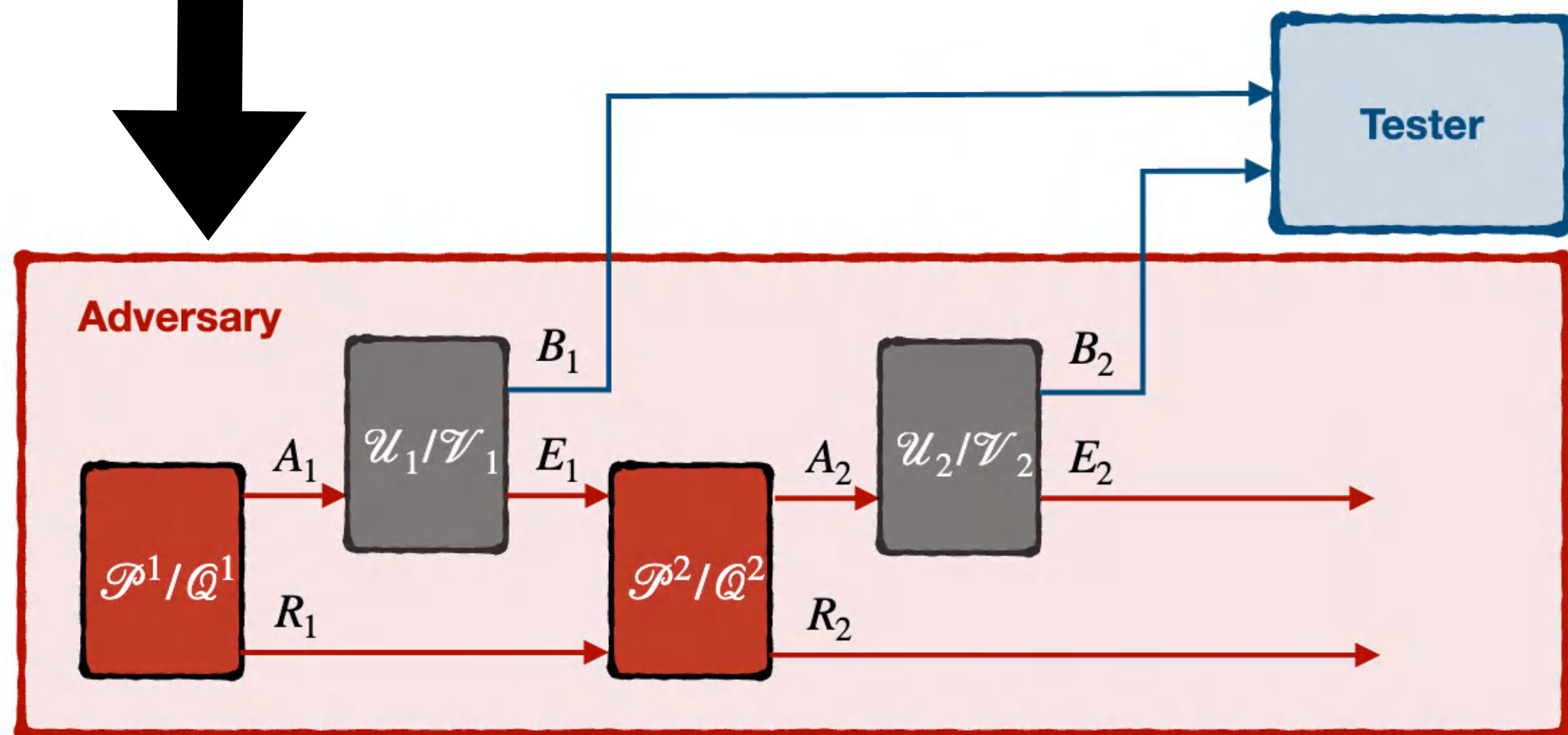
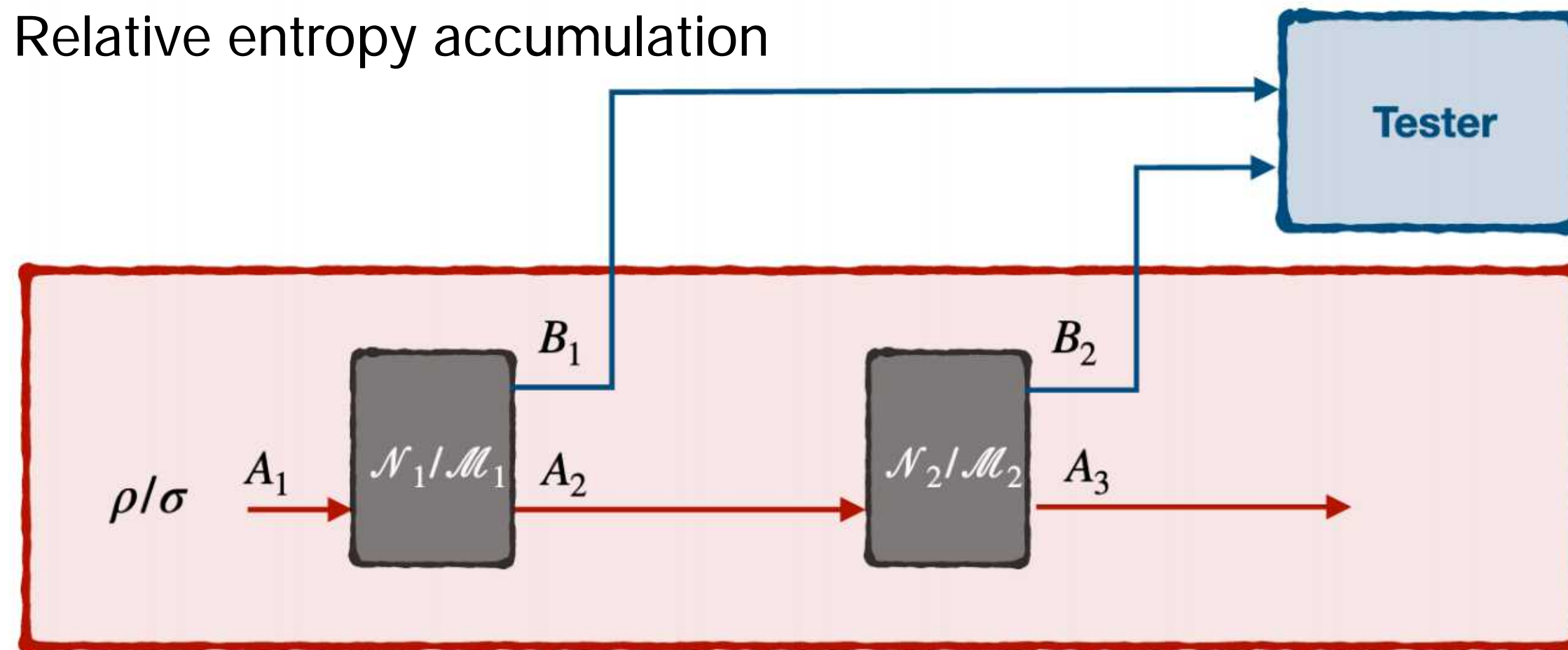
$$H_{\max}^{\varepsilon}(S_1 \dots S_n | C_1 \dots C_n)_{\mathcal{N}_n \circ \dots \circ \mathcal{N}_1(\rho_{R_0})} \leq \sum_{i=1}^n \sup_{\omega_{R_{i-1}}} H(S_i | C_i)_{\mathcal{N}_i(\omega)} + O(\sqrt{n})$$

How to generalize from conditional entropy to relative entropy?

Open question in [Metger, Fawzi, Sutter, Renner, 2022, FOCS] for $D_{\max, \varepsilon}$

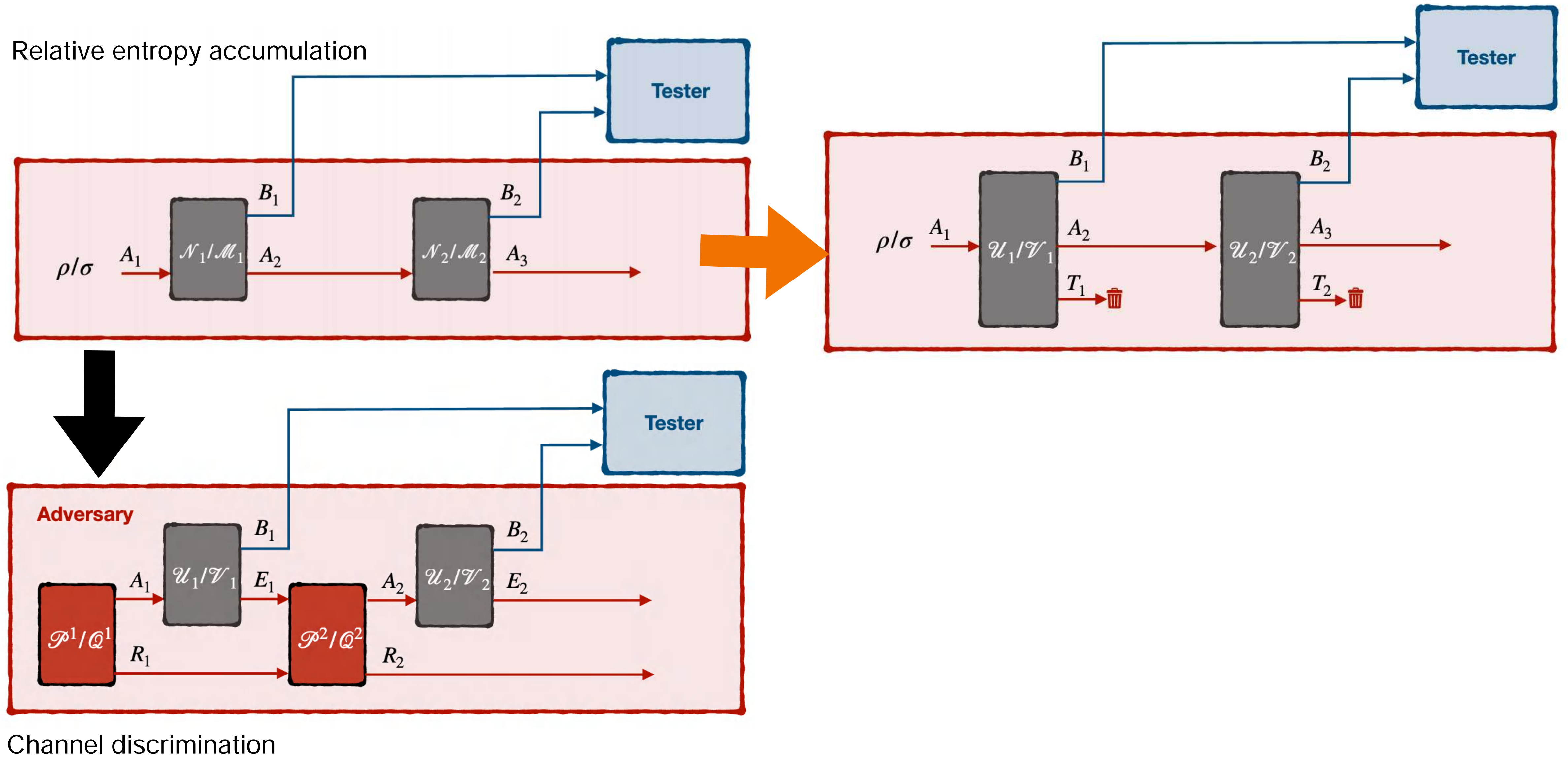
Application 3: a relative entropy accumulation theorem

Relative entropy accumulation

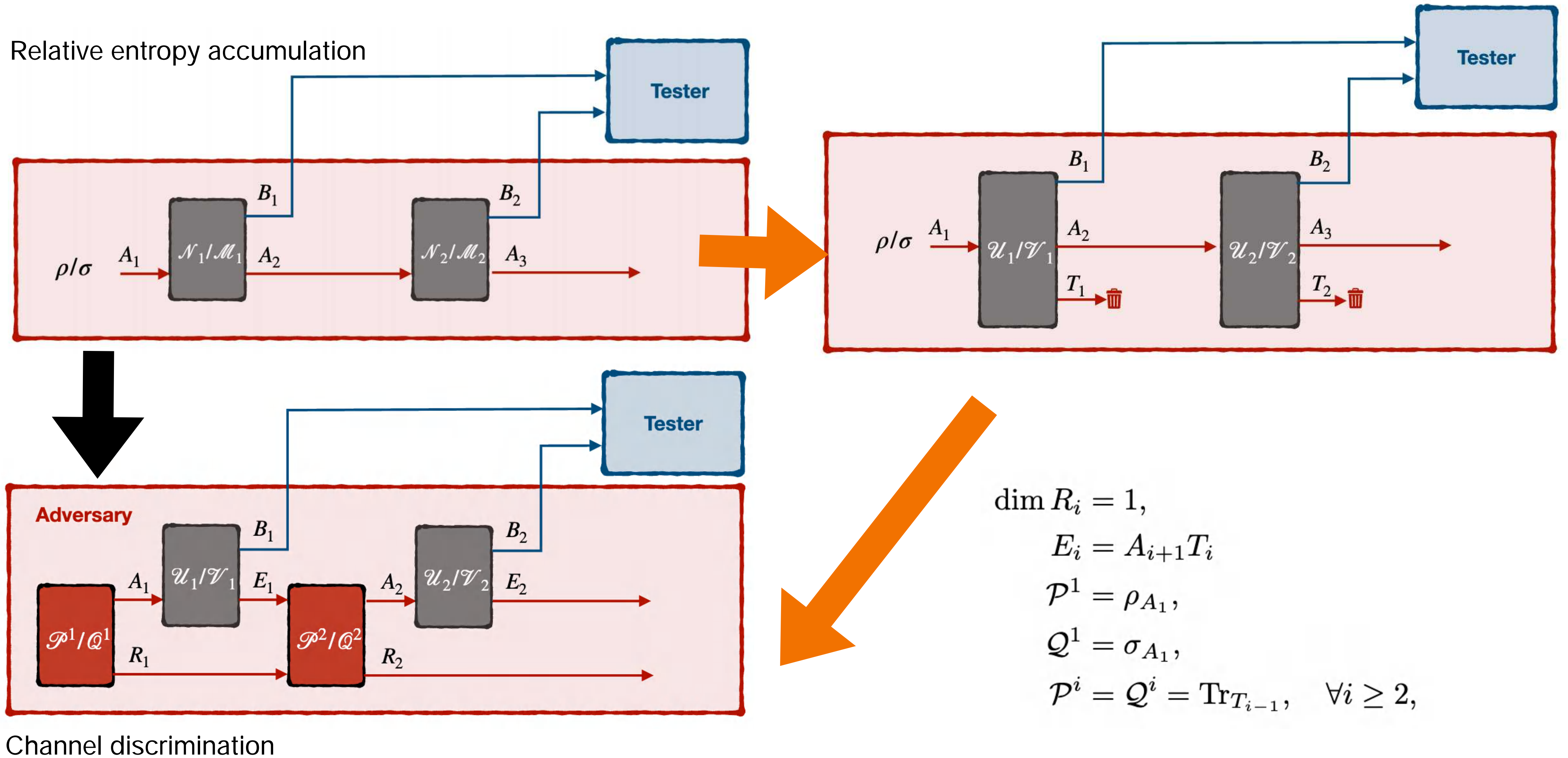


Channel discrimination

Application 3: a relative entropy accumulation theorem



Application 3: a relative entropy accumulation theorem



Application 3: a relative entropy accumulation theorem

How entropy accumulate for sequential operations on a state?


[Dupuis, Fawzi, Renner, 2020, CMP] Find plenty of applications in quantum cryptography

$$H_{\max}^{\varepsilon}(S_1 \dots S_n | C_1 \dots C_n)_{\mathcal{N}_n \circ \dots \circ \mathcal{N}_1(\rho_{R_0})} \leq \sum_{i=1}^n \sup_{\omega_{R_{i-1}}} H(S_i | C_i)_{\mathcal{N}_i(\omega)} + O(\sqrt{n})$$

How to generalize from conditional entropy to relative entropy?

Open question in [Metger, Fawzi, Sutter, Renner, 2022, FOCS] for $D_{\max, \varepsilon}$

Recover with a
slightly weaker
second order



Our answer

$$D_{H, \varepsilon} \left(\text{Tr}_{R_n} \circ \prod_{i=1}^n \mathcal{N}_i(\rho_{R_0}) \parallel \text{Tr}_{R_n} \circ \prod_{i=1}^n \mathcal{M}_i(\sigma_{R_0}) \right) \geq \sum_{i=1}^n D^{\text{inf}, \infty}(\text{Tr}_{R_i} \circ \mathcal{N}_i \parallel \text{Tr}_{R_i} \circ \mathcal{M}_i) - O(n^{2/3} \log n)$$

Application 4: efficient bounds for quantum resource theory

(A.1) Each \mathcal{A}_n is convex and compact;

(A.2) Each \mathcal{A}_n is permutation-invariant;

(A.3) $\mathcal{A}_m \otimes \mathcal{A}_k \subseteq \mathcal{A}_{m+k}$, for all $m, k \in \mathbb{N}$;

(A.4) $(\mathcal{A}_m)_+^\circ \otimes (\mathcal{A}_k)_+^\circ \subseteq (\mathcal{A}_{m+k})_+^\circ$, for all $m, k \in \mathbb{N}$;

If (A.4) is not directly satisfied, we do relaxation!!!

Note that $D^\infty(\mathcal{A} \parallel \mathcal{B}) := \lim_{n \rightarrow \infty} \frac{1}{n} D(\mathcal{A}_n \parallel \mathcal{B}_n)$ is efficiently computable

Improvement (even for the first level of approximation)

- Entanglement cost of quantum states and channels
- Entanglement distillation
- Magic state distillation

Refer to arXiv: 2502.15659 for more details

Application 4: efficient bounds for quantum resource theory

Entanglement cost for quantum states and channels

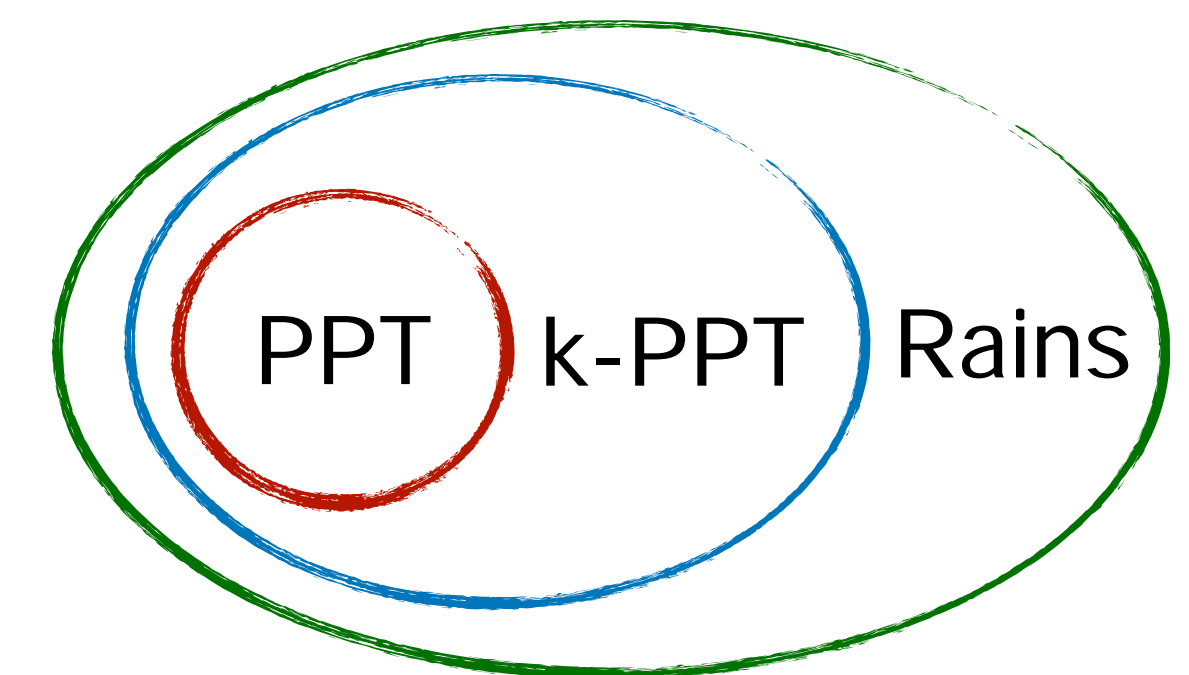
Using the minimum number of Bell states to prepare one copy of a state under LOCC operations

$$E_{C,\text{LOCC}}(\rho) \geq D^\infty(\rho \| \text{PPT}) := \lim_{n \rightarrow \infty} \frac{1}{n} D(\rho^{\otimes n} \| \text{PPT}(A^n : B^n))$$

Hard to evaluate in general

SDP lower bounds [Wang, Duan, 2017, PRL; Wang, Duan, 2017, PRA; Wang, Jing, Zhu, 2023]

$$E_{C,\text{LOCC}}(\rho) \geq D^\infty(\rho \| \text{PPT}) \geq \max \{ E_{\text{WD},1}(\rho), E_{\text{WD},2}(\rho), E_{\text{WJZ}}(\rho) \}$$



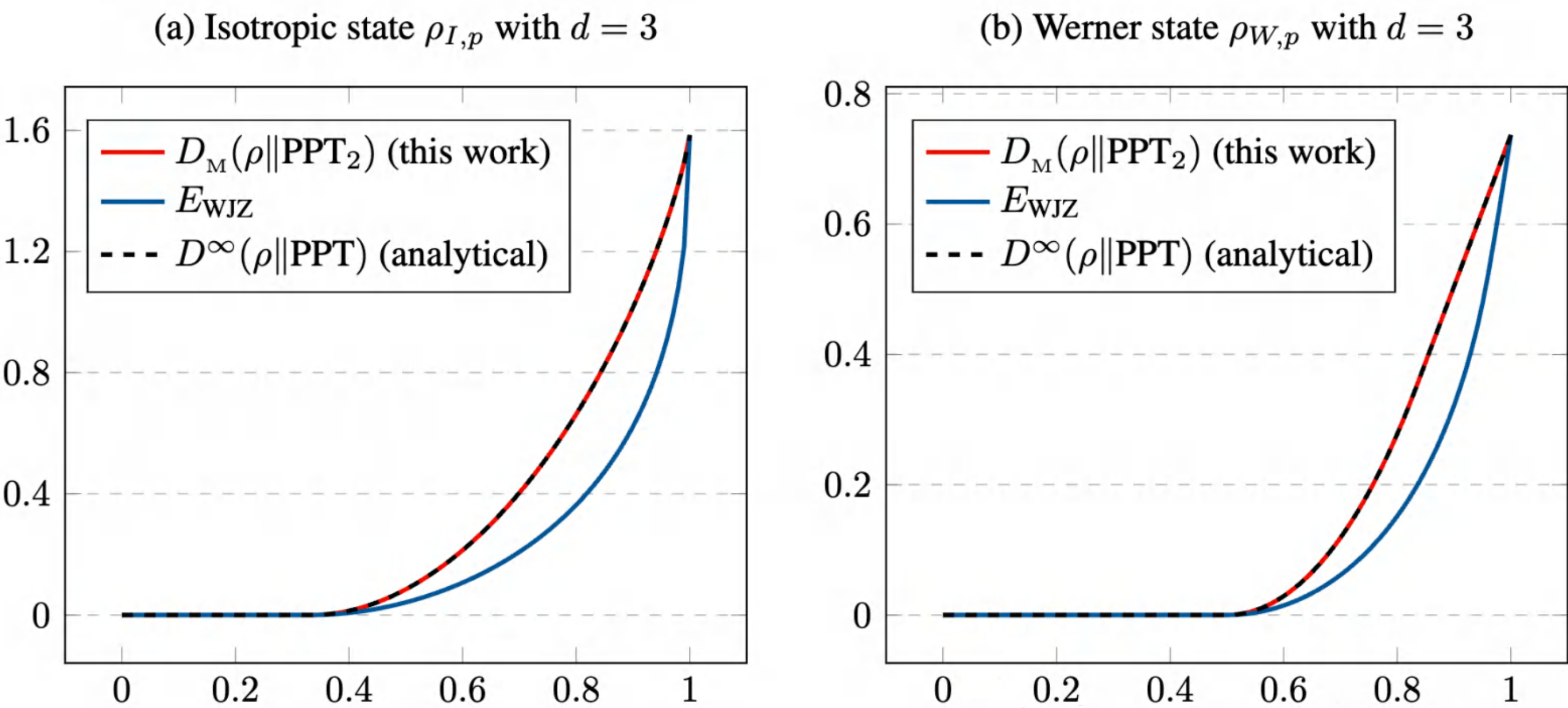
$$E_{C,\text{LOCC}}(\rho) \geq D^\infty(\rho \| \text{PPT}) \geq D^\infty(\rho \| \text{PPT}_k) \geq D_M(\rho \| \text{PPT}_k) \geq (*)$$

Improved bounds and still efficiently computable via convex programs

*Similar result holds for entanglement cost of quantum channels

Application 4: efficient bounds for quantum resource theory

Entanglement cost for quantum states and channels

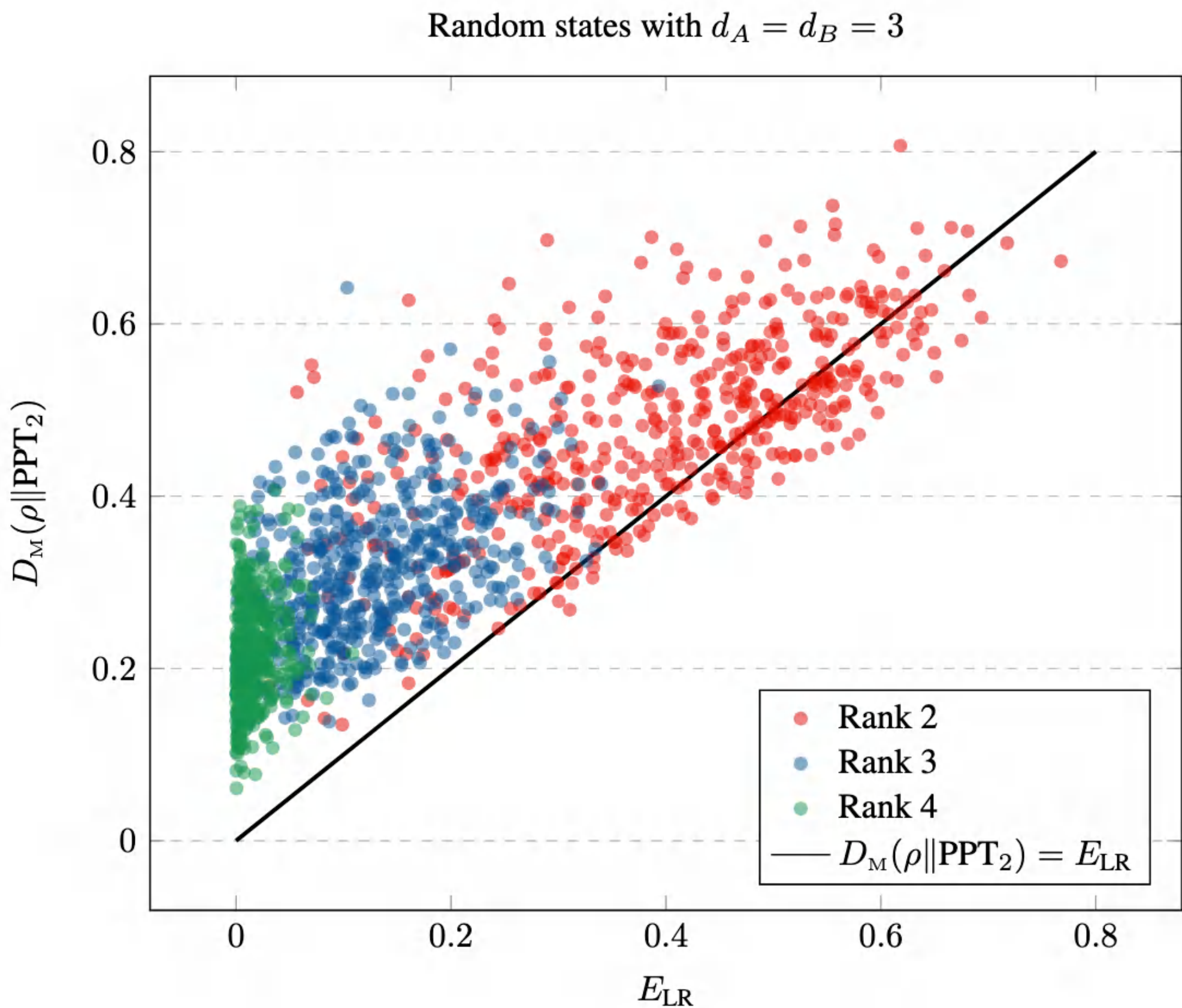


$$E_{C,LOCC}(\rho) \geq E_{LR}(\rho) \quad \text{SDP lower bound by [Lami, Regula, 2023, NP]}$$

Quantitative improvement, even at the first level

Match the analytical result for isotropic and Werner states

Outperform [Lami, Regula, 2023, NP] in most random cases



Application 4: efficient bounds for quantum resource theory

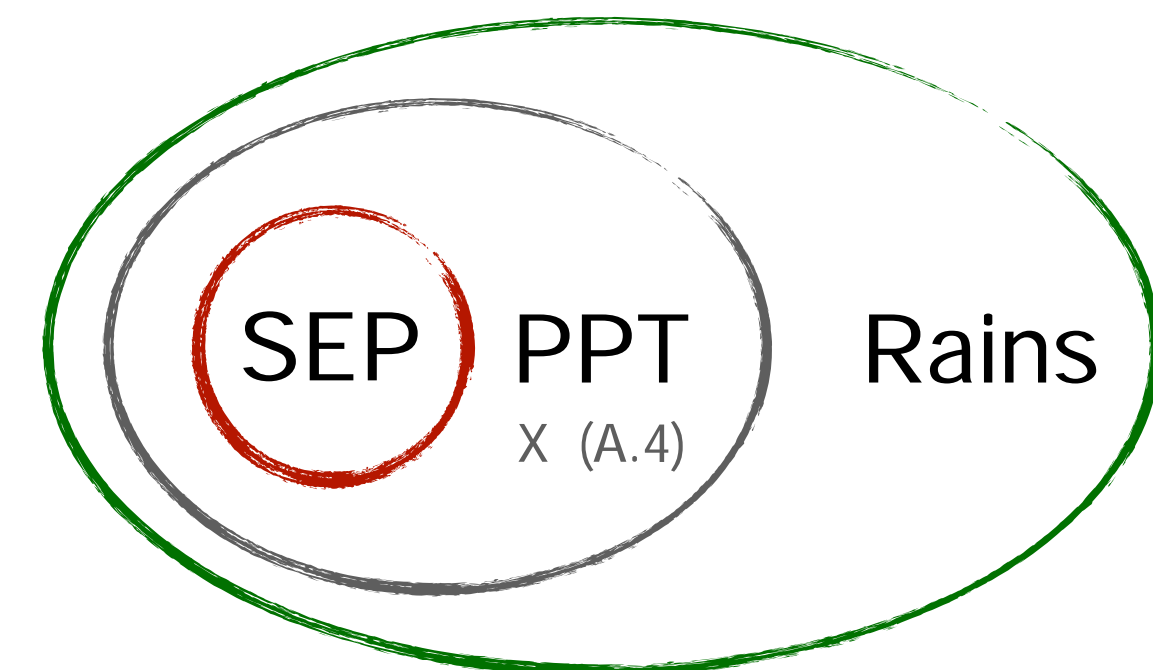
Entanglement distillation

The maximum number of Bell states that can be extracted from the given state with asymptotically vanishing error under the asymptotically non-entanglement generating operations (ANE) [Brandão, Plenio, 2010]

$$E_{D,ANE}(\rho_{AB}) = D^\infty(\rho_{AB} \| \text{SEP}) := \lim_{n \rightarrow \infty} \frac{1}{n} D(\rho_{AB}^{\otimes n} \| \text{SEP}(A^n : B^n))$$

Hard to evaluate in general

↓
Separable states



As $D(\rho \| \text{SEP})$ is minimization problem, any feasible solution gives an upper bound

$$D^\infty(\rho_{AB} \| \text{SEP}) \geq D^\infty(\rho_{AB} \| \text{Rains})$$

Can be **efficiently computed**

Operational meaning: distillable entanglement under *Rains-preserving operations*

[Regula, KF, Wang, Gu, 2019, NJP]

$$\text{Rains}(A : B) := \{\sigma \geq 0 : \|\sigma^{T_B}\|_1 \leq 1\}$$

[Rains, 2001; Audenaert et.al 2002]

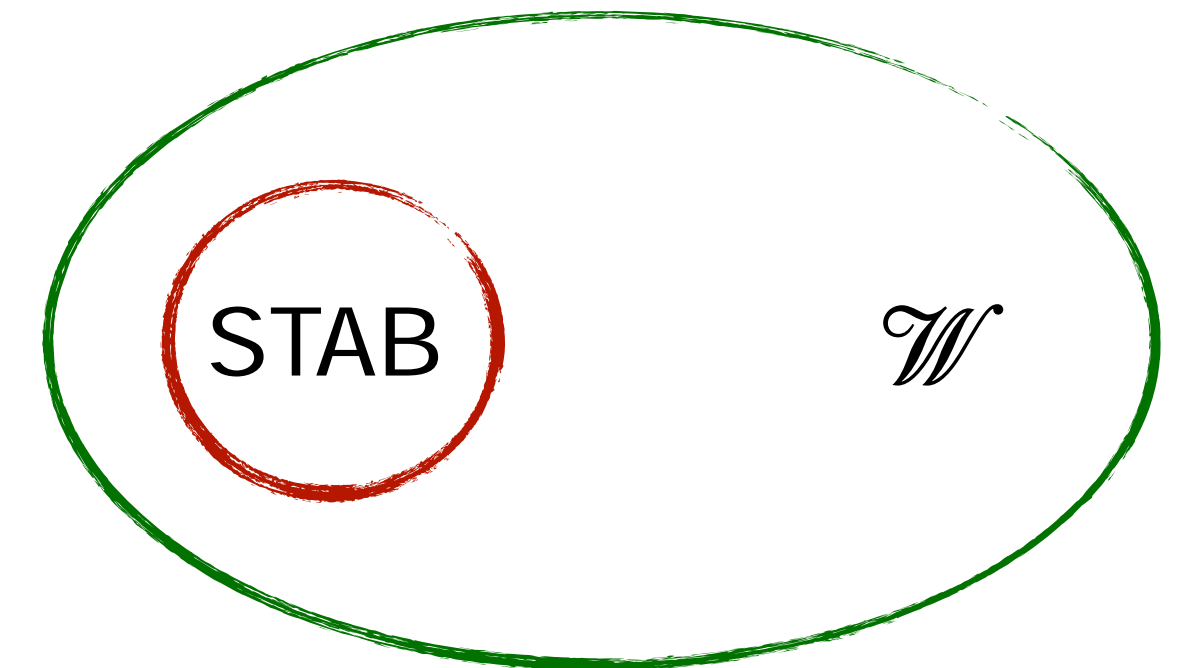
Application 4: efficient bounds for quantum resource theory

Magic state distillation

Extract as many copies of the target magic state as possible using stabilizer operations (STAB)

Thauma measure: [Wang, Wilde, Su, 2020, PRL]

Hard to evaluate in general $M_{D,\text{STAB}}(\rho) \leq D(\rho \| \mathcal{W}) c(T)$ \nearrow constant



Regularized Thauma measure, but remains **efficiently computable**

$$M_{D,\text{STAB}}(\rho) \leq D^\infty(\rho \| \mathcal{W}) c(T) \leq D(\rho \| \mathcal{W}) c(T)$$

$$\mathcal{W} := \{\sigma \geq 0 : \|\sigma\|_{\mathcal{W},1} \leq 1\}$$

Sub-normalized states
with non-positive mana

Summary

Generalized quantum AEP

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B})$$

Generality/efficiency/finite / estimate

(A.1) Each \mathcal{A}_n is convex and compact;

(A.2) Each \mathcal{A}_n is permutation-invariant;

(A.3) $\mathcal{A}_m \otimes \mathcal{A}_k \subseteq \mathcal{A}_{m+k}$, for all $m, k \in \mathbb{N}$;

(A.4) $(\mathcal{A}_m)_+^\circ \otimes (\mathcal{A}_k)_+^\circ \subseteq (\mathcal{A}_{m+k})_+^\circ$, for all $m, k \in \mathbb{N}$;

Technical tools (variational, superadditivity & chain rule):

$$D_M(\mathcal{A} \| \mathcal{B}) = \sup_{W \in \mathcal{B}_{++}^\circ} -h_{\mathcal{A}}(-\log W)$$

$$D_{M,\alpha}(\mathcal{A}_{12} \| \mathcal{B}_{12}) \geq D_{M,\alpha}(\mathcal{A}_1 \| \mathcal{B}_1) + D_{M,\alpha}(\mathcal{A}_2 \| \mathcal{B}_2)$$

$$D_{M,\alpha}(\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \| \mathcal{M}_{A \rightarrow B}(\sigma_{RA})) \geq D_{M,\alpha}(\rho_R \| \sigma_R) + D_{M,\alpha}^{\text{inf}}(\mathcal{N}_{A \rightarrow B} \| \mathcal{M}_{A \rightarrow B})$$

As AEP is in the heart of information theory, we expect further studies and applications.

Already been used in [2502.02563] by Arqand and Tan for quantum cryptography

I am hiring

One Brand, Two Campuses



香港中文大學
The Chinese University of Hong Kong



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen



Looking for postdocs, PhDs, research assistants...

Quantum Information Theory, Quantum Computation

kunfang.info

✉ kunfang@cuhk.edu.cn



I am hiring

One Brand, Two Campuses



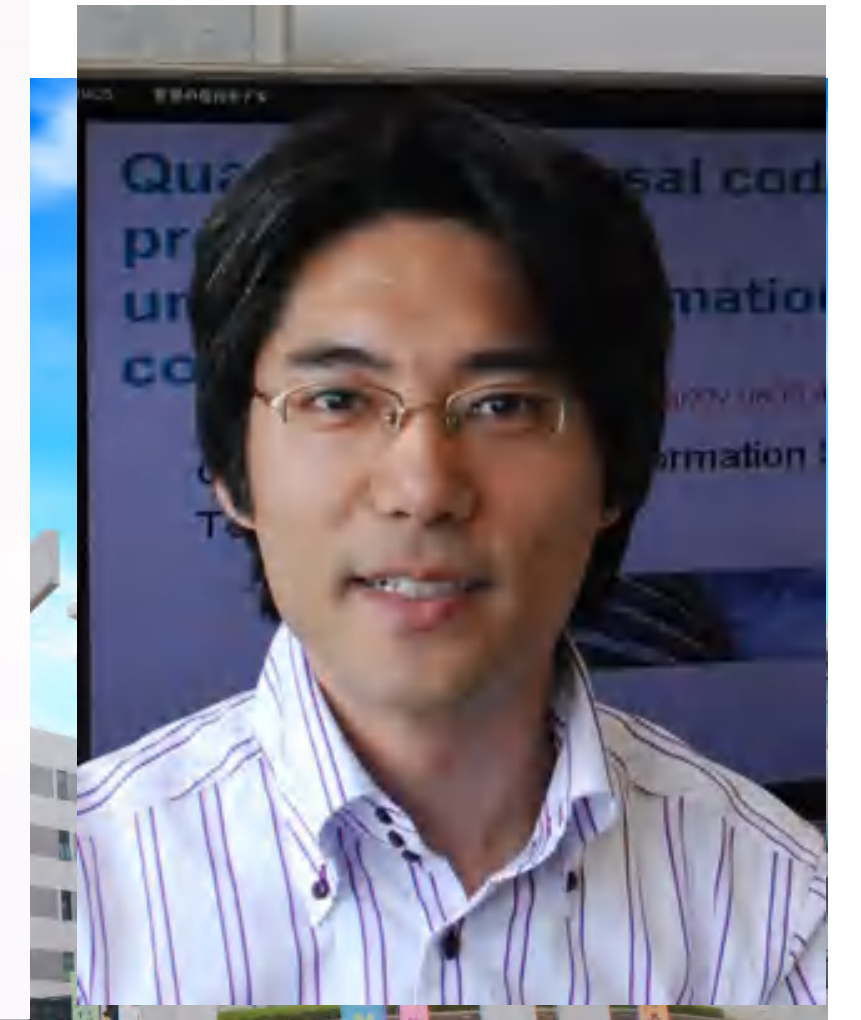
香港中文大學

The Chinese University of Hong Kong



香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen



Looking for postdocs, PhDs, research assistants...

Quantum Information Theory, Quantum Computation

kunfang.info

✉ kunfang@cuhk.edu.cn



BIID 2026



香港中文大學(深圳)

The Chinese University of Hong Kong, Shenzhen

Venue: CUHK-Shenzhen

Local Organizers:

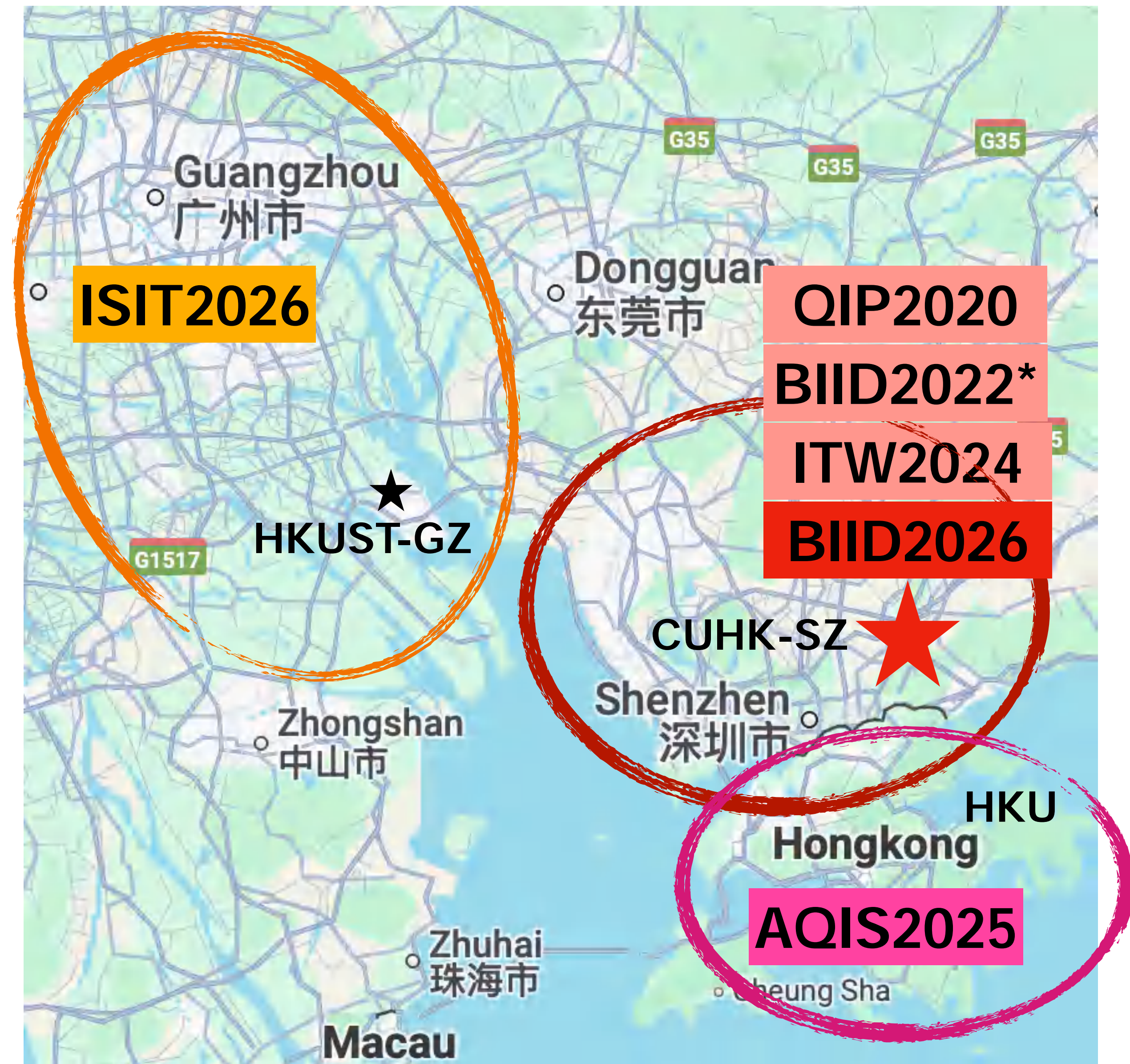
Masahito Hayashi, Kun Fang, Xin Wang ...

PC Chair:

Frédéric Dupuis

Planned Date: June 22-26, 2026 (one week before ISIT 2026 in Guangzhou)

Active place for information theory
& generous **visa-free** policy for
international travelers



Thanks for your attention!

arXiv: 2411.04035 & 2502.15659 & 2506.03060

