# New Rényi divergence families defined via convex optimization and their applications

Peter Brown, Hamza Fawzi and Omar Fawzi

Feb 01, 2021

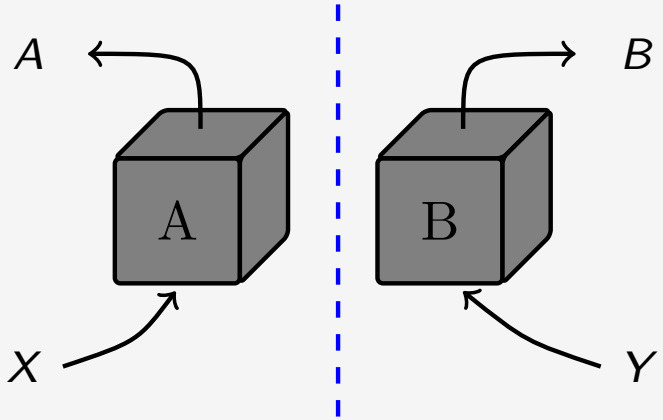
# Part 1

The *iterated mean* divergences and their application to device-independent cryptography

Based on *Brown, P., Fawzi, H. and Fawzi, O., Computing conditional entropies for quantum correlations, Nat Commun 12, 575 (2021), arXiv:2007.12575.*

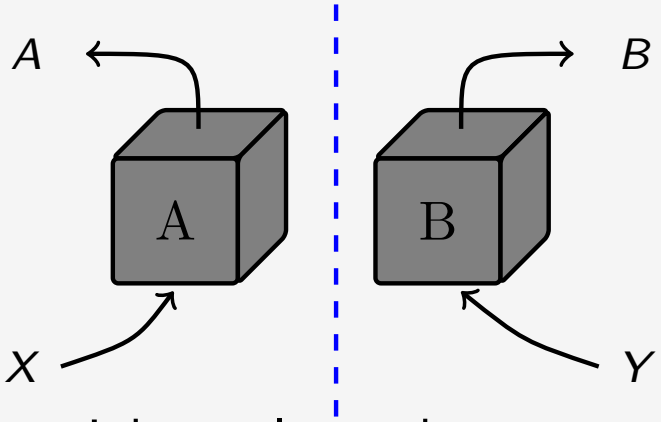
# Motivation I

## Bell-nonlocality



# Motivation I

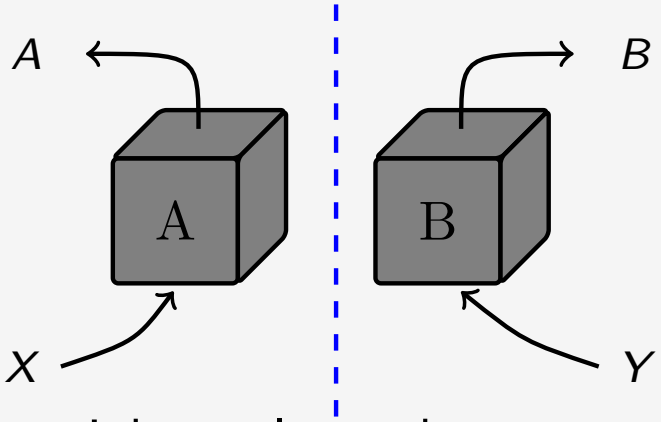
## Bell-nonlocality




- Nonlocal correlations are inherently random.

# Motivation I

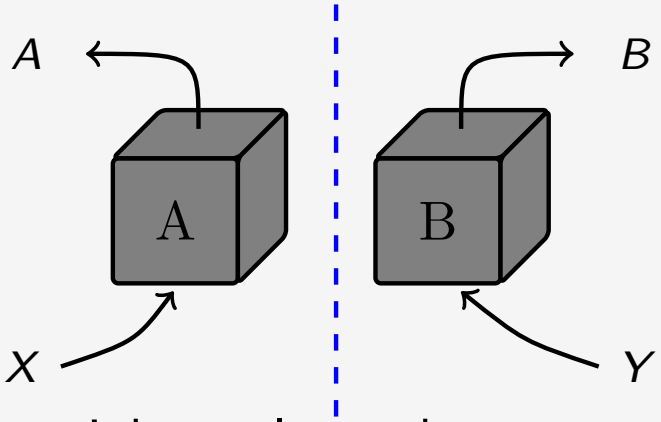
**Bell-nonlocality**




- Nonlocal correlations are inherently random.
- Foundation for randomness expansion / key-distribution protocols!

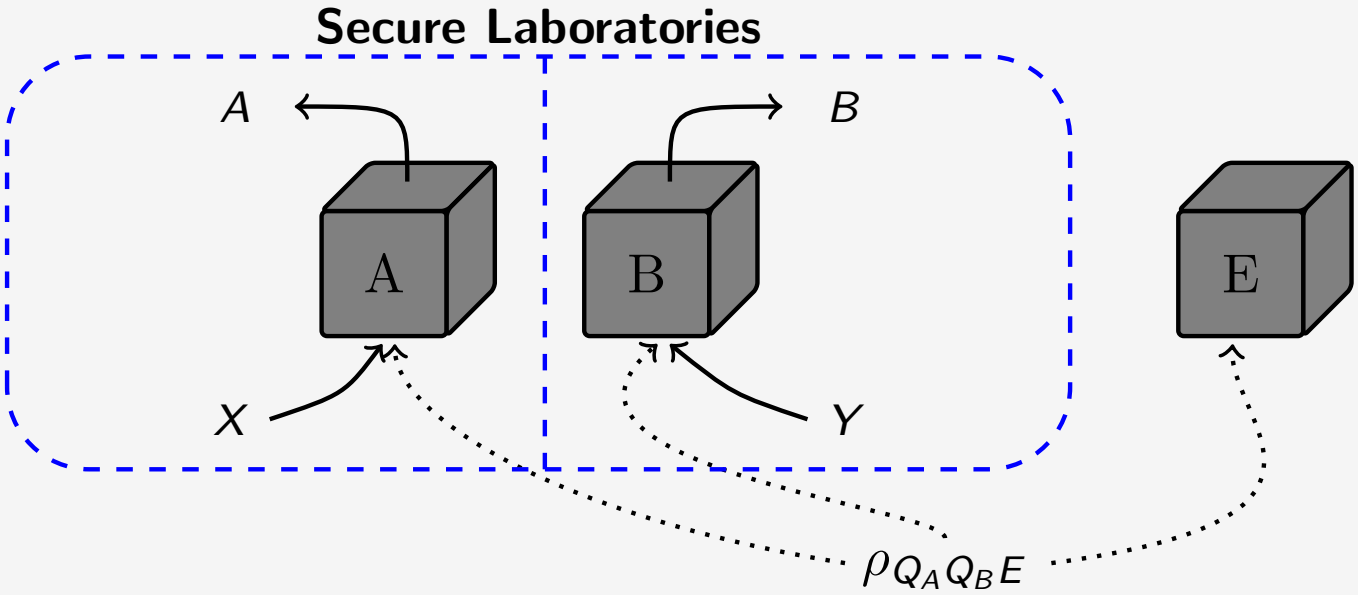
# Motivation I

**Bell-nonlocality**

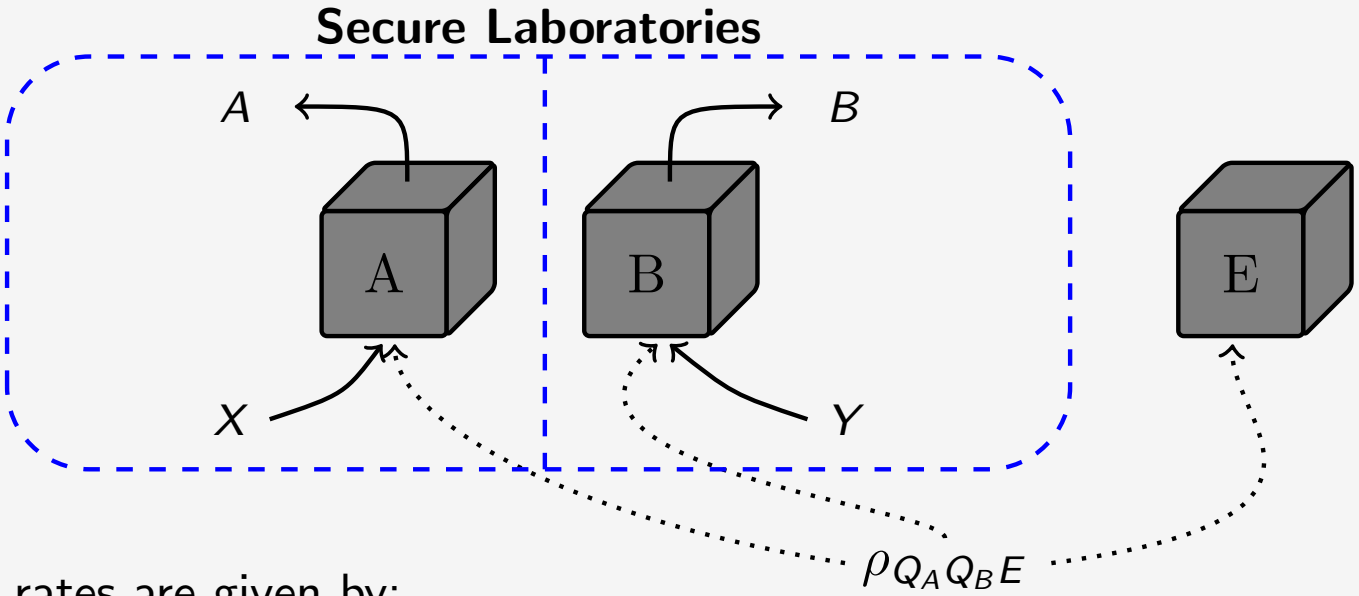



- Nonlocal correlations are inherently random.
- Foundation for randomness expansion / key-distribution protocols!
- Security and analysis relies on being able to calculate the *rates* of such protocols (bits per round).

# Randomness generated per round



# Randomness generated per round




Asymptotic rates are given by:

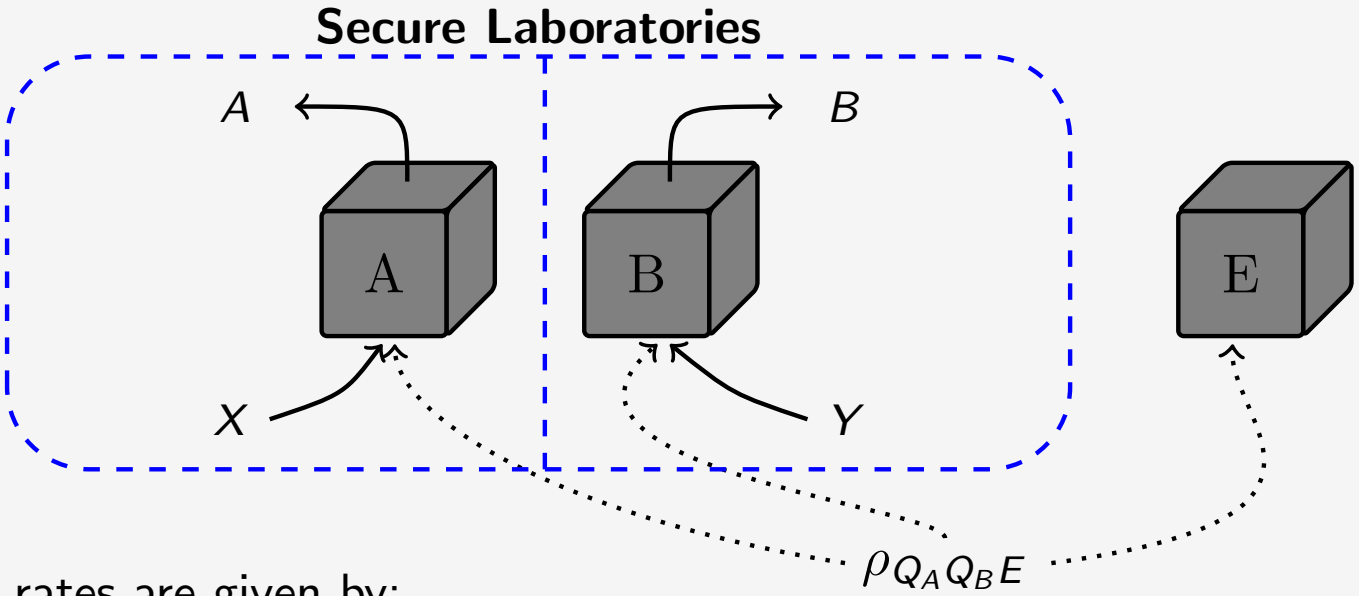
- **Randomness expansion**

$$H(AB|X = x^*, Y = y^*, E)$$

- **QKD**

$$H(A|X = x^*, E) - H(A|X = x^*, Y = y^*, B)$$

# Randomness generated per round




Asymptotic rates are given by:

- **Randomness expansion**

$$H(AB|X=x^*, Y=y^*, E)$$

- **QKD**

$$H(A|X=x^*, E) - H(A|X=x^*, Y=y^*, B)$$

Want device-independent lower bounds

# Example

We want lower bounds on

$$\begin{aligned} \inf \quad & H(A|X=x^*,E) \\ \text{s.t.} \quad & \sum_{abxy} c^i_{abxy}\, p(ab|xy) = w^i \end{aligned}$$

where the infimum is over all finite dimensional **states** $\rho_{Q_A Q_B E}$, **POVMs** $\{\{M_{a|x}\}_a\}_x, \{\{N_{b|y}\}_b\}_y$ and joint **Hilbert spaces** $Q_A \otimes Q_B \otimes E$.

# Example

We want lower bounds on

Difficult to solve
nonconvex / unbounded dimension

$$\begin{aligned} \inf \quad & H(A|X=x^*,E) \\ \text{s.t.} \quad & \sum_{abxy} c^i_{abxy}\, p(ab|xy) = w^i \end{aligned}$$

where the infimum is over all finite dimensional **states** $\rho_{Q_A Q_B E}$, **POVMs** $\{\{M_{a|x}\}_a\}_x, \{\{N_{b|y}\}_b\}_y$ and joint **Hilbert spaces** $Q_A \otimes Q_B \otimes E$.

# Example

We want lower bounds on

Difficult to solve
nonconvex / unbounded dimension

$$\begin{aligned} \inf \quad & H(A|X=x^*,E) \\ \text{s.t.} \quad & \sum_{abxy} c^i_{abxy}\, p(ab|xy) = w^i \end{aligned}$$

where the infimum is over all finite dimensional **states** $\rho_{Q_A Q_B E}$, **POVMs** $\{\{M_{a|x}\}_a\}_x, \{\{N_{b|y}\}_b\}_y$ and joint **Hilbert spaces** $Q_A \otimes Q_B \otimes E$.

**Known approaches**

- Analytical bounds [PAB+09] – tight bounds / restricted scope
- Numerical bounds on $H_{\min}$ – easy to compute / poor bounds
- Recent work [TSG+19] – good bounds / computationally intensive

# Example

We want lower bounds on

Difficult to solve
nonconvex / unbounded dimension

$$\begin{aligned} \inf \quad & H(A|X=x^*,E) \\ \text{s.t.} \quad & \sum_{abxy} c^i_{abxy}\, p(ab|xy) = w^i \end{aligned}$$

where the infimum is over all finite dimensional **states** $\rho_{Q_A Q_B E}$, **POVMs** $\{\{M_{a|x}\}_a\}_x, \{\{N_{b|y}\}_b\}_y$ and joint **Hilbert spaces** $Q_A \otimes Q_B \otimes E$.

## Known approaches

- Analytical bounds [PAB+09] – tight bounds / restricted scope
- Numerical bounds on $H_{\min}$ – easy to compute / poor bounds
- Recent work [TSG+19] – good bounds / computationally intensive

## Our approach

- Define new conditional entropies that are easy to bound device-independently and lower bound $H(A|E)$.

# The IM divergences

Entropies are special cases of divergences

$$\mathbb{H}^{\uparrow}(A|B)_{\rho} := \sup_{\sigma_B} -\mathbb{D}(\rho_{AB}\|I \otimes \sigma_B)$$

or

$$\mathbb{H}^{\downarrow}(A|B)_{\rho} := -\mathbb{D}(\rho_{AB}\|I \otimes \rho_B).$$

# The IM divergences

Entropies are special cases of divergences

$$\mathbb{H}^{\uparrow}(A|B)_{\rho} := \sup_{\sigma_B} -\mathbb{D}(\rho_{AB}\|I \otimes \sigma_B)$$

or

$$\mathbb{H}^{\downarrow}(A|B)_{\rho} := -\mathbb{D}(\rho_{AB}\|I \otimes \rho_B).$$

**We define our conditional entropies via a divergence**

# The IM divergences

## Definition (Iterated mean divergences)

Let $\alpha_k = 2^k/(2^k - 1)$ for $k = 1, 2, \ldots$. Then the **iterated mean** divergences are defined as

$$D_{(\alpha_k)}(\rho\|\sigma) := \frac{1}{\alpha_k - 1} \log Q_{(\alpha_k)}(\rho\|\sigma) \ , \tag{1}$$

with

$$\begin{aligned} Q_{(\alpha_k)}(\rho\|\sigma) := \max_{V_1,\ldots,V_k,Z} \ & \alpha_k \mathrm{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] - (\alpha_k - 1)\mathrm{Tr}\left[\sigma Z\right] \\ \text{s.t.} \quad & V_1 + V_1^* \geq 0 \\ & \begin{pmatrix} I & V_1 \\ V_1^* & \frac{(V_2+V_2^*)}{2} \end{pmatrix} \geq 0 \begin{pmatrix} I & V_2 \\ V_2^* & \frac{(V_3+V_3^*)}{2} \end{pmatrix} \geq 0 \cdots \begin{pmatrix} I & V_k \\ V_k^* & Z \end{pmatrix} \geq 0, \end{aligned} \tag{2}$$

where the optimization varies over $V_1, \ldots, V_k, Z \in \mathscr{L}(\mathcal{H})$.

# The IM divergences

**Discrete family** – $(2, \frac{4}{3}, \frac{8}{7}, \frac{16}{15}, \ldots$

## Definition (Iterated mean divergences)

Let $\alpha_k = 2^k/(2^k - 1)$ for $k = 1, 2, \ldots$. Then the **iterated mean** divergences are defined as

$$D_{(\alpha_k)}(\rho\|\sigma) := \frac{1}{\alpha_k - 1} \log Q_{(\alpha_k)}(\rho\|\sigma) \ , \tag{1}$$

with

$$\begin{aligned} Q_{(\alpha_k)}(\rho\|\sigma) := \max_{V_1,\ldots,V_k,Z} \ & \alpha_k \mathrm{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] - (\alpha_k - 1)\mathrm{Tr}\left[\sigma Z\right] \\ \text{s.t.} \quad & V_1 + V_1^* \geq 0 \\ & \begin{pmatrix} I & V_1 \\ V_1^* & \frac{(V_2+V_2^*)}{2} \end{pmatrix} \geq 0 \begin{pmatrix} I & V_2 \\ V_2^* & \frac{(V_3+V_3^*)}{2} \end{pmatrix} \geq 0 \cdots \begin{pmatrix} I & V_k \\ V_k^* & Z \end{pmatrix} \geq 0, \end{aligned} \tag{2}$$

where the optimization varies over $V_1, \ldots, V_k, Z \in \mathscr{L}(\mathcal{H})$.

# The IM divergences

**Discrete family** – $(2, \frac{4}{3}, \frac{8}{7}, \frac{16}{15}, \ldots$

## Definition (Iterated mean divergences)

Let $\alpha_k = 2^k/(2^k - 1)$ for $k = 1, 2, \ldots$. Then the **iterated mean** divergences are defined as

$$D_{(\alpha_k)}(\rho\|\sigma) := \frac{1}{\alpha_k - 1} \log Q_{(\alpha_k)}(\rho\|\sigma) \ , \tag{1}$$

with

$$Q_{(\alpha_k)}(\rho\|\sigma) := \max_{V_1,\ldots,V_k,Z} \ \alpha_k \mathrm{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] - (\alpha_k - 1)\mathrm{Tr}\left[\sigma Z\right]$$

$$\text{s.t.} \quad V_1 + V_1^* \geq 0$$

**Defined via SDP**

$$\begin{pmatrix} I & V_1 \\ V_1^* & \frac{(V_2+V_2^*)}{2} \end{pmatrix} \geq 0 \begin{pmatrix} I & V_2 \\ V_2^* & \frac{(V_3+V_3^*)}{2} \end{pmatrix} \geq 0 \cdots \begin{pmatrix} I & V_k \\ V_k^* & Z \end{pmatrix} \geq 0, \tag{2}$$

where the optimization varies over $V_1, \ldots, V_k, Z \in \mathscr{L}(\mathcal{H})$.

# The IM divergences

**Discrete family** – $(2, \frac{4}{3}, \frac{8}{7}, \frac{16}{15}, \ldots$

## Definition (Iterated mean divergences)

Let $\alpha_k = 2^k/(2^k - 1)$ for $k = 1, 2, \ldots$. Then the **iterated mean** divergences are defined as

$$D_{(\alpha_k)}(\rho\|\sigma) := \frac{1}{\alpha_k - 1} \log Q_{(\alpha_k)}(\rho\|\sigma) \ , \tag{1}$$

with

**Linear in $\rho$ and $\sigma$**

$$Q_{(\alpha_k)}(\rho\|\sigma) := \max_{V_1,\ldots,V_k,Z} \ \alpha_k \mathrm{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] - (\alpha_k - 1)\mathrm{Tr}\left[\sigma Z\right]$$

$$\text{s.t.} \quad V_1 + V_1^* \geq 0$$

**Defined via SDP**

$$\begin{pmatrix} I & V_1 \\ V_1^* & \frac{(V_2+V_2^*)}{2} \end{pmatrix} \geq 0 \begin{pmatrix} I & V_2 \\ V_2^* & \frac{(V_3+V_3^*)}{2} \end{pmatrix} \geq 0 \cdots \begin{pmatrix} I & V_k \\ V_k^* & Z \end{pmatrix} \geq 0, \tag{2}$$

where the optimization varies over $V_1, \ldots, V_k, Z \in \mathscr{L}(\mathcal{H})$.

# The IM divergences

**Discrete family** – $(2, \frac{4}{3}, \frac{8}{7}, \frac{16}{15}, \ldots$

## Definition (Iterated mean divergences)

Let $\alpha_k = 2^k/(2^k - 1)$ for $k = 1, 2, \ldots$. Then the **iterated mean** divergences are defined as

$$D_{(\alpha_k)}(\rho\|\sigma) := \frac{1}{\alpha_k - 1} \log Q_{(\alpha_k)}(\rho\|\sigma) \ , \tag{1}$$

with

**Linear in $\rho$ and $\sigma$**

$$Q_{(\alpha_k)}(\rho\|\sigma) := \max_{V_1,\ldots,V_k,Z} \ \alpha_k \mathrm{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] - (\alpha_k - 1)\mathrm{Tr}\left[\sigma Z\right]$$

$$\text{s.t.} \quad V_1 + V_1^* \geq 0$$

**Defined via SDP**

$$\begin{pmatrix} I & V_1 \\ V_1^* & \frac{(V_2+V_2^*)}{2} \end{pmatrix} \geq 0 \ \begin{pmatrix} I & V_2 \\ V_2^* & \frac{(V_3+V_3^*)}{2} \end{pmatrix} \geq 0 \ \cdots \ \begin{pmatrix} I & V_k \\ V_k^* & Z \end{pmatrix} \geq 0, \tag{2}$$

where the optimization varies over $V_1, \ldots, V_k, Z \in \mathscr{L}(\mathcal{H})$.

**Structure independent of the dimension!**

# IM divergence properties

- Satisfies data processing

$$D_{(\alpha_k)}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_{(\alpha_k)}(\rho\|\sigma) \qquad \forall \text{ channels } \mathcal{E}.$$

# IM divergence properties

- Satisfies data processing

$$D_{(\alpha_k)}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_{(\alpha_k)}(\rho\|\sigma) \qquad \forall \text{ channels } \mathcal{E}.$$

- Lies between geometric and sandwiched

$$\widetilde{D}_{\alpha_k}(\rho\|\sigma) \leq D_{(\alpha_k)}(\rho\|\sigma) \leq \widehat{D}_{\alpha_k}(\rho\|\sigma)$$

**Conditional entropies will lower bound $H$**

# IM divergence properties

- Satisfies data processing

$$D_{(\alpha_k)}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_{(\alpha_k)}(\rho\|\sigma) \qquad \forall \text{ channels } \mathcal{E}.$$

- Lies between geometric and sandwiched

$$\widetilde{D}_{\alpha_k}(\rho\|\sigma) \leq D_{(\alpha_k)}(\rho\|\sigma) \leq \widehat{D}_{\alpha_k}(\rho\|\sigma)$$

**Conditional entropies will lower bound $H$**

- Decreasing in $k$

$$D_{(\alpha_k)}(\rho\|\sigma) \leq D_{(\alpha_{k-1})}(\rho\|\sigma)$$

and so for the corresponding conditional entropies

$$H_{(\alpha_k)}(A|B) \geq H_{(\alpha_{k-1})}(A|B)$$

**Improving lower bounds on $H$**

# IM conditional entropies

Using the IM divergences we can construct a conditional entropy. Given a bipartitie state $\rho_{AB}$ we have

$$H^{\uparrow}_{(\alpha_k)}(A|B)_\rho = \frac{\alpha_k}{1-\alpha_k} \log Q^{\uparrow}_{(\alpha_k)}(\rho) \tag{3}$$

where

$$\begin{aligned} Q^{\uparrow}_{(\alpha_k)}(\rho) = \max_{V_1,\ldots,V_k} \ & \mathrm{Tr}\left[\rho \frac{(V_1+V_1^*)}{2}\right] \\ \text{s.t.} \quad & \mathrm{Tr}_A\left[V_k^* V_k\right] \leq I_B \\ & V_1 + V_1^* \geq 0 \\ & \begin{pmatrix} I & V_1 \\ V_1^* & \frac{(V_2+V_2^*)}{2} \end{pmatrix} \geq 0 \quad \begin{pmatrix} I & V_2 \\ V_2^* & \frac{(V_3+V_3^*)}{2} \end{pmatrix} \geq 0 \quad \cdots \\ & \begin{pmatrix} I & V_{k-1} \\ V_{k-1}^* & \frac{(V_k+V_k^*)}{2} \end{pmatrix} \geq 0 \, . \end{aligned} \tag{4}$$

# IM conditional entropies

Using the IM divergences we can construct a conditional entropy. Given a bipartitie state $\rho_{AB}$ we have

$$H^{\uparrow}_{(\alpha_k)}(A|B)_\rho = \frac{\alpha_k}{1-\alpha_k} \log Q^{\uparrow}_{(\alpha_k)}(\rho) \tag{3}$$

where

$$\begin{aligned} Q^{\uparrow}_{(\alpha_k)}(\rho) = \max_{V_1,\ldots,V_k} \ & \mathrm{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] \\ \text{s.t.} \quad & \mathrm{Tr}_A\left[V_k^* V_k\right] \leq I_B \\ & V_1 + V_1^* \geq 0 \\ & \begin{pmatrix} I & V_1 \\ V_1^* & \frac{(V_2+V_2^*)}{2} \end{pmatrix} \geq 0 \quad \begin{pmatrix} I & V_2 \\ V_2^* & \frac{(V_3+V_3^*)}{2} \end{pmatrix} \geq 0 \quad \cdots \\ & \begin{pmatrix} I & V_{k-1} \\ V_{k-1}^* & \frac{(V_k+V_k^*)}{2} \end{pmatrix} \geq 0 \, . \end{aligned} \tag{4}$$

**Form still suitable for DI optimization!**

# IM conditional entropies II

For example

$$H_{(2)}^{\uparrow}(A|B)_\rho = -2\log \max_{V_1} \operatorname{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] \quad \text{s.t.} \quad \operatorname{Tr}_A\left[V_1^* V_1\right] \leq I_B, \quad V_1 + V_1^* \geq 0 \tag{5}$$

Compare with

$$H_{\min}(A|B)_\rho = -\log \max \operatorname{Tr}\left[\rho M\right] \quad \text{s.t.} \quad \operatorname{Tr}_A\left[M\right] \leq I_B, \quad M \geq 0 \tag{6}$$

# IM conditional entropies II

For example

$$H^{\uparrow}_{(2)}(A|B)_\rho = -2\log \max_{V_1} \mathrm{Tr}\left[\rho \frac{(V_1 + V_1^*)}{2}\right] \quad \text{s.t.} \quad \mathrm{Tr}_A\left[V_1^* V_1\right] \leq I_B, \quad V_1 + V_1^* \geq 0 \tag{5}$$

Compare with

$$H_{\min}(A|B)_\rho = -\log \max \mathrm{Tr}\left[\rho M\right] \quad \text{s.t.} \quad \mathrm{Tr}_A\left[M\right] \leq I_B, \quad M \geq 0 \tag{6}$$

For DI applications we can rewrite this in terms of the initial entangled state $|\psi\rangle\langle\psi|$ and the POVM operators used by Alice.

Can then be optimized in the Navascués Pironio Acín hierarchy [NPA07].

# Application: DIRNG/DIQKD setup

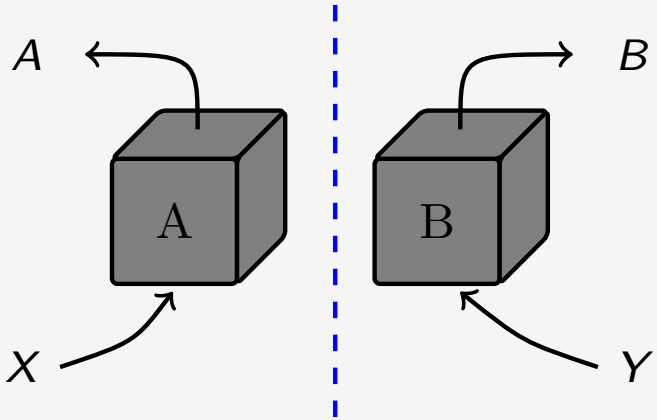
- DIRNG – Lower bound

$$H(AB|X=x, Y=y, E)$$

.

- DIQKD – Lower bound

$$H(A|X=x, E) - H(A|B, X=x, Y=y)$$



# Application: DIRNG/DIQKD setup

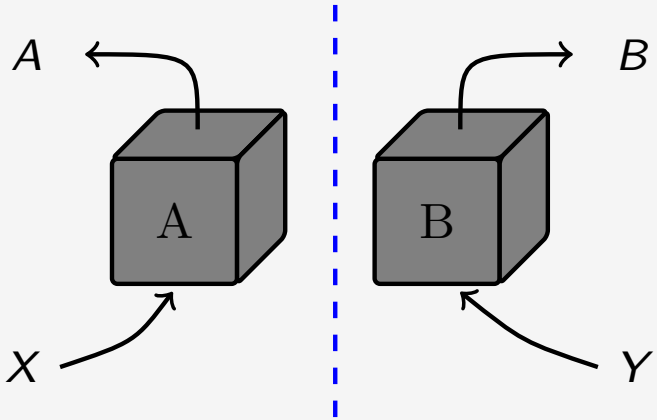
- DIRNG – Lower bound

$$H(AB|X=x, Y=y, E)$$

.

- DIQKD – Lower bound

$$H(A|X=x, E) - H(A|B, X=x, Y=y)$$




- Constrain devices by some full joint probability distribution $p_{AB|XY}$.

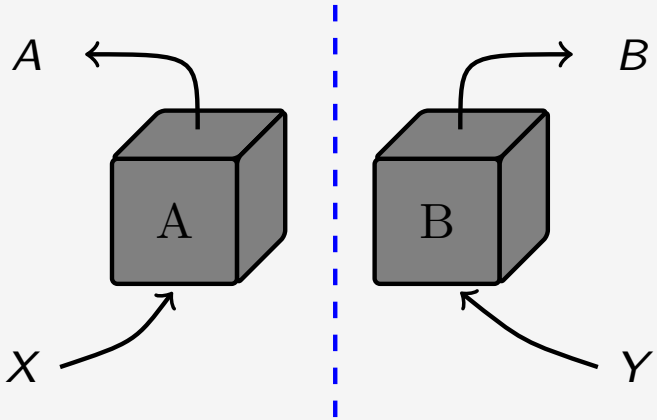
# Application: DIRNG/DIQKD setup

- DIRNG – Lower bound

$$H(AB|X=x, Y=y, E)$$

.

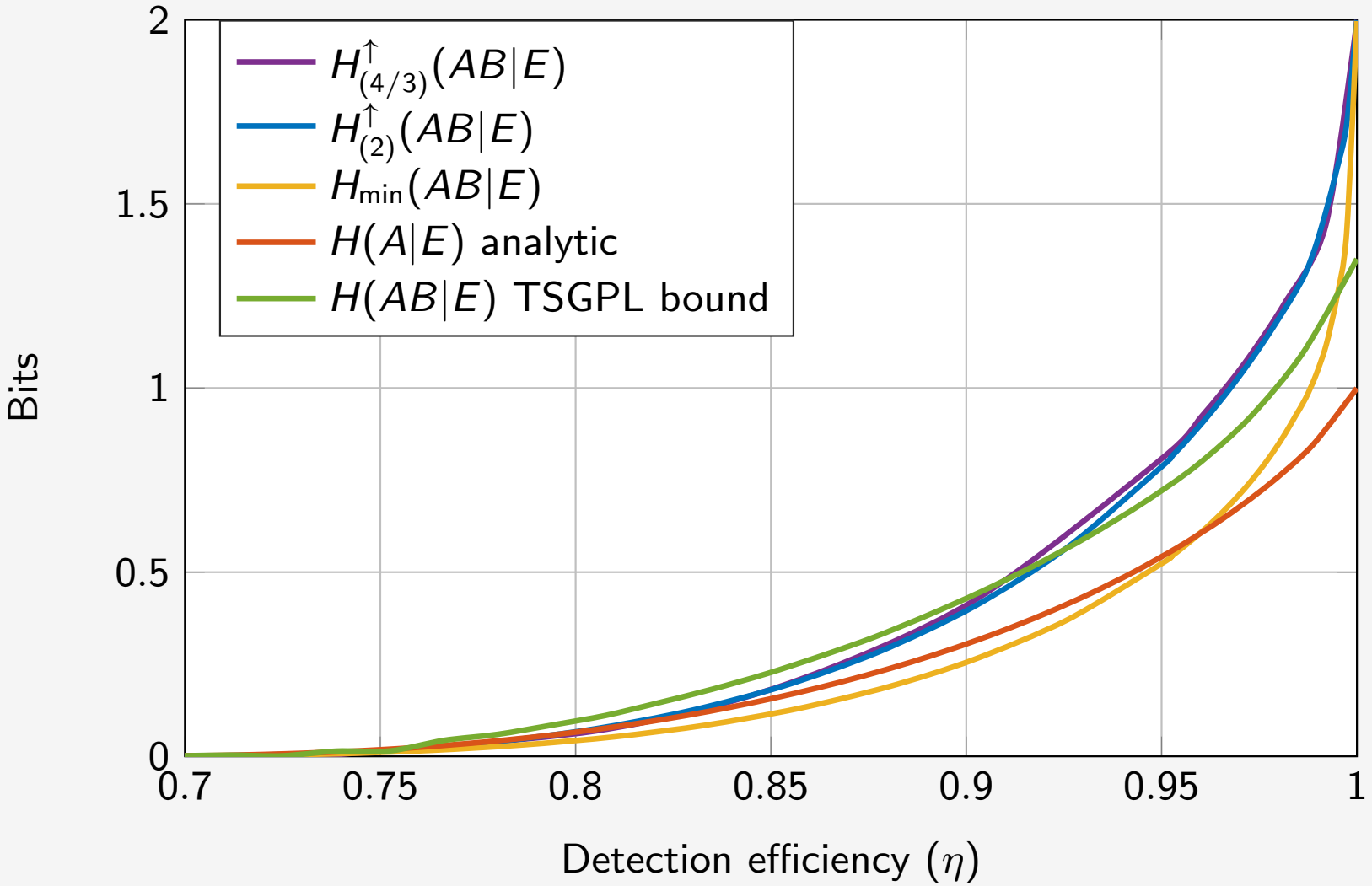
- DIQKD – Lower bound




$$H(A|X=x, E) - H(A|B, X=x, Y=y)$$

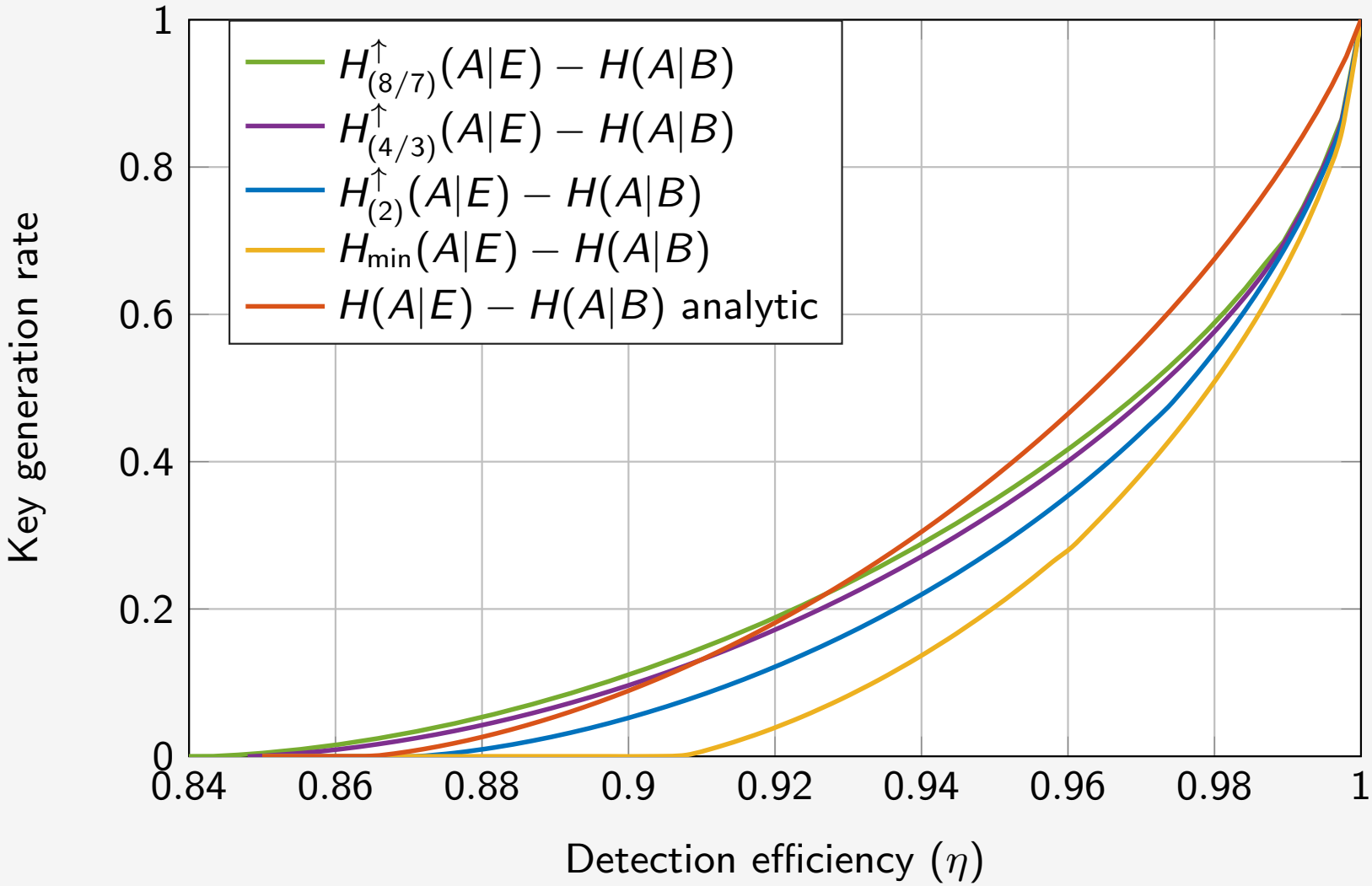
- Constrain devices by some full joint probability distribution $p_{AB|XY}$.
- Assume devices have *detection inefficiencies*. With probability $\eta$ device measures correctly and with probability $1-\eta$ device deterministically outputs 0.

# Application: DIRNG - full statistics / inefficient detectors

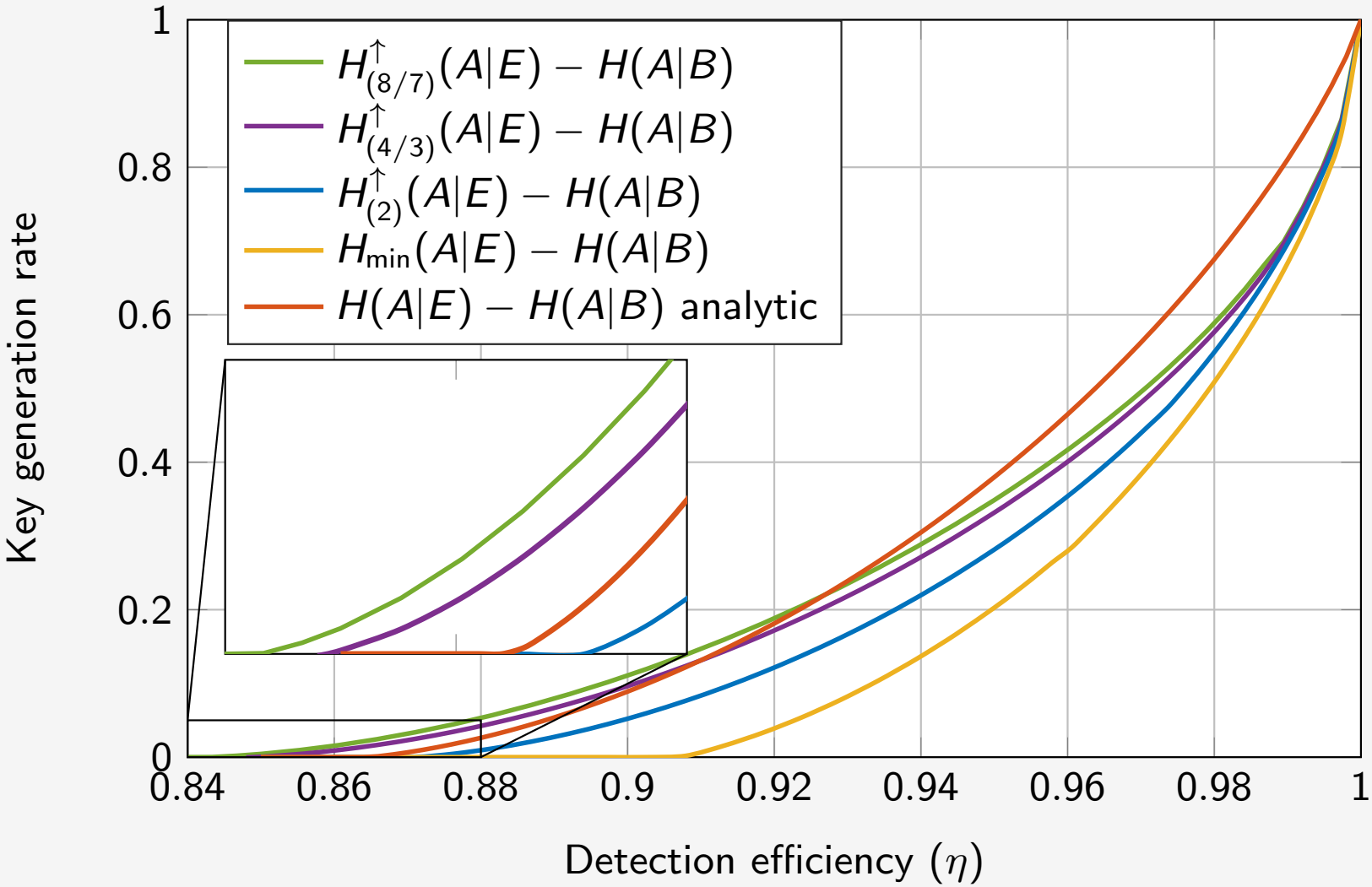



$H(A|E)$ bound from [PAB$^+$09].
TSGPL bound from [TSG$^+$19].

# Application: DIQKD - full statistics / inefficient detectors



# Application: DIQKD - full statistics / inefficient detectors



# Part 2

Divergences defined via convex optimization with applications to quantum Shannon theory

Based on *Fawzi, H. and Fawzi, O., Defining quantum divergences via convex optimization, Quantum, 2021, arXiv:2007.12576*.

# Motivation

Divergences are useful quantities in both classical and quantum Shannon theory.

- Can be used to define other important entropic quantities – entropies / mutual information.

# Motivation

Divergences are useful quantities in both classical and quantum Shannon theory.

- Can be used to define other important entropic quantities – entropies / mutual information.
- Find direct operational meanings in rates for hypothesis testing – measures of distinguishability.

# Motivation

Divergences are useful quantities in both classical and quantum Shannon theory.

- Can be used to define other important entropic quantities – entropies / mutual information.
- Find direct operational meanings in rates for hypothesis testing – measures of distinguishability.
- This work introduces another family of divergences $D_\alpha^\#$ which provide new insights for the sandwiched divergences

$$\widetilde{D}_\alpha(\rho\|\sigma) = \frac{1}{\alpha - 1} \log \mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right].$$

# Definition

Given two PSD matrices $A \gg B$ and $\beta \in [0,1]$, let

$$A\#_\beta B := A^{1/2}(A^{-1/2}BA^{-1/2})^\beta A^{1/2}.$$

# Definition

Given two PSD matrices $A \gg B$ and $\beta \in [0,1]$, let

$$A\#_\beta B := A^{1/2}(A^{-1/2}BA^{-1/2})^\beta A^{1/2}.$$

## Definition

For $\alpha > 1$ let

$$D_\alpha^\#(\rho\|\sigma) := \frac{1}{\alpha - 1} \log Q_\alpha^\#(\rho\|\sigma)$$

where

$$Q_\alpha^\#(\rho\|\sigma) := \min_{A \geq 0} \quad \mathrm{Tr}\,[A]$$
$$\text{s.t.} \quad \rho \leq \sigma\#_{1/\alpha}A$$

# Definition

Given two PSD matrices $A \gg B$ and $\beta \in [0,1]$, let

$$A\#_\beta B := A^{1/2}(A^{-1/2}BA^{-1/2})^\beta A^{1/2}.$$

## Definition

For $\alpha > 1$ let

$$D_\alpha^\#(\rho\|\sigma) := \frac{1}{\alpha - 1} \log Q_\alpha^\#(\rho\|\sigma)$$

where

$$Q_\alpha^\#(\rho\|\sigma) := \min_{A \geq 0} \quad \mathrm{Tr}\,[A]$$
$$\text{s.t.} \quad \rho \leq \sigma\#_{1/\alpha}A$$

**SDP when $\alpha \in \mathbb{Q}$**

# Definition

Given two PSD matrices $A \gg B$ and $\beta \in [0,1]$, let

$$A\#_\beta B := A^{1/2}(A^{-1/2}BA^{-1/2})^\beta A^{1/2}.$$

## Definition

For $\alpha > 1$ let

$$D_\alpha^\#(\rho\|\sigma) := \frac{1}{\alpha - 1} \log Q_\alpha^\#(\rho\|\sigma)$$

where

$$Q_\alpha^\#(\rho\|\sigma) := \min_{A \geq 0} \quad \mathrm{Tr}\,[A]$$
$$\text{s.t.} \quad \rho \leq \sigma\#_{1/\alpha}A$$

**SDP when $\alpha \in \mathbb{Q}$**

**Same as IM divergence when $\alpha = 2$**

# Channel divergence

We can also define a corresponding divergence for channels $\mathcal{N}, \mathcal{M} : \mathcal{L}(X') \to \mathcal{L}(Y)$ in the usual way

$$D_\alpha^\#(\mathcal{N}\|\mathcal{M}) = \sup_{\rho_{XX'}} D_\alpha^\#((\mathcal{I} \otimes \mathcal{N})(\rho_{XX'})\|(\mathcal{I} \otimes \mathcal{M})(\rho_{XX'})).$$

# Channel divergence

We can also define a corresponding divergence for channels $\mathcal{N}, \mathcal{M} : \mathcal{L}(X') \to \mathcal{L}(Y)$ in the usual way

$$D_\alpha^\#(\mathcal{N}\|\mathcal{M}) = \sup_{\rho_{XX'}} D_\alpha^\#((\mathcal{I}\otimes\mathcal{N})(\rho_{XX'})\|(\mathcal{I}\otimes\mathcal{M})(\rho_{XX'})).$$

For $D_\alpha^\#$ this can be reformulated as a *convex optimization problem*

$$D_\alpha^\#(\mathcal{N}\|\mathcal{M}) = \frac{1}{\alpha-1}\log Q_\alpha^\#(\mathcal{N}\|\mathcal{M})$$

with

$$Q_\alpha^\#(\mathcal{N}\|\mathcal{M}) = \inf_{A_{XY}\geq 0} \|\mathrm{Tr}_Y[A_{XY}]\|_\infty$$
$$\text{s.t.} \quad J_{XY}^{\mathcal{N}} \leq J_{XY}^{\mathcal{M}} \#_{1/\alpha} A_{XY}$$

# Channel divergence

We can also define a corresponding divergence for channels $\mathcal{N}, \mathcal{M} : \mathcal{L}(X') \to \mathcal{L}(Y)$ in the usual way

$$D_\alpha^\#(\mathcal{N}\|\mathcal{M}) = \sup_{\rho_{XX'}} D_\alpha^\#((\mathcal{I}\otimes\mathcal{N})(\rho_{XX'})\|(\mathcal{I}\otimes\mathcal{M})(\rho_{XX'})).$$

For $D_\alpha^\#$ this can be reformulated as a *convex optimization problem*

$$D_\alpha^\#(\mathcal{N}\|\mathcal{M}) = \frac{1}{\alpha-1}\log Q_\alpha^\#(\mathcal{N}\|\mathcal{M})$$

with

$$Q_\alpha^\#(\mathcal{N}\|\mathcal{M}) = \inf_{A_{XY}\geq 0} \|\mathrm{Tr}_Y[A_{XY}]\|_\infty$$
$$\text{s.t.} \quad J_{XY}^{\mathcal{N}} \leq J_{XY}^{\mathcal{M}} \#_{1/\alpha} A_{XY}$$

**Choi matrices**

# Properties

- Satisfies data processing

$$D_\alpha^\#(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_\alpha^\#(\rho\|\sigma) \qquad \forall \text{ channels } \mathcal{E}.$$

# Properties

- Satisfies data processing

$$D_\alpha^\#(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_\alpha^\#(\rho\|\sigma) \qquad \forall \text{ channels } \mathcal{E}.$$

- Relation to other divergences

$$\widetilde{D}_\alpha(\rho\|\sigma) \leq D_\alpha^\#(\rho\|\sigma) \leq \widehat{D}_\alpha(\rho\|\sigma).$$

# Properties

- Satisfies data processing

$$D_\alpha^\#(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_\alpha^\#(\rho\|\sigma) \qquad \forall \text{ channels } \mathcal{E}.$$

- Relation to other divergences

$$\widetilde{D}_\alpha(\rho\|\sigma) \leq D_\alpha^\#(\rho\|\sigma) \leq \widehat{D}_\alpha(\rho\|\sigma).$$

- Regularizes to sandwiched divergence

$$\lim_{n\to\infty} \frac{1}{n} D_\alpha^\#(\rho^{\otimes n}\|\sigma^{\otimes n}) = \widetilde{D}_\alpha(\rho\|\sigma).$$

# Application I: Computing $\lim_{n\to\infty} \frac{1}{n}\widetilde{D}_\alpha(\mathcal{M}^{\otimes n}\|\mathcal{N}^{\otimes n})$

We can use $D_\alpha^\#$ to compute

$$\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) := \lim_{n\to\infty} \frac{1}{n}\widetilde{D}_\alpha(\mathcal{M}^{\otimes n}\|\mathcal{N}^{\otimes n})$$

to arbitrary accuracy.

**Useful quantity in channel discrimination**

# Application I: Computing $\lim_{n\to\infty} \frac{1}{n}\widetilde{D}_\alpha(\mathcal{M}^{\otimes n}\|\mathcal{N}^{\otimes n})$

We can use $D_\alpha^{\#}$ to compute

$$\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) := \lim_{n\to\infty} \tfrac{1}{n}\widetilde{D}_\alpha(\mathcal{M}^{\otimes n}\|\mathcal{N}^{\otimes n})$$

to arbitrary accuracy.

**Useful quantity in channel discrimination**

## Theorem (Informal)

*For all* $\alpha > 1$ *and* $m \geq 1$

$$\frac{1}{m} D_\alpha^{\#}(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}) - g(m,\alpha) \leq \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$$

*and*

$$\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) \leq \frac{1}{m} D_\alpha^{\#}(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}).$$

# Application I: Computing $\lim_{n\to\infty} \frac{1}{n}\widetilde{D}_\alpha(\mathcal{M}^{\otimes n}\|\mathcal{N}^{\otimes n})$

We can use $D_\alpha^\#$ to compute

$$\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) := \lim_{n\to\infty} \tfrac{1}{n}\widetilde{D}_\alpha(\mathcal{M}^{\otimes n}\|\mathcal{N}^{\otimes n})$$

to arbitrary accuracy.

**Useful quantity in channel discrimination**

## Theorem (Informal)

*For all* $\alpha > 1$ *and* $m \geq 1$

$$\frac{1}{m} D_\alpha^\#(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}) - g(m,\alpha) \leq \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$$

*and*

$$\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) \leq \frac{1}{m} D_\alpha^\#(\mathcal{N}^{\otimes m}\|\mathcal{M}^{\otimes m}).$$

**Can also be used to compute bounds on the relative entropy analogue!**

# Application II: A new chain rule for $\widetilde{D}_\alpha$

## Theorem (Chain rule for $\widetilde{D}_\alpha$)

*Let $\alpha > 1$, $\rho, \sigma \geq 0$ and $\mathcal{N}, \mathcal{M} : \mathscr{L}(X) \to \mathscr{L}(Y)$ be quantum channels. Then*

$$\widetilde{D}_\alpha(\mathcal{N}(\rho)\|\mathcal{M}(\sigma)) \leq \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) + \widetilde{D}_\alpha(\rho\|\sigma)$$

# Application II: A new chain rule for $\widetilde{D}_\alpha$

## Theorem (Chain rule for $\widetilde{D}_\alpha$)

*Let $\alpha > 1$, $\rho, \sigma \geq 0$ and $\mathcal{N}, \mathcal{M} : \mathscr{L}(X) \to \mathscr{L}(Y)$ be quantum channels. Then*

$$\widetilde{D}_\alpha(\mathcal{N}(\rho)\|\mathcal{M}(\sigma)) \leq \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) + \widetilde{D}_\alpha(\rho\|\sigma)$$

- Generalization of the DPI

# Application II: A new chain rule for $\widetilde{D}_\alpha$

## Theorem (Chain rule for $\widetilde{D}_\alpha$)

*Let $\alpha > 1$, $\rho, \sigma \geq 0$ and $\mathcal{N}, \mathcal{M} : \mathscr{L}(X) \to \mathscr{L}(Y)$ be quantum channels. Then*

$$\widetilde{D}_\alpha(\mathcal{N}(\rho)\|\mathcal{M}(\sigma)) \leq \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) + \widetilde{D}_\alpha(\rho\|\sigma)$$

- Generalization of the DPI
- Same chain rule already known for the relative entropy [FFRS20]

# Application II: A new chain rule for $\widetilde{D}_\alpha$

## Theorem (Chain rule for $\widetilde{D}_\alpha$)

*Let $\alpha > 1$, $\rho, \sigma \geq 0$ and $\mathcal{N}, \mathcal{M} : \mathscr{L}(X) \to \mathscr{L}(Y)$ be quantum channels. Then*

$$\widetilde{D}_\alpha(\mathcal{N}(\rho)\|\mathcal{M}(\sigma)) \leq \widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) + \widetilde{D}_\alpha(\rho\|\sigma)$$

- Generalization of the DPI
- Same chain rule already known for the relative entropy [FFRS20]
- Ex: useful for bounding repeated channel applications

$$\widetilde{D}_\alpha(\mathcal{N}^t(\rho)\|\mathcal{M}^t(\sigma)) \leq t\widetilde{D}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) + \widetilde{D}_\alpha(\rho\|\sigma)$$

# Application III: Channel discrimination

**Task:** Given black box access to one of the channels $\mathcal{N}, \mathcal{M} : \mathscr{L}(X') \to \mathscr{L}(Y)$, determine if you received $\mathcal{N}$.

# Application III: Channel discrimination

**Task:** Given black box access to one of the channels $\mathcal{N}, \mathcal{M} : \mathscr{L}(X') \to \mathscr{L}(Y)$, determine if you received $\mathcal{N}$.

- Recent work [WBHK20] introduced the *amortized divergence*

$$\mathbb{D}^a(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{XX'},\sigma_{XX'} \in \mathscr{D}(XX')} [\mathbb{D}(\mathcal{N}(\rho_{XX'})\|\mathcal{M}(\sigma_{XX'})) - \mathbb{D}(\rho_{XX'}\|\sigma_{XX'})]$$

  as a tool for computing rates of this task.

# Application III: Channel discrimination

**Task:** Given black box access to one of the channels $\mathcal{N}, \mathcal{M} : \mathscr{L}(X') \to \mathscr{L}(Y)$, determine if you received $\mathcal{N}$.

- Recent work [WBHK20] introduced the *amortized divergence*

$$\mathbb{D}^a(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{XX'},\sigma_{XX'}\in\mathscr{D}(XX')} [\mathbb{D}(\mathcal{N}(\rho_{XX'})\|\mathcal{M}(\sigma_{XX'})) - \mathbb{D}(\rho_{XX'}\|\sigma_{XX'})]$$

  as a tool for computing rates of this task.

- Using the chain rule one can prove

$$\widetilde{D}^a_\alpha(\mathcal{N}\|\mathcal{M}) = \widetilde{D}^{\mathrm{reg}}_\alpha(\mathcal{N}\|\mathcal{M}).$$

# Application III: Channel discrimination

**Task:** Given black box access to one of the channels $\mathcal{N}, \mathcal{M} : \mathscr{L}(X') \to \mathscr{L}(Y)$, determine if you received $\mathcal{N}$.

- Recent work [WBHK20] introduced the *amortized divergence*

$$\mathbb{D}^a(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{XX'},\sigma_{XX'}\in\mathscr{D}(XX')} [\mathbb{D}(\mathcal{N}(\rho_{XX'})\|\mathcal{M}(\sigma_{XX'})) - \mathbb{D}(\rho_{XX'}\|\sigma_{XX'})]$$

  as a tool for computing rates of this task.

- Using the chain rule one can prove

$$\widetilde{D}^a_\alpha(\mathcal{N}\|\mathcal{M}) = \widetilde{D}^{\mathrm{reg}}_\alpha(\mathcal{N}\|\mathcal{M}).$$

**We can compute this!**

# Application III: Channel discrimination

**Task:** Given black box access to one of the channels $\mathcal{N}, \mathcal{M} : \mathscr{L}(X') \to \mathscr{L}(Y)$, determine if you received $\mathcal{N}$.

- Recent work [WBHK20] introduced the *amortized divergence*

$$\mathbb{D}^a(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{XX'},\sigma_{XX'}\in\mathscr{D}(XX')} [\mathbb{D}(\mathcal{N}(\rho_{XX'})\|\mathcal{M}(\sigma_{XX'})) - \mathbb{D}(\rho_{XX'}\|\sigma_{XX'})]$$

  as a tool for computing rates of this task.

- Using the chain rule one can prove

$$\widetilde{D}^a_\alpha(\mathcal{N}\|\mathcal{M}) = \widetilde{D}^{\mathrm{reg}}_\alpha(\mathcal{N}\|\mathcal{M}).$$

  **We can compute this!**

- It can also be shown in certain new regimes that adaptive strategies do not help!

# Application III: Channel discrimination

**Task:** Given black box access to one of the channels $\mathcal{N}, \mathcal{M} : \mathscr{L}(X') \to \mathscr{L}(Y)$, determine if you received $\mathcal{N}$.

- Recent work [WBHK20] introduced the *amortized divergence*

$$\mathbb{D}^a(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_{XX'},\sigma_{XX'}\in\mathscr{D}(XX')} [\mathbb{D}(\mathcal{N}(\rho_{XX'})\|\mathcal{M}(\sigma_{XX'})) - \mathbb{D}(\rho_{XX'}\|\sigma_{XX'})]$$

  as a tool for computing rates of this task.

- Using the chain rule one can prove

$$\widetilde{D}^a_\alpha(\mathcal{N}\|\mathcal{M}) = \widetilde{D}^{\mathrm{reg}}_\alpha(\mathcal{N}\|\mathcal{M}).$$

**We can compute this!**

- It can also be shown in certain new regimes that adaptive strategies do not help!

**Strong converse exponent**

# Further work

- Use to design better DI protocols / apply to different DI tasks. Can we include preprocessing in DIQKD? [HST$^+$20, WAP20]

# Further work

- Use to design better DI protocols / apply to different DI tasks. Can we include preprocessing in DIQKD? [HST$^+$20, WAP20]
- Analyze finite round key rates (feasibility of DIQKD).

# Further work

- Use to design better DI protocols / apply to different DI tasks. Can we include preprocessing in DIQKD? [HST$^+$20, WAP20]
- Analyze finite round key rates (feasibility of DIQKD).
- Can we make the computations more efficient? (Symmetries/dilations?)

# Further work

- Use to design better DI protocols / apply to different DI tasks. Can we include preprocessing in DIQKD? [HST$^+$20, WAP20]
- Analyze finite round key rates (feasibility of DIQKD).
- Can we make the computations more efficient? (Symmetries/dilations?)
- What are the limiting cases as $\alpha \to 1$

$$\lim_{\alpha \to 1} D_{(\alpha)}(\rho \| \sigma) = ?$$

$$\lim_{\alpha \to 1} D^{\#}_{\alpha}(\rho \| \sigma) = ?$$

# Further work

- Use to design better DI protocols / apply to different DI tasks. Can we include preprocessing in DIQKD? [HST$^+$20, WAP20]
- Analyze finite round key rates (feasibility of DIQKD).
- Can we make the computations more efficient? (Symmetries/dilations?)
- What are the limiting cases as $\alpha \to 1$

$$\lim_{\alpha \to 1} D_{(\alpha)}(\rho \| \sigma) = ?$$

$$\lim_{\alpha \to 1} D^{\#}_{\alpha}(\rho \| \sigma) = ?$$

- Other applications to $\widetilde{D}_\alpha$?

# Further work

- Use to design better DI protocols / apply to different DI tasks. Can we include preprocessing in DIQKD? [HST$^+$20, WAP20]
- Analyze finite round key rates (feasibility of DIQKD).
- Can we make the computations more efficient? (Symmetries/dilations?)
- What are the limiting cases as $\alpha \to 1$

$$\lim_{\alpha \to 1} D_{(\alpha)}(\rho \| \sigma) = ?$$

$$\lim_{\alpha \to 1} D^{\#}_{\alpha}(\rho \| \sigma) = ?$$

- Other applications to $\widetilde{D}_\alpha$?
- Can we construct other families in a similar way?

# Bibliography


Frédéric Dupuis, Omar Fawzi, and Renato Renner.

Entropy accumulation.

*arXiv preprint arXiv:1607.01796*, 2016.

Kun Fang, Omar Fawzi, Renato Renner, and David Sutter.

Chain rule for the quantum relative entropy.

*Phys. Rev. Lett.*, 124:100501, Mar 2020.

M Ho, P Sekatski, EY-Z Tan, R Renner, J-D Bancal, and N Sangouard.

Noisy preprocessing facilitates a photonic realization of device-independent quantum key distribution.

*Physical Review Letters*, 124(23):230502, 2020.

Miguel Navascués, Stefano Pironio, and Antonio Acín.

Bounding the set of quantum correlations.

*Physical Review Letters*, 98:010401, 2007.

Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani.

Device-independent quantum key distribution secure against collective attacks.

*New Journal of Physics*, 11(4):045021, 2009.

Ernest Y-Z Tan, René Schwonnek, Koon Tong Goh, Ignatius William Primaatmaja, and Charles C-W Lim.

Computing secure key rates for quantum key distribution with untrusted devices.

*arXiv preprint arXiv:1908.11372*, 2019.

Erik Woodhead, Antonio Acín, and Stefano Pironio.

Device-independent quantum key distribution based on asymmetric chsh inequalities.

*arXiv preprint arXiv:2007.16146*, 2020.

Mark M Wilde, Mario Berta, Christoph Hirche, and Eneet Kaur.

Amortized channel divergence for asymptotic quantum channel discrimination.

*Letters in Mathematical Physics*, 110(8):2277–2336, 2020.