# Parallelization of Sequential Quantum Channel Discrimination in the Non-Asymptotic Regime

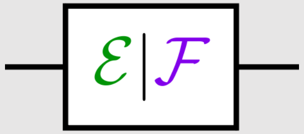
**Bjarne Bergh**, Nilanjana Datta, Robert Salzmann, Mark M. Wilde

BIID 2022

arXiv: 2206.08350

# The task

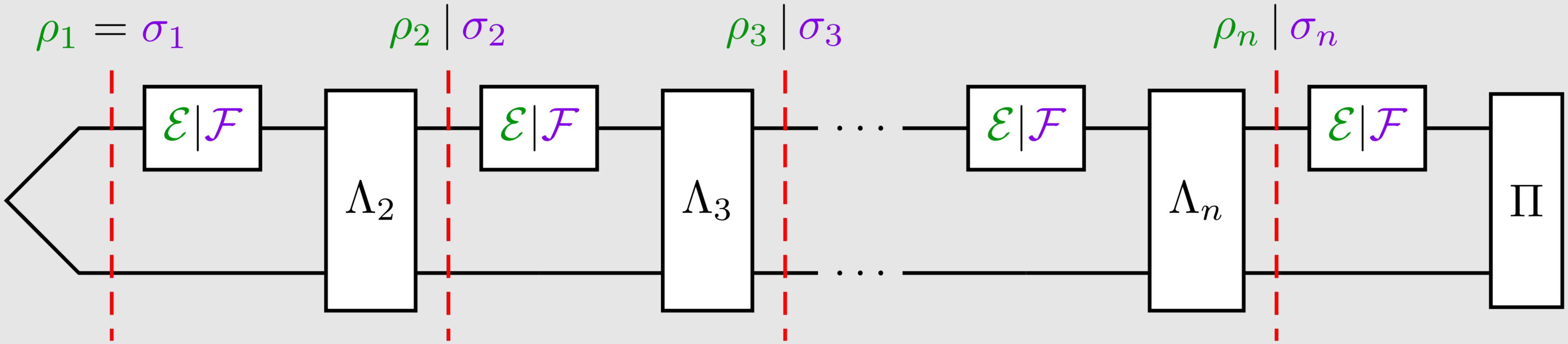
- Black-box promised to be one of two possible channels




- Find out which one

- You are allowed to use the black-box a finite number of times (n)

# The most general setup




- Input states will in general depend on previous channel outputs

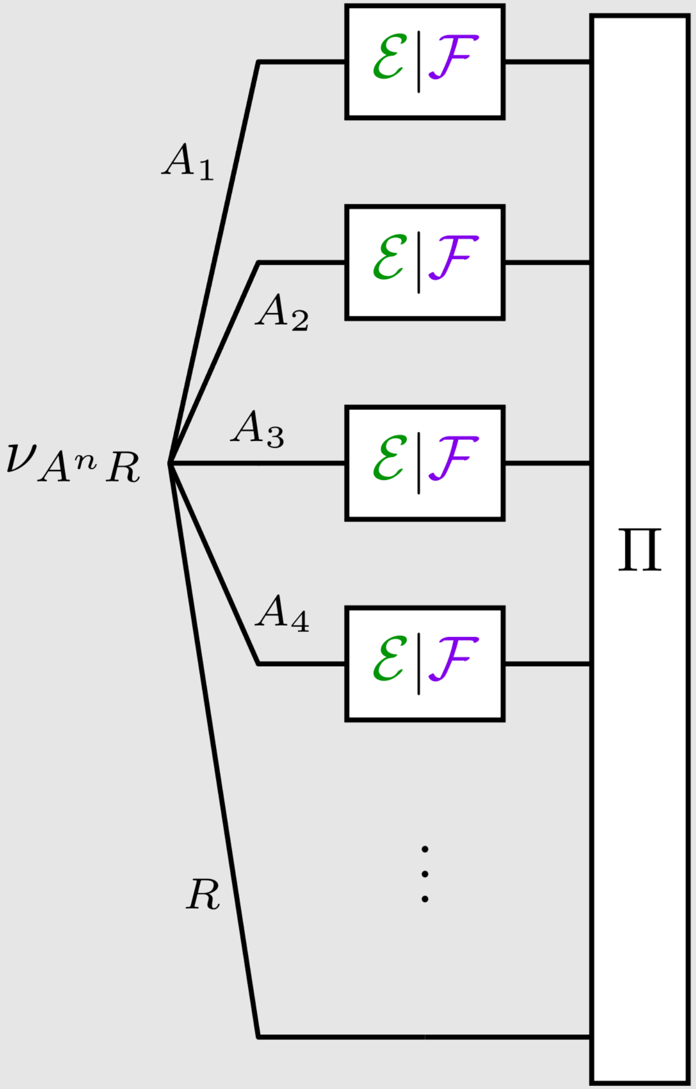
→ *adaptive strategy*

# Way simpler – A *parallel strategy*

Joint input state chosen at the beginning.

Reference system can be chosen isomorphic to $A^n$

Every parallel strategy can be written as an adaptive strategy.



# Interlude – Discriminating States

To discriminate between states $\rho$ and $\sigma$, do a binary POVM $\{\Pi, \mathbb{1} - \Pi\}$ which will have errors:

$$\alpha(\rho, \Pi) = \mathrm{Tr}(\rho(\mathbb{1} - \Pi))$$

$$\beta(\sigma, \Pi) = \mathrm{Tr}(\sigma\Pi)$$

# The asymmetric setting

Best type II error when type I error stays below $\varepsilon$:

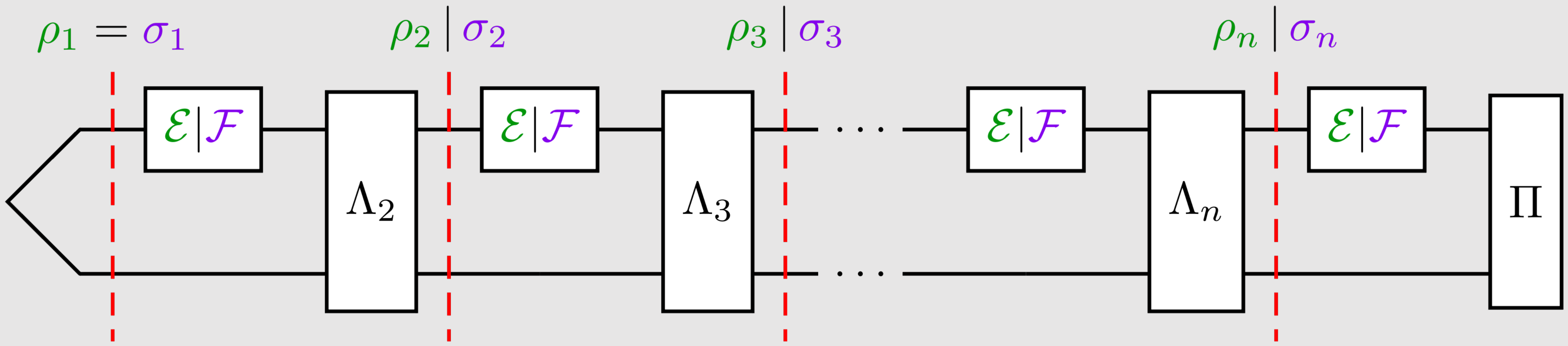
$$\beta(\varepsilon,\rho,\sigma) := \min\{\,\mathrm{Tr}(\Pi\sigma) \mid \Pi : 0 \le \Pi \le \mathbb{1}, \mathrm{Tr}(\rho(\mathbb{1}-\Pi)) \le \varepsilon\,\}$$

$$D_H^{\varepsilon}(\rho\|\sigma) := -\log\beta(\varepsilon,\rho,\sigma)$$

With n copies of the state: Type II error rate per used copy

$$\frac{1}{n}D_H^{\varepsilon}(\rho^{\otimes n}\|\sigma^{\otimes n}) = -\frac{1}{n}\log\beta(\varepsilon,\rho^{\otimes n},\sigma^{\otimes n})$$

# Error Exponents – Adaptive Strategy




Error rate per channel use of this strategy:

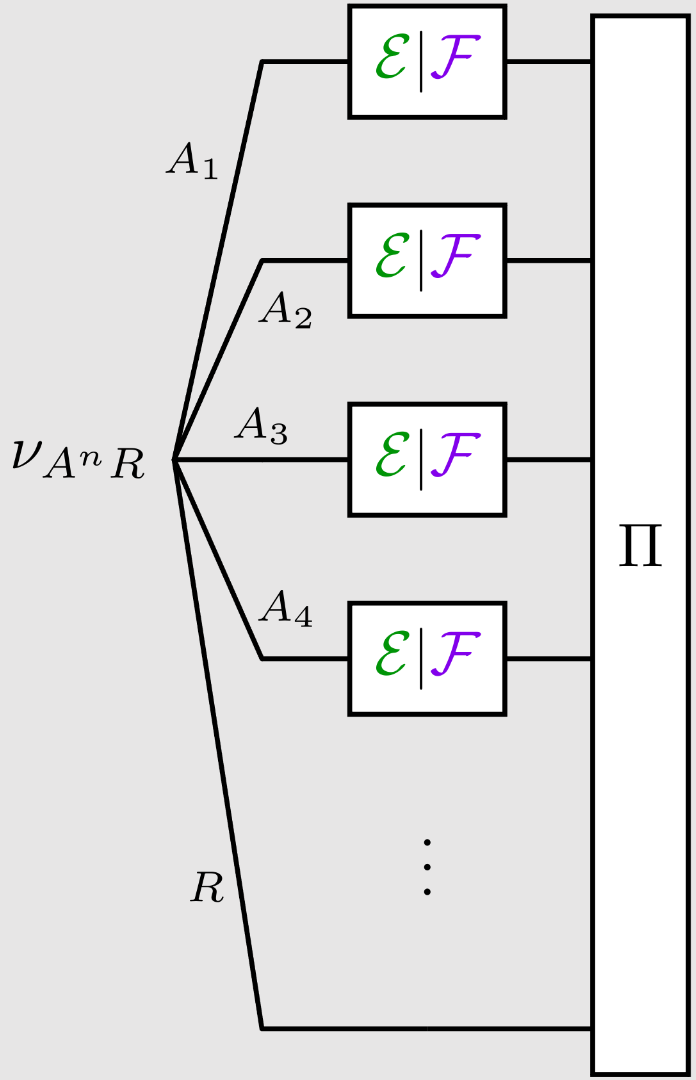
$$\frac{1}{n} D_H^{\varepsilon}((\mathrm{id}_R \otimes \mathcal{E})(\rho_n) \| (\mathrm{id}_R \otimes \mathcal{F})(\sigma_n))$$

# Error Exponents – Parallel strategy

$$\frac{1}{n} D_H^{\varepsilon}\big((\mathrm{id}_R \otimes \mathcal{E}^{\otimes n})(\nu) \| (\mathrm{id}_R \otimes \mathcal{F}^{\otimes n})(\nu)\big)$$

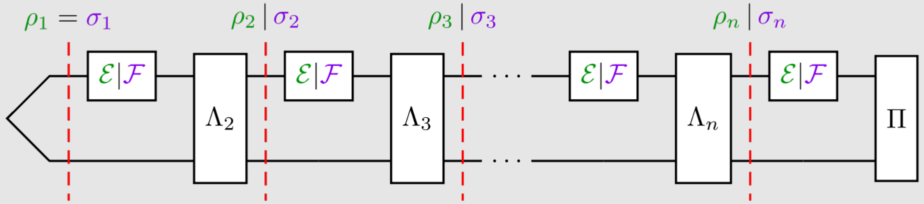
From now on, drop identities, i.e. write:

$$\frac{1}{n} D_H^{\varepsilon}\big(\mathcal{E}^{\otimes n}(\nu) \| \mathcal{F}^{\otimes n}(\nu)\big)$$



# What was known already – Stein Exponents

Optimization over all sequential strategies




$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\sup_{\rho_1,\{\Lambda_i\}}\frac{1}{n}D_H^{\varepsilon}(\mathcal{E}(\rho_n)\|\mathcal{F}(\sigma_n)) = D_A(\mathcal{E}\|\mathcal{F})$$

$$= D_{\text{reg}}(\mathcal{E}\|\mathcal{F}) = \lim_{\varepsilon\to 0}\lim_{m\to\infty}\sup_{\nu}\frac{1}{m}D_H^{\varepsilon}(\mathcal{E}^{\otimes m}(\nu)\|\mathcal{F}^{\otimes m}(\nu))$$

Best achievable type II error rate such that the type I error goes to zero as $n \to \infty$.

Optimization over all parallel strategies


Fang, Fawzi, Renner, Sutter. A chain rule for the quantum relative entropy. *Phys. Rev. Lett.* **124**, 100501 (2020).
Wang, Wilde. Resource theory of asymmetric distinguishability for quantum channels. *Phys. Rev. Research* **1**, 033169 (2019).

# The question

- What happens at finite n?

- How many channel uses does a parallel strategy need to achieve a similar rate as a (given) adaptive strategy?

# Our result

**Theorem.** Let $\mathcal{E}, \mathcal{F}$ s.t. $D_{\max}(\mathcal{E}\|\mathcal{F}) < \infty$. Given $\{\rho_i, \sigma_i\}_{i=1}^n$, $\forall \alpha_S, \alpha_P \in (0,1)$, $\forall m \in \mathbb{N} : \exists \nu \in \mathcal{D}\left(R^m A^m\right)$ :

$$\frac{1}{m} D_H^{\alpha_P}\left(\mathcal{E}^{\otimes m}(\nu) \| \mathcal{F}^{\otimes m}(\nu)\right) \geq \frac{1-\alpha_S}{n} D_H^{\alpha_S}(\mathcal{E}(\rho_n) \| \mathcal{F}(\sigma_n)) - \frac{Cn}{\sqrt{m}} \log\left(\frac{8}{\alpha_P}\right) - \frac{1}{n}$$

with:

$$C := 7 \log\left(2^{D_2(\mathcal{E}\|\mathcal{F})} + 2\right) \leq 7 \log\left(2^{D_{\max}(\mathcal{E}\|\mathcal{F})} + 2\right)$$

# Prerequisites – A chain rule

$$D_{\max}(\rho\|\sigma) := \log\inf\{\, \lambda \in \mathbb{R} \mid \rho \leq \lambda\sigma \,\}$$

$$D_{\max}(\mathcal{E}(\rho)\|\mathcal{F}(\sigma)) \leq D_{\max}(\rho\|\sigma) + D_{\max}(\mathcal{E}(\rho)\|\mathcal{F}(\rho))$$

$$D^{\varepsilon}_{\max}(\rho\|\sigma) := \inf_{\tilde{\rho}\in B^{\circ}_{\varepsilon}(\rho)} D_{\max}(\tilde{\rho}\|\sigma)$$

ε-ball of *normalized* states in purified distance

$$D^{\varepsilon+\varepsilon'}_{\max}(\mathcal{E}(\rho)\|\mathcal{F}(\sigma)) \leq D^{\varepsilon}_{\max}(\rho\|\sigma) + D^{\varepsilon'}_{\max}(\mathcal{E}(\nu)\|\mathcal{F}(\nu))$$

# Prerequisites – A chain rule

$$D_{\max}(\rho\|\sigma) := \log \inf \{ \lambda \in \mathbb{R} \mid \rho \leq \lambda\sigma \}$$

$$D^{\varepsilon}_{\max}(\rho\|\sigma) := \inf_{\tilde{\rho} \in B^{\circ}_{\varepsilon}(\rho)} D_{\max}(\tilde{\rho}\|\sigma)$$

$$D^{\varepsilon+\varepsilon'}_{\max}(\mathcal{E}(\rho)\|\mathcal{F}(\sigma)) \leq D^{\varepsilon}_{\max}(\rho\|\sigma) + D^{\varepsilon'}_{\max}(\mathcal{E}(\nu)\|\mathcal{F}(\nu))$$

$$\nu \leq 2^{D^{\varepsilon}_{\max}(\rho\|\sigma)}\sigma \quad \Rightarrow \quad \mathcal{F}(\nu) \leq 2^{D^{\varepsilon}_{\max}(\rho\|\sigma)}\mathcal{F}(\sigma)$$

$$\tau \leq 2^{D^{\varepsilon'}_{\max}(\mathcal{E}(\nu)\|\mathcal{F}(\nu))}\mathcal{F}(\nu)$$

$$\tau \leq 2^{D^{\varepsilon'}_{\max}(\mathcal{E}(\nu)\|\mathcal{F}(\nu)) + D^{\varepsilon}_{\max}(\rho\|\sigma)}\mathcal{F}(\sigma)$$

$$P(\tau, \mathcal{E}(\rho)) \leq P(\tau, \mathcal{E}(\nu)) + P(\mathcal{E}(\nu), \mathcal{E}(\rho)) \leq \varepsilon' + P(\nu, \rho) \leq \varepsilon' + \varepsilon\,.$$

Related previous result: Fang, Fawzi, Renner, Sutter. A chain rule for the quantum relative entropy. *Phys. Rev. Lett.* **124**, 100501 (2020).

# Prerequisites – Hypothesis Test. Rel. Entropy

$$D_{\max}^{\varepsilon}(\rho\|\sigma) \leq D_{H}^{1-\varepsilon^2}(\rho\|\sigma) - \log\big(1 - \varepsilon^2\big)$$

Anshu, Berta, Jain, Tomamichel, A Minimax Approach to One-Shot Entropy Inequalities. *J. Math. Phys.* **2019**, *60* (12), 122201.

$$D_{H}^{\varepsilon}(\rho\|\sigma) \leq \frac{1}{1-\varepsilon}\big(D(\rho\|\sigma) + h(\varepsilon)\big)$$

Well-known, e.g. Hiai and Petz (1991), Hayashi (2006), Wang and Renner (2012)

# Prerequisites – AEP

$$\frac{1}{n} D^{\varepsilon}_{\max}(\rho^{\otimes n} \| \sigma^{\otimes n}) \geq D(\rho\|\sigma) - \frac{4c_\gamma(\rho\|\sigma)}{\sqrt{n}} \log\left(\frac{2}{1-\varepsilon}\right)$$

$$\frac{1}{n} D^{\varepsilon}_{\max}(\rho^{\otimes n} \| \sigma^{\otimes n}) \leq D(\rho\|\sigma) + \frac{4c_\gamma(\rho\|\sigma)}{\sqrt{n}} \log\left(\frac{2}{\varepsilon}\right) + \frac{1}{n} \log\left(\frac{1}{1-\varepsilon^2}\right)$$

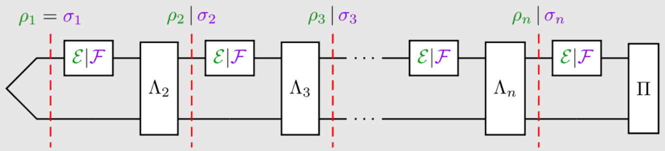
$$c_\gamma(\rho\|\sigma) := \frac{1}{\gamma} \log\left(2^{\gamma D_{1+\gamma}(\rho\|\sigma)} + 2^{-\gamma D_{1-\gamma}(\rho\|\sigma)} + 1\right) \qquad \gamma \in (0,1]$$

Related previous results: Tomamichel, A framework for non-asymptotic quantum information theory. (ETH Zurich, 2012).

# Proof of our Theorem

$$D_H^{\varepsilon}(\rho\|\sigma) \leq \frac{1}{1-\varepsilon}\big(D(\rho\|\sigma)+h(\varepsilon)\big)$$

$$\frac{1-\alpha_S}{n} D_H^{\alpha_S}(\mathcal{E}(\rho_n)\|\mathcal{F}(\sigma_n)) \leq \frac{1}{n} D(\mathcal{E}(\rho_n)\|\mathcal{F}(\sigma_n)) + \dots$$




$$\leq \frac{1}{n}\sum_{k=1}^{n}\left[D(\mathcal{E}(\rho_k)\|\mathcal{F}(\sigma_k)) - D(\rho_k\|\sigma_k)\right] + \dots$$

$$\leq D(\mathcal{E}(\rho_\ell)\|\mathcal{F}(\sigma_\ell)) - D(\rho_\ell\|\sigma_\ell) + \dots$$

$$D^{\varepsilon}_{\max}(\rho\|\sigma) \leq D_H^{1-\varepsilon^2}(\rho\|\sigma) - \log\big(1-\varepsilon^2\big)$$

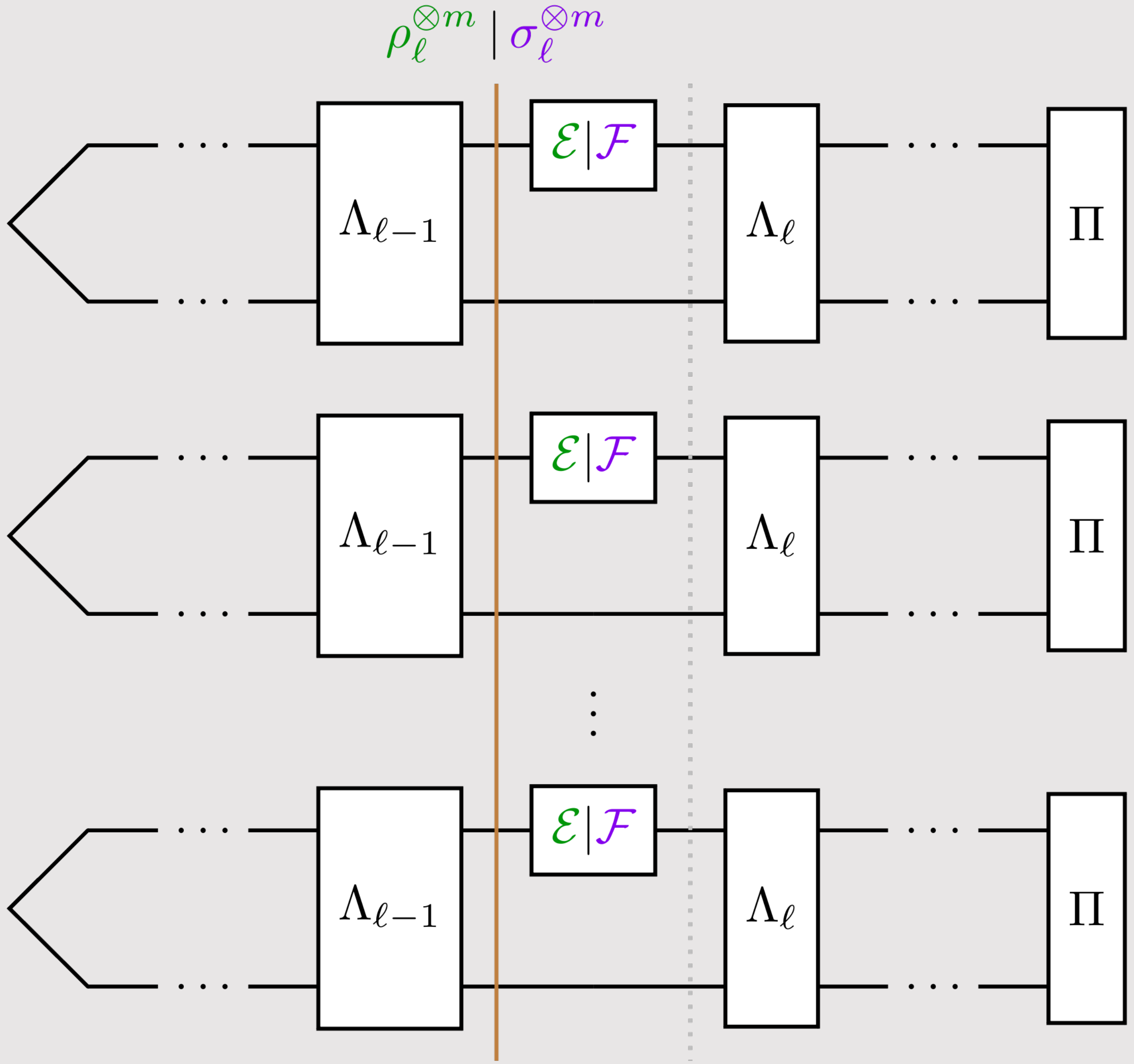
$$\frac{1}{n} D^{\varepsilon}_{\max}(\rho^{\otimes n}\|\sigma^{\otimes n}) \geq D(\rho\|\sigma) - \dots$$

$$\frac{1}{n} D^{\varepsilon}_{\max}(\rho^{\otimes n}\|\sigma^{\otimes n}) \leq D(\rho\|\sigma) + \dots$$

$$\leq \frac{1}{m}\left[D^{\varepsilon''}_{\max}(\mathcal{E}^{\otimes m}(\rho_\ell^{\otimes m})\|\mathcal{F}^{\otimes m}(\sigma_n^{\otimes m})) - D^{\varepsilon'}_{\max}(\rho_\ell^{\otimes m}\|\sigma_\ell^{\otimes m})\right] + \dots$$

$$D^{\varepsilon+\varepsilon'}_{\max}(\mathcal{E}(\rho)\|\mathcal{F}(\sigma)) \leq D^{\varepsilon}_{\max}(\rho\|\sigma) + D^{\varepsilon'}_{\max}(\mathcal{E}(\nu)\|\mathcal{F}(\nu))$$

$$\leq \frac{1}{m} D^{\varepsilon}_{\max}(\mathcal{E}^{\otimes m}(\nu)\|\mathcal{F}^{\otimes m}(\nu)) + \dots \leq \frac{1}{m} D_H^{\alpha_P}(\mathcal{E}^{\otimes m}(\nu)\|\mathcal{F}^{\otimes m}(\nu)) + \dots$$




$$\nu \approx^{\alpha_P} \rho_\ell^{\otimes m} \quad \text{s.t.} \quad D_{\max}(\nu \| \sigma_\ell^{\otimes m}) = \text{minimal}$$

# Example

$$\mathcal{E}, \mathcal{F} : \mathcal{B}\left(\mathbb{C}^2 \otimes \mathbb{C}^2\right) \to \mathcal{B}\left(\mathbb{C}^2\right)$$

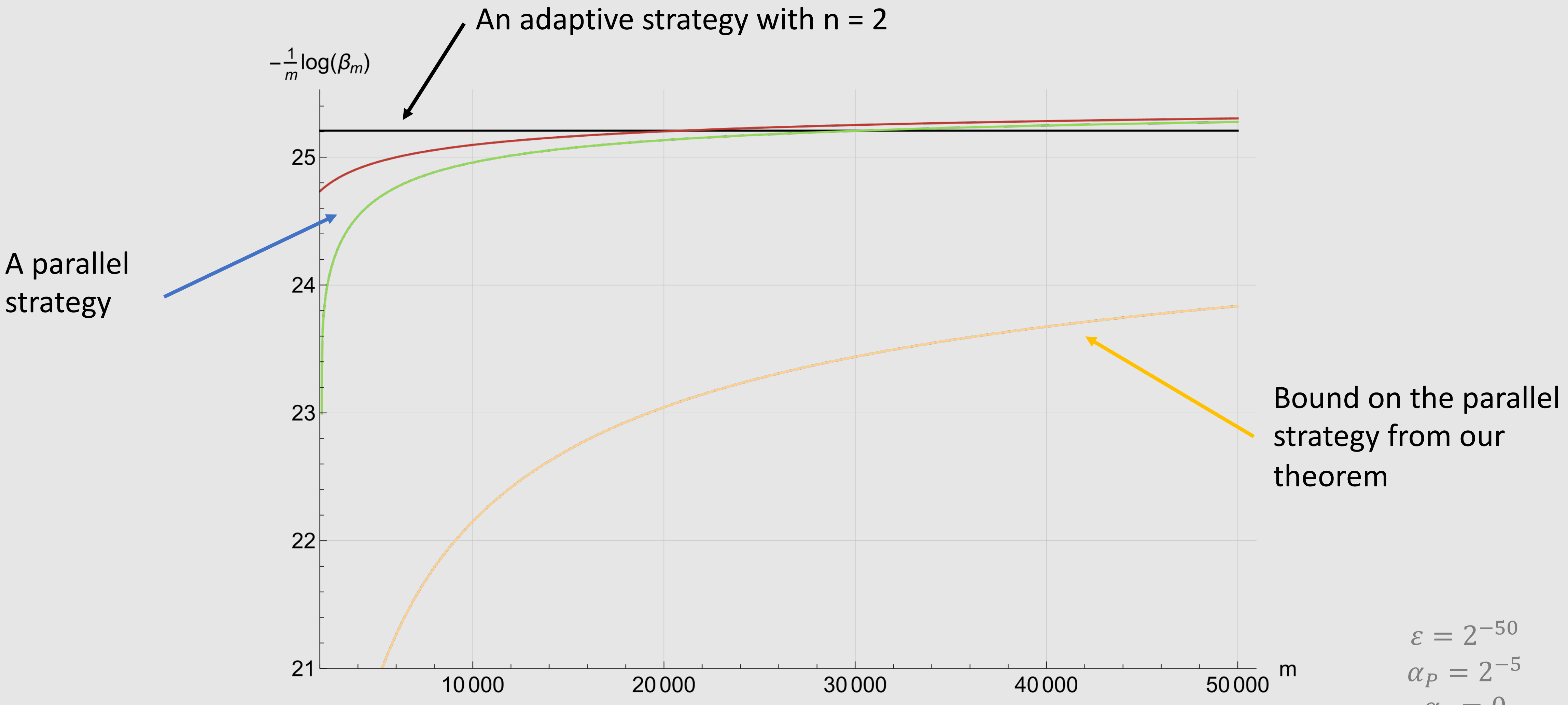
$$\mathcal{E}(\rho \otimes \omega) \coloneqq |0\rangle\langle 0| \langle 0|\omega|0\rangle + |0\rangle\langle 0| \langle 0|\rho|0\rangle \langle 1|\omega|1\rangle + \frac{\mathbb{1}}{2} \langle 11|\rho \otimes \omega|11\rangle$$

$$\mathcal{F}(\rho \otimes \omega) \coloneqq (1-\varepsilon)\left[|+\rangle\langle +| \langle 0|\omega|0\rangle + |1\rangle\langle 1| \langle +|\rho|+\rangle \langle 1|\omega|1\rangle + \frac{\mathbb{1}}{2} \langle -1|\rho \otimes \omega| -1\rangle\right]$$
$$+ \varepsilon \frac{\mathbb{1}}{2} \operatorname{Tr}[\rho \otimes \omega]$$

$\mathcal{E}$ and $\mathcal{F}$ almost perfectly distinguishable *adaptively* with just two uses

Slightly adapted from: Harrow, A. W., Hassidim, A., Leung, D. W. & Watrous, J.
Adaptive versus nonadaptive strategies for quantum channel discrimination. *Phys. Rev. A* **81**, 032339 (2010).

# Example




$\varepsilon = 2^{-50}$
$\alpha_P = 2^{-5}$
$\alpha_S = 0$

# Open questions

- Do we really need $\mathcal{O}(n^2)$ parallel channel uses to achieve the rate of a sequential strategy with $n$ uses?
  - Helpful would be a convergence bound on $D_{reg}$ or $D_A$
- Second order expansions for channel discrimination tasks
  - Requires controlling $V(\rho_n \| \sigma_n) = \mathcal{O}(n)$ or $V\big(\mathcal{E}^{\otimes n}(\nu) \big\| \mathcal{F}^{\otimes n}(\nu)\big) = \mathcal{O}(n)$
  - Would prove strong converse