The Role of Statistical Comparison Theory in the Study of Open Quantum Systems

Francesco Buscemi (Nagoya U)

Advances in open systems and fundamental tests of quantum mechanics

684th WE-Heraeus-Seminar, Bad Honnef (Germany), 4 December 2018

Classical Markov chains: some nomenclature

Time convention: $t_0 \leq t_1 \leq \cdots \leq t_N$.

- classical Markov chain:

$$P(\boldsymbol{x}_{t_i}|\boldsymbol{x}_{t_{i-1}},\boldsymbol{x}_{t_{i-2}},\ldots,\boldsymbol{x}_{t_0}) = P(\boldsymbol{x}_{t_i}|\boldsymbol{x}_{t_{i-1}}), \quad \forall i \in [1,N]$$

- **physical divisibility** (Markov equation):

$$P(\boldsymbol{x}_{t_N},\boldsymbol{x}_{t_{N-1}},\ldots,\boldsymbol{x}_{t_0}) = P(\boldsymbol{x}_{t_N}|\boldsymbol{x}_{t_{N-1}})\cdots P(\boldsymbol{x}_{t_1}|\boldsymbol{x}_{t_0})P(\boldsymbol{x}_{t_0})$$

- **stochastic divisibility** (Chapman-Kolmogorov equation):

$$P(\boldsymbol{x}_{t_k}|\boldsymbol{x}_{t_i}) = \sum_{\boldsymbol{x}_{t_j}} P(\boldsymbol{x}_{t_k}|\boldsymbol{x}_{t_j})P(\boldsymbol{x}_{t_j}|\boldsymbol{x}_{t_i}), \quad \forall k \geq j \geq i$$

- physical divisibility $\underset{\not\Longleftarrow}{\Longrightarrow}$ stochastic divisibility

The Problem with Quantum Systems

Quantum stochastic processes are like sealed black boxes: an observation at some time t_1 generally disturbs the process thus "spoiling" any subsequent observation made at later times $t_2 \geq t_1$.

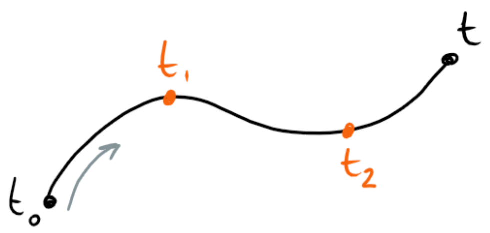


Figure 1: Here t_0 is an initial time, at which the quantum system can be prepared (fully controlled). There is no *direct* quantum analogue of the N-time joint distribution $P(\boldsymbol{x}_{t_N}, \ldots, \boldsymbol{x}_{t_0})$.

Quantum Dynamical Mappings

How to describe *quantum* stochastic processes then?

- time convention: $t_0 \leq t_1 \leq \cdots \leq t_N$
- open quantum systems formalism:
  $\rho_S(t_i) := \mathrm{Tr}_E \left\{ U_{t_0 \to t_i} \left[\rho_S(0) \otimes \rho_E(0) \right] U^\dagger_{t_0 \to t_i} \right\}$
- if the system is fully controlled at t_0, we obtain a sequence of CPTP linear maps by discarding the bath:
  $\Phi_i(\bullet_S) := \mathrm{Tr}_E \left\{ U_{t_0 \to t_i} \left[\bullet_S \otimes \rho_E(0) \right] U^\dagger_{t_0 \to t_i} \right\}$

Definition

A quantum dynamical mapping (QDM) is a sequence of CPTP linear maps $(\Phi_i)_{0 \leq i \leq N}$ satisfying $\Phi_0 = \mathrm{id}_S$ (consistency condition).

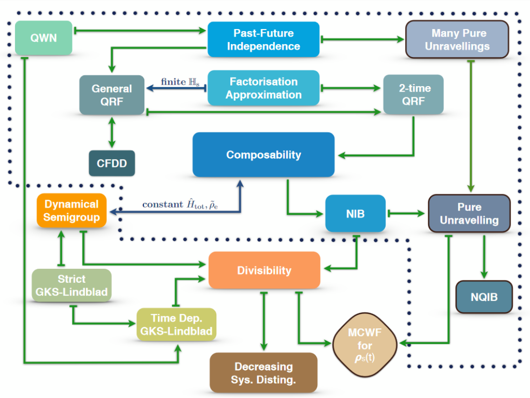
Two Approaches to Quantum Markoviania

- **Global (extrinsic) picture**: Markovianity is a property of the whole system+bath compound (like, e.g., singular coupling regime, approximate factorizability, etc)

- **Reduced (intrinsic) picture**: Markovianity is a property of the resulting quantum dynamical mapping alone (like, e.g., information decrease, divisibility, etc)

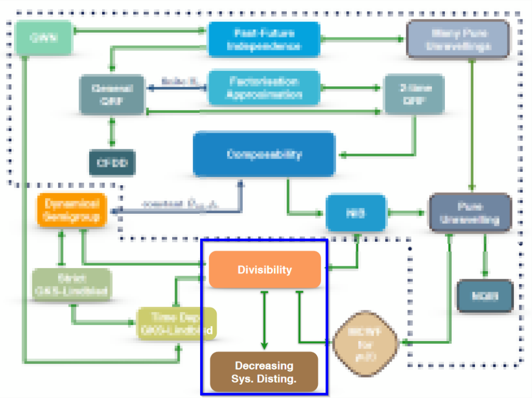
A “Zoo” of Quantum Markovianities

From: Li Li, Michael J. W. Hall, Howard M. Wiseman. *Concepts of quantum non-Markovianity: a hierarchy*. (arXiv:1712.08879 [quant-ph])



A “Zoo” of Quantum Markovianities

From: Li Li, Michael J. W. Hall, Howard M. Wiseman. *Concepts of quantum non-Markovianity: a hierarchy*. (arXiv:1712.08879 [quant-ph])



Decreasing System Distinguishability (DSD)

- introduced in [Breuer, Laine, Piilo; PRL 2009], it provides the bridge between physical and information-theoretic Markovianity
- for any pair of possible initial states of the system, say, $\rho_S^1(0)$ and $\rho_S^2(0)$, consider the same pair evolved at later times $t_i > t_0$:

$$\rho_S^{1,2}(t_i) := \Phi_i\left[\rho_S^{1,2}(0)\right]$$

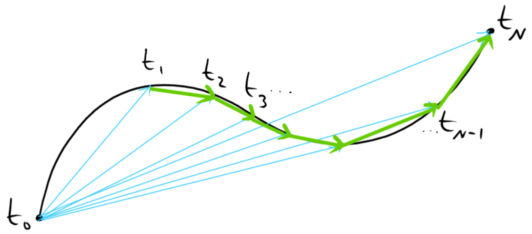
- DSD condition:

$$\|\rho_S^1(t_i) - \rho_S^2(t_i)\|_1 \geq \|\rho_S^1(t_j) - \rho_S^2(t_j)\|_1 \ , \quad \forall i \leq j$$

- **interpretation**: revival of distinguishability $\implies$ back-flow of information $\implies$ memory effects $\implies$ non-Markovianity

Divisibility (DIV)

- extends the idea of dynamical semigroups: $t \mapsto \Phi_t$ such that $\Phi_s \circ \Phi_t = \Phi_{t+s}$
- a QDM $(\Phi_i)_i$ is CPTP divisible if there exist CPTP linear maps $(\mathcal{E}_{i\to j})_{i\le j}$, which we call propagators, such that $\Phi_j = \mathcal{E}_{i\to j} \circ \Phi_i$, for all $0 \le i \le j \le N$



- DIV constitutes a quantum analogue of the Chapman-Kolmogorov equation (i.e., stochastic divisibility)
- very well captures the property of being memoryless, which is a crucial (*the* crucial?) property of Markovian processes

$$\mathrm{DIV} \underset{\not\Leftarrow}{\Longrightarrow} \mathrm{DSD}$$

can we make these equivalent?

DSD, DIV, and Reverse Data-Processing Theorems

- DIV is equivalent to the property of **degradability**: channel Φ is said to be degradable into channel Φ' whenever there exists a third channel $\mathcal{E}$ such that $\Phi' = \mathcal{E} \circ \Phi$
- hence, "DIV $\implies$ DSD" is a consequence of the **data-processing inequality for the trace norm**: for any pair of states (ρ_S^1, ρ_S^2),

$$\begin{aligned} \|\Phi'(\rho_S^1) - \Phi'(\rho_S^2)\|_1 &= \|(\mathcal{E} \circ \Phi)(\rho_S^1) - (\mathcal{E} \circ \Phi)(\rho_S^2)\|_1 \\ &\leq \|\Phi(\rho_S^1) - \Phi(\rho_S^2)\|_1 \end{aligned}$$

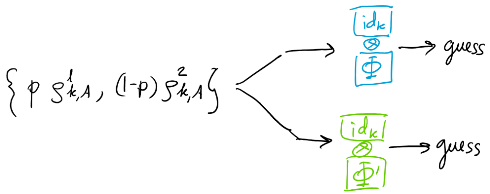
- in fact, the data-processing inequality is satisfied by *most* (all?) distinguishability measures
- hence, it is interesting to seek for possible **alternative (stronger) definitions of DSD**, maintaining the same "intuitive meaning", but leading to the sought after equivalence: DSD $\iff$ DIV

Strengthening DSD

Theorem (Chruściński, Kossakowski, and Rivas, 2011; Chruściński and Maniscalco, 2014; Wißman, Breuer, Vacchini, 2015)

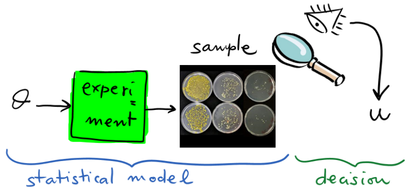
Let $\Phi : A \to B$ and $\Phi' : A \to B'$ be two quantum channels, with Φ invertible (as a linear map). Then, Φ is degradable into Φ' with a k-positive TP map $\mathcal{E} : B \to B'$, if and only if, for all $p \in [0,1]$ and all pairs of k-extended states $\rho^1_{k,A}, \rho^2_{k,A} \in \mathsf{L}(\mathbb{C}^k \otimes \mathcal{H}_A)$,

$$
\begin{aligned}
&\|p(\mathrm{id}_k \otimes \Phi')(\rho^1_{k,A}) - (1-p)(\mathrm{id}_k \otimes \Phi')(\rho^2_{k,A})\|_1 \\
&\leq \|p(\mathrm{id}_k \otimes \Phi)(\rho^1_{k,A}) - (1-p)(\mathrm{id}_k \otimes \Phi)(\rho^1_{k,A})\|_1
\end{aligned}
$$



Statistical Distinguishability in Mathematical Statistics

Statistical Models and Decision Problems



$$\begin{tikzcd}[column sep=large] \Theta \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{X} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{U} \arrow[d, squiggly] \\ \theta \arrow[r, "w(x|\theta)"'] & x \arrow[r, "d(u|x)"'] & u \end{tikzcd}$$

Formal Definitions

- A statistical model is given by: a parameter set Θ, a sample set $\mathcal{X}$, and a family of PDs $\{w_\theta(x)\} \equiv w(x|\theta)$.
- A statistical decision problem is given by: a parameter set Θ, an "action" set $\mathcal{U}$, and a payoff function $\ell : \Theta \times \mathcal{U} \to \mathbb{R}$.

How Much Is a Statistical Model Worth?

Each decision problem implicitly defines a **statistical distinguishability measure** for the PDs $\{w_\theta(x)\}$.

- the model $\mathbf{w} = \{w_\theta(x)\}$ represents info in X about θ
- the decision $d(u|x)$ optimally extracts from X information about θ, and uses this to decide the best action

$$\begin{array}{ccccc} \Theta & \xrightarrow{\text{experiment}} & \mathcal{X} & \xrightarrow{\text{decision}} & \mathcal{U} \\ \wr & & \wr & & \wr \\ \theta & \xrightarrow[w(x|\theta)]{} & x & \xrightarrow[d(u|x)]{} & u \end{array}$$

Definition (Expected Payoff)

The expected payoff of statistical model $\mathbf{w} = \{w_\theta(x)\}$ w.r.t. decision problem $\boldsymbol{\ell} = \{\ell(\theta, u)\}$ is given by

$$\mathbb{E}_{\boldsymbol{\ell}}[\mathbf{w}] \triangleq \max_{d(u|x)} \sum_{u,x,\theta} \ell(\theta, u) d(u|x) w(x|\theta) |\Theta|^{-1} .$$

Intuition: the higher the payoff, the more information about θ the PDs $w_\theta(x)$ carry, **the more "distinguishable" they are**.

Comparing Statistical Models 1/2

First model: $\mathbf{w} = \{w_\theta(x)\}$

$$\begin{tikzcd}
\Theta \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{X} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{U} \arrow[d, squiggly] \\
\theta \arrow[r, "w(x|\theta)"'] & x \arrow[r, "d(u|x)"'] & u
\end{tikzcd}$$

Second model: $\mathbf{w}' = \{w'_\theta(y)\}$

$$\begin{tikzcd}
\Theta \arrow[r, "\text{experiment}"] \arrow[d, squiggly] & \mathcal{Y} \arrow[r, "\text{decision}"] \arrow[d, squiggly] & \mathcal{U} \arrow[d, squiggly] \\
\theta \arrow[r, "w'(y|\theta)"'] & y \arrow[r, "d'(u|y)"'] & u
\end{tikzcd}$$

Given a statistical decision problem $\boldsymbol{\ell} = \{\ell(\theta, u)\}$, if $\mathbb{E}_{\boldsymbol{\ell}}[\mathbf{w}] \geq \mathbb{E}_{\boldsymbol{\ell}}[\mathbf{w}']$, then one says that model $\mathbf{w}$ is "more informative" (or "broader", or "more distinguishable") than model $\mathbf{w}'$, with respect to problem $\boldsymbol{\ell}$.

Comparing Statistical Models 2/2

Definition (Information Preorder)

If model $\mathbf{w} = \{w_\theta(x)\}$ is more informative than model $\mathbf{w}' = \{w'_\theta(y)\}$ **for all decision problems $\boldsymbol{\ell} = \{\ell(\theta, u)\}$**, then we say that $\mathbf{w}$ is *(always) more informative* than $\mathbf{w}'$, and write

$$\mathbf{w} \succeq \mathbf{w}' \ .$$

Intuition: $\mathbf{w} \succeq \mathbf{w}'$ means that the PDs $\{w_\theta(x)\}$ are always more distinguishable than $\{w'_\theta(y)\}$.

Problem. The information preorder is operational, but not really "concrete". Can we visualize this better?

The Fundamental Equivalence

Blackwell-Sherman-Stein (1948-1953)

Given two statistical models $\mathbf{w} = \{w_\theta(x)\}$ and $\mathbf{w}' = \{w'_\theta(y)\}$, the following are equivalent:

1. **w** is more informative than **w**$'$, i.e., $\mathbf{w} \succeq \mathbf{w}'$;
2. **w** is sufficent for **w**$'$, i.e., there exists a conditional PD $\varphi(y|x)$ such that $w'(y|\theta) = \sum_x \varphi(y|x) w(x|\theta)$.

$$\begin{array}{ccccccccc}
\Theta & \longrightarrow & \mathcal{Y} & & \Theta & \longrightarrow & \mathcal{X} & \xrightarrow{\text{noise}} & \mathcal{Y} \\
\wr\downarrow & & \wr\downarrow & = & \wr\downarrow & & \wr\downarrow & & \wr\downarrow \\
\theta & \xrightarrow[w'(y|\theta)]{} & y & & \theta & \xrightarrow[w(x|\theta)]{} & x & \xrightarrow[\varphi(y|x)]{} & y
\end{array}$$

David H. Blackwell (1919-2010)

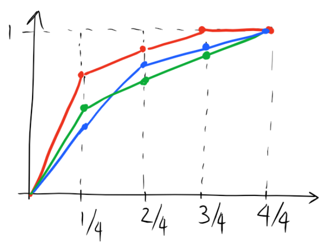
statistical sufficiency

$$\approx$$

better distinguishability w.r.t. all operational distinguishability measures

Paramount Example: Majorization and Lorenz Curves

- two probability distributions, $\boldsymbol{p}$ and $\boldsymbol{q}$, of the same dimension n
- truncated sums $P(k) = \sum_{i=1}^{k} p_i^{\downarrow}$ and $Q(k) = \sum_{i=1}^{k} q_i^{\downarrow}$, for all $k = 1, \ldots, n$
- $\boldsymbol{p}$ majorizes $\boldsymbol{q}$, i.e., $\boldsymbol{p} \succeq \boldsymbol{q}$, whenever $P(k) \geq Q(k)$, for all k
- minimal element: uniform distribution $\boldsymbol{e} = n^{-1}(1, 1, \cdots, 1)$



$(x_k, y_k) = (k/n, P(k)), \quad 1 \leq k \leq n$

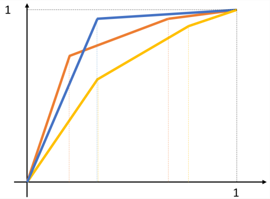
Hardy, Littlewood, and Pólya (1934)

$\boldsymbol{p} \succeq \boldsymbol{q} \iff \boldsymbol{q} = M\boldsymbol{p}$, for some bistochastic matrix M.

Intuition: $\boldsymbol{p} \succeq \boldsymbol{q}$ means that $\boldsymbol{p}$ is always more distinguishable than $\boldsymbol{q}$ from the uniform $\boldsymbol{e}$.

Generalization: Relative Majorization

- two *pairs* of probability distributions, $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ and $(\boldsymbol{q}_1, \boldsymbol{q}_2)$, of dimension m and n, respectively
- relabel entries such that ratios p_1^i/p_2^i and q_1^j/q_2^j are nonincreasing
- construct the truncated sums $P_{1,2}(k) = \sum_{i=1}^k p_{1,2}^i$ and $Q_{1,2}(k)$
- $(\boldsymbol{p}_1, \boldsymbol{p}_2) \succeq (\boldsymbol{q}_1, \boldsymbol{q}_2)$ iff the curve of the former is never below that of the latter



Relative Lorenz curves:

$(x_k, y_k) = (P_2(k), P_1(k))$

Blackwell Thm for Dichotomies, 1953

$(\boldsymbol{p}_1, \boldsymbol{p}_2) \succeq (\boldsymbol{q}_1, \boldsymbol{q}_2) \iff \boldsymbol{q}_i = M\boldsymbol{p}_i$, for some stochastic matrix M.

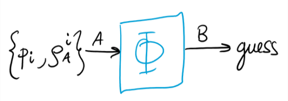
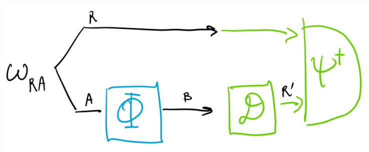
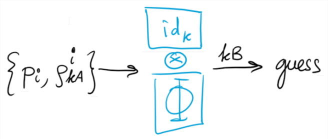
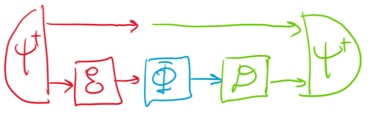
Intuition: $(\boldsymbol{p}_1, \boldsymbol{p}_2) \succeq (\boldsymbol{q}_1, \boldsymbol{q}_2)$ means that $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ are always more distinguishable than $(\boldsymbol{q}_1, \boldsymbol{q}_2)$.

observation: discrete noisy channels

$$\Phi : \begin{cases} \mathcal{X} \to \mathcal{Y} \\ x \mapsto p_x(y) \end{cases}$$

are equivalent to statistical models

Statistical Distinguishability Measures for Noisy Channels

<table>
<tr><th>guessing problems</th><th>quantum decoding problems</th></tr>
<tr><td>simple guess
</td><td>singlet extraction
</td></tr>
<tr><td>extended guess
</td><td>encoding-decoding
</td></tr>
</table>

Ordering Channels by Guessing Problems

- given two quantum channels (CPTP linear maps) $\Phi : A \to B$ and $\Phi' : A \to B'$, we say that Φ is less noisy than Φ', i.e., $\Phi \supseteq \Phi'$, whenever, for any input ensemble $\{p_x, \rho_A^x\}$,

$$P_{\text{guess}}(\{p_x, \Phi(\rho_A^x)\}) \geq P_{\text{guess}}(\{p_x, \Phi'(\rho_A^x)\})$$

- $\Phi \supseteq_k \Phi' \iff \mathsf{id}_k \otimes \Phi \supseteq \mathsf{id}_k \otimes \Phi'$ (id_k: identity channel on $\mathsf{L}(\mathbb{C}^k)$)
- $\Phi \supseteq_\infty \Phi' \iff \mathsf{id}_{B'} \otimes \Phi \supseteq \mathsf{id}_{B'} \otimes \Phi'$
- the identity channel can be replaced by any fixed, though arbitrary, invertible channel (possibly entanglement-breaking)

Theorem

- $\Phi \supseteq_k \Phi' \iff \exists$ *k-statistical morphism $\mathcal{M}$:* $\Phi' = \mathcal{M} \circ \Phi$
- $\Phi \supseteq_\infty \Phi' \iff \exists$ *quantum channel $\mathcal{E}$:* $\Phi' = \mathcal{E} \circ \Phi$

Ordering Channels by Quantum Decoding Problems

- given a quantum channel $\Phi : A \to B$, for any input bipartite state ω_{RA}, the **transmitted singlet fraction** is defined as

$$\mathscr{F}(\omega_{RA}|\Phi_A) := \sup_{\mathcal{D}:\text{CPTP}} \langle \Phi^+_{RR'}|(\mathrm{id}_R \otimes \mathcal{D}_B \circ \Phi)(\omega_{RA})|\Phi^+_{RR'}\rangle \,,$$

  where $|\Phi^+_{RR'}\rangle$ denotes the maximally entangled state $(R' \cong R)$
- given two quantum channels $\Phi : A \to B$ and $\Phi' : A \to B'$, we write $\Phi \succeq \Phi'$, whenever $\mathscr{F}(\omega_{RA}|\Phi_A) \geq \mathscr{F}(\omega_{RA}|\Phi'_A)$ for all ω_{RA} $(R \cong B'$ is enough)

Theorem

- $\Phi \succeq \Phi' \iff \Phi \succeq \Phi'$ *only for* **separable** $\omega_{RA} \iff \exists$ *quantum channel* $\mathcal{E}$*:* $\Phi' = \mathcal{E} \circ \Phi$
- $\Phi \succeq \Phi'$ *only for* **classical-quantum** $\omega_{RA} = \sum_x p_x |x\rangle\langle x|_R \otimes \rho^x_A \iff \exists$ *statistical morphism* $\mathcal{M}$*:* $\Phi' = \mathcal{M} \circ \Phi$

Application to Quantum Dynamical Mappings

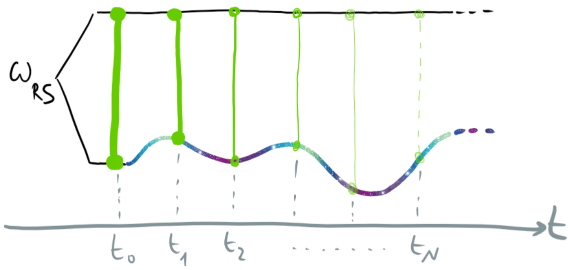


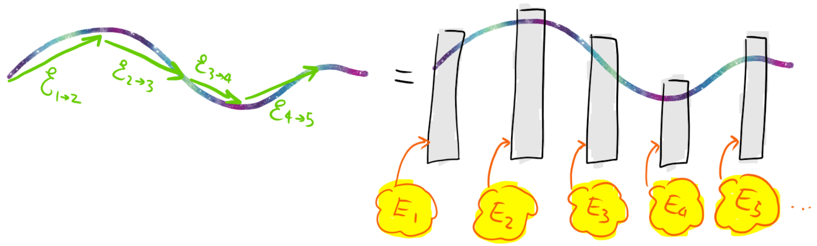
Figure 2: The varying thickness of the green lines depict the singlet fraction at any time.

- a QDM $(\Phi_i)_i$ is **CP-divisible** iff $\Phi_i \succeq \Phi_j$ for all $j \geq i$ and all initial **separable** states
- a QDM $(\Phi_i)_i$ **P-divisible** iff $\Phi_i \succeq \Phi_j$ for all $j \geq i$ and all initial **classical-quantum** states
- in terms of entropies: $H_{\min}(R|S_i) \leq H_{\min}(R|S_j)$, for all $j \geq i$

Some Final Remarks

Meaning of DIV

Why the propagators $(\mathcal{E}_{i\to j})_{i\le j}$ are assumed to be CPTP?



CP-divisibility is equivalent to saying that the open evolution is "collisional," in the sense that it can be realized by summoning a "fresh environment" at each time step.

To Strengthen DSD or to Relax DIV?

Do the propagators $(\mathcal{E}_{i\to j})_{i\leq j}$ *really* need to be linear CPTP?

- linearity is necessary (QDMs are linear)
- trace-preservation (a linear constraint) also
- CP perhaps not: propagators could be just P or even less (e.g., statistical morphisms), and yet be related to important physical/computational/thermodynamical properties (like, e.g., the "locality" or "causality" of the evolution)

Possible Ideas to Think About

- classical correlations can witness P-indivisibility but not CP-indivisibility
- for that, separable states are required: **discord/coherence**, anyone?
- it is known that CP-DIV can be decided by SDP: way to design **efficient tests**?
- robustness to **small deviations** (ϵ-DIV $\iff$ ϵ-DSD)
- to impose extra properties to DIV, e.g., **thermality or group-covariance**
- to understand P-DIV in a **generalized circuit formalism** (no extension possible, however no problem, because not in the black-box picture)
- to understand the **information-theoretic and computational capabilities** of such generalized circuit models, e.g., data-processing inequalities, computational/thermodynamical aspects, etc