Quantum conditional expectations

- A finite-dimensional introduction -

Tommaso Grigoletto

January 10th, 2025

Università degli Studi di Padova

Agenda

- Classical Conditional Expectations (CE) - a brief review:
  - Definitions,
  - Properties,
  - Their use - Maximum likelihood estimator;

- Quantum CEs:
  - Definition,
  - Properties (**Tomita-Takesaki Theorem**),
  - Their use:
    - Convergence to common fixed points by alternating projections;
    - Bayesian parameter estimation (maybe next time);
    - **Model reduction** (next time).

Disclaimer: We here consider ONLY discrete random variables.

Classical CEs

Motivating example - dice games

Imagine tossing two fair dice and take the sum of their outcomes.

Let X, Y be the random variable associated to each die toss, i.e. the alphabet $\mathcal{X} = \mathcal{Y} = \{1, 2, 3, 4, 5, 6\}$, and each outcome has probability $1/6$.
The sum, is a random variable $Z = X + Y$ distributed according to

z	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}[Z = z]$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

We would like to know:

1. What is the expected (weighted average) value of the sum (Z) assuming that one die came out as a 6 (e.g. $y = 6$)?
2. How can we represent this information for any value of one die (Y)?
3. Assume that, to win this turn of the dice game, you need to score at least a 7 (included) and depending on the actual sum you score points. What is the expected value of the sum above and below this threshold?

Different notions of conditional expectations

When studying (classical) probability theory, one encounters different definitions associated to the notion of “conditional expectations”.

These definitions are often categorized as:

- CE (value) given an event;
- CE given a random variable;
- CE given a σ-algebra.

We shall see that these are actually a renaming of the same object.

CE (value) given an event - Definition

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be a $\mathcal{F}$-measurable random variable taking values in the alphabet $\mathcal{X}$ and let E be an event $E \in \mathcal{F}$.

Definition 1

The conditional expectation (value) of X given that the event $E \in \mathcal{F}$ was observed, $\mathbb{E}[X|E]$, is defined as the value:

$$\mathbb{E}[X|E] \equiv \sum_{x \in \mathcal{X}} x \, \mathbb{P}[X = x|E].$$

Intuition: usual definition of expected value with the use of the probability conditioned on the event E.

Observation on the naming

According to this definition, $\mathbb{E}[X|E]$ is just a scalar value.

However, to keep agreement with the definitions that follow we can extend this definition to obtain a random variable by simply multiplying the conditional expectation value by the indicator function associated to the event.

That is, $\mathbb{E}[X|E]\mathbf{1}_E(\omega)$ is now an $\mathcal{F}$-measurable random variable.

Case with two random variables

Let X and Y be random variables defined over the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in the alphabets $\mathcal{X}, \mathcal{Y}$.

Events of interest in this case include when one of the two random variables assume a value, e.g. $Y = y$. In that case, The conditional expectation (value) of X given that Y has taken value $y \in \mathcal{Y}$ is

$$\mathbb{E}[X|Y = y] \equiv \sum_{x \in \mathcal{X}} x \, \mathbb{P}[X = x|Y = y].$$

Note that, we have constructed a function $g(\cdot)$ that, given an outcome $y \in \mathcal{Y}$ returns the conditional expectation (value) $g(y) = \mathbb{E}[X|Y = y]$.

CE given a random variable - Definition

Let X and Y be random variables defined over the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in the alphabets $\mathcal{X}, \mathcal{Y}$.

Definition 2

The conditional expectation of X given Y is then the **random variable** $g(Y)$ which takes the value $E[X|Y = y]$ with probability $p_Y(y)$, i.e.

$$\mathbb{E}[X|Y] \equiv \sum_{y \in \mathcal{Y}} \underbrace{\left(\sum_{x \in \mathcal{X}} x \, \mathbb{P}[X = x | Y = y] \right)}_{\mathbb{E}[X|Y=y]} \mathbf{1}_{Y=y}(\omega)$$

where $\mathbf{1}_{Y=y}(\omega)$ is the indicator function of the event $Y = y$.

CE given a σ-algebra - Definition

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be a $\mathcal{F}$-measurable random variable taking values in the alphabet $\mathcal{X}$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let atom$(\mathcal{G})$ be the set of atomic events of $\mathcal{G}$.
Note: Here we need to assume that $\mathcal{G}$ is generated by a finite set of events, atom$(\mathcal{G})$.

Definition 3
The conditional expectation of X given a σ-algebra $\mathcal{G}$, $\mathbb{E}[X|\mathcal{G}]$, is defined as the $\mathcal{G}$-measurable **random variable**:

$$\mathbb{E}[X|\mathcal{G}] \equiv \sum_{E \in \text{atom}(\mathcal{G})} \mathbb{E}[X|E] 1_E(\omega).$$

Note on the equivalence of the three definitions

CE given an event (r.v.s not values) are simply CE given the σ-algebra $\mathcal{G} = \{\emptyset, E, E^c, \Omega\}$.

CE given a random variable Y are simply CE given the smallest σ-algebra that makes the r.v. Y measurable, i.e. $\mathcal{G} = \sigma(\{Y = y\}_{y \in \mathcal{Y}})$ and atom$(\mathcal{G}) = \{Y = y\}_{y \in \mathcal{Y}}$.

The three different naming we saw do not refer to different notions of conditional expectations rather to different ways of specifying a σ-algebra.

Motivating example - dice games

Imagine tossing two fair dice and take the sum of their outcomes.

Let X, Y be the random variable associated to each die toss, i.e. the alphabet $\mathcal{X} = \mathcal{Y} = \{1, 2, 3, 4, 5, 6\}$, and each outcome has probability $1/6$.
The sum, is a random variable $Z = X + Y$ distributed according to

z	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}[Z = z]$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

We would like to know:

1. What is the expected (weighted average) value of the sum (Z) assuming that one die came out as a 6 ($y = 6$)?
2. How can we represent this information for any value of the die (Y)?
3. Assume that, to win this turn of the dice game, you need to score at least a 7 (included). What is the expected value of the sum above and below this threshold?

Dice games - Solution

$$\mathbb{E}[Z|Y=6] = \sum_{z\in\mathcal{Z}} z\,\mathbb{P}[Z=z|Y=6] = \sum_{z\in\mathcal{Z}} z\,\frac{\mathbb{P}[Z=z\&Y=6]}{\mathbb{P}[Y=6]} = 9.5$$

$W = E[Z|Y]$ is distributed according to

w	4.5	5.5	6.5	7.5	8.5	9.5
$\mathbb{P}[W=w]$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Define $\mathcal{G} = \{\emptyset, \{2,3,4,5,6\}, \{7,8,9,10,11,12\}, \{2,3,4,5,6,7,8,9,10,11,12\}\}$, and $Q = \mathbb{E}[Z|\mathcal{G}]$

q	4.667	8.667
$\mathbb{P}[Q=q]$	$\frac{5}{12}$	$\frac{7}{12}$

Q here is a **coarse** representation of the information contained in Z NOT a conditioning that is necessary after an observation. Often called a **coarse-graining**.

Properties

1. If $\mathbb{P}[A] \neq 0$ then $\mathbb{E}[X|A] = \mathbb{E}[X\mathbf{1}_A]/\mathbb{E}[\mathbf{1}_A]$

$$\mathbb{E}[X|A] = \sum_{x\in\mathcal{X}} x\mathbb{P}[X = x|A] = \sum_{x\in\mathcal{X}} x\frac{\mathbb{P}[X = x\&A]}{\mathbb{P}[A]} = \frac{\sum_{x\in\mathcal{X}} x\mathbb{E}[\mathbf{1}_{X=x}\mathbf{1}_A]}{\mathbb{P}[A]}$$
$$= \frac{\mathbb{E}[X\mathbf{1}_A]}{\mathbb{P}[A]} = \frac{\mathbb{E}[X\mathbf{1}_A]}{\mathbb{E}[\mathbf{1}_A]}$$

2. **Linearity:** $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha\mathbb{E}[X|\mathcal{G}] + \beta\mathbb{E}[Y|\mathcal{G}]$

$$\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \sum_{E\in\text{atom}(\mathcal{G})} \frac{\mathbb{E}[(\alpha X + \beta Y)\mathbf{1}_E]}{\mathbb{P}[E]}\mathbf{1}_E(\omega)$$
$$= \sum_{E\in\text{atom}(\mathcal{G})} \frac{\alpha\mathbb{E}[X\mathbf{1}_E] + \beta\mathbb{E}[Y\mathbf{1}_E]}{\mathbb{P}[E]}\mathbf{1}_E(\omega) = \alpha\mathbb{E}[X|\mathcal{G}] + \beta\mathbb{E}[Y|\mathcal{G}]$$

Properties

3. If X, Y are independent r.v. $\mathbb{E}[X|Y=y] = \mathbb{E}[X]$ for all $y \in \mathcal{Y}$:

$$\mathbb{E}[X|Y=y] = \sum_{x\in\mathcal{X}} x \frac{\mathbb{P}[X=x]\mathbb{P}[Y=y]}{\mathbb{P}[Y=y]} = \sum_{x\in\mathcal{X}} x\mathbb{P}[X=x] = \mathbb{E}[X].$$

   Hence $\mathbb{E}[X|Y]$ is a constant random variable.

4. Let $X = f(Y)$, then $\mathbb{E}[X|Y] = X$.

$$\mathbb{E}[X|Y] = \sum_{y\in\mathcal{Y}} \mathbb{E}[f(Y)|Y=y]\mathbf{1}_{Y=y}(\omega) = \sum_{y\in\mathcal{Y}} f(y)\mathbf{1}_{Y=y}(\omega) = f(Y)$$

Example 4 (Dice games)

$Z = X + Y$, X, Y independent then

$$\mathbb{E}[Z|Y] = \mathbb{E}[X+Y|Y] = \mathbb{E}[X|Y] + \mathbb{E}[Y|Y] = \mathbb{E}[X] + Y = 3.5 + Y$$

Properties

5. **Idempotence:** $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$

$$
\begin{aligned}
\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{G}] &= \sum_{E\in\text{atom}(\mathcal{G})} \frac{\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}_E]}{\mathbb{P}[E]}\mathbf{1}_E(\omega) \\
&= \sum_{E\in\text{atom}(\mathcal{G})} \frac{\mathbb{E}[\sum_{F\in\text{atom}(\mathcal{G})} \mathbb{E}[X|F]\mathbf{1}_F\mathbf{1}_E]}{\mathbb{P}[E]}\mathbf{1}_E(\omega) \\
&= \sum_{E\in\text{atom}(\mathcal{G})} \sum_{F\in\text{atom}(\mathcal{G})} \frac{\mathbb{E}[X|F]\overbrace{\mathbb{E}[\mathbf{1}_F\mathbf{1}_E]}^{\delta_{F,E}}}{\mathbb{P}[E]}\mathbf{1}_E(\omega) \\
&= \sum_{E\in\text{atom}(\mathcal{G})} \frac{\mathbb{E}[X|E]}{\mathbb{P}[E]}\mathbf{1}_E(\omega) = \mathbb{E}[X|\mathcal{G}]
\end{aligned}
$$

Properties

6. **Expectation preservation:** $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$

$$\begin{aligned}\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] &= \sum_{E\in\mathsf{atom}(\mathcal{G})} \mathbb{E}[X|E]\mathbb{P}[E] = \sum_{E\in\mathsf{atom}(\mathcal{G})}\sum_{x\in\mathcal{X}} x\,\mathbb{P}[X=x|E]\mathbb{P}[E] \\ &= \sum_{x\in\mathcal{X}}\sum_{E\in\mathsf{atom}(\mathcal{G})} x\,\mathbb{P}[X=x\&E] = \sum_{x\in\mathcal{X}} x\,\mathbb{P}[X=x] = \mathbb{E}[X]\end{aligned}$$

7. $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}_E] = \mathbb{E}[X\mathbf{1}_E]$ for all $E \in \mathcal{G}$

$$\begin{aligned}\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}_E] &= \sum_{G\in\mathsf{atom}(\mathcal{G})} \mathbb{E}[\mathbb{E}[X|G]\underbrace{\mathbf{1}_G\mathbf{1}_E}_{\delta_{E,G}}] = \mathbb{E}[\mathbb{E}[X|E]\mathbf{1}_E] = \mathbb{E}[X|E]\mathbb{E}[\mathbf{1}_E] \\ &= \frac{\mathbb{E}[X\mathbf{1}_E]}{\mathbb{E}[\mathbf{1}_E]}\mathbb{E}[\mathbf{1}_E] = \mathbb{E}[X\mathbf{1}_E]\end{aligned}$$

Alternative definition

Property 7. is very important. Often times it is used as the defining property for CEs.

Definition 5
The random variable denoted by $\mathbb{E}[X|\mathcal{G}]$ is the conditional expectation of X with respect to $\mathcal{G}$ if:

- it is $\mathcal{G}$-measurable;
- $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}_E] = \mathbb{E}[X\mathbf{1}_E]$ for all $E \in \mathcal{G}$.

This definition is nice because does not require the use of atom$(\mathcal{G})$ hence we need no assumptions on $\mathcal{G}$ and plays nice with continuous-random variables (more rigorous).

Note: With this definition one can prove that $\mathbb{E}[X|\mathcal{G}]$ exist, is unique $\mathbb{P}$-almost surely, and has the properties that we saw up to this point.

Properties

8. **Positivity:** If $X \geq 0$ almost surely, then $\mathbb{E}[X|\mathcal{G}] \geq 0$.
   Let E denote the event $E = \{\mathbb{E}[X|\mathcal{G}] < 0\} \in \mathcal{G}$.
   Since $X \geq 0$ then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbf{1}_E] = \mathbb{E}[X\mathbf{1}_E] \geq 0$ therefore $\mathbb{P}[E] = 0$.

8 bis. **Monotonicity:** If $X \geq Y$ almost surely, then $\mathbb{E}[X|\mathcal{G}] \geq \mathbb{E}[Y|\mathcal{G}]$.

9. Let Y be a $\mathcal{G}$-measurable r.v, then $\mathbb{E}[Y|\mathcal{G}] = Y$.
   Since Y is $\mathcal{G}$-measurable, we can write $Y = \sum_{E \in \text{atom}(\mathcal{G})} y_E \mathbb{1}_E(\omega)$, i.e. Y is constant for all $E \in \text{atom}(\mathcal{G})$.

$$\mathbb{E}[Y|\mathcal{G}] = \sum_{E \in \text{atom}(\mathcal{G})} \frac{\mathbb{E}[Y\mathbf{1}_E]}{\mathbb{E}[\mathbf{1}_E]} \mathbf{1}_E(\omega) = \sum_{E \in \text{atom}(\mathcal{G})} \sum_{E \in \text{atom}(\mathcal{G})} \frac{y_F \mathbb{E}[\mathbf{1}_F \mathbf{1}_E]}{\mathbb{E}[\mathbf{1}_E]} \mathbf{1}_E(\omega)$$

$$= \sum_{E \in \text{atom}(\mathcal{G})} \frac{y_E \mathbb{E}[\mathbf{1}_E]}{\mathbb{E}[\mathbf{1}_E]} \mathbf{1}_E(\omega) = Y$$

Properties

10. **Multiplicativity on $\mathcal{G}$:** Let Y be a $\mathcal{G}$-measurable r.v, then $\mathbb{E}[Xf(Y)|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]f(Y)$.

$$\begin{aligned}
\mathbb{E}[XY|\mathcal{G}] &= \sum_{E \in \text{atom}(\mathcal{G})} \mathbb{E}[XY|E]\mathbf{1}_E(\omega) = \sum_{E \in \text{atom}(\mathcal{G})} \frac{\mathbb{E}[XY\mathbf{1}_E]}{\mathbb{E}[\mathbf{1}_E]}\mathbf{1}_E(\omega) \\
&= \sum_{E \in \text{atom}(\mathcal{G})} \sum_{F \in \text{atom}(\mathcal{G})} \frac{y_F \mathbb{E}[X\mathbf{1}_F\mathbf{1}_E]}{\mathbb{E}[\mathbf{1}_E]}\mathbf{1}_E(\omega) = \sum_{E \in \text{atom}(\mathcal{G})} \frac{y_E \mathbb{E}[X\mathbf{1}_E]}{\mathbb{E}[\mathbf{1}_E]}\mathbf{1}_E(\omega) \\
&= \sum_{E \in \text{atom}(\mathcal{G})} y_E \mathbb{E}[X|E]\mathbf{1}_E(\omega) = \mathbb{E}[X|\mathcal{G}]Y
\end{aligned}$$

Algebraic representation of probability theory.

Recall from the first lecture that we can represent a probability space as a vector algebra $\mathscr{A} \subseteq \mathbb{R}^n$ (closed w.r.t. linear combinations and element-wise product $\wedge$) and a probability vector $\boldsymbol{p} \in \mathbb{R}^n$, s.t. $\boldsymbol{p} \geq 0$ and $\mathbf{1}^T \boldsymbol{p} = 1$.

We can then represent random variables as vectors , e.g. $\boldsymbol{x} \in \mathscr{A}$ and indicator functions as idempotent vectors, i.e. $\boldsymbol{f}_E \wedge \boldsymbol{f}_E = \boldsymbol{f}_E$.

To compute probabilities and expectation values we can then take inner product e.g. $\mathbb{E}[X] = \langle \boldsymbol{p}, \boldsymbol{x} \rangle$ and $\mathbb{P}[E] = \mathbb{E}[\mathbf{1}_E] = \langle \boldsymbol{p}, \boldsymbol{f}_E \rangle$.

We now wonder: How do CEs look like in this setting?

CE in this setting

1. We represent the sub-σ-algebra $\mathcal{G}$ as the sub-algebra $\mathscr{B} \subseteq \mathscr{A}$.

2. The CE $\mathbb{E}[\cdot|\mathcal{G}]$ is a linear map that takes rvs into rvs (prop. 2). We can thus represent the CE as a matrix $\in \mathbb{R}^{n\times n}$.

3. The CE is taken w.r.t. a specific probability measure, hence we denote it by $\mathbb{E}_{\mathscr{B},\boldsymbol{p}}$.

4. $\mathbb{E}[\cdot|\mathcal{G}]$ takes $\mathcal{F}$-measurable rvs into $\mathcal{G}$-measurable rvs hence

$$\begin{aligned} \mathbb{E}_{\mathscr{B},\boldsymbol{p}} :& \mathscr{A} \rightarrow \mathscr{B} \\ & \boldsymbol{x} \rightarrow \boldsymbol{y} = \mathbb{E}_{\mathscr{B},\boldsymbol{p}} \boldsymbol{x} \end{aligned}$$

where $\boldsymbol{y} \in \mathscr{B}$ is the vector representation of $Y = \mathbb{E}[X|\mathcal{G}]$ thus $\mathrm{Im}\mathbb{E}_{\mathscr{B},\boldsymbol{p}} = \mathscr{B}$.

5. Since (prop.5) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$, then $\mathbb{E}_{\mathscr{B},\boldsymbol{p}}\mathbb{E}_{\mathscr{B},\boldsymbol{p}} = \mathbb{E}_{\mathscr{B},\boldsymbol{p}}$.

CEs are projectors onto sub-algebras!

CE in this setting - other properties

- $\mathbb{E}_{\mathscr{B},\boldsymbol{p}}$ is not necessarily an orthogonal projector, i.e. $\mathbb{E}^T_{\mathscr{B},\boldsymbol{p}} \neq \mathbb{E}_{\mathscr{B},\boldsymbol{p}}$.
- It still is orthogonal wrt a different inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{p}} \equiv \mathbb{E}_{\boldsymbol{p}}[\boldsymbol{x} \wedge \boldsymbol{y}] = \langle \boldsymbol{p}, \boldsymbol{x} \wedge \boldsymbol{y} \rangle$:

$$\left\langle \boldsymbol{x}, \mathbb{E}_{\mathscr{B},\boldsymbol{p}} \boldsymbol{y} \right\rangle_{\boldsymbol{p}} = \left\langle \mathbb{E}_{\mathscr{B},\boldsymbol{p}} \boldsymbol{x}, \boldsymbol{y} \right\rangle_{\boldsymbol{p}}.$$

- (Prop. 6) implies $\left\langle \boldsymbol{p}, \mathbb{E}_{\mathscr{B},\boldsymbol{p}} \boldsymbol{x} \right\rangle = \left\langle \mathbb{E}^T_{\mathscr{B},\boldsymbol{p}} \boldsymbol{p}, \boldsymbol{x} \right\rangle = \langle \boldsymbol{p}, \boldsymbol{x} \rangle$ hence $\mathbb{E}^T_{\mathscr{B},\boldsymbol{p}}[\boldsymbol{p}] = \boldsymbol{p}$, or, in other words, the state $\boldsymbol{p}$ is preserved by the dual of the CE $\mathbb{E}^T_{\mathscr{B},\boldsymbol{p}}$.
  The above does not hold wrt other probability measures, i.e. $\mathbb{E}^T_{\mathscr{B},\boldsymbol{p}}[\boldsymbol{q}] \neq \boldsymbol{q}$.
- For all $\boldsymbol{f}_E \in \text{idem}(\mathscr{B})$ we have $\left\langle \boldsymbol{p}, (\mathbb{E}_{\mathscr{B},\boldsymbol{p}} \boldsymbol{x}) \wedge \boldsymbol{f}_E \right\rangle = \langle \boldsymbol{p}, \boldsymbol{x} \wedge \boldsymbol{f}_E \rangle$ (prop. 7).
- Due to monotonicity (prop. 8), $\mathbb{E}_{\mathscr{B},\boldsymbol{p}}$ is a non-negative matrix s.t. $\mathbb{E}_{\mathscr{B},\boldsymbol{p}} \mathbf{1} = \mathbf{1}$ and $\mathbb{E}^T_{\mathscr{B},\boldsymbol{p}}$ is stochastic.
- For all $\boldsymbol{y} \in \mathscr{B}$ we have $\mathbb{E}_{\mathscr{B},\boldsymbol{p}}(\boldsymbol{y}) = \boldsymbol{y}$ and $\mathbb{E}_{\mathscr{B},\boldsymbol{p}}(\boldsymbol{x} \wedge \boldsymbol{y}) = (\mathbb{E}_{\mathscr{B},\boldsymbol{p}} \boldsymbol{x}) \wedge \boldsymbol{y}$ (prop. 9, 10).

Use - Optimal least square estimator

Assume you have a rv X and a sub-σ-algebra $\mathcal{G}$ (given for example by a series of observations or by other random variables).

We would like to find a $\mathcal{G}$-measurable rv Y that is the best approximation of X in the least square sense, that is we want to solve the minimization problem

$$Y^* = \arg \min_{Y, \mathcal{G}\text{-meas.}} \mathbb{E}[(X - Y)^2].$$

Optimal least square estimator - Algebraic proof

In algebraic terms, given $\boldsymbol{x} \in \mathscr{A}$ and a sub-algebra $\mathscr{B} \subseteq \mathscr{A}$ we want to find $\boldsymbol{y}^* \in \mathscr{B}$, s.t.

$$\boldsymbol{y}^* = \arg\min_{\boldsymbol{y} \in \mathscr{B}} \left\langle \boldsymbol{p}, (\boldsymbol{x} - \boldsymbol{y})^2 \right\rangle$$

Note that

$$\left\langle \boldsymbol{p}, (\boldsymbol{x} - \boldsymbol{y})^2 \right\rangle = \langle \boldsymbol{p}, (\boldsymbol{x} - \boldsymbol{y}) \wedge (\boldsymbol{x} - \boldsymbol{y}) \rangle = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle_{\boldsymbol{p}} = ||\boldsymbol{x} - \boldsymbol{y}||_{\boldsymbol{p}}^2 .$$

It is now trivial to see that $\boldsymbol{y}^*$ is the orthogonal projection of $\boldsymbol{x}$ onto $\mathscr{B}$ where the notion of orthogonality is given wrt $\langle \cdot, \cdot \rangle_{\boldsymbol{p}}$. Thus $\boldsymbol{y}^* = \mathbb{E}_{\mathscr{B}, \boldsymbol{p}} \boldsymbol{x}$.

Optimal least square estimator - Usual proof

Alternatively,

$$\begin{aligned}
\mathbb{E}[(X-Y)^2] &= \mathbb{E}[(X-\mathbb{E}[X|\mathcal{G}]+\mathbb{E}[X|\mathcal{G}]-Y)^2] \\
&= \mathbb{E}[(X-\mathbb{E}[X|\mathcal{G}])^2]+\mathbb{E}[(\mathbb{E}[X|\mathcal{G}]-Y)^2]+2\mathbb{E}[(X-\mathbb{E}[X|\mathcal{G}])\underbrace{(\mathbb{E}[X|\mathcal{G}]-Y)}_{\mathcal{G}\text{-meas.}}] \\
&= \mathbb{E}[(X-\mathbb{E}[X|\mathcal{G}])^2]+\mathbb{E}[(\mathbb{E}[X|\mathcal{G}]-Y)^2]
\end{aligned}$$

where we used the fact that for any $\mathcal{G}$-measurable rv Z we have

$$\mathbb{E}[(X-\mathbb{E}[X|\mathcal{G}])Z]=\mathbb{E}[XZ]-\mathbb{E}[\mathbb{E}[XZ|\mathcal{G}]]=\mathbb{E}[XZ]-\mathbb{E}[XZ]=0.$$

Since $\mathbb{E}[(\mathbb{E}[X|\mathcal{G}]-Y)^2]\geq 0$, then for any $\mathcal{G}$-measurable rv Y:

$$\mathbb{E}[(X-Y)^2]\geq \mathbb{E}[(X-\mathbb{E}[X|\mathcal{G}])^2].$$

Quantum CEs

Brief review of $*$-algebras

We define a $*$-algebra $\mathscr{A}$ as an operator space closed under matrix multiplication and adjoint action.

$$X, Y \in \mathscr{A} \Rightarrow X + Y \in \mathscr{A} \qquad X^\dagger, Y^\dagger \in \mathscr{A} \qquad \text{and} \qquad XY \in \mathscr{A}$$

Remember: It is the fundamental mathematical structure that supports a quantum probability space.

Examples

$$\mathscr{A}_1 = \text{span}\{\sigma_j, \quad j = 0, x, y, z\} = \mathbb{C}^{2\times 2}, \qquad \dim(\mathscr{A}_1) = 4$$

$$\mathscr{A}_2 = \text{span}\{\sigma_j \otimes \sigma_k, \quad j, k = 0, x, y, z\} = \mathbb{C}^{4\times 4}, \qquad \dim(\mathscr{A}_2) = 16$$

$$\mathscr{A}_3 = \text{span}\{\sigma_j \otimes \sigma_k, \quad j = 0, x, y, z, k = 0, z\} \subsetneq \mathbb{C}^{4\times 4}, \qquad \dim(\mathscr{A}_3) = 8$$

$$\mathscr{A}_4 = \text{span}\{\sigma_j \otimes \sigma_j, \quad j = 0, x, y, z\} \subsetneq \mathbb{C}^{4\times 4}, \qquad \dim(\mathscr{A}_4) = 4$$

$$\mathscr{A}_5 = \text{span}\{|j\rangle\langle j|\,, \quad j = 0, 1, 2, 3\} \subsetneq \mathbb{C}^{4\times 4}, \qquad \dim(\mathscr{A}_5) = 4$$

Question: $\mathscr{A}_4$ and $\mathscr{A}_5$ have the same dimension. Are they the same algebra?
If not how do we distinguish them?

Wedderburn decomposition

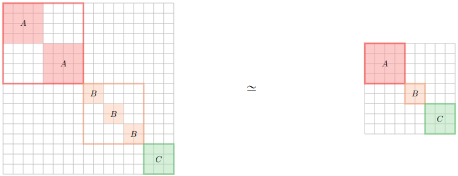
For any algebra $\mathscr{A}$, there exist an Hibert space decomposition

$$\mathcal{H} = \bigoplus_k \mathcal{H}_{S,k} \otimes \mathcal{H}_{F,k} \oplus \mathcal{H}_R$$

and a unitary operator U such that

$$\mathscr{A} = U \left(\bigoplus_k \mathcal{B}(\mathcal{H}_{S,k}) \otimes \mathbb{1}_{F,k} \oplus 0_R \right) U^\dagger.$$

Note: some values are repeated multiple times hence we can find a smaller isomorphic representation $\mathscr{A} \simeq \bigoplus_k \mathcal{B}(\mathcal{H}_{S,k}) \equiv \check{\mathscr{A}}$.



Examples

$$\mathscr{A}_1 = \mathrm{span}\{\sigma_j, \quad j = 0, x, y, z\} = \mathbb{C}^{2\times 2}$$

$$\mathscr{A}_2 = \mathrm{span}\{\sigma_j \otimes \sigma_k, \quad j, k = 0, x, y, z\} = \mathbb{C}^{4\times 4}$$

$$\mathscr{A}_3 = \mathrm{span}\{\sigma_j \otimes \sigma_k, \quad j = 0, x, y, z, k = 0, z\} \simeq \mathbb{C}^{2\times 2} \bigoplus \mathbb{C}^{2\times 2}$$

$$\mathscr{A}_4 = \mathrm{span}\{\sigma_j \otimes \sigma_j, \quad j = 0, x, y, z\} \simeq \mathbb{C}^{2\times 2}$$

$$\mathscr{A}_5 = \mathrm{span}\{|j\rangle\langle j|\,, \quad j = 0, 1, 2, 3\} \simeq \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}$$

Quantum CE - Definition

Definition 6

A *conditional expectation* $\mathbb{E}_{\mathscr{B},\rho}[\cdot]$ is a **positive projector** onto a unital algebra $\mathscr{B}$

$$\mathbb{E}_{\mathscr{B},\rho} : \mathscr{A} \to \mathscr{B}.$$

A conditional expectation preserves a state $\rho \in \mathfrak{D}(\mathcal{H})$ when, for all $X \in \mathcal{B}(\mathcal{H})$

$$\mathbb{E}_\rho[\mathbb{E}_{\mathscr{B},\rho}[X]] = \mathrm{tr}[\rho \mathbb{E}_{\mathscr{B},\rho}[X]] = \mathrm{tr}[\rho X] = \mathbb{E}_\rho[X].$$

Properties

Proposition 1

Let $\mathcal{E} : \mathscr{A} \to \mathscr{B}$ be a linear, and unital map, i.e. $\mathcal{E}(\mathbb{1}) = \mathbb{1}$. Then $\mathcal{E}$ is contractive in the operator norm if an only if $\mathcal{E}$ is positive, that is:

$$||\mathcal{E}[A]||_{op} \leq ||A||_{op} \forall A \in \mathscr{A} \quad \Leftrightarrow \quad \mathcal{E}(A^\dagger A) \geq 0, \forall A \in \mathscr{A}.$$

This proposition is quite useful because there exists a very powerful theorem that talks about projectors between $*$-algebras that are contractive in the operator norm.

Properties - Tomiyama

Theorem 7 (Tomiyama 1957)
Let $\Pi_{\mathscr{B}} : \mathscr{A} \to \mathscr{B}$ be a projection from $\mathscr{A}$ onto $\mathscr{B}$, contractive in the operator norm, i.e. $||\Pi_{\mathscr{B}}(X)||_{op} \leq ||X||_{op}$.
Then $\Pi_{\mathscr{B}}$ enjoys the following properties:

1. $\Pi_{\mathscr{B}}$ is positive and unital, i.e. $\Pi_{\mathscr{B}}(A^\dagger A) \geq 0$ for all $A \in \mathscr{A}$, and $\Pi(\mathbb{1}) = \mathbb{1}$;
2. $\Pi_{\mathscr{B}}$ is B-linear, i.e. for all $A \in \mathscr{A}$ and $B_1, B_2 \in \mathscr{B}$, $\Pi_{\mathscr{B}}(B_1 A B_2) = B_1 \Pi_{\mathscr{B}}(A) B_2$;
3. $\Pi_{\mathscr{B}}$ satisfies the Kadison-Schwartz inequality $\Pi_{\mathscr{B}}(A)^\dagger \Pi_{\mathscr{B}}(A) \leq \Pi_{\mathscr{B}}(A^\dagger A)$ for all $A \in \mathscr{A}$.

Corollary 8
Let $\Pi_{\mathscr{B}} : \mathscr{A} \to \mathscr{B}$ be a positive projector between $*$-algebras, then $\Pi_{\mathscr{B}}[\cdot]$ is completely positive.

Note: CEs are automatically CP and unital.

Properties - Orthogonality

Let consider a full rank density operator ρ and let us define the inner product

$$\langle X, Y \rangle_\rho \equiv \mathrm{tr}[\rho X^\dagger Y]$$

Proposition 2

Consider a CE $\mathbb{E}_{\mathscr{B},\rho}$ that preserves the state ρ. Then $\mathbb{E}_{\mathscr{B},\rho}$ is orthogonal wrt the inner product $\langle \cdot, \cdot \rangle_\rho$, i.e. $\left\langle X, \mathbb{E}_{\mathscr{B},\rho}(Y) \right\rangle_\rho = \left\langle \mathbb{E}_{\mathscr{B},\rho}(X), Y \right\rangle_\rho$.

Proof.

Using state preservation, B-linearity and positivity we get

$$\left\langle X, \mathbb{E}_{\mathscr{B},\rho}(Y) \right\rangle_\rho = \mathrm{tr}[\rho X^\dagger \mathbb{E}_{\mathscr{B},\rho}(Y)] = \mathrm{tr}[\rho \mathbb{E}_{\mathscr{B},\rho}(X^\dagger \mathbb{E}_{\mathscr{B},\rho}(Y))] = \mathrm{tr}[\rho \mathbb{E}_{\mathscr{B},\rho}(X)^\dagger \mathbb{E}_{\mathscr{B},\rho}(Y)],$$

$$\left\langle \mathbb{E}_{\mathscr{B},\rho}(X), Y \right\rangle_\rho = \mathrm{tr}[\rho \mathbb{E}_{\mathscr{B},\rho}(X)^\dagger Y] = \mathrm{tr}[\rho \mathbb{E}_{\mathscr{B},\rho}(\mathbb{E}_{\mathscr{B},\rho}(X)^\dagger Y)] = \mathrm{tr}[\rho \mathbb{E}_{\mathscr{B},\rho}(X)^\dagger \mathbb{E}_{\mathscr{B},\rho}(Y)].$$

$\square$

Properties - Existence

Theorem 9 (Tomita-Takesaki theorem)

Suppose $\rho \in \mathfrak{D}(\mathcal{H})$ is an invertible density matrix. The following conditions are equivalent:

1. A conditional expectation $\mathbb{E}_{\mathscr{B},\rho}[\cdot] : \mathscr{A} \to \mathscr{B}$, that preserves ρ exists;
2. $\mathscr{B}$ is $\mathcal{M}_{\rho,\lambda}(\cdot) \equiv \rho^{\lambda} \cdot \rho^{-\lambda}$ invariant $\forall \lambda \in \mathbb{C}$, i.e. for every $B \in \mathscr{B}$ it holds

$$\rho^{\lambda} B \rho^{-\lambda} \in \mathscr{B}.$$

The proof that follows is inspired by Petz's "Quantum Information Theory and Quantum Statistics".

Here we prove it in the finite-dimensional case but it holds also for infinite-dimensional algebras.

Tomita-Takesaki Proof of (1.) $\Rightarrow$ (2.)

This comes directly from the block structure of fixed points of CP and unital maps whose dual admits a full-rank equilibria (seen in the previous lecture).

$$\mathrm{fix}(\mathbb{E}|_{\mathscr{B},\rho}) = \mathscr{B} = \bigoplus_k \mathfrak{B}(\mathcal{H}_{S,k}) \otimes \mathbb{1}_{F,k} \quad \text{implies} \quad \rho = \bigoplus_k \tau_k \otimes \sigma_k \in \mathrm{fix}(\mathbb{E}|_{\mathscr{B},\rho}^{\dagger})$$

Tomita-Takesaki Proof of (1.) $\Rightarrow$ (2.)

Alternative proof from Petz:

- Let us denote by $\mathcal{S}$ the conjugate linear operator, i.e. $\mathcal{S}(X) = X^\dagger$, $\forall X \in \mathfrak{B}(\mathcal{H})$.
- Let us also denote by $*$ the adjoint operator wrt $\langle \cdot, \cdot \rangle_\rho$, e.g. $\mathbb{E}^*_{\mathscr{B},\rho} = \mathbb{E}_{\mathscr{B},\rho}$.
- Then $\mathcal{S}^*(X) = \rho X^\dagger \rho^{-1}$ since:

$$\langle \mathcal{S}(X), Y \rangle_\rho = \mathrm{tr}[\rho X Y] = \overline{\mathrm{tr}[Y^\dagger \rho^{-1} \rho X^\dagger \rho]} = \overline{\mathrm{tr}[\rho X^\dagger \rho Y^\dagger \rho^{-1}]} = \overline{\langle X, \rho Y^\dagger \rho^{-1} \rangle_\rho} = \langle \mathcal{S}^*(Y), X \rangle_\rho .$$

- Since $\mathbb{E}_{\mathscr{B},\rho}$ is positive, $\mathbb{E}_{\mathscr{B},\rho}\mathcal{S} = \mathcal{S}\mathbb{E}_{\mathscr{B},\rho}$.
- Furthermore, since $\mathbb{E}^*_{\mathscr{B},\rho} = \mathbb{E}_{\mathscr{B},\rho}$, then $\mathbb{E}_{\mathscr{B},\rho}\mathcal{S}^* = \mathcal{S}^*\mathbb{E}_{\mathscr{B},\rho}$.
- If we then define $\Delta \equiv \mathcal{S}^*\mathcal{S}$, $\Delta(X) = \rho X \rho^{-1}$, we have $\Delta\mathbb{E}_{\mathscr{B},\rho} = \mathbb{E}_{\mathscr{B},\rho}\Delta$ as well as $\Delta^{it}\mathbb{E}_{\mathscr{B},\rho} = \mathbb{E}_{\mathscr{B},\rho}\Delta^{it}$ with $\Delta^{it}(X) = \rho^{it} X \rho^{-it}$.
- To conclude we have

$$\Delta^{it}\mathscr{B} = \Delta^{it}\mathbb{E}_{\mathscr{B},\rho}\mathscr{A} = \mathbb{E}_{\mathscr{B},\rho}\Delta^{it}\mathscr{A} = \mathbb{E}_{\mathscr{B},\rho}\mathscr{A} = \mathscr{B}.$$

Tomita-Takesaki Proof of (2.) $\Rightarrow$ (1.)

Step 1: Prove that $\mathbb{E}_{\mathscr{B},\mathbb{1}/n}$, a positive projector onto $\mathscr{B}$ preserving $\mathbb{1}/n$, exists.

• $\mathcal{E} : \mathscr{A} \to \mathscr{B}$ be the orthogonal (wrt $\langle \cdot, \cdot \rangle_{HS}$) projector onto $\mathscr{B}$, i.e. $\mathcal{E}^2 = \mathcal{E} = \mathcal{E}^\dagger$.
• It exists as it is just a projector onto an operator subspace.
• Since $\mathscr{B}$ is unital, then $\mathcal{E}(\mathbb{1}) = \mathbb{1}$, hence $\mathcal{E}$ preserves $\mathbb{1}/n$ in the sense

$$\mathrm{tr}[\mathbb{1}/n\mathcal{E}(X)] = \mathrm{tr}[\mathbb{1}/nX].$$

• It then remains to prove that $\mathcal{E}$ is also positive (in the sense that maps positive semidefinite operators into positive semidefinite operators). Recall that X is positive semidefinite iff $\mathrm{tr}[Y^\dagger X] \geq 0$ for all $Y \geq 0$. Consider the $X \geq 0$.

$$\langle B, \mathcal{E}(X) \rangle_{HS} = \langle \mathcal{E}(B), X \rangle_{HS} = \mathrm{tr}[B^\dagger X] \geq 0$$

for all positive semidefinite $B \in \mathscr{B}$. Hence $\mathcal{E}(X) \in \mathscr{B}$ is positive semidefinite.

Tomita-Takesaki Proof of (2.) $\Rightarrow$ (1.) continues

Step 2: Assumption 2. holds if and only if ρ **is compatible with** $\mathscr{B}$ that is

$$\rho = \bigoplus_k \tau_k \otimes \sigma_k \quad \text{where} \quad \mathscr{B} = \bigoplus_k \mathfrak{B}(\mathcal{H}_{S,k}) \otimes \mathbb{1}_{F,k}.$$

$\Leftarrow$ is trivial. $\Rightarrow$ is quite technical, we just give an intuition.

- Note that $\mathcal{M}_{\rho,i\varphi}$ is a unitary super-operator group as $\rho^{i\varphi}$ is unitary: $\rho^{i\varphi}\rho^{i\varphi\dagger} = \rho^{i\varphi}\rho^{-i\varphi} = \mathbb{1}$. Let $\rho^{i\varphi} = e^{iH\varphi}$ and $\mathcal{L}(\cdot) = -i[H, \cdot]$
- The fact that $\mathscr{B}$ is $\mathcal{M}_{\rho,i\varphi}$ invariant implies that:
  - Taking the derivative, the off-diagonal blocks must be zero, i.e. $\mathcal{L}(\mathscr{B}) \subseteq \bigoplus_k \mathfrak{B}(\mathcal{H}_k)$.
  - In the diagonal-blocks, there can not be entanglement thus

$$\mathcal{L}(B) = \bigoplus_k \mathcal{L}_{S,k}(B_k) \otimes \mathbb{1}_{F,k} + \mathbb{1}_{S,k} \otimes \mathcal{L}_{F,k}(\mathbb{1}_{F,k}) \quad \Leftrightarrow \quad H = \bigoplus_k H_{S,k} \otimes \mathbb{1}_{F,k} + \mathbb{1}_{S,k} \otimes H_{F,k}.$$

  - $\mathcal{L}_{F,k}(\mathbb{1}_{F,k}) = 0$, already satisfied.

Tomita-Takesaki Proof of (2.) $\Rightarrow$ (1.) continues

Step 3: Define $\rho_0 \equiv \mathcal{E}(\rho) \in \mathscr{B}$. Under assumption 2. we want to prove that, for all $B \in \mathscr{B}$:

$$\rho^{\frac{1}{2}} B \rho^{-\frac{1}{2}} = \rho_0^{\frac{1}{2}} B \rho_0^{-\frac{1}{2}}.$$

As $B \in \mathscr{B}$, it admits a decomposition as $B = \bigoplus_k B_k \otimes \mathbb{1}_{F,k}$.
From step 2:

$$\rho = \bigoplus_k \tau_k \otimes \sigma_k \quad \text{then} \quad \rho_0 = \bigoplus \tau_k \otimes \mathbb{1}_{F,k}$$

thus:

$$\rho^{\frac{1}{2}} B \rho^{-\frac{1}{2}} = \bigoplus_k \tau_k^{\frac{1}{2}} B_k \tau_k^{-\frac{1}{2}} \otimes \sigma_k^{\frac{1}{2}} \sigma_k^{-\frac{1}{2}} = \bigoplus_k \tau_k^{\frac{1}{2}} B_k \tau_k^{-\frac{1}{2}} \otimes \mathbb{1}_{F,k} = \rho_0^{\frac{1}{2}} B \rho_0^{-\frac{1}{2}}$$

Tomita-Takesaki Proof of (2.) $\Rightarrow$ (1.) continues

Step 4: Define $\mathcal{F}_\rho(X) \equiv \rho_0^{-\frac{1}{2}}\mathcal{E}(\rho^{\frac{1}{2}}X\rho^{\frac{1}{2}})\rho_0^{-\frac{1}{2}}$ and observe:

- $\mathcal{F}_\rho$ is CP and preserves ρ:

$$\begin{aligned}\mathrm{tr}[\rho\mathcal{F}_\rho(X)] &= \mathrm{tr}[\rho\rho_0^{-\frac{1}{2}}\mathcal{E}(\rho^{\frac{1}{2}}X\rho^{\frac{1}{2}})\rho_0^{-\frac{1}{2}}] = \mathrm{tr}[\mathcal{E}(\rho_0^{-\frac{1}{2}}\rho\rho_0^{-\frac{1}{2}})\rho^{\frac{1}{2}}X\rho^{\frac{1}{2}}]\\ &= \mathrm{tr}[\rho_0^{-\frac{1}{2}}\mathcal{E}(\rho)\rho_0^{-\frac{1}{2}}\rho^{\frac{1}{2}}X\rho^{\frac{1}{2}}] = \mathrm{tr}[\rho_0^{-\frac{1}{2}}\rho_0\rho_0^{-\frac{1}{2}}\rho^{\frac{1}{2}}X\rho^{\frac{1}{2}}] = \mathrm{tr}[\rho X];\end{aligned}$$

- Under assumption 2. we have, for all $B \in \mathscr{B}$, that:

$$\begin{aligned}\mathcal{F}_\rho(B) &= \rho_0^{-\frac{1}{2}}\mathcal{E}(\rho^{\frac{1}{2}}B\rho^{\frac{1}{2}})\rho_0^{-\frac{1}{2}} = \rho_0^{-\frac{1}{2}}\mathcal{E}(\underbrace{\rho^{\frac{1}{2}}B\rho^{-\frac{1}{2}}}_{\in\mathscr{A}}\rho)\rho_0^{-\frac{1}{2}} = \rho_0^{-\frac{1}{2}}\rho^{\frac{1}{2}}B\rho^{-\frac{1}{2}}\mathcal{E}(\rho)\rho_0^{-\frac{1}{2}}\\ &= \rho_0^{-\frac{1}{2}}\rho_0^{\frac{1}{2}}B\rho_0^{-\frac{1}{2}}\mathcal{E}(\rho)\rho_0^{-\frac{1}{2}} = B\rho_0^{-\frac{1}{2}}\rho_0\rho_0^{-\frac{1}{2}} = B\end{aligned}$$

  hence $\mathcal{F}_\rho$ acts as the identity on $\mathscr{B}$ and, since $\mathrm{Im}\mathcal{F}_\rho = \mathscr{B}$, then $\mathcal{F}_\rho^2 = \mathcal{F}_\rho$. $\square$

Understanding the theorem

Attention! Given $\mathscr{B}$, there always exists a positive projector (CE) onto $\mathscr{B}$ that preserves a state, but, given $\mathscr{B}$ AND ρ, there might not exist a positive projector onto $\mathscr{B}$ that preserve that particular state ρ.

Take $\mathscr{A} = \mathfrak{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ and $\mathscr{B} = \mathfrak{B}(\mathcal{H}_1) \otimes \mathbb{1}_2$.

If $\rho = \rho_1 \otimes \rho_2$ then $\mathbb{E}_{\mathscr{B},\rho}[X] = \mathrm{tr}_{\mathcal{H}_2}[X(\mathbb{1}_1 \otimes \rho_2)]$.
If $\rho \neq \rho_1 \otimes \rho_2$ then $\nexists \mathbb{E}_{\mathscr{B},\rho}$, positive and such that ρ is preserved.

This is a **major departure** from the classical case - for every commuting subalgebra and distribution a conditional expectation fixing the latter exists!

CE and their block representation

Theorem 10

Let $\mathbb{E}_{\mathscr{B},\rho}[\cdot]$ be a conditional expectation onto $\mathscr{B}$, a unital $*$-subalgebra of $\mathscr{A}$. Then, there exists a set of density operators $\{\tau_k \in \mathfrak{D}(\mathcal{H}_{F,k})\}$ such that, for all $X \in \mathcal{B}(\mathcal{H})$

$$\mathbb{E}_{\mathscr{B},\rho}[X] = U \left(\bigoplus_{k=0}^{K-1} \mathrm{tr}_{F,k} \left[(W_k X W_k^\dagger)(\mathbb{1}_{d_k} \otimes \tau_k) \right] \otimes \mathbb{1}_{F,k} \right) U^\dagger$$

where W_k are partial isometries $W_k^\dagger : \mathcal{H}_k \to \mathcal{H}$, s.t. $W_k W_k^\dagger = \mathbb{1}_{\mathcal{H}_k}$ and $W_k^\dagger W_k = \Pi_{\mathcal{H}_k}$.

Lemma 11

Let $\mathbb{E}_{\mathscr{B},\rho}$ be a conditional expectation that preserves $\rho > 0$. Then $\mathbb{E}_{\mathscr{B},\rho}$ is orthogonal wrt the inner product $\langle X, Y \rangle_{\rho,\lambda} \equiv \mathrm{tr}[X^\dagger \rho^\lambda Y \rho^{1-\lambda}]$, for all $\lambda \in [0, 1]$.

State extensions - Definition

Definition 12
The adjoint operator of the conditional expectation $\mathbb{E}_{\mathscr{B},\rho}$ with respect to the Hilbert-Schmidt inner product takes the name of *state extension*, denoted as $\mathbb{J}_{\mathscr{B},\rho}[\cdot] = \mathbb{E}^{\dagger}_{\mathscr{B},\rho}[\cdot]$, i.e. $\left\langle X, \mathbb{E}_{\mathscr{B},\rho}[Y] \right\rangle_{HS} = \left\langle \mathbb{J}_{\mathscr{B},\rho}[X], Y \right\rangle_{HS}$ for all $X, Y \in \mathfrak{B}(\mathcal{H})$.

A conditional expectation $\mathbb{E}_{\mathscr{B},\rho}[\cdot]$, preserves ρ if and only if ρ is a fixed point of its state extension, i.e.

$$\mathbb{E}_{\rho}[\mathbb{E}_{\mathscr{B},\rho}[\cdot]] = \mathbb{E}_{\rho}[\cdot] \quad \Leftrightarrow \quad \mathbb{J}_{\mathscr{B},\rho}[\rho] = \rho.$$

Since $\mathbb{E}_{\mathscr{B},\rho}$ is CP and unital $\mathbb{J}_{\mathscr{B},\rho}$ is CPTP.
$\mathbb{J}_{\mathscr{B},\rho}$ is orthogonal wrt $\mathrm{tr}[X^{\dagger}\rho^{-\lambda}Y\rho^{\lambda-1}]$ for all $\lambda \in [0,1]$.

$\mathbb{E}_{\mathscr{B},\rho}$ produces a coarse-graining of observables (Heisenberg picture) while
$\mathbb{J}_{\mathscr{B},\rho}$ produces a coarse-graining of states (Schroedinger picture).

Factorization of state extensions

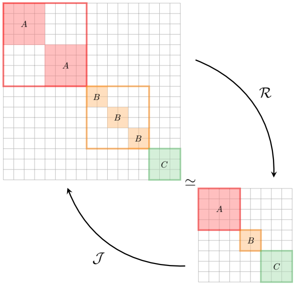
$\mathbb{J}_{\mathscr{B},\rho}[\cdot]$ can be factorized in two non-square CPTP factors

$$\mathbb{J}_{\mathscr{B},\rho}[\cdot] = \mathcal{J}\mathcal{R}$$

such that $\mathcal{R} : \mathscr{A} \to \check{\mathscr{B}}$ and $\mathcal{J} : \check{\mathscr{B}} \to \mathscr{B}$ and $\mathcal{R}\mathcal{J} = \mathcal{I}_{\check{\mathscr{B}}}$ (and also, $\mathbb{E}_{\mathscr{B},\rho} = \mathcal{R}^\dagger\mathcal{J}^\dagger$).

$$\mathcal{R}(X) = \bigoplus_{k=1}^{K} \mathrm{tr}_{\mathcal{H}_{F,k}}(V_k^* X V_k) = \bigoplus_{k=1}^{K} X_{S,k} = \check{X},$$

$$\mathcal{J}(\check{X}) = U\left(\bigoplus_{k=1}^{K} X_{S,k} \otimes \sigma_k\right)U^*.$$



$\mathcal{J}$ is a $*$-homomorphisms, i.e. is an isomorphisms
s.t. $\mathcal{J}(\check{X}\check{Y}) = \mathcal{J}(\check{X})\mathcal{J}(\check{Y})$ and $\mathcal{J}(\check{X}^\dagger) = \mathcal{J}(\check{X})^\dagger$.

Reduction of CP dynamics

Given a CPTP map $\mathcal{A}$, its restriction onto the algebra $\mathscr{B}$,

$$\mathcal{A}|_{\mathscr{B}} = \mathcal{R}\mathcal{A}\mathcal{J}$$

is CPTP.

Given a Lindblad-GKS generator $\mathcal{L}$, its restriction onto the algebra $\mathscr{B}$

$$\mathcal{L}|_{\mathscr{B}} = \mathcal{R}\mathcal{L}\mathcal{J}$$

is still a Lindblad-GKS generator.

Convergence to common fixed points by alternating projections

Alternating projections

Consider two $*$-subalgebras $\mathscr{B}_1, \mathscr{B}_2 \subseteq \mathscr{A}$. Assume there exists a density operator $\rho > 0$ such that $\exists \mathbb{E}_{\mathscr{B}_1,\rho}, \mathbb{E}_{\mathscr{B}_2,\rho}$, that is, ρ is compatible with both $\mathscr{B}_1$ and $\mathscr{B}_2$. Then, $\mathbb{E}_{\mathscr{B}_1,\rho}, \mathbb{E}_{\mathscr{B}_2,\rho}$ are orthogonal wrt the same inner product $\langle \cdot, \cdot \rangle_\rho$.

Theorem 13 (von Neumann-Halperin alternating projections)
Let $\mathcal{H}_1, \ldots, \mathcal{H}_r$ be closed subspaces in a Hilbert space $\mathcal{H}$ ad let $P_{\mathcal{H}_j}$ be the (co-)orthogonal projections. Then

$$\lim_{n\to\infty} (P_{\mathcal{H}_1} \ldots P_{\mathcal{H}_r})^n = P$$

where P is the orthogonal projection onto $\bigcap_{j=1}^r \mathcal{H}_j$.

This implies that

$$\lim_{n\to\infty} (\mathbb{E}_{\mathscr{B}_1,\rho}\mathbb{E}_{\mathscr{B}_2,\rho})^n = \mathbb{E}_{\mathscr{B}_1\cap\mathscr{B}_2,\rho}$$

where $\mathscr{B}_1 \cap \mathscr{B}_2$ is a $*$-subalgebra of $\mathscr{A}$.

Common fixed points

Consider now two CPTP maps $\mathcal{E}_1, \mathcal{E}_2$. Assume $\exists \rho > 0$ such that $\rho \in \mathrm{Fix}(\mathcal{E}_1) \cap \mathrm{Fix}(\mathcal{E}_2)$. Then, $\mathscr{B}_j = \mathrm{Fix}(\mathcal{E}_j^\dagger)$ and their Césaro means are

$$\mathbb{J}_{\mathscr{B}_j,\rho} = \lim_{N\to+\infty} \frac{1}{N} \sum_{k=1}^{N} \mathcal{E}_j^k$$

thus

$$\lim_{n\to\infty} (\mathbb{J}_{\mathscr{B}_1,\rho} \mathbb{J}_{\mathscr{B}_2,\rho})^n = \mathbb{J}_{\mathscr{B}_1\cap\mathscr{B}_2,\rho}$$

where $\mathbb{J}_{\mathscr{B}_1\cap\mathscr{B}_2,\rho}$ is the CPTP projector onto $\mathrm{Fix}(\mathcal{E}_1) \cap \mathrm{Fix}(\mathcal{E}_2)$.

If $\mathrm{Fix}(\mathcal{E}_1) \cap \mathrm{Fix}(\mathcal{E}_2) = \{\rho\}$ then, for all $\rho_0 \in \mathfrak{D}(\mathcal{H})$:

$$\lim_{n\to\infty} (\mathbb{J}_{\mathscr{B}_1,\rho} \mathbb{J}_{\mathscr{B}_2,\rho})^n \rho_0 = \rho.$$

Bayesian parameter estimation

- Classical case;
- Quantum Measurement;
- Hybrid classical-quantum.

Classical Example - Dungeon & Dragons

Assume you are playing D&D and your master, hidden from you, tosses two dice.
The first (fair) die has 6 faces and, depending on its outcome, he tosses a second (fair) die with a number of faces equal to the outcome of the first toss and tells you the outcome of the second toss.
We would like to estimate the value of the first die.

We model this problem by considering two discrete random variables:

- The first, X is hidden from us but we know that is uniformly distributed in $[1, N]$ (with $N = 6$), i.e. $X \sim \mathcal{U}([1, N])$;
- The second instead, Y is uniformly distributed in $[1, X]$, i.e. $Y \sim \mathcal{U}([1, X])$, and we know its outcome y;
- We would like to compute $\mathbb{E}[Z|Y]$ for some random variable $Z = g(X)$ (X, Z are measurable wrt the same σ-algebra).

Probability distribution

We know that $X \sim \mathcal{U}([1, N])$ thus $p_X(x) = \frac{1}{N}$ while $Y \sim \mathcal{U}([1, X])$ hence

$$p_{Y|X}(y|x) = \begin{cases} 1/x & \text{if } y \leq x \\ 0 & \text{otherwise} \end{cases}$$

which implies

$$p_{X,Y}(x, y) = p_{Y|X}(y|x) p_X(x) = \begin{cases} \frac{1}{Nx} & \text{if } y \leq x \\ 0 & \text{otherwise} \end{cases}.$$

Algebraic modelling

Consider an Hilbert spaces $\mathcal{H} = \mathbb{C}^N$. Let $\{|j\rangle\}$ form an orthonormal base for $\mathcal{H}$. We then construct two sub-algebras $\mathscr{B}_1 = \mathrm{alg}\{|j\rangle\langle j| \otimes \mathbb{1}_{\mathcal{H}}\}$ and $\mathscr{B}_2 = \mathrm{alg}\{\mathbb{1}_{\mathcal{H}} \otimes |j\rangle\langle j|\}$ which are subalgebras of $\mathscr{A} = \{|j\rangle\langle j| \otimes |k\rangle\langle k|\} \subset \mathfrak{B}(\mathcal{H} \otimes \mathcal{H})$.

The state of the system then is:

$$\rho = \sum_{x,y=1}^{N} p_{X,Y}(x,y)\,|y\rangle\langle y| \otimes |x\rangle\langle x|$$

and the observables X, Y and Z are

$$Y = \left(\sum_{y=1}^{N} y\,|y\rangle\langle y|\right) \otimes \mathbb{1}_N \in \mathscr{B}_1, \qquad X = \mathbb{1}_N \otimes \left(\sum_{x=1}^{N} x\,|x\rangle\langle x|\right) \in \mathscr{B}_2 \quad \text{and}$$

$$Z = \mathbb{1}_N \otimes \left(\sum_{x=1}^{N} g(x)\,|x\rangle\langle x|\right) \in \mathscr{B}_2,$$

Conditional expectation construction

Conditioning on observations of Y is equivalent to take a CE onto $\mathscr{B}_1$.
Trivially, $\exists \mathbb{E}_{\mathscr{B}_1,\rho}$, as this is just a complex way of representing classical probability theory. In fact, **ρ is compatible with $\mathscr{B}_1$** since they are both diagonal.

Note that $\mathscr{B}_1 = \bigoplus_{y=1}^{N} \mathbb{C} \otimes \mathbb{1}_N$ and

$$\rho = \bigoplus_{y=1}^{N} p_Y(y) \otimes \tau_y \qquad \text{where} \qquad \tau_y = \sum_{x=1}^{N} p_{X|Y}(x|y)\, |x\rangle\langle x| \in \mathbb{C}^{N\times N}$$

hence:

$$\mathbb{E}_{\mathscr{B}_1,\rho}[Z] = \bigoplus_{y=1}^{N} \mathrm{tr}\left[(W_y Z W_y^\dagger)(1 \otimes \tau_y) \right] \otimes \mathbb{1}_N$$

where $W_y = \langle y| \otimes \mathbb{1}_N$.

Double-check

$$\mathbb{E}_{\mathscr{B}_{1},\rho}[Z] = \bigoplus_{y=1}^{N} \mathrm{tr}\left[\langle y|y\rangle \otimes \left(\sum_{x=1}^{N} g(x)\,|x\rangle\langle x|\right)\left(\sum_{x'=1}^{N} p_{X|Y}(x'|y)\,|x'\rangle\langle x'|\right)\right] \otimes \mathbb{1}_N$$

$$= \bigoplus_{y=1}^{N} \underbrace{\sum_{x=1}^{N} g(x) p_{X|Y}(x|y)}_{\mathbb{E}[g(X)|Y=y]} \otimes \mathbb{1}_N = \bigoplus_{y=1}^{N} \mathbb{E}[Z|Y=y] \otimes \mathbb{1}_N$$

y	1	2	3	4	5	6
$\mathbb{P}[Y=y]$	.41	.24	.16	.1	.06	.03
$\mathbb{E}[X\|Y=y]$	2.44	3.44	4.21	4.86	5.45	6

Note:

If the CE exists.

Let us call $\hat{Z}(Y) = \mathbb{E}_{\mathscr{B}_1,\rho}[Z]$. Then:

- $\hat{Z}(Y)$ is an unbiased estimator of Z, i.e. $\mathbb{E}_\rho[Z] = \mathbb{E}_\rho[\hat{Z}(Y)]$;
- Is optimal in the least-square sense (minimal error variance), i.e.

$$\hat{Z} = \arg\min_{f(Y)} \mathbb{E}_\rho[(Z - f(Y))^\dagger (Z - f(Y))]$$

(easily proven by the orthogonality of $\mathbb{E}_{\mathscr{B}_1,\rho}$ wrt $\langle X, Y \rangle_\rho = \mathbb{E}_\rho[XY]$)

Quantum measurement

Let consider a finite-dimensional quantum system $\mathcal{H} \simeq \mathbb{C}^n$ and let $\mathscr{A} = \mathscr{B}(\mathcal{H})$.
Let assume that the system is in a state $\rho \in \mathfrak{D}(\mathcal{H})$ and that we perform the measurement of an observable $O \in \mathfrak{H}(\mathcal{H})$.
For simplicity let us assume that $O = \sum_{j=1}^{n} o_j \, |j\rangle\langle j|$ (no degeneracy and we work in the basis that diagonalizes O).

After the measurement we assume to have another observable, say $X \in \mathfrak{H}(\mathcal{H})$ and we would like to compute the optimal least-square estimator of the state of the quantum system after the measurement. That is, we want to find

$$\arg\min_{f(O)} \mathbb{E}_{?}[(X - f(O))^{\dagger}(X - f(O))].$$

Quantum measurement

After the measurement of O, the state of the quantum system is

$$\rho' = \sum_{j=1}^{n} \frac{|j\rangle\langle j|\,\rho\,|j\rangle\langle j|}{\sum_{k=1}^{n} \langle k|\,\rho\,|k\rangle} = \bigoplus_{j=1}^{n} p_j \quad \text{where} \quad p_j = \frac{\langle j|\,\rho\,|j\rangle}{\sum_{k=1}^{n} \langle k|\,\rho\,|k\rangle}.$$

Note that ρ is compatible with $\mathscr{B} = \mathrm{alg}(\{|j\rangle\langle j|\}_{j=1}^{n})$ hence $\exists\ \mathbb{E}_{\mathscr{B},\rho'}$.
Again:

- $\mathbb{E}_{\mathscr{B},\rho'}[X]$ is an unbiased estimator of X, i.e. $\mathbb{E}_{\rho'}[X] = \mathbb{E}_{\rho'}[\mathbb{E}_{\mathscr{B},\rho}[X]]$;
- is optimal in least-square sense:

$$\mathbb{E}_{\mathscr{B},\rho'}[X] = \arg\min_{f(O)} \mathbb{E}_{\rho'}[(X - f(O))^{\dagger}(X - f(O))].$$

Note that here X does not need to commute with O.

Quantum parameter estimation

Assume you have a hidden discrete random variable X that influences the state of a quantum system on which you can perform measurements. We would like to estimate the hidden random variable based on measurements of the quantum system. (Inspired by recent works by Mankei Tsang).

Consider two Hilbert spaces $\mathcal{H}_Q$ and $\mathcal{H}_C$. Our probability space is $(\mathscr{A}, \rho)$ where

$$\mathscr{A} = \mathfrak{B}(\mathcal{H}_Q) \otimes \operatorname{alg}(\{|j\rangle\langle j|\}_{j=1}^{n_C}) \subset \mathfrak{B}(\mathcal{H}_Q \otimes \mathcal{H}_C)$$

and, given a prior in the hidden variable $p_0 \in \mathbb{R}^{n_C}$ and the initial state of our quantum system $\tau_0 \in \mathfrak{D}(\mathcal{H}_Q)$ we have

$$\rho_0 = \tau_0 \otimes \operatorname{diag}(p_0).$$

Parameter influence

The random variable X then takes value x_j with probability p_j. We can thus model it as

$$X = \mathbb{1}_Q \otimes \left(\sum_{j=1}^{n_C} x_j \left|j\right\rangle\!\left\langle j\right| \right).$$

We here assume that the parameter influences the state of the system trough a parametric unitary rotation, i.e.

$$\tau_{|x} = e^{iHx} \tau_0 e^{-iHx}.$$

After this influence, the state of the joint (quantum-classical) system is

$$\rho_1 = \sum_{j=1}^{n_C} \tau_{|x_j} \otimes p_j \left|j\right\rangle\!\left\langle j\right|.$$

Measurement

We then assume to perform a measurement on the quantum system of a non-degenerate observable $O \in \mathfrak{H}(\mathcal{H}_Q)$, $O = \sum_{k=1}^{n_Q} o_j \left|k\right\rangle\!\left\langle k\right|$.
The state of the system after the measurement is then

$$\rho_2 = \sum_{k=1}^{n_Q} \frac{(\left|k\right\rangle\!\left\langle k\right| \otimes \mathbb{1}_C)\rho_1(\left|k\right\rangle\!\left\langle k\right| \otimes \mathbb{1}_C)}{*} = \sum_{k=1}^{n_Q}\sum_{j=1}^{n_C} \frac{\left\langle k\right| \tau_{|x_j} \left|k\right\rangle p_j}{\sum_{l,m} \left\langle l\right| \tau_{|x_m} \left|l\right\rangle p_m} \left|k\right\rangle\!\left\langle k\right| \otimes \left|j\right\rangle\!\left\langle j\right| .$$

Now, ρ_2 is diagonal, hence is compatible with $\mathscr{B} = \mathrm{alg}(\{\left|k\right\rangle\!\left\langle k\right| \otimes \mathbb{1}_C\})$.

$\exists\ \mathbb{E}_{\mathscr{B},\rho_2}$, and it's the same one we saw in the beginning (classical example).

$$\mathbb{E}_{|\mathscr{B},\rho_2}[X] = \bigoplus_{k=1}^{n_Q} \mathrm{tr}[W_k X W_k^\dagger(\mathbb{1} \otimes \tau_k)] \otimes \mathbb{1}_C$$

Università degli Studi di Padova

Dartmouth

Quantum Model Reduction

Based on [arXiv:2412.05102].

The model

$\mathcal{B}(\mathcal{H}) = \mathbb{C}^{n\times n}$
ρ are density operators:
$\rho \in \mathbb{C}^{n\times n}$, $\rho = \rho^\dagger \geq 0$, $\mathrm{tr}[\rho] = 1$

$\mathcal{L}$ is a Lindblad generator

ρ_0 is the initial condition

$$\begin{cases} \dot{\rho}(t) = \boxed{\mathcal{L}}\left[\,\boxed{\rho(t)}\,\right] \\ \boxed{Y(t)} = \boxed{\mathcal{C}}\,[\rho(t)] \end{cases}, \qquad \boxed{\rho_0 \in \mathfrak{D}(\mathcal{H})}$$

$Y(t) \in \mathbb{C}^{m\times m}$ is the output of interest, the one we want to preserve

$\mathcal{C}$ is a linear output map

The problem: quantum model reduction

Given a Quantum System $(\mathcal{L}, \mathcal{C})$ defined by a generator $\mathcal{L}$ and an output map $\mathcal{C}$ we want to find **another QS** $(\check{\mathcal{L}}, \check{\mathcal{C}})$ and a linear map $\Phi : \mathbb{C}^{n\times n} \to \mathbb{C}^{r\times r}$ such that for all $t \geq 0$ and all $\rho_0 \in \mathfrak{D}(\mathcal{H})$, $\check{\rho}_0 = \Phi[\rho_0]$

- **exact model reduction**
$$\mathcal{C}[e^{\mathcal{L}t}[\rho_0]] = \check{\mathcal{C}}[e^{\check{\mathcal{L}}t}[\check{\rho}_0]];$$
- approximate model reduction (future work)
$$\mathcal{C}[e^{\mathcal{L}t}[\rho_0]] \approx \check{\mathcal{C}}[e^{\check{\mathcal{L}}t}[\check{\rho}_0]].$$

Symmetries

Let U be a unitary operator, $UU^\dagger = \mathbb{1}$, and define $\mathcal{U}(\rho) = U\rho U^\dagger$.
Let $\{\mathcal{T}_t\}_{t\geq 0} = \{e^{\mathcal{L}t}\}_{t\geq 0}$ be the quantum dynamical semigroup generated by $\mathcal{L}$.

Definition 14

U is a symmetry for $\mathcal{L}$ if

$$[\mathcal{T}_t, \mathcal{U}] = 0, \quad \forall t \geq 0.$$

By continuity of the semigroup we have that U is a symmetry if an only if $[\mathcal{L}, \mathcal{U}] = 0$.

Definition 15

- **Strong symmetry** if $[H, U] = [L_k, U] = 0$;
- **Weak symmetry** if $[\mathcal{L}, \mathcal{U}] = 0$.

Note that: (1) Strong implies weak; (2) Symmetries form a group; (3) A symmetry for $\mathcal{L}$ is also a symmetry for $\mathcal{L}^\dagger$.

Symmetries and invariant subspaces

Proposition 3

If U is a symmetry for $\mathcal{L}$, operator eigenspaces of $\mathcal{U}$ are $\mathcal{L}$-invariant.

Proof.

Take $X \in \mathfrak{B}(\mathcal{H})$ such that $\mathcal{U}(X) = \nu X$ and denote by $Y = \mathcal{L}(X)$. Then

$$\mathcal{U}(Y) = \mathcal{U}\mathcal{L}(X) = \mathcal{L}\mathcal{U}(X) = \nu\mathcal{L}(X) = \nu Y$$

hence any ν-eigenoperator of $\mathcal{U}$ is mapped trough $\mathcal{L}$ to a ν-eigenoperator of $\mathcal{U}$, thus the ν-eigenspace is $\mathcal{L}$-invariant. □

Because $\mathcal{U}$ is normal, we can decompose the space of operators into $\mathcal{L}$-invariant subspaces as

$$\mathfrak{B}(\mathcal{H}) = \bigoplus_j \mathrm{eigsp}_{\nu_j}(\mathcal{U}).$$

Invariant algebras

Consider now a unitary subgroup $\mathscr{G}$ of weak symmetries for $\mathcal{L}$,
i.e. $\forall U \in \mathscr{G}$ we have $\mathcal{L}(U \cdot U^\dagger) = U\mathcal{L}(\cdot)U^\dagger$.

The commutant of $\mathscr{G}$, $\mathscr{G}' = \{X \in \mathfrak{B}(\mathcal{H}) | [X, U] = 0, \forall U \in \mathscr{G}\}$:

- is a **unital $*$-algebra**;
- is the intersection of the 1-eigenspaces of symmetries in $\mathscr{G}$,

$$\mathscr{G}' = \bigcap_{U \in \mathscr{G}} \mathrm{eigsp}_1(\mathcal{U});$$

- is $\mathcal{L}$- and $\mathcal{L}^\dagger$-invariant.

But then $\exists \mathbb{E}_{|\mathscr{G}', \mathbb{1}/n}$ which is a CP unital and orthogonal projector (and thus CPTP) onto an invariant subspace!

What does this mean?

Consider $\rho_0 \in \mathscr{G}'$. Because $\mathscr{G}'$ is $\mathcal{L}$-invariant we have that $\rho(t) = e^{\mathcal{L}t}\rho_0 \in \mathscr{G}'$, $\forall t \geq 0$. Then,

$$\rho(t) = \mathbb{E}_{|\mathscr{G},\mathbb{1}/n}[\rho(t)] = \mathbb{E}_{|\mathscr{G},\mathbb{1}/n} e^{\mathbb{E}_{|\mathscr{G},\mathbb{1}/n}\mathcal{L}\mathbb{E}_{|\mathscr{G},\mathbb{1}/n}t}\rho_0.$$

Using $\mathbb{E}_{|\mathscr{G},\mathbb{1}/n} = \mathcal{J}\mathcal{R}$, $\mathcal{R}\mathcal{J} = \mathcal{I}_{\check{\mathscr{A}}}$ the two CPTP factors defined last time, then we have

$$\begin{cases} \dot{\rho}(t) = \mathcal{L}[\rho(t)] \\ Y(t) = \mathcal{C}[\rho(t)] \end{cases} \quad \rho(0) = \rho_0 \qquad \equiv \qquad \begin{cases} \dot{\check{\rho}}(t) = \mathcal{R}\mathcal{L}\mathcal{J}[\check{\rho}(t)] \\ Y(t) = \mathcal{C}\mathcal{J}[\check{\rho}(t)] \end{cases} \quad \check{\rho}(0) = \mathcal{R}(\rho_0)$$

in the sense that

$$\mathcal{C}e^{\mathcal{L}t}(\rho_0) = \mathcal{C}\mathcal{J}e^{\mathcal{R}\mathcal{L}\mathcal{J}t}\mathcal{R}(\rho_0) \qquad \forall t \geq 0.$$

Furthermore, $\mathcal{R}\mathcal{L}\mathcal{J}$ is a Lindblad generator!

Observable symmetry-based reduction

Assume now that $\mathcal{C}(\rho) = \mathrm{tr}[O\rho]$ and $O \in \mathscr{G}'$. Because $\mathscr{G}'$ is ALSO $\mathcal{L}^\dagger$-invariant we have that $O(t) = e^{\mathcal{L}^\dagger t}O \in \mathscr{G}'$, $\forall t \geq 0$.
Then,

$$\mathrm{tr}[O\rho(t)] = \mathrm{tr}[Oe^{\mathcal{L}t}(\rho_0)] = \mathrm{tr}[e^{\mathcal{L}^\dagger t}(O)\rho_0] = \mathrm{tr}[\mathbb{E}_{|\mathscr{G},\mathbb{1}/n}[O(t)]\rho_0].$$

Using again $\mathbb{E}_{|\mathscr{G},\mathbb{1}/n} = \mathcal{J}\mathcal{R}$, $\mathcal{R}\mathcal{J} = \mathcal{I}_{\check{\mathscr{A}}}$ the two CPTP factors defined last time, then we have

$$\begin{cases} \dot{\rho}(t) = \mathcal{L}[\rho(t)] \\ Y(t) = \mathcal{C}[\rho(t)] \end{cases} \qquad \rho(0) = \rho_0 \qquad \equiv \qquad \begin{cases} \dot{\check{\rho}}(t) = \mathcal{R}\mathcal{L}\mathcal{J}[\check{\rho}(t)] \\ Y(t) = \mathcal{C}\mathcal{J}[\check{\rho}(t)] \end{cases} \qquad \check{\rho}(0) = \mathcal{R}(\rho_0)$$

in the sense that

$$\mathcal{C}e^{\mathcal{L}t}(\rho_0) = \mathcal{C}\mathcal{J}e^{\mathcal{R}\mathcal{L}\mathcal{J}t}\mathcal{R}(\rho_0) \qquad \forall t \geq 0.$$

Operator-dependent symmetries

What if we are interested in operators that are not in $\mathscr{G}'$?

Definition 16
A unitary U is an operator-dependent symmetry (ODS) for an observable O if $\mathcal{U}(\rho) = U\rho U^\dagger$ satisfies:

$$[\mathcal{T}_t^\dagger, \mathcal{U}](O) = 0, \quad \forall t \geq 0$$

$$\mathcal{U}(O) = O$$

Equivalent condition $\mathcal{U}\mathcal{L}^{\dagger k}(O) = \mathcal{L}^{\dagger k}\mathcal{U}(O)$ for all $k \in \mathbb{N}$.
Note that all symmetries are ODS for the observables in the commutant of their group.

Operator-dependent symmetries

If we now define $\mathscr{G}$ as a group of ALL ODS for the observable O and generator $\mathcal{L}$, we can prove that $\mathscr{G}'$ **is the minimal $*$-algebra containing** $e^{\mathcal{L}^\dagger t}[O]$ **for all** $t \geq 0$:

$$\mathscr{G}' = \mathrm{alg}\{\mathcal{L}^{\dagger k}[O],\ \forall k \in \mathbb{N}\}.$$

This gives us a numerical method to compute $\mathscr{G}'$.
Using the same procedure as before we can perform model reduction.

Note: $\mathscr{G}'$ is not necessarily $\mathcal{L}^\dagger$-invariant but contains the smallest $\mathcal{L}^\dagger$-invariant subspace that contains O, and this is the important fact.

Example - Central Spin System

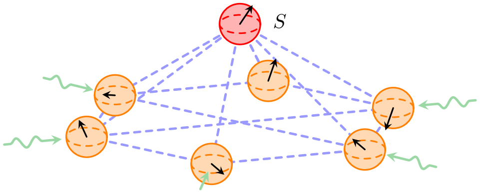
$$H = H_S \otimes \mathbb{1}_B + \frac{1}{2}\left(A_x \sigma_x^{(1)} J_x + A_y \sigma_y^{(1)} J_y + A_z \sigma_z^{(1)} J_z\right)$$

$$L_k^{loc} = \delta \sigma_+^{(k)}, \qquad \forall k \in \text{bath}$$

where $Y(t) = \mathrm{tr}[\rho(t)]$.

Weak symmetries are composed by bath permutations.

If we introduce a bath term
$H_B = \sum_{2 \leq j < k} B_{j,k} \left(\sigma_x^{(j)} \sigma_x^{(k)} + \sigma_y^{(j)} \sigma_y^{(k)} + \sigma_z^{(j)} \sigma_z^{(k)}\right)$
to H, the symmetries become ODS.



Reduction

The dimension of $\mathscr{O} = \mathscr{G}'$ scales with N^3 while the dimension of $\mathfrak{B}(\mathcal{H})$ is 4^N.

The dimension of the largest block grows with N^2.
We can efficiently parallelize the simulation if the symmetry is strong (case with collective dissipation).