Quantum Probability

Scott Morrison

ANU Undergraduate Colloquium

2018-04-18

(Kevin Walker)

Quantum probability

We'll describe quantum probability as a noncommutative analogue of classical probability.

Historically, this was one of the earliest instances of a pervasive theme in 20th century mathematics.

> If widgets are interesting, so are noncommutative (or "quantum") widgets.

Later examples:

(1) $C(X)$, the continuous functions on a (compact, Hausdorff) topological space,
is a commutative C^*-algebra, so
general C^*-algebras are "noncommutative topological spaces".

(2) $U\mathfrak{g}$, the enveloping algebra of a Lie group,
is a cocommutative Hopf algebra, so
general Hopf algebras are "quantum groups".

(3) $\operatorname{Rep} G$, the representation category of a finite group,
is a symmetric fusion category, so
general fusion categories are "quantum symmetries".

In classical probability, we have:

- a set X
- a σ-algebra $\Sigma \subset 2^X$ of "events"
- an assignment of probabilities to events, $\mathbb{P} : \Sigma \longrightarrow [0,1]$

In classical probability, we have:

- a set X
- a σ-algebra $\Sigma \subset 2^X$ of "events"
- an assignment of probabilities to events, $P: \Sigma \longrightarrow [0,1]$

Example:

- $X = \{1, 2, 3, 4\}$
- $\Sigma = 2^X$
- $P(\{1\}) = \frac{1}{2}$
  $P(\{2\}) = \frac{1}{3}$
  $P(\{3\}) = \frac{1}{6}$
  $P(\{4\}) = 0$

In classical probability, we have:

- a set X
- a σ-algebra $\Sigma \subset 2^X$ of "events"
- an assignment of probabilities to events, $P : \Sigma \longrightarrow [0,1]$

Example:

- $X = \{1, 2, 3, 4\}$
- $\Sigma = 2^X$
- $P(\{1\}) = \frac{1}{2}$
  $P(\{2\}) = \frac{1}{3}$
  $P(\{3\}) = \frac{1}{6}$
  $P(\{4\}) = 0$

Given a classical probability space $(X, \Sigma, \mathbb{P})$ we can define the "algebra of observables"

$$L^\infty(X) = \{\text{bounded } \Sigma\text{-measurable functions } X \to \mathbb{C}\}$$

and a "state" on it:

$$\mu : L^\infty(X) \longrightarrow \mathbb{C}$$
$$f \longmapsto \int_X f \, d\mathbb{P}$$

Example:

- $L^\infty(\{1,2,3,4\}) = \mathbb{C}^4$
- $\mu\left(\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix}\right) = \frac{1}{2}w + \frac{1}{3}x + \frac{1}{6}y + 0z$

In fact, we can talk about classical probability purely in terms of observables and states:

A classical probability space consists of:

- a commutative von Neumann algebra $\mathcal{A}$

  (recall an algebra is a vector space where you can also multiply vectors)

  (for today it's enough to think about finite dimensional algebras, and the only finite dimensional commutative von Neumann algebras are $\mathbb{C}^n$ with pointwise multiplication)

- a "state" $\mu : \mathcal{A} \to \mathbb{C}$ which is
  - linear
  - $\mu(1) = 1$
  - $\mu(A^* A) \geqslant 0$

Let's make sure we understand states on the commutative algebra $\mathbb{C}^n$.

- A linear map $\mu : \mathbb{C}^n \longrightarrow \mathbb{C}$ must be of the form

$$\mu(v) = v \cdot w = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$$

for some fixed $w \in \mathbb{C}^n$.

- In the algebra $\mathbb{C}^n$, the multiplicative identity is $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$,

so $\mu(1) = 1$ means $w_1 + w_2 + \cdots + w_n = 1$.

- The operation $* : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ is complex conjugation, so vectors of the form $A^* A$ are those with non-negative real entries.

$\mu(A^* A) \geq 0$ means each $w_i \geq 0$.

$\Longrightarrow$ a state on $\mathbb{C}^n$ is exactly a probability measure on $\{1, 2, \ldots, n\}$.

A classical probability space consists of:

- a commutative von Neumann algebra $\mathcal{A}$ e.g. $\mathbb{C}^n$
- a "state" $\mu : \mathcal{A} \to \mathbb{C}$ which is
  - linear
  - $\mu(1) = 1$
  - $\mu(A^* A) \geq 0$

e.g. $\mu(v) = v \cdot \omega$, $\omega \in \mathbb{R}^n_{\geq 0}$.

An element $f \in \mathcal{A}$ is an "observable", which we can secretly think of as a complex function on some set of underlying "pure states".

The state μ represents our knowledge of the system.

$\mu(f)$ "$= \int_X f \, d\mathbb{P}$" is the expected value of the observable f, given our current knowledge.

A ~~classical~~ quantum probability space consists of:

- a ~~commutative~~ von Neumann algebra $\mathcal{A}$
- a "state" $\mu : \mathcal{A} \to \mathbb{C}$ which is
  - linear
  - $\mu(1) = 1$
  - $\mu(A^*A) \geq 0$

e.g. $\mathcal{A} = M_2(\mathbb{C})$, the 2×2 matrices, with $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^* = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix}$.

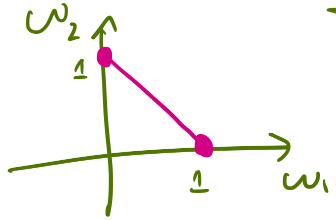
$$\mu\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{2}(a+d).$$

An element $f \in \mathcal{A}$ is still an "observable", but we can't necessarily interpret it as a function on some underlying space.

μ still represents our knowledge of the system, and $\mu(f)$ the expected value of an observable.

Before we study the fundamental example $A = M_2(\mathbb{C})$, lets do some more classical probability.

Recall states on $\mathbb{C}^2$ are probability measures on $\{1,2\}$:

$$\text{states}(\mathbb{C}^2) = \Big\{$$ 

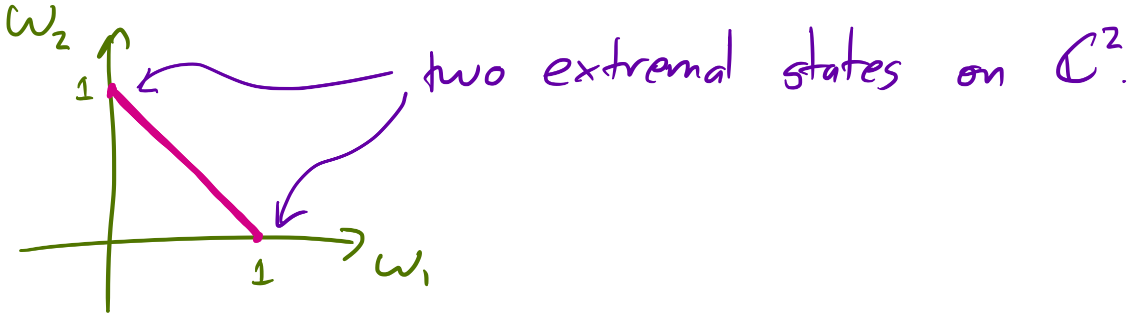
$$\Big\}$$

$$\mu\begin{pmatrix} a \\ b \end{pmatrix} = \omega_1 a + \omega_2 b$$

Notice that if μ and ν are states, so is $x\mu + (1-x)\nu$, for $x \in [0,1]$.

That is, the set of all states is convex: given any two states, the points on the interval between them are also states.

Extremal states (or "pure" states) are states which are not convex linear combinations of other states.



In classical probability, the extremal states correspond to the points of the underlying space — we're 100% certain the system is in a particular configuration.

More generally, **events** (that is subsets $Y \subset X$, so $Y \in \Sigma$) correspond to **projections** in the algebra A, that is, elements f so $f^2 = f = f^*$.

Given a subset $Y \subset X$, the characteristic function χ_Y is a projection.

Let's now work out all this theory in the simplest noncommutative setting: $\mathcal{A} = M_2(\mathbb{C})$.

An event in $M_2(\mathbb{C})$ is a 2×2 matrix E so $E^2 = E = E^*$.

Such a matrix has eigenvalues 0 and 1.

If both are 0, then $E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. (the "impossible" event) rank zero

If both are 1, then $E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. (the "certain" event) rank two

Otherwise rank one $\operatorname{tr} E = 0 + 1 = 1$ and $\det E = 0 \cdot 1 = 0$.

Using $E = E^*$ and $\operatorname{tr} E = 1$, we can write

$$E = \frac{1}{2}\begin{bmatrix} 1+z & x-iy \\ x+iy & 1-z \end{bmatrix} \quad \text{with } x, y, z \in \mathbb{R}.$$

Then $\det E = 0$ says $x^2 + y^2 + z^2 = 1$. Thus "rank one" projections/events in $M_2(\mathbb{C})$ are parametrised by the unit sphere in $\mathbb{R}^3$.

Every rank one projection E defines a state on $M_2(\mathbb{C})$ by the formula $\mu_E(X) = tr(EX)$.

(Notice this is linear in X, $\mu_E(I) = tr(E) = 1$ since E is rank one, and

$$\begin{aligned}\mu_E(A^*A) &= tr(EA^*A)\\ &= tr(E^*EA^*A)\\ &= tr(EA^*AE^*)\\ &= tr((EA^*)(EA^*)^*) \geq 0.\end{aligned}$$

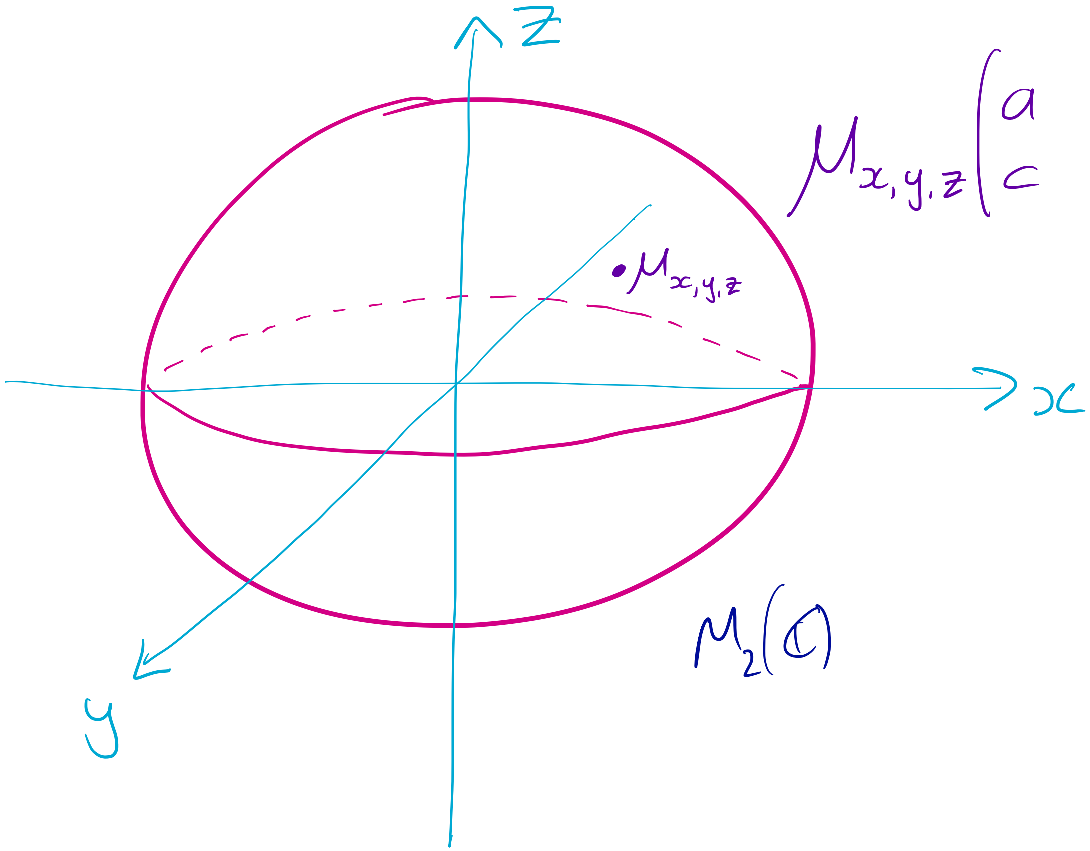
)

In fact, these are all extremal states, and every state is a convex linear combination of these. (Exercise!)

(This is a special case of the "state-operator correspondence":

If A is a <u>positive</u> operator with trace one, $X \longmapsto tr(AX)$ is a state, and every state is uniquely of this form.)

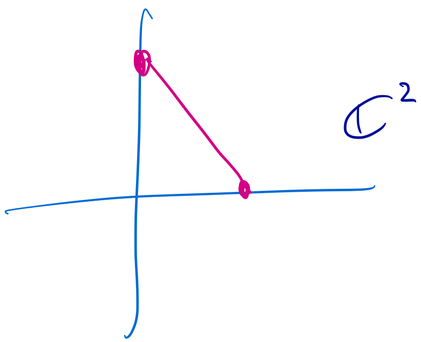
Thus the space of states for $M_2(\mathbb{C})$ is the unit ball in $\mathbb{R}^3$:

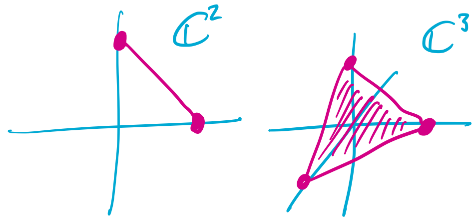


$$\mu_{x,y,z}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1+z}{2}\cdot a + \frac{x-iy}{2}\cdot b + \frac{x+iy}{2}\cdot c + \frac{1-z}{2}\cdot d$$

for $x, y, z \in \mathbb{R}$,

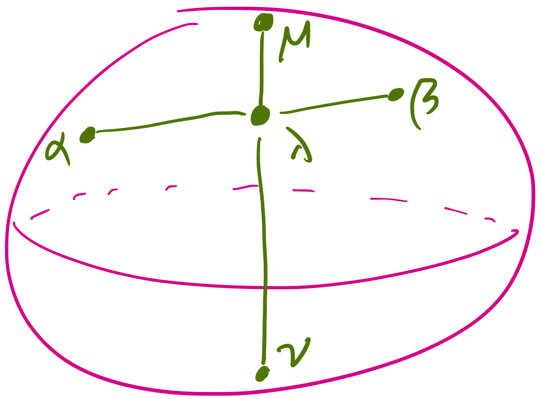
$$x^2+y^2+z^2 \leq 1.$$



Notice a significant difference from the classical setting 

There every state could be written uniquely as a convex linear combination of pure states.

Now we lose uniqueness: there is more than one way to express a point in the interior of the ball as a linear combination of points on the sphere:



$$\lambda = \frac{2}{3}\mu + \frac{1}{3}\nu = \frac{1}{2}\alpha + \frac{1}{2}\beta$$

States are not just probability measures on some underlying set!

Recall that we are thinking of projections $E^2 = E = E^*$ as events.

We think of these as observations that can take values

0 (the event doesn't occur) and

1 (the event does occur).

Recall $\mu(E)$ is the expected value of an observation, so we say E occurs with probability $\mu(E)$.

Remembering that a state represents our knowledge of a system, we should expect it to change after we observe that an event does or does not occur.

$\boxed{\text{How should the state change?}}$

In classical probability (μ is a probability measure on X, events E, F are characteristic functions on subsets of X.)

the rule for updating a state μ to a new state $\hat{\mu}$ after observing an event E is

$$\hat{\mu}(F) = \frac{\mu(EF)}{\mu(E)}$$

"The probability that F occurs given E has occured is the probability E & F occur, divided by the probability E occurs".

Since all classical observables can be written as (or approximated by) linear combinations of events, we have

$$\hat{\mu}(A) = \frac{\mu(EA)}{\mu(E)} \quad \text{for all observables } A.$$

This completely determines the new state.

Unfortunately, this rule doesn't quite work in the quantum world!

However, we can write the classical rule in an equivalent form

$$\hat{\mu}(A) = \frac{\mu(EA)}{\mu(E)} = \frac{\mu(EEA)}{\mu(E)} = \frac{\mu(EAE)}{\mu(E)}$$

events are projections! classically, observables commute!

This one works in the noncommutative setting too.

Let's check it defines a state:

$$\hat{\mu}(1) = \frac{\mu(E1E)}{\mu(E)} = \frac{\mu(E)}{\mu(E)} = 1$$

$$\hat{\mu}(A^*A) = \frac{\mu(EA^*AE)}{\mu(E)} = \frac{\mu(E^*A^*AE)}{\mu(E)} = \frac{\overbrace{\mu((AE)^*(AE))}^{\geq 0}}{\underbrace{\mu(E)}_{\in[0,1]}} \geq 0$$

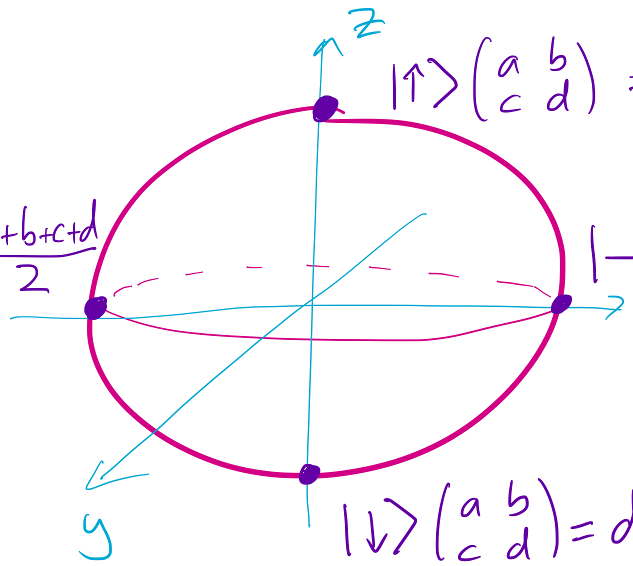
In many presentations of quantum mechanics, people use scary words like wave function collapse at this point.

$\boxed{\text{This is silly.}}$

Nothing more or less complicated is happening than what takes place when you flip a coin, and subsequently look at it.

Recall the space of states for $M_2(\mathbb{C})$ is the unit ball

$$\mu_{x,y,z}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1+z}{2}\cdot a + \frac{x-iy}{2}\cdot b + \frac{x+iy}{2}\cdot c + \frac{1-z}{2}\cdot d$$



$|\uparrow\rangle\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a$

$|\leftarrow\rangle\begin{pmatrix} a & b \\ c & d \end{pmatrix} = -\frac{a+b+c+d}{2}$

$|\rightarrow\rangle\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{a+b+c+d}{2}$

$|\downarrow\rangle\begin{pmatrix} a & b \\ c & d \end{pmatrix} = d$

Let's measure the observable $E = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ on the state $|\uparrow\rangle$.

$|\uparrow\rangle\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1$, so the event E certainly occurs.

When we observe this, the state doesn't change:

$$\hat{\mu}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{|\uparrow\rangle\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\cdot\begin{bmatrix} a & b \\ c & d \end{bmatrix}\cdot\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)}{|\uparrow\rangle\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)} = \frac{|\uparrow\rangle\left(\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}\right)}{|\uparrow\rangle\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)} = \frac{a}{1} = a$$

So $\hat{\mu} = |\uparrow\rangle$. (This is a good sign: if we observe something we're certain occurs, we don't need to update our knowledge.)

What if we measure the event $F = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$?

$|\uparrow\rangle\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\right) = \frac{1}{2}$, so we have a 50% chance of observing F (and hence also a 50% of "not F" $= I - F$)

If we observe F, the new state is

$$\hat{\mu}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \frac{|\uparrow\rangle\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\right)}{|\uparrow\rangle\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\right)} = \frac{|\uparrow\rangle\left(\begin{bmatrix} \frac{a+b+c+d}{4} & '' \\ '' & '' \end{bmatrix}\right)}{|\uparrow\rangle\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\right)} = \frac{a+b+c+d}{2}$$

So $\hat{\mu} = |\rightarrow\rangle$. If, on the other hand, we observe $I - F$, then the new state is $|\leftarrow\rangle$.

We can interpret $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ as "measuring in the z-direction" and $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ as "measuring in the x-direction".

After measuring $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, the state is either $|\uparrow\rangle$ or $|\downarrow\rangle$.

After measuring $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, the state is either $|\rightarrow\rangle$ or $|\leftarrow\rangle$.

In the classical setting, pure states are <u>definite</u>

(that is, $\mu(E) = 0$ or 1 for every event E)

and do <u>not change</u> after any measurement.

In the quantum setting, e.g. $M_2(\mathbb{C})$ there may be

<u>no definite states at all</u>!

It gets even stranger!

Bell's inequality

If A, B, C, D are classical observables taking values in $\{\pm 1\}$

(i.e. $\left(\frac{A+1}{2}\right)^2 = \frac{A+1}{2}$)

then for any state μ,

$$\mu(AB + BC + CD - AD) \leq 2$$

Proof Think of these in a square

$$\begin{tikzcd} A \arrow[r, dash] \arrow[d, dash] & D \arrow[d, dash] \\ B \arrow[r, dash] & C \end{tikzcd}$$

If $A = D$, then $AB + BC + CD - 1 \leq 3 - 1 = 2$

If $A \neq D$, then either $A \neq B$, $B \neq C$, or $C \neq D$.

In any case $AB + BC + CD + 1 \leq 2 - 1 + 1 = 2$.

However in the quantum world it is easy to violate this inequality

and this violation has been experimentally observed!

Example: $\mathcal{A} = M_4(\mathbb{C})$, $\mu(X) = \mathrm{tr}\left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} X\right)$

$$A = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}, \quad B = -\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & 1 \\ & & 1 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} & & 1 & 0 \\ & & 0 & 1 \\ 1 & 0 & & \\ 0 & 1 & & \end{bmatrix}, \quad D = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 & & \\ -1 & -1 & & \\ & & 1 & -1 \\ & & -1 & -1 \end{bmatrix}$$

and (exercise!) $\mu(AB + BC + CD - AD) = \mu(AB) + \mu(BC) + \mu(CD) - \mu(AD)$

$$= \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - -\frac{1}{\sqrt{2}} = 2\sqrt{2} > 2.$$