Begin 10-11-2023

Bayes rule and the updating of probabilities

The Bayesian Update Rule

Consider two stochastic variables A and B . The joint, conditional, and univariate probabilities are related as follows:

$$\left.\begin{aligned} P(A,B)&=P(A|B)\,P(B)\\ P(A,B)&=P(B|A)\,P(A)\end{aligned}\right\} \Rightarrow P(A|B)=\frac{P(B|A)\,P(A)}{P(B)}$$

Thus, with knowledge of $P(A)$ and $P(B)$ we can update our prior knowledge $P(B|A)$ when new information, $P(A|B)$, becomes available.

There are subtleties when working with a mix of probability densitity funticons (pdf's) and discrete data points. Let

α : continuous variable with pdf $p(\alpha)$

B : random discret data point

$p(B|\alpha)$: likelihood function

The Bayesian Update Rule generalizes like this:

$$p(\alpha|B)\,d\alpha=\frac{p(B|\alpha)\,p(\alpha)\,d\alpha}{P(B)}$$

where $P(B)=\int_{-\infty}^{\infty}p(B|\alpha)\,p(\alpha)\,d\alpha$ is a number.

Therefore, to within a normalization factor,

$$\boxed{p(\alpha|B)\propto p(B|\alpha)\,p(\alpha)}$$

See https://math.mit.edu/~dav/05.dir/clss13-slidesall.pdf Page 17

Begin 10-11-2023

Bayes rule and the updating of probabilities

The Bayesian Update Rule generalizes like this:

$$p(\alpha|B)\,d\alpha = \frac{p(B|\alpha)\,p(\alpha)\,d\alpha}{P(B)}$$

where $P(B) = \int_{-\infty}^{\infty} p(B|\alpha)\,p(\alpha)\,d\alpha$ is a number.

Therefore, to within a normalization factor,

$$\boxed{p(\alpha|B) \propto p(B|\alpha)\,p(\alpha)}$$

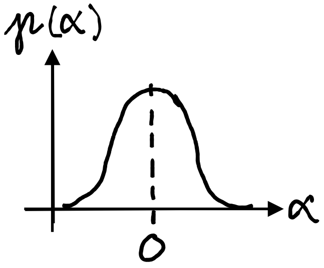
See https://math.mit.edu/~dav/05.dir/clss13-slidesall.pdf Page 17

Bayesian Update of Classical Information

Consider a classical particle located somewhere on the α– axis. The Bayesian interpretation holds that a probability distribution quantifies prior knowledge, in this example about the position of the particle.

Let $p(\alpha)$ be the probability density for finding the particle at position α. We assume this pdf is a Gaussian centered at $\alpha = 0$.

$$p(\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}}\,e^{-\alpha^2/2\sigma_\alpha^2}$$



Next, we measure the position of the particle without disturbing it. The measurement has finite resolution, i. e., there is a change of observing the particle at B even if the actual position is α. This resolution is quantified by the likelihood Function $p(B|\alpha)$

Begin 10-11-2023

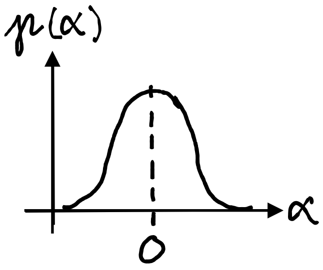
Bayes rule and the updating of probabilities

Bayesian Update of Classical Information

Consider a classical particle located somewhere on the α– axis. The Bayesian interpretation holds that a probability distribution quantifies prior knowledge, in this example about the position of the particle.

Let $p(\alpha)$ be the probability density for finding the particle at position α. We assume this pdf is a Gaussian centered at $\alpha = 0$.

$$p(\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} e^{-\alpha^2/2\sigma_\alpha^2}$$

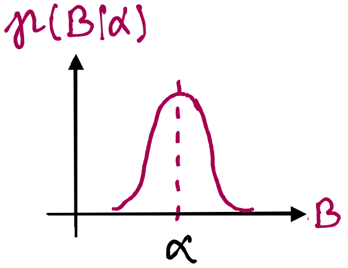


Next, we measure the position of the particle without disturbing it. The measurement has finite resolution, i. e., there is a change of observing the particle at B even if the actual position is α. This resolution is quantified by the likelihood Function $p(B|\alpha)$

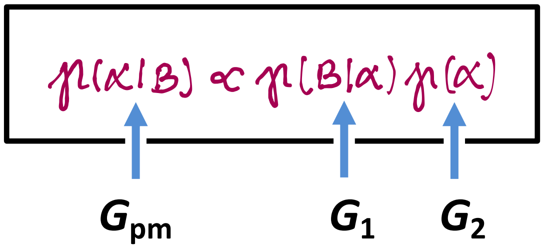
Bayesian Update of Classical Information, cont.

Let $p(B|\alpha)$ be a Gaussian,

$$P(B|\alpha) = \frac{1}{\sqrt{2\pi\sigma_B^2}} e^{-B^2/2\sigma_B^2}$$



Post-measurement, we can use Bayes Rule to update our knowledge of the position of the particle given that we observed B :



$$\boxed{p(\alpha|B) \propto p(B|\alpha)\, p(\alpha)}$$

The product of two Gaussians is a Gaussian, and therefore G_{pm} is also a Gaussian.

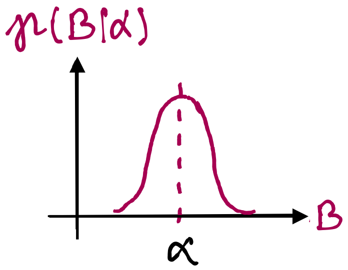
Furthermore, there are exact expressions for the means and σ's of the products, see, e. g.
http://www.lucamartino.altervista.org/2003-003.pdf

Bayes rule and the updating of probabilities

Bayesian Update of Classical Information, cont.

Let $p(B|\alpha)$ be a Gaussian,

$$P(B|\alpha) = \frac{1}{\sqrt{2\pi\sigma_B^2}}\, e^{-B^2/2\sigma_B^2}$$



Post-measurement, we can use Bayes Rule to update our knowledge of the position of the particle given that we observed B :

$$\boxed{\underset{G_{pm}}{\underset{\uparrow}{p(\alpha|B)}} \propto \underset{G_1}{\underset{\uparrow}{p(B|\alpha)}}\, \underset{G_2}{\underset{\uparrow}{p(\alpha)}}}$$

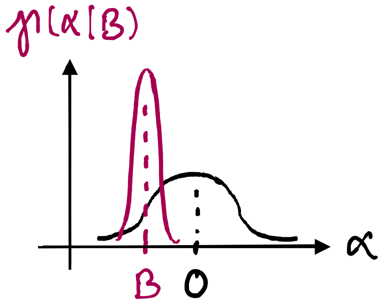
The product of two Gaussians is a Gaussian, and therefore G_{pm} is also a Gaussian.

Furthermore, there are exact expressions for the means and σ's of the products, see, e. g.
http://www.lucamartino.altervista.org/2003-003.pdf

Physical Interpretation, Sharp Measurement

Now let $\sigma_{B|\alpha} << \sigma_\alpha$. Then G_1 is ~ constant over the range where $G_2 \neq 0$. In that case the pdf's will look like this:

$$p(\alpha|B) \approx \frac{1}{\sqrt{2\pi\sigma_B^2}}\, e^{-\alpha^2/2\sigma_{B|\alpha}^2}$$



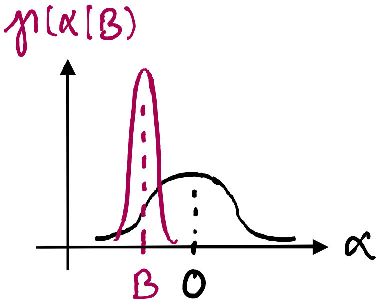
Here we learn a lot from the measurement, and this leads to a large update of our Prior. In this example there will be a large change in the mean and uncertainty that we assign post-measurement. The resulting pdf looks much more like the resolution function than the pdf for the original Gaussian $p(\alpha)$.

Bayes rule and the updating of probabilities

Physical Interpretation, Sharp Measurement

Now let $\sigma_{B|\alpha} << \sigma_{\alpha}$. Then G_1 is ~ constant over the range where $G_2 \neq 0$. In that case the pdf's will look like this:

$$p(\alpha|B) \approx \frac{1}{\sqrt{2\pi\sigma_B^2}} e^{-\alpha^2/2\sigma_{B|\alpha}^2}$$



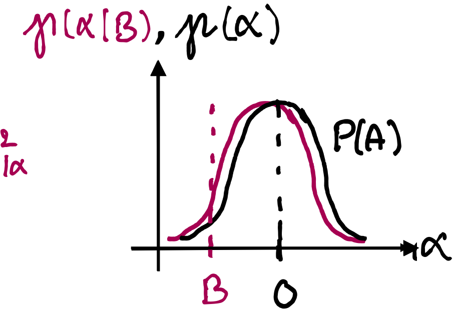
Here we learn a lot from the measurement, and this leads to a large update of our Prior. In this example there will be a large change in the mean and uncertainty that we assign post-measurement. The resulting pdf looks much more like the resolution function than the pdf for the original Gaussian $p(\alpha)$.

Physical Interpretation, Unsharp Measurement

Now let $\sigma_{B|\alpha} \approx \sigma_{\alpha}$. Then G_1 and G_2 are very similar and the pdf's will look like this:

$$p(\alpha|B) \approx \frac{1}{\sqrt{2\pi\sigma_B^2}} e^{-\alpha^2/2\sigma_{B|\alpha}^2}$$

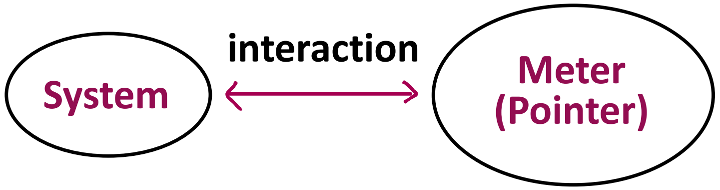
$$\approx p(\alpha)$$



Here we learn little from the measurement and this leads to at most a minor update of our Prior. In this example there will be at most a modest change in the mean and uncertainty that we assign post-measurement. The result looks like a slightly shifted and broadened version of the original.

General Theory of Quantum Measurement (Preskill ch. 3)

Von Neumann's Theory of Measurement



System Observable *M*

Pointer observable

(position *X* of a free particle)

Hamiltonian for the coupled System and Meter

$$H = H_0 + \frac{1}{2m}P^2 + \lambda M P$$

system

free particle

interaction

System-Meter interaction correlates *M* and *x*

➡ Measure *x* → indirect measurement of *M*

Standard Quantum Limit (example)

Heisenberg: $\Delta x \Delta p = \frac{\hbar}{2}$ ➡ $\Delta x(t)^2 \sim \Delta x(0)^2 + \left(\frac{\hbar t}{2m \cdot \Delta x(0)}\right)^2$

Interaction time t ➡ $\Delta x(t) \geq \Delta x_{SQL} \sim \sqrt{\frac{\hbar t}{m}}$

Heavy pointer,
Strong interaction ➡ $\boxed{H = \lambda M P}$

Note: *P* is the generator of translations along *x*

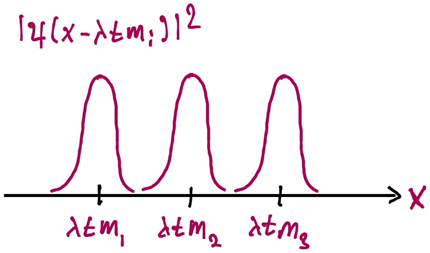
➡ Time evolution $U(t) = e^{-i\lambda t M P/\hbar}$

If	then
$M = \sum_a m_a \vert a\rangle\langle a\vert$	$U(t) = \sum_a \vert a\rangle\langle a\vert \, e^{-i\lambda t m_a P/\hbar}$

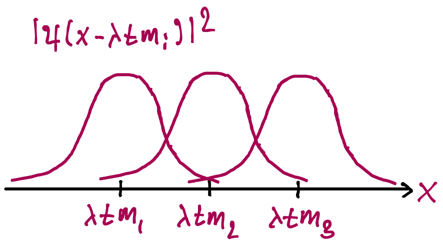
$$U(t) \sum_a \alpha_a |a\rangle \otimes |\psi(x)\rangle = \sum_a \alpha_a |a\rangle \otimes \underbrace{|\psi(x - \lambda t m_a)\rangle}$$

translation along *x* ∝ m_a

Projective



Non - Projective



General Theory of Quantum Measurement (Preskill ch. 3)

Orthogonal Measurement (OM)

Consider a set of measurements $\{E_a\}$ such that

$$E_a = E_a^\dagger \qquad E_a E_{a'} = \delta_{aa'} E_a \qquad \sum_a E_a = \mathbb{1}$$

orthogonal projectors ↑ complete set ↑

We can associate such a set with any observable $M = \sum_a m_a E_a$

This allows us to restate the measurement postulates:

An Orthogonal Measurement of an observable M is described by a collection of operators $\{E_a\}$,

$$E_a = E_a^\dagger \qquad E_a E_{a'} = \delta_{aa'} E_a \qquad \sum_a E_a = \mathbb{1}$$

The outcome m_a occurs w/prob. $P(m_a) = \langle\psi| E_a |\psi\rangle$

➡ the state collapses as $|\psi\rangle \to E_a|\psi\rangle / \sqrt{P(m_a)}$

Mixed state: $P(m_a) = \mathrm{Tr}[E_a \rho]$, $\rho \to E_a \rho E_a / P(m_a)$

m_a degenerate: E_a projects onto subspace

Can we generalize to a broader class? - Yes!

Consider:

$\sum_a E_a = \mathbb{1}$ is required (completeness)

$E_a E_{a'} = \delta_{aa'} E_a$ can be relaxed (orthogonality)

⬇

Concept of non-orthogonal measurements (POVMs)

POVM = Positive Operator Valued Measure

General Theory of Quantum Measurement (Preskill ch. 3)

Bob's OM has 3 outcomes m_a w/projectors $E_a \in \mathcal{H}$

If Alice only prepares states $\rho_A \in \mathcal{H}_A$ then

$$P(m_a) = \mathrm{Tr}[\rho_A E_a] = \mathrm{Tr}[E_A \rho_A E_A E_a]$$

$$= \mathrm{Tr}[\rho_A \underbrace{E_A E_a E_A}_{F_A}] \equiv \mathrm{Tr}[\rho_A F_A]$$

norm ≤1

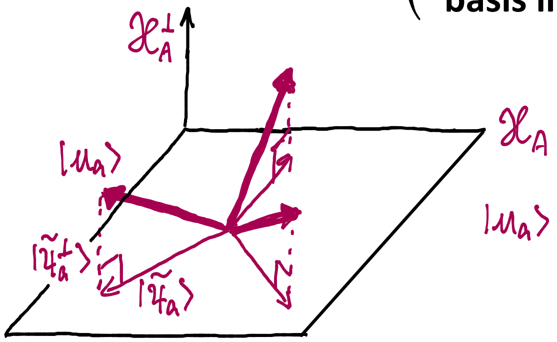
$$= \langle u_a | \rho_A | u_a \rangle = \langle \tilde{\psi}_A | \rho_A | \tilde{\psi}_A \rangle$$

$$= \lambda_a \langle \psi_a | \rho_A | \psi_a \rangle$$

number ≤1

normalized

Geometric visualization: (like an over complete basis in 2D subspace)



$$|u_a\rangle = |\tilde{\psi}_a\rangle + |\tilde{\psi}_a^\perp\rangle$$

We can now define effective measurement operators

$$F_a = E_A E_a E_A = |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a| = \lambda_a |\psi_a\rangle\langle\psi_a|$$

$$\Rightarrow P(m_a) = \mathrm{Tr}[E_a \rho_A] = \mathrm{Tr}[F_a \rho_A]$$

Properties:

- * Each F_a is Hermitian & non-negative ⇒ $P(m_a) \geq 0$
- * Individual F_A are not projectors unless $\lambda_a = 1$
- * $\sum_a F_a = E_A \sum_a E_a E_A = E_A \mathbb{1} E_A = \mathbb{1}_A$ ← identity on $\mathcal{H}_A$

POVM : **P**ositive **O**perator **V**alued **M**easure

A set of non-orthogonal meas. Operators $\{F_a\}$ such that the F_a 's are non-negative & $\sum_a F_a = \mathbb{1}$

General Theory of Quantum Measurement (Preskill ch. 3)

We can now define <u>effective</u> measurement operators

$$F_a = E_A E_a E_A = |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a| = \lambda_a |\psi_a\rangle\langle\psi_a|$$

$$\Rightarrow \quad P(m_a) = Tr[E_a \rho_A] = Tr[F_a \rho_A]$$

Properties:

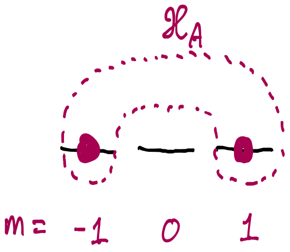
- * Each F_A is Hermitian & non-negative $\Rightarrow$ $P(m_a) \geq 0$
- * Individual F_A are not projectors unless $\lambda_a = 1$
- * $\sum_a F_a = E_A \sum_a E_a E_A = E_A \mathbb{1} E_A = \mathbb{1}_A$ ← identity on $\mathcal{H}_A$

POVM : Positive Operator Valued Measure

A set of non-orthogonal meas. Operators $\{F_a\}$
such that the F_a 's are non-negative & $\sum_a F_a = \mathbb{1}$

Example : POVM on Qubit encoded in Qutrit

$^{87}Rb(F=1)$ atomic HF state

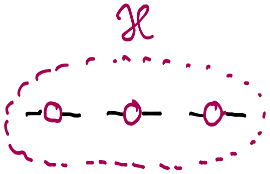


Unitary Transformation

$U: |\tilde{\psi}_1\rangle + |\tilde{\psi}_1^\perp\rangle \rightarrow |-1\rangle$

$U: |\tilde{\psi}_2\rangle + |\tilde{\psi}_2^\perp\rangle \rightarrow |0\rangle$

$U: |\tilde{\psi}_3\rangle + |\tilde{\psi}_3^\perp\rangle \rightarrow |1\rangle$



Alice prepares in $\mathcal{H}_A$

Bob measures

$E_m = |m\rangle\langle m| \in \mathcal{H}$

Choose the map U

$\Rightarrow$ any 1 qubit, 3 outcome POVM we want

Theorem: Any POVM can be realized by adding to $\mathcal{H}_A$ an orthogonal complement $\mathcal{H}_A^\perp$

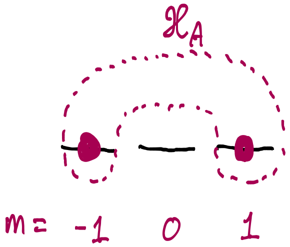
If N F_a's are desired, where $N > Dim\,\mathcal{H}_A$
then we need $Dim(\mathcal{H}_A + \mathcal{H}_A^\perp) \geq N$

(Preskill 3.1.4)

General Theory of Quantum Measurement (Preskill ch. 3)

Example : POVM on Qubit encoded in Qutrit

$^{87}Rb(F=1)$ atomic HF state

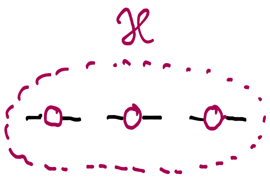


Unitary Transformation

$U: |\tilde{\psi}_1\rangle + |\tilde{\psi}_1^\perp\rangle \to |-1\rangle$

$U: |\tilde{\psi}_2\rangle + |\tilde{\psi}_2^\perp\rangle \to |0\rangle$

$U: |\tilde{\psi}_3\rangle + |\tilde{\psi}_3^\perp\rangle \to |1\rangle$



Alice prepares in $\mathcal{H}_A$

Bob measures

$E_m = |m\rangle\langle m| \in \mathcal{H}$

Choose the map U

- any 1 qubit, 3 outcome POVM we want

Theorem: Any POVM can be realized by adding to $\mathcal{H}_A$ an orthogonal complement $\mathcal{H}_A^\perp$

If N F_a's are desired, where $N > \mathcal{D}im\, \mathcal{H}_A$

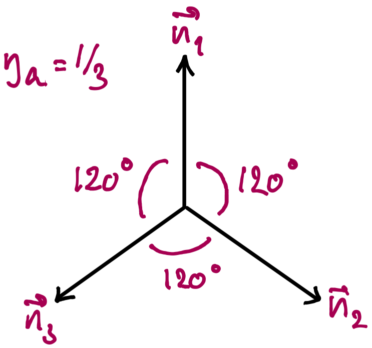
then we need $\mathcal{D}im(\mathcal{H}_A + \mathcal{H}_A^\perp) \geq N$

(Preskill 3.1.4)

Toy Example: One Qubit POVM, illustrates different capabilities of OM & non-OM POVM's

Pick 3 unit vectors s. t. $\sum_a y_a \vec{n}_a = 0, \ \sum_a y_a = 1$

$y_a = 1/3$



Measurement operators

$F_a = 2y_a |\uparrow_{\vec{n}_a}\rangle\langle\uparrow_{\vec{n}_a}| \Rightarrow \sum_a F_a = \mathbb{1}$

For the above & following, note that

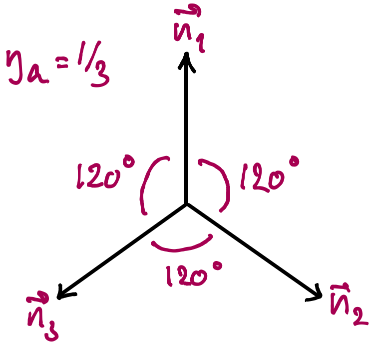
$$|\uparrow_{\vec{n}_2}\rangle = \cos(60°)\,|\uparrow_{\vec{n}_1}\rangle + \sin(60°)\,|\downarrow_{\vec{n}_1}\rangle = \frac{1}{2}|\uparrow_{\vec{n}_1}\rangle + \frac{\sqrt{3}}{2}|\downarrow_{\vec{n}_1}\rangle$$

$$|\uparrow_{\vec{n}_3}\rangle = \cos(60°)\,|\uparrow_{\vec{n}_1}\rangle + \sin(-60°)\,|\downarrow_{\vec{n}_1}\rangle = \frac{1}{2}|\uparrow_{\vec{n}_1}\rangle - \frac{\sqrt{3}}{2}|\downarrow_{\vec{n}_1}\rangle$$

General Theory of Quantum Measurement (Preskill ch. 3)

Toy Example: One Qubit POVM, illustrates different capabilities of OM & non-OM POVM's

Pick 3 unit vectors s. t. $\sum_a y_a \vec{n}_a = 0, \ \sum_a y_a = 1$



Measurement operators

$$F_a = 2 y_a |\uparrow_{\vec{n}_a}\rangle\langle\uparrow_{\vec{n}_a}| \quad \Rightarrow \quad \sum_a F_a = \mathbb{1}$$

For the above & following, note that

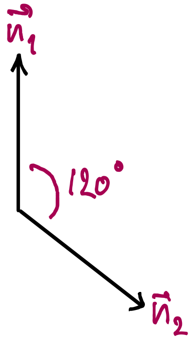
$$|\uparrow_{\vec{n}_2}\rangle = \cos(60^\circ)\,|\uparrow_{\vec{n}_1}\rangle + \sin(60^\circ)\,|\downarrow_{\vec{n}_1}\rangle = \frac{1}{2}|\uparrow_{\vec{n}_1}\rangle + \frac{\sqrt{3}}{2}|\downarrow_{\vec{n}_1}\rangle$$

$$|\uparrow_{\vec{n}_3}\rangle = \cos(60^\circ)\,|\uparrow_{\vec{n}_1}\rangle + \sin(-60^\circ)\,|\downarrow_{\vec{n}_1}\rangle = \frac{1}{2}|\uparrow_{\vec{n}_1}\rangle - \frac{\sqrt{3}}{2}|\downarrow_{\vec{n}_1}\rangle$$

Application: Discriminating between non-orthogonal states

Alice prepares $|\uparrow_{\vec{n}_1}\rangle, |\uparrow_{\vec{n}_2}\rangle$ w/equal probability

How can Bob best tell the difference ?



OM in $\{|\uparrow_{\vec{n}_1}\rangle, |\downarrow_{\vec{n}_1}\rangle\}$ basis ?

Bob's guess

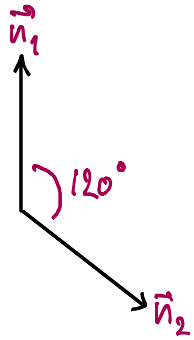
Alice sends
- $|\uparrow_{\vec{n}_1}\rangle$ → Bob gets $|\uparrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 1$ — Bob's guess $|\uparrow_{\vec{n}_1}\rangle$
- $|\uparrow_{\vec{n}_2}\rangle$ → Bob gets
  - $|\uparrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 1/4$ — Bob's guess $|\uparrow_{\vec{n}_1}\rangle$
  - $|\downarrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 3/4$ — Bob's guess $|\uparrow_{\vec{n}_2}\rangle$

Note: Bob can never know for sure he received $|\uparrow_{\vec{n}_1}\rangle$

General Theory of Quantum Measurement (Preskill ch. 3)

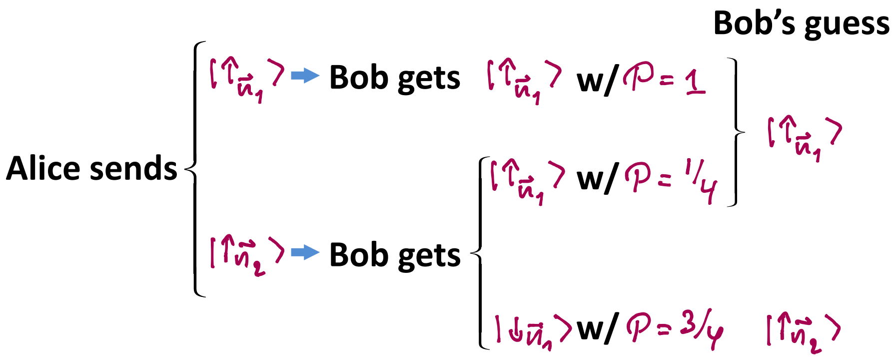
Application: Discriminating between non-orthogonal states

Alice prepares $|\uparrow_{\vec{n}_1}\rangle, |\uparrow_{\vec{n}_2}\rangle$
w/equal probability



How can Bob best tell the difference ?

OM in $\{|\uparrow_{\vec{n}_1}\rangle, |\downarrow_{\vec{n}_1}\rangle\}$ basis ?



Bob's guess

Alice sends
- $|\uparrow_{\vec{n}_1}\rangle$ → Bob gets $|\uparrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 1$ — Bob's guess $|\uparrow_{\vec{n}_1}\rangle$
- $|\uparrow_{\vec{n}_2}\rangle$ → Bob gets
  - $|\uparrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 1/4$ — Bob's guess $|\uparrow_{\vec{n}_1}\rangle$
  - $|\downarrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 3/4$ — Bob's guess $|\uparrow_{\vec{n}_2}\rangle$

Note: Bob can never know for sure he received $|\uparrow_{\vec{n}_1}\rangle$

Fidelity of Bob's guess (Prob. his guess if correct)

$$\mathcal{F}_{POVM} = \frac{1}{2}\times 1 + \frac{1}{2}\left(3/4 \times 1 + 1/4 \times 1/4\right) = \frac{29}{32} \simeq \underline{0.9063}$$

(a) (b) (c) (d) (Quite good)

(a) A sent $|\uparrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 1/2$, B guesses $|\uparrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 1$ $(\mathcal{F} = 1)$

(b) A sent $|\uparrow_{\vec{n}_2}\rangle$ w/ $\mathcal{P} = 1/2$

(c) Given $|\uparrow_{\vec{n}_2}\rangle$ B gets $|\downarrow_{\vec{n}_1}\rangle$ & guesses $|\uparrow_{\vec{n}_2}\rangle$ w/ $\mathcal{P} = 3/4$ $(\mathcal{F} = 1)$

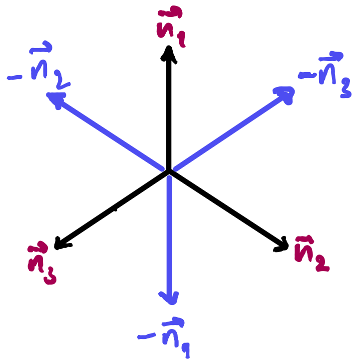
(d) Given $|\uparrow_{\vec{n}_2}\rangle$ B gets $|\uparrow_{\vec{n}_1}\rangle$ & guesses $|\uparrow_{\vec{n}_1}\rangle$ w/ $\mathcal{P} = 1/4$ $(\mathcal{F} = 1/4)$

General Theory of Quantum Measurement (Preskill ch. 3)

Instead

Bob does the POVM

$$\boxed{F_a = \frac{2}{3}\,|\downarrow_{\vec n_a}\rangle\langle\downarrow_{\vec n_a}|}$$



Alice sends

- $|\uparrow_{\vec n_1}\rangle$ → Bob gets
  - $|\downarrow_{\vec n_1}\rangle$ w/ $P = 0$
  - $|\downarrow_{\vec n_2}\rangle$ w/ $P = 1/2$
  - $|\downarrow_{\vec n_3}\rangle$ w/ $P = 1/2$
- $|\uparrow_{\vec n_2}\rangle$ → Bob gets
  - $|\downarrow_{\vec n_1}\rangle$ w/ $P = 1/2$
  - $|\downarrow_{\vec n_2}\rangle$ w/ $P = 0$
  - $|\downarrow_{\vec n_3}\rangle$ w/ $P = 1/2$

Bob gets

- $|\downarrow_{\vec n_1}\rangle$ → Bob knows Alice sent $|\uparrow_{\vec n_2}\rangle$
- $|\downarrow_{\vec n_2}\rangle$ → Bob knows Alice sent $|\uparrow_{\vec n_1}\rangle$
- $|\downarrow_{\vec n_3}\rangle$ → Bob is not sure

Fidelity of Bob's guess (Prob. his guess if correct)

$$\mathcal{F}_{POVM} = \frac{1}{2}\times 1 + \frac{1}{2}\left(\frac{1}{2}\times 1 + \frac{1}{2}\times\frac{1}{4}\right) = \frac{13}{16} = \underline{0.8125}$$

(a) (b) (c)

↑ P(know) ↑ P(don't know)

(a) A sent $|\uparrow_{\vec n_1}\rangle$ or $|\uparrow_{\vec n_2}\rangle$, B knows which one w/ $P = 1/2$ $(\mathcal{F} = 1)$

(b) A sent $|\uparrow_{\vec n_1}\rangle$ or $|\uparrow_{\vec n_2}\rangle$, B DK, correct guess w/ $P = 1/2$ $(\mathcal{F} = 1)$

(c) A sent $|\uparrow_{\vec n_1}\rangle$ or $|\uparrow_{\vec n_2}\rangle$, B DK, wrong guess w/ $P = 1/2$ $(\mathcal{F} = 1/4)$

Note: If in (c) Bob guesses $|\downarrow_{\vec n_3}\rangle$ w/ $\mathcal{F} = 3/4$ he gets a slightly better fidelity of

$$\mathcal{F}_{POVM} = \frac{14}{16} = \underline{0.8750}$$

However: if Bob sticks with **Heralded Success** he will have a subensemble w/ $\mathcal{F}_{POVM} = 1$!