ECE 586 Application: Tensor Products

Henry D. Pfister
Duke University

November 25th, 2022

1 A Few Questions

1.1 What is the transition-probability matrix for a Markov chain $Z_t = (X_t, Y_t)$ where X_t and Y_t are independently evolving Markov chains?

Let $P \in \mathbb{R}^{m\times m}$ be transition-probability matrix for X_t on the state space $[m]$ and let $Q \in \mathbb{R}^{n\times n}$ be transition-probability matrix for Y_t on the state space $[n]$. Let $\pi^{(t)}_{(x,y)} \coloneqq \Pr(X_t = x, Y_t = y)$ be the joint distribution of X_t and Y_t. Then, independence implies that

$$\begin{aligned}
\pi^{(t+1)}_{(x',y')} &= \Pr(X_{t+1} = x', Y_{t+1} = y') \\
&= \sum_{(x,y)\in[m]\times[n]} \Pr(X_t = x, Y_t = y) \Pr(X_{t+1} = x' | X_t = x) \Pr(Y_t = y' | Y_t = y) \\
&= \sum_{(x,y)\in[m]\times[n]} \pi^{(t)}_{(x,y)} P_{x,x'} Q_{y,y'} \\
&= \sum_{(x,y)\in[m]\times[n]} \pi^{(t)}_{(x,y)} T_{(x,y),(x',y')},
\end{aligned}$$

where $T_{(x,y),(x',y')} = P_{x,x'}Q_{y,y'}$ is the transition probability matrix indexed by (x, y) pairs.

Now, we can use the mapping $\tau\colon [m] \times [n] \to [mn]$ defined by $(x, y) \mapsto (x - 1)n + y$ to reindex $(x, y) \in [m] \times [n]$ by $z \in [mn]$. Using the change of variables (x', y') to $z' = (x' - 1)n + y'$, we see that

$$\begin{aligned}
\pi^{(t+1)}_{z'} &= \pi^{(t+1)}_{(x'-1)n+y'} \\
&= \sum_{(x,y)\in[m]\times[n]} \pi^{(t)}_{(x-1)n+y} T_{(x-1)n+y,(x'-1)n+y'} \\
&= \sum_{z=1}^{mn} \pi^{(t)}_z T_{z,z},
\end{aligned}$$

where $T_{z,z'} = P_{\lfloor (z-1)/n\rfloor+1,\lfloor (z'-1)/n\rfloor+1} Q_{(z-1) \bmod n+1,(z'-1) \bmod n+1}$ because τ is invertible and $\tau^{-1}(z) = \big(\lfloor (z-1)/n\rfloor + 1, (z-1) \bmod n + 1\big)$. The matrix $T \in \mathbb{R}^{mn\times mn}$ is the Kronecker product of P and Q.

For an arbitrary field F, the Kronecker product of $A \in F^{m\times n}$ and $B \in F^{p\times q}$ is the $mp \times nq$ matrix defined by the block matrix

$$A \otimes B \triangleq \begin{bmatrix} A_{1,1}B & A_{1,2}B & \cdots & A_{1,n}B \\ A_{2,1}B & A_{2,2}B & \cdots & A_{2,n}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1}B & A_{m,2}B & \cdots & A_{m,n}B \end{bmatrix}.$$

1.2 Consider closed quantum systems A and B with m and n distinguishable states, respectively. Suppose they are combined into a system AB with mn states so that joint operations on both systems are possible. If $G \in \mathbb{C}^{m\times m}$ and $H \in \mathbb{C}^{n\times n}$ define unitary operations on A and B, then what matrix describes their simultaneous independent action on AB?

A closed quantum system with m perfectly distinguishable states is completely characterized by a unit vector $\underline{u}$ in the Hilbert space $\mathbb{C}^m$. Since the combined system has a total of mn distinguishable states, it can be completely characterized by a vector $\underline{w}$ in the Hilbert space $\mathbb{C}^{mn}$. Let $\underline{u}_1, \ldots, \underline{u}_m \in \mathbb{C}^m$ and $\underline{v}_1, \ldots, \underline{v}_n \in \mathbb{C}^n$ be separate orthonormal bases for the two systems. One can naturally associate the distinguishable states of the combined system $\mathbb{C}^{mn}$ with ordered pairs $(\underline{u}_i, \underline{v}_j) \in \mathbb{C}^m \times \mathbb{C}^n$, for $(i,j) \in [m]\times[n]$, of the orthonormal basis vectors for the two systems . In particular, we can use them to define orthonormal basis vectors $\underline{w}_{i,j} = \underline{u}_i \otimes \underline{v}_j \in \mathbb{C}^{mn}$ for $(i,j) \in [m]\times[n]$, where $\otimes$ denotes the Kronecker product defined above.

The unitary action of G on an A state vector in $\mathbb{C}^m$ is defined bv $\underline{u}_i \mapsto G\underline{u}_i$. Similarly, the unitary action of H on a B state vector in $\mathbb{C}^n$ is defined by $\underline{v}_j \mapsto H\underline{v}_j$. Thus, the independent unitary action of G and H on an AB state vector in $\mathbb{C}^{mn}$ is defined by $\underline{w}_{i,j} \mapsto G\underline{u}_i \otimes H\underline{v}_j$. A well-known property of the Kronecker product is that

$$(G \otimes H)(\underline{u}_i \otimes \underline{v}_j) = G\underline{u}_i \otimes H\underline{v}_j.$$

Thus, the independent unitary action of G on A and H on B for an AB state vector in $\mathbb{C}^{mn}$ is defined by the unitary matrix $W = G \otimes H$.

1.3 For vector spaces $U = \mathbb{R}^m$ and $V = \mathbb{R}^n$, how can one represent the set of bilinear functionals mapping $U \times V \to \mathbb{R}$?

For any basis $\underline{u}_1, \ldots, \underline{u}_m$ of U, there exist linear coordinate functions $s_1(\underline{u}), \ldots, s_m(\underline{u})$ such that $\underline{u} = \sum_{i=1}^m s_i(\underline{u})\underline{u}_i$ for all $\underline{u} \in U$. It follows that any linear functional $f\colon U \to \mathbb{R}$ satisfies

$$f(\underline{u}) = f\left(\sum_{i=1}^m s_i(\underline{u})\underline{u}_i\right) = \sum_{i=1}^m s_i(\underline{u})\, f(\underline{u}_i)$$

and, thus, $f(\underline{u})$ is uniquely defined by $f(\underline{u}_1), \ldots, f(\underline{u}_m)$. Thus, the set of linear functionals on $\mathbb{R}^m$ is a vector space of dimension m.

A function $g\colon U \times V \to \mathbb{R}$ is called *bilinear* if it is separately linear in its arguments:

$$\begin{aligned} g(s\underline{u}_1 + \underline{u}_2, \underline{v}) &= sg(\underline{u}_1, \underline{v}) + g(\underline{u}_2, \underline{v}) \quad s \in \mathbb{R},\ \underline{u}_1, \underline{u}_2 \in U,\ \underline{v} \in V \\ g(\underline{u}, s\underline{v}_1 + \underline{v}_2) &= sg(\underline{u}, \underline{v}_1) + g(\underline{u}, \underline{v}_2) \quad s \in \mathbb{R},\ \underline{u} \in U,\ \underline{v}_1, \underline{v}_2 \in V. \end{aligned}$$

For example, the inner product on a real vector space is a bilinear function. For any basis $\underline{v}_1, \ldots, \underline{v}_m$ of U, there exist linear coordinate functions $t_1(\underline{v}), \ldots, t_n(\underline{v})$ such that $\underline{v} = \sum_{i=1}^m t_i(\underline{v})\underline{v}_i$ for all $\underline{v} \in V$. Similar to the linear case, any bilinear functional $g\colon U \times V \to \mathbb{R}$ satisfies

$$\begin{aligned} g(\underline{u}, \underline{v}) &= g\left(\sum_{i=1}^m s_i(\underline{u})\underline{u}_i, \underline{v}\right) \\ &= \sum_{i=1}^m s_i(\underline{u}) g\left(\underline{u}_i, \sum_{j=1}^n t_j(\underline{u})\underline{v}_j\right) \\ &= \sum_{i=1}^m \sum_{j=1}^n s_i(\underline{u}) t_j(\underline{u}) g\left(\underline{u}_i, \underline{v}_j\right) \end{aligned}$$

and, thus, the function $g(\underline{u}, \underline{v})$ is uniquely defined by $G_{i,j} = g(\underline{u}_i, \underline{v}_j)$ for $i, j \in [m]\times[n]$. This shows that the set of bilinear functionals on $\mathbb{R}^m \times \mathbb{R}^n$ is a vector space of dimension mn. Moreover, if standard bases are used for U and V, then we have $g(\underline{u},\underline{v}) = \underline{v}^T G\underline{u}$. A key difference here is that, while the space of linear functionals on the direct product $U \times V$ has dimension $m+n$, the space of bilinear functionals has dimension mn.

A bilinear functional is called *degenerate* if there is a non-zero $\underline{u} \in U$ such that $g(\underline{u},\underline{v}) = 0$ for all $\underline{v} \in V$ or if there is a non-zero $\underline{v} \in V$ such that $g(\underline{u},\underline{v}) = 0$ for all $\underline{u} \in U$. It is verify to see that g is degenerate if and only if it's G matrix is singular.

Remark 1. All of the above statements and derivations remain valid if $\mathbb{R}$ is replaced by any other field. Though, for the case of $\mathbb{C}$, it is common to see *sesquilinear* functions, which are linear in one argument and conjugate linear in the other. These functions arise naturally from the definition $g(\underline{u},\underline{v}) = \underline{v}^H G\underline{u}$.

2 Tensor Product Spaces

One challenge when discussing tensor products of vector spaces is that they have multiple equivalent definitions. Thus, different papers may use different definitions. Practitioners tend to prefer concrete definitions that are described in terms of the basis vectors of each space whereas mathematicians often prefer abstract definitions that do not refer to the basis vectors. Regardless, the tensor product space is independent of the bases used in the sense that there is an invertible change of basis that connects any two concrete constructions.

2.1 Concrete construction

For vector spaces $U = F^m$ and $V = F^n$ with fixed bases, the tensor product $U \otimes V$ (note: the symbol $\otimes$ here does not represent the Kronecker product) is a vector space of dimension $\dim(U)\dim(V) = mn$ that is spanned by all ordered pairs of basis vectors for U and V. In particular, if we let $\underline{u}_1, \ldots, \underline{u}_m \in F^m$ and $\underline{v}_1, \ldots, \underline{v}_n \in F^n$ be bases for U and V, then the induced basis for $U \otimes V$ is given by $\underline{w}_{i,j} = \underline{u}_i \otimes \underline{v}_j \in F^{mn}$ for $(i,j) \in [m] \times [n]$, where $\otimes$ represents the Kronecker product. Thus, any element of $U \otimes V$ can be written in the form

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}\underline{w}_{i,j} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}(\underline{u}_i \otimes \underline{v}_j),$$

for some $a_{i,j} \in F^{m\times n}$.

For $P \in F^{m\times m}$ and $Q \in F^{n\times n}$, consider a linear transform $T\colon U\otimes V \to U\otimes V$ defined by $T(\underline{u}_i \otimes \underline{v}_j) = P\underline{u}_i \otimes Q\underline{v}_j$. Such a linear transform T is called the tensor product of the linear transforms P and Q. Moreover, on the space F^{mn}, one can represent T concretely as a matrix $T = P \otimes Q$, where $\otimes$ denotes the Kronecker product. More generally, one has the following multiplication identity for Kronecker products. For $A_1 \in F^{m\times n}$, $A_2 \in F^{n\times p}$, $B_1 \in F^{q\times r}$, and $B_2 \in F^{r\times s}$, the matrix multiplication of Kronecker products satisfies

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2 \otimes B_1 B_2). \tag{1}$$

It also holds that $(A_1 \otimes B_1)^H = (A_1^H \otimes B_1^H)$. For vectors $\underline{u},\underline{v},\underline{w},\underline{x}$ where $\underline{u}^H\underline{w}$ and $\underline{v}^H\underline{x}$ are scalars, we can combine these rules to see that

$$(\underline{u}\otimes\underline{v})^H(\underline{w}\otimes\underline{x}) = (\underline{u}^H \otimes \underline{v}^H)(\underline{w}\otimes\underline{x}) = (\underline{u}^H\underline{w} \otimes \underline{v}^H\underline{x}) = (\underline{u}^H\underline{w})(\underline{v}^H\underline{x}).$$

Thus, if U and V are inner product spaces, then this implies that $U \otimes V$ can be seen as an inner product space where the inner product is defined by

$$\langle \underline{w}\otimes\underline{x}, \underline{u}\otimes\underline{v}\rangle_{U\otimes V} \coloneqq \langle \underline{w},\underline{u}\rangle_U \langle \underline{x},\underline{v}\rangle_V .$$

Exercise 1. (10 pts) Verify equation (1) explicitly using $m = 2$, $n = 3$,

$$A_1 = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix},\ A_2 = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix},\ B_1 = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 1 & -1 \\ -1 & 2 & 2 \end{bmatrix},\ \text{and } B_2 = \begin{bmatrix} 2 & -1 & 1 \\ 1 & -1 & -2 \\ -2 & 2 & 2 \end{bmatrix}.$$

2.2 Bilinear functionals

For the bilinear functional $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined by $g(\underline{w}, \underline{x}) = \underline{x}^T G \underline{w}$, can one change bases orthogonally so that the bilinear form becomes diagonal? Consider computing the SVD $G = U\Sigma V^T$ and notice that $\tilde{g}(\underline{y}, \underline{z}) = g(V\underline{y}, U\underline{z})$ satisfies

$$\tilde{g}(\underline{y}, \underline{z}) = \underline{z}^T U^T U \Sigma V^T V \underline{y} = \underline{z}^T \Sigma \underline{y}.$$

Exercise 2. (10 pts) Consider the bilinear functional defined by the matrix

$$G = \begin{bmatrix} 1 & -1 & 2 & 2 \\ 1 & 1 & -1 & 0 \\ -1 & 2 & 2 & 1 \end{bmatrix}.$$

Is this bilinear functional degenerate? Use the SVD of G to find an orthonormal change of basis for the domain and codomain so that the bilinear form is diagonal afterward. Define an equivalence relation $\sim$ on $\mathbb{R}^n$ so that $\underline{w} \sim \underline{w}'$ implies $g(\underline{w}, \underline{x}) = g(\underline{w}', \underline{x})$ for all $\underline{x} \in \mathbb{R}^m$. Use this equivalence relation to define a non-degenerate bilinear functional $h : \mathbb{R}^n / \sim \times \mathbb{R}^m \to \mathbb{R}$ such that $h(\underline{w}, \underline{x}) = g(\underline{w}, \underline{x})$ for all $\underline{w} \in \mathbb{R}^n$ and $\underline{x} \in \mathbb{R}^m$. Is there a natural way to define h in terms of the SVD of G?

2.3 Markov chains

Using the setup from Section 1.1, let $Q \in \mathbb{R}^{n \times n}$ be the transition-probability matrix for a Markov chain with n states and define $T = Q \otimes Q$ to be its Kronecker product with itself. Then, T is the transition-probability matrix for the Markov chain that jointly models two copies of the Markov chain defined by Q as they evolve independently. Specifically, T is is the transition-probability matrix for the joint Markov chain $(X_t, Y_t) \in [n]^2$ where X_t and Y_t evolve independently according to Q .

To make use of this setup, we need a few more tidbits from probability and Markov chains. First, for a function $f \colon [n] \to \mathbb{R}$, we have the simple conditional expectation formula

$$\begin{aligned} \mathbb{E}[f(X_{t+1}) \,|\, X_1 = i] &= \sum_{j=1}^{n} \Pr(X_{t+1} = j \,|\, X_1 = i)\, f(j) \\ &= \sum_{j=1}^{n} \left[Q^t\right]_{i,j} f(j) \\ &= \underline{e}_i Q^t \underline{f}, \end{aligned}$$

where $\underline{e}_i$ is the i-th standard basis vector and $\underline{f} = [f(1), f(2), \ldots, f(n)]^T$. If the Markov chain defined by Q has a unique stationary distribution $\underline{\pi} = [\pi_1, \ldots, \pi_n]^T$ satisfying $\underline{\pi}^T Q = \underline{\pi}^T$, then that implies

$$\lim_{t \to \infty} \mathbb{E}[f(X_{t+1}) \,|\, X_1 = i] = \underline{\pi}^T \underline{f} = \mathbb{E}\left[f(X)\right],$$

where X is drawn from the stationary distribution $\underline{\pi}$. Second, we have the general variance identity

$$\mathrm{Var}\left(f(Z)\right) = \frac{1}{2} \mathbb{E}\left[\left(f(Z) - f(Z')\right)^2\right],$$

where Z' is an i.i.d. copy of Z. This is easily verified by expanding both sides of the equality.

Now, we can combine these results to write the variance of $f(X_t)$ as the expected value of a linear function over the joint chain (X_t, Y_t). In particular, if the chains are initially synchronized (i.e., $X_1 = Y_1 = i$), then we have

$$\begin{aligned} \mathrm{Var}\left(f(X_t) \,|\, X_1 = i\right) &= \frac{1}{2} \mathbb{E}\left[\left(f(X_t) - f(Y_t)\right)^2 \,\middle|\, X_1 = Y_1 = i\right] \\ &= \mathbb{E}\left[g(X_t, Y_t) \,|\, X_1 = Y_1 = i\right] \\ &= (\underline{e}_i \otimes \underline{e}_i) T^{t-1} \underline{g}, \end{aligned}$$

where $g(x,y) = \frac{1}{2}(f(x) - f(y))^2$ and

$$\underline{g} = [g(1,1), g(1,2), \ldots, g(1,n), g(2,1), \ldots, g(2,n), g(3,1), \ldots, g(n,n)] .$$

Exercise 3. (45 pts) Consider the Markov chain defined by the matrix

$$Q = \begin{bmatrix} 0.90 & 0.10 & 0 & 0 \\ 0.15 & 0.80 & 0.05 & 0 \\ 0 & 0 & 0.95 & 0.05 \\ 0 & 0.05 & 0 & 0.95 \end{bmatrix} .$$

1. For $f(x) = 2x - 5$, use the ideas above to compute $m_{i,t} = \mathbb{E}\left[f(X_t) \,|\, X_1 = i\right]$ for all $i \in [n]$ and $t \in [10]$. Compare your your answer with $\lim_{t\to\infty} \mathbb{E}\left[f(X_t) \,|\, X_1 = i\right] = \underline{\pi}^T \underline{f}$ where $\underline{\pi}$ is the unique stationary distribution.

2. For $g(x,y) = f(x)f(y)$, use the ideas above with $T = Q \otimes Q$ to compute $\rho_{i,j,t} = \mathbb{E}\left[g(X_t, Y_t) \,|\, X_1 = i, Y_1 = j\right]$ for all $i,j \in [n]$ and $t \in [10]$. Compare your your answer with $\lim_{t\to\infty} \mathbb{E}\left[g(X_t,Y_t) \,|\, X_1 = i, Y_1 = j\right] = (\underline{\pi} \otimes \underline{\pi})^T (\underline{f} \otimes \underline{f})$. To what does this expression simplify?

3. For $g(x,y) = \frac{1}{2}(f(x) - f(y))^2$, use the ideas above with $T = Q \otimes Q$ to compute $v_{i,j,t} = \mathbb{E}\left[g(X_t, Y_t)\right]$ for all $i,j \in [n]$ and $t \in [10]$. Compare your answer with $\lim_{t\to\infty} \mathbb{E}\left[g(X_t,Y_t)\right] = (\underline{\pi} \otimes \underline{\pi})^T \underline{g}$ and a direct computation of variance based on the stationary distribution $\underline{\pi}$.

3 Quantum Mechanics

3.1 What does all of this have to do with quantum mechanics?

Suppose we have a quantum system with mn computational basis states. Then, the state of the system is defined by a complex vector $\underline{w}$ in the standard Hilbert space $\mathbb{C}^{mn}$ satisfying $\|\underline{w}\| = 1$. The primary operations on a quantum computer are:

- Unitary evolution: A unitary operation $\underline{U}$ is applied to the state $\underline{w}$ so that the new state is $U\underline{w}$,

- Projective measurement: Given k orthogonal subspaces defined by orthogonal projectors $P_1, \ldots, P_k$ such that $\sum_{i=1}^k P_i = I$, the measurement process randomly applies one of the projectors and reveals which projection occurred. The resulting vector (known as the *post-measurement state*) is renormalized to unit length and the probability that the i-th projection occurs is given by

$$p_i = \underline{w}^T P_i \underline{w}.$$

If the system is composed of a subsystem A with m computational basis states and a subsystem B with n computational basis states, then one can also associate the state with $\underline{w} \in \mathbb{C}^m \otimes \mathbb{C}^n$. This representation highlights the fact that someone (e.g, named Alice) with access to system A can apply physical operations to that system. For example, applying a unitary operation U_A to system A changes the state to $(U_A \otimes I_B)\underline{w}$, where I_B is the identity operator on subsystem B. Similarly, someone (e.g., named Bob) with access to system B can apply physical operations to that system. For example, applying a unitary operation V_B to system B changes the state to $(I_A \otimes V_B)\underline{w}$, where I_A is the identity operator on subsystem A. In particular, the set of bilinear unitary transformations on $\mathbb{C}^m \otimes \mathbb{C}^n$ contains all unitary quantum operations that can be implemented separately on the two subsystems.

Consider a pure state in the tensor product space $C^m \otimes \mathbb{C}^n$. Such a pure state can be defined by a vector in $\underline{w} \in \mathbb{C}^{mn}$ but one can also highlight its tensor product representation by writing it as an element of $A \in \mathbb{C}^{m\times n}$, where

$$\underline{w} = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{i,j} (\underline{e}_i \otimes \underline{e}_j).$$

In this case, it follows from (1) that

$$
\begin{aligned}
(U^H \otimes I)\underline{w} &= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}(U^H \underline{e}_i \otimes \underline{e}_j) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j} \sum_{k=1}^{m} \overline{U}_{i,k}(\underline{e}_k \otimes \underline{e}_j) \\
&= \sum_{k=1}^{m}\sum_{j=1}^{n} \underbrace{\left(\sum_{i=1}^{m} \overline{U}_{i,k} A_{i,j} \right)}_{[U^H A]_{k,j}} (\underline{e}_k \otimes \underline{e}_j).
\end{aligned}
$$

Thus, the action of $U^H \otimes I$ on $\underline{w}$ maps its matrix representation A to $U^H A$. Similarly, one can show that

$$
(I \otimes V)\underline{w} = \sum_{i=1}^{m}\sum_{k=1}^{n} \underbrace{\left(\sum_{j=1}^{m} A_{i,j} V_{j,k} \right)}_{[AV]_{i,k}} (\underline{e}_i \otimes \underline{e}_k)
$$

Exercise 4. (45 pts) Consider the pure state $\underline{w} \in \mathbb{C}^3 \otimes \mathbb{C}^4$ associated with the matrix representation

$$
A = \begin{bmatrix} \frac{18}{175} & \frac{81-112\sqrt{2}}{525} & \frac{162+56\sqrt{2}}{525} & -\frac{16\sqrt{2}}{75} \\ \frac{24}{175} & \frac{36+28\sqrt{2}}{175} & \frac{72-14\sqrt{2}}{175} & \frac{4\sqrt{2}}{25} \\ 0 & \frac{4\sqrt{2}}{15} & -\frac{2\sqrt{2}}{15} & \frac{4\sqrt{2}}{15} \end{bmatrix}
$$

in the tensor product space $W = \mathbb{C}^3 \otimes \mathbb{C}^4$. Using the SVD $A = U\Sigma V^H$, we observe applying the local unitaries $U^H \otimes I$ and $V \otimes I$ to $\underline{w}$ (in either order) maps its matrix representation A to $U^H AV = U^H(U\Sigma V^H)V = \Sigma$.

1. Let $\underline{u}_i$ be the i-th column of U. Suppose Alice implements the local projective measurement with 3 outcomes defined by $P_i = (\underline{u}_i \underline{u}_i^H \otimes I_B)$ for $i \in \{1, 2, 3\}$. Describe the resulting probability distribution over outcomes and post-measurement states.
2. Let $\underline{v}_i$ be the i-th column of V. Starting with the initial state, suppose Bob implements the local projective measurement with 4 outcomes defined by $P_i = (I_A \otimes \underline{v}_i \underline{v}_i^H)$ for $i \in \{1, 2, 3, 4\}$. Describe the resulting probability distribution over outcomes and post-measurement states.
3. Suppose Bob implements his measurement after Alice observes outcome i. Describe the resulting probability distribution over outcomes and post-measurement states.

4 Appendix: Abstract Approaches

4.1 Universal construction

Abstractly, the tensor product $U \otimes V$ of two vector spaces is a vector space defined by the following universal property: there exists a bilinear mapping $\phi \colon U \times V \to U \otimes V$ such that, for any vector space $W = \mathbb{R}^p$ and bilinear mapping $\psi \colon U \times V \to W$, there is a linear mapping $\rho \colon U \otimes V \to W$ satisfying

$$
\psi(\underline{u}, \underline{v}) = \rho\left(\phi(\underline{u}, \underline{v})\right).
$$

In words, this means that ϕ is a bilinear mapping of the $(m+n)$-dimensional space $U \times V$ to the mn-dimensional space $U \otimes V$ such that all bilinear forms on $U \times V$ become linear on $U \otimes V$. This is called a *universal property* because ϕ identifies the tensor product of U and V uniquely (up to isomorphism).

As a concrete example, we define

$$
\phi(\underline{u}, \underline{v}) = [u_1 v_1, u_1, v_2, \dots, u_1 v_n, u_2 v_1, u_2 v_2, \dots, u_2 v_n, \dots, u_m v_n]^T,
$$

which equals the Kronecker product $\underline{u} \otimes \underline{v}$ when vectors are column vectors. To see why this works, consider the previous question assuming the standard basis is used for U and V. Since coordinate of W can be treated separately, it suffices to consider the case where $p = 1$ and $W = \mathbb{R}$. For both U and V, we use $\underline{e}_1, \underline{e}_2, \ldots$ to represent the standard basis vectors since the correct space can be determined from the context. Then, for any bilinear mapping $\psi\colon U \times V \to W$, we can let $g(\underline{u}, \underline{v}) = \psi(\underline{u}, \underline{v})$ equal $\psi(\underline{u}, \underline{v})$ and write

$$
\begin{aligned}
g(\underline{u}, \underline{v}) &= \sum_{i=1}^{m} \sum_{j=1}^{n} u_i v_j g\left(\underline{e}_i, \underline{e}_j\right) \\
&= \underline{h}^T \phi(\underline{u}, \underline{v}),
\end{aligned}
$$

where $\underline{h} \in \mathbb{R}^{mn}$ is defined by $h_i = g\big(\underline{e}_{\lfloor (i-1)/n \rfloor + 1}, \underline{e}_{(i-1) \bmod n+1}\big)$ for $i \in [mn]$. Thus, $\rho(\underline{z}) = \underline{h}^T \underline{z}$ defines the linear functional on $U \otimes V$ that maps $\phi(\underline{u}, \underline{v})$ to $\psi(\underline{u}, \underline{v})$. For the extension to $p > 1$, one can define vectors $\underline{h}_k$ for $k \in [p]$ where $\rho_k(\underline{z}) = \underline{h}_k^T \underline{z}$ equals $[g(\underline{u}, \underline{v})]_k$.

5 The infinite dimensional case

For a complex Hilbert space $\mathcal{H}$, we define the inner product $\langle \psi, \phi \rangle_{\mathcal{H}}$ between $\psi, \phi \in \mathcal{H}$ to be linear in ψ. Let $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ denote the set of bounded linear operators mapping $\mathcal{H}_1$ to $\mathcal{H}_2$. For $A \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, the adjoint $A^* \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ satisfies $\langle \psi_1, A^* \psi_2 \rangle_{\mathcal{H}_1} = \langle A\psi_1, \psi_2 \rangle_{\mathcal{H}_2}$ for all $\psi_1 \in \mathcal{H}_1$ and $\psi_2 \in \mathcal{H}_2$.

By identifying $\phi \in \mathcal{H}$ with the linear functional $\langle \psi, \phi \rangle$ mapping $\psi \in \mathcal{H}$ to $\mathbb{C}$, we observe that its adjoint ϕ^* satisfies $\langle \phi^* \psi, 1 \rangle_{\mathbb{C}} = \langle \psi, \phi \rangle_{\mathcal{H}}$ for all $\psi \in \mathcal{H}$. This identification leads to the standard physics notation where $|\psi\rangle, |\phi\rangle$ represent elements of $\mathcal{H}$, the adjoint linear functional is denoted by reversing the decoration $\langle \phi| = \phi^*$, and the inner product is denoted by $\langle \phi|\psi \rangle = \phi^* \psi = \langle \psi, \phi \rangle$. It is important to observe that the order of the vectors in the inner product is reversed between the standard physics notation and the standard Hilbert space notation.

An operator $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ is called *square* and we use the shorthand notation $A \in \mathcal{B}(\mathcal{H})$ to denote this property. A square operator $A \in \mathcal{B}(\mathcal{H})$ is called *self-adjoint* if $A^* = A$ and this holds if and only if $\langle \phi, A\phi \rangle \in \mathbb{R}$ for all $\phi \in \mathcal{H}$. In any orthogonal basis, the matrix representation M of a self-adjoint operator is Hermitian symmetric (i.e., $M^H = M$). A square operator $A \in \mathcal{B}(\mathcal{H})$ is called *positive semidefinite* if $\langle \phi, A\phi \rangle \geq 0$ for all $\phi \in \mathcal{H}$ and this property is denoted by $A \geq 0$. It is called *positive definite* if the inequality is strict whenever $\phi \neq 0$ and this is denoted by $A > 0$. In a complex Hilbert space, a positive-semidefinite operator is assumed to be self-adjoint because the above inequality implicitly assumes that $\langle \phi, A\phi \rangle \in \mathbb{R}$ for all $\phi \in \mathcal{H}$.

For Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ over $\mathbb{C}$, the tensor product is formally defined using the following steps:

1. For all vectors $\psi \in \mathcal{H}_1$ and $\phi \in \mathcal{H}_2$, let $\psi \otimes \phi$ denote the formal tensor product of these vectors.

2. Define the bilinear (or sesquilinear) function $\langle \psi \otimes \phi, \psi' \otimes \phi' \rangle_{\mathcal{H}_1 \otimes \mathcal{H}_2} = \langle \psi, \psi' \rangle_{\mathcal{H}_1} \langle \phi, \phi' \rangle_{\mathcal{H}_2}$ for all $\psi, \psi' \in \mathcal{H}_1$ and $\phi, \phi' \in \mathcal{H}_2$.

3. Consider the set of finite linear combinations, $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} (\psi_i \otimes \phi_j)$, of formal tensor product elements where $a_{ij} \in \mathbb{C}$, $\psi_1, \ldots, \psi_n \in \mathcal{H}_1$, and $\phi_1, \ldots, \phi_n \in \mathcal{H}_2$.

4. Use linearity to extend the restricted definition of $\langle \cdot, \cdot \rangle_{\mathcal{H}_1 \otimes \mathcal{H}_2}$ to vectors that are finite linear combinations of formal tensor products. Due to representational degeneracy (i.e., there are distinct finite linear combinations that equal the same vector), this gives a *semi inner product* where a vector represented as a non-trivial linear combination of non-zero vectors (i.e., is formally non-zero) may have an inner product with itself that equals zero.

5. Let the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ denote the completion of the implied semi inner product space.

Taking the completion has two effects. First, it allows infinite linear combinations of formal tensor products. Second, the equivalence relation used by the completion operation naturally removes the degeneracy in the space and treats distinct representations of a vector as exactly the same vector.

In engineering and physics, one typically deals with separable Hilbert spaces that have countable orthonormal bases. Let $\psi_1, \psi_2, \ldots$ be a countable orthonormal basis for $\mathcal{H}_1$. If the singular values of $A \in \mathcal{B}(\mathcal{H}_1)$ are summable, then the trace

$$\operatorname{Tr}(A) := \sum_{i=1}^{\infty} \langle A\psi_i, \psi_i \rangle$$

is well defined and does not depend on the choice of basis. Using this, the set of density operators is defined by

$$\mathcal{D}(\mathcal{H}_1) = \{A \in \mathcal{B}(\mathcal{H}_1) \,|\, A \geq 0,\ \operatorname{Tr}(A) = 1\}\,.$$

Moreover, if $\phi_1, \phi_2, \ldots$ is a countable orthonormal basis for $\mathcal{H}_2$, then we can alternatively define the tensor product space $\mathcal{H}_1 \otimes \mathcal{H}_2$ as the Hilbert space spanned by the countable orthonormal basis $\{(\psi_i \otimes \phi_j) \,|\, i, j \in \mathbb{N}\}\,.$