Inner products and projections

To better prepare ourselves to explore the capabilities and limitations of quantum circuits, we now introduce some additional mathematical concepts — namely the *inner product* between vectors (and its connection to the Euclidean norm), the notions of *orthogonality* and *orthonormality* for sets of vectors, and *projection* matrices, which will allow us to introduce a handy generalization of standard basis measurements.

Inner products

Recall that when we use the Dirac notation to refer to an arbitrary column vector as a ket, such as

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

the corresponding bra vector is the *conjugate transpose* of this vector:

$$\langle\psi| = \big(|\psi\rangle\big)^\dagger = \begin{pmatrix} \overline{\alpha_1} & \overline{\alpha_2} & \cdots & \overline{\alpha_n} \end{pmatrix}. \tag{1}$$

Alternatively, if we have some classical state set Σ in mind, and we express a column vector as a ket, such as

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle,$$

then the corresponding row (or bra) vector is the conjugate transpose

$$\langle\psi| = \sum_{a\in\Sigma} \overline{\alpha_a} \langle a|. \tag{2}$$

We also have that the product of a bra vector and a ket vector, viewed as matrices either having a single row or a single column, results in a scalar. Specifically, if we have two column vectors

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \quad \text{and} \quad |\phi\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix},$$

so that the row vector $\langle\psi|$ is as in equation (1), then

$$\langle\psi|\phi\rangle = \langle\psi||\phi\rangle = \begin{pmatrix} \overline{\alpha_1} & \overline{\alpha_2} & \cdots & \overline{\alpha_n} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \overline{\alpha_1}\beta_1 + \cdots + \overline{\alpha_n}\beta_n$$

Alternatively, if we have two column vectors that we have written as

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle \quad \text{and} \quad |\phi\rangle = \sum_{b\in\Sigma} \beta_b |b\rangle,$$

so that $\langle\psi|$ is the row vector (2), we find that

$$\begin{aligned} \langle\psi|\phi\rangle &= \langle\psi||\phi\rangle \\ &= \left(\sum_{a\in\Sigma} \overline{\alpha_a}\langle a|\right)\left(\sum_{b\in\Sigma} \beta_b |b\rangle\right) \\ &= \sum_{a\in\Sigma}\sum_{b\in\Sigma} \overline{\alpha_a}\beta_b \langle a|b\rangle \\ &= \sum_{a\in\Sigma} \overline{\alpha_a}\beta_a, \end{aligned}$$

where the last equality follows from the observation that $\langle a|a\rangle = 1$ and $\langle a|b\rangle = 0$ for classical states a and b satisfying $a \neq b$.

The value $\langle\psi|\phi\rangle$ is called the *inner product* between the vectors $|\psi\rangle$ and $|\phi\rangle$. Inner products are critically important in quantum information and computation; we would not get far in understanding quantum information at a mathematical level without them.

Let us now collect together some basic facts about inner products of vectors.

1. **Relationship to the Euclidean norm.** The inner product of any vector

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle$$

with itself is

$$\langle\psi|\psi\rangle = \sum_{a\in\Sigma} \overline{\alpha_a}\alpha_a = \sum_{a\in\Sigma} |\alpha_a|^2 = \big\||\psi\rangle\big\|^2.$$

Thus, the Euclidean norm of a vector may alternatively be expressed as

$$\big\||\psi\rangle\big\| = \sqrt{\langle\psi|\psi\rangle}.$$

Notice that the Euclidean norm of a vector must always be a nonnegative real number. Moreover, the only way the Euclidean norm of a vector can be equal to zero is if every one of the entries is equal to zero, which is to say that the vector is the zero vector.

We can summarize these observations like this: for every vector $|\psi\rangle$ we have

$$\langle\psi|\psi\rangle \geq 0,$$

with $\langle\psi|\psi\rangle = 0$ if and only if $|\psi\rangle = 0$. This property of the inner product is sometimes referred to as *positive definiteness*.

2. **Conjugate symmetry.** For any two vectors

$$|\psi\rangle = \sum_{a\in\Sigma} \alpha_a |a\rangle \quad \text{and} \quad |\phi\rangle = \sum_{b\in\Sigma} \beta_b |b\rangle,$$

we have

$$\langle\psi|\phi\rangle = \sum_{a\in\Sigma} \overline{\alpha_a}\beta_a \quad \text{and} \quad \langle\phi|\psi\rangle = \sum_{a\in\Sigma} \overline{\beta_a}\alpha_a,$$

and therefore

$$\overline{\langle\psi|\phi\rangle} = \langle\phi|\psi\rangle.$$

3. **Linearity in the second argument (and conjugate linearity in the first).** Let us suppose that $|\psi\rangle$, $|\phi_1\rangle$, and $|\phi_2\rangle$ are vectors and α_1 and α_2 are complex numbers. If we define a new vector

$$|\phi\rangle = \alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle,$$

then

$$\langle\psi|\phi\rangle = \langle\psi|\big(\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle\big) = \alpha_1\langle\psi|\phi_1\rangle + \alpha_2\langle\psi|\phi_2\rangle.$$

That is to say, the inner product is *linear* in the second argument. This can be verified either through the formulas above or simply by noting that matrix multiplication is linear in each argument (and specifically in the second argument).

Combining this fact with conjugate symmetry reveals that the inner product is *conjugate linear* in the first argument. That is, if $|\psi_1\rangle$, $|\psi_2\rangle$, and $|\phi\rangle$ are vectors and α_1 and α_2 are complex numbers, and we define

$$|\psi\rangle = \alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle,$$

then

$$\langle\psi|\phi\rangle = \big(\overline{\alpha_1}\langle\psi_1| + \overline{\alpha_2}\langle\psi_2|\big)|\phi\rangle = \overline{\alpha_1}\langle\psi_1|\phi\rangle + \overline{\alpha_2}\langle\psi_2|\phi\rangle.$$

4. **The Cauchy–Schwarz inequality.** For every choice of vectors $|\phi\rangle$ and $|\psi\rangle$ having the same number of entries, we have

$$\big|\langle\psi|\phi\rangle\big| \leq \big\||\psi\rangle\big\|\big\||\phi\rangle\big\|.$$

This is an incredibly handy inequality that gets used quite extensively in quantum information (and in many other topics of study).

Orthogonal and orthonormal sets

Two vectors $|\phi\rangle$ and $|\psi\rangle$ are said to be *orthogonal* if their inner product is zero:

$$\langle\psi|\phi\rangle = 0.$$

Geometrically, we can think about orthogonal vectors as vectors at right angles to each other.

A set of vectors $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is called an *orthogonal set* if every vector in the set is orthogonal to every other vector in the set. That is, this set is orthogonal if

$$\langle\psi_j|\psi_k\rangle = 0$$

for all choices of $j, k \in \{1, \ldots, m\}$ for which $j \neq k$.

A set of vectors $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is called an *orthonormal* set if it is an orthogonal set and, in addition, every vector in the set is a unit vector.

Alternatively, this set is an orthonormal set if we have

$$\langle \psi_j | \psi_k \rangle = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{3}$$

for all choices of $j, k \in \{1, \ldots, m\}$.

Finally, a set $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is an *orthonormal basis* if, in addition to being an orthonormal set, it forms a basis. This is equivalent to $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ being an orthonormal set and m being equal to the dimension of the space from which $|\psi_1\rangle, \ldots, |\psi_m\rangle$ are drawn.

For example, for any classical state set Σ, the set of all standard basis vectors

$$\big\{|a\rangle \,:\, a \in \Sigma\big\}$$

is an orthonormal basis. The set $\{|+\rangle, |-\rangle\}$ is an orthonormal basis for the 2-dimensional space corresponding to a single qubit, and the Bell basis $\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$ is an orthonormal basis for the 4-dimensional space corresponding to two qubits.

Extending orthonormal sets to orthonormal bases

Suppose that $|\psi_1\rangle, \ldots, |\psi_m\rangle$ are vectors that live in an n-dimensional space, and assume moreover that $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is an orthonormal set. Orthonormal sets are always linearly independent sets, so these vectors necessarily span a subspace of dimension m. From this we conclude that $m \leq n$ because the dimension of the subspace spanned by these vectors cannot be larger than the dimension of the entire space from which they're drawn.

If it is the case that $m < n$, then it is always possible to choose an additional $n - m$ vectors $|\psi_{m+1}\rangle, \ldots, |\psi_n\rangle$ so that $\{|\psi_1\rangle, \ldots, |\psi_n\rangle\}$ forms an orthonormal basis. A procedure known as the *Gram–Schmidt orthogonalization process* can be used to construct these vectors.

Orthonormal sets and unitary matrices

Orthonormal sets of vectors are closely connected with unitary matrices. One way to express this connection is to say that the following three statements are logically equivalent (meaning that they are all true or all false) for any choice of a square matrix U:

1. The matrix U is unitary (that is, $U^\dagger U = \mathbb{I} = UU^\dagger$).
2. The rows of U form an orthonormal set.
3. The columns of U form an orthonormal set.

This equivalence is actually pretty straightforward when we think about how matrix multiplication and the conjugate transpose work. Suppose, for instance, that we have a 3×3 matrix like this:

$$U = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} \end{pmatrix}$$

The conjugate transpose of U looks like this:

$$U^\dagger = \begin{pmatrix} \overline{\alpha_{1,1}} & \overline{\alpha_{2,1}} & \overline{\alpha_{3,1}} \\ \overline{\alpha_{1,2}} & \overline{\alpha_{2,2}} & \overline{\alpha_{3,2}} \\ \overline{\alpha_{1,3}} & \overline{\alpha_{2,3}} & \overline{\alpha_{3,3}} \end{pmatrix}$$

Multiplying the two matrices, with the conjugate transpose on the left-hand side, gives us this matrix:

$$\begin{pmatrix} \overline{\alpha_{1,1}} & \overline{\alpha_{2,1}} & \overline{\alpha_{3,1}} \\ \overline{\alpha_{1,2}} & \overline{\alpha_{2,2}} & \overline{\alpha_{3,2}} \\ \overline{\alpha_{1,3}} & \overline{\alpha_{2,3}} & \overline{\alpha_{3,3}} \end{pmatrix} \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} \end{pmatrix}$$
$$= \begin{pmatrix} \overline{\alpha_{1,1}}\alpha_{1,1} + \overline{\alpha_{2,1}}\alpha_{2,1} + \overline{\alpha_{3,1}}\alpha_{3,1} & \overline{\alpha_{1,1}}\alpha_{1,2} + \overline{\alpha_{2,1}}\alpha_{2,2} + \overline{\alpha_{3,1}}\alpha \\ \overline{\alpha_{1,2}}\alpha_{1,1} + \overline{\alpha_{2,2}}\alpha_{2,1} + \overline{\alpha_{3,2}}\alpha_{3,1} & \overline{\alpha_{1,2}}\alpha_{1,2} + \overline{\alpha_{2,2}}\alpha_{2,2} + \overline{\alpha_{3,2}}\alpha \\ \overline{\alpha_{1,3}}\alpha_{1,1} + \overline{\alpha_{2,3}}\alpha_{2,1} + \overline{\alpha_{3,3}}\alpha_{3,1} & \overline{\alpha_{1,3}}\alpha_{1,2} + \overline{\alpha_{2,3}}\alpha_{2,2} + \overline{\alpha_{3,3}}\alpha \end{pmatrix}$$

If we form three vectors from the columns of U,

$$|\psi_1\rangle = \begin{pmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \alpha_{3,1} \end{pmatrix}, \quad |\psi_2\rangle = \begin{pmatrix} \alpha_{1,2} \\ \alpha_{2,2} \\ \alpha_{3,2} \end{pmatrix}, \quad |\psi_3\rangle = \begin{pmatrix} \alpha_{1,3} \\ \alpha_{2,3} \\ \alpha_{3,3} \end{pmatrix},$$

then we can alternatively express the product above as follows:

$$U^\dagger U = \begin{pmatrix} \langle\psi_1|\psi_1\rangle & \langle\psi_1|\psi_2\rangle & \langle\psi_1|\psi_3\rangle \\ \langle\psi_2|\psi_1\rangle & \langle\psi_2|\psi_2\rangle & \langle\psi_2|\psi_3\rangle \\ \langle\psi_3|\psi_1\rangle & \langle\psi_3|\psi_2\rangle & \langle\psi_3|\psi_3\rangle \end{pmatrix}$$

Referring to the equation (3), we now see that the condition that this matrix is equal to the identity matrix is equivalent to the orthonormality of the set $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$.

This argument generalizes to unitary matrices of any size. The fact that the rows of a matrix form an orthonormal basis if and only if the matrix is unitary then follows from the fact that a matrix is unitary if and only if its transpose is unitary.

Given the equivalence described above, together with the fact that every orthonormal set can be extended to form an orthonormal basis, we conclude the following useful fact: Given any orthonormal set of vectors $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ drawn from an n-dimensional space, there exists a unitary matrix U whose first m columns are the vectors $|\psi_1\rangle, \ldots, |\psi_m\rangle$. Pictorially, we can always find a unitary matrix having this form:

$$U = \begin{pmatrix} | & | & & | & | & & | \\ |\psi_1\rangle & |\psi_2\rangle & \cdots & |\psi_m\rangle & |\psi_{m+1}\rangle & \cdots & |\psi_n\rangle \\ | & | & & | & | & & | \end{pmatrix}.$$

Here, the last $n - m$ columns are filled in with any choice of vectors $|\psi_{m+1}\rangle, \ldots, |\psi_n\rangle$ that make $\{|\psi_1\rangle, \ldots, |\psi_n\rangle\}$ an orthonormal basis.

Projections and projective measurements

Projection matrices

A square matrix Π is called a *projection* if it satisfies two properties:

1. $\Pi = \Pi^\dagger$.
2. $\Pi^2 = \Pi$.

Matrices that satisfy the first condition — that they are equal to their own conjugate transpose — are called *Hermitian matrices*, and matrices that satisfy the second condition — that squaring them leaves them unchanged — are called *idempotent* matrices.

As a word of caution, the word *projection* is sometimes used to refer to any matrix that satisfies just the second condition but not necessarily the first, and when this is done the term *orthogonal projection* is typically used to refer to matrices satisfying both properties. In the context of quantum information and computation, however, the terms *projection* and *projection matrix* more typically refer to matrices satisfying both conditions.

An example of a projection is the matrix

$$\Pi = |\psi\rangle\langle\psi| \tag{4}$$

for any unit vector $|\psi\rangle$. We can see that this matrix is Hermitian as follows:

$$\Pi^\dagger = \big(|\psi\rangle\langle\psi|\big)^\dagger = \big(\langle\psi|\big)^\dagger\big(|\psi\rangle\big)^\dagger = |\psi\rangle\langle\psi| = \Pi.$$

Here, to obtain the second equality, we have used the formula

$$(AB)^\dagger = B^\dagger A^\dagger,$$

which is always true, for any two matrices A and B for which the product AB makes sense.

To see that the matrix Π in (4) is idempotent, we can use the assumption that $|\psi\rangle$ is a unit vector, so that it satisfies $\langle\psi|\psi\rangle = 1$. Thus, we have

$$\Pi^2 = \big(|\psi\rangle\langle\psi|\big)^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \Pi.$$

More generally, if $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ is any orthonormal set of vectors, then the matrix

$$\Pi = \sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k| \tag{5}$$

is a projection. Specifically, we have

$$\begin{aligned}
\Pi^\dagger &= \left(\sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k|\right)^\dagger \\
&= \sum_{k=1}^{m} \big(|\psi_k\rangle\langle\psi_k|\big)^\dagger \\
&= \sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k| \\
&= \Pi,
\end{aligned}$$

and

$$\begin{aligned}\Pi^2 &= \biggl(\sum_{j=1}^{m} |\psi_j\rangle\langle\psi_j|\biggr)\biggl(\sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k|\biggr)\\ &= \sum_{j=1}^{m}\sum_{k=1}^{m} |\psi_j\rangle\langle\psi_j|\psi_k\rangle\langle\psi_k|\\ &= \sum_{k=1}^{m} |\psi_k\rangle\langle\psi_k|\\ &= \Pi,\end{aligned}$$

where the orthonormality of $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$ implies the second-to-last equality.

In fact, this exhausts all of the possibilities: *every* projection Π can be written in the form (5) for some choice of an orthonormal set $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$. (Technically speaking, the zero matrix $\Pi = 0$, which is a projection, is a special case. To fit it into the general form (5) we must allow the possibility that the sum is empty, resulting in the zero matrix.)

Projective measurements

The notion of a measurement of a quantum system is more general than just standard basis measurements. *Projective measurements* are measurements that are described by a collection of projections whose sum is equal to the identity matrix. In symbols, a collection $\{\Pi_0, \ldots, \Pi_{m-1}\}$ of projection matrices describes a projective measurement if

$$\Pi_0 + \cdots + \Pi_{m-1} = \mathbb{I}.$$

When such a measurement is performed on a system X while it is in some state $|\psi\rangle$, two things happen:

1. For each $k \in \{0, \ldots, m-1\}$, the outcome of the measurement is k with probability equal to

$$\Pr\bigl(\text{outcome is } k\bigr) = \bigl\|\Pi_k|\psi\rangle\bigr\|^2.$$

2. For whichever outcome k the measurement produces, the state of X becomes

$$\frac{\Pi_k|\psi\rangle}{\bigl\|\Pi_k|\psi\rangle\bigr\|}.$$

We can also choose outcomes other than $\{0, \ldots, m-1\}$ for projective measurements if we wish. More generally, for any finite and nonempty set Σ, if we have a collection of projection matrices

$$\{\Pi_a : a \in \Sigma\}$$

that satisfies the condition

$$\sum_{a \in \Sigma} \Pi_a = \mathbb{I},$$

then this collection describes a projective measurement whose possible outcomes coincide with the set Σ, where the rules are the same as before:

1. For each $a \in \Sigma$, the outcome of the measurement is a with probability equal to

$$\Pr\big(\text{outcome is } a\big) = \big\|\Pi_a|\psi\rangle\big\|^2.$$

2. For whichever outcome a the measurement produces, the state of X becomes

$$\frac{\Pi_a|\psi\rangle}{\big\|\Pi_a|\psi\rangle\big\|}.$$

For example, standard basis measurements are equivalent to projective measurements, where Σ is the set of classical states of whatever system X we're talking about and our set of projection matrices is $\{|a\rangle\langle a| : a \in \Sigma\}$.

Another example of a projective measurement, this time on two qubits (X, Y), is given by the set $\{\Pi_0, \Pi_1\}$, where

$$\Pi_0 = |\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| \quad \text{and} \quad \Pi_1 = |\psi^-\rangle\langle\psi^-|.$$

If we have multiple systems that are jointly in some quantum state and a projective measurement is performed on just one of the systems, the action is similar to what we had for standard basis measurements — and in fact we can now describe this action in much simpler terms than we could before.

To be precise, let us suppose that we have two systems (X, Y) in a quantum state $|\psi\rangle$, and a projective measurement described by a collection $\{\Pi_a : a \in \Sigma\}$ is performed on the system X, while nothing is done to Y. Doing this is then equivalent to performing the projective measurement described by the collection

$$\big\{\Pi_a \otimes \mathbb{I} \,:\, a \in \Sigma\big\}$$

on the joint system (X, Y). Each measurement outcome a results with probability

$$\big\|(\Pi_a \otimes \mathbb{I})|\psi\rangle\big\|^2,$$

and conditioned on the result a appearing, the state of the joint system (X, Y) becomes

$$\frac{(\Pi_a \otimes \mathbb{I})|\psi\rangle}{\big\|(\Pi_a \otimes \mathbb{I})|\psi\rangle\big\|}.$$

Implementing projective measurements

Arbitrary projective measurements can be implemented using unitary operations, standard basis measurements, and an extra workspace system, as will now be explained.

Let us suppose that X is a system and $\{\Pi_0, \ldots, \Pi_{m-1}\}$ is a projective measurement on X. We can easily generalize this discussion to projective measurements having different sets of outcomes, but in the interest of convenience and simplicity we will assume the set of possible outcomes for our measurement is $\{0, \ldots, m-1\}$.

Let us note explicitly that m is not necessarily equal to the number of classical states of X — we'll let n be the number of classical states of X, which means that each matrix Π_k is an $n \times n$ projection matrix.

Because we assume that $\{\Pi_0 \ldots, \Pi_{m-1}\}$ represents a projective measurement, it is necessarily the case that

$$\sum_{k=0}^{m-1} \Pi_k = \mathbb{I}_n.$$

Our goal is to perform a process that has the same effect as performing this projective measurement on X, but to do this using only unitary operations and standard basis measurements.

We will make use of an extra workspace system Y to do this, and specifically we'll take the classical state set of Y to be $\{0, \ldots, m-1\}$, which is the same as the set of outcomes of the projective measurement. The idea is that we will perform a standard basis measurement on Y, and interpret the outcome of this measurement as being equivalent to the outcome of the projective measurement on X. We'll need to assume that

Y is initialized to some fixed state, which we'll choose to be $|0\rangle$. (Any other choice of fixed quantum state vector could be made to work, but choosing $|0\rangle$ makes the explanation to follow much simpler.)

Of course, in order for a standard basis measurement of Y to tell us anything about X, we will need to allow X and Y to interact somehow before measuring Y, by performing a unitary operation on the system (Y, X). First consider this matrix:

$$M = \sum_{k=0}^{m-1} |k\rangle\langle 0| \otimes \Pi_k.$$

Expressed explicitly as a so-called *block matrix*, which is essentially a matrix of matrices that we interpret as a single, larger matrix, M looks like this:

$$M = \begin{pmatrix} \Pi_0 & 0 & \cdots & 0 \\ \Pi_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{m-1} & 0 & \cdots & 0 \end{pmatrix}.$$

Here, each 0 represents an $n \times n$ matrix filled entirely with zeros, so that the entire matrix M is an $nm \times nm$ matrix.

Now, M is certainly not a unitary matrix (unless $m = 1$, in which case $\Pi_0 = \mathbb{I}$, giving $M = \mathbb{I}$ in this trivial case) because unitary matrices cannot have any columns (or rows) that are entirely 0; unitary matrices have columns that form orthonormal bases, and the all-zero vector is not a unit vector.

However, it is the case that the first n columns of M are orthonormal, and we get this from the assumption that $\{\Pi_0, \ldots, \Pi_{m-1}\}$ is a measurement. To verify this claim, notice that for each $j \in \{0, \ldots, n-1\}$, the vector formed by column number j of M is as follows:

$$|\psi_j\rangle = M|0, j\rangle = \sum_{k=0}^{m-1} |k\rangle \otimes \Pi_k |j\rangle.$$

Note that here we're numbering the columns starting from column 0. Taking the inner product of column i with column j when $i, j \in \{0, \ldots, n-1\}$ gives

$$\begin{aligned}\langle \psi_i | \psi_j \rangle &= \left(\sum_{k=0}^{m-1} |k\rangle \otimes \Pi_k |i\rangle \right)^\dagger \left(\sum_{l=0}^{m-1} |l\rangle \otimes \Pi_l |j\rangle \right) \\ &= \sum_{k=0}^{m-1} \sum_{l=0}^{m-1} \langle k|l\rangle \langle i|\Pi_k \Pi_l |j\rangle \\ &= \sum_{k=0}^{m-1} \langle i|\Pi_k \Pi_k |j\rangle \\ &= \sum_{k=0}^{m-1} \langle i|\Pi_k |j\rangle \\ &= \langle i|\mathbb{I}|j\rangle \\ &= \begin{cases} 1 & i = j \\ 0 & i \neq j, \end{cases}\end{aligned}$$

which is what we needed to show.

Thus, because the first n columns of the matrix M are orthonormal, we can replace all of the remaining zero entries by some different choice of complex number entries so that the entire matrix is unitary.

$$U = \begin{pmatrix} \Pi_0 & \boxed{?} & \cdots & \boxed{?} \\ \Pi_1 & \boxed{?} & \cdots & \boxed{?} \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{m-1} & \boxed{?} & \cdots & \boxed{?} \end{pmatrix}$$

If we're given the matrices $\Pi_0, \ldots, \Pi_{m-1}$, we can compute suitable matrices to fill in for the blocks marked $\boxed{?}$ in the equation — using the Gram–Schmidt process — but it does not matter specifically what these matrices are for the sake of this discussion.

Finally we can describe the measurement process: we first perform U on the joint system (Y, X) and then measure Y with respect to a standard basis measurement. For an arbitrary state $|\phi\rangle$ of X, we obtain the state

$$U\big(|0\rangle|\phi\rangle\big) = M\big(|0\rangle|\phi\rangle\big) = \sum_{k=0}^{m-1} |k\rangle \otimes \Pi_k |\phi\rangle,$$

where the first equality follows from the fact that U and M agree on their first n columns. When we perform a projective measurement on Y, we obtain each outcome k with probability

$$\big\| \Pi_k |\phi\rangle \big\|^2,$$

in which case the state of (Y, X) becomes

$$|k\rangle \otimes \frac{\Pi_k |\phi\rangle}{\left\| \Pi_k |\phi\rangle \right\|}.$$

Thus, Y stores a copy of the measurement outcome and X changes precisely as it would had the projective measurement described by $\{\Pi_0, \ldots, \Pi_{m-1}\}$ been performed directly on X.

Was this page helpful?

Yes No

Report a bug or request content on GitHub ↗.

Previous page Next page