Quantum Entropy and Accounting

John Fellingham
Ohio State University

Haijin Lin
University of Houston

Doug Schroeder
Ohio State University

2018

1 Introduction

The double entry system of accounting has endured and even thrived for over five centuries. There have been dramatic and profound changes in commerce and technology, yet the system of double entry remains recognizable.

Double entry accounting tracks stocks and flows through the march of time and does so in an elegant fashion. It is this tracking of stocks and flows we wish to emphasize. Further the scholarly intent is to connect the double entry accounting numbers to the concept of information in a fundamental way.

In recent decades many scientific disciplines have treated information as a core concept. Physics, for example, asks "How much information is in a black hole?" or "How much information is in the universe?". Biology, as well, asks information questions, such as "How much information is contained in generic material?"

In order to ask questions about the amount of information, an information metric is required. That was supplied by Claude Shannon in 1948 (Shannon 1948). Shannon's "entropy" can be treated as a measure of uncertainty and is constructed as a function of probabilities, where $p(y)$ is the probability of an observation of a random variable Y,

$$H(Y) = -\sum p(y) \ln p(y),$$

where ln is natural logarithm. Shannon's entropy has a very useful additive property,

$$H(X,Y) = H(Y|X) + H(X),$$

where $H(X,Y)$ is the joint entropy and $H(Y|X)$ is the conditional entropy constructed using joint and conditional probabilities, respectively. Since entropies can be added, they can also be subtracted yielding a useful definition of information as whatever it is that decreases entropy,

$$I(X;Y) = H(Y) - H(Y|X).$$

$I(X;Y)$ is termed mutual information and is the reduction of entropy in random variable Y if when a signal X is available.

Fellingham and Lin (2018) connect accounting numbers to information,

$$\ln\left(1 + \frac{income}{assets}\right) = r_f + I(X;Y), \tag{1}$$

where r_f is the risk free rate of returns. The accounting numbers, thus, provide a measure of "how much information" the reporting entity possesses, similar to other information sciences posing the "how much information?" question. Further, the relationship is an equality implying any accounting question, involving the assignment of numbers to income and assets, can be reframed as an information question. And, as the equality goes both ways, any information question can be reframed as an accounting question.

The basic accounting is done on an economic income basis; that is, assets are valued at discounted cash flow. Fellingham and Lin (2018) provide conditions under which alternative accounting methods yield the same results.

The environmental assumptions for the equality (1) to hold are three:

- long run decision perspective;
- arbitrage free prices;
- complete markets; that is, there exists an Arrow-Debreu security for every possible state realization.

It is the last assumption which is a bit problematic, as accountants are used to operating in an environment in which not all states can be traded. A purpose of this paper is to expand the domain of the relationship (1) to incomplete market settings. The way this is done is to access the mathematics of quantum processes and quantum information.

Since a risk free rate of return is a complete market concept, the equality (1) can be reformulated as

$$\ln\left(1+\frac{income}{assets}\right)=E\left[r|X_p\right]-H\left(Y|X\right), \tag{2}$$

where $E\left[r|X_p\right]$ is the expected return attainable if the information is perfect denoted by X_p (the reformulation is in Section 3). The main result of this paper is to derive an equality that applies to incomplete markets, written as

$$\ln\left(1+\frac{income}{assets}\right)=E\left[r|X_p\right]-S\left(Y|X\right). \tag{3}$$

The notable distinction between (2) and (3) is quantum entropy denoted by $S\left(Y|X\right)$ replaces the Shannon entropy expression $H\left(Y|X\right)$. Quantum entropy was developed by John von Neumann (Nielsen and Chuang 2004). When markets are complete, Shannon entropy and Von Neumann entropy compute to the same number, and the quantum relationship (3) reduces to the classical relationship (2).

Besides expanding the domain of the accounting information equivalence, there are other advantages to bringing quantum information into view. One is that quantum technological advances are occurring with increasing regularity. Examples are quantum computation, quantum clocks, and incredibly powerful quantum imaging techniques. It is reassuring that the power of the double entry accounting is well positioned to survive another significant technological change.

Another advantage is that quantum thinking brings into focus a powerful information resource known as non-locality (or as discussed in Section 4, entangled qubits). In the world of physics, non-locality is a troubling development. Einstein, for example, was disturbed by "spooky action at a distance." In an economic setting, however, non-locality is a natural and desirable phenomenon—individuals separated by distance can share information as well as act cooperatively. The results are not nearly as mysterious as distant quantum objects exhibiting non-local correlations.

The quantum mathematics illustrates the corrosive effects of local measurements of non-local objects. Local measurements strictly increases entropy which in turn by the equality (3) decreases expected return.

The remainder of the paper is organized as follows. Section 2 provides some preliminaries of the mathematics of quantum information and measurement including quantum entropy. Section 3 formulates the quantum decision problem and demonstrates the main result of the paper. Section 4 illustrates an application of the main result—using quantum concept of non-locality in an economic setting. The concluding remarks are in Section 5.

2 Quantum Preliminaries

2.1 Superposition

Quantum processes are generally considered "mysterious" related to the world we observe, and superposition is one mysterious quantum property. Excerpted from an article "Quantum Leaps; Subatomic Opportunities" (The Economists, March 11, 2017), *"Quantum mechanics...has a well-earned reputation for weirdness. That is because the world as humanity sees it is not, in fact, how the world works."*

A quantum unit can exist in a state of superposition when it is neither one attribute nor another, but both at once. The quantum attributes are things like positive or negative charge and both attributes can exist simultaneously.

A two element vector is required to represent a quantum unit (often referred to as "qubit"). Two commonly used states for a qubit, written in Dirac notation, are

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4}$$

A qubit, linearly combining the two states, can be written in superposition as

$$|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix} = a\,|0\rangle + b\,|1\rangle\,, \tag{5}$$

where the coefficients a and b are called amplitudes; these amplitudes can be positive or negative and scaled so that $a^2 + b^2 = 1$.[1] Examples of qubits are

$$\frac{3}{5}\,|0\rangle + \frac{4}{5}\,|1\rangle = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}; \text{ or}$$

$$\frac{1}{\sqrt{2}}\,|0\rangle - \frac{1}{\sqrt{2}}\,|1\rangle = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

Dirac notation for a transpose of qubit $|\psi\rangle$ is written as

$$\langle\psi| = \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}^T. \tag{6}$$

It is important to do linear algebra operation. The inner product of two qubits, $|\psi_1\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$ and $|\psi_2\rangle = \begin{bmatrix} c \\ d \end{bmatrix}$, is written as

$$\langle\psi_1|\;|\psi_2\rangle = \langle\psi_1|\,\psi_2\rangle = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd; \tag{7}$$

and the outer product is written as

$$|\psi_1\rangle\,\langle\psi_2| = \begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} ac & ad \\ bc & bd \end{bmatrix}. \tag{8}$$

[1] The coefficients a and b can be complex numbers. However, for our purpose in this paper, we restrict our attention to real coefficients.

2.2 Measurement

The state of a qubit can not be observed directly; all that can be accomplished directly is a quantum measurement. Quantum measurement is the use of two measurement vectors. A straightforward illustration is called the standard basis and consists of $|0\rangle$ and $|1\rangle$ (as defined by (4)).

When a qubit is measured, the post-measurement state is one of the measurement vectors probabilistically. The probabilities are determined by the square of the vector product of the measurement vector with the qubit. For example, measurement of an qubit $\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle$ by the standard basis results in a post-measurement qubit state of $|0\rangle$ with probability

$$\left[\langle 0| \left(\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle\right)\right]^2 = \left(\frac{3}{5}\right)^2 = \frac{9}{25},$$

and the probability the post-measurement qubit state is $|1\rangle$ is

$$\left[\langle 1| \left(\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle\right)\right]^2 = \frac{16}{25}.$$

Whenever the measurement vectors are orthogonal (termed projective measurement), the post-measurement probabilities will sum to one.[2] Any two orthogonal vectors can be used for projective measurements (this is true even for a quantum system with more than two qubit states).

Example 1 *Consider the two orthogonal vectors* $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ *and* $\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$ *and a qubit* $\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle$*. The post-measurement probability that the state is* $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ *is written as*

$$\begin{aligned} & \left[\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right)^T \left(\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle\right)\right]^2 \\ = & \left[\left(\frac{1}{\sqrt{2}}\langle 0| + \frac{1}{\sqrt{2}}\langle 1|\right) \left(\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle\right)\right]^2 \\ = & \left[\frac{3}{5\sqrt{2}}\langle 0|\,0\rangle + \frac{4}{5\sqrt{2}}\langle 1|\,1\rangle\right]^2 = \frac{49}{50}. \end{aligned}$$

[2] Two vectors are orthogonal if the inner product of the two vectors is zero. It is readily to check that the states $|0\rangle$ and $|1\rangle$ are orthogonal.

$$\langle 0|1\rangle = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.$$

Similarly, the probability that the post-measurement state is $\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$ *is computed to be* $\frac{1}{50}$.[3]

$$\begin{aligned} & \left[\left(\frac{1}{\sqrt{2}}\langle 0| - \frac{1}{\sqrt{2}}\langle 1|\right)\left(\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle\right)\right]^2 \\ = & \left[\frac{3}{5\sqrt{2}}\langle 0|\,0\rangle - \frac{4}{5\sqrt{2}}\langle 1|\,1\rangle\right]^2 = \frac{1}{50} \end{aligned}$$

Density Operator

An useful qubit representation is an outer product, called density operator, denoted by ρ. The density operator of a qubit $|\psi\rangle$ is written as

$$\rho = |\psi\rangle\langle\psi|\,. \tag{9}$$

Density operator is an alternative but equivalent tool for quantum mechanics including quantum measurement. Denote the measurement vectors by $|m_1\rangle$ and $|m_2\rangle$. Then the post-measurement probability of the state $|m_1\rangle$ is written as

$$\begin{aligned} [\langle m_1|\,\psi\rangle]^2 &= \langle m_1|\psi\rangle\langle m_1|\psi\rangle = \langle\psi|m_1\rangle\langle m_1|\psi\rangle = \langle\psi|M_1|\psi\rangle \\ &= tr\left(M_1\,|\psi\rangle\langle\psi|\right) = tr\left(M_1\rho\right), \end{aligned} \tag{10}$$

where $M_1 = |m_1\rangle\langle m_1|$ is the outer product of the measurement vector $|m_1\rangle$. The tr represents the trace of the matrix computed as the sum of the diagonals. With density operator, the post-measurement probabilities are determined by the trace of the product matrix between the density operator and the outer product of the measurement vector.

Example 2 *Continue Example 1. The density operator of qubit* $\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle$ *is computed as*

$$\begin{aligned} \rho &= \left(\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle\right)\left(\frac{3}{5}\langle 0| + \frac{4}{5}\langle 1|\right) \\ &= \frac{9}{25}|0\rangle\langle 0| + \frac{12}{25}|0\rangle\langle 1| + \frac{12}{25}|1\rangle\langle 0| + \frac{16}{25}|1\rangle\langle 1| \\ &= \frac{9}{25}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \frac{12}{25}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} + \frac{12}{25}\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix} + \frac{16}{25}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} \\ &= \begin{bmatrix}\frac{9}{25} & \frac{12}{25}\\ \frac{12}{25} & \frac{16}{25}\end{bmatrix}. \end{aligned}$$

[3] The measurement computation can be thought of as the R^2 (correlation) when the measured qubit is projected into one of the measurement qubits.

The outer product of the measurement vectors $|m_1\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ *is computed as*

$$\begin{aligned} M_1 &= |m_1\rangle\langle m_1| = \left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right)\left(\frac{1}{\sqrt{2}}\langle 0| + \frac{1}{\sqrt{2}}\langle 1|\right) \\ &= \left(\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|0\rangle\langle 1| + \frac{1}{2}|1\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|\right) \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

The post-measurement probability that the state is $|m_1\rangle$ *is again* $\frac{49}{50}$ *calculated based on (10).*

$$\begin{aligned} tr\left(M_1\rho\right) &= tr\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix}\right) \\ &= tr\left(\frac{1}{50}\begin{bmatrix} 21 & 28 \\ 21 & 28 \end{bmatrix}\right) = \frac{49}{50} \end{aligned}$$

2.3 von Neumann Entropy

The density operator formulation can be quite illuminating conceptually. In particular, quantum entropy, often called von Neumann entropy, is available with the density operator formulation. Consider a system, described by an ensemble $\{p_i,\ |\psi_i\rangle\}$, can be in one of the qubit states $|\psi_i\rangle$ and the probability of state $|\psi_i\rangle$ is p_i. The density operator is generalized to a mixture of qubits by

$$\rho = \Sigma_i\ p_i\rho_i, \tag{11}$$

where $\rho_i = |\psi_i\rangle\langle\psi_i|$ is density operator for qubit $|\psi_i\rangle$ (as defined in (9)). Let λ_j be the eigenvalues of the density operator matrix ρ in (11). The von Neumann entropy is then defined as

$$S\left(\rho\right) = -\Sigma_j\lambda_j \ln\left(\lambda_j\right). \tag{12}$$

von Neumann entropy incorporates the uncertainty inherent in quantum objects, just in the same spirit as (classical) Shannon entropy. Classical entropy as a measure of uncertainty was developed by Claude Shannon. The Shannon entropy of a random variable X, denoted by $H(X)$, is defined as a function of the probabilities associated with the possible realizations of the random variable.

$$H(X) = -\Sigma_x p(x) \ln [p(x)] \tag{13}$$

For example, if a random variable has two possible realizations, each with probability one-half, the Shannon entropy is computed as

$$H(X) = -\frac{1}{2} \ln \left(\frac{1}{2}\right) - \frac{1}{2} \ln \left(\frac{1}{2}\right) = \ln 2 \simeq 0.6931.$$

In parallel, von Neumann entropy is based on eigenvalues which will be shown (in the following example) to be the probabilities of projective measurements with the eigenvectors. Example 3 is illustrative.

Example 3 *A system contains a mixture of two qubits,* $|\psi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $|\psi_2\rangle = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$*, both with probability of one-half. The density operator is symmetric and positive, computed as*

$$\begin{aligned} \rho &= \frac{1}{2}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}\right) + \frac{1}{2}\left(\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \end{bmatrix}\right) \\ &= \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix} = \frac{1}{25}\begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix}. \end{aligned}$$

The eigenvalues of ρ *are computed as* $\lambda_1 = \frac{4}{5}$ *and* $\lambda_2 = \frac{1}{5}$.[4] *In general, the sum of the eigenvalues of the density operator is always one as long as the length of the qubits is normalized to one* ($\langle\psi|\,\psi\rangle = 1$)*. The corresponding eigenvectors are orthogonal and computed as* $|v_1\rangle = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$ *and* $|v_2\rangle = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}$.[5]

---

[4] The eigenvalues are calculated to ensure the characterisitc equation $\det(\rho - \lambda I) = 0$ holds. More specifically,

$$\begin{gathered} \det\left(\frac{1}{25}\begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0; \\ \det\left(\begin{bmatrix} \frac{17}{25} - \lambda & \frac{6}{25} \\ \frac{6}{25} & \frac{8}{25} - \lambda \end{bmatrix}\right) = 0; \\ \left(\frac{17}{25} - \lambda\right)\left(\frac{8}{25} - \lambda\right) - \frac{36}{625} = 0. \end{gathered}$$

The solutions are $\lambda_1 = \frac{4}{5}$ and $\lambda_2 = \frac{1}{5}$.

[5] The eigenvectors v are determined by $(\rho - \lambda I)\, v = 0$. For example, $|v_1\rangle$ is the eigenvector

If ρ is measured based on the two eigenvectors $|v_1\rangle$ and $|v_2\rangle$, the probability that the post-measurement state is $|v_i\rangle$ is λ_i, that is, $tr\left(|v_i\rangle\langle v_i|\,\rho\right) = \lambda_i$. For state $|v_1\rangle$,

$$\begin{aligned} tr\left(|v_1\rangle\langle v_1|\,\rho\right) &= tr\left(\begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \frac{1}{25} \begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix}\right) \\ &= tr\left(\begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} \frac{1}{25} \begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix}\right) \\ &= tr\left(\begin{bmatrix} \frac{16}{25} & \frac{8}{25} \\ \frac{8}{25} & \frac{4}{25} \end{bmatrix}\right) = \frac{4}{5}. \end{aligned}$$

The density operator ρ has a spectral decomposition as the sum of the outer products of the eigenvectors multiplied by the corresponding eigenvalues,

$$\begin{aligned} \rho &= \frac{4}{5}\left(\begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}\right) + \frac{1}{5}\left(\begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}\right) \\ &= \frac{4}{5}\begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} + \frac{1}{5}\begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{bmatrix} = \frac{1}{25}\begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix}. \end{aligned}$$

This is remarkable as the spectral decomposition essentially reframes the initial ensemble $\{\frac{1}{2}, \frac{1}{2}, |\psi_1\rangle, |\psi_2\rangle\}$ into a new ensemble $\{\frac{4}{5}, \frac{1}{5}, |v_1\rangle, |v_2\rangle\}$. In the new ensemble, the eigenvalues are the probabilities of eigenvectors. The von Neumann entropy is essentially the Shannon entropy, computed as

$$S(\rho) = -\frac{4}{5}\ln\left(\frac{4}{5}\right) - \frac{1}{5}\ln\left(\frac{1}{5}\right) = \ln(5) - \frac{4}{5}\ln(4) \simeq 0.5004.$$

It is noted that the von Neumann entropy in Example 3 is less than the Shannon entropy computed based on a variable with two equally likely realizations, $S(\rho) \simeq 0.5004 < H(\cdot) \simeq 0.6931$. Lemma 1 formally compares the two entropies.

Lemma 1 *Shannon entropy always weakly exceeds von Neumann entropy.*

---

associated with $\lambda_1 = \frac{4}{5}$.

$$\left(\frac{1}{25}\begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix} - \begin{bmatrix} \frac{4}{5} & 0 \\ 0 & \frac{4}{5} \end{bmatrix}\right)\begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = 0$$

Solving the linear system subject to the length of the vector $|v_1\rangle$ being one yields $|v_1\rangle = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$.

The eigenvector $|v_2\rangle$ associated with $\lambda_2 = \frac{1}{5}$ can be solved following the same steps.

Proof. Consider an ensemble $\{p_i,\ |\psi_i\rangle\}$, its density operator can be written as

$$\rho = \Sigma_i\ p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| . \tag{14}$$

Let λ_i be the eigenvalues of ρ and $|v_i\rangle$ be the respective eigenvectors. The density operator has a spectral decomposition

$$\rho = \Sigma_i\ \lambda_i \left|v_i\right\rangle \left\langle v_i\right| . \tag{15}$$

The von Neumann entropy is written as

$$S(\rho) = -\Sigma_i\ \lambda_i \ln(\lambda_i) . \tag{16}$$

Treating the state of the system as a random variable, the Shannon entropy is written as

$$H(|\psi_i\rangle) = -\Sigma_i\ p_i \ln(p_i) . \tag{17}$$

We prove that (16) is no more than (17) in two cases.

Suppose the states are orthogonal so that $\langle\psi_i|\,\psi_j\rangle = 0$ for any $i \neq j$. We show that (16) and (17) are equal. It is sufficient to show that $p_j = \lambda_j$. That is, for $\left|\psi_j\right\rangle$ nonzero, $\rho\left|\psi_j\right\rangle = p_j\left|\psi_j\right\rangle$. To see this,

$$\begin{aligned} \rho\left|\psi_j\right\rangle &= \left(\Sigma_i\ p_i \left|\psi_i\right\rangle\left\langle\psi_i\right|\right)\left|\psi_j\right\rangle \\ &= \Sigma_i\ p_i \left|\psi_i\right\rangle\left\langle\psi_i\right|\left.\psi_j\right\rangle \\ &= p_j\left|\psi_j\right\rangle\left\langle\psi_j\right|\left.\psi_j\right\rangle \\ &= p_j\left|\psi_j\right\rangle . \end{aligned} \tag{18}$$

The third equality in (18) is ensured by the assumption that the states are orthogonal; while the fourth equality is ensured by $\left\langle\psi_j|\psi_j\right\rangle = 1$. Since $|\psi_i\rangle$ are orthogonal, then the density operator ρ has a spectral decomposition as

$$\rho = \Sigma_i\ \lambda_i \left|\psi_i\right\rangle \left\langle\psi_i\right| . \tag{19}$$

Suppose the states are not orthogonal so that $\langle\psi_i|\,\psi_j\rangle \neq 0$ for any $i \neq j$. Denote the index K such that $p_K = \max_i\ p_i$. Consider a set of orthonormal measurement vectors $|m_j\rangle$ where $|\psi_K\rangle = |m_K\rangle$. Then $\{M_j = |m_j\rangle\langle m_j|\}$ is a set of orthogonal projectors, each of which satisfies $M_j^T = M_j$ and $\Sigma_j M_j = I$. The initial ensemble has qubit states so two orthonormal measurement vectors are sufficient, say $|m_K\rangle$ and $|m_N\rangle$, and $M_K + M_N = I$. The post-measurement probability that the state is $|m_j\rangle$ is

$$p'_j \quad = \quad tr\left(M_j\rho\right) = tr\left(M_j\Sigma_i\ p_i \left|\psi_i\right\rangle\left\langle\psi_i\right|\right)$$

$$= \quad \Sigma_i \; p_i \; tr\left(M_j \left|\psi_i\right\rangle \left\langle\psi_i\right|\right). \tag{20}$$

The new density operator after measurement takes the following form (see Nielsen and Chuang 2004, p101),

$$\rho' = M_K \rho M_K + M_N \rho M_N. \tag{21}$$

We proceed to show the following inequalities hold,

$$H\left(\left|\psi_i\right\rangle\right) \geq -p'_K \ln\left(p'_K\right) - \left(p'_N\right) \ln\left(p'_N\right) > S\left(\rho\right). \tag{22}$$

(i) To show the left side of (22), it is noted that

$$H\left(\left|\psi_i\right\rangle\right) = -\Sigma_i \; p_i \ln\left(p_i\right) \geq -p_K \ln\left(p_K\right) - \left(1 - p_K\right) \ln\left(1 - \; p_K\right), \tag{23}$$

where $p_K = \max_i \; p_i$. From (20),

$$\begin{aligned} p'_K &= \Sigma_i \; p_i \; Tr\left(M_K \left|\psi_i\right\rangle \left\langle\psi_i\right|\right) \\ &= p_K \; Tr\left(M_K \left|\psi_K\right\rangle \left\langle\psi_K\right|\right) + \Sigma_{i \neq K} \; p_i \; Tr\left(M_K \left|\psi_i\right\rangle \left\langle\psi_i\right|\right) \\ &= p_K + \Sigma_{i \neq K} \; p_i \; Tr\left(M_K \left|\psi_i\right\rangle \left\langle\psi_i\right|\right) > p_K; \end{aligned} \tag{24}$$

the last equality of (24) is implied by $\left|\psi_K\right\rangle = \left|m_K\right\rangle$ and $Tr\left(M_K \left|\psi_K\right\rangle \left\langle\psi_K\right|\right) = 1$; while the inequality is ensured as $Tr\left(M_K \left|\psi_i\right\rangle \left\langle\psi_i\right|\right) > 0$ is the post-measurement probability of the state $\left|m_K\right\rangle$. From the inequality in (24), it must be

$$p'_N < 1 - p_K. \tag{25}$$

Combining (23), (24), and (25) yields

$$H\left(\left|\psi_i\right\rangle\right) \geq -p_K \ln\left(p_K\right) - \left(1 - p_K\right) \ln\left(1 - p_K\right) > -p'_K \ln\left(p'_K\right) - p'_N \ln\left(p'_N\right). \tag{26}$$

(ii) To show the right side of (22), the expression (21) can be further written as

$$\begin{aligned} \rho' &= \Sigma_j \; M_j \left(\Sigma_i \; p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|\right) M_j = \Sigma_i \Sigma_j \; p_i \left(M_j \left|\psi_i\right\rangle \left\langle\psi_i\right| M_j\right) \\ &= \Sigma_i \Sigma_j \; p_i \left(\left|m_j\right\rangle \left\langle m_j \right| \psi_i\rangle \left\langle\psi_i \right| m_j\rangle \left\langle m_j\right|\right) \\ &= \Sigma_i \Sigma_j \; p_i \left(\left|m_j\right\rangle \left\langle\psi_i \right| m_j\rangle \left\langle m_j \right| \psi_i\rangle \left\langle m_j\right|\right) \\ &= \Sigma_i \Sigma_j \; p_i \left(\left|m_j\right\rangle \left\langle\psi_i\right| M_j \left|\psi_i\right\rangle \left\langle m_j\right|\right) \\ &= \Sigma_j \; \left[\Sigma_i \; p_i \left\langle\psi_i\right| M_j \left|\psi_i\right\rangle\right] \left|m_j\right\rangle \left\langle m_j\right| \\ &= p'_K \left|m_K\right\rangle \left\langle m_K\right| + p'_N \left|m_N\right\rangle \left\langle m_N\right|, \end{aligned} \tag{27}$$

where p'_j is defined in (20). The right side inequality of (22) is ensured by Theorem 11.9 in Nielsen and Chuang (2004)–which states that the Shannon entropy based on the post-measurement probabilities is the smallest as long as the measurement vectors are the eigenvectors. (This result is formally proved in Proposition 2.) Noted that ρ defined in (15) has the post-measurement eigenvector state $|v_i\rangle$ with probability λ_i; while ρ' defined in (27) has the post-measurement state $|m_j\rangle$ with probability p'_j. ■

Illustration

Consider a system with two qubits, $|\psi_1\rangle = \frac{1}{\sqrt{x^2+1}}\begin{bmatrix}-1\\ x\end{bmatrix}$ and $|\psi_2\rangle = \frac{1}{\sqrt{y^2+1}}\begin{bmatrix}y\\ 1\end{bmatrix}$ and their respective probabilities are p_1 and p_2. The density operator is

$$\begin{aligned}
\rho &= p_1\left(\frac{1}{x^2+1}\begin{bmatrix}-1\\ x\end{bmatrix}\begin{bmatrix}-1 & x\end{bmatrix}\right)+p_2\left(\frac{1}{y^2+1}\begin{bmatrix}y\\ 1\end{bmatrix}\begin{bmatrix}y & 1\end{bmatrix}\right)\\
&= \frac{p_1}{x^2+1}\begin{bmatrix}1 & -x\\ -x & x^2\end{bmatrix}+\frac{p_2}{y^2+1}\begin{bmatrix}y^2 & y\\ y & 1\end{bmatrix}\\
&= \begin{bmatrix}\frac{p_1}{x^2+1}+\frac{p_2y^2}{y^2+1} & \frac{-p_1x}{x^2+1}+\frac{p_2y}{y^2+1}\\ \frac{-p_1x}{x^2+1}+\frac{p_2y}{y^2+1} & \frac{p_1x^2}{x^2+1}+\frac{p_2}{y^2+1}\end{bmatrix}.
\end{aligned}$$

The eigenvalues of ρ can be calculated as follows,

$$\det\left(\begin{bmatrix}\frac{p_1}{x^2+1}+\frac{p_2y^2}{y^2+1}-\lambda & \frac{-p_1x}{x^2+1}+\frac{p_2y}{y^2+1}\\ \frac{-p_1x}{x^2+1}+\frac{p_2y}{y^2+1} & \frac{p_1x^2}{x^2+1}+\frac{p_2}{y^2+1}-\lambda\end{bmatrix}\right)=0,$$

$$\left[\frac{p_1}{x^2+1}+\frac{p_2y^2}{y^2+1}-\lambda\right]\left[\frac{p_1x^2}{x^2+1}+\frac{p_2}{y^2+1}-\lambda\right]-\left[\frac{-p_1x}{x^2+1}+\frac{p_2y}{y^2+1}\right]^2=0,$$

$$\lambda^2-(p_1+p_2)\lambda+\frac{p_1p_2(xy+1)^2}{(x^2+1)(y^2+1)}=0$$

which has the solutions

$$\lambda_1=\frac{p_1+p_2-\sqrt{\Delta}}{2} \text{ and } \lambda_2=\frac{p_1+p_2+\sqrt{\Delta}}{2}$$

where $\Delta=(p_1+p_2)^2-\frac{4p_1p_2(xy+1)^2}{(x^2+1)(y^2+1)}$. Note that $\lambda_1+\lambda_2=1$ as $p_1+p_2=1$. Without loss of generality, we assume $p_2>p_1$. It is readily shown that $\lambda_1\le p_1$ and $\lambda_2\ge p_2$.

$$\frac{p_1+p_2-\sqrt{\Delta}}{2}\le p_1 \Leftrightarrow \sqrt{\Delta}\ge p_2-p_1$$

$$
\begin{aligned}
&\Leftrightarrow \quad (p_1+p_2)^2 - \frac{4p_1p_2\,(xy+1)^2}{(x^2+1)\,(y^2+1)} \geq (p_2-p_1)^2 \\
&\Leftrightarrow \quad \left(x^2+1\right)\left(y^2+1\right) \geq (xy+1)^2 \\
&\Leftrightarrow \quad (x-y)^2 \geq 0
\end{aligned}
$$

The equality holds as long as $x = y$ which implies the two qubits are orthogonal. Note that $(x-y)^2 \geq 0$ also implies $\lambda_2 \geq p_2$. Then comparing the two entropies,

$$H(\psi) = -\Sigma_i p_i \ln(p_i) \geq -\Sigma_i \lambda_i \ln(\lambda_i) = S(\rho),$$

the equality holds as long as the two qubits are orthogonal.

Example 4 illustrates the proof of Lemma 1.

Example 4 *Consider a system with three qubits,* $|\psi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $|\psi_2\rangle = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$, *and* $|\psi_3\rangle = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. *Their respective probabilities are* 0.1, 0.5, *and* 0.4. *The density operator is computed as*

$$
\begin{aligned}
\rho &= 0.1\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) + 0.5\left(\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix}\right) + 0.4\left(\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\right) \\
&= \begin{bmatrix} \frac{12}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{13}{25} \end{bmatrix}.
\end{aligned}
$$

This is the expression (14) in the proof of Lemma 1. The Shannon entropy for this ensemble is

$$H(|\psi\rangle) = -0.1\ln(0.1) - 0.5\ln(0.5) - 0.4\ln(0.4) \simeq 0.9433.$$

The eigenvalues of ρ *can be calculated as follows,*

$$\det\left(\begin{bmatrix} \frac{12}{25} - \lambda & \frac{1}{25} \\ \frac{1}{25} & \frac{13}{25} - \lambda \end{bmatrix}\right) = 0,$$

$$\left[\frac{12}{25} - \lambda\right]\left[\frac{13}{25} - \lambda\right] - \frac{1}{625} = 0,$$

which has the solutions

$$\lambda_1 = \frac{25+\sqrt{5}}{50} \text{ and } \lambda_2 = \frac{25-\sqrt{5}}{50};$$

and the respective eigenvectors are $|v_1\rangle = \begin{bmatrix} \frac{\sqrt{10(5-\sqrt{5})}}{10} \\ \frac{\sqrt{10(5+\sqrt{5})}}{10} \end{bmatrix}$ *and* $|v_2\rangle = \begin{bmatrix} \frac{\sqrt{10(5+\sqrt{5})}}{10} \\ -\frac{\sqrt{10(5-\sqrt{5})}}{10} \end{bmatrix}$.
The spectral decomposition of ρ *is then written as*

$$\begin{aligned} \rho &= \left(\frac{25+\sqrt{5}}{50}\right)(|v_1\rangle\langle v_1|) + \left(\frac{25-\sqrt{5}}{50}\right)(|v_2\rangle\langle v_2|) \\ &= \left(\frac{25+\sqrt{5}}{50}\right)\begin{bmatrix} \frac{5-\sqrt{5}}{10} & \frac{\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} & \frac{5+\sqrt{5}}{10} \end{bmatrix} + \left(\frac{25-\sqrt{5}}{50}\right)\begin{bmatrix} \frac{5+\sqrt{5}}{10} & -\frac{\sqrt{5}}{5} \\ -\frac{\sqrt{5}}{5} & \frac{5-\sqrt{5}}{10} \end{bmatrix} \\ &= \begin{bmatrix} \frac{12}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{13}{25} \end{bmatrix}. \end{aligned}$$

This is the expression (15). The von Neumann entropy is then computed as

$$\begin{aligned} S(\rho) &= -\left(\frac{25+\sqrt{5}}{50}\right)\ln\left(\frac{25+\sqrt{5}}{50}\right) - \left(\frac{25-\sqrt{5}}{50}\right)\ln\left(\frac{25-\sqrt{5}}{50}\right) \\ &\simeq 0.6891. \end{aligned}$$

Consider two orthonormal measurement vectors, $|m_1\rangle = \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$ *and* $|m_2\rangle = |\psi_2\rangle = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$*. The post-measurement probability that the state is* $|m_1\rangle$ *is*

$$\begin{aligned} p_1' &= tr(|m_1\rangle\langle m_1|\rho) = tr\left(\begin{bmatrix} \frac{16}{25} & -\frac{12}{25} \\ -\frac{12}{25} & \frac{9}{25} \end{bmatrix}\begin{bmatrix} \frac{12}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{13}{25} \end{bmatrix}\right) \\ &= tr\left(\begin{bmatrix} \frac{36}{125} & -\frac{28}{125} \\ -\frac{27}{125} & \frac{21}{125} \end{bmatrix}\right) = \frac{57}{125}; \end{aligned}$$

and the post-measurement probability for the state $|m_2\rangle$ *is*

$$\begin{aligned} p_2' &= tr(|m_2\rangle\langle m_2|\rho) = tr\left(\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix}\begin{bmatrix} \frac{12}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{13}{25} \end{bmatrix}\right) \\ &= tr\left(\begin{bmatrix} \frac{24}{125} & \frac{33}{125} \\ \frac{32}{125} & \frac{44}{125} \end{bmatrix}\right) = \frac{68}{125}. \end{aligned}$$

These probabilities are (25) in the proof of Lemma 1. The inequalities (22) are confirmed as follows:

$$\begin{aligned} H(|\psi\rangle) &> -0.5\ln(0.5) - (0.4+0.1)\ln(0.4+0.1) \simeq 0.6932 \\ &> -\frac{57}{125}\ln\left(\frac{57}{125}\right) - \frac{68}{125}\ln\left(\frac{68}{125}\right) \simeq 0.6893 > S(\rho). \end{aligned}$$

3 Main Analysis

In this section, we analyze the relationship between accounting number and information number, the latter is described by the von Neumann entropy. This entails an economy setting that can be depicted by a quantum process. As we show, the relationship between accounting number and the von Neumann entropy takes the same form as the relationship between accounting number and the Shannon entropy in a classical Arrow-Debreu economy.

Fellingham and Lin (2018) state that in a general state-act-outcome decision problem with Arrow-Debreu securities and long run preferences, the accounting rate of return computed based on an economic income basis—which is also the expected long run return—is equal to a base amount minus the information number measured by the Shannon entropy. The base amount is the expected return with perfect information.

Lemma 2 *In an Arrow-Debreu economy with long run preferences, the following relationship holds:*

$$\ln\left(1+\frac{Income}{Assets}\right)=E\left[r|X_p\right]-H\left(Y|X\right), \tag{28}$$

where X_p denotes the perfect information; Y denotes the payoffs of interests, and X denotes the information signals about Y.

Proof. Under economic income accounting, the accounting rate of return is the continuously compounded rate of return based on information X, that is,

$$\frac{Income}{Assets}=e^{r(X)}-1\Rightarrow r\left(X\right)=\ln\left(1+\frac{Income}{Assets}\right). \tag{29}$$

Taking a long run perspective so that many returns have been gathered for an entity, the accounting rate of return converges to the expected rate of return given information X. This is the application of the law of large number. That is,

$$r\left(X\right)=E\left[r|X\right]. \tag{30}$$

The central result in Fellingham and Lin (2018) is reproduced as follows,

$$\ln\left(1+\frac{Income}{Assets}\right)=r_f+I\left(X;Y\right) \tag{31}$$

where r_f denotes the risk free rate; and $I\left(X;Y\right)=H\left(Y\right)-H\left(Y|X\right)$ is the mutual information which measures the reduction of uncertainty in the presence of information X. Combining (29) and (30), the relationship (31) is rewritten as

$$\ln\left(1+\frac{Income}{Assets}\right)=E\left[r|X\right]=r_f+H\left(Y\right)-H\left(Y|X\right). \tag{32}$$

With perfect information X_p, the decision maker learns perfectly about the underlying state. There is no uncertainty so that $H\left(Y|X_p\right)=0$, from (32),

$$\begin{aligned} E\left[r|X_p\right] &= r_f+H\left(Y\right)-H\left(Y|X_p\right)=r_f+H\left(Y\right) \\ \Rightarrow r_f &= E\left[r|X_p\right]-H\left(Y\right). \end{aligned} \tag{33}$$

Substituting r_f expression (33) in (32) yields

$$\begin{aligned} E\left[r|X\right] &= E\left[r|X_p\right]-H\left(Y\right)+H\left(Y\right)-H\left(Y|X\right) \\ &= E\left[r|X_p\right]-H\left(Y|X\right). \end{aligned} \tag{34}$$

In words, the expected return with information X is the expected return with perfect information minus the Shannon entropy given X. ■

In this paper, we show that the relationship, described in (28), stays the same when probabilities and transformations are governed (specified) by quantum process. In notation,

$$\ln\left(1+\frac{Income}{Assets}\right)=E\left[r|X_p\right]-S\left(\rho|X\right). \tag{35}$$

In words, the expected return of an entity (defined by accounting rate of return) is the expected return with perfect information minus the von Neumann entropy. Both perfect information and entropy are defined distinctively in the quantum setting. In particular, only the measured states can be perfectly observed in the quantum setting (35) while the underlying states can be perfectly observed in the classical setting (28). We next present an economic setting in which the relationship (35) is derived.

3.1 The quantum decision problem

We formulate a state-act-payoff decision problem where quantum mechanics apply. There are two distinct features that are absent in the classical setting depicted by Fellingham and Lin (2018). One, the state-act-payoff matrix does not have to be full rank in which case an Arrow-Debreu security does not exist in some states. Two, the underlying states can not be directly observed nor known.

Suppose a decision maker is confronted with uncertainty over the set of the states described by an ensemble of qubits. The underlying state of the qubit can not be observed but can be measured with some prescribed measurement basis. The decision maker's resource allocation decision is then based on the measured state. In this setting, the decision maker has two decisions to make: (i) choosing an optimal measurement basis; and (ii) choosing an optimal resource allocation conditional on the measurement basis. To proceed, we first solve the optimal resource allocation decision for a given measurement basis, and then characterize the optimal measurement basis.

In particular, the economy is described by an ensemble $\{p_j, \ |\psi_j\rangle\}$ where the state is denoted by $|\psi_j\rangle$ and its respective probability is p_j. The density operator of the ensemble ρ has two eigenvectors $|v_1\rangle$ and $|v_2\rangle$ and the respective eigenvalues are λ_1 and λ_2. The ensemble is common knowledge but the (realized state) cannot be observed. Only the measured state is observable. The decision maker chooses some measurement basis $\{|m_i\rangle\}$ to measure the ensemble and observes the state $|m_i\rangle$ with probability g_i. Let y be the payoff for every dollar invested in the (measured) state. The sequence of events is as follows.

- At $t = 0$, the decision maker chooses a measurement basis.
- At $t = 1$, the decision maker chooses the fraction of the initial wealth invested in each measured state.
- At $t = 2$, the measured state is observed and the payoff is realized.

As the decision maker repeats the sequence of events for many rounds, the realized return converges to the expected rate of return. This is the application of the law of large number.

We model the decision maker as a Kelly decision maker who repeatedly invests a fraction b_i of the initial wealth in state $|m_i\rangle$. The initial wealth is normalized to one. Using continuous compounding, the rate of return in state $|m_i\rangle$ is written as

$$b_i y = e^{r_i} \Rightarrow r_i = \ln (b_i y). \tag{36}$$

Maximizing long run wealth is equivalent to maximizing expected log return (net of any measurement costs). The Kelly decision maker's maximization problem is written as follows.

$$\max_b E[r|X] - C \quad = \quad \sum_i g_i \ln (b_i y) - C \tag{37}$$

$$\text{Subject to } \sum_i b_i \quad = \quad 1$$

where C is the cost of measurement scaled by the initial wealth; and X is the information available at the time of measurement.

Proposition 1 *In an economy described by an ensemble* $\{p_i,\ |\psi_i\rangle\}$*, given a measurement basis* $\{|m_i\rangle\}$*, the post-measurement probability that the state of the economy is state* $|m_i\rangle$ *is* g_i*. The Kelly decision maker's optimal decision is to invest* $b_i = g_i$ *portion of the initial wealth in state* $|m_i\rangle$*.*

Proof. The Lagrangian for the decision maker's program (37) is defined as

$$L(b_i;\mu) = \sum_i g_i \ln(b_i y) - C - \mu \left[\sum_i b_i - 1\right] \tag{38}$$

where μ is the Lagrange multiplier. Then

$$\frac{\partial L(b_i;\mu)}{\partial b_i} = 0 \ \Rightarrow\ b_i = \frac{g_i}{\mu}. \tag{39}$$

Since $\sum_i b_i = 1$, $\sum_i \frac{g_i}{\mu} = 1 \ \Rightarrow\ \mu = 1$. Then it must be $b_i = g_i$. This is the Kelly "bet your beliefs" criterion (Kelly 1956). ■

Proposition 1 describes the Kelly decision maker's optimal resource allocation decision for a given measurement basis. In the next section, we solve for the optimal measurement basis.

3.2 Measurement basis optimality

The decision maker chooses the optimal measurement basis to maximize his expected return (as the measurement cost C is exogenous). Incorporating the optimal resource allocation decision in Proposition 1, the expected return for the decision maker can be further written as

$$\begin{aligned} E[r|X] &= \sum_i g_i \ln(g_i y) \qquad (40)\\ &= \sum_i g_i \ln(g_i) + \sum_i g_i \ln(y) \\ &= \sum_i g_i \ln(g_i) + \ln(y). \end{aligned}$$

The second term $\ln(y)$ is the return with perfect information. To see why, if the decision maker knows which state $|m_i\rangle$ is observed with certainty, he will optimally invest all of the initial wealth to the state $|m_i\rangle$ and not invest in other states ($b_i = 1$ and $b_j = 0$ for $i \neq j$). In this case, the expected return is written as

$$E[r|X_p] = \ln(y). \tag{41}$$

Here perfect information refers to the measured states while in the classical setting perfect information refers to the underlying states. Nevertheless, both the measured states in the quantum setting and the underlying states in the classical setting are payoff-relevant. In quantum setting the underlying states are not observable and therefore irrelevant to payoffs. The expected return (40) can be further written as

$$E[r|X] = E[r|X_p] + \sum_i g_i \ln(g_i). \tag{42}$$

Example 5 explains how the choice of measurement basis affects the expected return.

Example 5 *Continue Example 3. Suppose the economy is described by an ensemble*

$$\left\{ \frac{1}{2}, \frac{1}{2}, |\psi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |\psi_2\rangle = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} \right\}.$$

The density operator of the economy is

$$\rho = \frac{1}{2} |\psi_1\rangle \langle\psi_1| + \frac{1}{2} |\psi_2\rangle \langle\psi_2| = \frac{1}{25} \begin{bmatrix} 17 & 6 \\ 6 & 8 \end{bmatrix}.$$

The eigenvectors are computed in Example 3 as

$$|v_1\rangle = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \text{ and } |v_2\rangle = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix};$$

and the post-measurement probabilities are $\lambda_1 = \frac{4}{5}$ and $\lambda_2 = \frac{1}{5}$ respectively. The payoff is $y = 2.5$ and the expected return with perfect information is $\ln(y) = 0.9163$.

If the economy is measured by the eigenvectors, the optimal bids are $b_1 = 0.8$ and $b_2 = 0.2$. Then the expected return is computed as

$$E[r] = 0.8 \times \ln(0.8 \times 2.5) + 0.2 \times \ln(0.2 \times 2.5) = 0.4159.$$

The entropy is $\frac{4}{5}\ln\left(\frac{4}{5}\right)+\frac{1}{5}\ln\left(\frac{1}{5}\right)\simeq -0.5004$. *Then we derive*

$$E\left[r\right]=-0.5004+\ln\left(y\right)=-0.5004+0.9163=0.4159.$$

If the economy is measured by the standard basis, $|m_1\rangle=\begin{bmatrix}1\\0\end{bmatrix}$ *and* $|m_2\rangle=\begin{bmatrix}0\\1\end{bmatrix}$. *The post-measurement probability that the state is* $|m_1\rangle=\begin{bmatrix}1\\0\end{bmatrix}$ *is*

$$\begin{aligned} g_1 &= tr\left(\begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}1 & 0\end{bmatrix}\frac{1}{25}\begin{bmatrix}17 & 6\\6 & 8\end{bmatrix}\right) \\ &= tr\left(\frac{1}{25}\begin{bmatrix}17 & 6\\0 & 0\end{bmatrix}\right)=\frac{17}{25}; \end{aligned}$$

while the post-measurement probability for the state $|m_2\rangle=\begin{bmatrix}0\\1\end{bmatrix}$ *is*

$$\begin{aligned} g_2 &= tr\left(\begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix}\frac{1}{25}\begin{bmatrix}17 & 6\\6 & 8\end{bmatrix}\right) \\ &= tr\left(\frac{1}{25}\begin{bmatrix}0 & 0\\6 & 8\end{bmatrix}\right)=\frac{8}{25}. \end{aligned}$$

The optimal bids are $b_1=\frac{17}{25}=0.68$ *and* $b_2=\frac{8}{25}=0.32$. *Then the expected return is computed as*

$$E\left[r\right]=\frac{17}{25}\times\ln\left(0.68\times 2.5\right)+\frac{8}{25}\times\ln\left(\frac{8}{25}\times 2.5\right)=0.2894.$$

The entropy is $\frac{17}{25}\ln\left(\frac{17}{25}\right)+\frac{8}{25}\ln\left(\frac{8}{25}\right)\simeq -0.6269$. *Then we derive*

$$E\left[r\right]=-0.6269+\ln\left(y\right)=-0.6269+0.9163=0.2894.$$

The decision maker chooses the optimal measurement basis to maximize the expected return. Since the expected return with perfect information $E\left[r|X_p\right]$ is not affected by the measurement basis, maximizing the expected return is equivalent to maximizing the following expression,

$$\max_{\{|m_i\rangle\}}\sum_i g_i\ln\left(g_i\right). \tag{43}$$

In Example 5, the measurement basis $\{|v_i\rangle\}$ yields higher expected return than the standard basis $\{|m_i\rangle\}$. As we prove later, the measurement basis $\{|v_i\rangle\}$ is the optimal basis.

To understand the intuition, recall the measurement basis $\{|v_i\rangle\}$ is the eigenbasis for the ensemble and the probabilities λ_i are the eigenvalues. In particular, eigenvectors $|v_i\rangle$ solve $\rho\,|v_i\rangle = \lambda_i\,|v_i\rangle$. This allows spectral decomposition of the density operator ρ as

$$\begin{aligned}\rho &= \lambda_1\,|v_1\rangle\,\langle v_1| + \lambda_2\,|v_2\rangle\,\langle v_2| \\ &= \frac{4}{5}\begin{bmatrix}\frac{4}{5} & \frac{2}{5}\\ \frac{2}{5} & \frac{1}{5}\end{bmatrix} + \frac{1}{5}\begin{bmatrix}\frac{1}{5} & \frac{2}{5}\\ -\frac{2}{5} & \frac{4}{5}\end{bmatrix} = \frac{1}{25}\begin{bmatrix}17 & 6\\ 6 & 8\end{bmatrix}.\end{aligned}$$

In this form it is seen that ρ is also the density operator for another ensemble,

$$\left\{\frac{4}{5},\ \frac{1}{5},\ |v_1\rangle = \begin{bmatrix}\frac{2}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}\end{bmatrix},\ |v_2\rangle = \begin{bmatrix}\frac{1}{\sqrt{5}}\\ -\frac{2}{\sqrt{5}}\end{bmatrix}\right\}.$$

Since the density operator for the modified ensemble is the same as that of the initial ensemble, the optimal measurement basis should be the same for both ensembles. To understand why the eigenbasis $\{|v_i\rangle\}$ is the optimal measurement basis, consider the measurement vectors $|m_i\rangle$ and the respective probabilities g_i. A density operator can be formulated as

$$\begin{aligned}\rho' &= g_1\,|m_1\rangle\,\langle m_1| + g_2\,|m_2\rangle\,\langle m_2| \\ &= \frac{17}{25}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \frac{8}{25}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} = \frac{1}{25}\begin{bmatrix}17 & 0\\ 0 & 8\end{bmatrix}.\end{aligned}$$

The new density operator ρ' is a projective measurement of the initial density operator ρ (as shown in the proof of Proposition 2), that is,

$$\rho' = \sum_i M_i \rho M_i.$$

The term $\sum_i g_i \ln(g_i) = -S(\rho')$ is the negative entropy of ρ' while $\sum_i \lambda_i \ln(\lambda_i) = -S(\rho)$ is the negative entropy of ρ. However, as Theorem 11.9 in Nielsen and Chuang (2014) states, projective measurement will never decrease entropy so that $S(\rho) \le S(\rho') \Rightarrow \sum_i \lambda_i \ln(\lambda_i) \ge \sum_i g_i \ln(g_i)$. Proposition 2 formally states the result.

Proposition 2 *In an economy described by an ensemble* $\{p_i,\ |\psi_i\rangle\}$ *and its density operator* ρ*, the optimal measurement basis that maximizes long run expected return is the set of eigenvectors for the density operator,* $\{|v_i\rangle\}$*.*

Proof. First, we show that for any measurement vectors $|m_i\rangle$, the respective probabilities g_i and the density operator $\rho' = \sum_i g_i |m_i\rangle \langle m_i|$, the following equality holds.

$$\sum_i g_i \ln(g_i) = tr(\rho' \ln(\rho')) \tag{44}$$

Two facts about ρ' are useful in proving (44). One, $|m_i\rangle$ and g_i are the respective eigenvectors and eigenvalues for ρ' so that $\langle m_i| \rho' = g_i \langle m_i|$. Two, the density operator ρ' is diagonalizable in the form of $\rho' = D^{-1}AD$. Matrix D has the eigenvectors $|m_i\rangle$ as the columns $D = \begin{bmatrix} |m_1\rangle & |m_2\rangle \end{bmatrix}$ and $D^{-1} = \begin{bmatrix} \langle m_1| \\ \langle m_2| \end{bmatrix}$.[6] Matrix A is a diagonal matrix whose diagonal elements are probabilities g_i. Then $\ln(A)$ is a diagonal matrix whose diagonal elements are $\ln(g_i)$. The natural log of a diagonalizable matrix is defined as

$$\begin{aligned} \ln(\rho') &= D^{-1} \ln(A) D \qquad (45)\\ &= \begin{bmatrix} \langle m_1| \\ \langle m_2| \end{bmatrix} \begin{bmatrix} \ln(g_1) & 0 \\ 0 & \ln(g_2) \end{bmatrix} \begin{bmatrix} |m_1\rangle & |m_2\rangle \end{bmatrix} \\ &= \ln(g_1) |m_1\rangle \langle m_1| + \ln(g_2) |m_2\rangle \langle m_2| = \sum_j \ln(g_j) |m_j\rangle \langle m_j| . \end{aligned}$$

The first fact explains the second equality in (46); and (45) ensures the third equality in (46),

$$\begin{aligned} tr(\rho' \ln(\rho')) &= \sum_i \langle m_i| \rho' \ln(\rho') |m_i\rangle = \sum_i g_i \langle m_i| \ln(\rho') |m_i\rangle \qquad (46)\\ &= \sum_i g_i \langle m_i| \left(\sum_j \ln(g_j) |m_j\rangle \langle m_j| \right) |m_i\rangle \\ &= \sum_i g_i \left(\sum_j \ln(g_j) \langle m_i| m_j\rangle \langle m_j| m_i\rangle \right) = \sum_i g_i \ln(g_i) . \end{aligned}$$

[6] Matrix D is orthonormal as

$$D^T D = \begin{bmatrix} \langle m_1| \\ \langle m_2| \end{bmatrix} \begin{bmatrix} |m_1\rangle & |m_2\rangle \end{bmatrix} = \begin{bmatrix} \langle m_1| m_1\rangle & \langle m_1| m_2\rangle \\ \langle m_2| m_1\rangle & \langle m_2| m_2\rangle \end{bmatrix} = I.$$

For an orthonormal matrix, its inverse is equal to its transpose.

The last equality of (46) is ensured as $\{|m_i\rangle\}$ is orthonormal basis.

Next, to prove

$$\sum_i \lambda_i \ln(\lambda_i) \geq \sum_i g_i \ln(g_i), \tag{47}$$

(44) suggests that it is sufficient to show the following inequality holds,

$$tr(\rho \ln(\rho)) \geq tr(\rho' \ln(\rho')) \tag{48}$$

as the density operator of the initial ensemble is written as $\rho = \sum_i \lambda_i |v_i\rangle \langle v_i|$. We take two steps. In the first step, we show $tr(\rho' \ln(\rho')) = tr(\rho \ln(\rho'))$; and in the second step, $tr(\rho \ln(\rho)) \geq tr(\rho \ln(\rho'))$.

Step 1. For $g_i = \langle m_i| \rho |m_i\rangle$ and $M_i = |m_i\rangle \langle m_i|$, ρ' can be written as

$$\begin{aligned} \rho' &= \sum_i g_i |m_i\rangle \langle m_i| = \sum_i \langle m_i| \rho |m_i\rangle |m_i\rangle \langle m_i| \\ &= \sum_i |m_i\rangle \langle m_i| \rho |m_i\rangle \langle m_i| = \sum_i M_i \rho M_i. \end{aligned} \tag{49}$$

Note that (49) suggests the density operator ρ' is a projective measurement of ρ. Now we prove the result in Step 1:

$$\begin{aligned} tr(\rho' \ln(\rho')) &= tr\left(\sum_i M_i \rho M_i \ln(\rho')\right) = tr\left(\sum_i M_i \rho \ln(\rho') M_i\right) \\ &= tr\left(\sum_i M_i^2 \rho \ln(\rho')\right) = tr(\rho \ln(\rho')). \end{aligned} \tag{50}$$

The first equality in (50) is ensured by (49), the second equality is ensured by the observation that $\ln(\rho')$ commutes with M_i as shown in (51). The third equality in (50) is to apply the cyclic property of trace and the fourth equality is ensured by $M_i^2 = M_i$ and $\sum_i M_i = I$. In addition, the definition (45) and the commutative property of M_i yield

$$\begin{aligned} M_i \ln(\rho') &= M_i \left(\sum_j \ln(g_j) M_j\right) = \sum_j \ln(g_j) M_i M_j \\ &= \sum_j \ln(g_j) M_j M_i = \ln(\rho') M_i. \end{aligned} \tag{51}$$

Step 2. Based on (44), we derive

$$tr(\rho \ln(\rho)) - tr(\rho \ln(\rho')) = \sum_i \lambda_i \ln(\lambda_i) - \sum_i \langle v_i| \rho \ln(\rho') |v_i\rangle. \tag{52}$$

Substituting $\langle v_i| \rho = \lambda_i \langle v_i|$ and the definition (45) in (52) yields

$$\begin{aligned} tr(\rho \ln(\rho)) - tr(\rho \ln(\rho')) &= \sum_i \lambda_i \ln(\lambda_i) - \sum_i \lambda_i \langle v_i| \ln(\rho') |v_i\rangle \\ &= \sum_i \lambda_i \ln(\lambda_i) - \sum_i \lambda_i \langle v_i| \left[\sum_j \ln(g_j) |m_j\rangle \langle m_j| \right] |v_i\rangle \\ &= \sum_i \lambda_i \left[\ln(\lambda_i) - \sum_j \ln(g_j) \langle v_i |m_j\rangle \langle m_j| v_i\rangle \right]. \end{aligned} \tag{53}$$

Because $\sum_j \langle v_i |m_j\rangle \langle m_j| v_i\rangle = 1$ and $\ln(\cdot)$ is a concave function, applying Jensen's inequality, (53) yields the following inequality,

$$\begin{aligned} & tr(\rho \ln(\rho)) - tr(\rho \ln(\rho')) \\ \geq \; & \sum_i \lambda_i \left[\ln(\lambda_i) - \ln\left(\sum_j g_j \langle v_i |m_j\rangle \langle m_j| v_i\rangle \right) \right] \\ = \; & \sum_i \lambda_i \left[\ln\left(\frac{\lambda_i}{\sum_j g_j \langle v_i |m_j\rangle \langle m_j| v_i\rangle} \right) \right] = -\sum_i \lambda_i \left[\ln\left(\frac{\sum_j g_j \langle v_i |m_j\rangle \langle m_j| v_i\rangle}{\lambda_i} \right) \right]. \end{aligned} \tag{54}$$

Applying Jensen's inequality can further reduce the right-hand side expression in (54) and yields the inequality in (55),

$$\begin{aligned} & tr(\rho \ln(\rho)) - tr(\rho \ln(\rho')) \\ \geq \; & -\ln\left(\sum_i \lambda_i \left[\frac{\sum_j g_j \langle v_i |m_j\rangle \langle m_j| v_i\rangle}{\lambda_i} \right] \right) \\ = \; & -\ln\left(\sum_i \sum_j g_j \langle v_i |m_j\rangle \langle m_j| v_i\rangle \right) = -\ln\left(\sum_i \langle v_i| \left[\sum_j g_j |m_j\rangle \langle m_j| \right] |v_i\rangle \right) \\ = \; & -\ln\left(\sum_i \langle v_i| \rho' |v_i\rangle \right) = 0. \end{aligned} \tag{55}$$

The last equality in (55) is ensured as $\langle v_i| \rho' |v_i\rangle$ is the post-measurement probability of the state being $|v_i\rangle$ and all the probabilities sum up to one.

In a nutshell, we have shown $tr(\rho \ln(\rho)) \geq tr(\rho \ln(\rho')) = tr(\rho' \ln(\rho')) \Rightarrow \sum_i \lambda_i \ln(\lambda_i) \geq \sum_i g_i \ln(g_i)$ so that the measurement basis $\{|v_i\rangle\}$ yields the greatest expected return. The equality holds as long as $g_i = \lambda_i$. ■

The following example illustrate Proposition 2 for an ensemble with three qubits.

Example 6 *Consider a system with three qubits,* $|\psi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $|\psi_2\rangle = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$, *and* $|\psi_3\rangle = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{bmatrix}$. *Their respective probabilities are* 0.3, 0.5, *and* 0.2. *The density operator is computed as*

$$\begin{aligned} \rho &= 0.3 \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) + 0.5 \left(\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix} \right) + 0.2 \left(\begin{bmatrix} \frac{1}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{9}{10} \end{bmatrix} \right) \\ &= \begin{bmatrix} \frac{1}{2} & \frac{3}{10} \\ \frac{3}{10} & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

The eigenvalues of ρ *can be calculated as follows,*

$$\det \left(\begin{bmatrix} \frac{1}{2} - \lambda & \frac{3}{10} \\ \frac{3}{10} & \frac{1}{2} - \lambda \end{bmatrix} \right) = 0,$$

$$\left[\frac{1}{2} - \lambda \right] \left[\frac{1}{2} - \lambda \right] - \frac{9}{100} = 0,$$

which has the solutions

$$\lambda_1 = 0.8 \text{ and } \lambda_2 = 0.2;$$

and the respective eigenvectors are $|v_1\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ *and* $|v_2\rangle = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. *The payoff is* $y = 2.5$ *and the expected return with perfect information is* $\ln(y) = 0.9163$. *When measured by the eigenvectors, the optimal bids are* $b_1 = 0.8$ *and* $b_2 = 0.2$. *Then the expected return is computed as*

$$E[r] = 0.8 \times \ln(0.8 \times 2.5) + 0.2 \times \ln(0.2 \times 2.5) = 0.4159.$$

The entropy is $0.8 \ln(0.8) + 0.2 \ln(0.2) \simeq -0.5004$. *Then we derive*

$$E[r] = -0.5004 + \ln(y) = -0.5004 + 0.9163 = 0.4159.$$

If the economy is measured by the standard basis, $|m_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $|m_2\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. *The post-measurement probability that the state is* $|m_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *is*

$$g_1 = tr \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{3}{10} \\ \frac{3}{10} & \frac{1}{2} \end{bmatrix} \right)$$

$$= \quad tr\left(\begin{bmatrix} \frac{1}{2} & \frac{3}{10} \\ 0 & 0 \end{bmatrix}\right) = \frac{1}{2};$$

while the post-measurement probability for the state $|m_2\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *is*

$$\begin{aligned} g_2 &= tr\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{3}{10} \\ \frac{3}{10} & \frac{1}{2} \end{bmatrix}\right) \\ &= tr\left(\begin{bmatrix} 0 & 0 \\ \frac{3}{10} & \frac{1}{2} \end{bmatrix}\right) = \frac{1}{2}. \end{aligned}$$

The optimal bids are $b_1 = 0.5$ *and* $b_2 = 0.5$*. Then the expected return is computed as*

$$E[r] = \frac{1}{2} \times \ln(0.5 \times 2.5) + \frac{1}{2} \times \ln(0.5 \times 2.5) = 0.2231.$$

The entropy is $\frac{1}{2}\ln\left(\frac{1}{2}\right) + \frac{1}{2}\ln\left(\frac{1}{2}\right) \simeq -0.6932$*. Then we derive*

$$E[r] = -0.6932 + \ln(y) = -0.6932 + 0.9163 = 0.2231.$$

Combining Propositions 1 and 2, the relationship between accounting information and quantum entropy is immediate.

Proposition 3 *In an economy described by an ensemble* $\{p_i,\ |\psi_i\rangle\}$ *and its density operator* ρ*, the following relationship holds,*

$$\ln\left(1 + \frac{Income}{Assets}\right) = E[r|X_p] - S(\rho|X), \tag{56}$$

where X_p *denotes perfect information regarding the measured states.*

Proof. Proposition 2 suggests that the expected return (42) is written as

$$E[r|X] = E[r|X_p] + \sum_i \lambda_i \ln(\lambda_i) = E[r|X_p] - S(\rho|X), \tag{57}$$

where $S(\rho|X)$ denotes the von Neumann entropy of the density operator for a given information X. The accounting rate of return under economic income accounting converges to the expected return (see the discussions in Lemma 2) so that

$$\ln\left(1 + \frac{Income}{Assets}\right) = E[r|X]. \tag{58}$$

The result is immediate while combining (57) and (58). ■

Proposition 3 applies to all economic settings, in particular, the settings in which complete market does not exist. If there is no trading in some states of the economy (that is, the qubit states in the ensemble are not orthogonal, please refer to Example 5), taking the return with perfect information as a benchmark, the expected return is reduced by the amount equal to the von Neumann entropy. On the other hand, if the qubit states are orthogonal, the quantum decision problem is reduced to a classical problem which is depicted by an Arrow-Debreu economy. Then the von Neumann entropy of the initial ensemble is the same as the Shannon entropy, the latter is defined as $H\left(\{|\psi_i\rangle\}\right) = -\sum_i p_i \ln(p_i)$. In this sense, the quantum decision problem essentially stretches the domain of the classical decision problem by incorporating incomplete markets. Corollary 1 formally states the result.

Corollary 1 *In an economy described by an ensemble $\{p_i,\ |\psi_i\rangle\}$ and its density operator ρ, assume the states $|\psi_i\rangle$ are orthogonal. Then the following relationship holds,*

$$\ln\left(1 + \frac{Income}{Assets}\right) = E\left[r|X_p\right] - H\left(\{|\psi_i\rangle\}\right). \tag{59}$$

Proof. It is sufficient to show that the eigenvectors of the ensemble are $|\psi_i\rangle$ and the eigenvalues are p_i. The density operator of the ensemble is $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|$. Then we derive

$$\rho \left|\psi_j\right\rangle = \left(\sum_i p_i |\psi_i\rangle \langle\psi_i|\right) \left|\psi_j\right\rangle = p_j \left|\psi_j\right\rangle. \tag{60}$$

That is, $\lambda_j = p_j$ and $|v_j\rangle = \left|\psi_j\right\rangle$. ■

There are two implications. One, accounting stocks and flows are useful in quantum setting. Two, classical problems are seen to be a special case of quantum decision problems. Example 7 illustrates how accounting information connects with the von Neumann entropy.

Example 7 *Consider a new asset is acquired at the beginning of each period and generates periodic cash flows for three periods after acquisition. The cash flows are defined as $CF_i = k_i \left(e^r\right)^i$, where $k = \{k_i\} = \left[\frac{1}{3}, \frac{1}{6}, \frac{1}{2}\right]$ and $r = 0.3955$ is the rate of return calculated in Example 5. The economic value of the acquired* *asset is*

$$C = \sum_i \frac{CF_i}{(e^r)^i} = \sum_i \frac{k_i (e^r)^i}{(e^r)^i} = \frac{1}{3} + \frac{1}{6} + \frac{1}{2} = 1.$$

In the steady state, there will always be three productive assets. The ending balance of the asset, denoted by B, converges to a constant so that

$$B + C - Depreciation = B.$$

The periodic income is the expected rate of return multiplied by the assets available at the beginning of each period $(e^r - 1)(B + C)$. The depreciation expense can be derived from the income statement,

Revenues	$\sum_i CF_i = \frac{e^r}{3} + \frac{e^{2r}}{6} + \frac{e^{3r}}{2}$
Expenses	*Depreciation*
Income	$(e^r - 1)(B + C)$

so that Depreciation $= \sum_i CF_i - (e^r - 1)(B + C)$. *The steady state asset value can be derived as*

$$B + C - \sum_i CF_i + (e^r - 1)(B + C) = B$$

$$\Rightarrow \quad B = \frac{\sum_i CF_i - e^r C}{e^r - 1} = \frac{-\frac{2e^r}{3} + \frac{e^{2r}}{6} + \frac{e^{3r}}{2}}{e^r - 1} = 2.0929.$$

The periodic income is income $= (e^r - 1)(B + C) = \sum_i CF_i - C = 1.5004$. *Then*

$$\ln\left(1 + \frac{Income}{Assets}\right) = \ln\left(1 + \frac{Income}{B + C}\right) = 0.3955 = E[r].$$

Example 8 *In this example, we incorporate (nontrivial) partial information X. Continue Example 5 in which the economy is described by an ensemble*

$$\left\{ \frac{1}{2},\ \frac{1}{2},\ |\psi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ |\psi_2\rangle = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} \right\}.$$

If the decision maker does not have any information (except the ensemble) at the measurement, $E[r] = 0.4159$.

Suppose there is information source X producing two signals, x_1 and x_2, and the joint probabilities are defined as

$pr(\lvert\psi_1\rangle, x)$	$\lvert\psi_1\rangle$	$\lvert\psi_2\rangle$
x_1	$\frac{1}{4}$	0
x_2	$\frac{1}{4}$	$\frac{1}{2}$

The marginal probabilities of the qubit states are consistent with those in Example 5. All the other information is also consistent with Example 5.

If x_1 is observed, the signal perfectly predicts the state of the ensemble is $|\psi_1\rangle$ and the density operator $\rho(x_1) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. There is no uncertainty. The expected return given signal x_1 is

$$E[r|x_1] = E[r|X_p] - S(\rho|x_1) = 0.9163 - 0 = 0.9163.$$

If x_2 is observed, the decision maker updates his belief and the probability that the qubit state is $|\psi_1\rangle$ ($|\psi_2\rangle$) is $\frac{1}{3}$ ($\frac{2}{3}$). The revised ensemble is written as

$$\left\{ \frac{1}{3},\ \frac{2}{3},\ |\psi_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ |\psi_2\rangle = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} \right\}$$

and the density operator is computed as

$$\begin{aligned} \rho(x_2) &= \frac{1}{3}|\psi_1\rangle\langle\psi_1| + \frac{2}{3}|\psi_2\rangle\langle\psi_2| \\ &= \frac{1}{3}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{2}{3}\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix} = \begin{bmatrix} \frac{43}{75} & \frac{8}{25} \\ \frac{8}{25} & \frac{32}{75} \end{bmatrix}. \end{aligned}$$

The eigenvalues of $\rho(x_2)$ are computed as $\lambda_1(x_2) = \frac{15+\sqrt{97}}{30} \simeq 0.8283$, and $\lambda_2(x_2) = \frac{15-\sqrt{97}}{30} \simeq 0.1717$. The respective eigenvectors are

$$|v_1\rangle = \begin{bmatrix} 0.7821 \\ 0.6231 \end{bmatrix} \text{ and } |v_2\rangle = \begin{bmatrix} 0.6231 \\ -0.7821 \end{bmatrix}.$$

It is checked that $\langle v_1 | v_2\rangle = 0$ so that the two vectors are orthogonal. The expected return given signal x_2 is

$$E[r|x_2] = 0.8283 \times \ln(0.8283) + 0.1717 \times \ln(0.1717) + 0.9163 = 0.4577.$$

The expected return is higher when information source X is available. This is evident as $E[r|x_i] > 0.4159$, the latter, calculated in Example 5, is the expected *return given the prior knowledge is merely the ensemble. The expected return prior learning information X can be calculated as*

$$E[r] = \frac{1}{4} \times E[r|x_1] + \frac{3}{4} \times E[r|x_2] \simeq 0.5724.$$

4 Application–Performance Measurement

The equivalence relationship makes it convenient to look at problems in one discipline from the perspective of another discipline. One such example is the issue of performance measurement. Consider a large organization in which individuals, although work remotely, share information and work as a team. How shall the individual be evaluated, based on individual performance or based on group performance?

We illustrate that in this setting, it is impossible to do individual measurement without having a strictly negative impact on expected return. Speculation is offered about implications of individual measure when individuals sharing information and coordinating their behaviors is natural and beneficial. In this case individual measures are quite destructive.[7]

To properly describe the setting, we introduce entangled qubits and its nonlocality property in quantum mechanics.

Entangled qubits

So far, the state of an ensemble is described by a qubit. Now consider the state of a system is described by two qubits, which can be represented by a four element vector. For example, the standard states for two qubits are written as

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \text{ and } |11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

These states can be written as a the Kronecker matrix product (or tensor prod-

[7]Demski, Fellingham, Lin and Schroeder (2008) also demonstrates the detrimental effect of individual measurement in a setting in which qubit entanglement describes positive productive interaction.

uct) of two qubits $|0\rangle$ and/or $|1\rangle$. For example,

$$|01\rangle = |0\rangle \otimes |1\rangle = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}.$$

An important two-qubit state is the Bell state or EPR qubit pair (EPR stands for Einstein, Podolsky, and Rosen), written as[8]

$$|\beta_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\0\\1\end{bmatrix}.$$

An unique feature of the Bell state is that it cannot be written as a Kronecker product of any two qubits. This property is called "entanglement." Entangled qubits are correlated to the extent that they instantaneously share each other's state even when separated by any distance. When the two entangled qubits are moved far apart, a measurement of either qubit will allow the prediction with certainty of the state of the other qubit. This "non-locality" directly violates Einstein's principle of local action—meaning distant objects do not have direct impact on each other and therefore is viewed as a disturbing phenomenon.

Measurement

In economic setting, non-locality is not at all mysterious—individuals in an entity often share information and work cooperatively. Now the question is whether, in such setting, it makes sense to adopt group measurement or individual measurement. Consider the state of the system is the Bell state $|\beta_{00}\rangle$. The measurement axiom and the entropy concept discussed in previous sections also apply here. Since the state is unique and there is no uncertainty, the von Neumann entropy is zero. To see this, the density operator of the Bell state $|\beta_{00}\rangle$ is computed as

$$\rho_{|\beta_{00}\rangle} = |\beta_{00}\rangle\langle\beta_{00}| = \frac{1}{2}\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1 & 0 & 0 & 1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\1 & 0 & 0 & 1\end{bmatrix}.$$

[8] Three other Bell states are $|\beta_{01}\rangle = \frac{|01\rangle+|10\rangle}{\sqrt{2}}$, $|\beta_{10}\rangle = \frac{|00\rangle-|11\rangle}{\sqrt{2}}$, and $|\beta_{11}\rangle = \frac{|01\rangle-|10\rangle}{\sqrt{2}}$.

The eigenvalues of $\rho_{|\beta_{00}\rangle}$ are 1 and 0.[9] This suggests the optimal measurement vector is $|\beta_{00}\rangle$. So the von Neumann entropy is $S\left(\rho_{|\beta_{00}\rangle}\right) = -\ln(1) = 0$.

Alternatively, it is possible to measure one qubit at a time. The measurement matrix is constructed using the Kronecker product of a measurement matrix and the identity matrix. Consider a set of orthonormal measurement vectors $\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}1\\x\end{bmatrix}$ and $\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}-x\\1\end{bmatrix}$ where x can be any real number. To measure the first qubit using the measurement vector $\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}1\\x\end{bmatrix}$ while leaving the second qubit unmeasured, the measurement matrix is written as

$$\begin{aligned} M_1 &= \frac{1}{\sqrt{x^2+1}}\begin{bmatrix}1\\x\end{bmatrix}\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}1 & x\end{bmatrix}\otimes\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \\ &= \frac{1}{x^2+1}\begin{bmatrix}1 & x\\x & x^2\end{bmatrix}\otimes\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = \frac{1}{x^2+1}\begin{bmatrix}1 & 0 & x & 0\\0 & 1 & 0 & x\\x & 0 & x^2 & 0\\0 & x & 0 & x^2\end{bmatrix}. \end{aligned}$$

---

[9] The eigenvalues λ satisfies the following expression,

$$\det\left(\begin{bmatrix}\frac{1}{2}-\lambda & 0 & 0 & \frac{1}{2}\\0 & -\lambda & 0 & 0\\0 & 0 & -\lambda & 0\\\frac{1}{2} & 0 & 0 & \frac{1}{2}-\lambda\end{bmatrix}\right) = 0.$$

It is more complex to determine the eigenvalues for a 4×4 matrix.

$$\begin{aligned} &\det\left(\begin{bmatrix}\frac{1}{2}-\lambda & 0 & 0 & \frac{1}{2}\\0 & -\lambda & 0 & 0\\0 & 0 & -\lambda & 0\\\frac{1}{2} & 0 & 0 & \frac{1}{2}-\lambda\end{bmatrix}\right) \\ =& \left(\frac{1}{2}-\lambda\right)\det\left(\begin{bmatrix}-\lambda & 0 & 0\\0 & -\lambda & 0\\0 & 0 & \frac{1}{2}-\lambda\end{bmatrix}\right) - \frac{1}{2}\det\left(\begin{bmatrix}0 & -\lambda & 0\\0 & 0 & -\lambda\\\frac{1}{2} & 0 & 0\end{bmatrix}\right) \\ =& \left(\frac{1}{2}-\lambda\right)\left[(-\lambda)\det\left(\begin{bmatrix}-\lambda & 0\\0 & \frac{1}{2}-\lambda\end{bmatrix}\right)\right] - \frac{1}{2}\left[\lambda\det\left(\begin{bmatrix}0 & -\lambda\\\frac{1}{2} & 0\end{bmatrix}\right)\right] \\ =& \left(\frac{1}{2}-\lambda\right)(-\lambda)(-\lambda)\left(\frac{1}{2}-\lambda\right) - \frac{1}{2}\lambda\left(\frac{\lambda}{2}\right) \\ =& \lambda^2(1-\lambda)(-\lambda) = 0. \end{aligned}$$

Therefore, there are two distinct eigenvalues 1 and 0. The eigenvector associated with $\lambda = 1$ is $|\beta_{00}\rangle$.

Similarly, to measure the second qubit using the measurement vector $\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}-x\\1\end{bmatrix}$ while leaving the second qubit unmeasured, the measurement matrix is written as

$$\begin{aligned} M_2 &= \frac{1}{\sqrt{x^2+1}}\begin{bmatrix}-x\\1\end{bmatrix}\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}-x & 1\end{bmatrix}\otimes\begin{bmatrix}1&0\\0&1\end{bmatrix} \\ &= \frac{1}{x^2+1}\begin{bmatrix}x^2 & -x\\ -x & 1\end{bmatrix}\otimes\begin{bmatrix}1&0\\0&1\end{bmatrix} = \frac{1}{x^2+1}\begin{bmatrix}x^2 & 0 & -x & 0\\ 0 & x^2 & 0 & -x\\ -x & 0 & 1 & 0\\ 0 & -x & 0 & 1\end{bmatrix}. \end{aligned}$$

The post-measurement probability that the first qubit is $\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}1\\x\end{bmatrix}$ can be computed as

$$\begin{aligned} tr\left(M_1\rho_{|\beta_{00}\rangle}\right) &= tr\left(\frac{1}{x^2+1}\begin{bmatrix}1&0&x&0\\0&1&0&x\\x&0&x^2&0\\0&x&0&x^2\end{bmatrix}\frac{1}{2}\begin{bmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{bmatrix}\right) \\ &= tr\left(\frac{1}{2(x^2+1)}\begin{bmatrix}1&0&0&1\\x&0&0&x\\x&0&0&x\\x^2&0&0&x^2\end{bmatrix}\right) = \frac{1}{2}; \end{aligned}$$

similarly, the post-measurement probability that the first qubit is $\frac{1}{\sqrt{x^2+1}}\begin{bmatrix}-x\\1\end{bmatrix}$ can be computed as

$$\begin{aligned} tr\left(M_2\rho_{|\beta_{00}\rangle}\right) &= tr\left(\frac{1}{x^2+1}\begin{bmatrix}x^2&0&-x&0\\0&x^2&0&-x\\-x&0&1&0\\0&-x&0&1\end{bmatrix}\frac{1}{2}\begin{bmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{bmatrix}\right) \\ &= tr\left(\frac{1}{2(x^2+1)}\begin{bmatrix}x^2&0&0&x^2\\-x&0&0&-x\\-x&0&0&-x\\x^2&0&0&1\end{bmatrix}\right) = \frac{1}{2}. \end{aligned}$$

That is, for any measurement vector (described by x), the post-measurement probabilities would be one-half. The entropy is then calculated as

$$-\frac{1}{2}\ln\left(\frac{1}{2}\right)-\frac{1}{2}\ln\left(\frac{1}{2}\right)=\ln 2>0.$$

What we just illustrated is an application of Proposition 2—that is, any projective measurement that is not eigenbasis would increase entropy. In our economic setting, measuring using the entangled Bell state $|\beta_{00}\rangle$ can be viewed as group measurement as the two qubits are measured simultaneously. Measuring one qubit at a time can be viewed as individual measurement. Clearly, individual measurement increases the entropy than group measurement. Proposition 3 suggests in this case individual measurement also decreases the expected rate of return and therefore is corrosive.

5 Concluding Remarks

The main result of the paper is the equality between accounting number and quantum entropy, $S(Y|X)$,

$$\ln\left(1+\frac{Income}{Assets}\right)=E\left[r|X_p\right]-S\left(Y|X\right).$$

In an attempt to establish accounting as an information discipline, a previous paper had established conditions for the accounting-entropy equality using Shannon or classical entropy. One of the conditions for the equality was the existence of complete market. This paper demonstrates that, when quantum entropy is substituted for Shannon entropy, the domain of the equality is expanded to include incomplete market. (When markets are complete, quantum entropy equals Shannon entropy.)

An accounting question, that is, what number to assign to income and assets, can be reframed as an information question, that is, how much information does the accounting entity possesses. Since the equality goes both ways, an information question can be reframed as an accounting question.

There are advantages to the quantum approach besides expanding the domain of the accounting-entropy equality. One advantage is that quantum mathematics capture non-local effects. That is, distant units can share information instantaneously. While this has been a source of debate in physics, it does seem more plausible in an economic setting, where coordination and information sharing among units are routine and valuable. We illustrate in the paper that individual measurement of units can be destructive: entropy is increased and rate of return declines.

Another advantage of the quantum approach is that it anticipates an imminent technological change. A number of quantum technologies, such as quantum computers and quantum clocks, are already beginning to appear. Double entry accounting has proven useful over the certain notwithstanding dramatic technological shifts. It is somewhat comforting that accounting remains well positioned to adapt to another such shift.

References

Demski, Joel, John Fellingham, Haijin Lin, and Douglas Schroeder. 2008. "Interaction between productivity and measurement." *Journal of Management Accounting Research* 20: 169-190.

Fellingham, John, and Haijin Lin. "Is accounting an information science?" *Accounting, Economics, and Law: A Convivium.* Forthcoming

Kelly, John. 1956. "A New Interpretation of Information Rate." *The Bell System Technical Journal*, 35: 917-926.

Nielsen, Michael A. and Isaac L. Chuang. 2004. *Quantum Computation and Quantum Information*. Cambridge University Press.

Pacioli, Luca. 1494. *Summa de Arithmetica, Geometrica, Proporcioni et Proporcionalita*. Trans. by J. B. Geijsbeek as *Ancient Double-Entry Bookkeeping*. Reprinted by Houston: Scholars.

Shannon, Claude E. 1948. "A mathematical theory of communication." *The Bell System Technical Journal*, 27: 379-423.