Local Realistic Hidden-Variable Model for the States and Dynamics of Liquid-State NMR Information Processing

Nick Menicucci
Faculty Mentor: Carlton M. Caves

Department of Physics and Astronomy,
University of New Mexico, Albuquerque, NM 87131–1156
(Dated: April 19, 2002)

Abstract

Quantum computation exploits entanglement to speed up a certain class of information processing algorithms. Platforms for practical implementation of quantum computation are still being developed. High-temperature, liquid-state nuclear magnetic resonance (NMR) information processing may be used to emulate quantum computations, but it has been recently found that no entanglement is present in such NMR experiments up to about 12 nuclear spins. Still, it was hoped that the dynamics of such an NMR experiment would defy description by a local realistic model, thus establishing the "quantum" character of the technology. In this paper, a local realistic model of the states and dynamics of a bulk-ensemble NMR experiment is constructed explicitly, eliminating this possibility. This model applies to all bulk-ensemble information processing that accesses only separable states, including liquid-state NMR.

1. INTRODUCTION

The purpose of this chapter is to motivate research in quantum computation, to describe relevant background information relating to entanglement, Bell inequalities, and the potential for liquid-state NMR to act as a quantum computer, and to motivate the research reported in this thesis. There are many potential applications of quantum computation, most of which deal with reducing classically "impossible" problems to tractable ones by exploiting the quantum property of entanglement. The work in this thesis focuses on characterizing the information processing capabilities of liquid-state NMR experiments as a candidate for a viable quantum computer. Important background, therefore, includes an understanding of the motivation for developing quantum algorithms and realizing them in the physical world. To accomplish this, an example of a quantum factoring algorithm will be presented to illustrate the power of a quantum computer to perform an operation that becomes intractable by classical means very quickly as the number to be factored gets large (Sec. 1.1). This has vast implications in the area of public-key cryptography, where the impossibility of quickly factoring large numbers ensures the security of encrypted data transmissions. In the quantum formulation, the security inherent to large-number factoring is diminished by the ability of a quantum algorithm to factor large numbers efficiently. During the course of this description, some quantum computing terminology and concepts will be defined and used.

Once this motivation is established and some vocabulary obtained, we move on to a discussion of the potential for liquid-state NMR to act as a quantum computer (Sec. 1.2). High-temperature, liquid-state NMR, as a realization of bulk-ensemble quantum computation, never accesses entangled states, which makes it questionable as a possible quantum computing technology, since entanglement is thought to be of primary importance in quantum computation. Still, there remains the possibility that the dynamics of such NMR experiments might defy a realistic description, which could indicate an inherent "quantumness" to the technology.

After this discussion, density-operator notation will be introduced, along with definitions for separable and entangled states (Sec. 1.3). This discussion will be followed by a description of traditional and temporal Bell inequalities and their relationship to this research (Sec. 1.4). All Bell inequalities are formulated on two assumptions: locality and reality. These terms will be defined both in language and mathematically, and their relation to entanglement will be discussed. A traditional Bell inequality is a relationship that must hold if the subsystems of a composite system are to be considered to have definite properties that can be measured independently of measurements made on their sibling systems. A temporal Bell inequality places restrictions on successive measurements of a single system, assuming that the measurements do not disturb the system and assuming that the system is in a definite state at all times (i.e., assuming realistic dynamics). Most of reality can be described in this way, but some entangled states and certain quantum dynamics cannot. This will be explored.

Once this background has been illuminated for the reader, we will move on to discussing three successively improved local realistic hidden-variable models for liquid-state NMR experiments up to about 12 nuclear spins (Ch. 2), with the third model (Sec. 2.4) providing a complete description for all states and dynamics in any bulk-ensemble information processing platform, as long as the number of information processing units (e.g., nuclear spins, in the case of NMR) remains within an established bound. The existence of this model rules out

violation of any Bell inequality, traditional or temporal, for all bulk-ensemble information processing that accesses only separable states, including liquid-state NMR.

1.1. Quantum Computation—Motivation, Terminology, and An Example

Since the pioneering work of David Deutsch in 1985 [1], quantum computation algorithms have been searching for a practical implementation in physical reality. One motivation for such research has been the realization that a quantum computational algorithm [2], which exploits the entanglement of quantum states, may be used to significantly speed up the factoring of large numbers—an operation used in public-key cryptography [3, 4]. In order to break most modern public-key encryption codes, it is necessary to factor a large number into its two (large) prime factors. The security of this scheme lies in the inability to perform this factorization in a reasonable amount of time. Even with many computers working in parallel, the time it takes to factor a number grows exponentially in the number of bits used to represent the number, reaching many times the age of the universe after a relatively small number of bits [5]. However, if this factorization were to be performed on a quantum computer using a quantum algorithm, the time it would take to factor a number would grow only *polynomially* in the number of bits [6].

Before discussing the details of how this quantum magic works, some background on factoring is in order. Let's assume we wish to factor a number N, which is known to have only two prime factors. The most straightforward way of doing this is to try all prime factors beginning with 2 until one is found that divides into N evenly. This method is horribly inefficient, as the number of factors that must be attempted grows exponentially with the number of bits n used to represent the number N (i.e., it grows as 2^n).

A better method involves a bit of number theory. (This description follows that given in Ref. [6]. Several brief introductions to this method and its applications may also be found on-line [7].) In the following, a *residue* is the integer remainder obtained after division by N. Two integers a and b are said to be *congruent modulo* N if and only if the residue (remainder) of a/N equals the residue of b/N. This relation is written

$$a \equiv b \pmod{N} . \tag{1.1}$$

(Note that the symbol $\equiv$ as used above does *not* indicate equivalence but rather, congruence.) Suppose now that we wish to factor $N = ab$ into its two prime factors, a and b. The first step is to find two positive integers u and v that satisfy

$$uv \equiv 0 \pmod{N} , \quad \text{and} \tag{1.2}$$

$$u, v \not\equiv 0 \pmod{N} . \tag{1.3}$$

From these requirements, we know that

$$u \equiv \alpha a \pmod{N} \quad \text{and} \quad v \equiv \beta b \pmod{N} , \tag{1.4}$$

where the integers α and β satisfy

$$0 < \alpha < b \quad \text{and} \quad 0 < \beta < a . \tag{1.5}$$

We know that u and v have the form (1.4) because when multiplied together, they must give a multiple of N (1.2), but each alone must not be a multiple of N (1.3). This means

that one factor must be contained in the residue of u and one in that of v. Finally, since we are assuming that a and b are prime, then taking the greatest common factor (GCF) of u and N will necessarily give a, and taking the GCF of v and N will necessarily give b [8, 9].

If we write $u = s + t$ and $v = s - t$, then the requirements (1.2) and (1.3) become

$$(s+t)(s-t) \equiv 0 \pmod{N} \quad \Longrightarrow \quad s^2 \equiv t^2 \pmod{N}\,, \quad \text{and} \tag{1.6}$$

$$(s \pm t) \not\equiv 0 \pmod{N} \quad \Longrightarrow \quad s \not\equiv \pm t \pmod{N}\,, \tag{1.7}$$

respectively. If we can find two numbers s and t that satisfy these requirements, the we can extract the two prime factors of N as described above—by evaluating the greatest common factors of $(s \pm t)$ and N, the calculation time of which increases only polynomially in the number of bits [6].

We may implement this scheme (still in a classical sense) by looking at the sequence of residues obtained by evaluating the function

$$F_N(x) = a^x \bmod N\,, \tag{1.8}$$

where the right-hand side is an explicit instruction to divide the quantity a^x by N and keep only the remainder; the integer a satisfies $0 < a < N$, and x is a nonnegative integer. We find that for each chosen a, this sequence of residues has a period of repetition, which we will denote as r_a, such that $F_N(x + r_a) = F_N(x)$. Noting that

$$a^{r_a} \bmod N \; = F_N(r_a) = F_N(0) = 1\,, \tag{1.9}$$

we realize that

$$a^{r_a} \equiv 1 \pmod{N}\,. \tag{1.10}$$

If we are lucky enough to have an even r_a, and we are further fortunate enough to have $a^{r_a/2} \not\equiv \pm 1 \pmod{N}$, then this equation satisfies the requirements (1.6) and (1.7) with $s = a^{r_a/2}$ and $t = 1$. With N given to have only 2 prime factors, we may now obtain one of these factors by evaluating the GCF of $(a^{r_a/2} + 1)$ and N or that of $(a^{r_a/2} - 1)$ and N. It is important to note that not all choices of a will produce a viable r_a to satisfy the conditions mentioned above. However, as N becomes large, the chances of obtaining a viable period from a particular choice of a increases to more than 50% [6]. This is enough to ensure that the polynomial time-scaling of this method holds in the regime of large N, since we should not have to repeat this algorithm more than twice (on average) to obtain the factors of N.

Up to this point, we have not yet introduced quantum mechanics. All of the logic used so far has been classical in nature. Quantum computation takes the forefront in this algorithm when we consider that the time it takes for a classical computer to evaluate enough values of $F_N(x)$ in order to obtain a period r_a grows just as fast as in the brute-force case of dividing the number by all primes—i.e., exponentially in the number of bits representing N [6].

To illustrate how this method may be implemented and enhanced with a quantum algorithm, we introduce two state vectors, $|0\rangle$ and $|1\rangle$, that represent the binary digits, 0 and 1, respectively. These state vectors may be taken to be the orthogonal eigenstates of any two-state system (e.g., the spin-state of a spin-$\frac{1}{2}$ nucleus), which is called a *qubit* or "quantum bit" in this context. If we use two qubits, we can represent the integers 0–3 as follows:

$$\begin{aligned} 0: \quad |00\rangle &= |0\rangle \otimes |0\rangle \\ 1: \quad |01\rangle &= |0\rangle \otimes |1\rangle \\ 2: \quad |10\rangle &= |1\rangle \otimes |0\rangle \\ 3: \quad |11\rangle &= |1\rangle \otimes |1\rangle\,, \end{aligned} \tag{1.11}$$

where the first ket in the tensor product is the state of the designated first qubit, and the second ket in the tensor product is the state of the second qubit. The qubits must be distinguishable so that $|01\rangle$ is not the same as $|10\rangle$. With three qubits, the integers from 0–7 may be represented, and on to an arbitrary number of n qubits, representing 2^n possible integers.

With these definitions, the quantum factorization algorithm proceeds as follows. An initial state $|X\rangle$ is prepared that is a superposition of all possible values for the 2^n values represented by the qubits. For example, if $n = 2$, then

$$|X\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big) \otimes \frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big) = \frac{1}{2}\Big(|00\rangle + |01\rangle + |10\rangle + |11\rangle\Big) \; .$$

With our initial state $|X\rangle$ as a superposition of n-qubit representations of all possible values x, we can now perform all evaluations of the function $F_N(x)$ at once, instead of individually. This is accomplished by first tensor-multiplying our initial state with another initial state of n qubits initialized to $|0\rangle$, i.e., $|X\rangle \otimes |0\rangle$, where

$$|X\rangle \;=\; \frac{1}{\sqrt{2^n}}\Big(|0\rangle + |1\rangle\Big) \otimes \cdots \otimes \Big(|0\rangle + |1\rangle\Big) \; , \quad \text{and} \tag{1.12}$$

$$|0\rangle \;=\; |0\rangle \otimes \cdots \otimes |0\rangle \; , \tag{1.13}$$

where both repeated tensor multiplications occur n times. We now have a $2n$-qubit *quantum register*, comprised of $|X\rangle$ tensor-multiplied with $|0\rangle$, with which to perform calculations.

It can be shown that any classical operation, which must consist only of the gates AND, OR, and NOT, can be implemented on a quantum register consisting of a fixed number of qubits by a combination of unitary operations on individual qubits and specific 2- and 3-qubit unitary operations such as the controlled not (CNOT) gate and Toffoli gate, respectively [6, 10]. While it will not be shown explicitly here, with this combination of quantum gates at our disposal, we can simulate any classical reversible operation on our hypothetical quantum computer [6]. Thus, by some combination of unitary transformations in the classes listed above, we can perform an operation on our input state $|X\rangle \otimes |0\rangle$ that transforms it into

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |F_N(x)\rangle \; , \tag{1.14}$$

where $F_N(x)$ is defined in Eq. (1.8) using a suitable value of a. The input and output registers then become *entangled* in that the value of the second register $|F_N(x)\rangle$ is inexorably linked to that of the first register $|x\rangle$, even though the first register was not given a definite value of x to begin with. Now, if a projective measurement were to be made on the second register, that register would be forced into one of its eigenstates: a randomly chosen $|F_N(k)\rangle$, where k is a random nonnegative integer. Because of the entanglement between the two registers, this measurement would then force the input register $|x\rangle$ to assume a *superposition over all possible values* of x that will give that particular function value $F_N(k)$. Thus, the two-register system would become (ignoring normalization)

$$\Big(|k\rangle + |k + r_a\rangle + |k + 2r_a\rangle + \cdots\Big) \otimes |F_N(k)\rangle \; . \tag{1.15}$$

Finally, a *quantum Fourier transform* (also polynomial in n, the details of which may be found in Ref. [6]) may then be performed on the input register, from which the period r_a of

$F_N(x)$ may be easily obtained and subsequently used to extract the prime factors of N as described above.

The efficiency of the quantum factoring method lies in the simultaneous "evaluation" of all $F_N(x)$ using a $2n$-qubit register, although none of the individual values has been explicitly obtained except for one. Technically, therefore, we have not received all of the information that may be obtained by doing each individual calculation of $F_N(x)$ for every possible value of x, but this is irrelevant, since this is not the information that we are interested in. In fact, it is only the period of the function that we care about—not the individual values it takes. Thus, the quantum algorithm has succeeded in bypassing much unnecessary work that would otherwise have been done if the computation were performed using a classical algorithm, and it is this bypass that makes the quantum algorithm much more efficient (polynomial in n) than its classical counterpart (exponential in n). For more details on the quantum Fourier transform and this factoring method in general, see Refs. [2, 6, 11].

1.2. NMR as a Possible Quantum Computing Platform

With this motivation for a physical realization of a quantum computer, we examine the possibility of using liquid-state NMR for such an application. Liquid-state NMR is a mature technology [12], so using it as a quantum computing platform would make the technology very accessible to experimenters. As it stands, high-temperature, liquid-state NMR provides a testing ground for the new ideas for information processing that are being developed in quantum information science [13]. The qubits used in NMR are two-level nuclear spins, which are bound together in a single molecule. A liquid NMR sample contains a macroscopic number of molecules, each of which functions as an independent information-processing unit—i.e., an independent qubit. The molecules are initially in thermal equilibrium at high enough temperature that the nuclear spins are only weakly polarized along the direction of a strong magnetic field. NMR techniques cannot control the quantum states of individual molecules, and the measurements performed in NMR detect the average magnetization of the entire sample. For these reasons the use of high-temperature, liquid-state NMR to emulate quantum computation is called *bulk-ensemble quantum computation*.

The original proposals [14, 15] for quantum information processing using NMR were greeted with enthusiasm tempered by skepticism. The enthusiasm led to a remarkable series of experiments in which NMR techniques have been used to implement the operations for a variety of quantum-information-processing jobs involving up to seven qubits (for reviews of NMR information processing, see Refs. [16–21]). The persistent skepticism has to do with questions about the "quantumness" of NMR information processing. Initially based on doubt that the highly mixed states used in NMR could be used to achieve genuinely quantum-mechanical effects, these questions were made concrete by the realization that all the quantum states accessed in present experiments are *unentangled* [22]. (Entanglement is often thought to be an essential feature of quantum computation [23, 24] and will be discussed further in Secs. 1.3 and 1.4.) Arguments for an essential "quantumness" in NMR information processing are presented in Ref. [25], and an entirely different method for characterizing the "quantumness" of NMR is developed by Poulin [26].

The absence of entanglement in present NMR experiments means that the statistics of measurements made at any time during the experiments can be understood in terms of a *local realistic* hidden-variable (LRHV) model in which each spin has objective properties that determine the results of the measurements (more on this in Sec. 1.4). In a local realistic

model, the correlations observed in experiments can be attributed to classical correlations between realistic properties of the component qubits. The ability to describe the correlations observed in NMR experiments in terms of classical correlations between the qubits casts doubt on the "quantumness" of the experiments.

Modelling the statistics of the measurements made in NMR experiments is not sufficient for understanding the experiments. One must also be able to model the dynamics of the nuclear spins in a realistic way. NMR experimenters can implement with high accuracy any polynomially specifiable unitary operation, including the nonfactorizable unitary operations—those that cannot be written as a tensor-product of unitaries for each qubit—that produce entanglement when applied to pure quantum states. As discussed in the previous section, a complete set of quantum gates may be constructed out of such unitary operations to implement any quantum computing algorithm. Thus, it has been suggested that this ability of NMR to implement a complete set of quantum gates was an argument for its potential as a quantum computer. In fact, previous attempts [27] to devise a local realistic description of the dynamics were only partially successful in that they did not provide a local realistic description of the changes produced by nonfactorizable unitaries which reproduced all the predictions of quantum mechanics (see Sec. 2.3). This left open the possibility that one might not be able to describe the correlations observed in successive measurements separated by nonfactorizable unitary operations in terms of local realistic properties and thus that present NMR experiments might violate *temporal* Bell inequalities [28, 29] for successive measurements. Traditional Bell inequalities place restrictions on the correlations of measurements made on composite systems that admit a local, realistic description (i.e., a hidden-variable model). Temporal Bell inequalities generalize this result to correlations between successive measurements on a single system. Entanglement, Bell inequalities (traditional and temporal), and hidden-variable models are discussed further in Secs. 1.3 and 1.4.

1.3. Density Operator Notation and Entanglement

In the previous two sections, quantum computation research is motivated, and the potential for high-temperature, liquid-state NMR experiments to function as quantum computers is outlined as known to date, with previous research establishing that there is no entanglement in NMR experiments up to about 12 qubits [22, 27]. From this follows the basis of this thesis: to further characterize the ability of liquid-state NMR to function as a quantum computing platform by testing for violation of temporal Bell inequalities. It is important therefore to define what we mean by "entangled states" or "entanglement" and to discuss the traditional and temporal Bell inequalities.

As we discuss entanglement, we will establish some mathematical formalism to be used throughout the remainder of this document. The Dirac bra-ket notation is frequently used to label a quantum state in a Hilbert space. However, If we do not know the actual state of a system or—as in the case of liquid-state NMR—the system is an ensemble of smaller component systems, each of which could be in a different state, then we say that the system is in a "mixed state." In this case, we must use a *density operator* (or density matrix) to describe the system. This operator has the form of a convex combination of projectors onto the possible (or component) states of the system. If we label these states $|\psi_i\rangle$, then our

density operator $\hat{\rho}$ has the form

$$\hat{\rho} = \sum_{i=1}^{N} c_i |\psi_i\rangle\langle\psi_i| \ , \tag{1.16}$$

where $c_i \geq 0$, and $\sum_{i=1}^{N} c_i = 1$, and the sum is over all possible—or component, as in the case of an ensemble—states of the system. If all c_i are 0 except for one particular coefficient $c_k = 1$, then the density operator represents a pure state, and it has the simple form

$$\hat{\rho} = |\psi_k\rangle\langle\psi_k| \ , \tag{1.17}$$

which is just the projector onto the state $|\psi_k\rangle$.

The trace operation is very useful when working in density operators, so we will begin with an overview of its most important uses in this context. First, we establish that the trace of any density operator is 1. This can easily be seen by writing $\hat{\rho}$ as in Eq. (1.16) and using the fact that the trace is linear, that the trace is invariant under cyclic permutation of its multiplicative arguments (i.e., tr (ABC) = tr (BCA) = tr (CAB)), and that the states $|\psi_i\rangle$ are normalized:

$$\text{tr}\ \hat{\rho} \ = \ \text{tr} \sum_{i=1}^{N} c_i |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^{N} c_i \ \text{tr} \left(|\psi_i\rangle\langle\psi_i|\right) = \sum_{i=1}^{N} c_i \ \text{tr} \left(\langle\psi_i|\psi_i\rangle\right) = \sum_{i=1}^{N} c_i = 1 \ . \tag{1.18}$$

The trace has a further use in calculating expectation values for an operator $\hat{A}$ acting on a system with density operator $\hat{\rho}$. In Dirac notation, the expectation value for a given operator, given a particular state $|\psi_i\rangle$ is given by $\langle\hat{A}\rangle = \langle\psi_i|\hat{A}|\psi_i\rangle$. This quantity being a scalar (i.e., a 1×1 matrix), we do no harm by taking the trace of it: $\langle\hat{A}\rangle = \text{tr}\ \langle\psi_i|\hat{A}|\psi_i\rangle$. If we have a mixed state, then we must average over the expectation values of the possible states using the coefficients c_i as in Eq. (1.16):

$$\begin{aligned}
\langle\hat{A}\rangle \ &= \ \sum_{i=1}^{N} c_i \ \text{tr} \left(\langle\psi_i|\hat{A}|\psi_i\rangle\right) \\
&= \ \sum_{i=1}^{N} c_i \ \text{tr} \left(|\psi_i\rangle\langle\psi_i|\hat{A}\right) \\
&= \ \text{tr} \ \left[\left(\sum_{i=1}^{N} c_i |\psi_i\rangle\langle\psi_i|\right) \hat{A}\right] \\
&= \ \text{tr}\ (\hat{\rho}\hat{A}) = \text{tr}\ (\hat{A}\hat{\rho}) \ .
\end{aligned} \tag{1.19}$$

Thus, the expectation value of any operator, given a density operator $\hat{\rho}$ may be written as tr $(\hat{\rho}\hat{A})$ = tr $(\hat{A}\hat{\rho})$. Notice that this expectation value is manifestly independent of the particular decomposition (1.16) of $\hat{\rho}$. Thus, it is *the density operator itself* that determines all results of measurements made on a system. No particular decomposition of $\hat{\rho}$ may be argued to better represent the system's "actual" state; any decomposition will do.

Similarly, it is the density operator itself that evolves under unitary transformations. If a state $|\psi_i\rangle$, when acted upon by a unitary transformation $\hat{U}$, becomes $\hat{U}|\psi_i\rangle$, then the associated density operator $\hat{\rho}$ becomes $\hat{U}\hat{\rho}\hat{U}^{\dagger}$. For a single transformation, unitarity is required

to preserve the trace of $\hat{\rho}$:

$$\begin{aligned} \text{tr}\,(\hat{U}\hat{\rho}\hat{U}^\dagger) &= \text{tr}\,(\hat{\rho}\hat{U}^\dagger\hat{U}) \\ &= \text{tr}\,(\hat{\rho}\hat{1}) \\ &= \text{tr}\,(\hat{\rho}) = 1\,, \end{aligned} \tag{1.20}$$

which holds only if $\hat{U}^\dagger\hat{U} = \hat{1}$. A more general type of evolution may be used to evolve the density operator, which acts on ρ in the following way:

$$\hat{\rho}' = \sum_j \hat{A}_j\hat{\rho}\hat{A}_j^\dagger\,, \tag{1.21}$$

where the $\hat{A}_j$ operators are not unitary, but a similar condition is enforced to preserve the trace, namely $\sum_j \hat{A}_j^\dagger\hat{A}_j = \hat{1}$ [30].

We now turn our attention to a two-state system (i.e., a qubit) to establish some further properties of density operators related to representing states and how to characterize states as being entangled. Any qubit may have its spin measured to be either up or down along any spatial direction (unit-vector) $\boldsymbol{n}$. The operator that effects this measurement is $\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}$ [31]. If we know for sure that a measurement along a direction $\boldsymbol{m}$ will produce an "up" result, then quantum mechanics tells us that

$$\langle\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}\rangle = \text{tr}\left(\hat{\rho}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}\right) = \boldsymbol{n}\cdot\boldsymbol{m}\,. \tag{1.22}$$

The rules of the Pauli algebra tell us that $\text{tr}\,(\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{m}) = 2\boldsymbol{n}\cdot\boldsymbol{m}$, which tells us that the form of the density operator must be

$$\hat{\rho} = \frac{1}{2}(\hat{1} + \hat{\boldsymbol{\sigma}}\cdot\boldsymbol{m})\,, \tag{1.23}$$

where the extra term of $\hat{1}/2$ has been added to ensure that $\text{tr}\,\hat{\rho} = 1$. Eq. (1.23) gives the density operator corresponding to the pure state "up along $\boldsymbol{m}$," which is also often written as $|\boldsymbol{m}\rangle\langle\boldsymbol{m}|$. If we have more than one qubit, we may write the multiple-projector density operator as tensor products of this density operator (1.23), one for each qubit:

$$|\boldsymbol{n}_1\rangle\langle\boldsymbol{n}_1|\otimes\cdots\otimes|\boldsymbol{n}_N\rangle\langle\boldsymbol{n}_N| = \frac{1}{2^N}(\hat{1}+\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}_1)\otimes\cdots\otimes(\hat{1}+\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}_N)\,, \tag{1.24}$$

where N is the number of qubits.

We are now ready to define entanglement. All states—mixed or pure—may be written as a convex combination of possible (or component) states as in Eq. (1.16). For an *entangled state* of a multi-particle system, these states *cannot* be tensor products of states of the individual constituent systems. Returning to our qubit notation of Sec. 1.1, an example of such a two-qubit state is

$$|\psi_e\rangle = \frac{1}{\sqrt{2}}\Big(|00\rangle + |11\rangle\Big)\,, \tag{1.25}$$

with an associated density operator

$$\hat{\rho}_e = |\psi_e\rangle\langle\psi_e| = \frac{1}{2}\Big(|00\rangle + |11\rangle\Big)\Big(\langle 00| + \langle 11|\Big)\,. \tag{1.26}$$

Recall that the symbol $|ab\rangle = |a\rangle \otimes |b\rangle$. Now, we attempt to write this density operator (1.26) as

$$\hat{\rho} \;=\; \sum_j \sum_k c_{jk} |\phi_j\rangle\langle\phi_j| \otimes |\phi_k\rangle\langle\phi_k| \;, \tag{1.27}$$

where the sums run over a number of vectors $|\phi_n\rangle$ that form a general overcomplete basis (i.e., a basis with more vectors than necessary to span the space) for the space of one qubit. The possible tensor products of these vectors then form an overcomplete basis for the composite system.

The connection with Eq. (1.16) is made by writing $|\psi_i\rangle = |\phi_j\rangle \otimes |\phi_k\rangle$ and $c_i = c_{jk}$. This decomposition is, in general, possible since we have defined the projectors as forming an overcomplete basis for the 2-particle system, but for the case of $\hat{\rho}_e$ we find that at least one $c_{jk} < 0$ for any choice of *tensor-product* basis vectors $\{|\phi_j\rangle \otimes |\phi_k\rangle\}$ [32], meaning that this density operator represents a state of entanglement of the two constituent qubits [33].

Conversely, a density operator is called *separable* if writing it in the form of Eq. (1.27) can be done with all $c_{jk} \geq 0$. An example of such a state is

$$|\psi_s\rangle = \frac{1}{\sqrt{2}}\Big(|00\rangle + |01\rangle\Big) \;, \tag{1.28}$$

with an associated density operator

$$\hat{\rho}_s = |\psi_s\rangle\langle\psi_s| = \frac{1}{2}\Big(|00\rangle + |01\rangle\Big)\Big(\langle 00| + \langle 01|\Big) \;, \tag{1.29}$$

which may be written as

$$\hat{\rho}_s = |\boldsymbol{z}\rangle\langle\boldsymbol{z}| \otimes |\boldsymbol{x}\rangle\langle\boldsymbol{x}| \;, \tag{1.30}$$

where $|\boldsymbol{z}\rangle = |1\rangle$, and $|\boldsymbol{x}\rangle = (|1\rangle + |0\rangle)/\sqrt{2}$, which corresponds to one qubit in the "up along $\boldsymbol{z}$" state and one in the "up along $\boldsymbol{x}$" state (i.e., +1 eigenstates of $\hat{\sigma}_z$ and $\hat{\sigma}_x$, respectively). All separable density operators may be written in the form of Eq. (1.27) with all $c_{jk} \geq 0$ for some choice of tensor-product basis $\{|\phi_j\rangle \otimes |\phi_k\rangle\}$ [32].

The most important difference between entangled and separable states is that because separable states may be written as a convex combination of tensor-product states, the constituent qubits may be modelled as a collection of classical tops whose spin-axes are oriented in individual, particular directions. The "true" configuration is determined stochastically, but this stochasticity necessarily assumes that the system has a definite configuration at all times, even though we do not know for sure what it is. Because of our ignorance of the system's true configuration, all we can assign to the system are probabilities for the various possible configurations, but it is always assumed that one of these configurations truly represents the state of affairs in the system. The property described here is called *local realism*. It is called *realistic* because each qubit can be modelled as a top whose spin-axis actually points in a given direction; it is called *local* because the spin-axis of each qubit exists independently of that of the other qubits, and thus measurements of the spin-axis of one qubit do not affect measurements of the spin-axes of the other qubits. Correlations may exist between these spin directions, but such correlations can be determined by averaging over the various choices of spin configurations, weighted by some probability distribution that is given by the (nonnegative) coefficients c_{ij} in the decomposition of the density operator (1.27). Any state that is separable has no entanglement.

1.4. Bell Inequalities and Hidden-Variable Models

In the previous section, we discuss entanglement and separability and define the mathematical formalism that we use to describe a quantum system using density operators. At the very end of the section, local realism is defined. In this section, the properties of locality and reality are formalized, which leads to a description of traditional and temporal Bell inequalities.

The property of *local realism* mentioned at the end of Sec. 1.3 is the statement that subsystems of a composite system each have real properties that can be measured independently, without such measurements affecting or depending on other subsystems. This is not an unreasonable assumption about reality, considering it is true of most of our everyday experiences. In fact, in 1935, Einstein, Podolsky, and Rosen (EPR) asserted that quantum mechanics, in its probabilistic wave-function description, is incomplete [34] because it does not separately account for the apparently real properties of two spatially separated particles if such particles are in an entangled state. Bohr's reply to this assertion followed quickly [35] in an attempt to argue that the wave function is all that is physically predictable for a system. Nonetheless, Einstein held fast to a belief in a local realistic description of nature that retains the intuitive notions that all composite systems whose components are spatially separated should admit a real state for each component that can be measured independently of the other components. Einstein asserted that if the wave function in some cases cannot describe the real state of one subsystem independently of the real state of another, then the wave-function description of nature must be incomplete and should be supplemented with other variables [36]. This belief in so-called "hidden variables" remained at least plausible until 1964 when Bell formulated his famous inequality [37], which ruled out a local realistic description of entangled states that is consistent with the statistical predictions of quantum mechanics. Subsequent experiments consistently supported quantum mechanics, thus ruling out any local realistic description of entangled states [38]. What follows is a description of Bell's original inequality [37]. (See also Refs. [38–40] for more information and examples of other Bell inequalities.) This description also serves to illustrate the way locality and reality may be formalized in any hidden-variable model, including the models described in this thesis.

We begin by positing a system of two qubits that are in the (entangled) singlet state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|01\rangle - |10\rangle\Big) \,, \tag{1.31}$$

with an associated density operator

$$\hat{\rho} = |\psi\rangle\langle\psi| \;= \frac{1}{2}\Big(|01\rangle - |10\rangle\Big)\Big(\langle 01| - \langle 10|\Big) \,. \tag{1.32}$$

A measurement of the spins of the two constituent qubits along any single axis $\boldsymbol{a}$—i.e., $\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a} \;\otimes \hat{1}$ for the first qubit and $\hat{1}\otimes\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}$ for the second qubit—will result in perfect anticorrelation of the measurement results (i.e., one will be up, while the other will be down). Furthermore, we may define the correlation function,

$$C_{QM}(\boldsymbol{a},\boldsymbol{b}) \equiv \langle\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}\otimes\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{b}\rangle = \mathrm{tr}\,(\hat{\rho}\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}\otimes\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{b}) = -\boldsymbol{a}\cdot\boldsymbol{b} \,, \tag{1.33}$$

(where $\equiv$ once again indicates equivalence) for arbitrary choices of spin directions, $\boldsymbol{a}$ for the first qubit and $\boldsymbol{b}$ for the second qubit (note that $C(\boldsymbol{a},\boldsymbol{a}) = -1$, as required). Eq. (1.33) is

the prediction made by quantum mechanics for the correlation of measurement results along arbitrary axes of a two-qubit system in the (entangled) singlet state (1.31).

Following in the vein of Einstein [36], we wish to model this correlation function in terms of two deterministic measurement functions A and B, which are each functions only of the spin-measurement direction, $\boldsymbol{a}$ or $\boldsymbol{b}$, for the associated qubit (1 or 2, respectively) and a general hidden variable λ. This hidden variable may be any mathematical object or collection of such objects (scalars, vectors, tensors, functions, matrices, etc.), but we will treat it formally as if it were a continuous single parameter. If we wish to model the behavior of a quantum system in terms of these measurement functions, then the functions can only give out values of ± 1; thus,

$$A(\boldsymbol{a}, \lambda) = \pm 1 \ , \quad B(\boldsymbol{b}, \lambda) = \pm 1 \ . \tag{1.34}$$

It is through this way of writing of the measurement functions that locality is enforced: the measurement of the spin of qubit 1 does not depend on the direction of the measurement being made on qubit 2, and vice versa. If we had instead written the measurement function for the first qubit as $A(\boldsymbol{a}, \boldsymbol{b}, \lambda)$, then we would be constructing a nonlocal model. Although we cannot know the exact value of λ (which is why it is called a hidden variable), it is assumed to have a definite value, which is known probabilistically, and it is this probability distribution $P(\lambda)$ for the hidden variable that results in different "states" of the system. The expectation value of the correlated measurements of A and B is then given by

$$C_{HV}(\boldsymbol{a}, \boldsymbol{b}) = \int d\lambda \, P(\lambda) A(\boldsymbol{a}, \lambda) B(\boldsymbol{b}, \lambda) \ , \tag{1.35}$$

where the "HV" subscript indicates that this is the correlation predicted by the hidden-variable model. By requiring that $P(\lambda)$ be a nonnegative, normalized probability distribution over the possible actual values of λ, we enforce our second requirement, reality. The nonnegativity of this distribution implies that the hidden variable λ *actually has* one of its possible values; we don't know which one exactly, so we use a probability to describe our level of surety for each possible value of being the actual value.

Presuming quantum mechanics to give correct statistical predictions, the above expectation value must equal the quantum mechanically predicted correlation function for this system, i.e.,

$$C_{HV}(\boldsymbol{a}, \boldsymbol{b}) = C_{QM}(\boldsymbol{a}, \boldsymbol{b}) = -\boldsymbol{a} \cdot \boldsymbol{b} \tag{1.36}$$

for all choices of measurement directions, $\boldsymbol{a}$ and $\boldsymbol{b}$. It will be shown presently that this is in general impossible. Contradiction is demonstrated through the example of the singlet state given by Eq. (1.31). The requirement that $C_{HV}(\boldsymbol{a}, \boldsymbol{a}) = -1$ for agreement with quantum mechanics and the restrictions placed on $P(\lambda)$ require that

$$A(\boldsymbol{a}, \lambda) = -B(\boldsymbol{a}, \lambda) \tag{1.37}$$

for all choices of $\boldsymbol{a}$. Thus, we may rewrite Eq. (1.35) as

$$C_{HV}(\boldsymbol{a}, \boldsymbol{b}) = -\int d\lambda \, P(\lambda) A(\boldsymbol{a}, \lambda) A(\boldsymbol{b}, \lambda) \ . \tag{1.38}$$

If $\boldsymbol{c}$ is another measurement direction, then it follows from this result that

$$\begin{aligned} C_{HV}(\boldsymbol{a},\boldsymbol{b}) - C_{HV}(\boldsymbol{a},\boldsymbol{c}) &= -\int d\lambda\, P(\lambda)[A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda) - A(\boldsymbol{a},\lambda)A(\boldsymbol{c},\lambda)] \\ &= \int d\lambda\, P(\lambda)A(\boldsymbol{a},\lambda)[A(\boldsymbol{c},\lambda) - A(\boldsymbol{b},\lambda)] \\ &= \int d\lambda\, P(\lambda)A(\boldsymbol{a},\lambda)[A(\boldsymbol{b},\lambda)A(\boldsymbol{b},\lambda)A(\boldsymbol{c},\lambda) - A(\boldsymbol{b},\lambda)] \\ &= \int d\lambda\, P(\lambda)A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda)[A(\boldsymbol{b},\lambda)A(\boldsymbol{c},\lambda) - 1]\ , \end{aligned} \tag{1.39}$$

where the next-to-last equality follows from Eq. (1.34), from which we obtain

$$|C_{HV}(\boldsymbol{a},\boldsymbol{b}) - C_{HV}(\boldsymbol{a},\boldsymbol{c})| \leq \int d\lambda\, P(\lambda)[1 - A(\boldsymbol{b},\lambda)A(\boldsymbol{c},\lambda)]\ . \tag{1.40}$$

The right-hand side is just $1 + C_{HV}(\boldsymbol{b},\boldsymbol{c})$, which implies that

$$1 + C_{HV}(\boldsymbol{b},\boldsymbol{c}) \geq |C_{HV}(\boldsymbol{a},\boldsymbol{b}) - C_{HV}(\boldsymbol{a},\boldsymbol{c})|\ . \tag{1.41}$$

It is straightforward to see that this inequality is violated by the quantum mechanical correlation function (1.33). If we choose our unit vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ to be perpendicular, and we choose $\boldsymbol{c}$ to bisect the angle made by $\boldsymbol{a}$ and $\boldsymbol{b}$, then $C_{HV}(\boldsymbol{a},\boldsymbol{b}) = 0$, and $C_{HV}(\boldsymbol{b},\boldsymbol{c}) = C_{HV}(\boldsymbol{a},\boldsymbol{c}) = -\sqrt{2}/2 \approx -.707$. In this case the inequality of Eq. (1.41) is violated because

$$1 - .707 = .203 \ngeq .707\ . \tag{1.42}$$

The inequality in Eq. (1.41) is called a *Bell inequality* after its creator, John Bell. This inequality applies specifically to the (entangled) singlet state of two qubits [41], but all inequalities of this sort—i.e., those that place restrictions on correlations established through a local realistic hidden-variable model—are also called "Bell inequalities." Notice that in no way have we restricted what the hidden variable λ is. Thus, this result applies to *all* hidden variable models for a singlet state that may be formulated with a correlation function as in Eq. (1.35) with the further requirements that $P(\lambda) \geq 0$ and $\int d\lambda P(\lambda) = 1$. Models of this sort automatically satisfy the requirements of locality through the fact that $A(\boldsymbol{a},\lambda) \neq A(\boldsymbol{a},\boldsymbol{b},\lambda)$ (and similarly for B) and reality through the restriction that $P(\lambda)$ be a true probability distribution. Because all traditional Bell inequalities are founded on the assumptions of locality and reality, any system that admits a local realistic hidden-variable model description automatically satisfies any Bell inequality that may be derived for that system.

So far, our discussion of Bell inequalities has focussed on correlations between the subsystems of a composite system at a single moment in time. Another kind of Bell inequality called a *temporal Bell inequality* places restrictions on the correlations between successive measurements of the same system and may be violated for some quantum mechanical processes [28, 29]. These inequalities also employ the assumption of reality in that a system is assumed to have actual values for its measurable properties at any given time; thus it admits a hidden-variable description (although it may not be a local one), and such hidden variables follow well-defined trajectories. The assumption of spatial locality in traditional Bell inequalities is replaced in the temporal version by a sort of "temporal locality" in the form of noninvasive measurement functions. The assumption is that the measurement of

a property of a system at any time t_0 will not affect the result of a measurement of that realistic property at a later time t_1.

If we already have a local realistic model for a system, then temporal Bell inequalities place restrictions on the *dynamics* of the hidden variables in such a model. Satisfaction of temporal Bell inequalities requires a consistent updating of the hidden variables that allows them to follow well-defined trajectories. Updating based on the probability distribution over the hidden variables is therefore not allowed, because our usage of a probability distribution is based solely on our lack of information about the real state of the system, which is given by the actual value of λ. If the hidden variable λ were known, all randomness would disappear, and we would have a deterministic model with the ability to predict the result of any projective measurement simply by evaluating the appropriate $A_i(\boldsymbol{a}_i, \lambda)$ measurement function and using the correct value of λ. Thus the hidden variables must be updated based on their current configuration, either deterministically or stochastically.

Stochastic evolution involves transition probabilities that give the probability $T(\lambda'|\lambda) \geq 0$ for a new value λ' of the hidden variable, given the current value λ. Normalization requires that $\int d\lambda'\, T(\lambda'|\lambda) = 1$. Both deterministic and stochastic evolution (using transition probabilities) are inherently realistic because they assume an actual value of λ to begin with, which is mapped, either deterministically or stochastically, to a new value λ'. Deterministic evolution can be modelled as a special case of stochastic evolution where the transition probabilities are δ-functions that link (with 100% certainty) every value of λ to the corresponding value of λ'.

Previous analysis of the states used in bulk-ensemble NMR information processing [22] found that all states used in such experiments are separable (i.e., unentangled) and thus are known to admit a local realistic hidden-variable model description (although no such model was explicitly constructed until now). An attempt was made [27] to devise a consistent dynamical model that would satisfy the aforementioned requirements for temporal Bell inequalities, but the results were not satisfactory because they did not completely duplicate the predictions of quantum mechanics. This is the launching point for the current work on this project—creation of a complete local realistic model for bulk-ensemble NMR information processing that includes realistic dynamics and noninvasive measurement functions, guaranteeing satisfaction of all traditional and temporal Bell inequalities.

2. LOCAL REALISTIC MODELS FOR LIQUID-STATE NMR EXPERIMENTS

This chapter presents an overview of the methods and formalism behind NMR information processing (Sec. 2.1), followed by three successively better local realistic hidden-variable models for a general bulk-ensemble quantum computation that never accesses entangled states. High-temperature, liquid-state NMR experiments up to about 12 nuclear spins fall into this category. The first model (Sec. 2.2) is simply a model for the states and is based on previous work in this area [22]. The second (Sec. 2.3) is a model that corresponds to the partially successful attempt by Schack and Caves [27] to model the dynamics of an NMR experiment in a local realistic way. The final model, presented in Sec. 2.4, is the culmination of this work, an overview of which is given in [42]. It is a complete local realistic hidden-variable model for all states and dynamics of high-temperature liquid-state NMR experiments that remain in the separable regime (i.e., up to about 12 nuclear spins). The aforementioned requirements of reality, spatial locality, well-defined hidden-variable trajectories, and noninvasive measurement functions are all satisfied by this model, thus

ensuring that the model will satisfy any Bell inequality, traditional or temporal, that might be created for it. This result applies to all bulk-ensemble information processing that uses separable states exclusively.

2.1. NMR Information Processing Overview

All NMR quantum computing experiments performed so far work in the following way. The state of each molecule, consisting of N active spin-$\frac{1}{2}$ nuclei, is described by a density operator

$$\hat{\rho} = (1-\epsilon)\hat{1}/2^N + \epsilon\hat{\rho}_1 \ , \tag{2.1}$$

which is a mixture of the desired state of the quantum computer, $\hat{\rho}_1$, with the maximally mixed state for N qubits, $\hat{1}/2^N$, $\hat{1}$ being the unit operator. When $\hat{\rho}_1$ is a pure state, $\hat{\rho}$ is called a *pseudopure state* [14].

The molecules in an NMR sample begin in thermal equilibrium, with a weak polarization $\alpha = h\nu/2kT \sim 2 \times 10^{-5}$ at room temperature, where $\nu \sim 300\,\mathrm{MHz}$ is the average resonant frequency of the active spins in the strong longitudinal magnetic field. The first step in NMR information processing is to transform the molecules from equilibrium to a pseudopure state [15, 43]. A consequence of pseudopure state synthesis, in contrast to distillation of a pure state [44], is that the mixing parameter scales like $\epsilon = \alpha N/2^N$.

After synthesis of the desired initial state, the computation begins. The unitary operations required for the computation can be constructed from sequences of radio-frequency pulses alternating with periods of continuous evolution under the nuclear-spin Hamiltonian [16–21]. A unitary operator $\hat{U}$ takes an input state $\hat{\rho}$ to an output state

$$\hat{U}\hat{\rho}\,\hat{U}^\dagger = (1-\epsilon)\hat{1}/2^N + \epsilon\hat{U}\hat{\rho}_1\hat{U}^\dagger \ . \tag{2.2}$$

The maximally mixed state is unaffected by the unitary transformation. The output state retains the form (2.1) with the same value of ϵ, and—this is the essence of the bulk-ensemble paradigm for quantum computation—$\hat{\rho}_1$ undergoes the desired unitary transformation.

The computation completed, the last step is to read out the answer. By applying radio-frequency pulses and then measuring the transverse magnetization of the sample, an NMR experimenter can determine the expectation value of any product of spin components, one for each qubit [16–21]. These expectation values have the form

$$\begin{aligned} C(\tilde{\boldsymbol{a}}) &\equiv C(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N) \\ &= \mathrm{tr}\,(\hat{\rho}\,\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}_1 \otimes \cdots \otimes \hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}_N) \\ &= \epsilon\,\mathrm{tr}\,(\hat{\rho}_1\,\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}_1 \otimes \cdots \otimes \hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}_N) \ . \end{aligned} \tag{2.3}$$

Here and throughout, a tilde over a quantity denotes a collection of N such quantities, one for each spin. In Eq. (2.3) the tensor product includes one operator for each spin; the vector operator $\hat{\boldsymbol{\sigma}} \equiv \hat{1}\boldsymbol{e}_0 + \hat{\sigma}_x\boldsymbol{e}_x + \hat{\sigma}_y\boldsymbol{e}_y + \hat{\sigma}_z\boldsymbol{e}_z$, where $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$ are the Pauli operators; and $\boldsymbol{a}_r$ is either a spatial unit vector, in which case $\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{a}_r$ is the component of spin r along the direction $\boldsymbol{a}_r$, or the unit vector $\boldsymbol{e}_0$ in the "zero" direction, in which case spin r does not contribute to the expectation value. The last equality in Eq. (2.3) assumes at least one of the vectors $\boldsymbol{a}_r$ is a spatial direction.

The expectation values (2.3) express the correlations between spin components of different spins. The maximally mixed state does not contribute to the expectation values, which are

determined by the state $\hat{\rho}_1$ that undergoes the desired evolution. The mixing parameter ϵ measures the strength of the magnetization signal. The scaling $\epsilon = \alpha N/2^N$ that comes from pseudopure state synthesis thus leads to an in-principle demand for an exponentially increasing number of molecules as the number of qubits increases [14, 45], implying that bulk-ensemble quantum computation is not suitable for large-scale quantum computation.

Before constructing our local realistic hidden-variable (LRHV) model, we recall that in Sec. 1.3, we claimed that any N-qubit density operator $\hat{\tau}$ can be written in the form

$$\hat{\tau} \;=\; \sum_i \sum_j c_{ij} |\phi_i\rangle\langle\phi_i| \otimes |\phi_j\rangle\langle\phi_j| \;, \tag{2.4}$$

even though the c_{ij} are not in general positive [32]. Still we can choose an appropriate basis and group these coefficients together into a *quasidistribution*, $w_{\hat{\tau}}$, that is associated with the density operator $\hat{\tau}$. This distribution $w_{\hat{\tau}}$ is called a quasidistribution because for some (separable) states it can be nonnegative, thus qualifying as an traditional probability distribution, but for entangled states it necessarily takes on some negative values. A natural, though not unique quasidistribution may be extracted from a general N-qubit density operator $\hat{\tau}$ as follows [32]:

$$w_{\hat{\tau}}(\tilde{\boldsymbol{n}}) \equiv \mathrm{tr}\left(\hat{\tau}\,\hat{Q}(\tilde{\boldsymbol{n}})\right) , \tag{2.5}$$

where the vectors in the set $\tilde{\boldsymbol{n}} \equiv (\boldsymbol{n}_1, \ldots, \boldsymbol{n}_N)$ are spatial unit vectors, and

$$\hat{Q}(\tilde{\boldsymbol{n}}) \equiv \frac{1}{\mathcal{N}^N}(\hat{1} + 3\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}_1) \otimes \cdots \otimes (\hat{1} + 3\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}_N) \;. \tag{2.6}$$

For each spin, the unit vector $\boldsymbol{n}$ can point in $\mathcal{N}$ different directions satisfying

$$0 \;=\; \sum_{\boldsymbol{n}} n_j \;, \tag{2.7}$$

$$\frac{1}{3}\delta_{jk} \;=\; \frac{1}{\mathcal{N}} \sum_{\boldsymbol{n}} n_j n_k \;, \tag{2.8}$$

where the sums are over the possible directions, and the subscripts indicate spatial components of $\boldsymbol{n}$. Such a collection of directions forms an appropriate overcomplete basis for separating $\hat{\tau}$, as was outlined in Sec. 1.3. Condition (2.8) means that the vectors $\sqrt{3/\mathcal{N}}\boldsymbol{n}$ form a resolution of the 3-dimensional unit tensor; condition (2.7) places an additional constraint on the placement of the vectors. The vertices of a tetrahedron give the minimum number, $\mathcal{N} = 4$, of possible directions. The six vectors along the cardinal directions make up another simple possibility.

The density operator is then given by [32]

$$\hat{\tau} = \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\tau}}(\tilde{\boldsymbol{n}}) |\tilde{\boldsymbol{n}}\rangle\langle\tilde{\boldsymbol{n}}| \;, \tag{2.9}$$

where $|\tilde{\boldsymbol{n}}\rangle\langle\tilde{\boldsymbol{n}}| \equiv |\boldsymbol{n}_1\rangle\langle\boldsymbol{n}_1| \otimes \cdots \otimes |\boldsymbol{n}_N\rangle\langle\boldsymbol{n}_N|$ and $|\boldsymbol{n}\rangle\langle\boldsymbol{n}| = \frac{1}{2}(\hat{1} + \hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n})$ is the $+1$ eigenstate of $\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{n}$, or "up along $\boldsymbol{n}$," and the sum is over all possible sets of vectors $\tilde{\boldsymbol{n}}$. In terms of $w_{\hat{\tau}}(\tilde{\boldsymbol{n}})$, the expectation values (2.3) take the form

$$C(\tilde{\boldsymbol{a}}) = \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\tau}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \boldsymbol{a}_j \cdot \boldsymbol{n}_j \;. \tag{2.10}$$

This correlation function works even if $\boldsymbol{a}_r = \boldsymbol{e}_0$ for a particular direction if each vector $\boldsymbol{n}_j$ is supplemented with a unit-component in the $\boldsymbol{e}_0$-direction (i.e., $\boldsymbol{n}_j \to \boldsymbol{n}_j + \boldsymbol{e}_0$). This "0-component" never changes or evolves under transformations; it is simply a formal supplement to the 3 spatial components of $\boldsymbol{n}_j$ to account for the possibility of ignoring the spin of a particular qubit.

Under a unitary transformation the quasidistribution evolves according to

$$w_{\hat{U}\hat{\tau}\hat{U}^\dagger}(\tilde{\boldsymbol{n}}') = \sum_{\tilde{\boldsymbol{n}}} T^{\hat{U}}_{\tilde{\boldsymbol{n}}'\tilde{\boldsymbol{n}}} w_{\hat{\tau}}(\tilde{\boldsymbol{n}}) \; . \tag{2.11}$$

Here the transformation matrix $T^{\hat{U}}$ has matrix elements

$$T^{\hat{U}}_{\tilde{\boldsymbol{n}}'\tilde{\boldsymbol{n}}} \equiv \langle \tilde{\boldsymbol{n}} | \hat{U}^\dagger \hat{Q}(\tilde{\boldsymbol{n}}') \hat{U} | \tilde{\boldsymbol{n}} \rangle = w_{\hat{U}|\tilde{\boldsymbol{n}}\rangle\langle\tilde{\boldsymbol{n}}|\hat{U}^\dagger}(\tilde{\boldsymbol{n}}') \; , \tag{2.12}$$

given by the quasidistribution for $\hat{U}|\tilde{\boldsymbol{n}}\rangle$.

We now recall that a *separable* density operator is one that has an ensemble decomposition in terms of product states (Sec. 1.3). Such a state has no entanglement. If the quasidistribution $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$ is everywhere nonnegative, then $\hat{\rho}$ is definitely separable, and the statistics of all measurements can be understood in terms of classical tops whose probability to point in the directions $\tilde{\boldsymbol{n}}$ is $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$.

For any density operator, the quasidistribution satisfies [27, 32]

$$w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \geq [\text{minimum eigenvalue of } \hat{Q}(\tilde{\boldsymbol{n}})] = \frac{(-2)4^{N-1}}{\mathcal{N}^N} = \frac{-2^{2N-1}}{\mathcal{N}^N} \; . \tag{2.13}$$

Thus for density operators of the form (2.1), the quasidistribution is everywhere nonnegative if [22, 32]

$$\epsilon \leq \frac{1}{1+2^{2N-1}} \equiv \eta \; . \tag{2.14}$$

Such states are *unentangleable* by any unitary transformation. For the polarization $\alpha \sim 2 \times 10^{-5}$ of present NMR experiments, all states up to about 12 qubits are unentangleable. It is known [46] that entangled states of the form (2.1) exist for $\epsilon > (1+2^{N-1})^{-1} \equiv \eta'$, i.e., $N \gtrsim 2/\alpha$, but whether there are entangled states for $\eta < \epsilon \leq \eta'$ is an open question.

2.2. Simple Hidden-Variable Model

We turn now to constructing a LRHV model for unentangleable states, i.e., for $\epsilon \leq \eta$. A straightforward model regards the directions $\tilde{\boldsymbol{n}}$ as hidden spin directions that determine the results of measurements stochastically. To do this, we write the density operator as

$$\hat{\rho} = \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) |\tilde{\boldsymbol{n}}\rangle\langle\tilde{\boldsymbol{n}}| \; , \tag{2.15}$$

where $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$ is the quasidistribution for $\hat{\rho}$ obtained from Eq. (2.5). If $\hat{\rho}$ is unentangleable, then $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$ is guaranteed to be nonnegative. This gives a correlation function

$$C(\tilde{\boldsymbol{a}}) = \langle \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{a}_1 \otimes \cdots \otimes \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{a}_N \rangle = \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \boldsymbol{a}_j \cdot \boldsymbol{n}_j \; . \tag{2.16}$$

Notice that this works even when $\boldsymbol{a} = \boldsymbol{e}_0$ if we allow the vectors $\boldsymbol{n}$ not only to be unit vectors in the "spatial dimensions," but also always to have a fixed unit component in the zero direction, as stated above in the discussion following Eq. (2.10).

As it stands, this hidden-variable model is stochastic because the hidden spin directions, $\boldsymbol{n}_j$, determine the probabilities for measurement results, instead of determining the results themselves. We can always put a deterministic model underneath a stochastic one. In this case that job is accomplished by associating an additional hidden variable Λ_j with each spin, so that the whole collection of hidden variables, called λ in Sec. 1.4, is $(\tilde{\boldsymbol{n}}, \tilde{\Lambda}) \equiv (\boldsymbol{n}_1, \ldots, \boldsymbol{n}_N, \Lambda_1, \ldots, \Lambda_N)$. Now we allow each Λ_j to be uniformly distributed between -1 and $+1$, meaning that the entire probability distribution for λ becomes

$$P(\lambda) = P(\tilde{\boldsymbol{n}}, \tilde{\Lambda}) = \frac{1}{2^N} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \; . \tag{2.17}$$

We now further define our measurement functions to be

$$A_j(\boldsymbol{a}, \lambda) = A_j(\boldsymbol{a}, \Lambda_j, \boldsymbol{n}_j) = \begin{cases} +1, & \text{if } \Lambda_j \geq -\boldsymbol{a} \cdot \boldsymbol{n}_j \; , \\ -1, & \text{if } \Lambda_j < -\boldsymbol{a} \cdot \boldsymbol{n}_j \; . \end{cases} \tag{2.18}$$

Notice that for $\boldsymbol{a} = \boldsymbol{e}_0$, this definition gives $A_j(\boldsymbol{e}_0, \lambda) = +1$, as required if a measurement is not being performed on qubit j. All this works because

$$\frac{1}{2} \int_{-1}^{+1} d\Lambda_j \, A_j(\boldsymbol{a}, \Lambda_j, \boldsymbol{n}_j) = -\frac{1}{2} \int_{-1}^{-\boldsymbol{a}\cdot\boldsymbol{n}_j} d\Lambda_j + \frac{1}{2} \int_{-\boldsymbol{a}\cdot\boldsymbol{n}_j}^{+1} d\Lambda_j = \boldsymbol{a} \cdot \boldsymbol{n}_j \; , \tag{2.19}$$

thus giving

$$\begin{aligned} C_{HV}(\tilde{\boldsymbol{a}}) \; &= \; \int d\lambda \, P(\lambda) \prod_{j=1}^{N} A_j(\boldsymbol{a}_j, \lambda) \\ &= \; \int d\tilde{\Lambda} \, \frac{1}{2^N} \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} A_j(\boldsymbol{a}_j, \lambda) \\ &= \; \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \frac{1}{2} \int_{-1}^{+1} d\Lambda_j \, A_j(\boldsymbol{a}_j, \Lambda_j, \boldsymbol{n}_j) \\ &= \; \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \boldsymbol{a}_j \cdot \boldsymbol{n}_j \; . \end{aligned} \tag{2.20}$$

Since the deterministic hidden-variable model gives the same correlation function as that given by quantum mechanics (2.16), we can use it to reconstruct that density operator in order to derive from it statistical predictions for measurements that can't be formulated in terms of the hidden variables, e.g., measurements in an entangled-state basis for two qubits. This reconstruction is accomplished via

$$\hat{\rho} = \frac{1}{2^N} \sum_{\alpha_1, \ldots, \alpha_N} C_{HV}(\boldsymbol{e}_{\alpha_1}, \ldots, \boldsymbol{e}_{\alpha_N}) \, \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{e}_{\alpha_1} \otimes \cdots \otimes \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{e}_{\alpha_N} \; . \tag{2.21}$$

Here the Greek indices run from 0 to 3, with $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$ being orthogonal spatial vectors and, as before, $\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{e}_0 = \hat{1}$. As before, one should note that the density operator places constraints on the coefficients $C_{HV}(\boldsymbol{e}_{\alpha_1}, \ldots, \boldsymbol{e}_{\alpha_N})$ extracted from the correlation function,

so arbitrary coefficients generally lead to an operator in Eq. (2.21) that is not a density operator. If one knows, however, that the coefficients are consistent with a density operator (for example, because they are determined by evaluation of a correlation function derived from a hidden-variable model that duplicates the predictions of quantum mechanics for some density operator), then one can construct a density operator (2.21) from these coefficients and derive from it statistical predictions for measurements that can't be formulated in terms of the hidden variables, e.g., measurements in an entangled-state basis for two spins.

This model succeeds in providing a local realistic description of all states accessed in present high-temperature liquid-state NMR experiments. Reality is enforced through the use of a hidden variable $\lambda \equiv (\tilde{\boldsymbol{n}}, \tilde{\Lambda})$ to describe the "real" state of the system, with an associated nonnegative probability distribution $P(\lambda)$ as given by Eq. (2.17), and locality is enforced through the definition of the measurements functions (2.18) as depending only on the associated measurement direction and not on the measurement directions for the other qubits. Thus, it is impossible that any experiment described by such a model will violate any traditional Bell inequality. Although the form of the measurement functions (2.18) is noninvasive in that they do nothing to change λ, but we have said nothing about the dynamics of the hidden variables. Updating λ in a realistic (trajectory-oriented) way might be problematic, and thus, it is conceivable that experiments that obey this model may still violate temporal Bell inequalities.

2.3. Improved Hidden-Variable Model

In their attempt to close this loophole by providing a realistic description of NMR dynamics, Schack and Caves [27] realized that a general entangling unitary operation $\hat{U}$—even when applied to an unentangleable state—resulted in an associated transition matrix $T^{\hat{U}}$ (2.12) that generally took on negative values and thus could not be used as a transition probability for $w_{\hat{\rho}}$ as in Eq. (2.11). This meant that the hidden variables in the previous section could not be updated in a realistic (trajectory-oriented) fashion, and thus that that model did not satisfy the requirement of realistic dynamics.

The improved model presented here is based on the work by Schack in Caves [27] to improve on this dynamical description. The authors were able to obtain nonnegative transition probabilities by modifying Eq. (2.12), but these transitions left the spins noisier than they were before the operation. Still, these noise-making transition probabilities proved useful in the following way: a second set of spins were added to the model that were as tightly correlated with the actual state $\hat{\rho}_1$ of the NMR system as was allowed by the requirement of unentangleability (i.e., they had a signal-to-noise ratio of η). Then, the measured spins were allowed to correlate loosely with these "hidden spins" to get the right mixing parameter ϵ in the predicted results of the correlation function. After an entangling unitary operation, the hidden spins would lose some coherence due to the noisy transition probabilities, but simultaneously, the correlation between the measured spins and their hidden counterparts would become tighter by a calculated amount so as to maintain the correct signal-to-noise ratio ϵ in the predicted measurements. After a finite number of gates, however, this model would break down because the measured spins could not become any more tightly correlated with the hidden spins. Still, it is useful here to show how this sort of multi-layered dynamics and correlation could be written in the formalism of an explicit LRHV model.

Although Schack and Caves used a completely new set of spins in their model, this improved model adds only two new hidden variables, τ and n, to the total hidden variable

set λ to effect the same result. The hidden variable τ is binary, taking on values 0 and 1; its value determines whether the spin components are entirely random (if $\tau = 0$) or have the properties of the simple hidden-variable model of Sec. 2.2 (if $\tau = 1$). The hidden variable n takes on integer values starting at 0, indicating the total number of entangling gates through which the system has passed. It is used in the evolution of the system to determine the new weighting of the probability of τ.

The values of the spin components for the jth spin are denoted by a new function,

$$B_j(\boldsymbol{a}, \lambda) = B_j(\boldsymbol{a}, \tilde{\boldsymbol{n}}, \tilde{\Lambda}, n, \tau) \ , \tag{2.22}$$

defined by

$$B_j(\boldsymbol{a}, \tilde{\boldsymbol{n}}, \tilde{\Lambda}, n, \tau = 1) = A_j(\boldsymbol{a}, \Lambda_j, \boldsymbol{n}_j) \tag{2.23}$$

and

$$B_j(\boldsymbol{a}, \tilde{\boldsymbol{n}}, \tilde{\Lambda}, n, \tau = 0) \equiv B_j^0(\boldsymbol{a}, \Lambda_j) = \begin{cases} +1, & \text{if } \boldsymbol{a} \cdot \boldsymbol{e}_0 = 1, \\ +1, & \text{if } \boldsymbol{a} \cdot \boldsymbol{e}_0 = 0 \text{ and } \Lambda_j \geq 0, \\ -1, & \text{if } \boldsymbol{a} \cdot \boldsymbol{e}_0 = 0 \text{ and } \Lambda_j < 0. \end{cases} \tag{2.24}$$

(Note that these B's have no relation to the B measurement function of Sec. 1.4.) The probability distribution for the hidden variables can be written as

$$P(\lambda) = P(\tilde{\Lambda}) P(\tilde{\boldsymbol{n}}) \delta_{nk} P(\tau | k) = \frac{1}{2^N} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \delta_{nk} P(\tau | k) \ , \tag{2.25}$$

where k is the step number that gets incremented after each entangling gate. Notice that

$$\frac{1}{2} \int_{-1}^{+1} d\Lambda_j \, B_j^0(\boldsymbol{a}, \Lambda_j) = \begin{cases} 1, & \text{if } \boldsymbol{a} \cdot \boldsymbol{e}_0 = 1, \\ 0, & \text{if } \boldsymbol{a} \cdot \boldsymbol{e}_0 = 0. \end{cases} \tag{2.26}$$

Now we can evaluate the correlation functions as for the simple model, replacing the A's with the B's:

$$\begin{aligned} C_{HV}(\tilde{\boldsymbol{a}}) \ &= \ \int d\lambda \, P(\lambda) \prod_{j=1}^{N} B_j(\boldsymbol{a}_j, \lambda) \\ &= \ \sum_{\tau} P(\tau | k) \int d\tilde{\Lambda} \, \frac{1}{2^N} \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} B_j(\boldsymbol{a}_j, \lambda) \\ &= \ P(\tau = 0 | k) \prod_{j=1}^{N} \frac{1}{2} \int_{-1}^{+1} d\Lambda_j \, B_j^0(\boldsymbol{a}_j, \Lambda_j) \\ &\quad + P(\tau = 1 | k) \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \frac{1}{2} \int_{-1}^{+1} d\Lambda_j \, A_j(\boldsymbol{a}_j, \Lambda_j, \boldsymbol{n}_j) \\ &= \ P(\tau = 0 | k) \prod_{j=1}^{N} \frac{1}{2} \int_{-1}^{+1} d\Lambda_j \, B_j^0(\boldsymbol{a}_j, \Lambda_j) \\ &\quad + P(\tau = 1 | k) \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \boldsymbol{a}_j \cdot \boldsymbol{n}_j \ . \end{aligned} \tag{2.27}$$

If all the $\boldsymbol{a}$'s are equal to $\boldsymbol{e}_0$, then this result gives 1, as it must. If any of the $\boldsymbol{a}$'s is not equal to $\boldsymbol{e}_0$, the first term in the sum vanishes, and we get

$$C_{HV}(\tilde{\boldsymbol{a}}) = P(\tau = 1|k) \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \boldsymbol{a}_j \cdot \boldsymbol{n}_j \ . \tag{2.28}$$

This is the correct correlation function because it assumes that the quasidistribution $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$ has a higher signal-to-noise ratio than it should, which is compensated for by multiplying by the probability that the random variable τ is a 1. Specifically, in the Schack-Caves model, the probability $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$ is updated at each entangling gate by a classical transition probability that leaves a new distribution whose signal strength is reduced relative to quantum mechanics by a factor of η. This is compensated for in the Schack-Caves model by having the probability $P(\tau = 1|k)$ increase by a factor of η^{-1} until the counter index k reaches a maximum value K, at and above which $P(\tau = 1|k) = 1$, from which point on the model predicts an exponential decrease in signal-to-noise. As presented by Schack and Caves, this is accomplished by having two sets of spins: one set has its signal strength reduced at each entangling gate and the other set, which determines the measurement statistics, becomes more tightly correlated with the first set at each entangling gate. The presentation here is much simpler, dispensing with the second set of vector-valued hidden variables in favor of a single binary parameter τ. In the dynamics the probability $P(\tau = 1|k)$ would be determined anew after the gate: increasing by a factor of η^{-1} if $k \leq K$, and thereafter always remaining at 1. This corresponds to reweighting and reflipping the τ coin at each entangling gate, the new weighting determined by the total number of gates k through which the system has passed.

The key point for the dynamics here is that the reweighting and reflipping is independent both of the kind of entangling gate and of the system's state, although it requires knowledge of how many entangling gates have been performed. The analogous point in the original Schack-Caves presentation is that the tightening of the correlation is independent of the kind of entangling gate and of the system's state. In contrast, a similar reweighting and reflipping of $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$ in the previous model (Sec. 2.2) requires knowledge of the kind of entangling gate *and* of the initial system state (i.e., the quasidistribution). This is unacceptable if we want realistic dynamics based on the assumption that one particular configuration of $\tilde{\boldsymbol{n}}$ is the actual state of the system at all times. Thus, this model is an improvement over that of Sec. 2.2, but it does not completely satisfy the requirements for satisfaction of all temporal Bell inequalities because it breaks down after a finite number of entangling gates.

2.4. Local Realistic Dynamical Model

In this section is described a complete local realistic hidden-variable (LRHV) model for the states and dynamics of bulk-ensemble NMR information processing up to about 12 qubits. The existence of such a model rules out violation of any Bell-type inequality in present NMR experiments. This conclusion applies only to the bulk-ensemble model of information processing realized in present high-temperature, liquid-state NMR experiments; it does not apply to NMR methods based on distilling a pure state from a thermal state [17, 44, 47].

This third hidden-variable model is based conceptually on the original model presented by Schack and Caves [27], which, in its original formulation, includes another set of hidden

spins in the total hidden-variable set λ. These spins stochastically determine the probability distribution of the observed spins, which in turn directly determine the results of individual spin measurements through the additional variables $\tilde{\Lambda} \equiv (\Lambda_1, \ldots, \Lambda_N)$. While such a physical picture was unnecessary in duplicating the dynamics of this model, as was demonstrated in the previous section, the original idea has inspired the creation of a new DeBroglie-Bohm-type hidden-variable model [48] for the entirety of an NMR experiment. Instead of another set of spins, this model includes a suitable representation of the density operator in the set of hidden variables. However, unlike the Bohmian formulation of quantum mechanics in general, this model is locally real.

Using Eq. (2.5) for the quasidistribution for a given density operator, each density operator $\hat{\eta}$ generates $\mathcal{N}^N$ values of the associated function $w_{\hat{\eta}}(\tilde{\boldsymbol{n}})$ as the unit vectors range over their possible directions. We collect these values into an $\mathcal{N}^N$-component vector denoted by $\bar{w}_{\hat{\eta}}$, which is associated in a 1-to-1 fashion with its generating function $w_{\hat{\eta}}(\tilde{\boldsymbol{n}})$. With the function $w_{\hat{\eta}}(\tilde{\boldsymbol{n}})$ as a bridge, what we have set up is a 1-to-1 correspondence between density operators $\hat{\eta}$ and vectors $\bar{w}_{\hat{\eta}}$. The vectors $\bar{w}_{\hat{\eta}}$ corresponding to density operators span a 4^N-dimensional subspace of the $\mathcal{N}^N$-dimensional vector space of all vectors $\bar{w}$. (See the Appendix for further discussion of the vectors $\bar{w}$.)

In the following, $\hat{\eta}$ denotes an arbitrary density operator of the form (2.1) satisfying Eq. (2.14), and $\hat{\rho}$ denotes the "actual" system density operator. The new set of hidden variables is

$$\lambda = (\bar{w}, \tilde{\boldsymbol{n}}, \tilde{\Lambda}) \; . \tag{2.29}$$

The probability distribution over λ is now chosen to be

$$P(\lambda) = P(\bar{w}, \tilde{\boldsymbol{n}}) P(\tilde{\Lambda}) \; , \tag{2.30}$$

with

$$P(\bar{w}, \tilde{\boldsymbol{n}}) = \delta(\bar{w} - \bar{w}_{\hat{\rho}}) w(\tilde{\boldsymbol{n}}) \tag{2.31}$$

and

$$P(\tilde{\Lambda}) = \frac{1}{2^N} \; . \tag{2.32}$$

From the particular form of (2.31), it is to be understood that the probability distribution of the hidden spins is determined by the choice of the hidden vector. Thus, this distribution of the spins is identically the function $w(\tilde{\boldsymbol{n}})$ associated with the hidden vector $\bar{w}$ picked out by the δ-function. Notice that $w(\tilde{\boldsymbol{n}})$ must be nonnegative for this model to work.

We now use the A-functions of Eq. (2.18) to write the correlation function for a given set of spin-directions $\tilde{\boldsymbol{a}}$ as

$$\begin{aligned} C_{HV}(\tilde{\boldsymbol{a}}) \; &= \; \int d\lambda \, P(\lambda) \prod_{j=1}^{N} A_j(\boldsymbol{a}_j, \lambda) \\ &= \; \sum_{\tilde{\boldsymbol{n}}} \int d\bar{w} \, d\tilde{\Lambda} \, P(\bar{w}, \tilde{\boldsymbol{n}}) P(\tilde{\Lambda}) \prod_{j=1}^{N} A_j(\boldsymbol{a}_j, \lambda) \\ &= \; \sum_{\tilde{\boldsymbol{n}}} \int d\bar{w} \, \delta(\bar{w} - \bar{w}_{\hat{\rho}}) \, w(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \frac{1}{2} \int_{-1}^{+1} d\Lambda_j \, A_j(\boldsymbol{a}_j, \Lambda_j, \boldsymbol{n}_j) \\ &= \; \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \prod_{j=1}^{N} \boldsymbol{a}_j \cdot \boldsymbol{n}_j \; . \end{aligned} \tag{2.33}$$

Comparison with Eq. (2.10) shows this gives the right correlation coefficients.

Dynamical evolution is handled in the following way: the vector $\bar{w} = \bar{w}_{\hat{\eta}}$ is mapped deterministically to a new vector $\bar{w}' = \bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger}$, followed by "spinning a roulette wheel," weighted by $\bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger}$, to get new hidden-spin directions $\tilde{\boldsymbol{n}}$ that are distributed according to the distribution $\bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger}$. Thus the distribution over the spins is determined anew after each unitary operation by a spin of a roulette wheel that is weighted by the vector selected after the gate operation completes. The dynamics of a transition becomes simply a map of initial vectors $\bar{w}_{\hat{\eta}}$ to final vectors $\bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger}$, corresponding to initial and final density operators, $\hat{\eta}$ and $\hat{U}\hat{\eta}\hat{U}^\dagger$, followed by a roulette wheel spin to get $w(\tilde{\boldsymbol{n}})$. The linear transformation for this transition,

$$w_{\hat{U}\hat{\eta}\hat{U}^\dagger}(\tilde{\boldsymbol{n}}') = \mathrm{tr}\left(\hat{U}\hat{\eta}\hat{U}^\dagger\hat{Q}(\tilde{\boldsymbol{n}}'\right) = \mathrm{tr}\left(\hat{\eta}\hat{U}^\dagger\hat{Q}(\tilde{\boldsymbol{n}}')\hat{U}\right) = \sum_{\tilde{\boldsymbol{n}}} w_{\hat{\eta}}(\tilde{\boldsymbol{n}})\langle\tilde{\boldsymbol{n}}|\hat{U}^\dagger\hat{Q}(\tilde{\boldsymbol{n}}')\hat{U}|\tilde{\boldsymbol{n}}\rangle \ , \quad (2.34)$$

is effected by the matrix (2.12) as in Eq. (2.11). As shown in the Appendix, this matrix acts in the 4^N-dimensional subspace spanned by vectors $\bar{w}$ that correspond to density operators. In that subspace it has an inverse given by $T^{\hat{U}^{-1}}$. Schack and Caves [27] wanted to use this as a transition probability but were unable to do so because the matrix isn't nonnegative. In the present context, that doesn't matter. What matters is the property they liked: the transition function for each unitary operator $\hat{U}$ is simply a different 1-to-1 linear map on the vector space and is manifestly independent of the hidden-variable distribution, $P(\lambda)$.

To ensure that our dynamics is explicitly realistic, we will write out this process in terms of transition probabilities. Transition probabilities give the probability that, given a particular value of λ for the system, a new value λ' will result after a gate operation. This kind of dynamics is inherently realistic because a nonnegative probability of transition exists for every possible initial value of λ. Thus, using transition probabilities for the dynamics assumes right from the start that the hidden variable has a definite value to begin with; then, given this value, the transition probability matrix gives the probability that a new particular value of the hidden variable will result after the gate operation completes. Formally, this is the requirement that any new distribution $P'(\lambda') = P(\bar{w}', \tilde{\boldsymbol{n}}', \tilde{\Lambda}')$ should be determined by the original distribution $P(\lambda) = P(\bar{w}, \tilde{\boldsymbol{n}}, \tilde{\Lambda})$ as follows:

$$P'(\lambda') = \int d\lambda \, T_{\hat{U}}(\lambda'|\lambda)P(\lambda) \ . \quad (2.35)$$

where $T_{\hat{U}}(\lambda'|\lambda)$ is the transition probability, associated with a unitary operation $\hat{U}$ and independent of the system state $\hat{\rho}$, for obtaining a final hidden variable λ', given an initial hidden variable λ.

We now define a candidate $T_{\hat{U}}(\lambda'|\lambda)$, show that it is properly normalized when integrated over λ', and finally show that it generates the correct transition from $P(\lambda)$ to $P'(\lambda')$ when applied as in Eq. (2.35). Let

$$T_{\hat{U}}(\lambda'|\lambda) \equiv \frac{1}{2^N}\int d\bar{w}_{\hat{\eta}} \, \delta(\bar{w}' - \bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger}) \, \delta(\bar{w} - \bar{w}_{\hat{\eta}}) \, w'(\tilde{\boldsymbol{n}}') \ , \quad (2.36)$$

where the integration is over all vectors $\bar{w}_{\hat{\eta}}$ associated with unentangleable density operators $\hat{\eta}$—i.e., those density operators (2.1) satisfying (2.14). This transition probability generates δ-function transitions from initial vectors $\bar{w}$ to final vectors $\bar{w}'$ whenever these vectors correspond to initial and final states connected by the unitary operator $\hat{U}$, while simultaneously

extracting a new distribution $w'(\tilde{\boldsymbol{n}}')$ from the resulting vector associated with the final quantum state of the system. Note that the transition probability is determined wholly by $\hat{U}$ and manifestly has no dependence on the initial state $\hat{\rho}$ or the initial distribution of the hidden spins $w_{\hat{\rho}}(\tilde{\boldsymbol{n}})$, as required by a realistic dynamical model. The initial correlations of the hidden spins have no effect on the hidden-variables' transitions (eliminating the troublesome decrease in signal-to-noise, as in the model of Sec. 2.3), since the final correlations of the hidden spins are extracted anew from the vector obtained through the 1-to-1 map associated with $\hat{U}$.

It is straightforward to verify that $T_{\hat{U}}(\lambda'|\lambda)$ is normalized:

$$\begin{aligned}
\int d\lambda'\, T_{\hat{U}}(\lambda'|\lambda) &= \sum_{\tilde{\boldsymbol{n}}'} \int d\bar{w}'\, d\tilde{\Lambda}'\, T_{\hat{U}}(\lambda'|\lambda) \\
&= \int d\bar{w}'\, d\bar{w}_{\hat{\eta}}\, \delta(\bar{w}' - \bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger})\, \delta(\bar{w} - \bar{w}_{\hat{\eta}}) \sum_{\tilde{\boldsymbol{n}}'} w'(\tilde{\boldsymbol{n}}') \frac{1}{2^N} \int d\tilde{\Lambda}' \\
&= 1\ . \qquad (2.37)
\end{aligned}$$

We now verify that $T_{\hat{U}}(\lambda'|\lambda)$ satisfies (2.35):

$$\begin{aligned}
P'(\lambda') &= \int d\lambda\, T_{\hat{U}}(\lambda'|\lambda) P(\lambda) \\
&= \sum_{\tilde{\boldsymbol{n}}} \int d\bar{w}\, d\tilde{\Lambda}\, \frac{1}{2^N} \int d\bar{w}_{\hat{\eta}}\, \delta(\bar{w}' - \bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger})\, \delta(\bar{w} - \bar{w}_{\hat{\eta}})\, w'(\tilde{\boldsymbol{n}}') \frac{1}{2^N}\, \delta(\bar{w} - \bar{w}_{\hat{\rho}}) w(\tilde{\boldsymbol{n}}) \\
&= \frac{1}{2^N}\, w'(\tilde{\boldsymbol{n}}') \int d\bar{w}\, d\bar{w}_{\hat{\eta}}\, \delta(\bar{w}' - \bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger})\, \delta(\bar{w} - \bar{w}_{\hat{\eta}})\, \delta(\bar{w} - \bar{w}_{\hat{\rho}}) \\
&= \frac{1}{2^N}\, w'(\tilde{\boldsymbol{n}}') \int d\bar{w}_{\hat{\eta}}\, \delta(\bar{w}' - \bar{w}_{\hat{U}\hat{\eta}\hat{U}^\dagger})\, \delta(\bar{w}_{\hat{\eta}} - \bar{w}_{\hat{\rho}}) \\
&= \frac{1}{2^N}\, \delta(\bar{w}' - \bar{w}_{\hat{U}\hat{\rho}\hat{U}^\dagger})\, w'(\tilde{\boldsymbol{n}}')\ , \qquad (2.38)
\end{aligned}$$

which is consistent with Eqs. (2.31) and (2.32). In the final step we use the fact that our map from density operators to vectors is 1-to-1 to conclude that the integral identically sets $\hat{\eta} = \hat{\rho}$.

With a simple local realistic hidden-variable model for the states of a high-temperature, liquid-state NMR experiment, such as we created in Sec. 2.2, it is guaranteed that no such experiment may violate any traditional Bell inequality because it satisfies the premises on which all traditional Bell inequalities are founded: reality and locality. It also has noninvasive measurement functions given by Eq. (2.18), but the dynamics are not explicitly realistic. An attempt to generate realistic dynamics in Sec. 2.3 fails because the model breaks down after a finite number of entangling gates. The model in this section, however, also has realistic dynamics with no such breakdown. The existence of such a model—one with realistic dynamics and noninvasive measurements—ensures that every NMR experiment will satisfy all temporal Bell inequalities, as well. It is also worth noting that this model is not restricted to NMR experiments. Any bulk-ensemble information processing implementation that accesses only unentangleable states—i.e., those satisfying Eq. (2.14)—may be modelled as shown in this section and thus cannot violate any type of Bell inequality.

2.5. Discussion and Implications

The previous section presented a complete local realistic hidden-variable model that correctly describes all aspects of liquid-state NMR information processing, including modelling unitary evolution of states as the deterministic evolution of a hidden vector $\bar{w}$. This vector is basically a stand-in for the quantum density operator, and we could just as well regard the components of the density matrix itself as the hidden variables. In either case, the state lies in a 4^N-dimensional subspace (see the Appendix) and determines the probabilities for the hidden spins and thus for measurement results. The dynamics is a map of this subspace onto itself.

It is not really so surprising that this model for the unitary dynamics of an NMR experiment must include another mathematical object (i.e., another hidden variable) beneath the hidden spin set $\tilde{\boldsymbol{n}}$. Stochastic unitary evolution of a probability distribution is limited by an inability to make stronger correlations than were contained in the original configuration (which was the problem encountered by Schack and Caves [27]), thus limiting the prospects for evolution of the spins-distribution itself. However, when the spin directions are stochastically determined by another *more fundamental* hidden variable, which itself undergoes the evolution, we no longer are required to evolve the spins themselves based on their current configuration.

This dynamics of the model described in Sec. 2.4 applies to a discrete unitary transformation, but it is easy to generalize it to a quasicontinuous hidden-variable dynamics: each molecule updates randomly in the fashion just described; i.e., each has a probability $\gamma\, dt$ to update within each time interval dt. The only requirement on this quasicontinuous dynamics is that the mean time γ^{-1} between updates be large compared to the precession time ν^{-1} of the nuclear spins in the strong magnetic field, but shorter than the duration of the radio-frequency pulses that are used to produce the desired dynamics.

It is trivial to generalize the LRHV model presented here to nonunitary evolutions (1.21), since these evolutions are, like unitary transformations, linear in the density operator. It is likely that the model could be extended to include all dynamics that accesses only separable states—i.e., states that have an expansion like Eq. (2.9), but with more general nonnegative quasidistributions than the canonical form (2.5).

The LRHV model developed here achieves the purpose of determining whether present NMR experiments can violate Bell inequalities. Traditional Bell inequalities [37–40] are founded on two assumptions: realism and locality. Realism is the assumption that systems have objective properties that determine measurement results, and locality assumes that measurements made on one subsystem of a composite system do not depend on what measurements are made on any other subsystem. Temporal Bell inequalities [28, 29], which involve successive measurements on a single system, are based on realistic dynamics and noninvasive measurements. Realistic dynamics implies that at all times, the hidden variables of a system are in a definite configuration, and these configurations evolve according to transition probabilities. The assumption of noninvasive measurements means that measurements made on the system at a particular time do not influence the actual values of realistic properties, which may be subsequently measured at a later time. This LRHV model for NMR experiments (Sec. 2.4) describes the statistics of all measurements in terms of realistically evolving classical correlations between realistic properties of the constituent nuclear spins, and measurements are modelled so that they do not alter the system in any way. Thus, reality, locality, realistic dynamics, and noninvasive measurements are all incorporated in the

model. The conclusion is, therefore, that NMR experiments up to about 12 qubits cannot violate any Bell inequality, whether traditional or temporal.

This purpose achieved, it should be acknowledged that this LRHV model is terribly contrived. It succeeds in giving a local realistic description of the dynamics by the brute force device of including an encoding of the entire density operator among the hidden variables. As a result, there are an exponentially increasing number of hidden variables, $\sim 4^N$ in the most efficient version of the model, which is to be expected for a complete local realistic description [49], and updating them is exponentially inefficient. It leaves open the possibility that [27] the "quantumness" of NMR information processing lies in the ability to implement nonfactorizable unitary operations that do not have an efficient local realistic description.

APPENDIX A: FURTHER DISCUSSION OF THE HIDDEN VECTORS

While a given $\mathcal{N}^{\mathcal{N}}$-component vector $\bar{w}$ generally lives in a $\mathcal{N}^{\mathcal{N}}$-dimensional vector space, we wish to show that such vectors that correspond to density operators really live in a 4^N-dimensional subspace. We also wish to show that the transition matrix (2.12) first projects a general vector $\bar{w}$ on this subspace and then maps this subspace onto itself, resulting in a new vector $\bar{w}'$ that corresponds to unitary evolution of the density operator represented by the projection of the original $\bar{w}$ onto the aforementioned subspace.

For simplicity we restrict attention first to a single qubit, generalizing below to the case of many qubits. An arbitrary vector $\bar{w}$ has $\mathcal{N}$ components $w(\boldsymbol{n})$. If the vector is associated with a density operator $\hat{\rho} = \frac{1}{2}(\hat{1} + \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{S})$, the components are given by

$$w_{\hat{\rho}}(\boldsymbol{n}) = \mathrm{tr}\,\Big(\hat{\rho}\,\hat{Q}(\boldsymbol{n})\Big) = \frac{1}{\mathcal{N}}\mathrm{tr}\,\Big(\hat{\rho}\,(\hat{1} + 3\hat{\boldsymbol{\sigma}} \cdot \boldsymbol{n})\Big) = \frac{1}{\mathcal{N}}(1 + 3\boldsymbol{n} \cdot \boldsymbol{S})\ . \tag{A1}$$

Although vectors associated with density operators can have $\mathcal{N}$ nonzero components, they span only a four-dimensional subspace, since the density operator $\hat{\rho}$ lies in a four-dimensional vector space. Formally, one can display the four-dimensional subspace by defining four vectors $\bar{0}$, $\bar{x} = \bar{1}$, $\bar{y} = \bar{2}$, and $\bar{z} = \bar{3}$, whose components are given by

$$0(\boldsymbol{n}) = \frac{1}{\sqrt{\mathcal{N}}}\ , \quad x(\boldsymbol{n}) = \sqrt{\frac{3}{\mathcal{N}}}n_x\ , \quad y(\boldsymbol{n}) = \sqrt{\frac{3}{\mathcal{N}}}n_y\ , \quad z(\boldsymbol{n}) = \sqrt{\frac{3}{\mathcal{N}}}n_z\ . \tag{A2}$$

The conditions (2.7) and (2.8) imply that these four vectors are orthonormal. A vector associated with a density operator can be written as

$$\begin{aligned} \bar{w}_{\hat{\rho}} &= \frac{1}{\sqrt{\mathcal{N}}}\Big(\bar{0} + \sqrt{3}S_x\bar{x} + \sqrt{3}S_y\bar{y} + \sqrt{3}S_z\bar{z}\Big) \\ &= \frac{1}{\sqrt{\mathcal{N}}}\left(\bar{0} + \sqrt{3}\sum_{j=1}^{3} S_j\bar{j}\right)\ , \end{aligned} \tag{A3}$$

with its components given by

$$\begin{aligned} w_{\hat{\rho}}(\boldsymbol{n}) &= \frac{1}{\sqrt{\mathcal{N}}}\Big(0(\boldsymbol{n}) + \sqrt{3}S_x x(\boldsymbol{n}) + \sqrt{3}S_y y(\boldsymbol{n}) + \sqrt{3}S_z z(\boldsymbol{n})\Big) \\ &= \frac{1}{\sqrt{\mathcal{N}}}\left(0(\boldsymbol{n}) + \sqrt{3}\sum_{j=1}^{3} S_j j(\boldsymbol{n})\right)\ . \end{aligned} \tag{A4}$$

Using $\hat{U}\hat{\boldsymbol{\sigma}}\hat{U}^\dagger = R^{-1}\hat{\boldsymbol{\sigma}}$, where R is the rotation matrix associated with $\hat{U}$ (R projects into the spatial directions), we find that the transition matrix (2.12) for a single qubit becomes

$$\begin{aligned} T^{\hat{U}}_{\boldsymbol{n}'\boldsymbol{n}} &= w_{\hat{U}|\boldsymbol{n}\rangle\langle\boldsymbol{n}|\hat{U}^\dagger}(\boldsymbol{n}') \\ &= w_{|R\boldsymbol{n}\rangle\langle R\boldsymbol{n}|}(\boldsymbol{n}') \\ &= \frac{1}{\mathcal{N}}(1 + 3\boldsymbol{n}'\cdot R\boldsymbol{n}) \\ &= \frac{1}{\mathcal{N}}(1 + 3R^{-1}\boldsymbol{n}'\cdot \boldsymbol{n}) \\ &= \frac{1}{\sqrt{\mathcal{N}}}\left(0(\boldsymbol{n}) + \sqrt{3}\sum_{k=1}^{3}(R^{-1}\boldsymbol{n}')_k k(\boldsymbol{n})\right) . \end{aligned} \tag{A5}$$

Now we notice that

$$\sum_{\boldsymbol{n}} T^{\hat{U}}_{\boldsymbol{n}'\boldsymbol{n}} 0(\boldsymbol{n}) = \frac{1}{\sqrt{\mathcal{N}}} \tag{A6}$$

$$\sum_{\boldsymbol{n}} T^{\hat{U}}_{\boldsymbol{n}'\boldsymbol{n}} j(\boldsymbol{n}) = \sqrt{\frac{3}{\mathcal{N}}}(R^{-1}\boldsymbol{n}')_j \tag{A7}$$

$$\sum_{\boldsymbol{n}} T^{\hat{U}}_{\boldsymbol{n}'\boldsymbol{n}} v(\boldsymbol{n}) = 0 , \tag{A8}$$

where $\bar{v}$ is any vector orthogonal to $\bar{0}$, $\bar{1}$, $\bar{2}$, and $\bar{3}$. Thus for any vector corresponding to a density operator, we have, using Eqs. (A4) and (A5),

$$\sum_{\boldsymbol{n}} T^{\hat{U}}_{\boldsymbol{n}'\boldsymbol{n}} w_{\hat{\rho}}(\boldsymbol{n}) = \frac{1}{\mathcal{N}}(1 + 3R^{-1}\boldsymbol{n}'\cdot\boldsymbol{S}) = \frac{1}{\mathcal{N}}(1 + 3\boldsymbol{n}'\cdot R\boldsymbol{S}) . \tag{A9}$$

Moreover, any vector not in the subspace spanned by the vectors corresponding to density operators is first projected into that subspace before undergoing a transformation like that in Eq. (A9). Notice that if $\hat{U} = \hat{1}$, the transformation matrix $T^{\hat{1}}_{\boldsymbol{n}'\boldsymbol{n}}$ is the projector onto this subspace.

Consider now N qubits. We begin by writing the operator $\hat{Q}(\tilde{\boldsymbol{n}})$ of Eq. (2.6) in the form

$$\hat{Q}(\tilde{\boldsymbol{n}}) = \left(\frac{3}{\mathcal{N}}\right)^N \sum_{\alpha_1,\ldots,\alpha_N} 3^{-N_{\alpha_1\ldots\alpha_N}} (n_1)_{\alpha_1}\cdots(n_N)_{\alpha_N}\hat{\sigma}_{\alpha_1}\otimes\cdots\otimes\hat{\sigma}_{\alpha_N} , \tag{A10}$$

where

$$N_{\alpha_1\ldots\alpha_N} \equiv \delta_{\alpha_1 0} + \cdots + \delta_{\alpha_N 0} \tag{A11}$$

is the number of zeroes in the list $\alpha_1\ldots\alpha_N$ (recall that each unit vector has an additional unit component in the $\boldsymbol{e}_0$-direction). Notice that from Eq. (2.8),

$$\sum_{\boldsymbol{n}} n_\alpha n_\beta = \frac{\mathcal{N}}{3}\, 3^{\delta_{\alpha 0}}\delta_{\alpha\beta} = \begin{cases} \mathcal{N} , & \alpha = \beta = 0, \\ \mathcal{N}/3 , & \alpha = \beta = 1,2,3, \\ 0 , & \alpha \neq \beta. \end{cases} \tag{A12}$$

An arbitrary vector $\bar{w}$ has $\mathcal{N}^N$ nonzero components. If the vector is associated with a density operator $\hat{\rho}$, using Eqs. (2.5) and (A10), the components are given by

$$\begin{aligned} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) &= \mathrm{tr}\left(\hat{\rho}\,\hat{Q}(\tilde{\boldsymbol{n}})\right) \\ &= \left(\frac{3}{\mathcal{N}}\right)^N \sum_{\alpha_1,\ldots,\alpha_N} 3^{-N_{\alpha_1\ldots\alpha_N}} (n_1)_{\alpha_1}\cdots(n_N)_{\alpha_N} \mathrm{tr}\left(\hat{\rho}\,\hat{\sigma}_{\alpha_1}\otimes\cdots\otimes\hat{\sigma}_{\alpha_N}\right). \end{aligned} \tag{A13}$$

Now define $4^{\mathcal{N}}$ vectors $\overline{\alpha_1\ldots\alpha_N}$ whose components are given by

$$\alpha_1\ldots\alpha_N(\tilde{\boldsymbol{n}}) = \left(\frac{3}{\mathcal{N}}\right)^{N/2} 3^{-N_{\alpha_1\ldots\alpha_N}/2} (n_1)_{\alpha_1}\cdots(n_N)_{\alpha_N}\,. \tag{A14}$$

This is the generalization of the vectors (A2). The conditions (2.7) and (2.8) imply that these vectors are orthonormal:

$$\begin{aligned} \sum_{\tilde{\boldsymbol{n}}} \alpha_1\ldots\alpha_N(\tilde{\boldsymbol{n}})\beta_1\ldots\beta_N(\tilde{\boldsymbol{n}}) &= \left(\frac{3}{\mathcal{N}}\right)^N 3^{-N_{\alpha_1\ldots\alpha_N}/2} 3^{-N_{\beta_1\ldots\beta_N}/2} \\ &\qquad \times \sum_{\tilde{\boldsymbol{n}}} (n_1)_{\alpha_1}\cdots(n_N)_{\alpha_N}(n_1)_{\beta_1}\cdots(n_N)_{\beta_N} \\ &= \left(\frac{3}{\mathcal{N}}\right)^N 3^{-N_{\alpha_1\ldots\alpha_N}/2} 3^{-N_{\beta_1\ldots\beta_N}/2} \\ &\qquad \times \underbrace{\sum_{\boldsymbol{n}_1}(n_1)_{\alpha_1}(n_1)_{\beta_1}\cdots\sum_{\boldsymbol{n}_N}(n_1)\alpha_1(n_N)\beta_N}_{\left(\frac{\mathcal{N}}{3}\right)^N 3^{N_{\alpha_1\ldots\alpha_N}}\delta_{\alpha_1\beta_1}\cdots\delta_{\alpha_N\beta_N}} \\ &= \delta_{\alpha_1\beta_1}\cdots\delta_{\alpha_N\beta_N}\,. \end{aligned} \tag{A15}$$

Now we can write any vector corresponding to a density operator as

$$\bar{w}_{\hat{\rho}} = \left(\frac{3}{\mathcal{N}}\right)^{N/2} \sum_{\alpha_1,\ldots,\alpha_N} 3^{-N_{\alpha_1\ldots\alpha_N}/2} \mathrm{tr}\left(\hat{\rho}\,\hat{\sigma}_{\alpha_1}\otimes\cdots\otimes\hat{\sigma}_{\alpha_N}\right)\overline{\alpha_1\ldots\alpha_N}\,, \tag{A16}$$

with its components given by

$$w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) = \left(\frac{3}{\mathcal{N}}\right)^{N/2} \sum_{\alpha_1,\ldots,\alpha_N} 3^{-N_{\alpha_1\ldots\alpha_N}/2} \mathrm{tr}\left(\hat{\rho}\,\hat{\sigma}_{\alpha_1}\otimes\cdots\otimes\hat{\sigma}_{\alpha_N}\right)\alpha_1\ldots\alpha_N(\tilde{\boldsymbol{n}})\,, \tag{A17}$$

which are the generalizations of Eqs. (A3) and (A4), respectively. Thus the $4^{\mathcal{N}}$-dimensional subspace spanned by the vectors corresponding to density operators is the same as the subspace spanned by the $4^{\mathcal{N}}$ vectors $\overline{\alpha_1\ldots\alpha_N}$.

The effect of a unitary transformation $\hat{U}$ on a tensor product of Pauli operators can be written as

$$\hat{U}^\dagger\hat{\sigma}_{\alpha_1}\otimes\cdots\otimes\hat{\sigma}_{\alpha_N}\hat{U} = \sum_{\beta_1,\ldots,\beta_N} R_{\alpha_1\ldots\alpha_N,\beta_1\ldots\beta_N}\hat{\sigma}_{\beta_1}\otimes\cdots\otimes\hat{\sigma}_{\beta_N}\,, \tag{A18}$$

where $R_{\alpha_1\ldots\alpha_N,\beta_1\ldots\beta_N}$ is a generalized rotation matrix for N qubits, which implies that

$$\hat{U}^\dagger \hat{Q}(\tilde{\boldsymbol{n}})\hat{U} \;=\; \left(\frac{3}{\mathcal{N}}\right)^N \sum_{\substack{\alpha_1,\ldots,\alpha_N\\ \beta_1,\ldots,\beta_N}} 3^{-N_{\alpha_1\ldots\alpha_N}} (n_1)_{\alpha_1}\cdots(n_N)_{\alpha_N} R_{\alpha_1\ldots\alpha_N,\beta_1\ldots\beta_N} \hat{\sigma}_{\beta_1}\otimes\cdots\otimes\hat{\sigma}_{\beta_N}\ . \tag{A19}$$

The transition matrix (2.12) now takes the form

$$\begin{aligned} T^{\hat{U}}_{\tilde{\boldsymbol{n}}'\tilde{\boldsymbol{n}}} \;&=\; \langle\tilde{\boldsymbol{n}}|\hat{U}^\dagger\hat{Q}(\tilde{\boldsymbol{n}}')\hat{U}|\tilde{\boldsymbol{n}}\rangle \\ &=\; \left(\frac{3}{\mathcal{N}}\right)^N \sum_{\substack{\alpha_1,\ldots,\alpha_N\\ \beta_1,\ldots,\beta_N}} 3^{-N_{\alpha_1\ldots\alpha_N}} (n'_1)_{\alpha_1}\cdots(n'_N)_{\alpha_N} R_{\alpha_1\ldots\alpha_N,\beta_1\ldots\beta_N} (n_1)_{\beta_1}\cdots(n_N)_{\beta_N} \\ &=\; \left(\frac{3}{\mathcal{N}}\right)^{N/2} \sum_{\substack{\alpha_1,\ldots,\alpha_N\\ \beta_1,\ldots,\beta_N}} 3^{-N_{\alpha_1\ldots\alpha_N}} 3^{N_{\beta_1\ldots\beta_N}/2} (n'_1)_{\alpha_1}\cdots(n'_N)_{\alpha_N} R_{\alpha_1\ldots\alpha_N,\beta_1\ldots\beta_N} \beta_1\ldots\beta_N(\tilde{\boldsymbol{n}})\ . \end{aligned} \tag{A20}$$

From this form of the transition matrix, it is clear that when acting on any vector that is orthogonal to the subspace spanned by vectors corresponding to density operators, the transition matrix gives zero. Thus the transition matrix projects into this subspace. When acting on a vector corresponding to a density operator, the transition matrix gives

$$\begin{aligned} w_{\hat{U}\hat{\rho}\hat{U}^\dagger}(\tilde{\boldsymbol{n}}') \;&=\; \sum_{\tilde{\boldsymbol{n}}} T^{\hat{U}}_{\tilde{\boldsymbol{n}}'\tilde{\boldsymbol{n}}} w_{\hat{\rho}}(\tilde{\boldsymbol{n}}) \\ &=\; \left(\frac{3}{\mathcal{N}}\right)^N \sum_{\substack{\alpha_1,\ldots,\alpha_N\\ \beta_1,\ldots,\beta_N}} 3^{-N_{\alpha_1\ldots\alpha_N}} (n'_1)_{\alpha_1}\cdots(n'_N)_{\alpha_N} R_{\alpha_1\ldots\alpha_N,\beta_1\ldots\beta_N} \mathrm{tr}\,(\hat{\rho}\,\hat{\sigma}_{\beta_1}\otimes\cdots\otimes\hat{\sigma}_{\beta_N}) \\ &=\; \left(\frac{3}{\mathcal{N}}\right)^N \sum_{\alpha_1,\ldots,\alpha_N} 3^{-N_{\alpha_1\ldots\alpha_N}} (n'_1)_{\alpha_1}\cdots(n'_N)_{\alpha_N} \mathrm{tr}\,(\hat{U}\hat{\rho}\hat{U}^\dagger\hat{\sigma}_{\alpha_1}\otimes\cdots\otimes\hat{\sigma}_{\alpha_N})\ . \end{aligned} \tag{A21}$$

This leads us back to a result that we could have written down right from the start, but what we have learned is that the transition matrix first projects into the subspace spanned by the vectors corresponding to density operators and then does just what it ought to do. If $\hat{U}=\hat{1}$, the transition matrix is just the projector onto the desired subspace.

---

[1] D. Deutsch, Proc. Roy. Soc. London Ser. A **400** (1985), 96–117.

[2] P. W. Shor, SIAM J. Computing **26** (1997), 1484–1509.

[3] R. L. Rivest, A. Shamir and L. Adleman, Comm. Assoc. Comput. Mach. **21** (1978), 120–126.

[4] A brief introduction to RSA encryption and related concepts may be found at `http://mathworld.wolfram.com/RSAEncryption.html` .

[5] Public-key encryption on most web browsers in the USA is 128-bit, but even lower strength—e.g. 40-bit—encryption gives adequate security for most applications.

[6] P. W. Shor, "Introduction to Quantum Algorithms," notes for talk given for the short course at the January 2000 American Math Society meeting, `arXiv.org e-print quant-ph/0005003`.

[7] Several tutorials on basic quantum computation, introductory cryptoanalysis, and factoring algorithms may be found at `http://www.qubit.org` .

[8] Since u has a residue that is a multiple of a, then $u = cN + \alpha a = (cb + \alpha)a$. Since a and b are prime and $\alpha < b$, then α can have no common factors with b other than 1. This means that the GCF of $(cb + \alpha)$ and N is either 1 or a, with a resulting from $c, \alpha \propto a$. This means that, definitely, the CGF of u and N is a. A similar analysis holds for v.

[9] If N has more than two prime factors, then at least one of the factors a and b is not prime. It is then possible that in taking the GCF of u and N, the GCF of α and b will be extracted instead of a (and similarly for v). Still, any factor of a or b is a factor of N by our definition, $N = ab$. The only difference this would make is that multiplying the two factors obtained would not produce N, but this is unimportant. The purpose was to factor N completely, and to do this, this process need only be repeated for one of the factors extracted and the new number obtained from dividing N by that factor.

[10] The CNOT (controlled not) gate reverses the second (target) qubit if the first (control) qubit is $|1\rangle$ and leaves it alone if the control qubit is $|0\rangle$. The Toffoli gate is similar to the controlled not, but it reverses the spin of the target qubit only if both of two control qubits are $|1\rangle$.

[11] A. Ekert, "Quantum Cryptoanalysis—Introduction," found on-line at `http://www.qubit.org/intros/cryptana.html` .

[12] S. Braun, H.-O. Kalinowski, S. Berger, *100 and More Basic NMR Experiments: A Practical Course* (VCH, New York, NY, 1996).

[13] M. A. Nielsen and I. L.Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[14] D. G. Cory, A. F. Fahmy, and T. F. Havel, Proc. Natl. Acad. Sci. U.S.A. **94**, 1634 (1997).

[15] N. A. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997).

[16] T. F. Havel, S. S. Somaroo, C.-H. Tseng, and D. G. Cory, Appl. Alg. Eng. Commun. Comp. **10**, 339 (2000).

[17] D. G. Cory *et al.*, Fortschr. Phys. **48**, 875 (2000).

[18] J. A. Jones, Fortschr. Phys. **48**, 909 (2000).

[19] J. A. Jones, Prog. Nucl. Mag. Res. Spec. **38**, 325 (2001).

[20] J. A. Jones, PhysChemComm **11**, 1 (2001).

[21] L. M. K. Vandersypen, C. S. Yannoni, and I. L. Chuang, to be published in *Encyclopedia of Nuclear Magnetic Resonance*, edited by D. M. Grant and R. K. Harris (Wiley, New York, 2001), `arXiv.org e-print quant-ph/0012108`.

[22] S. L. Braunstein *et al.*, Phys. Rev. Lett. **83**, 1054 (1999).

[23] R. Jozsa, *The Geometric Universe: Science, Geometry, and the Work of Roger Penrose*, edited by S. A. Huggett, L. J. Mason, K. P. Tod, S. T. Tsou, and N. M. J. Woodhouse (Oxford University, Oxford, England, 1998), p. 369.

[24] A. Ekert and R. Jozsa, Phil. Trans. Roy. Soc. London A **356**, 1769 (1998). The authors argue that entanglement is the "essential non-classical feature providing the computational speed-up in the known quantum algorithms."

[25] R. Laflamme, D. G. Cory, C. Negrevergne, and L. Viola, Quant. Inf. Comp., to be published, `arXiv.org e-print quant-ph/0110029`.

[26] D. Poulin, unpublished, `arXiv.org e-print quant-ph/0108102`.

[27] R. Schack and C. M. Caves, Phys. Rev. A **60**, 4354 (1999).

[28] A. J. Leggett and A. Garg, Phys. Rev. Lett. **54**, 857 (1985).

[29] J. P. Paz and G. Mahler, Phys. Rev. Lett. **71**, 3235 (1993).

[30] C. M. Caves, J. Superconductivity **12**, 707–718 (1999). Operations that involve a portion of the environment not modelled by $\hat{\rho}$ must obey the alternative condition $\sum_j \hat{A}_j^\dagger \hat{A}_j \leq 1$. The new density operator is then given by $\rho' = \sum_j \hat{A}_j \hat{\rho} \hat{A}_j^\dagger / \mathrm{tr}\,(\sum_j \hat{A}_j \hat{\rho} \hat{A}_j^\dagger)$.

[31] Throughout this document, we treat the Pauli operators as spin measurement operators themselves, ignoring the factor of $\hbar/2$ needed in front of them to properly call them angular momentum projectors. This is an overall factor and is not necessary to discuss the information content (up or down; 0 or 1) of a particular qubit.

[32] R. Schack and C. M. Caves, J. Mod. Opt. **47**, 387 (2000).

[33] Recall that it is always possible to write the density operator as a convex combination of a general basis set as in Eq. (1.16). What makes a state entangled is that in such a representation, this basis can never be a *tensor-product* basis. If we were to attempt expansion of an entangled state in any tensor-product basis, as in Eq. (1.27), the expansion coefficients would not all be nonnegative.

[34] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777–780 (1935).

[35] N. Bohr, Phys. Rev. **48**, 696 (1935).

[36] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambride, England, 1993), pp. 89–91.

[37] J. S. Bell, Physics **1**, 195 (1964).

[38] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, The Netherlands, 1993), Part II.

[39] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **26**, 880 (1969).

[40] E. P. Wigner, Am. J. Phys. **38**, 1005 (1970).

[41] The original EPR state was a correlated state of the positions and momenta of particles. The corresponding singlet argument was given by Bohm: D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1951). Sections 15–19 of chapter 22 (pp. 611–623) contain Bohm's spin version of the EPR paradox.

[42] N. C. Menicucci and C. M. Caves, Phys. Rev. Lett. **88**, 167901 (2002).

[43] D. G. Cory, M. D. Price, and T. F. Havel, Physica D **120**, 82 (1998); I. L. Chuang, N. Gershenfeld, M. G. Kubinec, and D. W. Leung, Proc. Roy. Soc. London A **454**, 447 (1998); E. Knill, I. Chuang, and R. Laflamme, Phys. Rev. A **57**, 3348 (1998).

[44] L. J. Schulman and U. V. Vazirani, in *Proceedings of the 31st Annual ACM Symposium on the Theory of Computation (STOC)* (ACM Press, 1999), p. 322; see also `arXiv.org e-print quant-ph/9804060`.

[45] W. S. Warren, Science **277**, 1688 (1997).

[46] W. Dür, J. I. Cirac, and R. Tarrach, Phys. Rev. Lett. **83**, 3562 (1999).

[47] E. Knill and R. Laflamme, Phys. Rev. Lett. **81**, 5672 (1998).

[48] D. Bohm, Phys. Rev. **85**, 166 and 180 (1952).

[49] A. Peres, Found. Phys. **29**, 589 (1999).