

## Macroscopic observables

David Poulin\*

*Perimeter Institute for Theoretical Physics, Waterloo, Ontario, N2L 2Y5, Canada;*  
*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1 Canada;*  
*and Department of Physics, University of Waterloo, Waterloo, Ontario N2L 3G1 Canada*  
 (Received 2 June 2004; published 7 February 2005)

We study macroscopic observables defined as the total value of a physical quantity over a collection of quantum systems. We show that previous results obtained for an *infinite* ensemble of identically prepared systems lead to incorrect conclusions for finite ensembles. In particular, exact measurement of a macroscopic observable significantly disturbs the state of any finite ensemble. However, we show how this disturbance can be made arbitrarily small when the measurements are of finite accuracy. We demonstrate a general trade-off between state disturbance and measurement coarseness as a function of the size of the ensemble. Using this trade-off, we show that the histories generated by any sequence of finite accuracy macroscopic measurements always generate a consistent family in the absence of large-scale entanglement for sufficiently large ensembles. Hence, macroscopic observables behave “classically” provided that their accuracy is coarser than the quantum correlation length scale of the system. The role of these observable is also discussed in the context of NMR quantum information processing and bulk ensemble quantum state tomography.

DOI: 10.1103/PhysRevA.71.022102

PACS number(s): 03.65.Ta, 03.67.–a

### I. OVERVIEW

Macroscopic observables correspond to physical quantities which are accessible to our senses. Since the physical scale of individual quanta is generally tiny, macroscopic observables arise when a collection of quantum systems are measured jointly. Formally, they can be described by *type projectors*, which reveal information about the average population of single-particle states. For example, the total magnetization of an ensemble of spin- $\frac{1}{2}$  particles provides some information about the relative occupation number of the *spin-up* and *spin-down* states. We will derive several general properties of these measurement and discuss how they lead to the emergence of a quasiclassical domain in the absence of large-scale entanglement.

The effect of macroscopic observations on *infinite* ensemble of identically prepared quantum systems has been studied in various contexts [1–4]. The main conclusion of these studies is that the state  $|\psi\rangle^{\otimes N}$  describing such an ensemble is an eigenstate of type projectors when  $N=\infty$ . However, for finite ensembles, things change dramatically. The measurement of a macroscopic observable induces a disturbance which *increases* as the size of the ensemble grows, in apparent contradiction with the infinite-copy result. This discrepancy follows from the ambiguous extension of finite-copy considerations to the nonseparable Hilbert space of an infinite-copy ensemble [5]. In this article, we show how the essence of the infinite-copy result can be recovered for finite ensembles by “smoothing” the type projectors into coarse-grained positive-operator-valued measurements (POVM’s) (essentially going from the strong to the weak law of large numbers).

The paper is organized as follows. The central mathematical objects of the present study are defined in Sec. II. We first

summarize the method of type and define type projectors. These are projectors on the degenerate eigensubspaces of macroscopic observables of the form  $A_N = \sum_{k=1}^N a_{(k)}$ , where  $a_{(k)}$  is a physical observable acting on the  $k$ th system of the ensemble. Using the theory of generalized measurements, we also define coarse-grained POVMs corresponding to finite accuracy estimation of a macroscopic observable.

Section III contains the core mathematical analysis of our study. We first recapitulate the well-known facts about type projectors acting on infinite ensembles and show how they dramatically break down for finite ensembles. Then, we show how the result is approximately recovered when the measurements are of finite accuracy and study the general trade-off between measurement coarseness and state disturbance—measured in terms of fidelity—as a function of the size of the ensemble. In short, we demonstrate that a measurement of coarseness  $\sigma \gg 1/\sqrt{N}$  leaves the systems essentially unchanged; i.e. the fidelity  $F$  between the pre- and post-measurement states of the ensemble satisfies  $1-F \propto \ln(N\sigma^2)/N\sigma^2$ .

Section IV is a discussion of the de Finetti representation theorem which provides a wide class of state—exchangeable states—for the study of macroscopic observables and is of prime importance to bridge the gap between the subjective nature of quantum states and the objective classical reality [7]. Exchangeable states have recently been employed for the discussion of quantum-state tomography based on single-system measurements followed by Bayesian update [6]. We will show how macroscopic observables offer an alternative perspective on quantum tomography. Moreover, this approach offers interesting applications for quantum information theory [8,9] and is a more accurate description of experimental spectroscopy-based implementations of tomography—e.g., as achieved in Ref. [10].

Macroscopic observables also provide an explanation for the emergence of the classical world we perceive from the

\*Electronic address: dpoulin@iqc.ca

underlying quantum theory. Indeed, we demonstrate in Sec. V that in the absence of large-scale entanglement, one of the main characteristics of the classical domain follows naturally from the general properties of coarse grained type POVM's: they behave *as if* they were revealing information about an underlying reality. This is done using the consistent histories formalism which we briefly summarize. We demonstrate that the histories generated by *any* sequence of macroscopic observables of accuracy  $\sigma \gg \sqrt{\xi/N}$  are consistent, where  $\xi$  is the quantum correlation length scale of the system. This generalizes some ideas introduced by Halliwell [11] on how to achieve classicality in closed quantum systems.

Finally, Sec. VI discusses the role of macroscopic observables in NMR quantum information processing. In this context, macroscopic observables are used to extract the output of the computation, but also, since the measurement device cannot be “turned off”—i.e., the state of the processor can always be read off from the spectrometer—they constantly perturb the computation. Following the results of Sec. III and a measurement model introduced in Ref. [12], we show that the measurements used in NMR can in principle be sufficiently precise to extract useful information about the computation but yet so coarse grained that they induce a negligible perturbation. However, as we will demonstrate, NMR measurements may not follow our optimal measurement coarseness-state disturbance trade-off when performed at room temperature; caution is advised when applying our conclusions. Finally, Sec. VII summarizes our results and discusses some open questions.

## II. DEFINITIONS

This section contains all the mathematical definitions required for our study. Our general setting consists of an ensemble of  $N$  quantum systems of the same nature. For sake of clarity, we adopt the vocabulary of NMR. Therefore, we shall refer to individual systems of an ensemble as *molecules* and to the ensemble of  $N$  molecules itself as the *sample*. Thus, the word “molecule” should not be taken literally in what follows; it could be any elementary constituent of a larger system.

### A. Method of types

The method of type is a very powerful statistical tool with applications ranging from large deviation theory, universal coding, and hypothesis testing. We will only scratch the surface of this theory here; more details and applications can be found in [13], for instance.

Let  $X = x_{j_1} x_{j_2} \cdots x_{j_N} \in \mathcal{X}^N$  be a string of  $N$  letters drawn from a  $d$ -letter alphabet  $\mathcal{X} = \{x_1, x_2, \dots, x_d\}$ . The *type* (or empirical probability distribution) of  $X$  is a vector of positive numbers summing to one defined by

$$\mathbf{L}(X) = (L_1(X), L_2(X), \dots, L_d(X)),$$

where  $L_j(X)$  is the relative frequency of the letter  $x_j$  in the string  $X$

$$L_j(X) = \frac{1}{N} \sum_{k=1}^N \delta_{j,j_k};$$

it is simply the number of occurrences of the letter  $x_j$  in  $X$ , divided by the length of  $X$ . For example, if  $\mathcal{X} = \{a, b, c\}$  and  $N=4$ , then  $\mathbf{L}(cbaa) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ . We also define a type class  $T$  to be the set of strings of a given type:

$$T[\mathbf{L}] = \{X \in \mathcal{X}^N : \mathbf{L}(X) = \mathbf{L}\}.$$

For example, using the same alphabet as above, we have  $T[(\frac{1}{4}, 0, \frac{3}{4})] = \{acc, cacc, ccac, ccca\}$ . The class  $T[\mathbf{L}]$  can be generated by applying all permutations to any single string of type  $\mathbf{L}$ . Hence, the number of elements in  $T[\mathbf{L}]$  is given by the multinomial coefficient:

$$|T[\mathbf{L}]| = \binom{N}{NL_1, NL_2, \dots, NL_d} = \frac{N!}{(NL_1)! (NL_2)! \dots (NL_d)!}.$$

Let  $\mathbf{R} = (R_1, R_2, \dots, R_d)$  be a probability distribution over  $\mathcal{X}$ . The probability of the string of outputs  $X = x_{j_1} \cdots x_{j_N}$  of  $N$  letters, each drawn independently according to the distribution  $\mathbf{R}$ , is  $P(X) = R_{j_1} R_{j_2} \cdots R_{j_N}$ . This can also be written as

$$P(X) = R_1^{NL_1(X)} R_2^{NL_2(X)} \dots R_d^{NL_d(X)},$$

so given a fixed distribution  $\mathbf{R}$ , the probability of a string  $X \in \mathcal{X}^N$  depends only on its type. Intuitively, the type of the observed outcome  $X$  is very likely to be close to the probability distribution of the random variable—i.e.,  $L_j(X) \approx R_j$ , as  $N$  increases. This is the substance of the *typical sequence theorem* [13],

$$P(\|\mathbf{L} - \mathbf{R}\|_1^2 > \epsilon) \leq e^{-N[\epsilon/2 - d \ln(N+1)/N]} \approx e^{-N\epsilon/2}, \quad (1)$$

where the “difference” between the type  $\mathbf{L}$  and the probability distribution  $\mathbf{R}$  is quantified by the variational distance ( $L_1$  norm):

$$\|\mathbf{L} - \mathbf{R}\|_1 = \sum_j |L_j(X) - R_j|.$$

The typical sequence theorem takes on various forms. It can be formulated in a stronger version using the relative entropy, which is an upper bound to the variational distance. Nevertheless, for our considerations, this simple version will be sufficient.

### B. Macroscopic observable

Using this notation, we now formally define macroscopic observables. Consider a Hermitian operator (i.e., *observable*)  $a$  acting on the  $d$ -dimensional Hilbert space of a single molecule  $\mathcal{H}_m$ . Let  $\{|x_1\rangle, |x_2\rangle, \dots, |x_d\rangle\}$  and  $\{\alpha_1, \alpha_2, \dots, \alpha_d\}$  denote its eigenvectors and eigenvalues:  $a|x_j\rangle = \alpha_j|x_j\rangle$ . We will assume that  $a$  is *nondegenerate*; generalization is straightforward. The macroscopic observable  $A_N$  corresponds to the sum of observable  $a$  over all the  $N$  molecules of the sample,

$$A_N = \sum_{k=1}^N a_{(k)}, \quad (2)$$

where  $a_{(k)}$  is the operator  $a$  acting on the  $k$ th molecule:

$$a_{(k)} = \underbrace{1 \otimes \dots \otimes 1}_{k-1} \otimes a \otimes \underbrace{1 \otimes \dots \otimes 1}_{N-k}.$$

The operator  $A_N$  acts on the joint Hilbert space of the  $N$  molecules  $\mathcal{H}_s = \mathcal{H}_m^{\otimes N}$ —the Hilbert space of the sample—which has dimension  $d^N$ . We use the standard abbreviation  $|X\rangle = |x_{j_1}\rangle \otimes |x_{j_2}\rangle \otimes \dots \otimes |x_{j_N}\rangle$  for each string  $X \in \mathcal{X}^N$ . Clearly, the states  $\{|X\rangle\}$  form an orthonormal basis for  $\mathcal{H}_s$ . Moreover, they are eigenstates of the macroscopic observable  $A_N$ :

$$\begin{aligned} A_N |X\rangle &= A_N |x_{j_1}\rangle \otimes |x_{j_2}\rangle \otimes \dots \otimes |x_{j_N}\rangle \\ &= \sum_{k=1}^N a_{(k)} |x_{j_1}\rangle \otimes |x_{j_2}\rangle \otimes \dots \otimes |x_{j_N}\rangle \\ &= \sum_{k=1}^N \alpha_{j_k} |x_{j_1}\rangle \otimes |x_{j_2}\rangle \otimes \dots \otimes |x_{j_N}\rangle = \sum_{k=1}^N \alpha_{j_k} |X\rangle \\ &= \left( \sum_{j=1}^d NL_j(X) \alpha_j \right) |X\rangle. \end{aligned} \quad (3)$$

Thus, we see that the eigenvalue associated to a basis state  $|X\rangle$  depends only on its type  $\mathbf{L}(X)$ . As a consequence, the degenerate eigensubspaces of  $A_N$  are those subspaces spanned by the vectors  $|X\rangle$  belonging to the same type class.

This brings us to the definition of *type measurements* which are von Neumann measurements composed of the projection operators on the subspaces of a given type:

$$Q_{\mathbf{L}}^{(N)} = \sum_{X \in \mathcal{T}[\mathbf{L}]} |X\rangle\langle X|. \quad (4)$$

Each of these *type projectors* is labeled by a vector of  $d$  positive numbers  $L_j$  which correspond to the type  $\mathbf{L}(X)$  of the basis vectors  $|X\rangle$  spanning the subspace. Obviously, the projectors  $Q_{\mathbf{L}}^{(N)}$  depend on the choice of basis  $|x_j\rangle$  over  $\mathcal{H}_m$ —i.e., on the eigenvectors of the observable  $a$ —so we could explicitly note  $Q_{\mathbf{L}}^{(N,a)}$ . Moreover, we would like to stress that the spectral projectors  $Q_{\mathbf{L}}^{(N,a)}$  and  $Q_{\mathbf{L}'}^{(N,b)}$  associated with two *distinct* macroscopic observables  $A_N = \sum_k a_{(k)}$  and  $B_N = \sum_k b_{(k)}$  do not commute, unless the underlying single-molecule observables  $a$  and  $b$  happen to commute. To avoid cumbersome notation, however, we will only use an extra superscript when necessary (cf. Sec. IV A). For the time being, we will consider a fixed arbitrary macroscopic observable  $A_N$ . In this case, it is straightforward to verify that the type projectors are mutually orthogonal and that they sum to the identity

$$Q_{\mathbf{L}}^{(N)} Q_{\mathbf{L}'}^{(N)} = \delta_{\mathbf{L},\mathbf{L}'} Q_{\mathbf{L}}^{(N)}, \quad \sum_{\mathbf{L}} Q_{\mathbf{L}}^{(N)} = 1. \quad (5)$$

In other words, these projectors correspond to the exact measurement of the population of the levels  $|x_j\rangle$  over an ensemble of  $N$  molecules, without distinguishing between the molecules of the sample. The type projectors  $Q_{\mathbf{L}}^{(N)}$  allows us to express the operator  $A_N$  in a simple form:

$$A_N = \sum_{\mathbf{L}} A_{\mathbf{L}} Q_{\mathbf{L}}^{(N)}, \quad (6)$$

where we have defined  $A_{\mathbf{L}} = \sum_{j=1}^d NL_j \alpha_j$ . This decomposition follows straightforwardly from Eq. (3) as all the states  $|X\rangle$  with  $\mathbf{L}(X) = \mathbf{L}$  composing the projector  $Q_{\mathbf{L}}^{(N)}$  have eigenvalue  $A_{\mathbf{L}}$ . Similarly, any macroscopic observable of the form of Eq. (2) has a spectral decomposition involving only type projectors, as in Eq. (6). Hence, following textbook quantum mechanics, when measuring a macroscopic observable—or measuring the “expectation value” of a physical observable over a macroscopic sample—one is really performing a projective von Neumann measurement composed of type projectors.

These type projectors have been studied under many different forms [1–3] and take on many different names. Among other formulations are the frequency operators. Recall that  $L_j(X)$  is the relative frequency of the symbol  $x_j$  in the string  $X$ . We can define a *frequency operator*

$$F_j^{(N)} = \sum_X L_j(X) |X\rangle\langle X|.$$

This operator is a macroscopic physical observable whose eigenvalues are  $f_j = 0, 1/N, 2/N, \dots, 1$ . Indeed,  $F_j^{(N)}$  takes on the form of Eq. (2) by setting the single-molecule observable  $a$  to  $(1/N)|x_j\rangle\langle x_j|$ . Following textbook quantum mechanics, when the measurement associated with  $F_j^{(N)}$  is performed and eigenvalue  $f_j$  is observed, the state of the system gets collapsed to the subspace spanned by the states  $|X\rangle$  for which  $L_j(X) = f_j$ . Hence, the eigenvalue  $f_j$  indicates the relative population of the single-molecule state  $|x_j\rangle$  in the sample of  $N$  molecules.

The above construction yields  $d$  commuting physical observables  $\{F_j^{(N)}\}_{j=1,\dots,d}$ , one for each single-molecule state  $\{|x_j\rangle\}_{j=1,\dots,d}$ . Regrouping these observable into a  $d$ -component observable yields

$$\mathbf{F}^{(N)} = (F_1^{(N)}, F_2^{(N)}, \dots, F_d^{(N)}) = \sum_{\mathbf{L}} \mathbf{L} Q_{\mathbf{L}}^{(N)}, \quad (7)$$

which takes on the form of Eq. (6), with a  $d$ -component eigenvalue  $A_{\mathbf{L}} = \mathbf{L}$ . The value of any macroscopic observable of the form of Eq. (6) can be deduced straightforwardly from the value of  $\mathbf{F}^{(N)}$ . Hence, a great deal of attention has been focused on the macroscopic observable  $\mathbf{F}^{(N)}$ , without loss of generality.

We illustrate macroscopic observables for a sample of  $N$  spin- $\frac{1}{2}$  particles. We choose the basis  $|x_1\rangle = |\uparrow\rangle$  and  $|x_2\rangle = |\downarrow\rangle$  corresponding, respectively, to  $+\frac{1}{2}$  and  $-\frac{1}{2}$  units of magnetization in the  $z$  direction:

$$\sigma^z |x_1\rangle = \frac{1}{2} |x_1\rangle \quad \text{and} \quad \sigma^z |x_2\rangle = -\frac{1}{2} |x_2\rangle.$$

We can use a single positive number  $L \in \{0, 1/N, 2/N, \dots, 1\}$  to label the type of a binary string  $X$ , which corresponds to the fraction of  $x_1$ 's (or *spin up*'s) in  $X$ . Hence, a type  $L$  is a shorthand for  $\mathbf{L} = (L, 1-L)$ . The bulk (or total) magnetization of the sample is equal to the sum of the magnetization of each molecules: the corresponding operator

is therefore  $M^z = \sum_{k=1}^N \sigma_{(k)}^z$ , where  $\sigma_{(k)}^z$  is the Pauli operator in the  $z$  direction acting on the  $k$ th molecule. When the sample is in a state of a definitive type  $L$ , its bulk magnetization is equal to  $\frac{1}{2}N(L_1 - L_2) = \frac{1}{2}N(2L - 1)$ , which is simply the number of spins pointing up minus the number of spins pointing down, times  $\frac{1}{2}$ . Hence, the observable corresponding to the bulk magnetization can be written as

$$M^z = N \frac{1}{2} \sum_L (2L - 1) Q_L^{(N)},$$

where the sum is over all types. The type projectors  $Q_L^{(N)}$  are projectors on the degenerate eigensubspaces of the bulk magnetization operator. Clearly, an exact measurement of the magnetization  $M^z$  would reveal the type of the state of the sample—i.e., the relative frequency of *up* and *down* spins.

### C. Coarse-grained macroscopic POVM's

We will now present how finite accuracy macroscopic observables can be expressed in terms of type projectors. Before we do so, we briefly recall some basic concepts of the theory of generalized measurements. Generalized measurements (POVM's) are described by a set positive operators  $E_j$  summing to identity. The generalized Born rule for the probability of getting outcome  $E_j$  given initial state  $\rho$  is the same as for von Neumann measurements:

$$P(E_j|\rho) = \text{Tr}\{E_j\rho\}. \quad (8)$$

After the measurement outcome  $E_j$  is observed, the state of the system gets updated to

$$\rho \rightarrow \rho_j = \frac{\sum_i A_{ji}^\dagger \rho A_{ji}}{P(E_j|\rho)}, \quad (9)$$

where the *Kraus operators*  $A_{ji}$  can be any set of operators satisfying  $\sum_i A_{ji}^\dagger A_{ji} = E_j$ . Here, we will often consider *ideal* quantum measurements where the disturbance inflicted to the system is in some sense minimal [14]. This restriction is necessary if we want to study the optimal trade-off between information gathering and state disturbance. To each measurement outcome  $E_j$  of an ideal measurement is associated a *single* Kraus operator  $A_{j0} = \sqrt{E_j}$ . In this case, the state update rule, Eq. (9), simplifies to

$$\rho \rightarrow \rho_j = \frac{\sqrt{E_j}^\dagger \rho \sqrt{E_j}}{P(E_j|\rho)}, \quad (10)$$

which reduced to the regular state update rule when  $E_j$  are projection operators. Hence, von Neumann measurements are minimally disturbing POVM's with an extra orthogonality constraint. Any generalized measurement can be realized physically by coupling the system of interest to a larger system and performing a von Neumann measurement on the larger system; an example of such a physical construction will be presented in Sec. VI. Similarly, any such “indirect” measurement corresponds to a POVM. Hence, POVM's do not add anything extra to plain textbook quantum mechanics, beside conciseness.

Continuing with our example, *finite accuracy* measurement of the bulk magnetization of a sample of  $N$  spin- $\frac{1}{2}$  molecules can be described in terms of *coarse-grained* type operators  $\tilde{Q}_\ell^{(N)}$ . When the state of the sample is of a definite type  $L$ , the *observed* value of the bulk magnetization will not necessarily be equal to  $\frac{1}{2}N(2L - 1)$ , but due to the uncertainty of the measurement apparatus, may take different values  $\frac{1}{2}N(2\ell - 1)$ , with respective probabilities  $q_L(\ell)$ . The function  $q_L(\ell)$  should be centered around  $L$  and have a certain width  $\sigma$  corresponding to the coarseness of the measurement.

Hence, the coarse-grained type measurements can be defined by “smoothing” the exact type projectors:

$$\tilde{Q}_\ell^{(N)} = \sum_L \sqrt{q_L(\ell)} Q_L^{(N)}, \quad (11)$$

where  $q_L(\ell)$  is some probability distribution over  $\ell$  centered roughly at  $L$  and has the interpretation given above. In principle,  $\ell$  could be any real  $d$ -dimensional vector, as it contains statistical fluctuations. For example,  $q_L(\ell)$  could be a  $d$ -dimensional Gaussian

$$q_L(\ell) = \left( \frac{1}{2\pi\sigma^2} \right)^{d/2} \exp \left\{ - \frac{\|\ell - \mathbf{L}\|_2^2}{2\sigma^2} \right\}, \quad (12)$$

which is properly normalized  $\int q_L(\ell) d\ell = 1$  and where the  $L_2$  norm is  $\|\ell - \mathbf{L}\|_2^2 = \sum_j (\ell_j - L_j)^2$ . The operators  $E_\ell = \tilde{Q}_\ell^{(N)} \tilde{Q}_\ell^{(N)\dagger}$  form a POVM (with a continuous outcome) since they are all positive operators and satisfy

$$\int E_\ell d\ell = \int \tilde{Q}_\ell^{(N)} \tilde{Q}_\ell^{(N)\dagger} d\ell = 1. \quad (13)$$

These coarse-grained type operators describe a situation where our measurement apparatus is not sufficiently precise to measure the exact population of each level, but rather provides an estimation of it within a finite accuracy  $\sigma$ .

We have assumed that the measurement outcome  $\ell$  takes on a continuous spectrum. However, several measurement apparatus, like those equipped with a numerical output display, have a discrete spectrum of outcomes. This can be taken into account by choosing a smoothing function

$$q_L(\ell) = \sum_{\ell_j} \delta(\ell - \ell_j) f_j(\mathbf{L}),$$

where  $\{\ell_j\}$  is the set of possible outcomes. Thus, we will henceforth consider the more general continuous case, but all our analysis carries through for discrete measurement outcomes by performing the above substitution.

### III. TYPE MEASUREMENT ON IDENTICALLY PREPARED SYSTEMS

Type projectors were first studied by Finkelstein [1], Hartle [2], and Graham [3] as part of discussions on the interpretation of probabilities in quantum theory. The main characteristic of type projectors identified by these authors can be summarized as follows. Let  $|\psi\rangle = \sum_j \beta_j |x_j\rangle$  be an arbitrary pure state of a  $d$ -level molecule, with associated density

matrix  $\nu = |\psi\rangle\langle\psi|$ . Consider a sample of  $N$  identically prepared molecules, such that the state of the sample is  $|\Psi_N\rangle = |\psi\rangle^{\otimes N}$ . Upon measurement of the type of the sample, we expect a result close to the probability distribution  $\mathbf{R} = (\langle x_1| \nu |x_1\rangle, \dots, \langle x_d| \nu |x_d\rangle) = (|\beta_1|^2, \dots, |\beta_d|^2)$ . Indeed, it follows from the strong law of large numbers that

$$\lim_{N \rightarrow \infty} |F_j^{(N)}|\Psi_N\rangle - |\beta_j|^2|\Psi_N\rangle|^2 = 0, \quad (14)$$

where  $F_j^{(N)}$  is the  $j$ th component of the frequency operator defined in Eq. (7). In other words,  $\mathbf{F}^{(N)}|\Psi_N\rangle = \mathbf{R}|\Psi_N\rangle$  with probability 1 in the limit of infinite  $N$ . This led Hartle to the conclusion that an infinite number of identically prepared molecules are in an eigenstate  $|\Psi_\infty\rangle$  of the frequency operator  $\mathbf{F}^{(\infty)}$  with eigenvalue  $\mathbf{R}$ . Finkelstein, on the other hand, concluded from Eq. (14) that for finite  $N$ ,  $|\Psi_N\rangle$  is “close” to an eigenstate of  $\mathbf{F}^{(N)}$  with eigenvalue  $\mathbf{R}$ . Thus, a measurement of the frequency operator reveals the probabilities  $R_j = \langle x_j| \nu |x_j\rangle$ , in the standard Copenhagen sense, of observing a single molecule of the sample in the state  $|x_j\rangle$ .

However, the conclusions reported above can be quite misleading. There are really two distinct issues here. The first one concerns the validity of the argument as a derivation of Born’s rule to assign probabilities in quantum theory. The main complication comes from the definition of  $\mathbf{F}^{(\infty)}$  as the limit of a finite operator. This limit does not uniquely define the operator on the *nonseparable* Hilbert space  $\mathcal{H}_m \otimes \mathcal{H}_m \otimes \dots$  of the infinite sample: specifying the action of  $\mathbf{F}^{(\infty)}$  on all states of the form  $|x_1\rangle \otimes |x_2\rangle \otimes \dots$  is not enough to define it. This was realized in [4] where an alternative derivation of the probability rule was presented. Nevertheless, the proposed solution is still not satisfactory as it relies itself on probability theory. An up-to-date and rather critical discussion of the status of the frequency operator and the related programs can be found in a recent paper of Caves and Schack [5]. We will not address these issues any further and do not claim to offer an alternative derivation of Born’s rule.

The second difficulty which is directly relevant to the present study concerns state disturbance. When a system is prepared in an eigenstate of a physical observable, the act of measurement does not disturb it. While Eq. (14) does not grant this for any finite  $N$ , one naturally expects (like Finkelstein did) that, as  $N$  grows, the disturbance caused by the measurement should decrease and eventually become negligible for all practical purpose.

In what follows, we will show that the *measurement of macroscopic observables induces an important disturbance to the state of the sample*. In fact, *this disturbance increases as the size  $N$  of the sample grows*. This is in apparent contradiction with the conclusion that one might intuitively draw from Eq. (14) by extending it to finite  $N$ . However, we will show how the above conclusion can be recovered when the measurement of macroscopic observables are of finite accuracy: *sufficiently coarse-grained type measurements induce a negligible disturbance to the state of the sample*. We are interested in the trade-off between measurement accuracy and state disturbance.

### A. State disturbance

The state of the sample can be rearranged as follows:

$$\begin{aligned} |\Psi_N\rangle &= \left( \sum_{j=1}^d \beta_j |x_j\rangle \right)^{\otimes N} = \sum_{X \in \mathcal{X}^N} \left( \prod_{j=1}^d \beta_j^{NL_j(X)} \right) |X\rangle \\ &= \sum_{\mathbf{L}} \left[ \prod_{j=1}^d \beta_j^{NL_j} \right] \sum_{X \in T[\mathbf{L}]} |X\rangle = \sum_{\mathbf{L}} \left( \prod_{j=1}^d \beta_j^{NL_j} \right) \sqrt{|T[\mathbf{L}]|} |\mathbf{L}\rangle, \end{aligned} \quad (15)$$

where we have defined the normalized state

$$|\mathbf{L}\rangle = \frac{1}{\sqrt{|T[\mathbf{L}]|}} \sum_{X \in T[\mathbf{L}]} |X\rangle$$

and

$$|T[\mathbf{L}]| = \binom{N}{NL_1, \dots, NL_d}$$

denotes the cardinality of the type class  $T[\mathbf{L}]$ . The density operator associated to this state will be denoted  $\rho_N = |\Psi_N\rangle\langle\Psi_N| = \nu^{\otimes N}$ .

Upon measurement of the coarse-grained operators of Eq. (11), the probability of observing an outcome within an infinitesimal volume range  $d\ell$  of  $\ell$  is  $P(\tilde{Q}_\ell^{(N)}|\rho_N)d\ell$  where [see Eq. (8)]

$$P(\tilde{Q}_\ell^{(N)}|\rho_N) = \text{Tr}\{\tilde{Q}_\ell^{(N)}\rho_N\tilde{Q}_\ell^{(N)}\} \quad (16)$$

$$\begin{aligned} &= \sum_{\mathbf{L}, \mathbf{L}'} \sqrt{q_{\mathbf{L}}(\ell)q_{\mathbf{L}'}(\ell)} \text{Tr}\{Q_{\mathbf{L}}^{(N)}\rho_N Q_{\mathbf{L}'}^{(N)}\} \\ &= \sum_{\mathbf{L}} q_{\mathbf{L}}(\ell) \langle\Psi_N|Q_{\mathbf{L}}^{(N)}|\Psi_N\rangle \\ &= \sum_{\mathbf{L}} q_{\mathbf{L}}(\ell) m(\mathbf{L}, \mathbf{R}) \end{aligned} \quad (17)$$

and  $m(\mathbf{L}, \mathbf{R})$  denotes the multinomial distribution

$$m(\mathbf{L}, \mathbf{R}) = \binom{N}{NL_1, \dots, NL_d} \prod_j R_j^{NL_j}.$$

Following Eq. (10), the *conditional post-measurement state* of the ensemble given measurement outcome  $\ell$  is

$$\begin{aligned} \rho_{N|\ell} &= \frac{\tilde{Q}_\ell^{(N)}\rho_N\tilde{Q}_\ell^{(N)}}{P(\tilde{Q}_\ell^{(N)}|\rho_N)} = \frac{1}{P(\tilde{Q}_\ell^{(N)}|\rho_N)} \sum_{\mathbf{L}} \sum_{\mathbf{L}'} \prod_{j,j'} \beta_j^{NL_j} \beta_{j'}^{*NL_{j'}} \\ &\quad \times \sqrt{q_{\mathbf{L}}(\ell)q_{\mathbf{L}'}(\ell)} \sqrt{|T[\mathbf{L}]| \cdot |T[\mathbf{L}']|} |\mathbf{L}\rangle\langle\mathbf{L}'|. \end{aligned} \quad (18)$$

The post-measurement state is obtained by averaging the conditional post-measurement states over all measurements outcomes:

$$\begin{aligned}\rho'_N &= \int P(\tilde{Q}_\ell^{(N)}|\rho_N)\rho_{N|\ell}d\ell = \int \tilde{Q}_\ell^{(N)}|\Psi\rangle\langle\Psi|\tilde{Q}_\ell^{(N)} \\ &= \sum_{\mathbf{L}} \sum_{\mathbf{L}'} \prod_{j,j'} \beta_j^{NL_j} \beta_{j'}^{*NL_{j'}} G(\mathbf{L},\mathbf{L}') \sqrt{|T[\mathbf{L}]| \cdot |T[\mathbf{L}']|} |\mathbf{L}\rangle\langle\mathbf{L}'|,\end{aligned}\quad (19)$$

where we have defined the *decoherence kernel*

$$G(\mathbf{L},\mathbf{L}') = \int \sqrt{q_{\mathbf{L}}(\ell)q_{\mathbf{L}'}(\ell)}d\ell. \quad (20)$$

Notice that setting  $G(\mathbf{L},\mathbf{L}')=1$  in Eq. (19) would yield a density matrix  $\rho'_N$  identical to  $\rho_N$ . Finally, the post-measurement state of a single molecule of the sample is obtained by taking a partial trace over  $N-1$  molecules  $\rho'_1 = \text{Tr}_{N-1}\{\rho'_N\}$  and similarly for the *conditional* post-measurement state  $\rho_{1|\ell} = \text{Tr}_{N-1}\{\rho_{N|\ell}\}$ .

The disturbance caused by the measurement is evaluated with the *fidelity* between the pre- and post-measurement states. A fidelity of 1 indicates that the two states are identical—i.e., the measurement did not cause disturbance—while a fidelity 0 indicates maximal disturbance. The fidelity between two states  $\rho$  and  $\nu$  is

$$F(\rho,\nu) = (\text{Tr}\{\sqrt{\rho^{1/2}\nu\rho^{1/2}}\})^2. \quad (21)$$

If one of the state is pure—say,  $\nu=|\phi\rangle\langle\phi|$ —this reduces to the familiar “overlap”  $F(\rho,|\phi\rangle\langle\phi|) = \langle\phi|\rho|\phi\rangle$ .

It is instructive to first consider the case where the measurement are perfectly accurate,  $\sigma=0$  in Eq. (12), which implies  $q_{\mathbf{L}}(\ell) = \delta(\ell-\mathbf{L})$  and  $G(\mathbf{L},\mathbf{L}') = \delta_{\mathbf{L}\mathbf{L}'}$ . In this case, the post-measurement density matrix is

$$\rho'_N = \sum_{\mathbf{L}} m(\mathbf{L},\mathbf{R})|\mathbf{L}\rangle\langle\mathbf{L}|, \quad (22)$$

so it has completely decohered in the type basis  $|\mathbf{L}\rangle$ ; i.e., there are no off-diagonal terms of the form  $|\mathbf{L}\rangle\langle\mathbf{L}'|$  like in Eq. (19). The fidelity between the pre- and post-measurement states is then

$$\begin{aligned}F_{\sigma=0}(\rho_N,\rho'_N) &= \sum_{\mathbf{L}} [m(\mathbf{L},\mathbf{R})]^2 \\ &\leq \underbrace{\sum_{\mathbf{L}} m(\mathbf{L},\mathbf{R})}_{1} \times (\max_{\mathbf{L}} m(\mathbf{L},\mathbf{R})), \\ &\approx \frac{1}{(2\pi N)^{d-1/2} \prod_j |\beta_j|},\end{aligned}\quad (23)$$

where the subscript  $\sigma=0$  indicates that the measurement are perfectly accurate, and we have used Stirling’s approximation in the last line. Clearly, *exact type measurements greatly disturb the system*, since fidelity goes to zero as the size of the sample increases, except in the case where  $\beta_j = \delta_{j0}$ . A similar conclusion based on different considerations was reached by Squires [15]. It follows from the concavity of fidelity  $F(\rho, \sum_j p_j \nu_j) \geq \sum_j p_j F(\rho, \nu_j)$  that the conditional post-

measurement state  $\rho_{N|\ell}$  also has, with high probability, a vanishing fidelity with the original state  $\rho_N$ .

The disturbance caused by an exact type measurement is most obvious when considering the conditional post-measurement state of a *single* molecule from the sample. As shown in Appendix A,

$$\rho_{1|\ell} = \sum_{j=1}^d \ell_j |x_j\rangle\langle x_j|, \quad (24)$$

the conditional post-measurement state of a single molecule is diagonal in the  $|x_j\rangle$  basis with eigenvalues given by the *observed* type of the sample  $\ell$ , independently of its state  $\nu$  prior to the measurement. In words, an ensemble measurement of the relative occupation number of the single-molecule states  $\{|x_j\rangle\}$  with outcome  $\ell = (\ell_1, \ell_2, \dots, \ell_d)$  “resets” the reduced state of a single molecule to a statistical mixture of the states  $\{|x_j\rangle\}$  with associated probabilities  $\ell_j$ . However, following the typical sequence theorem, Eq. (1), the observed coefficients  $\ell_j$  are very likely to be close to  $R_j = \langle x_j | \nu | x_j \rangle$ . When averaging over measurement outcomes, we recover the state  $\nu \rightarrow \rho'_1 = \sum_j R_j |x_j\rangle\langle x_j|$  which has no off-diagonal terms—i.e.,  $|x_i\rangle\langle x_j|$ . Thus, the exact measurement of a macroscopic observable *completely decoheres individual molecules of the sample*; it leaves the diagonal elements of  $\nu$  unchanged while suppressing all off-diagonal terms. (This situation might appear worrisome for bulk-ensemble quantum computing; we will return to this in Sec. VI). Moreover, the measurement creates correlation between the molecules, so  $\rho'_N \neq (\rho'_1)^{\otimes N}$  and  $\rho_{N|\ell} \neq (\rho_{1|\ell})^{\otimes N}$  in general. The conditional post-measurement state  $\rho_{N|\ell}$  can even have entanglement across the different molecules of the sample.

## B. Gaussian smoothing

We now turn our attention to the case where the smoothing function  $q_{\mathbf{L}}(\ell)$  has a finite width  $\sigma$ . In the case of interest, the initial state of the sample  $|\Psi_N\rangle$  is pure, so combining Eqs. (15) and (19) we get

$$\begin{aligned}F(\rho_N,\rho'_N) &= \langle\Psi_N|\rho'_N|\Psi_N\rangle \\ &= \sum_{\mathbf{L},\mathbf{L}'} m(\mathbf{L},|\beta_j|^2)m(\mathbf{L}',|\beta_j|^2)G(\mathbf{L},\mathbf{L}').\end{aligned}\quad (25)$$

For the sake of clarity, we will first consider the Gaussian distribution  $q_{\mathbf{L}}(\ell)$  defined in Eq. (12). The decoherence kernel defined in Eq. (20) is then given by

$$\begin{aligned}G(\mathbf{L},\mathbf{L}') &= \int \left(\frac{1}{2\pi\sigma^2}\right)^{d/2} e^{-\|\ell-\mathbf{L}\|_2^2 + \|\ell-\mathbf{L}'\|_2^2 / 4\sigma^2} d\ell \\ &= \exp\left\{-\frac{\|\mathbf{L}-\mathbf{L}'\|_2^2}{2(2\sigma)^2}\right\}.\end{aligned}$$

This is not surprising as the decoherence kernel is the convolution of the smoothing function with itself. The convolution of two distribution of width  $\sigma_1$  and  $\sigma_2$  gives a distribution of width  $\sigma' = \sigma_1 + \sigma_2$ , so  $G(\mathbf{L},\mathbf{L}')$  is a function of width  $2\sigma$ .

We can find a lower bound to the fidelity by truncating the sum in Eq. (25). By restricting  $\mathbf{L}$  and  $\mathbf{L}'$  to the domain  $D$

$=\{\mathbf{L}:\|\mathbf{L}-\mathbf{R}\|_2\leq\Delta\}$  where  $R_j=\langle x_j|v|x_j\rangle$ , we can lower bound the kernel by  $G(\mathbf{L},\mathbf{L}')\geq\exp\{-\Delta^2/2\sigma^2\}$  using the triangle inequality. This yields the inequality

$$F(\rho_N,\rho'_N)\geq\exp\left\{-\frac{\Delta^2}{2\sigma^2}\right\}\left(\sum_{L\in\mathcal{D}}b(L)\right)^2.$$

The quantity in the parentheses is a sum over the range  $\mathcal{D}$  of a multinomial probability distribution. It is equal to  $P(\|\mathbf{L}-\mathbf{R}\|_2\leq\Delta)\geq P(\|\mathbf{L}-\mathbf{R}\|_1\leq d\Delta)\geq(1-e^{-Nd\Delta/2})$  by the Cauchy-Schwartz inequality and the typical sequence theorem, Eq. (1). Thus, we get

$$F(\rho_N,\rho'_N)\geq\exp\left\{-\frac{\Delta^2}{2\sigma^2}\right\}(1-e^{-Nd\Delta/2})^2. \quad (26)$$

Since this bound holds for all  $\Delta$  (which is an arbitrary cut-off), we can maximize the right-hand side of Eq. (26)—the optimal value turn out to be attained when  $\Delta^2=2\ln(1+2N\sigma^2d)/Nd$ —to get the tightest bound:

$$F(\rho_N,\rho'_N)\geq 1 - \frac{1 + \ln(2N\sigma^2d)}{N\sigma^2d}. \quad (27)$$

Hence, as the size of the sample increases, the measurement accuracy  $\sigma$  can decrease as fast as  $1/\sqrt{N}$  while maintaining a constant fidelity  $F(\rho_N,\rho'_N)=1-\epsilon$  between the pre- and post-measurement states. If  $\sigma$  decreases less rapidly than  $1/\sqrt{N}$ —e.g.,  $N^{-s}$  for  $0<s<1/2$ —the fidelity will go to 1 as  $N$  grows. In particular, if  $\sigma$  is constant,  $F(\rho_N,\rho'_N)\sim 1-\epsilon\ln N/N$ .

The fidelity between the pre- and *conditional* post-measurement states—i.e.,  $\rho_N$  and  $\rho_{N|\ell}$  respectively—can be computed using similar techniques. The computation is illustrated in Appendix B. While the mathematical details are slightly more involved, the essence of the result is identical: as long as  $\sigma\gg 1/\sqrt{N}$ —i.e.,  $\sigma=N^{-s}$  for  $0\leq s<1/2$ —the fidelity  $F(\rho_N,\rho_{N|\ell})$  goes to 1 as  $N$  increases. Moreover, this is true *independently of the measurement outcome*  $\tilde{Q}_\ell^{(N)}$ , in contrast with the result obtained in Ref. [9], where due to a sharp smoothing function a low fidelity was obtained with some small probability.

### C. General smoothing

We now wish to argue that the essence of our measurement accuracy-state disturbance trade-off applies to arbitrary smoothing function  $q_{\mathbf{L}}(\ell)$  introduced in Eq. (11), provided that it is actually smooth with respect to  $\mathbf{L}$ . Let us be more precise. Intrinsic to the smoothing function is a notion of distance on the  $d$ -dimensional probability simplex. One can define various distance measures on this space; e.g., our choice of smoothing function, Eq. (12), in the previous section relied on the distance  $\|\mathbf{L}-\mathbf{L}'\|_2$  induced by the  $L_2$  norm. The exact statement of the trade-off will obviously depend on the choice of distance measure. However, the essence of the result is independent of this choice, as all good distance measures are equivalent on small distances. Thus, a good smoothing function  $q_{\mathbf{L}}(\ell)$  should satisfy

$$|q_{\mathbf{L}}(\ell)-q_{\mathbf{L}'}(\ell)|\leq c\left(\frac{\|\mathbf{L}-\mathbf{L}'\|_1}{\sigma}\right)^s \quad (28)$$

for sufficiently small  $\|\mathbf{L}-\mathbf{L}'\|_1$  and some positive constants  $c$  and  $s$  [Eq. (28) is known as the Lipschitz condition]. In general,  $c$  depends on the dimension  $d$  of the molecules. Therefore, the dependence of the bound, Eq. (27), on the dimension  $d$  (which may seem awkward) only reflects our choice of the  $L_2$  norm in the smoothing function; it is not universal. Given this assumption, we can derive the general result. It should be mentioned that, ultimately,  $q_{\mathbf{L}}(\ell)$  depends on the details of the measurement procedure of the corresponding macroscopic observable (see for example the model of Sec. VI). However, if this measurement is of finite accuracy, then the smoothing function must have a certain width and should satisfy the above assumption.

We see from Eq. (25) that fidelity between the pre- and post-measurement states only depends on the decoherence kernel  $G(\mathbf{L},\mathbf{L}')=\int\sqrt{q_{\mathbf{L}}(\ell)}q_{\mathbf{L}'}(\ell)d\ell$ . Thus, the procedure used in the previous section carries through straightforwardly. We can truncate the sum, Eq. (25), to the domain  $\mathcal{D}$  where  $\|\mathbf{L}-\mathbf{R}\|_1\leq\Delta$ , with  $R_j=|\beta_j|^2$ . On this domain, the fluctuations of the kernel are bounded by Eq. (28) using the triangle inequality. Moreover, as  $G(\mathbf{L},\mathbf{L})=1$  by the normalization condition of the smoothing function, we obtain

$$G(\mathbf{L},\mathbf{L}')\geq 1 - c(\Delta/\sigma)^s \text{ on } \mathcal{D}.$$

The bound

$$F(\rho_N,\rho'_N)\geq\left\{1-c\left(\frac{\Delta}{2\sigma}\right)^s\right\}(1-e^{-N\Delta^2/2})^2 \quad (29)$$

follows straightforwardly from the typical sequence theorem, Eq. (1). Given the value of  $c$  and  $s$ , one can perform an optimization with respect to  $\Delta$  to get the tightest bound. However, this depends on the details of the smoothing function.

Similarly, we can derive a bound for the fidelity of the conditional post-measurement state when  $\sigma\gg 1/\sqrt{N}$ . The technique illustrated in Appendix B for a specific choice of smoothing function can indeed be applied straightforwardly to any type of smoothing function satisfying Eq. (28). The basic steps, starting from the expression

$$F(\rho_N,\rho_{N|\ell})=\frac{\left[\sum_{\mathbf{L}}\sqrt{q_{\mathbf{L}}(\ell)}m(\mathbf{L},\mathbf{R})\right]^2}{\sum_{\mathbf{L}}q_{\mathbf{L}}(\ell)m(\mathbf{L},\mathbf{R})},$$

are the following. The sum in the numerator is truncated to the domain  $\mathcal{D}$ . Using the typical sequence theorem and Eq. (28), we can bound the nominator of this expression by

$$[q_{\mathbf{R}}(\ell)-c(\Delta/\sigma)^s](1-e^{-N\Delta^2/2})^2.$$

The denominator is bounded by Eq. (28):

$$\sum_{\mathbf{L}}q_{\mathbf{L}}(\ell)m(\mathbf{L},\mathbf{R})\leq q_{\mathbf{R}}(\ell)+c(\Delta/\sigma)^s.$$

Combining these two inequalities yields, to first order,

$$F(\rho_N, \rho_{N|\ell}) \geq 1 - 2c(\Delta/\sigma)^s - 2e^{-N\Delta^2/2}.$$

Again, the free parameter  $\Delta$  should be optimized to find the tightest bound; the details of the trade-off depend on the form of the smoothing function.

Finally, the scaling  $\sigma \sim 1/\sqrt{N}$  is optimal. A higher precision would considerably disturb the state of the system. This is because the multinomial distribution  $m(\mathbf{L}, \mathbf{R})$  has a width  $1/\sqrt{N}$ . Consider the expression of Eq. (25). If the kernel has a width smaller than the binomial distribution, the sum, and hence the fidelity  $F(\rho_N, \rho_{N'})$ , will be roughly equal to  $\text{erf}(\sigma\sqrt{N}) \approx 2\sigma/\sqrt{N\pi}$  for  $\sigma \ll 1/\sqrt{N}$ . The bound is also tight for the conditional post-measurement fidelity  $F(\rho_N, \rho_{N|\ell})$  as fidelity is a convex function. This can also be seen intuitively by considering the behavior of two consecutive measurements. Upon fine-grained measurement  $Q_L^{(N)}$ , the variance of the outcome  $\mathbf{L}$  is  $1/\sqrt{N}$ . However, if we first perform a coarse-grained measurement  $\tilde{Q}_\ell^{(N)}$  of width  $\sigma \ll 1/\sqrt{N}$  and then perform a fine-grained measurement on the updated state  $\rho_{N|\ell}$ , the variance of the second measurement outcome will be  $\sigma$ : performing the coarse-grained measurement has altered its statistics. This means that the coarse-grained measurement has appreciably disturbed the state of the sample, so  $F(\rho_N, \rho_{N|\ell})$  is far from 1.

#### D. Mixed states

The results established in the present section hold unchanged when the molecules of the sample are all prepared in the same *mixed* state  $\nu = \sum_{i=1}^d \lambda_i |\psi_i\rangle\langle\psi_i|$ . The argument proceeds in three steps. First, we can construct a purification of the state  $\nu$ ,

$$|\phi\rangle = \sum_{i=1}^d \sqrt{\lambda_i} |\psi_i\rangle |i\rangle,$$

by appending to each molecule an ancillary system of dimension  $d$  with orthonormal basis  $\{|i\rangle\}$ . Clearly, the reduced state of the molecule—obtained by tracing out the ancilla—is  $\text{Tr}_{\text{ancilla}}\{|\phi\rangle\langle\phi|\} = \nu$ . Second, the vectors  $\{|x_j\rangle|i\rangle\}_{i,j=1,\dots,d}$  form a basis for the Hilbert space of the pair molecule+ancilla. The type projectors  $Q_L^{(N)}$  associated with the molecule only measure the type of the prefix  $x_j$ , so are a coarse-grained version of the type projectors associated with the pair: the disturbance they cause to the state of the sample can only be less than the disturbance caused by the complete type projectors. Thus, the bound, Eq. (26), can be applied to  $F(\Phi_N, \Phi'_N)$  where  $\Phi_N = |\phi\rangle\langle\phi|^{\otimes N}$  and

$$\Phi'_N = \int (\tilde{Q}_\ell^{(N)} \otimes 1) \Phi_N (\tilde{Q}_\ell^{(N)} \otimes 1) d\ell.$$

Finally, by monotonicity of the fidelity— $F(\mathcal{E}(\rho), \mathcal{E}(\nu)) \geq F(\rho, \nu)$  for any trace preserving quantum operation  $\mathcal{E}$ —the bound applies directly to the pre- and post-measurement states of the sample of molecules by tracing out the ancillas. By similar considerations, all of the above conclusions can be extended to mixed states.

#### IV. EXCHANGEABILITY

Before proceeding with the applications of the above results, we present an important result that will—under certain assumptions—lead to an objective perception of the state of macroscopic systems. The concept of *exchangeability* was introduced in the classical theory of probability by de Finetti [16] to substitute the incorrect use of “unknown probabilities.” A probability assignment is the expression of one’s subjective knowledge about the possible outcomes of an experiment. Hence, it is not a property of a physical system itself but, rather, a property of the agent assigning the probability, so it cannot be unknown to him.

There are also several good reasons to believe that quantum states are subjective; see, for example, [14,18–20] and references therein. The state of a quantum system is a mathematical construct which allows one to compute probabilities for various measurements outcomes.<sup>1</sup> As a consequence of the subjective nature of quantum states, the concept of an *unknown quantum state* is in general an oxymoron, for essentially the same reasons which lead to this conclusion for classical probability assignment.

However, unknown quantum states turn out to be quite useful for the description of certain physical settings. Of particular interest to us is the description of a sample of  $N$  “molecules.” Under certain circumstances—e.g., thermal equilibrium—one can arrive at the conclusion that all the molecules of the sample are equivalent, so they should all be describe by the same state  $\nu$ , which is itself unknown. This is a very common state of affairs in nuclear, atomic, or molecular physics where spectral quantities—which are formally described by macroscopic observables—are measured over a large collection of quantum systems. In fact, in almost all physical experiments where ensemble measurements are performed, the components of the sample are assumed to be in the “same unknown state” and the purpose of the measurement is to (partially) determine this state. Moreover, note that macroscopic observables do not allow one to discriminate between the molecules: all molecules of the sample are treated on equal footing. Hence, information gathered by macroscopic measurements naturally leads to a state assignment where all molecules are in the same, but perhaps partially unknown, state.

To arrive at an appropriate description of the sample without referring to the unknown quantum state of individual molecules, we must clearly state the assumption of the agent assigning the state. His assumption is that the *arbitrary number of molecules are all equivalent*, which can be formalized by demanding that the state  $\rho_N$  assigned to the sample by the observer satisfies the following requirements.

(i) For any permutation  $\pi$  of  $N$  molecules,  $\pi[\rho_N] = \rho_N$ . Such a state is called symmetric.

(ii) For any positive integer  $M$ , there exists a symmetric state  $\rho_{N+M}$  such that  $\rho_N = \text{Tr}_M\{\rho_{N+M}\}$ , where  $\text{Tr}_M$  denotes the partial trace over  $M$  molecules.

A state  $\rho_N$  satisfying these two conditions is called *ex-*

<sup>1</sup>To quote Robert Griffiths, “If probabilities are not real, then pre-probabilities [quantum states] are even less real” [17].



*changeable*. The quantum de Finetti representation theorem [6,21,22] asserts that any exchangeable quantum state  $\rho_N$  of a sample of  $N$  molecules can be written as

$$\rho_N = \int \nu^{\otimes N} \text{Pr}(\nu) d\nu, \quad (30)$$

where  $\nu$  are density operators of a single molecule and  $\text{Pr}(\nu)$  is a probability distribution over the quantum states of a single molecule.

The interpretation of this theorem is that it is *mathematically correct* to look upon  $\nu$  as an objective element of reality about which we have incomplete knowledge: hence we assign it some probability distribution  $\text{Pr}(\nu)$ . For example, when the POVM  $\{E_j\}$  is measured on the sample, the outcome  $E_j$  is observed with probability

$$\begin{aligned} P(E_j|\rho_N) &= \text{Tr}\{E_j\rho_N\} \\ &= \int \text{Tr}\{E_j\nu^{\otimes N}\}\text{Pr}(\nu)d\nu = \int P(E_j|\nu^{\otimes N})\text{Pr}(\nu)d\nu. \end{aligned} \quad (31)$$

We can think of  $P(E_j|\nu^{\otimes N})$  as the probability of  $E_j$  given a value of the *real parameter*  $\nu$ , but since  $\nu$  is unknown, we average this probability over the possible values of  $\nu$  distributed according to  $\text{Pr}(\nu)$ . However, it must be emphasized that it is the assumption of exchangeability which leads to the form of Eq. (30), which in turn legitimizes the term “unknown state” for mathematical convenience.

### A. Bulk tomography

Quantum-state tomography is an experimental procedure which transforms an exchangeable state of the form of Eq. (30) into a product state  $\rho = \nu^{\otimes N}$  through repeated state updates Eq. (9). According to the de Finetti representation theorem, we can equivalently say—and this is how tomography is conventionally formulated—that the purpose of tomography is to determine which is the *real yet unknown* state  $\nu$  describing the  $N$  identical molecules of the sample.

In Ref. [6], quantum-state tomography was studied in the context where the molecules of the sample are measured individually and the state of the *rest* of the sample was updated. Here, we present how quantum-state tomography can be performed through bulk measurements. A similar description was recently and independently developed in [9]. Let  $A_N = \sum_k a(k)$  be a macroscopic observable deriving from the single-molecule observable  $a$  as in Eq. (2). Exceptionally, we denote the eigenstates and eigenvalues of  $a$  with a superscript  $a|x_j^{(a)}\rangle = \alpha_j^{(a)}|x_j^{(a)}\rangle$  for later convenience. The finite accuracy measurement of the macroscopic observable  $A_N$  is defined through the POVM  $\{\tilde{Q}_\ell^{(N,a)}\}$ .

The conditional state of the sample after the measurement of  $\{\tilde{Q}_\ell^{(N,a)}\}$  with outcome  $\ell^{(a)}$  is

$$\rho_{N|\ell^{(a)}} = \int (\tilde{Q}_\ell^{(N,a)} \nu^{\otimes N} \tilde{Q}_\ell^{(N,a)}) \frac{\text{Pr}(\nu)}{P(\tilde{Q}_\ell^{(N,a)}|\rho_N)} d\nu.$$

The quantity in parentheses is proportional to the conditional post-measurement state of the sample, given that it was ini-

tially in state  $\nu^{\otimes N}$ . As demonstrated in Sec. III, this measurement has very high fidelity with the original state. Mathematically, this means

$$\tilde{Q}_\ell^{(N,a)} \nu^{\otimes N} \tilde{Q}_\ell^{(N,a)} \approx P(\tilde{Q}_\ell^{(N,a)}|\nu^{\otimes N}) \nu^{\otimes N}.$$

Therefore, we get

$$\rho_{N|\ell^{(a)}} \approx \int \nu^{\otimes N} \frac{\text{Pr}(\nu) P(\tilde{Q}_\ell^{(N,a)}|\nu^{\otimes N})}{P(\tilde{Q}_\ell^{(N,a)}|\rho_N)} d\nu. \quad (32)$$

Comparing this updated state with the initial state of the sample, Eq. (30), it is *as if* we had updated the probability distribution  $\text{Pr}(\nu)$  of the real yet unknown state  $\nu$  according to Bayes’ rule  $P(y|x) = P(x|y)P(y)/P(x)$ . However, this is strictly a mathematical identity; all we did was to apply the state update rule, Eq. (9), to an exchangeable state. This is of capital importance. Given the exchangeability assumption and given that the observer can only gather information through the measurement of macroscopic quantities, the quantum state update of Eq. (9) behaves just like the classical Bayes update rule. For these observers, the sample behaves *as if* the reduced state of a single molecule  $\nu$  was an objective element of reality. Hence, the usual experimentalist’s objection to the subjective nature of quantum states, “Of course quantum states are real, I measure them in my lab!” is mathematically justified by his limitations to measure macroscopic observables on exchangeable states (or more generally on states with no entanglement on macroscopic scales as we shall soon establish).

We can repeat the procedure with different macroscopic observables  $B_N, C_N, \dots$  derived from the single-molecule observables  $b, c, \dots$ , which do not necessarily commute with each other. If the sets of observables are sufficiently informative—i.e., if their eigenstates  $\{|x_j^{(\mu)}\rangle\}_{\mu=a,b,c,\dots}$  contain  $d^2 - 1$  linearly independent elements—the updated probability distribution will converge with very high probability to a  $\delta$  function for sufficiently large  $N$ ,  $\text{Pr}(\nu|\ell^{(a)}, \ell^{(b)}, \dots) \approx \delta(\hat{\nu})$ , so the final state will be  $\rho_{N|\ell^{(a)}, \ell^{(b)}, \dots} \approx \hat{\nu}^{\otimes N}$ . This is because the functions  $P(\tilde{Q}_\ell^{(N,\mu)}|\nu^{\otimes N})$  are centered around  $\ell^{(\mu)} = (\langle x_1^{(\mu)}|\nu|x_1^{(\mu)}\rangle, \langle x_2^{(\mu)}|\nu|x_2^{(\mu)}\rangle, \dots)$  and have a width  $\sigma$ . The state  $\hat{\nu}$  is the only one satisfying all the linear constraints  $\langle x_j^{(\mu)}|\hat{\nu}|x_j^{(\mu)}\rangle = \ell_j^{(\mu)}$  for all  $\mu = a, b, c, \dots$  up to accuracy  $\sigma$ . Again, this is *as if* the measurements simply inform us of the identity of the “real but unknown”  $\hat{\nu}^{\otimes N}$ , without disturbing it in the limit of large  $N$ .

### V. CLASSICALITY

We have seen at the end of the last section that—under the exchangeability assumption—macroscopic observables behave as classical. Our goal here is to formalize as well as extend this result. We will demonstrate that macroscopic observables define a classical limit for closed quantum systems; i.e., they do not require interaction with any “environment.” Observations play very different roles in classical and quantum theory. In the classical setting, we can think of measurements as unveiling an underlying “real” state of affairs: ob-

servations reveal information about the state of the world without affecting it. On the other hand, quantum measurements disturb or “collapse” the state of the system.

When states, either quantum or classical, are regarded as subjective judgments of the world, both of the above descriptions need revision. Let  $P(x_i, y_j)$  be the joint probability distribution, or state, that the agent assigns to the classical sequence of events,  $X=x_i$  and  $Y=y_j$ . Upon the observation  $X=x_i$ , the agent updates her predictions for event  $Y$  according to Bayes’ rule

$$P(y_j|x_i) = \frac{P(x_i, y_j)}{P(x_i)},$$

where  $P(x_i) = \sum_j P(x_i, y_j)$ . This state generally differs from the pre-measurement state assigned to  $Y$ :

$$P(y_j|x_i) \neq P(y_j) = \sum_i P(x_i, y_j).$$

Hence, the act of observing  $X$  modifies the state assigned to  $Y$ . However, disregarding the observed value of  $X$  for later probability assignments is like not measuring the value of  $X$  at all:

$$P(y_j) = \sum_i P(y_j|x_i)P(x_i). \quad (33)$$

Indeed, we can interpret the observation as revealing the “real” value of  $X$  which was there all along: the agent simply did not know about it prior to her observation. In this sense,  $X=x_i$  is a real state of affairs about which the agent learns through the act of measurement. Thus, the state she assigns to  $Y$  prior to her observation of  $X$  is the mixture of the state it would have given the different value of  $X$ , weighted by the probability of  $X$ ; cf. Eq. (33). This reasoning extends in an obvious way to any sequence of events  $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ . We can consider that the system follows a fixed history  $x_{j_1}^{(1)}, x_{j_2}^{(2)}, \dots, x_{j_n}^{(n)}$  of which the agent has incomplete knowledge, resulting in a joint distribution  $P(x_{j_1}^{(1)}, x_{j_2}^{(2)}, \dots, x_{j_n}^{(n)})$ .

Quantum measurements behave quite differently. A quantum event corresponds to a “click” on a measurement apparatus at some instant of time  $t$ . Hence, each event is associated a POVM element  $E_{j_k}^{(k)}(t_k)$  in the Heisenberg picture at a given time  $t_k$ . (We will henceforth drop the explicit time label  $t_k$ .) In general, assigning definite yet unknown outcomes to these events leads to incorrect predictions; e.g., the sum rule

$$P(E_{j_2}^{(2)}) = \sum_{j_1} P(E_{j_2}^{(2)}|E_{j_1}^{(1)})P(E_{j_1}^{(1)})$$

does not hold in general. This is most obvious in Young’s double-slit experiment where the events  $E_{j_1}^{(1)}$  correspond to the particle going through slit  $j_1=1$  or  $2$  and  $j_2$  label the various positions on the detector. Reasonings involving the particle going through a definite yet unknown slit lead to incorrect predictions.

There are, however, sequences of quantum events which do behave classically, as if the observations were revealing an underlying reality, the typical example being when all the

POVM elements describing the events commute. The consistent histories approach to quantum theory [23] lays down a set of conditions under which such behavior occurs. A complete list of alternative events  $\zeta^{(k)} = \{E_{j_k}^{(k)}\}$  at time  $t_k$  defines a POVM. A history is a list of POVM elements  $H = (E_{j_1}^{(1)}, E_{j_2}^{(2)}, \dots, E_{j_n}^{(n)})$  at distinct times  $t_1, t_2, \dots, t_n$ . When the initial state of the system is  $\rho$ , the probability of an history  $H$  is <sup>2</sup>

$$P(H|\rho) = \text{Tr}\{(E_{j_n}^{(n)})^{1/2} \dots (E_{j_1}^{(1)})^{1/2} \rho (E_{j_1}^{(1)})^{1/2} \dots (E_{j_n}^{(n)})^{1/2}\} \quad (34)$$

following Eqs. (8) and (10). A complete family of histories is the set of all combination of POVM elements from the sets  $\zeta^{(k)}$  at all times,  $\mathcal{F} = \{\zeta^{(1)}, t_1; \zeta^{(2)}, t_2; \dots; \zeta^{(n)}, t_n\}$ . A complete family of histories is thus a sample space on which a probability distribution  $P(H|\rho)$  is defined. The family is said to be *consistent* when the sum rule approximately holds for  $P(H|\rho)$ . This conditions is the simplest version of all consistency conditions but will be sufficient for our purposes. In this sense, consistent histories define a quasiclassical domain of familiar experience.

As was observed by Halliwell [11], histories corresponding to a sequence of finite accuracy measurement of macroscopic observables generate a consistent family if the system is a sufficiently large sample of identically prepared molecules—i.e., when  $\rho = \nu^{\otimes N}$ . It should be stressed that the single-molecule observables  $a, b, c, \dots$  making up the histories *do not need to commute*. For example, the coarse measurement of the magnetization of a sample of spin  $\frac{1}{2}$  along the  $z$  axis followed by a measurement along the  $y$  axis can generate a consistent family if the sample is sufficiently large.

A simple argument to build our intuition in this direction is to consider the commutator of any two *normalized* macroscopic observables. Let  $a$  and  $b$  be two arbitrary single-molecule observables and define  $c$  to be their commutator  $c = [a, b]$ . These operator can be suitably normalized so that they satisfy  $\|a\|, \|b\|, \|c\| \approx 1$ . The normalized macroscopic observable  $A_N$  is defined as  $A_N = (1/N) \sum_{k=1}^N a_{(k)}$  and similarly for  $B_N$  and  $C_N$ ; hence,  $\|A_N\|, \|B_N\|, \|C_N\| \approx 1$ . A straightforward calculation shows that the commutator of the normalized macroscopic observables obeys

$$[A_N, B_N] = \frac{1}{N} C_N, \quad (35)$$

which implies  $\|[A_N, B_N]\| \approx 1/N$ . Thus, all macroscopic observables commute in the limit of infinite-size sample and commuting observables systematically generate consistent histories: measuring the value of one observable does not affect the outcome statistics of other commuting observables.

However, the infinite-sample considerations cannot be applied straightforwardly to finite ensembles (this is the recurrent theme of this paper). In particular, Eq. (35) does not involve any coarseness, which is essential to achieve consis-

<sup>2</sup>We assume for simplicity that the POVM are *ideal*; see Sec. II B.

tency in finite ensembles. Our analysis of Sec. III provides the right tools to formally study the consistency of macroscopic histories operating on finite ensembles. Indeed, Halliwell's result is a straightforward consequence of the fact that these measurements leave the state  $\nu^{\otimes N}$  of the ensemble essentially unchanged, so they do not alter the statistics of subsequent measurement outcomes. Moreover, our general analysis will allow us to extend the conclusions reached by Halliwell to a much wider set of initial states.

The first generalization is straightforward: by linearity of Eq. (34), such families are automatically consistent for initial *exchangeable* states. Indeed, for any initial exchangeable state  $\rho_N$  of the form Eq. (30), the probability of history  $H$  reads

$$P(H|\rho_N) = \int P(H|\nu^{\otimes N})\Pr(\nu)d\nu. \quad (36)$$

Clearly, if the sum rule is satisfied for the  $P(H|\nu^{\otimes N})$  individually, it is also satisfied for their convex combination. This is very much in the spirit of the de Finetti representation theorem as one can interpret the outcome of the macroscopic measurements as revealing partial information about the *real* quantum state  $\nu^{\otimes N}$  of the sample, of which we have incomplete knowledge. The consistency of such ‘‘macroscopic histories’’ for initial exchangeable state can also be seen as a consequence of the fact that in those conditions the quantum state update behaves approximately like a Bayesian update; cf. Eq. (32). Indeed, the identity  $P(x) = \sum_y P(x|y)P(y)$  of classical probability theory (or more precisely its continuous version) applied to Eq. (32) implies that the average post-measurement state of the sample is approximately equal to the initial state,  $\rho'_N = \int \rho_N|\ell^{(a)}\rangle P(\tilde{Q}_\ell^{(N,a)}|\rho_N)d\ell^{(a)} \approx \rho_N$ , for any initial exchangeable state  $\rho_N$ .

Moreover, consider an arbitrary product state of the sample  $\rho = \nu_1 \otimes \nu_2 \otimes \dots \otimes \nu_N$ . We will construct a state  $\bar{\nu}^{\otimes N}$  whose measurement outcomes, for coarse-grained macroscopic observables, are statistically indistinguishable from those obtained from the product state  $\rho$ . This will prove that Halliwell's result applies to arbitrary product states as well. Consider the symmetrized version of  $\rho$ :

$$\Pi[\rho] = \frac{1}{N!} \sum_{\pi} \nu_{\pi(1)} \otimes \nu_{\pi(2)} \otimes \dots \otimes \nu_{\pi(N)},$$

where the sum is over all permutations of  $N$  elements. The reduced state of a single molecule is

$$\text{Tr}_{N-1}\{\Pi[\rho]\} = \frac{1}{N} \sum_{k=1}^N \nu_k = \bar{\nu}. \quad (37)$$

The states  $\Pi[\rho]$  and  $\bar{\nu}^{\otimes N}$  are in some sense very similar: they are both symmetric, yield the same reduced single-molecule state  $\bar{\nu}$ , and yield the same expectation value of the frequency operator  $\langle \mathbf{F}^{(N)} \rangle = (\langle x_1 | \bar{\nu} | x_1 \rangle, \langle x_1 | \bar{\nu} | x_1 \rangle, \dots)$ . However, they are not identical. To illustrate this, consider a sample of  $N$  two-dimensional molecule in the state

$$\rho = \underbrace{|x_1\rangle\langle x_1| \otimes \dots \otimes |x_1\rangle\langle x_1|}_{N/2} \otimes \underbrace{|x_2\rangle\langle x_2| \otimes \dots \otimes |x_2\rangle\langle x_2|}_{N/2}:$$

the first half of the molecules are in state  $|x_1\rangle$  while the second half are in state  $|x_2\rangle$ . The measurement of the frequency operator  $\mathbf{F}^{(N)}$  of Eq. (7) yields the outcome  $(\frac{1}{2}, \frac{1}{2})$  with certainty when the state of the system is  $\Pi[\rho]$ . The average single-molecule state is  $\bar{\nu} = \frac{1}{2}|x_1\rangle\langle x_1| + \frac{1}{2}|x_2\rangle\langle x_2|$ , so  $\bar{\nu}^{\otimes N}$  also yields an average result  $(\frac{1}{2}, \frac{1}{2})$  of the frequency operator. But as opposed to the state  $\Pi[\rho]$ , the outcome of the measurement of  $\mathbf{F}^{(N)}$  can fluctuate away from  $(\frac{1}{2}, \frac{1}{2})$  when the sample is in state  $\bar{\nu}^{\otimes N}$ . However, according to the typical sequence theorem, Eq. (1), the size of these fluctuations will be of order  $1/\sqrt{N}$  and so can only be perceived by macroscopic measurements of accuracy  $\sigma \lesssim 1/\sqrt{N}$ . For macroscopic observables of coarseness  $\sigma \gg 1/\sqrt{N}$ , the two states  $\Pi[\rho]$  and  $\bar{\nu}^{\otimes N}$  will yield the same statistics up to order  $\sigma\sqrt{N} \ll 1$ .

Moreover, the states  $\rho$  and  $\Pi[\rho]$  yield exactly the same statistics for measurement outcomes of macroscopic observables: this follows straightforwardly from the permutation invariance of the type projectors, Eq. (4). We have thus established the chain of equality,

$$P(\tilde{Q}_\ell^{(N)}|\rho) = P(\tilde{Q}_\ell^{(N)}|\Pi[\rho]) \approx P(\tilde{Q}_\ell^{(N)}|\bar{\nu}^{\otimes N}),$$

so the states  $\rho$  and  $\bar{\nu}^{\otimes N}$  yield almost identical predictions when  $\sigma \gg 1/\sqrt{N}$ . It follows that a sequence of finite accuracy macroscopic measurements performed on a state of the form  $\rho = \nu_1 \otimes \nu_2 \otimes \dots \otimes \nu_N$  generates a consistent family of histories for sufficiently large samples. This simple result can be summarized as follows: when the sample is assigned a state  $\rho = \nu_1 \otimes \nu_2 \otimes \dots \otimes \nu_N$ , coarse-grained macroscopic measurements behave—to a high accuracy—as if the average state of the molecules  $\bar{\nu} = (1/N)\sum_j \nu_j$  was an objective element of physical reality being discovered.

In fact, *any separable state leads to consistency of macroscopic histories*. Indeed, when the molecules of the sample are *not entangled* with each other, their state can be written as

$$\rho = \int \nu_1 \otimes \nu_2 \otimes \dots \otimes \nu_N \Pr(\nu_1, \nu_2, \dots, \nu_N) d\nu_1 d\nu_2 \dots d\nu_N. \quad (38)$$

[This is the definition of entanglement: a state is said to be entangled if it cannot be written as Eq. (38).] Now, consider the state

$$\bar{\rho} = \int \bar{\nu}^{\otimes N} \Pr(\bar{\nu}) d\bar{\nu}, \quad (39)$$

where we have defined

$$\Pr(\bar{\nu}) = \int \Pr(\nu_1, \dots, \nu_N) \delta\left(\bar{\nu} - \frac{1}{N} \sum_k \nu_k\right) d\nu_1 \dots d\nu_N.$$

By linearity and using the result established above, we have  $P(H|\rho) \approx P(H|\bar{\rho})$  provided that the histories  $H$  are generated by macroscopic observables of accuracy  $\sigma \gg 1/\sqrt{N}$ . Since  $\bar{\rho}$  is exchangeable, the probabilities  $P(H|\bar{\rho})$ , and therefore the probabilities  $P(H|\rho)$ , approximately satisfy the sum rule. Again, this has a simple interpretation. The average state of

the sample  $\bar{\nu}$  behaves like an objective element of reality [7]. When the molecules of the sample are correlated with each other, there is no well-defined average state. The observer therefore assigns a probability distribution  $P(\bar{\nu})$  over the possible values of this “objective element of reality”  $\bar{\nu}$ , hence Eq. (39).

To put it simply, when macroscopic measurements are coarse with respect to the quantum correlation length scale of the system, they behave classically. Indeed, assume that a sample of  $N$  molecules has quantum correlation length  $\xi$ ; i.e., there is  $\xi$ -molecule entanglement in the system. Then, all of the above construction can be applied to the  $N/\xi$  collections of  $\xi$  molecules. We simply have to treat each block of entangled  $\xi$  molecules as one big molecule. There is no entanglement between these big molecules, so the previous analysis applies, as long as the measurement accuracy is larger than  $\sqrt{\xi/N}$ . Thus, we see that only entanglement on “macroscopic” scales can cause quantum effects to the measurement of coarse-grained macroscopic observable.

We have demonstrated in this section that a classical limit can be obtained for closed quantum systems under certain assumptions about the system’s initial state. This approach is complementary to the *decoherence* program [24], where classicality arises from the interaction between the system of interest and its environment. Moreover, the combination of these two approaches extends the conclusions reached in the present section. When the sample interacts with an environment, it will typically end up in a state that behaves classically under macroscopic measurements. For example, when the effect of this interaction is to dephase or depolarize the molecules of the sample, decoherence will destroy entanglement on macroscopic scales [25], so coarse-grained “macroscopic histories” will be consistent. An other possibility is that the environment interacts with the system through a “collective coupling”; i.e., the system observables appearing in the coupling Hamiltonian are of the form of Eq. (2). This type of interaction will be studied in the next section. Under such a coupling, the environment is effectively measuring the macroscopic observables appearing in the coupling Hamiltonian. Thus, after the interaction, the system is insensitive to further measurement of the same macroscopic observables. For example, an environment consisting of charged particles interacts with the total magnetization of a ferromagnet, effectively measuring the average value of the Pauli operators  $\langle\sigma_x\rangle$ ,  $\langle\sigma_y\rangle$ , and  $\langle\sigma_z\rangle$  to finite accuracy. Therefore, after decoherence the ferromagnet will not be further disturbed by such measurements.

## VI. NMR INFORMATION PROCESSING

Room-temperature nuclear magnetic resonance (NMR) has been for several years a benchmark for quantum information processing [26]. The sample contains  $N \approx 10^{20}$  molecules which are to good approximation noninteracting due to dynamical decoupling caused by thermal excitations. Hence, the total Hamiltonian is the sum of the single-molecule Hamiltonians  $H = \sum_k \hbar h_{(k)}$ ; it takes the form of Eq. (2), so it is a macroscopic observable. Initially, the sample is in a thermal state

$$\rho_N = \frac{e^{-\beta H}}{Z} = \frac{e^{-\sum_k \hbar h_{(k)}}}{Z} = \left( \frac{e^{-\beta h}}{z} \right)^{\otimes N},$$

where  $Z = \text{Tr}\{e^{-\beta H}\}$  and  $z = \text{Tr}\{e^{-\beta h}\}$  are the partition functions of the sample and of a single molecule, respectively. Each molecule contains a certain number of nuclei which carry a spin, and it is these spin degrees of freedom which are used to perform the computation. Due to their different chemical environment, the various spins of a molecule can have different Larmor frequencies  $\omega_j$ , which makes it possible to address them individually. To do so, the sample is placed in a coiled wire through which a sequences of externally controlled radio frequency (rf) current pulses can be applied. By properly tuning the frequency of the rf pulse, we can address all the spins with the same Larmor frequency, so all the  $N$  molecules are addressed in parallel. Therefore, a sequence of pulses transforms the state of the sample according to

$$\rho_N \rightarrow U^{\otimes N} \rho_N U^{\dagger \otimes N},$$

where  $U$  is a unitary matrix acting on the Hilbert space of a single molecule. This transformation preserves the tensor product structure of the density matrix  $\rho_N = \nu^{\otimes N}$ ; it collectively changes the state of individual molecules  $\nu$ .<sup>3</sup>

It has been known for a long time [27] that the coupling between the nuclear spins and the coil can considerably disturb the state of the sample in certain regimes through *back-action*. This noise is *not* fundamentally irreversible; it is only due to our neglecting of high-order terms in the coupling Hamiltonian. However, since the coil is also used to *read out* the state of the sample, it must unavoidably induce extra *fundamentally irreversible* noise, of the kind discussed in Sec. III. Indeed, in quantum theory, any measurement that reveals some information must unavoidably perturb the state of the system [28]. This result is puzzling because the coil is present throughout the computation, not only during the measurement phase, so should in principle disturb the computation.

A simple model to study the effect of this noise was presented in Ref. [12]. The current in the coil can be modeled by a continuous quantum variable  $\Phi_j = \int \phi_j | \phi_j \rangle \langle \phi_j | dr_j$  where  $j$  labels the modes of the field in the coil. Each field mode  $\Phi_j$  couples to the resonant magnetization of the sample—i.e., to the spins of Larmor frequency  $\omega_j$ —through its conjugate momentum  $\Pi_j$ ,  $[\Pi_j, \Phi_j] = i$  ( $\Pi$  is the “generator of translations” for  $\Phi$ ). The coupling Hamiltonian takes the form  $H_c = \gamma \sum_j \Pi_j M_j^x$  where  $\gamma$  is some coupling constant (that absorbs the magnetization units  $\frac{1}{2} \hbar g$ ) and

<sup>3</sup>When the sequence of pulses generates a complex transformation, it is practically impossible to keep track of  $\nu$ , as this would require an exponential amount of computation. Hence, given our limited computational capacities, the sample should really be described by an exchangeable state of the form, Eq. (30). Indeed, if we had sufficient computational power to have complete knowledge of the state  $\nu^{\otimes N}$  after the pulse sequences, it would mean that the quantum computation was useless since we are able to predict its outcome.

$$M_j^x = \sum_{k=1}^N \sigma_{j(k)}^x = \frac{1}{2} N \sum_L (2L-1) Q_L^{(N)} \quad (40)$$

is the total transverse magnetization of the  $j$ th specie of nuclei in their rotating frame of Larmor frequency  $\omega_j$ . Assume for simplicity that each molecule contain a single spin- $\frac{1}{2}$  nucleus which couples to the field mode  $\Phi_0$  (we will thus drop the explicit 0 subscript). This field mode is initially in state  $|\Psi\rangle = \int \Psi(\phi) |\phi\rangle d\phi$  and the sample is in state  $\rho_N$ . After a time  $t$ , simple calculations show that the joint state of the field and the sample is

$$\rho(t) = \sum_{L,L'} \int d\phi d\phi' \Psi(\phi) \Psi^*(\phi') \\ \times Q_L^{(N)} \rho_N Q_{L'}^{(N)} \otimes |\phi + f(L)\rangle \langle \phi' + f(L')|,$$

where  $f(L) = \gamma t N (2L-1)$ . The field mode  $\Phi_0$  is subsequently observed to be in state  $|\phi\rangle$ , and accordingly the state of the sample is updated to

$$\rho_{N \rightarrow \rho_{N|\phi}} = \frac{\text{Tr}_{\text{coil}}\{(|\phi\rangle\langle\phi| \otimes \mathbb{1})\rho(t)\}}{P(\phi)} = \frac{\tilde{Q}_\phi^{(N)} \rho_N \tilde{Q}_\phi^{(N)\dagger}}{P(\phi)},$$

where  $\tilde{Q}_\phi^{(N)} = \sum_{L'} \Psi(\phi - f(L')) Q_{L'}^{(N)}$  are coarse-grained type measurements like those of Eq. (11). The initial field configuration  $\Psi(\phi)$  plays the role of the smoothing function and has width (in the  $L$  domain)  $\sigma \approx \sqrt{\langle \Psi | \Phi_0^2 | \Psi \rangle - \langle \Psi | \Phi_0 | \Psi \rangle^2} / N \gamma t$ .

This model may appear overcomplicated, but it is in fact quite simple. The field variable  $\Phi_0$ —which we can think of as the amplitude of the  $\omega_0$  Fourier component of the current in the coil—serves as a measurement apparatus: it is by reading the value  $\phi$  of  $\Phi_0$  that we learn about the magnetization of the spins with Larmor frequency  $\omega_0$ . The coupling Hamiltonian is such that it “shifts” the value of  $\Phi_0$  by an amount that is proportional to the value of  $M_0^x$  (the rotating magnetization induces a current in the coil). Thus, determining the amount by which  $\Phi_0$  got shifted allows us to infer the value of  $M_0^x$ . However, the coil may initially *not* be in an eigenstate of  $\Phi_0$ , so  $\langle \Psi | \Phi_0^2 | \Psi \rangle - \langle \Psi | \Phi_0 | \Psi \rangle^2 > 0$  (it will typically be in a coherent state). Thus, our final measure of the value of  $\Phi_0$  does not allow us to determine exactly by what amount it got shifted: it can only do so up to accuracy  $\sqrt{\langle \Psi | \Phi_0^2 | \Psi \rangle - \langle \Psi | \Phi_0 | \Psi \rangle^2}$ , which sets the accuracy  $\sigma$  on the measurement of  $M_0^x$ . Following the results established in Sec. III C, a width  $\sigma \gg 1/\sqrt{N} \approx 10^{-10}$  ensures us that the measurement does not significantly perturb the computation.

Of course, the measurements achieved in the laboratory are much coarser than  $10^{-10}$ . Given the results presented in this paper, we could follow [12] and conclude that the presence of the coil (or the NMR measurements in general) induces a negligible disturbance to the state of the sample. However, our analysis does not apply here straightforwardly since NMR measurements are not *ideal* (see Sec. II B). This is because the coil is not in a pure state at room temperature. As a consequence, not only is the coil not in an eigenstate of  $\Phi_0$  (such as a coherent state), but it is in a statistical mixture

of such states. Consider, for example, the initial state of the coil in a Gaussian *mixture* of Gaussian-like field modes  $\rho_{\text{coil}} \propto \int e^{-q^2/2\lambda^2} |\Psi_q\rangle \langle \Psi_q| dq$  where  $|\Psi_q\rangle \propto \int e^{-(\phi-q)^2/4\sigma^2} |\phi\rangle d\phi$  (the  $\Psi_q$  are like coherent states centered around the field value  $q$ ). After a coupling time  $t$ , the observation of the field mode in state  $|\phi\rangle$  updates the state of the sample to

$$\rho_{N \rightarrow \rho_{N|\phi}} = \frac{\text{Tr}_{\text{coil}}\{(|\phi\rangle\langle\phi| \otimes \mathbb{1})\rho(t)\}}{P(\phi)} \\ = \int e^{-q^2/2\lambda^2} \tilde{Q}_{\phi-q}^{(N)} \rho_N \tilde{Q}_{\phi-q}^{(N)} dq,$$

where  $\tilde{Q}_\phi^{(N)} \propto \sum_{L'} e^{-[\phi - f(L')]^2/4\sigma^2} Q_{L'}^{(N)}$ . This is the continuous version of the general-state update rule, Eq. (9), for nonideal measurements. The corresponding POVM elements  $E_\phi = \int e^{-q^2/2\lambda^2} [\tilde{Q}_{\phi-q}^{(N)}]^2$  have width  $\lambda + \sigma$ , which determines the accuracy of the measurement outcomes following Eq. (8). However, the Kraus operators  $A_{qr} = \tilde{Q}_{\phi-q}^{(N)}$  have width  $\sigma$ . Following Eq. (9), it is this width which governs the disturbance caused to the state. Thus, it is not the measurement coarseness  $\lambda + \sigma$  which ultimately determines the disturbance caused to the state, but the details of the measurement process. In this example, the statistical mixture caused by the finite temperature of the coil added an extra source of uncertainty, characterized by the parameter  $\lambda$ .

It is therefore necessary to have a detailed model of the interaction between the coil and sample and of the initial state of the coil to evaluate its contribution to decoherence of the state of the molecules. We suspect that, in actual NMR settings, the measurement coarseness is largely due to statistical (thermal) fluctuations of the type of  $\lambda$ . However, we also suspect the coherent spread of the coil’s wave function  $\sigma$  to be much larger than  $10^{-10}$ , since coherent manipulation of the molecules appears to be possible despite the coupling to the coil. These questions, however, deserve a separate study.

## VII. CONCLUSION

We have demonstrated a trade-off between macroscopic measurement accuracy and state disturbance for sample of identically prepared quantum systems. A measurement coarseness smaller than  $1/\sqrt{N}$  causes a disturbance to the state of the system which increases as the size of the ensemble grows, which is in apparent contradiction with the infinite-copy result. However, a measurement coarseness  $\sigma \gg 1/\sqrt{N}$  induces a negligible disturbance to the state of the sample. This demonstrates that coarseness is an essential feature of the macroscopic limit. The type projectors of Eq. (4) will unavoidably lead to macroscopic quantum fluctuations regardless of the size of the ensemble. Thus, the classical observables encountered in our everyday life (e.g., position and velocity of a baseball, bulk magnetization of a ferromagnet, etc.) *cannot* be suitably described in terms of von Neumann macroscopic type projectors Eq. (4)—the spectral projectors of average observables Eq. (2)—such as prescribed by textbook quantum theory, but require coarse-grained

POVM's, Eq. (11). Moreover, as illustrated with the measurement model of Sec. VI, realistic measurements settings give rise to coarse grained POVM's, not projective measures.

Using these results, we have argued that any sequence of macroscopic observations behaves essentially classically provided that there is no large-scale entanglement in the sample. More precisely, the measurement of macroscopic observables generates consistent families of histories provided that their coarseness is larger than  $\sqrt{\xi/N}$  where  $\xi$  is the quantum correlation length scale of the system. Under this condition, the quantum state update rule behaves *as if* the measurements were revealing information about an objective element of reality: namely, the average reduced state of a single molecule  $\bar{\nu}$ . Hence, many independent observers acquiring their information through these measurements will arrive to a common assignment of  $\bar{\nu}$ . In this operational sense,  $\bar{\nu}$  becomes an objective element of reality [7].

Our analysis does not apply when entanglement becomes present on macroscopic scales. Clearly, not all such entanglement will yield quantum effects on the measurement of macroscopic observables. For example, the macroscopic quantum superposition (“cat”) state  $\frac{1}{\sqrt{2}}(|x_1\rangle^{\otimes N} + |x_2\rangle^{\otimes N})$ , which is *not* exchangeable, behaves just like the exchangeable state  $\frac{1}{2}[(|x_1\rangle\langle x_1|)^{\otimes N} + (|x_2\rangle\langle x_2|)^{\otimes N}]$  for all macroscopic observables. Thus, it will be interesting to determine what type of entangled states, if any, will manifest their quantum nature in the measurement of macroscopic observables.

An interesting question arises from the study of the relation between exchangeable states and macroscopic observations. We have seen in Sec. V that applying a random permutation to the molecules in a separable state yields a state which is not exchangeable, but possesses similar characteristics. We do not know what type of operation can transform a generic quantum state into an exchangeable one. We suspect that performing a tomographically complete set of macroscopic measurements on subsets of the sample followed by a random permutation of the molecules would do the trick. Physically, this would mean that a collective coupling to the environment and a diffusion process would map any state to an exchangeable state. We also suspect that a random subset of  $\sqrt{N}$  molecules out of the  $N$  molecules of the sample would also be in an exchangeable state, regardless of the initial state of the sample. This would be very interesting as it would extend the reach of our classicality analysis. Moreover, understanding under what circumstances can a sample of physical systems be treated *as if* they were all in the same unknown state is of crucial importance since this is assumed in most quantum experiments performed on macroscopic samples (e.g., any type of spectroscopy).

Finally, we have related our study to a NMR measurement model introduced in [12]. We have extended their analysis to the case where the coil is not in a pure mode state but rather in a statistical mixture of such states, like a thermal state. In this case, there are two parameters describing the macroscopic measurements: the width  $\sigma + \lambda$  of the POVM elements describes the accuracy of the measurements and width  $\sigma$  of the Kraus operators governs the disturbance caused to the state of the sample. Therefore, a measurement accuracy  $\sigma + \lambda \gg 1/\sqrt{N}$  does not guaranty a negligible disturbance except

when the measurement is ideal. The NMR measurement process therefore deserves a detailed study.

## ACKNOWLEDGMENTS

We thank Charles Bennett, Carl Caves, Raymond Laflamme, Camille Negrevergne, and Harold Ollivier for stimulating discussions. This work was supported in part by Canada's NSERC.

## APPENDIX A: SINGLE-MOLECULE POST-MEASUREMENT STATE

We will show how to compute the post-measurement state of a single molecule: namely, Eq. (24). For this, it will be useful to alter our notation a bit. In this subsection only, we will consider *non-normalized types*: If  $\mathbf{L}(X)$  denotes the normalized type of  $X$ , then its non-normalized type is  $N\mathbf{L}(X)$ . Thus, for this section only,  $\mathbf{L}(X)$  is a  $d$ -component vector whose  $j$ th component  $L_j$  equals the number of occurrences of the letter  $x_j$  in  $X$ . Adding to the notation, for the type  $\mathbf{L} = (L_1, \dots, L_d)$  of a  $N$ -letter string  $X$ , we denote by  $\mathbf{L}^{-x_j} = (L_1, \dots, L_j - 1, \dots, L_d)$  the type of the string obtained by removing one occurrence of  $x_j$  from  $X$ . Of course, this is a well-defined type only when  $L_j \geq 1$ .

Given this notation, we can write

$$Q_{\mathbf{L}}^{(N)} = \sum_j |x_j\rangle\langle x_j| \otimes Q_{\mathbf{L}^{-x_j}}^{(N-1)}, \quad (\text{A1})$$

where the  $Q_{\mathbf{L}^{-x_j}}^{(N-1)} = 0$  when  $\mathbf{L}^{-x_j}$  is not a well-defined type. Applying the state update rule and tracing out all but a single molecule, we get

$$\begin{aligned} \rho_{1|\mathbf{L}} &= \text{Tr}_{N-1} \left\{ \frac{Q_{\mathbf{L}}^{(N)} \rho_N Q_{\mathbf{L}}^{(N)}}{P(Q_{\mathbf{L}}^{(N)}|\rho_N)} \right\} \\ &= \sum_{ij} |x_i\rangle\langle x_i| \nu |x_j\rangle\langle x_j| \text{Tr} \left\{ \frac{Q_{\mathbf{L}^{-x_i}}^{(N-1)} \nu^{\otimes N-1} Q_{\mathbf{L}^{-x_j}}^{(N-1)}}{P(Q_{\mathbf{L}}^{(N)}|\rho_N)} \right\} \\ &= \sum_j R_j |x_j\rangle\langle x_j| \frac{P(Q_{\mathbf{L}^{-x_j}}^{(N-1)}|\rho_{N-1})}{P(Q_{\mathbf{L}}^{(N)}|\rho_N)}; \end{aligned} \quad (\text{A2})$$

in the last line, we used the definition of the probability, Eq. (16), and the orthogonality of the type projectors, Eq. (5). The ratio appearing in the last line can easily be computed as it involves multinomial distributions; it is equal to

$$\frac{R_1^{L_1} \dots R_j^{L_j-1} \dots R_d^{L_d} \binom{N}{L_1, \dots, L_j-1, \dots, L_d}}{R_1^{L_1} \dots R_j^{L_j} \dots R_d^{L_d} \binom{N}{L_1, \dots, L_j, \dots, L_d}} = \frac{1}{R_j} \frac{L_j}{N}.$$

Inserting this into Eq. (A2) (and keeping in mind the different definitions of  $L_j$ ) yields the result, Eq. (24). Averaging this state over measurement outcomes  $\mathbf{L}$  gives

$$\rho'_1 = \sum_j R_j |x_j\rangle\langle x_j|. \quad (\text{A3})$$

The effect of *coarse-grained* measurements  $\tilde{Q}_\ell^{(N)}$  on the state of a single molecule can be studied by straightforward modifications of the method outlined above. The results are easily predictable: while the off-diagonal elements  $|x_i\rangle\langle x_j|$  of  $\rho'_1$  are completely suppressed when  $\sigma=0$ , they only get damped by a factor proportional to  $\text{Tr}\{\tilde{Q}_{\mathbf{L}^{-x_i}}^{(N-1)}\rho_{N-1}\tilde{Q}_{\mathbf{L}^{-x_j}}^{(N-1)}\}$  with  $\mathbf{L}\approx N\mathbf{R}$  when the measurement is coarse. Since  $\mathbf{L}^{-x_i}$  and  $\mathbf{L}^{-x_j}$  are very close to each other on the probability simplex, this decoherence factor is close to unity when the smoothing function  $q_{\mathbf{L}}(\ell)$  is sufficiently wide.

For example, when we choose a Gaussian smoothing function

$$q_{\mathbf{L}}(\ell) = \left(\frac{1}{2\pi\sigma^2}\right)^{d/2} \exp\left[-\frac{\|\ell - \mathbf{L}\|_2^2}{2\sigma^2 N^2}\right]$$

(the extra  $N^2$  in the denominator of the exponent is due to the special normalization of the types used in this section), elementary algebra can be used to show that the matrix element  $\nu_{ij}$  of the single molecule density matrix  $\nu$  get updated according to

$$\nu_{ij} \xrightarrow{\ell} \nu_{ij|\ell} = \nu_{ij} \frac{\sum_{\mathbf{L}} q_{\mathbf{L}}(\ell) P(Q_{\mathbf{L}}^{(N)}|\rho_N) L_j/R_j e^{(-\ell_j + L_j + \ell_i - L_i + 1)/N^2 \sigma^2}}{\sum_{\mathbf{L}} q_{\mathbf{L}}(\ell) P(Q_{\mathbf{L}}^{(N)}|\rho_N)}.$$

The only nontrivial factors in this expression are  $L_j/R_j$  and  $e^{(-\ell_j + L_j + \ell_i - L_i + 1)/N^2 \sigma^2}$ . However, the probability  $P(Q_{\mathbf{L}}^{(N)}|\rho_N)$  is—according to the typical sequence theorem, Eq. (1)—very sharply peaked around  $L_j=R_j$ , so the factor  $L_j/R_j$  is nearly trivial [i.e., it differs from 1 by  $O(1/N)$ ] in dominant terms of the sum. Similarly, the smoothing function  $q_{\mathbf{L}}(\ell)$  becomes very small when  $L_j$  differs from  $\ell_j$  by more than roughly  $N\sigma$ . Thus, the argument of the exponential  $e^{(-\ell_j + L_j + \ell_i - L_i + 1)/N^2 \sigma^2}$  is of order  $1/N$  in the dominant terms of the sum, so this factor also differs from unity by  $O(1/N)$ . Thus, the disturbance caused to the matrix elements  $\nu_{ij}$  are of *relative* order  $1/N$ . A more quantitative comparison of the original and final single-molecule state is obtained straightforwardly from our general fidelity trade-off and the monotonicity of fidelity.

## APPENDIX B: CONDITIONAL FIDELITY

We will compute, for an ensemble of  $N$  molecule initially in state  $|\Psi_N\rangle = \sum_{j=1}^d \beta_j |x_j\rangle$ , the fidelity between the pre- and *conditional* post-measurement states  $\rho_N$  and  $\rho_{N|\ell}$ , respectively. Starting from Eqs. (17) and (18), we can express the fidelity

$$F(\rho_N, \rho_{N|\ell}) = \langle \Psi_N | \rho_{N|\ell} | \Psi_N \rangle = \frac{\left[ \sum_{\mathbf{L}} \sqrt{q_{\mathbf{L}}(\ell)} m(\mathbf{L}, \mathbf{R}) \right]^2}{\sum_{\mathbf{L}} q_{\mathbf{L}}(\ell) m(\mathbf{L}, \mathbf{R})},$$

where  $R_j = \langle x_j | \nu | x_j \rangle$ , and the multinomial coefficient

$$m(\mathbf{L}, \mathbf{R}) = \binom{N}{NL_1, \dots, NL_d} \prod_j R_j^{NL_j}.$$

Assuming that the smoothing function is Gaussian,

$$q_{\mathbf{L}}(\ell) = \left(\frac{1}{2\pi\sigma^2}\right)^{d/2} \exp\left[-\frac{\|\ell - \mathbf{L}\|_2^2}{2\sigma^2}\right],$$

the fidelity takes the form

$$F(\rho_N, \rho_{N|\ell}) = \frac{\left( \sum_{\mathbf{L}} \exp\left[-\frac{\|\ell - \mathbf{L}\|_2^2}{4\sigma^2}\right] m(\mathbf{L}, \mathbf{R}) \right)^2}{\sum_{\mathbf{L}} \exp\left[-\frac{\|\ell - \mathbf{L}\|_2^2}{2\sigma^2}\right] m(\mathbf{L}, \mathbf{R})}. \quad (\text{B1})$$

The lower bound on this expression is found by similar means that led to the bound on the fidelity between  $\rho_N$  and  $\rho'_N$  in Sec. III, except that the presence of a denominator—absent in Eq. (25)—creates extra complications. Hence, let us start by considering the numerator. First, we use the triangle inequality  $\|\mathbf{L} - \ell\|_2 \leq \|\ell - \mathbf{R}\|_2 + \|\mathbf{R} - \mathbf{L}\|_2$  to get an upper bound

$$e^{-\|\mathbf{L} - \ell\|_2^2/2\sigma^2} \geq e^{-\|\mathbf{R} - \ell\|_2 + \|\mathbf{R} - \mathbf{L}\|_2)^2/2\sigma^2}.$$

Then, just like we did in Sec. III, we truncate the sum to the domain  $\mathcal{D} = \{\mathbf{L} : \|\mathbf{L} - \mathbf{R}\|_2 \leq \Delta\}$ . Clearly on this domain

$$e^{-\|\mathbf{L} - \ell\|_2^2/2\sigma^2} \geq e^{-\|\mathbf{R} - \ell\|_2 + \Delta)^2/2\sigma^2} \text{ on } \mathcal{D}. \quad (\text{B2})$$

Combining these two steps yields the lower bound

$$e^{-\|\mathbf{R} - \ell\|_2 + \Delta)^2/2\sigma^2} \left( \sum_{\mathbf{L} \in \mathcal{D}} m(\mathbf{L}, \mathbf{R}) \right)^2 \geq e^{-\|\mathbf{R} - \ell\|_2 + \Delta)^2/2\sigma^2} \left( 1 - e^{-Nd\Delta^2/2} \right)^2 \quad (\text{B3})$$

for the numerator of Eq. (B1), where we appealed to the Cauchy-Schwartz inequality and the typical sequence theorem, Eq. (1), to get the second line.

Since we are interested in finding a lower bound to the fidelity, we must now find an *upper* bound to the denominator of Eq. (B1). To do this, we decompose the sum into  $\sum_{\mathbf{L} \in \mathcal{D}} + \sum_{\mathbf{L} \notin \mathcal{D}}$ . The sum outside the domain  $\mathcal{D}$  can easily be bounded:

$$\sum_{\mathbf{L} \notin \mathcal{D}} e^{-\|\ell - \mathbf{L}\|_2^2/2\sigma^2} m(\mathbf{L}, \mathbf{R}) \leq \sum_{\mathbf{L} \notin \mathcal{D}} m(\mathbf{L}, \mathbf{R}) \leq e^{-Nd\Delta^2/2}, \quad (\text{B4})$$

where the second inequality follows from the Cauchy-Schwartz inequality and the typical sequence theorem. For the sum inside  $\mathcal{D}$ , we use the triangle inequality  $\|\ell - \mathbf{R}\|_2 \leq \|\ell - \mathbf{L}\|_2 + \|\mathbf{R} - \mathbf{L}\|_2$ , which translates into

$$\|\mathbf{L} - \ell\|_2 \geq \|\ell - \mathbf{R}\|_2 - \Delta \quad (\text{B5})$$

on the domain  $\mathcal{D}$ . Before squaring this quantity, it is important to determine whether  $\|\ell - \mathbf{R}\|_2 - \Delta$  is a positive or a negative quantity. We shall thus distinguish two cases.

*Case 1.*  $\|\ell - \mathbf{R}\|_2 \leq \Delta$ . In this case, we use the straightforward bound

$$\sum_{\mathbf{L} \in \mathcal{D}} e^{-\|\ell - \mathbf{L}\|_2^2/2\sigma^2} m(\mathbf{L}, \mathbf{R}) \leq 1.$$

Combining this with Eq. (B4) yields the upper bound  $1 + e^{-Nd\Delta^2/2}$  for the denominator of Eq. (B1), so together with

the numerator, Eq. (B3), we get a bound on the fidelity:

$$F(\rho_N, \rho_{N|\ell}) \geq \frac{e^{-(\|\mathbf{R} - \ell\|_2 + \Delta)^2/2\sigma^2}(1 - e^{-Nd\Delta^2/2})^2}{1 + e^{-Nd\Delta^2/2}} \geq \frac{e^{-2\Delta^2/\sigma^2}(1 - e^{-Nd\Delta^2/2})^2}{1 + e^{-Nd\Delta^2/2}}.$$

Case 2.  $\|\ell - \mathbf{R}\|_2 > \Delta$ . In this case, the inequality, Eq. (B5), gives  $\|\mathbf{L} - \ell\|_2^2 \geq (\|\ell - \mathbf{R}\|_2 - \Delta)^2$ , so the sum on the domain  $\mathcal{D}$  is bounded by

$$\sum_{\mathbf{L} \in \mathcal{D}} e^{-\|\ell - \mathbf{L}\|_2^2/2\sigma^2} m(\mathbf{L}, \mathbf{R}) \leq e^{-(\|\ell - \mathbf{R}\|_2 - \Delta)^2/2\sigma^2} \sum_{\mathbf{L} \in \mathcal{D}} m(\mathbf{L}, \mathbf{R}) \leq e^{-(\|\ell - \mathbf{R}\|_2 - \Delta)^2/2\sigma^2},$$

so together with Eq. (B4), we get the upper bound  $e^{-(\|\ell - \mathbf{R}\|_2 - \Delta)^2/2\sigma^2} + e^{-Nd\Delta^2/2}$  for the denominator of Eq. (B1). Combining this result with the bound on the numerator, Eq. (B3), yields

$$F(\rho_N, \rho_{N|\ell}) \geq \frac{e^{-(\|\mathbf{R} - \ell\|_2 - \Delta)^2/2\sigma^2}(1 - e^{-Nd\Delta^2/2})^2}{e^{-(\|\ell - \mathbf{R}\|_2 - \Delta)^2/2\sigma^2} + e^{-Nd\Delta^2/2}} = \frac{e^{-\|\ell - \mathbf{R}\|_2 \Delta/\sigma^2}(1 - e^{-Nd\Delta^2/2})^2}{1 + e^{-Nd\Delta^2/2 + (\|\ell - \mathbf{R}\|_2 - \Delta)^2/2\sigma^2}},$$

or using the fact that the  $L_2$  distance is bounded by  $\|\mathbf{p} - \mathbf{q}\|_2 \leq 2$  for any two probability distributions  $\mathbf{p}$  and  $\mathbf{q}$ ,

$$F(\rho_N, \rho_{N|\ell}) \geq \frac{e^{-2\Delta/\sigma^2}(1 - e^{-Nd\Delta^2/2})^2}{1 + e^{-Nd\Delta^2/2 + 2/\sigma^2}}. \quad (\text{B6})$$

This second case gives the worst bound, so Eq. (B6) turns out to be universal, independent of the sign of  $\|\ell - \mathbf{R}\|_2 - \Delta$ . The cutoff  $\Delta$  is a free parameter, so we should again try to optimize it in order to achieve the tightest bound. We have not found a closed form expression for this optimum. However, any assignment  $\sigma = N^{-a}$  and  $\Delta = N^b$  with  $2b < 1$ ,  $2a < b$ , and  $2a < 1 - 2b$  yields

$$F(\rho_N, \rho_{N|\ell}) \geq \frac{e^{-AN^{-\alpha}}(1 - e^{-BN^\beta})^2}{1 + e^{-CN^\gamma}} \approx 1 - \frac{A}{N^\gamma} - 2e^{-BN^\beta} - e^{-CN^\gamma}$$

for some positive constants  $A$ ,  $B$ ,  $C$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$ . For example, when the accuracy of the measurement apparatus is fixed to a constant  $\sigma$ , setting the arbitrary cutoff parameter  $\Delta = N^{-1/3}$  gives

$$F(\rho_N, \rho_{N|\ell}) \geq 1 - \frac{2}{\sigma^2 N^{1/3}} - 3e^{-dN^{1/3/2}}$$

to first order.

- 
- [1] D. Finkelstein, *Trans. N. Y Acad. Sci.* **25**, 621 (1963).
  - [2] J. B. Hartle, *Am. J. Phys.* **36**, 704 (1968).
  - [3] N. Graham, in *The Many Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and N. Graham (Princeton University Press, Princeton, 1973).
  - [4] E. Farhi, J. Goldstone, and S. Gutmann, *Ann. Phys. (N.Y.)* **192**, 368 (1989).
  - [5] C. M. Caves and R. Schack, e-print quant-ph/0409144.
  - [6] C. M. Caves, C. A. Fuchs, and R. Schack, *J. Math. Phys.* **43**, 4537 (2002).
  - [7] D. Poulin, Ph.D. thesis, University of Waterloo, 2004.
  - [8] M. Hayashi and K. Matsumoto, e-print quant-ph/0209124.
  - [9] C. H. Bennett, A. W. Harrow, and S. Lloyd, e-print quant-ph/0403078.
  - [10] L. Viola, E. M. Fortunato, M. A. Pravia, E. Knill, R. Laflamme, and D. G. Cory, *Science* **293**, 2059 (2001).
  - [11] J. J. Halliwell, *Phys. Rev. D* **58**, 105015 (1998); *Phys. Rev. Lett.* **83**, 2481 (1999).
  - [12] S. Lloyd and J.-J. E. Slotine, *Phys. Rev. A* **62**, 12307 (2000).
  - [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
  - [14] C. A. Fuchs, e-print quant-ph/0205039.
  - [15] E. J. Squires, *Phys. Lett. A* **145**, 67 (1990).
  - [16] B. de Finetti, *Theory of Probability* (Wiley, New York, 1990).
  - [17] R. Griffiths, Perimeter Institute Lecture, March 2004.
  - [18] C. M. Caves and C. A. Fuchs, in *The Dilemma of Einstein, Podolsky and Rosen—60 Years Later*, edited by A. Mann and M. Revzen [*Ann. Isr. Phys. Soc.* **12**, 226 (1996)].
  - [19] C. A. Fuchs and A. Peres, *Phys. Today* **53**(3), 70 (2000).
  - [20] N. D. Mermin, in *Quantum (Un)speakables: Essays in Commemoration of John S. Bell*, edited by R. Bertlmann and A. Zeilinger (Springer-Verlag, Berlin, 2001).
  - [21] R. L. Hudson and G. R. Moody, *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* **33**, 343 (1976).
  - [22] R. L. Hudson, *Found. Phys.* **11**, 805 (1981).
  - [23] R. B. Griffiths, *J. Stat. Phys.* **36**, 219 (1984); R. Omnès R., *ibid.* **53**, 893 (1988); *Rev. Mod. Phys.* **64**, 339 (1992); M. Gell-Mann and J. B. Hartle, in *Complexity, Entropy and the Physics of Information*, edited by W. H. Zurek, Santa Fe Institute Studies in the Science of Complexity, Vol. 8 (Westview Press, Redwood City, CA, 1990), p. 425.
  - [24] D. Giulini *et al.*, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin, 1996); W. H. Zurek, *Rev. Mod. Phys.* **75**, 715 (2003).
  - [25] D. Janzing and T. Beth, *Phys. Rev. A* **61**, 52308 (2000).
  - [26] D. G. Cory *et al.*, *Fortschr. Phys.* **48**, 875 (2000); R. Laflamme *et al.*, e-print quant-ph/0207172.
  - [27] N. Bloembergen and R. V. Pound, *Phys. Rev.* **95**, 8 (1954).
  - [28] C. H. Bennett, G. Brassard, and N. D. Mermin, *Phys. Rev. Lett.* **68**, 557 (1992).