Dynamics of statistical distance: Quantum limits for two-level clocks

Samuel L. Braunstein
Department of Chemical Physics, The Weizmann Institute of Science, Rehovot 76100, Israel

G. J. Milburn
Department of Physics, University of Queensland, St. Lucia 4072, Australia
(Received 9 September 1994)

We study the evolution of statistical distance on the Bloch sphere under unitary and nonunitary dynamics. This corresponds to studying the limits to clock precision for a clock constructed from a two-state system. We find that the initial motion away from pure states under nonunitary dynamics yields the greatest accuracy for a "one-tick" clock; in this case the clock's precision is not limited by the largest frequency of the system.

PACS number(s): 03.65.Bz, 06.30.Ft, 42.50.Lc

I. INTRODUCTION

In a recent paper Braunstein and Caves [1] introduced a natural metric on density operators based on the statistical distinguishability of nearby states. This metric defines a statistical distance between density operators which generalizes earlier results for pure states [2]. If density operators are smoothly related by the change in a continuous parameter then this metric of statistical distance tells us the absolute limits quantum theory places on our ability to determine the value of the parameter based on arbitrary measurements of N members of an ensemble with a single value of this parameter. If this parameter is time then we can study the absolute quantum limits to clock design.

We start by reviewing some of the consequences of this metric and the limits it places on our ability to extract the values of parameters from measurements on a finite number of systems. The body of this paper is technical and consists of studying the evolution of statistical distance on the Bloch sphere. Finally, we ask how this result throws light on the limits to clock design. In particular, we consider the limitations to precision available in just a single cycle ("one tick" of the clock) when the clock is constructed from ensembles of two-level systems—for which our analysis of statistical distance on the Bloch sphere may be applied. We find that for nonunitary dynamics the clock's precision is apparently not limited by the largest frequency of the system.

There is a simple way to see why our results predict higher precision for nonunitary behavior over unitary behavior in the design of a one-tick (single-cycle) clock. Consider the two model probabilities for finding a two-level system in a state orthogonal to its initial pure state:

$$p_{\mathrm{U}} = 1 - \cos^2(\omega t), \qquad p_{\mathrm{NU}} = 1 - \exp(-\gamma t)\ , \tag{1.1}$$

where the former probability p_{U} corresponds to typical unitary evolution and the latter probability p_{NU} corresponds to typical nonunitary evolution. For times short compared to the reciprocal rates, $t < \omega^{-1}, \gamma^{-1}$, these probability laws may be approximated by

$$p_{\mathrm{U}} \simeq (\omega t)^2, \qquad p_{\mathrm{NU}} \simeq \gamma t\ . \tag{1.2}$$

Now suppose that we plan to extract the value for the time t elapsed between preparation and measurement by measuring the relative frequency for a transition to occur. This relative frequency is a direct estimate of the transition probability, and from this Eq. (1.3) allows us to determine the value of t (assuming we already know the rate constant involved). In particular, we have

$$t_{\mathrm{est}} \simeq \omega^{-1}\sqrt{f_{\mathrm{U}}}, \qquad t_{\mathrm{est}} \simeq \gamma^{-1} f_{\mathrm{NU}}\ , \tag{1.3}$$

where f_{U} and f_{NU} are the observed relative frequencies for a transition to occur for the unitary and nonunitary evolutions, respectively. Elementary error analysis tells us what confidence to place in our estimates. All other things being equal we see from Eq. (1.4) that the unitary evolution yields an inferior (square-root) dependence on the measured quantity. (Fewer transitions occur in the same time for the same rate constants leading to a larger relative error from counting statistics). This inferior behavior shows up in broader confidence intervals and hence a less precise determination of the elapsed time t. Statistical distance allows us to relax the assumption of the type of measurement made to extract information because it places bounds on the efficiency of measurement.

II. STATISTICAL DISTANCE

We now review the limitations to determining parameters through measurement based on statistical distance

[1]. Consider N replicas of a quantum system. Each replica is prepared in the same quantum state (density operator) $\hat{\rho}(X)$, which is parametrized by the single parameter X. A general smooth path on the space of density operators may be described by

$$\hat{\rho}(X) = \sum_j p_j |j\rangle\langle j| , \tag{2.1}$$

where both the eigenvalues p_j and the eigenvectors $|j\rangle$ can change along the path. By choosing optimal measurements and methods of data analysis on the N replicas our job is to optimally distinguish the $\hat{\rho}(X)$ for differing X and hence the value of X.

The most general measurement permitted by quantum mechanics [3,4] can be described by a set of bounded, non-negative, Hermitian operators $\hat{E}(\xi)d\xi$ (generalizations of projection operators), which are complete in the sense that

$$\int d\xi\, \hat{E}(\xi) = \hat{1\!\!1} = (\text{unit operator}) . \tag{2.2}$$

The quantity ξ labels the "results" of the measurement; written here as a single continuous real variable, it could be discrete or multivariate. In the discrete case the operators $\hat{E}(\xi)$ are called "effects." In the continuous case $\hat{E}(\xi)$ is an "effect density," and $\hat{E}(\xi)d\xi$ is an "effect-valued measure" or a "positive-operator-valued measure." The probability distribution for result ξ, given the parameter X, is

$$p(\xi|X) = \mathrm{tr}\left[\hat{E}(\xi)\hat{\rho}(X)\right] . \tag{2.3}$$

Let $\xi_1, \ldots, \xi_N$ denote the results of measurements on the N replicas of our quantum system. Consider a general form of data analysis in which a function

$$X_{\text{est}} = X_{\text{est}}(\xi_1, \ldots, \xi_N) \tag{2.4}$$

yields an estimate X_{est} for X, based on the data $\xi_1, \ldots, \xi_N$ and nothing else. The estimate's deviation away from the parameter can be quantified by [1]

$$\delta X \equiv \frac{X_{\text{est}}}{|d\langle X_{\text{est}}\rangle_X/dX|} - X , \tag{2.5}$$

where the derivative $d\langle X_{\text{est}}\rangle_X/dX$ is included to remove the local difference in the "units" of the estimator and the parameter.

There is a lower bound [1] on the second moment of δX:

$$\langle (\delta X)^2 \rangle \geq \frac{1}{N(ds/dX)^2} . \tag{2.6}$$

This inequality holds for any effect density $\hat{E}(\xi)$ and is written in terms of a line element $(ds)^2$, which defines a "statistical distance" that measures the distinguishability of neighboring quantum states and provides a natural Riemannian geometry on the space of density operators. The explicit form that Braunstein and Caves [1] (see also [3,5]) find for the line element is

$$\left(\frac{ds}{dX}\right)^2 = \left\langle \left[\mathcal{L}_{\hat{\rho}}(d\hat{\rho}/dX)\right]^2 \right\rangle = \mathrm{tr}\left[\mathcal{L}_{\hat{\rho}}(d\hat{\rho}/dX)\, d\hat{\rho}/dX\right] , \tag{2.7}$$

where $\mathcal{L}_{\hat{\rho}}$ is a super operator which, in the basis that diagonalizes $\hat{\rho}$, takes the form

$$\mathcal{L}_{\hat{\rho}}(\hat{O}) = \sum_{\{j,k|p_j+p_k\neq 0\}} \frac{2}{p_j + p_k} O_{jk}|j\rangle\langle k| . \tag{2.8}$$

If $\hat{\rho}$ has no zero eigenvalues, $\mathcal{L}_{\hat{\rho}}$ is the inverse of the super operator defined by $\mathcal{R}_{\hat{\rho}}(\hat{O}) \equiv \frac{1}{2}(\hat{\rho}\hat{O} + \hat{O}\hat{\rho})$; hence, $\mathcal{L}_{\hat{\rho}}$ was denoted $\mathcal{R}_{\hat{\rho}}^{-1}$ by Braunstein and Caves [1]. Helstrom [5] and Holevo [3] call $\mathcal{L}_{\hat{\rho}}(\hat{\rho}')$ the "symmetric logarithmic derivative" of $\hat{\rho}$. The lower bound in the inequality (2.6) can be achieved by using effects $\hat{E}(\xi)$ that are one-dimensional projection operators onto the orthonormal eigenstates of the Hermitian operator $\mathcal{L}_{\hat{\rho}}(\hat{\rho}')$ [1,6]. For pure states $\hat{\rho}(X) = |\psi(X)\rangle\langle\psi(X)|$ the distinguishability metric [2] reduces to

$$(ds)^2 = 4\left[1 - |\langle\psi(X+dX)|\psi(X)\rangle|^2\right] , \tag{2.9}$$

which is also known as the Fubini-Study metric.

Note that in Eq. (2.6) there is an overall $\sqrt{N}$ improvement in the precision to which we may determine the parameter X as we increase the number of identically prepared systems we can make measurements upon. This is familiar to us as the typical improvement upon increasing our sample size; here we see that it is a general limit to how well we can determine a parameter from quantum systems. It is worth pointing out the generality of the derivation of this quantum bound: it allows for any type of measurement on the N replicas of the system (even entangled measurements on all replicas simultaneously [6]) and allows for arbitrary methods of data analysis in analyzing the results to determine the value of X. In general, this bound is achievable for N large enough so long as an optimal measurement is chosen [1]. We will discuss later how large N must be to achieve the performance described by statistical distance.

III. EVOLUTION OF STATISTICAL DISTANCE

We will now restrict our attention to using statistical distance to determine the parameter of time t. For this parameter Eq (2.6) becomes

$$N\langle(\delta t)^2\rangle \left(\frac{ds}{dt}\right)^2 \geq 1 . \tag{3.1}$$

Since we are now interested in the rate of change of statistical distance with respect to time we shall study the evolution of statistical distance both under unitary and nonunitary evolution. The faster this distance changes the greater intrinsic precision is available for us to exploit in determining the parameter of time.

Under unitary evolution, the density operator changes in time as

$$\frac{d\hat{\rho}}{dt} = \frac{-i}{\hbar}\left[\hat{H}, \hat{\rho}\right] . \tag{3.2}$$

If the system starts in a pure state it will remain pure. The rate of change of statistical distance is then given by

$$\left(\frac{ds}{dt}\right)^2 = \frac{4}{\hbar^2}\langle\psi(0)|(\Delta\hat{H})^2|\psi(0)\rangle . \tag{3.3}$$

Note that this is independent of t, depending only on the initial state. In fact, apart from the factor of 4, this is simply the probability that the system will not be found in the initial state a time dt later. The bound to determining the value of t for a state $|\psi(t)\rangle$ from arbitrary measurements is simply

$$\langle(\delta t)^2\rangle \geq \frac{\hbar^2}{4\langle\psi(0)|(\Delta\hat{H})^2|\psi(0)\rangle} . \tag{3.4}$$

The conditions required to realize the lower bound are given in Ref. [6].

Now consider the case of purely nonunitary dynamics. If we restrict the discussion to Markov open systems, the most general form for the evolution equation is

$$\frac{d\hat{\rho}}{dt} = \sum_{j=1}^{K}\left[\hat{A}_j\hat{\rho}\hat{A}_j^\dagger - \mathcal{R}_{\hat{\rho}}(\hat{A}_j^\dagger\hat{A}_j)\right] , \tag{3.5}$$

where K is the number of channels "open" (i.e., coupled directly) to the environment. For simplicity we only treat the case of one open channel (K=1). In this case the rate of change of statistical distance is given by

$$\left(\frac{ds}{dt}\right)^2 = \mathrm{tr}[\hat{A}\hat{\rho}\hat{A}^\dagger\,\mathcal{L}_{\hat{\rho}}(\hat{A}\hat{\rho}\hat{A}^\dagger)] - 2\langle(\hat{A}^\dagger)^2\hat{A}^2\rangle + \langle(\hat{A}^\dagger\hat{A})^2\rangle . \tag{3.6}$$

We shall now restrict our attention to two-level systems where we can derive explicit expressions for the statistical distance and its rate of change.

IV. DYNAMICS ON THE BLOCH SPHERE

In the case of two-state systems there is a geometric representation of states in terms of the Bloch sphere. Unitary evolution corresponds to motion in the sphere at a constant radius (for pure states this is motion on the sphere's surface—at unit radius) while nonunitary evolution in general corresponds to a motion within the sphere. The rate of change of statistical distance limits how well we can measure time by making the measurements on the evolving states.

The most general density operator for a two-state system may be written as

$$\hat{\rho} = \frac{1}{2}(\hat{\mathbb{1}} + x^1\hat{\sigma}_1 + x^2\hat{\sigma}_2 + x^3\hat{\sigma}_3) , \tag{4.1}$$

where $\vec{x}$ is the Bloch vector and the operators $\hat{\sigma}_j$ are Pauli "matrices" defined by

$$\begin{aligned}\hat{\sigma}_1 &= |1\rangle\langle 2| + |2\rangle\langle 1|,\\ \hat{\sigma}_2 &= -i(|1\rangle\langle 2| - |2\rangle\langle 1|),\\ \hat{\sigma}_3 &= |2\rangle\langle 2| - |1\rangle\langle 1| ,\end{aligned} \tag{4.2}$$

and we require

$$x^j x^j \leq 1 . \tag{4.3}$$

Note that repeated indices are summed over with Latin indices used for the range 1, 2, 3. The equality defines the pure states. Thus pure states lie on the surface of the unit sphere while mixed states lie inside the sphere. It is, however, possible to define a new coordinate system so that all states lie on the surface of a hemisphere embedded in a four-dimensional Euclidean space. Define the new coordinate as

$$x^0 = \sqrt{2[1 - \mathrm{tr}(\hat{\rho}^2)]} \tag{4.4}$$

where, by convention, we take the positive square root. Then it is clear that $x^\mu x^\mu = 1$, with Greek indices ranging over 0, 1, 2, 3.

We now prove that the statistical distinguishability is just the metric on the surface of the four sphere, that is [7]

$$ds^2 = dx^\mu dx^\mu . \tag{4.5}$$

Let

$$\hat{\rho} = \frac{1}{2}(\hat{\mathbb{1}} + x^j\hat{\sigma}_j), \tag{4.6}$$

$$d\hat{\rho} = \frac{1}{2}\,dx^j\hat{\sigma}_j . \tag{4.7}$$

We then write the most general form for $\mathcal{L}_{\hat{\rho}}(d\hat{\rho})$ as

$$\mathcal{L}_{\hat{\rho}}(d\hat{\rho}) = \alpha^0\hat{\mathbb{1}} + \alpha^j\hat{\sigma}_j , \tag{4.8}$$

which may be rewritten as $d\hat{\rho} = \mathcal{R}_{\hat{\rho}}(\alpha^0\hat{\mathbb{1}} + \alpha^j\hat{\sigma}_j)$. Substituting in Eqs. (4.6) and (4.7) and equating coefficients we find

$$\alpha^0 + \alpha^j x_j = 0, \tag{4.9}$$

$$\alpha^j + \alpha^0 x^j = dx^j . \tag{4.10}$$

Now $x^0 dx^0 = -2\,\mathrm{tr}\,(\hat{\rho}\,d\hat{\rho}) = -x^j dx^j$. Combining this last result with Eqs. (4.9) and (4.10) and substituting into Eq. (2.7) we find the appealing result

$$ds^2 = dx^\mu dx^\mu . \tag{4.11}$$

The rate of change of statistical distance then becomes

$$\left(\frac{ds}{dt}\right)^2 = \sum_{\mu=0}^{3}\left(\frac{dx^\mu}{dt}\right)^2 . \tag{4.12}$$

As an example of unitary motion consider the Hamiltonian

$$\hat{H} = \hbar\omega\hat{\sigma}_2 . \tag{4.13}$$

This describes a rotation of the Bloch vector about the x^2 axis. Let the initial state of the system be an eigenstate

of $\hat{\sigma}_1$. This corresponds to a point on the equator of the Bloch sphere. This point moves with time along the surface of the sphere towards the south pole, all the while remaining in the x^1-x^3 plane. Using Eq. (3.4), the bound to the error in time estimation is then

$$\langle(\delta t)^2\rangle \geq \frac{1}{4\omega^2}, \tag{4.14}$$

which is not a surprising result. The error in the clock is half a cycle.

We can integrate the rate of change in statistical distance as the initial Bloch vector rotates to the point opposite on the sphere's surface, that is, over one-half period. It is not difficult to see that this gives π. In other words the statistical distance between the two eigenstates of $\hat{\sigma}_1$, along a great circle path on the Bloch sphere, is π. Thus there is a very close correspondence between angles swept by paths on the Bloch sphere and the statistical distance between unitarily evolving pure states. In fact, it is exactly this: it is the integral of the angle increments subtended by any path. For evolution interior to the sphere the same holds, but now the angle increments are measured in the four-dimensional coordinate system.

We turn now to the change in statistical distance due to nonunitary evolution. We consider two models: the decay model and the diffusion model [8]. The decay model is defined by the evolution equation

$$\frac{d\hat{\rho}}{dt} = \gamma\left[\hat{\sigma}_-\hat{\rho}\hat{\sigma}_+ - \mathcal{R}_{\hat{\rho}}(\hat{\sigma}_+\hat{\sigma}_-)\right], \tag{4.15}$$

where $\hat{\sigma}_\pm \equiv \hat{\sigma}_1 \pm i\hat{\sigma}_2$ are the raising and lowering operators. Substituting Eq. (4.1) into this we obtain

$$\begin{aligned}\frac{dx^1}{dt} &= \frac{-\gamma x^1}{2},\\ \frac{dx^2}{dt} &= \frac{-\gamma x^2}{2},\\ \frac{dx^3}{dt} &= -\gamma(x^3+1).\end{aligned} \tag{4.16}$$

This can describe the spontaneous decay of a two-level atom, in which case $1/\gamma$ is the lifetime of the excited state. Any state on or inside the Bloch sphere will approach the south pole ($x^3 = -1$) as $t \longrightarrow \infty$. For example, an initial eigenstate of $\hat{\sigma}_1$ will move inside the Bloch sphere and approach the south pole while remaining in the x^1-x^3 plane.

The diffusion model is defined by the evolution equation,

$$\frac{d\hat{\rho}}{dt} = \gamma[\hat{\sigma}_3\,\hat{\rho}\,\hat{\sigma}_3 - \mathcal{R}_{\hat{\rho}}(\hat{\sigma}_3^2)] = \gamma\,(\hat{\sigma}_3\,\hat{\rho}\,\hat{\sigma}_3 - \hat{\rho}). \tag{4.17}$$

Using Eq. (4.1), this may be written as

$$\begin{aligned}\frac{dx^1}{dt} &= -2\gamma x^1,\\ \frac{dx^2}{dt} &= -2\gamma x^2,\\ \frac{dx^3}{dt} &= 0.\end{aligned} \tag{4.18}$$

All states eventually become diagonal in the eigenstates of $\hat{\sigma}_3$. States on (or inside) the Bloch sphere move directly towards the x^3 axis as $t \longrightarrow \infty$, on the plane $x^3(t) = x^3(0)$.

V. NET STATISTICAL DISTANCES

Suppose the initial state is an eigenstate of $\hat{\sigma}_1$,

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle). \tag{5.1}$$

In the case of the damped model we find the coordinates change by

$$\begin{aligned}x^1(t) &= e^{-t/2},\\ x^2(t) &= 0,\\ x^3(t) &= e^{-t} - 1,\\ x^0(t) &= e^{-t/2}\sqrt{1-e^{-t}}\end{aligned} \tag{5.2}$$

(here, and in what follows, time is measured in units of inverse decay constant).

The rate of change of statistical distance is then given by

$$\frac{ds}{dt} = \frac{1}{2}\sqrt{\frac{2-e^{-t}}{e^t-1}}. \tag{5.3}$$

Note that as $t \longrightarrow 0$, $ds/dt \longrightarrow \infty$, while in the limit $t \longrightarrow \infty$, $ds/dt \longrightarrow 0$. The interpretation of this is as follows. As the initial state is an eigenstate of $\hat{\sigma}_1$ we can monitor the evolution from this state by measuring $\hat{\sigma}_1$. As soon as we get a result other than the eigenvalue $+1$ we are certain that the system has left the initial state. As soon as the state moves in from the surface of the Bloch sphere it is perfectly distinguishable from the initial state and thus the rate of change of statistical distance goes to infinity. However as time goes on, it is less and less easy to see changes in the state and the rate of change of statistical distance goes to zero. Thus initially the estimate of time is perfect in that the uncertainty relation is $\langle(\delta t)^2\rangle_{t=0} \geq 0$ while for long times the determination of time becomes less and less accurate.

A plot of ds/dt for the damped model is shown in Fig. 1, for an initial state $x^1(0) = 1$. If we now integrate this result from $t = 0$ to $t = \infty$ we obtain the statistical distance between the initial state and the state at the south pole as the system moves along a path inside the Bloch sphere. The result is $s \simeq 1.910$. The important point to note about this result is that it is greater than $\pi/2$. The statistical distance between the points (1, 0, 0) and (0, 0, −1) is shortest for a path along a great circle joining these two points and in fact is exactly $\pi/2 \simeq 1.5708$. Thus, as expected, the path through the interior of the sphere connecting these points has a greater statistical distance.

We now turn to the diffusion model with the same initial eigenstate of $\hat{\sigma}_1$. The solutions are

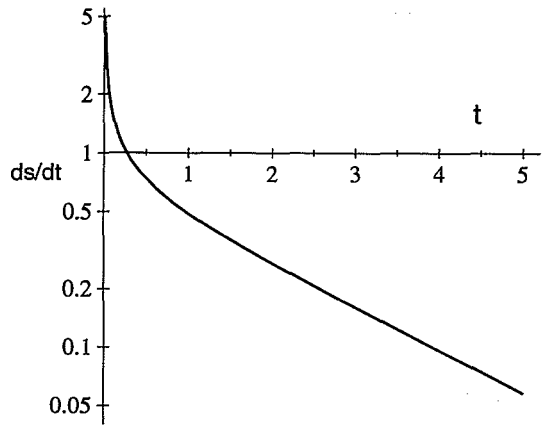

FIG. 1. Plot of the evolution of the rate of change of the statistical distance ds/dt versus time t for the decay model of Eq. (4.15), with an initial pure state given by $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$.

$$\begin{aligned} x^1(t) &= e^{-2t}, \\ x^2(t) &= 0, \\ x^3(t) &= 0, \\ x^0(t) &= \sqrt{1-e^{-4t}}\,. \end{aligned} \tag{5.4}$$

As before we take as an example the same initial eigenstate of σ_1 in Eq. (5.1). The result is

$$\frac{ds}{dt} = \frac{2}{\sqrt{e^{4t}-1}}\,. \tag{5.5}$$

Once again we find the rate of change of statistical distance diverges as $t \longrightarrow 0$, while it tends to zero as $t \longrightarrow \infty$. This has the same interpretation as the damped case. The integrated statistical distance along this path is found to be $\pi/2$. In terms of the Bloch sphere the state moves from the point on the equator $x^1 = 1$ directly to the origin of the sphere. In terms of the embedded hemisphere, the state moves along a great circle in the x^0-x^1 plane, corresponding to geodesic motion, terminating at the point $x^0 = 1$. In this case the minimum statistical distance between the initial and final point is $\pi/2$. So the change of statistical distance for this quantum nondemolition [8] measurement is a minimum; this is not a general feature.

VI. PRECISION MEASUREMENT OF TIME AND CLOCKS

When we think of a clock we tend to think of a device which gives a continual (or roughly so) readout of the current time. For a quantum-mechanical system, continual readouts will disturb the state of the clock and perturb its evolution unless the system is very large. Rather than the readout being exactly continual it is more likely to be in discrete steps—each cycle will correspond to observation free evolution followed by the "tick," the fast readout. We study here a paradigm for a very simple model of a clock where at the end of each cycle the length of the cycle will be determined as accurately as possible and the quantum state of the clock will be reset for the next cycle to proceed. Because we are primarily interested here in considering small systems with decidedly quantum behavior under monitoring we do not expect good long term stability necessarily; we except such considerations to be more fruitful for larger more "classical" systems.

In preceding sections we have been studying the precision with which such a single cycle could be performed. For each cycle we assume that there are N identically prepared two-level systems and after some short (not precisely specified period) a readout is made disturbing the system; the data from this readout is used to give a precise value to the length of the cycle though data analysis of the measurement results, which when added to the lengths of the preceding cycles will give a roughly continual readout of the current time. Because we reduce the action of a clock to clocks of single cycles we call these "one-shot" or "one-tick" clocks (or even clocks which never "tock," or "tockless" clocks).

The formalism of the previous sections is ideal for the analysis of one-tick clocks: the rate of change of statistical distance gives the limits to achievable precision for determining the length in time for a single cycle. We found that this rate of change grows in an unbounded manner as we approach the pure state boundary of the Bloch sphere through nonunitary evolution. This growth suggests that there is an in principle precision which is apparently not limited by the largest frequency in the system as one would expect for purely unitary evolution. Can we exploit this divergence? If so, what price must we pay? There are two problems: one is to find a suitable optimal measurement (this has been discussed in detail in Refs. [1,6] and will not be considered further here); further, one needs to use a sufficiently optimal method of data analysis and a sufficiently large number N of identically prepared systems to achieve the bound set by Eq. (2.6). Here we will determine what performance is achievable with the method of maximum likelihood estimation.

Let us restrict our attention to the "typical" evolution expected for unitary or nonunitary evolution displayed in Eq. (1.3). Each cycle uses N replicas of some two-level system which starts in a pure eigenstate and evolves with time. At the end of each cycle a measurement is made to see how many of the systems will be found in the initial state and how many n have undergone a transition to the orthogonal state. Our naive estimate for the probability of transition will then be the relative frequency of transition, i.e., n/N. Then from Eq. (1.3) we may extract the length in time of the cycle. Here we want to determine the expected precision we can assign to this value; to do this we will assume that the method of data analysis is maximum likelihood estimation (see [9] for a similar problem). If p_0 is the "actual" value of this transition probability then the chance of finding n out of N transitions is given by the binomial distribution

$$\text{prob}(n \,|\, N, p_0) = \binom{N}{n} p_0^n (1-p_0)^{N-n}\,. \tag{6.1}$$

Now Bayes' theorem tells us that the probability p for the occurrence of a transition given that we have measured n transitions out of N is distributed according to

$$\text{prob}(p\,|\,n,N) \propto p^n(1-p)^{N-n}\,, \tag{6.2}$$

if we have no initial prejudice for the value of p. The most likely value for p (the maximum likelihood estimate inferred from our data) is found by elementary calculus to be $p_{\text{est}} = n/N$ (which agrees with our naive expectations above). We now want to determine our experimental uncertainty in this value. In particular, the noise figure we shall quote for this uncertainty will be based on the standard deviation of the distribution of Eq. (6.2) which was inferred from the measurements. Using the integral identity [10]

$$\int_0^1 dt\, t^k(1-t)^\ell = \left[(k+\ell+1)\binom{k+\ell}{k}\right]^{-1}\,, \tag{6.3}$$

we calculate the variance of p over Eq. (6.2) as

$$(\Delta p)^2 = \frac{(N-n+1)(n+1)}{(N+2)^2(N+3)}\,, \tag{6.4}$$

so Δp would be the experimentally quoted uncertainty for the observed probability.

We are interested here in the typical experimental results, so we calculate the average values for the "observed" probability and its uncertainty. With an overall probability p_0 for obtaining a transition given by Eq. (6.1) the average experimental *estimate* of the probability is just $\overline{p_{\text{est}}} = p_0$, and the average experimental estimate of the variance is

$$(\Delta p_{\text{est}})^2 \equiv \overline{(\Delta p)^2} = \frac{N(N-1)p_0(1-p_0) + (N+1)}{(N+2)^2(N+3)}\,. \tag{6.5}$$

We note that for $N \gtrsim 1/p_0(1-p_0)$ this expected uncertainty in the measured probability simplifies to

$$(\Delta p_{\text{est}})^2 \simeq \frac{p_0(1-p_0)}{N}\,. \tag{6.6}$$

This variance agrees with Fisher's theorem [11] for the asymptotic behavior of maximum likelihood estimation.

Alternately, for $N \lesssim 1/p_0(1-p_0)$ the variance is independent of p_0 with

$$(\Delta p_{\text{est}})^2 \simeq \frac{N+1}{(N+2)^2(N+3)}\,. \tag{6.7}$$

The reason for this independence is clear: suppose $p_0 \ll 1$, then for fewer than about $N \simeq 1/p_0$ replica systems almost every one will be found in the initial state; it is only when there is some non-negligible chance to see a transition that we can begin to distinguish the size of p_0 and hence the length of the cycle.

Figure 2 shows a log-log plot of the uncertainty Δp_{est} [solid line, from Eq. (6.5)] versus sample size N with $p_0 = 0.99$, and the asymptotic result [dashed line, from Eq. (6.6)] of "ordinary" counting statistics; we are able

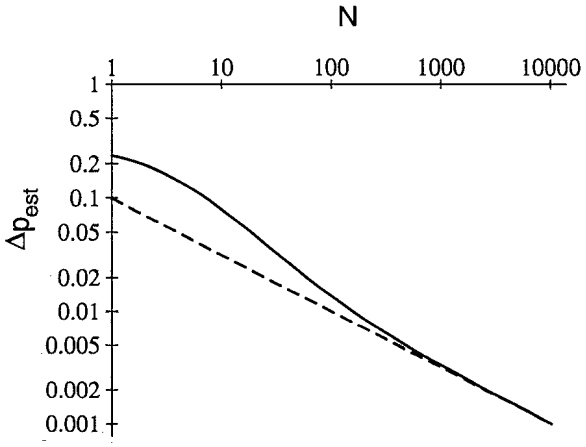

FIG. 2. Log-log plot showing the convergence of the statistical noise Δp_{est} [solid line, from Eq. (6.5)] with an increasing number of counts N for $p_0 = 0.01$; the ordinary $1/\sqrt{N}$ behavior of counting statistics [dashed line, from Eq. (6.6)] is attained only for $N \gtrsim 1/p_0(1-p_0) \simeq 100$.

to see a "knee" around $N \simeq 1/p_0(1-p_0) \simeq 100$ corresponding to where the asymptotic $1/\sqrt{N}$ behavior usually associated with counting statistics turns on.

After all this analysis how do the two types of typical unitary and nonunitary evolution in Eqs. (1.3) compare? For the short time unitary behavior described in Eqs. (1.3) we have

$$t \simeq \frac{1}{\omega}\sqrt{p_{\text{U}}}, \tag{6.8}$$

$$(\Delta t)^2_{\text{U}} \simeq \frac{1}{\omega^2}\frac{(\Delta p_{\text{U}})^2}{p_{\text{U}}}\,, \tag{6.9}$$

where the uncertainty in Eq. (6.9) follows from standard error analysis. Taking the number of replicas N to the "best" value (near the knee) $N = N_{\text{opt}} \simeq 1/p_{\text{U}}$ we have by Eq. (6.6)

$$(\Delta t)_{\text{U}} \simeq \frac{1}{\omega}\frac{1}{\sqrt{N}}\,, \tag{6.10}$$

i.e., the unitary evolution displays a sensitivity determined by the largest frequency of the system with the usual $1/\sqrt{N}$ improvement when operated optimally with $N \simeq 1/p_{\text{U}}$.

We now consider short time nonunitary behavior as described by Eq. (1.3), so now

$$t \simeq \frac{1}{\gamma}p_{\text{NU}}\,. \tag{6.11}$$

For a comparison to the unitary evolution we shall assume the two one-shot clocks, unitary and nonunitary, are being used to measure the cycle times t of equal duration and further that each has an equal number N of two-level systems; to deflate the advantage of the nonunitary clock we shall choose $N = N_{\text{opt}} \equiv 1/p_{\text{U}}$ so as to optimize the unitary clock's sensitivity.

We find that there are two regimes: $\gamma \lesssim \omega/\sqrt{N}$; and $\gamma \gtrsim \omega/\sqrt{N}$. For $\gamma \lesssim \omega/\sqrt{N}$ we see by equating Eqs. (6.8) and (6.11) that $p_{\text{NU}} \lesssim p_{\text{U}}$ so

$$N \lesssim \frac{1}{p_{\rm NU}} . \tag{6.12}$$

Now by Eqs. (6.7) and (6.11) we may use this result to calculate the sensitivity of the nonunitary one-shot clock; in this case it is always worse than the behavior expressed by Eq. (6.10). Instead, when $\gamma \gtrsim \omega/\sqrt{N}$ we find $p_{\rm NU} \gtrsim p_{\rm U}$ so

$$N \gtrsim \frac{1}{p_{\rm NU}}, \tag{6.13}$$

and now Eq. (6.6) applies. To summarize the behavior we find

$$(\Delta t)_{\rm NU} \simeq \begin{cases} \dfrac{1}{\gamma}\dfrac{1}{N} \gtrsim (\Delta t)_{\rm U} & \text{for } \gamma \lesssim \omega/\sqrt{N} , \\ \dfrac{1}{\sqrt{\gamma\omega}}\dfrac{1}{N^{3/4}} \lesssim (\Delta t)_{\rm U} & \text{for } \gamma \gtrsim \omega/\sqrt{N} , \end{cases} \tag{6.14}$$

assuming the number N of two-level systems is optimized for the unitary case with $N \equiv N_{\rm opt} \simeq 1/p_{\rm U}$.

A looser comparison can be made by taking $\gamma = \omega$ in which case we find

$$(\Delta t)_{\rm NU} \simeq \frac{(\Delta t)_{\rm U}}{N^{1/4}} \simeq \frac{1}{\omega}\frac{1}{N^{3/4}}, \tag{6.15}$$

with $N \simeq 1/p_{\rm U}$; and so our nonunitary one-shot clock is a simple device beating one-over-root-N behavior. We can do even better than this if we optimize the number of two-level systems for the nonunitary clock. Again taking $\gamma = \omega$ we now find

$$(\Delta t)_{\rm NU} \simeq \frac{(\Delta t)_{\rm U}}{N^{1/2}} \simeq \frac{1}{\omega}\frac{1}{N} , \tag{6.16}$$

having taken $N \simeq 1/p_{\rm NU}$.

VII. CONCLUSION

We have described here only the most preliminary investigation into quantum-limited clock design based on two-state systems. By using the formalism of statistical distance we were able to study the types of precision accessible by utilizing various elements. In this way, we found that there is a singular rate of change in the statistical distance as one moves away from the pure state boundary of the Bloch sphere under nonunitary evolution. A more detailed investigation into achievable performance suggests that the use of nonunitary evolving subsystems have significant advantages over purely unitary ones.

The advantage of our approach is that it shows us in which directions increased clock precision might be attainable, in principle. The disadvantages are that it does not consider the practicalities involved in any particular scheme.

[1] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[2] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).

[3] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982), Chap. VI.2.

[4] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, Berlin, 1983).

[5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976), Chap. VIII.4.

[6] S. L. Braunstein, C. M. Caves, and G. J. Milburn (unpublished).

[7] M. Hübner, Phys. Lett. A **163**, 239 (1992).

[8] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer-Verlag, Berlin, 1994).

[9] S. L. Braunstein and A. Mann, Phys. Rev. A **47**, R2427 (1993).

[10] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, London, 1980), p. 950.

[11] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946), pp. 500–504.