# Macroscopic Quantum Measurements of noncommuting observables

Tomer Jack Barnea,\* Marc-Olivier Renou,\* Florian Fröwis,† and Nicolas Gisin Group of Applied Physics, University of Geneva, 1211 Geneva 4, Switzerland (Dated: 1 October 2018)

Assuming a well-behaving quantum-to-classical transition, measuring large quantum systems should be highly informative with low measurement-induced disturbance, while the coupling between system and measurement apparatus is "fairly simple" and weak. Here, we show that this is indeed possible within the formalism of quantum mechanics. We discuss an example of estimating the collective magnetization of a spin ensemble by simultaneous measuring three orthogonal spin directions. For the task of estimating the direction of a spin-coherent state, we find that the average guessing fidelity and the system disturbance are nonmonotonic functions of the coupling strength. Strikingly, we discover an intermediate regime for the coupling strength where the guessing fidelity is quasi-optimal, while the measured state is almost not disturbed.

#### I. INTRODUCTION

In everyday life we continuously perform measurements. For instance, to locate our friends we perform some kind of position measurements; similarly to read this text. Presumably all this can be described with the quantum formalism. But, obviously, these measurements are not of the standard von Neumann projective kind. They are highly noninvasive while still collecting a large amount of information in a global, single shot. Additionally, from a physical point of view, we expect a fairly simple coupling between system and observer. We call measurements that fulfill these requirements "macroscopic quantum measurements". Our goal is to see if and how macroscopic quantum measurement can be realized.

For concreteness, think of a magnet whose N atoms are idealized by spin 1/2 particles and assume that all spins are aligned parallel to an unknown direction  $\vec{u} \in \mathbb{R}$ . By a freely selectable measurement, we are asked to estimate  $\vec{u}$ . We define the score function  $\cos^2\theta/2$ , where  $\theta$  is the angle between  $\vec{u}$  and the best guess  $\vec{w}$ . This task can be seen as a simultaneous measurement of the magnetization in the three principal directions x,y and z. The corresponding measurement operators are collective spin operators  $S_x = \frac{1}{2} \sum_i \sigma_x^{(i)}$  (and similarly for y and z). The maximal average guessing fidelity is  $\mathcal{F}_{\mathrm{opt}} = (N+1)/(N+2) = 1 - 1/N + O(N^{-2})$ , which is achievable either with nontrivial, entangling projective measurements or with a continuous general measurement with elements  $\Omega(\vec{r}) \propto |\vec{r}\rangle\langle\vec{r}|^{\otimes N}$  [1]. The same fidelity can be asymptotically reached by random measurements on individual spins [2].

Here, in order to realize the simultaneous measurement of x, y and z, we build upon von Neumann's measurement model by coupling all three observables  $S_x$ ,  $S_y$  and  $S_z$  to three independent pointers. Each pointer is then individually measured and the results are processed to guess  $\vec{w}$  [3]. (This is in the spirit of the Arthurs-Kelly

model [4], where position and momentum are coupled to two individual pointers.) In the strong coupling regime, this measurement corresponds to randomly choosing a direction  $\vec{n}$  and to strongly measure along this axis [3]. This measurement model leads to an average fidelity of 3/4 (i.e., far away from the optimal value) for large N [5].

The coupling strength between system and pointers influences the guessing fidelity and the measurement-induced disturbance. As shown by Poulin [6], the measurement of a single observable (e.g., only  $S_x$ ) is more informative and more disturbing as the coupling strength increases for a natural choice of the initial pointer state and the final measurement. For N parallel spins measured in any direction, it turns out that the system is asymptotically undisturbed if the coupling is much less than  $N^{-1/2}$  [6]. Poulin then suggested to perform many measurements in this weak coupling regime; potentially for several noncommuting observables. He conjectured that this should eventually give maximal information about the initial state without disturbing it. However, apart from practical questions, the limit of information

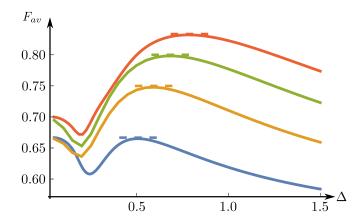


FIG. 1. Average fidelity  $\mathcal{F}_{\rm av}$  for guessing the unknown state as a function of the inverse coupling strength  $\Delta$  for  $N=1,\ldots,4$  (from bottom to top). The dashed lines correspond to the optimal value  $\mathcal{F}_{\rm opt}$ . For  $N\geq 2$ , the global maximum is reached by some nonzero  $\Delta$  and is close to the optimal value.

<sup>\*</sup> The first two authors contributed equally to this work.

<sup>†</sup> Electronic address: florian.froewis@unige.ch

gain and the measurement-induced disturbance of "infinitely many and infinitely weak" measurements is unclear so far. Other contributions, like Ref. [7], discuss the influence of the coupling strength, but do not compare to optimal strategies.

In this paper, we present an example of the surprising nonmonotonic relation between information gain, measurement-induced disturbance and the coupling strength when noncommuting observables are simultaneously measured. By studying the example of estimating the direction of parallel spins, we find an intermediate regime for the coupling strength where the average guessing fidelity is numerically very close to the optimal value  $\mathcal{F}_{\text{opt}}$ ; both for small and large N. In the very same coupling regime, the measurement-induced disturbance is small, in particular as N increases. More specifically, the averaged Bloch vector of the post-measured states is almost identical with the initial state, meaning that repeating the protocol by independent observers gives almost the same estimates. We claim that the quasi-optimal information gain and small measurement-induced disturbance is reminiscent of classical measurements. In this sense, the presented model is an example of a quantum measurement that behaves classically for large system sizes.

# II. GENERAL MEASUREMENT MODEL AND AVERAGE FIDELITY

In the following, we denote a normalized spin state orientated in the direction of a general vector  $\vec{x} \in \mathbb{R}^3$ by  $|\vec{x}\rangle \in \mathbb{C}^2$ . Generally, we consider a scheme where the system state  $|\vec{u}\rangle^{\otimes N}$  is coupled to a measurement apparatus  $|\phi\rangle$  -called pointer- through virtue of an interaction Hamiltonian  $H_{\text{int}}$ . Afterwards, the pointer is measured. This procedure is often referred to as the von Neumann measurement model [8–10]. For simplicity, the coupling with strength  $\mu$  is uniformly switched on for a unit time and dominates meanwhile all other processes; thus  $U = \exp(-i\mu H_{\rm int})$ . After this the joint state of system and pointer reads  $U|\phi\rangle\otimes|\vec{u}\rangle^{\otimes N}$ . The subsequent measurement of the pointer is modeled by a projective measurement with outcome  $\vec{r}$  for the eigenvectors  $|\psi_{\vec{r}}\rangle$ . (The dimension of  $\vec{r}$  depends on the pointer space.) This outcome is then classically processed. Knowledge about the initial pointer state, the coupling and the final measurement allows one to calculate the optimal guess state  $|\vec{w}_{\vec{r}}\rangle$ . So far, this is a fully general measurement procedure and is linked to the Kraus operator formalism via

$$E(\vec{r}) = \langle \psi_{\vec{r}} | U | \phi \rangle. \tag{1}$$

To measure the accuracy of the measurement procedure we choose the average fidelity  $|\langle \vec{w}_{\vec{r}}| \ \vec{u} \rangle|^2$ . The average score is then given by

$$\mathcal{F}_{\text{av}} = \int \frac{d\vec{u}}{4\pi} \int d\vec{r} \ p(\vec{r}|\vec{u}) \ |\langle \vec{w}_{\vec{r}}|\vec{u}\rangle|^2, \tag{2}$$

where  $p(\vec{r}|\vec{u}) = ||E(\vec{r})|\vec{u}\rangle^{\otimes N}||^2$  is the probability to get outcome  $\vec{r}$  given that the initial state was  $|\vec{u}\rangle^{\otimes N}$ .

A typical instance is a pointer with one spatial degree of freedom [6]. For example, the initial spatial wave function of the pointer is a Gaussian function with spread  $\Delta$ . In order to measure an observable A, one then defines  $H_{\rm int} = p \otimes A$ , where p represents the displacement operator in the pointer space, which is formally equivalent to the momentum operator. Thus, the coupling induces a momentum kick on the pointer whose strength depends on the initial system state. Finally, a position measurement of the pointer allows some inference about the system. Information gain and invasiveness of this procedure manifest themselves in the relationship between the coupling strength  $\mu$ , the spread of the Gaussian,  $\Delta$ , and the spectrum of the eigenvalues of the operator A. By redefining  $\Delta$ , one can set  $\mu = 1$  without loss of generality. Thus, the coupling strength of the measurement is proportional to  $\Delta^{-1}$ . A  $\Delta$  small compared to the spectral gap of A corresponds to a strong coupling that resolves the individual eigenvalues. The momentum kicks induced by two nearby eigenstates of A are distinguishable. On the other side, a  $\Delta$  large compared to the spectral gap of A means that system states prepared in neighboring eigenstates cannot be well distinguished. This is the regime of fuzzy measurements [10]. Note that the maximal information gain and the maximal disturbance are typically realized in the limit of a vanishing  $\Delta$ , that is, when the pointer model approaches the ideal projective measurement of A. Hence, even for large N, the basic idea of small invasiveness and large information gain does not seem to be reachable with this simple construction.

#### III. SPECIFIC MODELING

To measure the direction of  $|\vec{u}\rangle^{\otimes N}$ , a pointer with three spatial degrees of freedom seems to be more appropriate. We choose

$$|\phi\rangle = \frac{1}{(2\pi\Delta^2)^{3/4}} \int dx dy dz \ e^{-\frac{(x^2+y^2+z^2)}{4\Delta^2}} |x\rangle |y\rangle |z\rangle, \quad (3)$$

where x, y and z denote the coordinates of the pointer. The spatial distribution is a rotationally invariant Gaussian function with spread  $\Delta$ . The direction of  $|\vec{u}\rangle^{\otimes N}$  is determined by the three expectation values of the collective spin operators  $S_x, S_y$  and  $S_z$ . Thus, a classically inspired interaction Hamiltonian reads

$$H_{\rm int} = p_x \otimes S_x + p_y \otimes S_y + p_z \otimes S_z = \vec{p} \cdot \vec{S}, \quad (4)$$

which follows the pattern of the Arthurs-Kelly model. The  $p_k$  for k=x,y,z represent the displacement operators in the x,y and z directions, respectively. Finally, a position measurement with outcome  $\vec{r} \in \mathbb{R}^3$  is performed on the pointer. We simply guess  $|\vec{w}_{\vec{r}}\rangle = |\vec{r}\rangle[11]$ . The model is rotationally symmetric. Without loss of generality, we hence consider a fixed input state  $|\vec{u}\rangle = |\hat{e}_z\rangle \equiv |\uparrow\rangle$ .

With  $p(\vec{r}) \equiv p(\vec{r}|\hat{e}_z)$ , Eq. (2) then simplifies to

$$\mathcal{F}_{\rm av} = 2\pi \int r^2 \sin\theta dr d\theta p(\vec{r}) \cos^2\theta/2, \tag{5}$$

since  $p(\vec{r})$  does not depend on the azimuthal angle.

In the limit  $\Delta \to 0$ , the measurement model is suboptimal and the average fidelity is  $\mathcal{F}_{\rm av} = \frac{1}{4}(3N+2)/(N+1)$  for even N and  $\mathcal{F}_{\rm av} = \frac{1}{4}(3N+5)/(N+2)$  for odd N [3]. These values are bounded by  $\frac{3}{4}$  and are hence far away from the optimal fidelity. As we will show now, the three spatial dimensions of the pointer combined with noncommuting operators in Eq. (4) involve a rich behavior for nonvanishing  $\Delta$ . In contrast to the one-dimensional pointer, information gain and disturbance are not monotonically related to the coupling strength.

#### IV. SMALL NUMBER OF SPINS

We start by calculating the average fidelity  $\mathcal{F}_{\rm av}$  for one to four spins as a function of the spread  $\Delta$  of the Gaussian pointer. In these cases the expression in Eq. (5) is easy to work with due to the small number of spins. The Kraus operators  $E(\vec{r})$  of the interaction, Eq. (1), can be derived analytically. The average fidelity is analytically calculated in the case of one single qubit, while for more spins the integration over the radial component of  $\vec{r}$  is performed numerically. The results are shown in Fig. 1. For two or more spins the optimal  $\Delta$  is clearly distinct from zero and  $\mathcal{F}_{\rm opt}$  (dashed lines in Fig. 1) can almost be reached.

Analyzing the graph shown in Fig. 1 in more detail reveals three distinct regions. For very small values of  $\Delta$  (i.e., in the strong coupling regime) the limits predicted in [3] are recovered, yielding an average fidelity of 2/3 for N=1 and 2 and  $\mathcal{F}_{av}=7/10$  for N=3 and N=4, respectively. On the other extreme, for  $\Delta>1$ , we see that the average fidelity starts to decline rapidly. This can be understood if one notices that after a certain value of  $\Delta$  the coupling is so weak that the procedure is essentially equivalent to randomly guessing, therefore yielding an average fidelity of 1/2. The intermediate region is particularly intriguing because in the case of two and more spins the average fidelity is superior to what can be achieved when  $\Delta \to 0$ . This means that in this case a lesser coupling strength can achieve better results than a strong coupling.

## V. LARGE NUMBER OF SPINS

For N>4 the full calculation becomes cumbersome without approximations. To continue with larger N, we calculate a lower bound on  $\mathcal{F}_{\rm av}$ . For this, we resolve the identity  $\mathbb{1}=|\vec{r}\rangle\langle\vec{r}|^{\otimes N}+(\mathbb{1}-|\vec{r}\rangle\langle\vec{r}|^{\otimes N})$  and insert it in the expression for  $p(\vec{r})$ . The reason for this choice is that  $E(\vec{r})$  is diagonal in the eigenbasis of  $\vec{r}\cdot\vec{S}$  and the Kraus

operator leading to the optimal  $\mathcal{F}_{av}$  should have a large overlap with  $|\vec{r}\rangle\langle\vec{r}|^{\otimes N}$  [12]. In the following, we bound  $p(\vec{r}) \geq |\langle\vec{r}|^{\otimes N} E(\vec{r})|\uparrow\rangle|^{2N}$ .

For general N, one has to find further simplifications for  $p(\vec{r})$ . For an  $\mathcal{F}_{\rm av}$  scaling as 1-O(1/N) (i.e., the optimal scaling), it is necessary that  $|\langle \vec{r}| \uparrow \rangle|^2 = 1-O(1/N)$  for almost all  $\vec{r}$  which have O(1) support from  $p(\vec{r})$ . Hence, the most important  $|\vec{r}\rangle$  are  $O(1/\sqrt{N})$ -close to  $|\uparrow\rangle$ . Inspecting the matrix element  $E_{\vec{r}} = \langle \vec{r}|^{\otimes N} E(\vec{r}) |\uparrow\rangle^{\otimes N}$  in the momentum basis of the pointer,

$$E_{\vec{r}} \propto \int d\vec{p} e^{i\vec{r}\cdot\vec{p}} e^{-\Delta^2 p^2} \langle \vec{r} |^{\otimes N} e^{-i\vec{p}\cdot\vec{S}} | \uparrow \rangle^{\otimes N}, \qquad (6)$$

one notices that this implies  $\vec{p} = O(1/\sqrt{N})$ , which in turn means that  $\Delta = O(\sqrt{N})$ . In this regime, it is advantageous to perform a Holstein-Primakoff transformation [13] to express  $S_+ = S_x + iS_y$  as  $S_+ = \sqrt{1-a^\dagger a/N}a$  with  $[a^\dagger,a]=1$ ; similarly for  $S_- = S_+^\dagger$  and  $S_z = \frac{1}{2}[S_+,S_-] = \frac{N}{2} - a^\dagger a$ . The input state  $|\uparrow\rangle^{\otimes N}$  is transformed to the vacuum  $|0\rangle$ , which is an eigenstate of a with eigenvalue 0. The state  $|r\rangle^{\otimes N}$  corresponds to a coherent state, since rotations on the Bloch sphere are approximated by planer displacements. In summary, we have the mapping  $\langle \vec{r}|^{\otimes N} e^{-i\vec{p}\cdot\vec{S}} |\uparrow\rangle^{\otimes N} \to \langle 0| \exp[f(a,a^\dagger)] |0\rangle$  with a lengthy function f in a and  $a^\dagger$ .

The next step is to find the optimal value of  $\Delta$ . For this, we write  $\exp[f(a,a^{\dagger})]$  in a Taylor series for  $N\to\infty$ . Keeping terms up to order O(1), one can analytically calculate  $\mathcal{F}_{\rm av}$  and find a maximum for  $\Delta_{\rm opt}=\sqrt{N/8}$ . Heuristic arguments indicate that the following results are asymptotically unchanged by adding  $o(\sqrt{N})$  terms to  $\Delta_{\rm opt}$ .

In order to have more precise results for  $\mathcal{F}_{\rm av}$ , we take into account terms up to order O(1/N) in the power series of  $\exp[f(a,a^{\dagger})]$ . We are still able to analytically integrate over  $\vec{p}$  in Eq. (6) and the radial part of  $\vec{r}$  in Eq. (5). The final integration over  $\theta$  is done numerically. We find that  $\mathcal{F}_{\rm av} \gtrsim 1 - \epsilon_N/N$ , where  $\epsilon_N$  seems to asymptotically converge to a value slightly above one (see Fig. 2). This indicates that the approximate lower bound does asymptotically not coincide with  $\mathcal{F}_{\rm opt} = 1 - 1/N + O(N^{-2})$ . However, the relative difference between the two clearly goes to zero. Moreover, taking into account more matrix elements in addition to Eq. (6) could reduce the gap.

## VI. MEASUREMENT-INDUCED DISTURBANCE

Interestingly, it turns out that the presented measurement scheme is hardly invasive. To show this, we calculate the quantum fidelity between the pre- and post-measured state averaged over all measurement outcomes. The post-measured state reads  $\rho_{\rm post}={\rm Tr}_{\rm pointer}U\,|\phi\rangle\langle\phi|\otimes|\uparrow\rangle\langle\uparrow|^{\otimes N}\,U^{\dagger}.$  The disturbance is then defined to be  $D=1-\langle\uparrow|^{\otimes N}\,\rho_{\rm post}\,|\uparrow\rangle^{\otimes N}.$  For calculating the trace

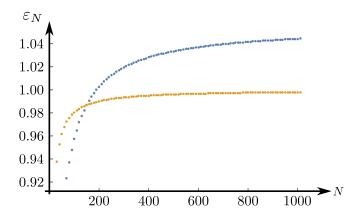


FIG. 2. Scaling factor  $\epsilon_N \lesssim N(1-\mathcal{F}_{\rm av})$  from the approximate lower bound on  $\mathcal{F}_{\rm av}$  (upper, blue curve) compared to the optimal scaling factor  $N(1-\mathcal{F}_{\rm opt})=(1+2/N)^{-1}$  (lower, orange curve). For large N,  $\epsilon_N$  seems to stay slightly above one, which indicates a nonoptimal lower bound. For N<150, the approximation clearly does not hold.

of  $\rho_{\rm post}$  it is most convenient to work in the momentum basis of the pointer space. Then, simple manipulations give

$$D = 1 - \int d^3p \, |\phi(\vec{p})|^2 \left[ 1 - \sin^2(p/2) \sin^2 \theta \right]^N.$$
 (7)

For Gaussian pointers, this can in principle be calculated for any N.

For small N, we observe that the maximal D is found for some  $\Delta > 0$ , which does however not correspond to the  $\Delta$  leading to the maximal information (see Fig. 3).

For  $N \gg 1$ , a closed formula is only found for an approximated integrand. Similar as before, we set  $\Delta = c\sqrt{N/8}$  and do a Taylor series for  $N \to \infty$ . In lowest order, one finds the Lorentzian-like function  $D = (1+c^2)^{-1}$ , which does not indicate any particular behavior at c=1, for which  $\mathcal{F}_{\rm av}$  is maximal in the large N limit. Higher orders exhibit simple, but long expressions in c, which simplify for c=1 to

$$D = \frac{1}{2} + \frac{23}{1440N^2} + O\left(\frac{1}{N^3}\right). \tag{8}$$

Hence, in the limit of large N, the disturbance is close to the minimal disturbance  $D_{\min} = (N+1)/(2N+1)$  in the case of optimal guessing fidelity [14]. A nonunit value seems to indicate a rather severe change in the state. We note, however, that the overlap between two states  $|\vec{u}\rangle^{\otimes N}$  and  $|\vec{w}\rangle^{\otimes N}$  generally decays exponentially in N unless the angle between  $\vec{u}$  and  $\vec{w}$  is at most  $O(1/\sqrt{N})$ . A constant overlap for large N indicates an  $O(1/\sqrt{N})$  closeness between  $|\uparrow\rangle^{\otimes N}$  and  $\rho_{\rm post}$  on the Bloch sphere.

This becomes even more evident when we consider the Bloch vector of the spin ensemble before and after measurement. The expectation value  $\langle \vec{S} \rangle$  for both states can be calculated in a straightforward manner. One has

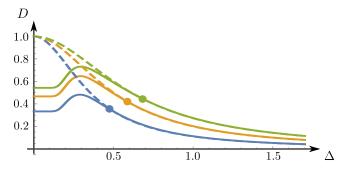


FIG. 3. Disturbance measured by the quantum fidelity between the pre- and post-measured states as function of  $\Delta$  for N=1,2,3 (from bottom to top). The solid lines correspond to the exact expression, Eq. (7), while the dashed curves are the lowest-order approximations for large  $N,\,D\approx \left(1+8\Delta^2/N\right)^{-1}$ . The points on the curves indicate where  $\Delta=\Delta_{\rm opt}.$ 

 $\langle S_x \rangle = \langle S_y \rangle = 0$  for both  $|\uparrow\rangle^{\otimes N}$  and  $\rho_{\rm post}$ . For the z component, we find  $\langle S_z \rangle_{|\uparrow\rangle^{\otimes N}} = N/2$  and  $\langle S_z \rangle_{\rho_{\rm post}} = \frac{1}{6}N\left[1 + e^{-1/(8\Delta^2)}\left(2 - 1/(2\Delta^2)\right)\right] = N/2 - 1 + O(1/N)$ , where the last expression holds for  $\Delta = \Delta_{\rm opt}$ . In words, the length of the post-measured Bloch vector is only minimally reduced. A repetition of the measurement will give almost the same information about the Bloch vector than the first one. This is in stark contrast to some other optimal measurement strategies. For example, in Ref. [2], the single-spin measurements in a random direction leaves behind a completely depolarized product state; therefore making it impossible for a different pointer to gain any information about the initial state in a subsequent measurement.

#### VII. DISCUSSION AND OPEN QUESTIONS

In summary, the simultaneous measurement of the noncommuting observables  $S_x, S_y$  and  $S_z$  can be made highly sensitive and hardly disturbing in a regime where the coupling to the pointers is relatively weak. This holds true in particular for large spin ensembles. Via an approximate Trotter expansion, our results also guide a quantitative analysis of proposals [6] for sequential measurements of noncommuting observables in the weak coupling regime.

Given the features of the model together with a fairly simple scheme, the discussed model is reminiscent of a classical measurement. This observation allows us to speculate about a general class of "macroscopic quantum measurements" as measurements within the quantum formalism with high information gain, low disturbance and simple physical pointer models. In other words, the present model could be a starting point to understand the emergence of classical behaviour of quantum measurements for large systems. It might be interesting to

study a connection to recent work on classifying different levels of simultaneous measurements [15]. Further research will be dedicated to other schemes of simultaneous measurements as a function of coupling strength, in particular when the observables have well defined macroscopic limits such as position and momentum.

It is likely that the very same measurement model can also be used to illustrate the classical appearance of macroscopic quantum states by changing the input state. Suppose a superposition of macroscopically distinct states (i.e., a Schrödinger-cat state) [16] is perfectly prepared. We expect that the coherence between these distinct states cannot be witnessed if the system is subject to a coupling as in Eq. (4). This reflects the basic idea of einselection [17], where the measurement apparatus takes the role of the environment. In addition, it

would be interesting to see whether the presented coupling could also serve as an illustrating example of quantum Darwinism [18].

Further open questions include a better characterization of  $\mathcal{F}_{\rm av}$  for  $N\gg 1$  in terms of the precise scaling with N as well as the robustness with respect to variations of  $\Delta_{\rm opt}$ . Also, more realistic initial states, that is, ones that better match the properties of real-world examples, are interesting to study. A first attempt would include nonmaximally polarized and thermal states.

Acknowledgments.— We thank Pavel Sekatski and Serge Massar for stimulating discussions. This work was supported by the National Swiss Science Foundation (SNSF), the Austrian Science Fund (FWF), grant number J3462, the COST Action No. MP1006 and the European Research Council (ERC MEC).

- S. Massar and S. Popescu, Phys. Rev. Lett. 74, 1259 (1995).
- [2] E. Bagan, A. Monras, and R. Muñoz-Tapia, Phys. Rev. A 71, 062318 (2005).
- [3] G. D'Ariano, P. L. Presti, and M. Sacchi, Physics Letters A 292, 233 (2002).
- [4] E. Arthurs and J. L. Kelly, Bell System Technical Journal 44, 725 (1965).
- [5] The average fidelity of 3/4 is exactly the same value as estimating a classical spin with a Stern-Gerlach-type experiment with randomly chosen measurement axis.
- [6] D. Poulin, Phys. Rev. A 71, 022102 (2005).
- [7] M. Bußhardt and M. Freyberger, Phys. Rev. A 82, 042117 (2010).
- [8] J. von Neumann, Mathematical Foundations of Quantum Mechanics (Princeton University Press, Princeton, New Jersey, 1955).
- [9] P. Busch, P. J. Lahti, and P. Mittelstaedt, in *The Quantum Theory of Measurement*, Lecture Notes in Physics

- Monographs No. 2 (Springer Berlin Heidelberg, 1991) pp. 27-98
- [10] A. Peres, Quantum Theory: Concepts and Methods (Kluwer Academic Publishers, New York, 2002).
- [11] Note that this choice is not always optimal. For some (nonoptimal) values of  $\Delta$ , the optimal choice is  $|\vec{w}_{\vec{r}}\rangle = |-\vec{r}\rangle$ . For symmetry reasons the optimal guess has to lie on the axis of  $\vec{r}$ .
- [12] M.-O. Renou et al., in prep.
- [13] T. Holstein and H. Primakoff, Phys. Rev. 58, 1098 (1940).
- [14] M. F. Sacchi, Phys. Rev. A 75, 012306 (2007).
- [15] T. Heinosaari, Phys. Rev. A 93, 042118 (2016).
- [16] A. O. Caldeira and A. J. Leggett, Phys. Rev. Lett. 46, 211 (1981).
- [17] W. H. Zurek, Rev. Mod. Phys. **75**, 715 (2003).
- [18] W. H. Zurek, Nat Phys 5, 181 (2009).