

On the power of weak measurements in separating quantum states

Boaz Tamir · Eliahu Cohen · Avner Priel

Received: 16 November 2014 / Accepted: 8 January 2015 / Published online: 24 January 2015
© Chapman University 2015

Abstract We investigate the power of weak measurements in the framework of quantum state discrimination. First, we define and analyze the notion of weak consecutive measurements. Our main result is a convergence theorem whereby we demonstrate when and how a set of consecutive weak measurements converges to a strong measurement. Second, we show that for a small set of consecutive weak measurements, long before their convergence, one can separate close states without causing their collapse. We thus demonstrate a tradeoff between the success probability and the bias of the original vector towards collapse. Next, we use post-selection within the two-state vector formalism and present the non-linear expansion of the expectation value of the measurement device's pointer to distinguish between two predetermined close vectors.

Keywords Weak measurements · Quantum state discrimination · Two-state vector formalism

1 Introduction

Weak measurement [1] has already been proven to be very helpful in several experimental tasks [2–5], as well as in revealing fundamental concepts [6–10]. Tasks traditionally believed to be self-contradictory by nature such as determining a particle's state between two measurements prove to be perfectly possible with the aid of this technique. Within the framework of the two-state vector formalism (TSVF), weak measurements reveal several new and sometime puzzling phenomena. For a general discussion on weak measurements see [1, 11–14]. In this paper, we analyze the strength of weak measurements by addressing the question of quantum state discrimination.

There are several known methods for quantum state discrimination, i.e., for the task of deciding which vector was chosen out of a predetermined set of (possibly close) state vectors (for a general review see [15]). Discriminating

B. Tamir
Faculty of Interdisciplinary Studies, Bar-Ilan University, Ramat-Gan, Israel
e-mail: boaz_tamir@post.bezalel.ac.il

E. Cohen (✉)
School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
e-mail: eliahuco@post.tau.ac.il

A. Priel
Department of Physics, University of Alberta, Edmonton, AB, Canada
e-mail: apriel@ualberta.ca

between predetermined non-orthogonal vectors can be seen as quantum hypothesis testing [15]. Suppose we use projective measurements, then the question is what are the best projections to choose so that the error probability would be as small as possible. This was first discussed by Helstrom in [16] for the case of two predetermined vectors. The density matrix version of the problem was developed by Osaki in [17]. A different scheme was presented by Ivanovic [18] where the discrimination is error free, but there is a probability for obtaining a non-conclusive result, i.e., a ‘don’t know’ result. A variant of this scheme was suggested in [19, 20] where the inconclusive outcomes have a fixed rate. Recently, Zilberberg et al. [21] have presented a new method, based on partial measurements followed by post-selections. The problem of state discrimination and the problem of cloning are deeply connected. A recent paper of Yao et al. [22] discusses approximate cloning and probabilistic cloning. Weak measurements were also proposed for the task of discrimination between two very close states [23]. The scheme we suggest here in Sect. 2 is similar, yet somewhat more general, while the scheme we use at Sect. 1, based on an sequential weak measurements, is quite different. The advantages and disadvantages of each method will be discussed.

In Sect. 1, we discuss the orbit of a two-dimensional state vector under the set of transformations induced by weak measurements. This process can be described as a biased Gaussian random walk on the unit circle. The probability amplitude that govern the next step (the ‘coin’ probability amplitude) is changing with each step. In this, our walk is similar to the one presented in [24], but differs from the quantum random walk presented in [25]. We thereby rotate two different vectors in opposite directions (therefore, in a non-unitary way), where we can use a single (strong) projective measurement to distinguish between them. We show that using enough weak measurements (the number of which is a function of the weak coupling), the overall success probability converges to the known optimal result for discrimination by projective measurements [16].

Next, we use a different approach; we reduce the number of measurements, thus compromising the success probability, however, gaining an advantage by avoiding the collapse.

In Sect. 2, we apply the TSVF of weak measurements. By choosing the right Hermitian operator and a proper post-selection, we can get imaginary weak values. Imaginary weak values are best suited for the analysis of the coordinate variable of the measurement space. We apply a non-linear expansion of the weak value and use it to compute the first and second moments of such a variable. These moments of the coordinate variable change as functions of the initial vector. We then pick two state vectors maximizing the difference between the two distributions of the coordinate variable.

2 A cloning protocol using iterative weak measurements

In this Section, we shall perform consecutive weak measurements (without post-selection) to show that it is possible in principle to differentiate between two non-orthogonal vectors. Suppose Alice is sending Bob one of two predetermined state vectors of a two-dimensional system S :

$$|\psi_1\rangle = \cos \phi |0\rangle + \sin \phi |1\rangle$$

$$|\psi_2\rangle = \sin \phi |0\rangle + \cos \phi |1\rangle$$

where $|0\rangle$ and $|1\rangle$ are the eigenvalues of S_z and $\frac{\pi}{2} > \phi > \frac{\pi}{4}$. Let θ be the angle between the two vectors:

$$\langle \psi_1 | \psi_2 \rangle = \cos \theta,$$

therefore, the two vectors have the same angle $\frac{\theta}{2}$ with respect to the vector $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ (see Fig. 1). We assume Alice is sending each of the vectors with the same probability. It is well known by [16] that the maximal success probability is:

$$P_S(\text{opt}) = \frac{1}{2} \left(1 + \sqrt{1 - 4\lambda_1\lambda_2|\langle \psi_1 | \psi_2 \rangle|^2} \right)$$

where Alice is sending $|\psi_1\rangle$ (resp. $|\psi_2\rangle$) with probability λ_1 (resp. λ_2). Since, we are using $\lambda_1 = \lambda_2 = 1/2$ we can write:

$$P_S(\text{opt}) = \frac{1 + \sin \theta}{2} = \cos^2 \phi \tag{1}$$

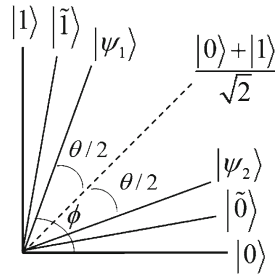


Fig. 1 The choice of axes, vectors and angles that is used throughout the paper

Below we shall show that one can reach the same limit using the following weak measurement protocol. By a series of weak measurements Bob will be able to ‘rotate’ the initial vector towards the direction of $|0\rangle$ or $|1\rangle$. Bob will stop the rotations after a predetermined number of iterations, by then he can assume with high probability that the vector would have crossed $|\tilde{0}\rangle$ or $|\tilde{1}\rangle$ which are close to the axes $|0\rangle$ or $|1\rangle$. He will then strongly measure the final ‘rotated’ vector in the standard S_z basis to get the result $|0\rangle$ or $|1\rangle$. If the result is $|1\rangle$, he can conclude the initial vector was $|\psi_1\rangle$, otherwise it was $|\psi_2\rangle$.

The error probability has two factors; the first originates from the weak ‘rotations’, i.e., the probability that the weak ‘rotations’ will take $|\psi_1\rangle$ (resp. $|\psi_2\rangle$) to $|\tilde{0}\rangle$ (resp. $|\tilde{1}\rangle$). The second factor originates from the strong final measurement, i.e., the probability that having ‘rotated’ $|\psi_1\rangle$ (resp. $|\psi_2\rangle$) in the correct direction towards $|1\rangle$ (resp. $|0\rangle$) the strong measurements will produce $|0\rangle$ (resp. $|1\rangle$) results.

Below we start by describing the process of weak measurement. Then, we describe the protocol in details.

2.1 A Gaussian-type random walk induced by weak coupling

Let S denote our two-dimensional system to be measured. Let \hat{S}_z be the Pauli spin matrix on the system S . Let $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ be a state vector in the eigenbasis $|0\rangle$ and $|1\rangle$ of \hat{S}_z .

Let $|\phi\rangle$ denote the wave function of a quantum measurement device. Then,

$$|\phi\rangle = |\phi(x)\rangle = \int_x \phi(x)|x\rangle dx \tag{2}$$

where $\hat{X}|x\rangle = x|x\rangle$ is the position operator of the measuring needle. Suppose $|\phi(x)|^2$ is normally distributed around 0 with variance σ^2 :

$$\phi(x) = (2\pi\sigma^2)^{-1/4} e^{-x^2/4\sigma^2}$$

The function $\phi(x)$ represents the device’s ‘needle’ distribution amplitude. Let \hat{P} be the momentum conjugate operator of the measuring device, such that $[\hat{X}, \hat{P}] = i\hbar$.

We shall start the measuring process with the vector:

$$|\psi\rangle \otimes |\phi(x)\rangle$$

in the tensor product space of the two systems. We will now couple the two systems by the interaction Hamiltonian \hat{H}_{int} :

$$\hat{H} = \hat{H}_{int} = g(t)\hat{S}_z \otimes \hat{P} \tag{3}$$

where $g(t)$ is the coupling function satisfying:

$$\int_0^T g(t)dt = g,$$

and T is the coupling time. We will use $g = 1$ throughout this section for simplicity.

Following the weak coupling the system and the measuring device are entangled:

$$\int_x [\alpha|0\rangle \otimes \phi(x-1) + \beta|1\rangle \otimes \phi(x+1)]|x\rangle dx \quad (4)$$

where the above functions $\phi(x \pm 1)$ are two normal functions with high variance, overlapping each other.

We can write the entangled (unnormalized) state of the measured vector and the measurement device as:

$$\int_x \left[e^{-\frac{(x-1)^2}{4\sigma^2}} \alpha|0\rangle \otimes |x\rangle + e^{-\frac{(x+1)^2}{4\sigma^2}} \beta|1\rangle \otimes |x\rangle \right] dx. \quad (5)$$

We will now strongly measure the needle. Suppose the needle collapses to the vector $|x_0\rangle$, then our system is now in the state:

$$\left[e^{-\frac{(x_0-1)^2}{4\sigma^2}} \alpha|0\rangle + e^{-\frac{(x_0+1)^2}{4\sigma^2}} \beta|1\rangle \right] \otimes |x_0\rangle. \quad (6)$$

The eigenvalue x_0 could be anywhere around -1 or 1 , or even further away, especially if σ is big enough, i.e., when the measurement is very weak. Note that the collapse of the needle biases the system's vector. However, if σ is very large with respect to the difference between the eigenvalues of \hat{S}_z then the bias will be very small and the resulting system's vector will be very similar to the original vector.

Note that, the projective measurement on the outer needle's space induces a unitary evolution on the inner space of the composite system. Thus, we control the evolution of the state by weak measurements. This resembles an adiabatic evolution of a state vector by strong measurements (see [26]).

Consider now the orbit of the initial state vector under the series of weak measurements. We couple the particle to the measuring device and then measure the needle. Next, we couple the biased vector to another measuring device and measure its needle. We repeat this process over and over again. The measuring needle is re-calibrated after each measurement, while the particle's state accumulates the successive biases one by one. The series of biased vectors describes an asymmetric random walk on the circle.

Note that, the random walk is continuous and 'weighted' in the sense that the distribution function for the next sampling step is changing as a function of the location on the circle. Near the axes $|0\rangle$ and $|1\rangle$ it looks like a Gaussian random walk. The following protocol goes through such a random walk trying to identify the point of start.

2.2 Distinguishing by consecutive iterations of weak measurements

Bob will perform a series of weak measurements to rotate the initial vector $|\psi\rangle$ towards $|0\rangle$ or $|1\rangle$. We use a numerical simulation to investigate the random walk. The vectors $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$ are very close to $|0\rangle$ and $|1\rangle$, respectively, and they will define the 'collapse'.

First, we address the task of quantifying the number of weak measurements needed to 'collapse' the initial vector as a function of the standard error σ of the needle. We simulate the probability distribution of the number of weak measurements needed to collapse the initial vector $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ (see Appendix). We start with $\sigma = 20$. The probability distribution is best fitted ($R^2 > 0.99$) by a log-normal distribution with $\tilde{\mu} = 2.8$ and $\tilde{\sigma} = 0.71$, see Fig. 2. When examining the form of q_m in the Appendix, and applying the Central Limit Theorem under weak dependence, the success of the fit turns obvious: q_m can be described approximately as an exponent of a normally distributed random variable and hence behaves like a log-normal random variable.

Next, we look at the median of the above results (which distribute log-normally) as a function of σ , see Fig. 3. By changing σ we observe a quadratic relation between the median and σ . In other words, Bob needs $O(\sigma^2)$ steps before he knows the vector had 'collapsed' to $|\tilde{0}\rangle$ or $|\tilde{1}\rangle$ with high probability. In this simulation, we chose $|\tilde{0}\rangle = \cos(10^\circ)|0\rangle + \sin(10^\circ)|1\rangle$; $|\tilde{1}\rangle = \cos(80^\circ)|0\rangle + \sin(80^\circ)|1\rangle$. The result is independent of the initial vector. Rotating $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$ toward the axes will only multiply the number of steps needed for the collapse by a constant factor.

Fig. 2 Distribution of the number of measurements until the collapse, given a fixed σ

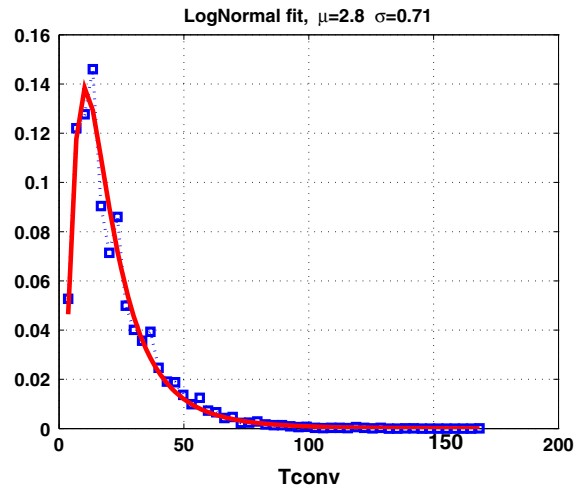
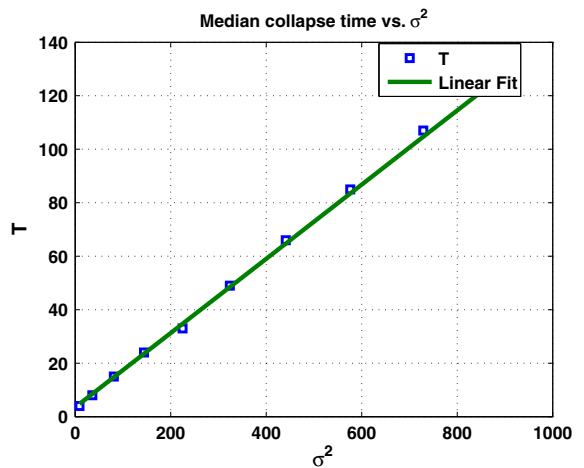


Fig. 3 Average number of measurements until the collapse as a function of σ



Second, we address the error probability in determining the vector’s original identity. Having started with $|\psi_1\rangle$, the error probability $Err(|\psi_1\rangle)$ is:

$$Err(|\psi_1\rangle) = P^w(|\psi_1\rangle \rightarrow |\tilde{0}\rangle)P^s(|\tilde{0}\rangle \rightarrow |0\rangle) + P^w(|\psi_1\rangle \rightarrow |\tilde{1}\rangle)P^s(|\tilde{1}\rangle \rightarrow |0\rangle).$$

The error probability for $|\psi_1\rangle$ is the probability that the series of weak measurement iterations will take $|\psi_1\rangle$ to $|\tilde{0}\rangle$ and the strong measurement will take $|\tilde{0}\rangle$ to $|0\rangle$, plus the probability that the weak process will take $|\psi_1\rangle$ correctly to $|\tilde{1}\rangle$, but the strong measurement will take $|\tilde{1}\rangle$ to $|0\rangle$.

In the next simulation, we compute the probability $P^w(|\psi_1\rangle \rightarrow |\tilde{1}\rangle)$ and $P^w(|\psi_2\rangle \rightarrow |\tilde{0}\rangle)$ for a set of initial vectors $|\psi_i\rangle$ where $|\tilde{0}\rangle = \cos 1^\circ|0\rangle + \sin 1^\circ|1\rangle$ and $|\tilde{1}\rangle = \cos 89^\circ|0\rangle + \sin 89^\circ|1\rangle$. Figure 4 presents the success probability as a function of θ . The simulation was performed 1000 times (see the pseudo-code in the appendix) where each time we followed the trajectory of the random walk until it crossed the boundaries defined by $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$. The success probabilities are higher than the best separation value in [16]. However, these simulations disregard the possible error in the strong measurements. As we increase the angle between $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$, we reduce the error of the strong measurement and the success probabilities of the weak process approach the limit in [16] from above.

The following conclusion is only natural:

$$Err(|\psi_1\rangle) = \sin^2 \alpha.$$

We can demonstrate the conjecture by extending the angle between $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$.

Fig. 4 The success probability $P^w(|\psi_1\rangle \rightarrow |\tilde{1}\rangle)$ as a function of θ . The *solid curve* describes the optimal success probability for projective measurements, $P^s(|\psi_1\rangle \rightarrow |1\rangle)$. Note that the success probabilities are with respect to different outcome vectors

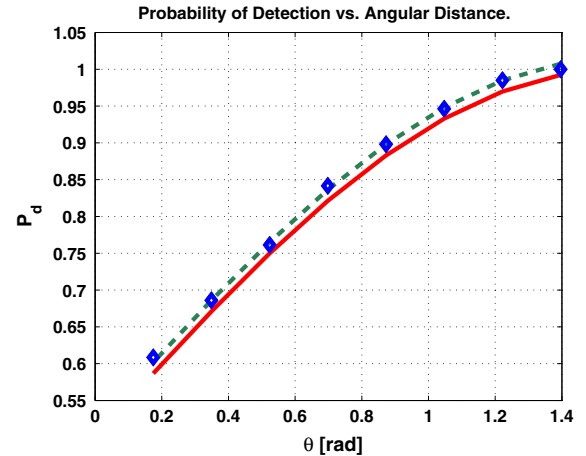
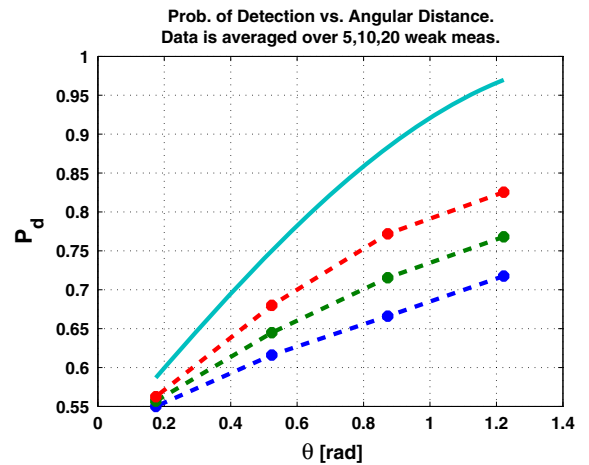


Fig. 5 Success probability for hypothesis testing with low number of weak measurements. The *solid curve* describes the optimal success probability for projective measurement (see Eq. 1 above)



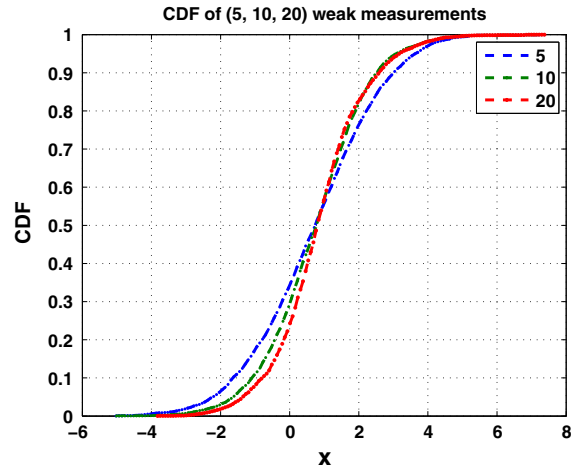
2.3 Hypothesis testing with weak measurements

So far, we have used the weak measurements to iteratively produce small biases of the initial vector to shift it towards one of the axes. The results we got for the pointer of the weak measurement apparatus were so far ignored. Suppose now we use a very small number of weak measurements. We could use the pointers' readings as samples from the vector's distribution. Moreover, we can average over the few values and use the result as a statistic. Notice, however, that we are sampling from a distribution that is changing following each pointer's reading. The advantage of such a protocol lies in the fact that the vector has not collapsed; if the standard error of the weak measurement is large and the number of weak measurement is small then the resulting vector is still in the neighborhood of the original one. We will show that the distributions of the averages behave as a function of the initial vector and, therefore, could be used to distinguish between the two.

In the next simulation (Fig. 5) we weakly measured the vectors for 5, 10, and 20 times, using $\sigma = 3$. This simulation was performed 5,000 times for several different angles θ (as in Fig. 1). It is expected that the average value (of weak measurements) for $|\psi_1\rangle$ (resp. $|\psi_2\rangle$) should be below 0 (resp. above 0). The success probability was computed by the number of times the average value did not cross 0.

Figure 6 describes the cumulative distribution function for the above average values (denoted by x) for the initial angle $\theta = 50^\circ$ and initial vector $|\psi_2\rangle$. It can be seen that the graphs are similar to a shifted cumulative

Fig. 6 Cumulative distribution function for the average of weak values for the initial angle $\theta = 50^\circ$ and initial vector $|\psi_2\rangle$



distribution function of a normal random variable. Moreover, the median is fixed regardless of the number of weak measurements. This median is supposed to coincide with the theoretical ‘strong’ one when σ becomes lower.

In practice, if we use such a protocol to distinguish between two vectors (separated by an angle θ , as in Fig. 1) we should be aware of two types of errors. The first comes from the fact that the vector is changing with each weak measurement. It could possibly drift following the first measurement towards the other vector, thereafter staying in that neighborhood for long. Thus, our readings will identify the wrong vector. The second type of error comes from hypothesis testing. The sample of the average could be very close to 0 making the decision tougher. The simulation above does not distinguish between the two types of errors. It only matches the true vector with the average value of weak measurements, to give an overall success probability.

3 State discrimination with post-selection: the non-linear computation

In this Section, we will use the TSVF of weak measurements to perform a non-linear analysis of quantum state discrimination for a general coupling strength. We will show that for the rare events where the post-selection is successful we have a high probability to identify the vector. The motivation for utilizing post-selection was already verified in many precision measurements [2–4], hence it is natural to examine it for this application.

In the pre- and post-selected set of weak measurement (see below), we will look at the distribution of the pointer’s coordinate variable \hat{X}_{fin} . The moments of \hat{X}_{fin} are functions of the pre-selected state $|\psi_{\text{in}}\rangle$, the post-selected state $|\psi_{\text{fin}}\rangle$, and the operator A . We will show that it is possible to pick two initial vectors $|\psi_{\text{in}}^i\rangle$ such that the corresponding distributions of \hat{X}_{fin} are easily distinguished (possibly by one sample). In particular, for one of the initial vectors \hat{X}_{fin} will be distributed around 0 with standard error σ while for the other initial vector \hat{X}_{fin} will be distributed around σ with very low standard error (almost 0). This will make it easy to differentiate between the two cases.

In Sect.3.1 we compute the expansion of the weak value $\langle\psi_{\text{fin}}|e^{-ig\hat{A}\hat{X}/\hbar}|\psi_{\text{in}}\rangle$ using all terms in \hat{A} and \hat{X} . In Sect.3.1 we show how to pick two initial vectors that maximizes the difference between the corresponding distributions of \hat{X}_{fin} . Our derivation is based on [27].

3.1 The non-linear approximation of weak values

Suppose \hat{A} is an Hermitian operator on the principle system S . Let $|\psi\rangle$ denote a state vector for that system. Let $\hbar = 1$. We will also assume $\hat{A}^2 = 1$, this will make it easy to write $e^{-ig\hat{A}\hat{X}/\hbar}$ as a power series. Assume also that:

$$\langle \hat{A} \rangle_w = \frac{\langle \psi_{\text{fin}} | \hat{A} | \psi_{\text{in}} \rangle}{\langle \psi_{\text{fin}} | \psi_{\text{in}} \rangle} = ib$$

Let $|\phi\rangle$ denote the wave function of a quantum measurement device. Then,

$$|\phi\rangle = |\phi(x)\rangle = \int_x \phi(x) |x\rangle dx \quad (7)$$

where $\hat{X}|x\rangle = x|x\rangle$ is the position operator of the measuring needle. We will also assume that $|\phi(x)|^2$ is normally distributed around 0 with variance σ^2 :

$$\phi(x) = (2\pi\sigma^2)^{-1/4} e^{-x^2/4\sigma^2}$$

The function $\phi(x)$ represents the device's 'needle' amplitude distribution. We will now couple the principle system and the measurement system by the interaction Hamiltonian \hat{H}_{int} :

$$\hat{H} = \hat{H}_{\text{int}} = g(t) \hat{A} \otimes \hat{X} \quad (8)$$

(we used the operator \hat{X} instead of \hat{P} since we need imaginary weak values [28]). Here, $g(t)$ is a coupling function satisfying:

$$\int_0^T g(t) dt = g,$$

where T is the coupling time.

We shall start the measuring process with the vector:

$$|\psi\rangle \otimes |\phi(x)\rangle$$

in the tensor product of the two systems. Then, we apply the Hamiltonian:

$$e^{-i\hat{A}\hat{X}/\hbar} |\psi\rangle \otimes |\phi(x)\rangle.$$

Let

$$|\Phi_{\text{fin}}(x)\rangle = \langle \psi_{\text{fin}} | e^{-ig\hat{A}\hat{X}/\hbar} | \psi_{\text{in}} \rangle | \phi(x) \rangle$$

be the wave function of the needle following the coupling and the post-selection. For an observable \hat{M} on the needle's space $|\phi(x)\rangle$, let:

$$\langle \hat{M} \rangle_{\text{in}} = \frac{\langle \phi | \hat{M} | \phi \rangle}{\langle \phi | \phi \rangle},$$

$$\langle \hat{M} \rangle_{\text{fin}} = \frac{\langle \Phi_{\text{fin}} | \hat{M} | \Phi_{\text{fin}} \rangle}{\langle \Phi_{\text{fin}} | \Phi_{\text{fin}} \rangle}.$$

Since $\hat{A}^2 = 1$ we can write $\langle \psi_{\text{fin}} | e^{-ig\hat{A}\hat{X}} | \psi_{\text{in}} \rangle$ as

$$\begin{aligned} & \sum_{n=0}^{\infty} \frac{(-ig\hat{X})^{2n}}{(2n)!} \langle \psi_{\text{fin}} | \psi_{\text{in}} \rangle + \sum_{n=0}^{\infty} \frac{(-ig\hat{X})^{2n+1}}{(2n+1)!} \langle \psi_{\text{fin}} | \hat{A} | \psi_{\text{in}} \rangle \\ & = \langle \psi_{\text{fin}} | \psi_{\text{in}} \rangle [\cos(g\hat{X}) - i \langle \hat{A} \rangle_w \sin(g\hat{X})] \end{aligned} \quad (9)$$

where

$$\langle \hat{A} \rangle_w = \frac{\langle \psi_{\text{fin}} | \hat{A} | \psi_{\text{in}} \rangle}{\langle \psi_{\text{fin}} | \psi_{\text{in}} \rangle}.$$

Consider now the average bias of the needle:

$$\langle \hat{X} \rangle_{\text{fin}} = \frac{\langle \Phi_{\text{fin}} | \hat{X} | \Phi_{\text{fin}} \rangle}{\langle \Phi_{\text{fin}} | \Phi_{\text{fin}} \rangle} = \frac{\langle \hat{X} | \cos(g\hat{X}) - i \langle \hat{A} \rangle_w \sin(g\hat{X}) |^2 \rangle_{\text{in}}}{\langle | \cos(g\hat{X}) - i \langle \hat{A} \rangle_w \sin(g\hat{X}) |^2 \rangle_{\text{in}}}.$$

Recall $\langle \hat{A} \rangle_w = ib$, and let $a^+ = \frac{1+b^2}{2}$ and $a^- = \frac{1-b^2}{2}$, then [29]

$$\langle \hat{X} \rangle_{\text{fin}} = \frac{a^+ \langle \hat{X} \rangle_{\text{in}} + a^- \langle \hat{X} \cos(2g\hat{X}) \rangle_{\text{in}} + b \langle \hat{X} \sin(2g\hat{X}) \rangle_{\text{in}}}{a^+ + a^- \langle \cos(2g\hat{X}) \rangle_{\text{in}} + b \langle \sin(2g\hat{X}) \rangle_{\text{in}}} \tag{10}$$

Now since the needle is symmetric (normally distributed) we can write:

$$\langle \hat{X} \cos(2g\hat{X}) \rangle_{\text{in}} = \langle \sin(2g\hat{X}) \rangle_{\text{in}} = \langle \hat{X} \rangle_{\text{in}} = 0.$$

Hence

$$\langle \hat{X} \rangle_{\text{fin}} = \frac{b \langle \hat{X} \sin(2g\hat{X}) \rangle_{\text{in}}}{a^+ + a^- \langle \cos(2g\hat{X}) \rangle_{\text{in}}}. \tag{11}$$

We shall now use the parametrization $b = \cot(\frac{\eta}{2})$, ($a^- = -\frac{\cos(\eta)}{2\sin^2(\frac{\eta}{2})}$ and $a^+ = \frac{1}{2\sin^2(\frac{\eta}{2})}$), therefore:

$$\langle \hat{X} \rangle_{\text{fin}}^\eta = \frac{\sin(\eta) \langle \hat{X} \sin(2g\hat{X}) \rangle_{\text{in}}}{1 - \cos(\eta) \langle \cos(2g\hat{X}) \rangle_{\text{in}}}. \tag{12}$$

Since the needle is Gaussian, we can write

$$\langle \cos(2g\hat{X}) \rangle_{\text{in}} = e^{-2(g\sigma)^2}$$

and

$$\langle \hat{X} \sin(2g\hat{X}) \rangle_{\text{in}} = 2g\sigma^2 e^{-2(g\sigma)^2},$$

[30] and therefore,

$$\langle \hat{X} \rangle_{\text{fin}}^\eta = \frac{\sin(\eta) 2g\sigma^2}{e^{2(g\sigma)^2} - \cos(\eta)}. \tag{13}$$

The above ratio has maximal value at $\cos(\eta) = e^{-2(g\sigma)^2}$ which is [31]:

$$\langle \hat{X} \rangle_{\text{fin}}^{\text{max}} = \frac{2g\sigma^2}{\sqrt{e^{4(g\sigma)^2} - 1}}. \tag{14}$$

If the measurement is weak, i.e., $g \cdot \sigma \ll 1$, then

$$\frac{2g\sigma}{\sqrt{e^{4(g\sigma)^2} - 1}} \sim 1,$$

hence

$$\langle \hat{X} \rangle_{\text{fin}}^{\text{max}} \sim \sigma. \tag{15}$$

This will be true for η close to 0. However, if $\eta = 0$ then $\langle \hat{X} \rangle_{\text{fin}}^\eta = 0$ for all $g \neq 0$.

To sum-up, $\langle \hat{X} \rangle_{\text{fin}}^\eta$ is a function of two variables η and g . If the two variables are correlated such that $\cos(\eta) = e^{-2(g\sigma)^2}$ then for g small enough such that $g \cdot \sigma \ll 1$ we can get $\langle \hat{X} \rangle_{\text{fin}}^\eta$ to be very close to σ . But if we fix g (however, small) and let η go to 0, we can decrease $\langle \hat{X} \rangle_{\text{fin}}^\eta$ to 0. The variable η is a function of the weak values of \hat{A} . This means that we can tune η such that for η_1 the expectation value $\langle \hat{X} \rangle_{\text{fin}}^{\eta_1}$ will be close to σ , and for η_2 the expectation value $\langle \hat{X} \rangle_{\text{fin}}^{\eta_2}$ will close to 0.

To compute the variance of the needle after the post-selection note that:

$$\langle \hat{X}^2 \rangle_{\text{fin}} = \frac{a^+ \langle \hat{X}^2 \rangle_{\text{in}} + a^- \langle \hat{X}^2 \cos(2g\hat{X}) \rangle_{\text{in}} + b \langle \hat{X}^2 \sin(2g\hat{X}) \rangle_{\text{in}}}{a^+ + a^- \langle \cos(2g\hat{X}) \rangle_{\text{in}} + b \langle \sin(2g\hat{X}) \rangle_{\text{in}}} \tag{16}$$

Since the needle is normally distributed it is easy to see that

$$\langle \hat{X}^2 \cos(2g\hat{X}) \rangle_{\text{in}} = \sigma^2 e^{-2(g\sigma)^2} [1 - 4g^2\sigma^2], \quad (17)$$

[32] and, therefore,

$$\langle \hat{X}^2 \rangle_{\text{fin}} = \sigma^2 \left\{ \frac{a^+ + a^- e^{-2(g\sigma)^2} [1 - 4g^2\sigma^2]}{a^+ + a^- e^{-2(g\sigma)^2}} \right\}. \quad (18)$$

For small enough $g\sigma$ the value of $1 - 4g^2\sigma^2$ will be close to 1 and, therefore, $\langle \hat{X}^2 \rangle_{\text{fin}}$ will be close to σ^2 . Note that $\langle \hat{X}^2 \rangle_{\text{fin}}$ does not depend on the weak value since the needle has symmetric distribution.

3.2 Distinguishing between two non-orthogonal vectors

We can now pick two small angles η_1 and η_2 such that the difference between $\langle \hat{X} \rangle_{\text{fin}}^{\eta_1}$ and $\langle \hat{X} \rangle_{\text{fin}}^{\eta_2}$ is almost σ . Since η_1 and η_2 correspond to two initial vectors, we can distinguish between the two vectors by estimating the value of $\langle \hat{X} \rangle_{\text{fin}}^\eta$. In particular, consider

$$|\psi_{\text{in}}^i\rangle = \alpha_i |0\rangle - \beta_i |1\rangle$$

for $i = 1, 2$. Also

$$|\psi_{\text{fin}}\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle)$$

Let \hat{A} be the Hermitian operator:

$$\hat{A} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Then, $\hat{A}^2 = 1$. It is easy to see that:

$$\langle \hat{A} \rangle_w^i = i \frac{\alpha_i + \beta_i}{\alpha_i - \beta_i}$$

Let:

$$\eta_1 = \arccos \left(e^{-2(g\sigma)^2} \right).$$

Then, we can choose $|\psi_{\text{in}}^1\rangle = \alpha_1 |0\rangle - \beta_1 |1\rangle$ such that:

$$\alpha_1 = \frac{1}{\sqrt{2}} \left(\cos \frac{\eta_1}{2} + \sin \frac{\eta_1}{2} \right)$$

$$\beta_1 = \frac{1}{\sqrt{2}} \left(\cos \frac{\eta_1}{2} - \sin \frac{\eta_1}{2} \right)$$

and therefore

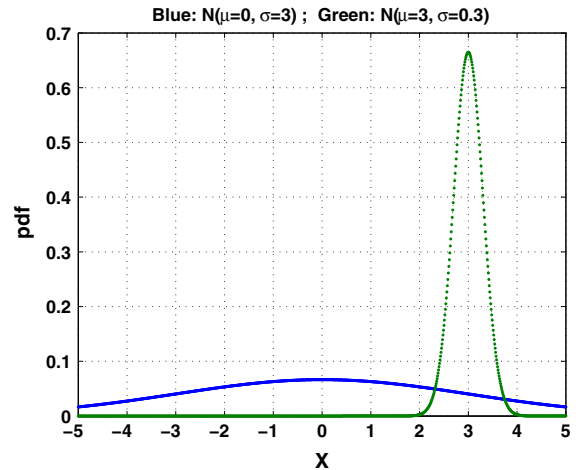
$$\frac{\alpha_1 + \beta_1}{\alpha_1 - \beta_1} = \cot \left(\frac{\eta_1}{2} \right).$$

Also, for η_2 close to 0 we can choose $|\psi_{\text{in}}^2\rangle = \alpha_2 |0\rangle - \beta_2 |1\rangle$ such that:

$$\alpha_2 = \frac{1}{\sqrt{2}} \left(\cos \frac{\eta_2}{2} + \sin \frac{\eta_2}{2} \right)$$

$$\beta_2 = \frac{1}{\sqrt{2}} \left(\cos \frac{\eta_2}{2} - \sin \frac{\eta_2}{2} \right)$$

Fig. 7 Illustration of the normal distribution of the pointer's first moment for each of the two initial vectors



hence

$$\frac{\alpha_2 + \beta_2}{\alpha_2 - \beta_2} = \cot\left(\frac{\eta_2}{2}\right).$$

Therefore, the difference between $\langle \hat{X} \rangle_{\text{fin}}^{\eta_1}$ and $\langle \hat{X} \rangle_{\text{fin}}^{\eta_2}$ will be close to σ .

The final variance of the needle in case $|\psi_{\text{in}}\rangle = |\psi_{\text{in}}^1\rangle$ is

$$\langle \hat{X}^2 \rangle_{\text{fin}} - \langle \hat{X} \rangle_{\text{fin}}^{\eta_1 2} \approx \sigma^2 - \sigma^2 = 0,$$

and, therefore, the needle is normally distributed around σ with very low standard error.

The final variance of the needle in case $|\psi_{\text{in}}\rangle = |\psi_{\text{in}}^2\rangle$ is

$$\langle \hat{X}^2 \rangle_{\text{fin}} - \langle \hat{X} \rangle_{\text{fin}}^{\eta_2 2} \approx \sigma^2 - 0 = \sigma^2,$$

hence the needle is normally distributed around 0 with standard error σ .

We can now easily distinguish between the two alternative distributions, possibly with a single sample using standard hypothesis testing (see for example the illustration in Fig. 7).

The success of the protocol depends solely on the post-selection probability, which is:

$$|\langle \psi_{\text{fin}} | \psi_{\text{in}}^i \rangle|^2 = \left| \frac{\alpha_i - \beta_i}{\sqrt{2}} \right|^2 = \left| \sin \frac{\eta_i}{2} \right|^2.$$

This probability will be low since each of the initial vectors is almost orthogonal to the final vector.

To sum-up, having post-selected the final vector, the probability to correctly guess the right vector could be high. Initially, the probability to post-select is low and, therefore, the complexity of the protocol depends mainly on the post-selection.

In the above example, the two initial states were represented in the same bases. Including an additional linear transformation in the scheme, we can generalize it to the case of two states chosen from two different mutually unbiased bases. This scheme might be suitable for quantum cryptography, and indeed, very recently a method based on sequential weak measurements was suggested for secure key distribution [33].

4 Discussion

Weak measurement theory challenges some of the most basic principles of quantum theory. In a nutshell, it allows the accumulation of information regarding the state vector without forcing its collapse. As we have shown in

Sect. 1, when performed many times, weak measurements are equivalent to a single strong one, thus approaching the well-known optimal success probability for discrimination performed by projective measurement. However, when performed only a limited number of times, they do not collapse the vector, but only rotate it. In such a case, we can still get some weak information about the state by reading (collapsing) the needle of the weak measurement apparatus. Moreover, when followed by post-selection, weak measurements can reveal underlying properties of the initial vector, allowing one to perform quantum state discrimination in retrospect as discussed in Sect. 2.

The gradual process of ‘getting information while determining the state’ which was demonstrated in Sect. 1 is strictly connected to the old measurement problem [34]. By weakly measuring the initial unknown vector we slowly collapse it, creating a continuous tradeoff between our knowledge and its superposition. This process can be thought of as a step-by-step decoherence in which the measured system ‘leaks’ through a small hole (the weak coupling to the measurement device) into the environment. As opposed to traditional decoherence this process is rigidly controlled and can be stopped at every stage, thus enabling much liberty to the experimenter. Therefore, weak measurement’s important contribution is its flexibility. With weak measurement one controls the tradeoff between success rate and collapse rate by choosing the strength of the coupling and the number of weak measurement.

Acknowledgments We thank Yakir Aharonov, Niv Cohen and Ariel Landau for helpful comments and discussions. E.C was partially supported by Israel Science Foundation Grant No. 1311/14.

Appendix: A pseudo-code describing the weak orbit

We wish to find the pointer final position q_m (the expectation of its distribution) when performing m successive weak measurements on a single particle prepared (for example) in the initial state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Let q_0 be distributed according to $N(1, \sigma^2)$ with probability 1/2 and according to $N(-1, \sigma^2)$ with probability 1/2. One can take $\sigma = m \gg 1$. Now let q_1 be distributed according to $N(1, \sigma^2)$ with probability:

$$\frac{\exp[(q_0 + 1)^2/2\sigma^2]}{\exp[(q_0 + 1)^2/2\sigma^2] + \exp[(q_0 - 1)^2/2\sigma^2]}$$

and according to $N(-1, \sigma^2)$ with probability:

$$\frac{\exp[(q_0 - 1)^2/2\sigma^2]}{\exp[(q_0 + 1)^2/2\sigma^2] + \exp[(q_0 - 1)^2/2\sigma^2]}.$$

Let q_2 be distributed according to $N(1, \sigma^2)$ with probability:

$$\frac{\exp[((q_0 + 1)^2 + (q_1 + 1)^2)/2\sigma^2]}{\exp[((q_0 + 1)^2 + (q_1 + 1)^2)/2\sigma^2] + \exp[((q_0 - 1)^2 + (q_1 - 1)^2)/2\sigma^2]}$$

and according to $N(-1, \sigma^2)$ with probability:

$$\frac{\exp[((q_0 - 1)^2 + (q_1 - 1)^2)/2\sigma^2]}{\exp[((q_0 + 1)^2 + (q_1 + 1)^2)/2\sigma^2] + \exp[((q_0 - 1)^2 + (q_1 - 1)^2)/2\sigma^2]}.$$

Then, q_m is distributed according to $N(1, \sigma^2)$ with probability:

$$\frac{\exp\left[\left(\sum_{i=0}^{m-1} (q_i + 1)^2\right)/2\sigma^2\right]}{\exp\left[\left(\sum_{i=0}^{m-1} (q_i + 1)^2\right)/2\sigma^2\right] + \exp\left[\left(\sum_{i=0}^{m-1} (q_i - 1)^2\right)/2\sigma^2\right]}$$

and according to $N(-1, \sigma^2)$ with probability:

$$\frac{\exp\left[\left(\sum_{i=0}^{m-1} (q_i - 1)^2\right)/2\sigma^2\right]}{\exp\left[\left(\sum_{i=0}^{m-1} (q_i + 1)^2\right)/2\sigma^2\right] + \exp\left[\left(\sum_{i=0}^{m-1} (q_i - 1)^2\right)/2\sigma^2\right]}.$$

References

1. Aharonov, Y., Albert, D., Vaidman, L.: How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.* **60**, 1351–1354 (1988)
2. Hosten, O., Kwiat, P.: Observation of the spin Hall effect of light via weak measurements. *Science* **319**, 787–790 (2008)
3. Starling, D.J., Dixon, P.B., Jordan, A.N., Howell, J.C.: Optimizing the signal-to-noise ratio of a beam-deflection measurement with interferometric weak values. *Phys. Rev. A* **80**, 041803 (2009)
4. Xu, X.Y., Kedem, Y., Sun, K., Vaidman, L., Li, C.F., Guo, G.C.: Phase estimation with weak measurement using a white light source. *Phys. Rev. Lett.* **111**, 033604 (2013)
5. Jordan, A.N., Tollaksen, J., Troupe, J.E., Dressel, J., Aharonov, Y.: Heisenberg scaling with weak measurement: A quantum state discrimination point of view (2014). [arXiv:1409.3488](https://arxiv.org/abs/1409.3488)
6. Aharonov, Y., Botero, A., Popescu, S., Reznik, B., Tollaksen, J.: Revisiting Hardy’s paradox: counterfactual statements, real measurements, entanglement and weak values. *Phys. Lett. A* **301**, 130–138 (2002)
7. Tollaksen, J., Aharonov, Y., Casher, A., Kaufherr, T., Nussinov, S.: Quantum interference experiments, modular variables and weak measurements. *New J. Phys.* **12**, 013023 (2010)
8. Aharonov, Y., Popescu, S., Rohrlich, D., Skrzypczyk, P.: Quantum cheshire cats. *New J. Phys.* **15**, 113015 (2013)
9. Vaidman, L.: Past of a quantum particle. *Phys. Rev. A* **87**, 052104 (2013)
10. Aharonov, Y., Cohen, E., Ben-Moshe, S.: Unusual interactions of pre- and post-selected particles. *EPJ Web Conf.* **70**, 00053 (2014)
11. Aharonov, Y., Vaidman, L.: The two-state vector formalism: an updated review. In: Muga, J.G., et al. (eds.) *Time in Quantum Mechanics*, pp. 399–447. Springer, Berlin, Heidelberg (2007)
12. Aharonov, Y., Rohrlich, D.: *Quantum Paradoxes: Quantum Theory for the Perplexed*. Wiley, Weinheim (2005)
13. Tamir, B., Cohen, E.: Introduction to weak measurements and weak values. *Quanta* **2**, 7–17 (2013)
14. Aharonov, Y., Cohen, E., Elitzur, A.C.: Foundations and applications of weak quantum measurements. *Phys. Rev. A* **89**, 052105 (2014)
15. Chefles, A.: Quantum state discrimination. *Contemp. Phys.* **41**, 401–424 (2000)
16. Helstrom, C.W.: *Quantum Detection and Estimation Theory*. Academic press, New York (1976)
17. Osaki, M., Ban, M., Hirota, O.: Derivation and physical interpretation of the optimum detection operators for coherent-state signals. *Phys. Rev. A* **54**, 1691 (1996)
18. Ivanovic, I.D.: How to differentiate between non-orthogonal states. *Phys. Lett. A* **123**, 257–259 (1987)
19. Bagan, E., Muñoz-Tapia, R., Olivares-Rentería, G.A., Bergou, J.A.: Optimal discrimination of quantum states with a fixed rate of inconclusive outcomes. *Phys. Rev. A* **86**, 030303 (2012)
20. Herzog, U.: Optimal state discrimination with a fixed rate of inconclusive results. Analytic solution and relation to state discrimination with a fixed error rate. *Phys. Rev. A* **86**, 032314 (2012)
21. Zilberberg, O., Romito, A., Starling, D.J., Howland, G.A., Broadbent, C.J., Howell, J.C., Gefen, Y.: Null values and quantum state discrimination. *Phys. Rev. Lett.* **110**, 170405 (2013)
22. Chiribella, G., Yang, Y., Yao, A.C.: Quantum replication at the Heisenberg limit. *Nat. Commun.* **4**, 2915 (2013)
23. Qiao, C., Wu, S., Chen, Z.B.: Unambiguous discrimination of extremely similar states by a weak measurement (2013). [arXiv:1302.5986](https://arxiv.org/abs/1302.5986)
24. Murch, K.W., Weber, S.J., Macklin, C., Siddiqi, I.: Observing single quantum trajectories of a superconducting quantum bit. *Nature* **502**, 211–214 (2013)
25. Kempe, J.: Quantum random walks: an introductory overview. *Contemp. Phys.* **44**, 307–327 (2003)
26. Childs, A.: *Quantum information processing in continuous time*, Ph.D. Thesis, M.I.T (2004)
27. Koike, T., Tanaka, S.: Limits on amplification by Aharonov–Albert–Vaidman weak measurement. *Phys. Rev. A* **84**, 062106 (2011)
28. Jozsa, R.: Complex weak values in quantum measurement. *Phys. Rev. A* **76**, 044103 (2007)
29. $|\cos(g\hat{X}) - i\langle \hat{A} \rangle_w \sin(g\hat{X})|^2 = |\cos(g\hat{X}) + b \sin(g\hat{X})|^2 = \cos^2(g\hat{X}) + b^2 \sin^2(g\hat{X}) + b \sin(2g\hat{X}) = (1 - b^2) \cos^2(g\hat{X}) + b^2 + b \sin(2g\hat{X}) = (1 - b^2)(\cos(2g\hat{X}) + 1)/2 + b^2 + b \sin(2g\hat{X}) = \frac{(1-b^2)}{2} \cos(2g\hat{X}) + \frac{(1+b^2)}{2} + b \sin(2g\hat{X})$
30. Use integration by parts on $\int e^{-\frac{x^2}{2\sigma^2}} \cos(kx) dx$
31. Note that $\cos(\eta) = e^{-2(g\sigma)^2}$, and $\sin(\eta) = \sqrt{1 - e^{-4(g\sigma)^2}}$, therefore $\frac{\sin(\eta)2g\sigma^2}{e^{2(g\sigma)^2} - \cos(\eta)} = \frac{\sqrt{1 - e^{-4(g\sigma)^2}} 2g\sigma^2}{e^{2(g\sigma)^2} (1 - e^{-4(g\sigma)^2})}$
32. Use integration by parts on $\int e^{-\frac{x^2}{2\sigma^2}} [\cos(2kx) - 2kx \sin(2kx)] dx$ and the previous formula for $\langle \hat{X} \sin(2g\hat{X}) \rangle_{in}$
33. Troupe, J.E.: Quantum key distribution using sequential weak values. *Quantum Stud. Math. Found.* **1**, 79–96 (2014)
34. von Neumann, J.: *Mathematical Foundation of Quantum Mechanics*. Princeton University Press, Princeton, NJ, USA (1955)