

Quantum Trajectories for Realistic Photodetection I: General Formalism

P. Warszawski and H. M. Wiseman*

Centre for Quantum Dynamics, School of Science, Griffith University, Brisbane 4111, Australia.

Quantum trajectories describe the stochastic evolution of an open quantum system conditioned on continuous monitoring of its output, such as by an ideal photodetector. In practice an experimenter has access to an output filtered through various electronic devices, rather than the microscopic states of the detector. This introduces several imperfections into the measurement process, of which only inefficiency has previously been incorporated into quantum trajectory theory. However, all electronic devices have finite bandwidths, and the consequent delay in conveying the output signal to the observer implies that the evolution of the conditional state of the quantum system must be non-Markovian. We present a general method of describing this evolution and apply it to avalanche photodiodes (APDs) and photoreceivers. We include the effects of efficiency, dead time, bandwidth, electronic noise, and dark counts. The essential idea is to treat the quantum system and classical detector jointly, and to average over the latter to obtain the conditional quantum state. The significance of our theory is that quantum trajectories for realistic detection are necessary for sophisticated approaches to quantum feedback, and our approach could be applied in many areas of physics.

PACS numbers: 03.65.Yz, 03.65.Ta, 42.50.Lc, 42.50.Ar

I. INTRODUCTION

A. Measurement in Quantum Mechanics

To obtain information about a system of interest a measurement has to be made. In experiments that probe the quantum nature of our world, the system itself is, in general, necessarily affected by the act of measurement. For experiments involving continuous monitoring, traditional (projective) measurement theory as axiomatized by von Neumann [1] is inadequate as a description. Such measurements would prevent any interesting evolution occurring, because of the quantum Zeno effect [2]. To avoid this, a generalized quantum theory of measurement must be used [3, 4]. Generalized quantum measurements can be derived by considering projective measurements on a second system interacting with the system of interest. Often the second system has a clear physical interpretation. For example, in quantum optics a bath (continuum) of electromagnetic field modes plays this role and the quantum system is said to be *open*.

If the system is weakly coupled to its bath and the dynamics are such that information concerning the system is spread throughout the many degrees of freedom of the bath, then a Markovian evolution equation for the system alone may be obtained. This equation is known as a *master equation* (ME). If, in addition, the bath is being measured then a stochastic master equation (SME) for the conditioned state of the system can be found. This is termed a *quantum trajectory* [5]. The relation of the ME to quantum trajectories is that the former results from summing all possible trajectories, weighted by their probability of occurrence. Thus, the ME represents av-

erage evolution and gives the evolution of the state of the system when the environment is not monitored. A quantum trajectory (which is said to ‘unravel’ the master equation) gives the evolution of the state of the system conditioned on the results of monitoring the environment. A quantum trajectory is necessarily stochastic, and can be jumpy [5, 6, 7] or diffusive [5, 8, 9, 10, 11, 12], depending on the nature of the measurement.

B. Realistic Quantum Trajectories

In experiments performed in the laboratory, a detector mediates between the detector input (for example, the electromagnetic field emitted by a fluorescent quantum system) and the experimenter. Realistic detectors are not perfect. Information is lost in the conversion of the quantum field to a signal that the observer can use. This loss may occur at the front end of the detector (characterized by an inefficiency), in the circuit containing the detector (described by a response time and electronic noise) or at the circuit output (electronic output noise). In order to obtain a true quantum trajectory for the experiment, the observer must condition the state of the quantum system on results that are available in the laboratory rather than on the microscopic absorption events considered previously in quantum trajectories.

Realistic quantum trajectories of this sort cannot be encompassed within the standard formalism of Refs. [3, 4], except for inefficiency [13]. This is because the electrical circuit, which filters the measured output of the quantum system, causes a statistical delay in the observer obtaining information about the system. Thus the conditioned evolution of the quantum system alone is non-Markovian. A method for treating this was first proposed by us and Mabuchi [14], in which the quantum system is embedded within a supersystem that also contains the

*Electronic address: H.Wiseman@gu.edu.au

state of the detector. If the set of (classical) detector states is \mathbb{S} , then the supersystem is described by the set $\{\rho_s : s \in \mathbb{S}\}$. Here $\text{Tr}[\rho_s]$ is the probability that the apparatus is in state s , and $\rho_s/\text{Tr}[\rho_s]$ is the system state given this event. The supersystem state, $\{\rho_s\}$, obeys a Markovian evolution equation. The detector states are classical in the sense that it is not necessary to consider superpositions between them.

In this paper we consider in detail two sorts of realistic detectors, an avalanche photodiode in geiger mode (a photon counter) and a photoreceiver (for homodyne detection). The detector states in the former case are discrete (we use three states) and in the latter continuous (although discretization is necessary for numerical simulation). For the photon counter, this model allows us to include the effects of detection efficiency, dark counts, response (or rise) time and dead time. For homodyne detection using a photoreceiver we investigate detector efficiency, output Johnson noise (which adds onto the measurement result linearly) and detector bandwidth. (Of course, this bandwidth relates not to the range of frequencies (colours) of light to which the detector will respond, but to the characteristic time it takes for the absorption of light by the detector to generate a response in the detector output.)

Before passing onto our method of derivation of realistic quantum trajectories, it is worth mentioning that a three-level model for a photon counter including many of the above effects has been considered by Gardiner [15]. However this theory is based upon quantum Langevin equations rather than quantum trajectories, and does not allow one to determine the system state conditioned on the experimenter's knowledge.

C. Derivation of Realistic Quantum Trajectories

The derivation of realistic quantum trajectories consists of a number of well defined steps, that are schematically illustrated in Fig. 1. These can be understood by considering the subsystems that make up the supersystem, namely the quantum system and the detector.

The basic philosophy that we adopt at the start of the derivation is that *in principle* all of the microscopic states of the detector could be monitored, without altering the quantum trajectory of the system. Essentially this is because the detector states consist of distinct configurations of an enormous number of particles. These particles are constantly interacting with their environment, thus making coherences between detector states unimportant. Thus in principle it is possible for an experimenter to have sufficient information to unravel the evolution of the system by a conventional stochastic master equation (SME). In practice, an experimenter has access to far less knowledge. This loss of knowledge is due to *classical* uncertainties introduced by the detector.

To describe the detector we consider some property of it that can be related to the output that the observer

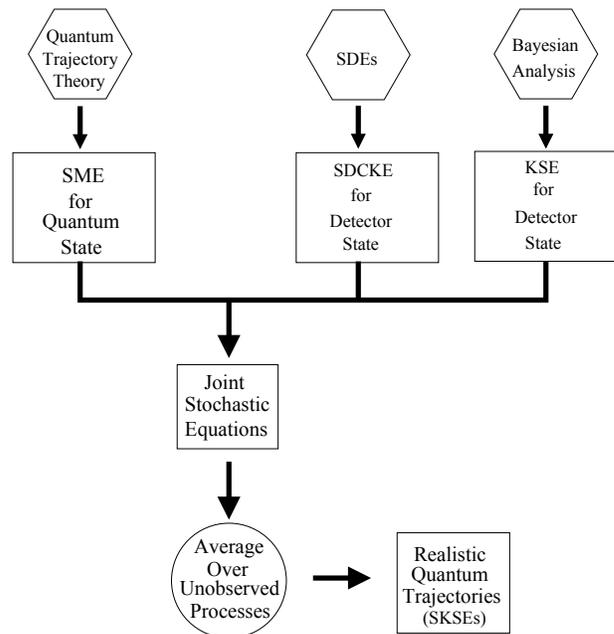


FIG. 1: Schematic diagram for the method that we use to derive realistic quantum trajectories. It should be noted that some of the steps shown above can be performed in a different order. This flexibility is useful for keeping the derivation of realistic quantum trajectories as simple as possible. The following abbreviations have been used: stochastic master equation (SME), stochastic differential equation (SDE), stochastic differential Chapman-Kolmogorov equation (SDCKE), Kushner-Stratonovich equation (KSE) and superoperator Kushner-Stratonovich equation (SKSE). Hexagons, rectangles and circles indicate bodies of knowledge, equations and procedures, respectively. See text for further details.

monitors. For example, in the case of homodyne detection the charge across one of the capacitors in the circuit represents the detector state. By considering all of the dynamic influences upon this circuit variable we formulate an evolution equation for it that is stochastic in general. Since finally we will not know the stochastic inputs, we transform this into a stochastic differential Chapman-Kolmogorov equation (SDCKE) that follows the probability of the various states being occupied. We use this terminology to indicate a generalization of the deterministic differential Chapman-Kolmogorov equation discussed in [16].

The realistic measurement results available to the observer give information about the detector state, without necessarily revealing it fully. Based on the result and the *a priori* (initial) probability distribution for the detector state, the best estimate of the updated state (the *a posteriori* distribution) is found. This is done via a Bayesian analysis [17]. The resulting equation for the probability distribution is, for the case of Gaussian white noise, usually referred to as a Kushner-Stratonovich equation (KSE) [18]. For convenience we will use this terminology even in the case of point-process noise.

Note that normally a KSE would also contain that unconditioned (deterministic) evolution due to the system dynamics. Since these dynamics are stochastic at this stage of the derivation (they are contained in the SDCKE), we do not include them in the KSE for the detector state. Note also that the SME, despite its name (introduced in Ref. [13]), is conceptually closer to the KSE than to the SDCKE: the stochasticity in both the KSE and the SME is solely due to measurement noise, and the stochastic terms are nonlinear in both cases. In fact, the SME can justifiably be called a quantum KSE [19].

The evolution of the SME, the SDCKE and the KSE are combined to give the joint stochastic equations governing the evolution of the supersystem. Up until this stage all the stochastic influences on the supersystem are still present. To link the evolution of the system with that of the detector this is necessary. Now, however, we average over processes that a realistic observer cannot monitor. This leaves only the stochasticity associated with the measurement and our derivation of the realistic quantum trajectory is then complete. We term this final stochastic trajectory for the supersystem a superoperator Kushner-Stratonovich equation (SKSE). Obviously the reader will gain much greater insight into this process by following the examples provided in the later sections.

D. Importance of Realistic Quantum Trajectories

Thus far, the main practical utility of quantum trajectory theory has been in improving the computational efficiency of simulations used to compare models with experimental data. For this purpose, the introduction of realistic quantum trajectories could only hinder simulations. However, quantum trajectory theory is now gaining increasing importance as the quantum generalization of Kalman filtering, which provides essential signal-processing methods in classical estimation, communication, and control engineering. Quantum trajectory theory should in principle play the same pivotal role for emerging quantum analogs of these technologies [20, 21, 22]. Before this can happen it is essential that the theory be extended to account for the imperfections of realistic measurement devices. Non-ideal detector dynamics can dramatically affect the proper inference from measured signals to the conditional quantum state of an observed system, as we will show.

E. Overview of this Paper

As stated above, the main thrust of this paper is to derive in detail the realistic quantum trajectories for photon counting and homodyne photodetection introduced in Ref. [14]. In Sec. II, we first briefly give the theory of classical trajectories and measurement, which is important for obtaining SDCKEs and KSEs for the detector

state. A discussion of non-realistic (that is, ideal) quantum trajectories follows this, which allows the SMEs for the quantum system to be found.

In Secs. III and IV, realistic quantum trajectories are derived for photon counters and photoreceivers respectively. Both of these derivations are based on the procedure outlined in Sec. IC. For the photon counter, we finally idealize the measurement process to show that our equations reduce to those associated with perfect detection (traditional quantum trajectories). An important and non-trivial consideration in the case of the photoreceiver is the bandwidth of the device. As the photoreceiver circuit adds electronic noise, a calculation of this bandwidth based on just the resistance-capacitance (RC) time constant of the circuit does not yield the correct answer. As a final point we consider the effect of letting the RC time constant go to zero but still maintaining the presence of the electronic noise which obscures the input signal. We conclude with a summary of our results and discussion of future directions in Sec. V. A numerical exploration of our theory is given in the following paper [23].

II. STOCHASTIC EQUATIONS

As discussed in Sec. IC, there are three types of stochastic equations that go into making a realistic quantum trajectory. In this section we examine the general theory for each.

A. Stochastic Differential Chapman-Kolmogorov Equation

A typical Langevin equation or stochastic differential equation (SDE) for a classical random variable X can be written as

$$dX = a(X)dt + \sqrt{D}dW(t) + edN(t). \quad (2.1)$$

Here $a(X)$ is an arbitrary function of X , and $dW(t)$ and $dN(t)$ are stochastic increments. The point process $dN(t)$ is either zero or one, and has an infinitesimal mean which is a positive function of X , say $g(X)dt$. The Wiener increment $dW(t)$ is related to Gaussian white noise $\xi(t)$ by $dW(t) = \xi(t)dt$. Gaussian white noise has statistics of [16]

$$E[\xi(t)] = 0, \quad E[\xi(t)\xi(t')] = \delta(t - t'). \quad (2.2)$$

Here the roman font E denotes expectation value. By writing the SDE in the form of Eq. (2.1) we imply that the equation is to be interpreted in the Itô [16] or explicit [24] sense. This means, for example, that the expected value of dX may be found by replacing dN and dW by their expected value. This is not true for the implicit [24] or Stratonovich [16] form. By the same token, the chain

rule of standard calculus does not apply in general to the Itô form, whereas it does for the Stratonovich form.

Langevin equations are appropriate for situations in which the stochastic increment is known, thus allowing X to be tracked. Often it is relatively easy to turn the physics of a problem into a Langevin equation. An example of this is a noisy electrical circuit. The analysis of such circuits is well understood and is based on simple principles such as Kirchhoff's laws.

In this paper, Langevin equations are first found on the *presumption* that the stochastic increment is known. In reality, however, this is not the case and we cannot track the variable X perfectly. For this reason it is necessary to turn the Langevin equation into an equation describing the evolution of the *probability distribution* for X , which is written as $P(x)$. Note that the transformations that we perform actually maintain the stochastic nature of this evolution equation, as the averaging over unobserved processes is done at a later stage. This is necessary because the noise in the equation for $P(x)$ may also appear in other parts of the overall derivation. Before this final averaging is done, the stochastic equation for $P(x)$, would preserve a Dirac δ -function as a solution. We now do this for Eq. (2.1), as an aid the reader.

If X is known at time t , then at that time $P(x) = \delta(x - X)$. This then evolves according to

$$d\delta(x - X) = \delta(x - X - dX) - \delta(x - X) \quad (2.3)$$

$$= \delta[x - X - a(X)dt - \sqrt{D}dW(t) - edN(t)] - \delta(x - X). \quad (2.4)$$

We can treat the diffusion and jump noise in a unified fashion by constructing a Taylor series of Eq. (2.4) to all orders of the first term on the right-hand side (RHS):

$$d\delta(x - X) = \left(\exp \left\{ \frac{\partial}{\partial x} \left[-a(X)dt - \sqrt{D}dW(t) - edN(t) \right] \right\} - 1 \right) \delta(x - X). \quad (2.5)$$

We now expand the exponential to all orders and use the Itô rules $dN^2 = dN$, $dW^2 = dt$, with all other products being zero. This gives

$$d\delta(x - X) = \left\{ -\frac{\partial}{\partial x} \left[a(X)dt + \sqrt{D}dW(t) \right] + \frac{D}{2} \frac{\partial^2}{\partial x^2} dt \right\} \delta(x - X) + dN(t) [\delta(x - X - e) - \delta(x - X)] \quad (2.6)$$

Using the delta function to change $a(X)$ to $a(x)$ and then averaging over the random variable X gives the Itô equation for the probability $P(x) \equiv \text{E}[\delta(x - X)]$,

$$dP(x) = \left\{ \frac{\partial}{\partial x} \left[-a(x)dt - \sqrt{D}dW(t) \right] + \frac{D}{2} \frac{\partial^2}{\partial x^2} dt \right\} \times P(x) + dN(t) [P(x - e) - P(x)]. \quad (2.7)$$

The first (dt) term is deterministic drift, the second (dW) is stochastic drift, the third (dt) is diffusion and the last (dN) is the jump term. We refer to this equation as a stochastic differential Chapman-Kolmogorov equation (SDCKE). If there were no jump term, then we would have a stochastic Fokker-Planck equation.

If averages over $dW(t)$ and $dN(t)$ are taken then we have a standard (deterministic) differential Chapman-Kolmogorov equation

$$\frac{\partial}{\partial t} P(x) = \left\{ -\frac{\partial}{\partial x} a(x) + \frac{D}{2} \frac{\partial^2}{\partial x^2} \right\} P(x) + g(x - e)P(x - e) - g(x)P(x), \quad (2.8)$$

where $g(x)$ enters from the definition $\text{E}[dN(t)f(X)] = \text{E}[g(X)f(X)]$.

B. Kushner-Stratonovich Equation

From the point of view of measurement, it useful to consider the state of a classical system to be defined by the probability distribution on *state space*. This is the space of all relevant physical quantities, called *state variables*. In this paper the classical system will be a detector and the associated electrical circuit. Only one variable is necessary to describe it so we can think of the distribution $P(x)$ that was introduced in the previous section as representing the state.

As the state of the system summarizes the observer's knowledge it is logical that the state will change on the basis of measurement upon the system. From the *a priori* state $P(x)$ of the system and the result r of measurement, the *a posteriori* state $P(x|r)$ can be found using Bayesian inference [17]

$$P(x|r) = \frac{P(r|x)P(x)}{P(r)}. \quad (2.9)$$

Here $P(x|r)$ reads as 'the probability of x given r '. The division by $P(r)$ is for normalization purposes, where

$$P(r) = \int dx P(r|x)P(x). \quad (2.10)$$

Here $P(r|x)$ represents the probability of obtaining result r , given that the state is x . The expression in Eq. (2.9) will be repeatedly used in this paper to obtain the Kushner-Stratonovich equations (KSEs) for the classical system. This is simply the continuous-in-time limit of Bayes theorem. Rather than provide the general form of the KSE here, the reader will be guided through the derivations of the KSEs (as they arise) from fundamental considerations.

The unconditioned system state which we would use if the measurement had been performed and ignored is given by

$$P'(x) = \int dr P(x|r)P(r), \quad (2.11)$$

where we have assumed that the results take on continuous values. The prime is included to remind the reader that this is an *a posteriori* state. In the above we have assumed that there is no measurement back-action on the system – that is, the measurement only affects the observer’s knowledge and not the dynamics of the system itself – so Eq. (2.11) reduces to $P(x)$ as one would expect. Of course, it is possible to consider classical measurements for which there is a measurement back-action. This is also the case for all quantum measurements because of the Heisenberg uncertainty principle.

C. Stochastic Master Equation

As noted in Sec. IC, the SME is a quantum version of the KSE, including the evolution unconnected with the measurement as well. The SME can be derived within the standard formalism of generalized quantum measurement (for simple introductions, see for example Refs. [25] and [26]). This formalism is simply the quantum analogue of Bayes’ theorem, generalized for measurement back-action, as shown recently in Ref. [27]. In this section we will not review the derivation of the SME from the standard formalism, but will rather just quote the final results.

1. Jumpy Trajectories

In many examples of continuous measurement on the output system, the result can only take on two possible values. This is the case in ideal direct photodetection in which the electromagnetic field is monitored with an ideal detection device that responds to the presence of photons. In any infinitesimal time interval a photon either is or is not detected. Such measurements give rise to jumps in the conditional system state because the rare event of a photon detection conveys to the observer a finite amount of information about quantum system. For example, if the fluorescent system were a TLA then a photon detection implies that the TLA now resides in the ground state. A null result still causes a non-unitary change to the system state, but it is of infinitesimal magnitude. That is not to say that it is unimportant, however, as over a finite length of time between jumps, it will also cause a finite change in the system state.

Let the fluorescent system have lowering operator c , and assume that the fluorescence can be formed into a beam of light. Also let an optical local oscillator (that is, an effectively classical light field from a laser) be added to the fluorescence before it is detected. This can be done using a highly transmitting beam-splitter. Let the local oscillator (LO) amplitude added to the fluorescent beam be μ . If we measure time in units such that the photon flux from the system is $\langle c^\dagger c \rangle$, then the unobserved system obeys the (deterministic) master equation

$$\dot{\rho}(t) = -i[H, \rho(t)] + \mathcal{D}[c]\rho(t) \equiv \mathcal{L}\rho, \quad (2.12)$$

where the superoperator \mathcal{D} is defined, for arbitrary B , as

$$\mathcal{D}[B]\rho = \mathcal{J}[B]\rho - \mathcal{A}[B]\rho \quad (2.13)$$

$$\equiv B\rho B^\dagger - \frac{1}{2}(B^\dagger B\rho + \rho B^\dagger B), \quad (2.14)$$

where the two terms in the last line define \mathcal{J} and \mathcal{A} respectively. This master equation is of the Lindblad form [28], where we have assumed for simplicity that the evolution apart from the radiation damping term ($\mathcal{D}[c]\rho$) can be described by a Hamiltonian H .

To give an explicit description of the conditional system state evolution under ideal photon detection, it is useful to define the point process, $dN(t)$, that is equal to 1 if there is a detection in the time interval $[t, t + dt)$ and 0 otherwise. Allowing for a detection efficiency η , the statistics of $dN(t)$ are defined by

$$dN(t) = dN(t)^2, \quad (2.15)$$

$$E[dN(t)] = \eta \text{Tr}[(c^\dagger + \mu^*)(c + \mu)\rho_N] dt. \quad (2.16)$$

The subscript N on the system state is to indicate that it is conditioned on the detection events. The evolution of the system state in terms of $dN(t)$ is given by [24]

$$\begin{aligned} d\rho_N = & -dt\mathcal{H}[iH + \frac{1}{2}\eta c^\dagger c + \eta\mu^*c + \frac{1}{2}\eta|\mu|^2]\rho_N \\ & + dN(t)\mathcal{G}[\sqrt{\eta}(c + \mu)]\rho_N \\ & + dt(1 - \eta)\mathcal{D}[c]\rho_N. \end{aligned} \quad (2.17)$$

This equation is a stochastic master equation (SME). Note that we have dropped the time argument of the conditioned system state, but maintained that of the increment, $dN(t)$. This is done for convenience and, also, to emphasize the stochastic nature of the detection events. Time arguments are only included in this paper when they will aid the reader’s understanding. The non-linear superoperators, \mathcal{G} and \mathcal{H} are defined by

$$\mathcal{G}[A]B = \frac{\mathcal{J}[A]B}{\text{Tr}[\mathcal{J}[A]B]} - B. \quad (2.18)$$

$$\mathcal{H}[A]B \equiv AB + BA^\dagger - \text{Tr}[AB + BA^\dagger]B. \quad (2.19)$$

It is possible to return to the ME (Eq. (2.12)) by replacing $dN(t)$ in Eq. (2.17) by its expectation value (Eq. (2.16)).

2. Diffusive Trajectories

In the limit of $\mu \rightarrow \infty$ the rate of quantum jumps goes to ∞ but the effect on the system of each one goes to zero. This leads to quantum trajectories of a diffusive nature, where the system evolves continuously but non-differentiably in time. In this limit the photocounts give way to a photocurrent, J . Including a detection efficiency η , and for a particular normalization, it is given by

$$Jdt = \eta \langle ce^{-i\Phi} + c^\dagger e^{i\Phi} \rangle dt + \sqrt{\eta}dW(t), \quad (2.20)$$

where $dW(t) = \xi(t)dt$ represents Gaussian white noise [16] having the properties given in Eqs. (2.2)–(2.2). Here

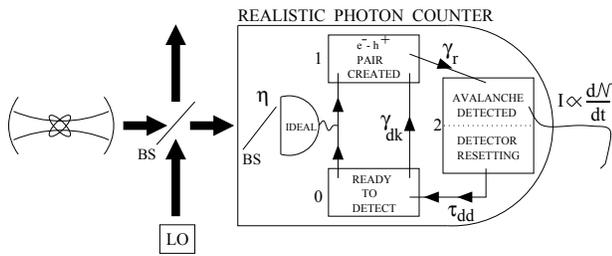


FIG. 2: Schematic diagram of the model used in this paper for a realistic photon counter monitoring a TLA placed in an optical cavity, the output of which is combined with that of a weak local oscillator (LO) at a low reflectivity beam splitter (BS) before direct detection. This will be the system on which investigation will focus in future numerical work. In this diagram we have used a BS of efficiency η followed by an ideal photon counter to represent the quantum efficiency of the realistic detector. The three states of the detector 0, 1, 2 are explained in the text. Single arrow heads are used for Poisson processes. The photocurrent produced by the detector I will consist of spikes at the time of avalanches ($dN = 1$).

$\Phi = \arg(\mu)$, and putting $\Phi = 0$ in Eq. (2.20) corresponds to measurement of the x quadrature $c + c^\dagger$, while $\Phi = \pi/2$ corresponds to measurement of the y quadrature $-i(c - c^\dagger)$. The system state evolution conditioned on this photocurrent is [5, 13]

$$d\rho_J = dt\mathcal{L}\rho_J + \{J(t)dt - E[J(t)]dt\} \mathcal{H}[e^{-i\Phi}c]\rho_J, \quad (2.21)$$

where again the subscript on ρ_J indicates that it is conditioned on the recording of the current J .

III. REALISTIC PHOTON COUNTING

An avalanche photodiode (APD) operating in Geiger mode produces a macroscopic current pulse in response to an incident photon. This allows the observer to detect the presence of photons and is the reason the device is also known as a photon counter. The size of the current pulse is large compared to the sources of noise within the detection circuitry, which allows the threshold current value for a detection to be set well above this noise.

The APD essentially consists of a p-n junction operated under a reverse bias greater than the breakdown voltage [29, 30, 31]. Under these conditions the diode can be described by just three classical states [30] (see Fig. 2 and Fig. 3). This will enable a mathematical model of the APD to be simply constructed.

The first state (0) is a stable low-current state in which there are no charge carriers in the depletion region of the junction. The transition from 0 to the ‘‘unstable’’ intermediate state (1) takes place when an electron-hole (e^-h^+) pair is created in the depletion region by an incident photon (with quantum efficiency η) or by thermally initiated ‘dark counts’ occurring at a rate γ_{dk} . Further impact ionization, under the influence of the electric

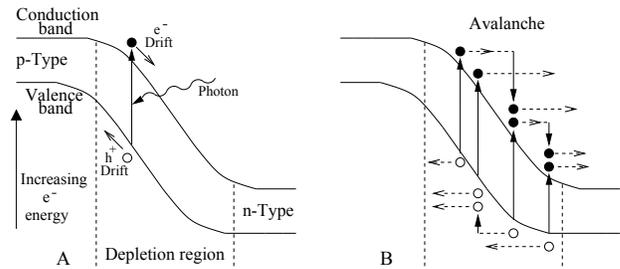


FIG. 3: Diagrams (A) and (B) show electron energy band diagrams for a p-n junction operated under a beyond breakdown reverse bias. In diagram (A) an incident photon causes the creation of an electron (filled circles) - hole (circles) pair that begins to separate due to the applied electric field. The APD would be in state 1. In (B) the electron and hole collide with other charges leading to an avalanche. Whether the avalanche has reached threshold or not determines if diagram (B) corresponds to the detector being in state 2 or 1, respectively.

field, leads to an avalanche forming. At this stage (before the avalanche has been detected) the APD is still in state 1. The avalanche continues to build until the current reaches some threshold value and a detection is registered, thus changing the state of the APD to 2. Due to the stochastic manner in which the avalanche spreads, the dwell time in state 1 is not deterministic. We model the 1 to 2 transition as a Poissonian process with rate γ_r . We call the mean dwell time, γ_r^{-1} , the ‘response time’. The avalanche, once detected, is arrested by the application of a negative-going voltage pulse that temporarily brings the bias voltage below the breakdown value [31]. This results in a fixed ‘dead time’, τ_{dd} , during which the APD cannot detect photons, after which it is restored to state 0.

A. Physical Explanation of Parameters

Now that our basic model for the APD has been introduced, a further physical explanation is given for the quantum efficiency, response time, the dead time and dark counts.

The quantum efficiency, η , is largely determined by two considerations. These are the fraction of incident photons that are absorbed in the active region of the junction and the fraction of e^-h^+ pairs that lead to an avalanche [30]. The first is determined by the absorption co-efficient of the material at the wavelength of interest and, also, the thickness of the depletion region. If the absorption co-efficient is too large then the photons will be absorbed near the surface, where there exists only a weak biasing electric field. Too small a co-efficient allows the light to pass straight through. The second consideration is a strong function of the overvoltage (beyond breakdown), with higher fields ensuring that the creation of a charged pair leads to an avalanche. Unfortunately, this also leads to approximately exponentially increased dark

counts [32].

The response rate, γ_r , is also a strong function of the overvoltage due to its effect on the rate at which charge carriers move through the junction. Other important considerations are the dimensions of the junction, which alter the time taken for the avalanche to spread throughout. Two important mechanisms that facilitate the spreading are the direct electrical diffusion of carriers into adjacent regions of the junction and the emission of photons from hot carriers that are absorbed in a different region [33].

The dead time, τ_{dd} , is generally quite long as all charge carriers must be swept out of the active region before the overvoltage can be re-applied without another avalanche resulting. The application of a negative-going voltage pulse (active quenching) is more effective than letting a resistance in series with the junction reduce the reverse bias below breakdown as the current increases (passive quenching). The active quenching of the avalanche can be initiated as soon as the current reaches a threshold value. Minimizing the circuit noise allows the threshold to be lowered thus reducing the response time and, also, the duration of the voltage pulse.

The dark count rate, γ_{dk} , can be reduced to insignificance with the use of a high quality Silicon APD [30]. However, such devices are not appropriate for detection at all wavelengths. In other materials the existence of free-carrier trapping centers and thermally generated carriers creates a considerable problem. During an avalanche, carriers can become trapped in a region of the junction, only to be released after the above breakdown voltage has been restored (causing an afterpulse). This leads to a dark count which is signal-dependent as the number of trapping centers generally increases with the frequency of avalanches. Cooling will reduce the thermal dark count. These complexities are not introduced into the APD model in this paper.

B. Realistic Quantum Trajectories

Our aim is to derive the quantum trajectories for the quantum system (the source of the light entering the APD) *conditioned on the observation of an avalanche*. In order that the derived equations be applicable to different measurement schemes, we consider the output field of the system of interest to be combined at a low reflectivity beam splitter with that of a weak local oscillator (LO), having transmitted amplitude μ . If the LO is not weak, and contributes a large photon flux, then a photon counter is not the appropriate type of detector. On the other hand, μ can be set to zero to obtain direct detection without problems.

In Fig. 1 the general method of deriving realistic quantum trajectories was given. The point that some of the steps in the method are interchangeable was made in the caption to that figure. For the APD, we will in fact perform the conditioning upon measurement after the un-

observed processes are averaged over. This will keep the equations that we manipulate during the derivation as simple as possible.

1. Stochastic Master Equation

As described in the introduction, the detection device can be treated in a classical manner. This implies that the presence of an observer monitoring the detailed behaviour of the charge-carriers in the p-n junction would not alter the detection results. The hypothetical observer could also, in principle, deduce the incidence of photons that were absorbed outside the active region of the junction. In this way, a list of the times at which photons arrived could be compiled. Of course, such sensitive measurements are currently out of reach due to technological reasons, but the point is that it is consistent with quantum measurement theory to assume a quantum jump unraveling of the ME. In fact, we even assume at this stage that the measurement can have unit efficiency, as it is conceptually more satisfying to explicitly include inefficiency in our description of the APD. Thus the fictitious conditional dynamics of the system are those of Eq. (2.17), with $\eta = 1$.

2. Stochastic Differential Chapman-Kolmogorov Equation

We can now formulate equations of motion for the probability of each of the three classical states of the APD being occupied. Because the state space is discrete, averaging over the stochastic processes would necessarily yield a classical master equation [16] and we could refer to the equations we derive in this section as a stochastic classical master equation. However, to solidify for the reader the processes involved in forming a realistic quantum trajectory (described in Sec. I C), we refer to the stochastic probability equations as SDCKEs. The stochastic equation would apply if all transitions were monitored, and that when an electron-hole pair is created a distinction can be drawn between those created via photon absorption and those via thermal fluctuations (dark counts). Because of the classical nature of the APD dynamics, this is logically possible. Eventually we will average over processes that a realistic observer cannot monitor.

There are a number of processes that contribute to the evolution of the occupation probabilities. Here, we will treat dark counts, avalanches, APD resetting and incident photons from the monitored quantum system. Since all transitions are being monitored by a fictitious observer and we assume, for simplicity, that the initial state of the APD is known the evolution of the probabilities is described by jump terms only. If we use $dN_{dk}(t) = 0, 1$ to define the occurrence of transitions due to dark counts, $dN'(t)$ to indicate transitions due to photons from the quantum system of interest and $dN''(t)$ for avalanches

then

$$dP_0 = -dN_{\text{dk}}(t) - dN'(t) + dN'(t - \tau_{\text{dd}}) \quad (3.1)$$

$$dP_1 = dN_{\text{dk}}(t) + dN'(t) - dN'(t) \quad (3.2)$$

$$dP_2 = dN'(t) - dN'(t - \tau_{\text{dd}}). \quad (3.3)$$

From these SDCKEs it can be seen that photons and dark counts, which lead to the creation of an electron-hole pair, take the detector from state 0 to state 1. Avalanches cause the transition $1 \rightarrow 2$ and the resetting of the detector, with the use of the delayed time $t - \tau_{\text{dd}}$, returns it to the ready state 0.

The use of the prime on $dN'(t)$ indicates that this point process is associated with, but not the same as, $dN(t)$ which appears in Eq. (2.17). In fact $dN'(t)$ is equal to $dN(t)$, except for having some of the ones removed. This leads to

$$dN(t)dN'(t) = dN'(t). \quad (3.4)$$

Their expectation values are related by

$$E[dN'(t)] = \eta P_0 E[dN(t)]. \quad (3.5)$$

This reflects the two ways in which it is possible to have a ‘‘missed detection’’ [$dN(t) = 1$, $dN'(t) = 0$], namely, inefficiency or the APD’s not being in the ready state.

The use of the prime on $dN'(t)$ is to distinguish an avalanche from an observed avalanche, denoted by $dN(t)$. We note that such a distinction could be necessary if the efficiency of avalanche detection was less than unity. Although we do not include an avalanche detection inefficiency here we feel that the prime may aid the reader in understanding the need for the conditioning of the equations upon measurement of the presence of avalanches, which will occur later. The statistics of $dN'(t)$ are defined by its expectation value

$$E[dN'(t)] = \gamma_r P_1 dt. \quad (3.6)$$

This comes from our assumption that the avalanche’s reaching its threshold value is a Poisson process. Finally, the expectation value of $dN_{\text{dk}}(t)$ is

$$E[dN_{\text{dk}}(t)] = \gamma_{\text{dk}} P_0 dt, \quad (3.7)$$

reflecting another Poisson assumption, here for the dark counts.

3. Joint Stochastic Equations

The evolution of the quantum system is linked to that of the detector through $dN'(t)$. To obtain realistic quantum trajectories for the supersystem, which includes the quantum system and the classical detector states, we form the quantity

$$\rho_i = P_i \rho_N, \quad (3.8)$$

where ρ_i is the unnormalized state of the system given that the detector is in state i . The normalized system state is

$$\rho = \rho_0 + \rho_1 + \rho_2. \quad (3.9)$$

It is only possible to assume that the quantum and detector states factorize if there are no correlations between them. This will be the case until we average over unobserved processes. Thus the evolution of ρ_i can be found from

$$\rho_i + d\rho_i = (P_i + dP_i)(\rho_N + d\rho_N), \quad (3.10)$$

with the use of the increments in Eq. (2.17) and Eqs. (3.1)–(3.3). Conditioning upon the measurement of $dN(t)$ will be done as a later step. Substitution into Eq. (3.10) with the use of Eq. (3.4) gives

$$\begin{aligned} \rho_0(t + dt) &= \rho_0 + P_0 [dN(t)\mathcal{G}[c + \mu] - dt\mathcal{H}[iH + \frac{1}{2}c^\dagger c + \mu^* c + \frac{1}{2}|\mu|^2]] \rho_N \\ &\quad - [dN_{\text{dk}}(t) + dN'(t) - dN'(t - \tau_{\text{dd}}) + dN'(t)\mathcal{G}[c + \mu]] \rho_N \end{aligned} \quad (3.11)$$

$$\begin{aligned} \rho_1(t + dt) &= \rho_1 + P_1 [dN(t)\mathcal{G}[c + \mu] - dt\mathcal{H}[iH + \frac{1}{2}c^\dagger c + \mu^* c + \frac{1}{2}|\mu|^2]] \rho_N \\ &\quad + [dN_{\text{dk}}(t) + dN'(t) - dN'(t) + dN'(t)\mathcal{G}[c + \mu]] \rho_N \end{aligned} \quad (3.12)$$

$$\begin{aligned} \rho_2(t + dt) &= \rho_2 + P_2 [dN(t)\mathcal{G}[c + \mu] - dt\mathcal{H}[iH + \frac{1}{2}c^\dagger c + \mu^* c + \frac{1}{2}|\mu|^2]] \rho_N \\ &\quad + [dN'(t) - dN'(t - \tau_{\text{dd}})] \rho_N, \end{aligned} \quad (3.13)$$

remembering that all time arguments not stated are t . The linking of the quantum system and the APD is manifest in the $dN'(t)\mathcal{G}[c + \mu]\rho_N$ terms. If we were now to average over the emissions from the quantum system a correlation would still develop between the quantum state

and the APD.

4. Average Over Unobserved Processes

In order to obtain trajectories that are based on information that is realistically available in the laboratory, we must now average over unobserved processes. These are the photon emissions and dark counts. Strictly speaking, we should also average over $d\mathcal{N}'(t)$ since $d\mathcal{N}(t)$ was stated as specifying observed avalanches. However, we have not yet performed the conditioning due to the detection of the avalanches so these transitions will be left for the moment. Doing the averages before the measurement does not change the final result, but does make the equations more simple.

The relevant averages are Eq. (2.16), Eq. (3.5) and Eq. (3.7). Substitution of these into Eqs. (3.11)–(3.13) gives

$$\rho_0(t+dt) = \rho_0 + dt(\mathcal{L} - \gamma_{\text{dk}} - \eta\mathcal{J}[c + \mu])\rho_0 + d\mathcal{N}'(t - \tau_{\text{dd}})\rho \quad (3.14)$$

$$\rho_1(t+dt) = \rho_1 + dt\mathcal{L}\rho_1 + dt(\gamma_{\text{dk}} + \eta\mathcal{J}[c + \mu])\rho_0 - d\mathcal{N}'(t)\rho \quad (3.15)$$

$$\rho_2(t+dt) = \rho_2 + dt\mathcal{L}\rho_2 + d\mathcal{N}'(t)\rho - d\mathcal{N}'(t - \tau_{\text{dd}})\rho. \quad (3.16)$$

Note that ρ is given by Eq. (3.9) and \mathcal{L} is specified in Eq. (2.12).

5. Kushner-Stratonovich Equation

The probabilities assigned to the occupation of the three APD states by the realistic observer will undergo jumps in the event of an avalanche. If there is no avalanche then this also provides information to the observer about the probabilities. Let us first consider the effect of an avalanche. The new states will be given by

$$\rho_{i|d\mathcal{N}(t)=1} = \frac{\rho_i}{\text{Tr}[\rho_i]} P_{i|d\mathcal{N}(t)=1}, \quad (3.17)$$

where $P_{i|d\mathcal{N}(t)=1}$ is the probability of the APD being in state i given that an avalanche was observed. These probabilities can be simply calculated using Bayes conditional probability theorem [17]

$$P_{i|d\mathcal{N}(t)=1} = \frac{P_{d\mathcal{N}(t)=1|i}P_i}{P_{d\mathcal{N}(t)=1}}. \quad (3.18)$$

Here, $P_{d\mathcal{N}(t)=1|i}$ is the probability of observing an avalanche given that the APD is in state i , P_i is the prior occupation of state i and $P_{d\mathcal{N}(t)=1}$ is the probability of observing an avalanche. The relevant probabilities are

$$P_{d\mathcal{N}(t)=1|0} = P_{d\mathcal{N}(t)=1|2} = 0 \quad (3.19)$$

$$P_{d\mathcal{N}(t)=1|1} = \gamma_r dt \quad (3.20)$$

$$P_{d\mathcal{N}(t)=1} = \gamma_r P_1 dt. \quad (3.21)$$

This leads to the following three results for the states

$$\rho_{0|d\mathcal{N}(t)=1} = 0 \quad (3.22)$$

$$\rho_{1|d\mathcal{N}(t)=1} = \frac{\rho_1}{\text{Tr}[\rho_1]} \quad (3.23)$$

$$\rho_{2|d\mathcal{N}(t)=1} = 0. \quad (3.24)$$

If there is no observed avalanche then we need to consider probabilities of the form $P_{i|d\mathcal{N}(t)=0}$. This can be calculated from

$$P_{i|d\mathcal{N}(t)=0} = \frac{P_{d\mathcal{N}(t)=0|i}P_i}{P_{d\mathcal{N}(t)=0}}, \quad (3.25)$$

with the probabilities interpreted in the usual fashion. The relevant expressions are

$$P_{d\mathcal{N}(t)=0|0} = P_{d\mathcal{N}(t)=0|2} = 1 \quad (3.26)$$

$$P_{d\mathcal{N}(t)=0|1} = 1 - \gamma_r dt \quad (3.27)$$

$$P_{d\mathcal{N}(t)=0} = 1 - \gamma_r P_1 dt. \quad (3.28)$$

The new states based on these probabilities are, to first order in dt ,

$$\rho_{0|d\mathcal{N}(t)=0} = (1 + \gamma_r P_1 dt)\rho_0 \quad (3.29)$$

$$\rho_{1|d\mathcal{N}(t)=0} = (1 - \gamma_r dt + \gamma_r P_1 dt)\rho_1 \quad (3.30)$$

$$\rho_{2|d\mathcal{N}(t)=0} = (1 + \gamma_r P_1 dt)\rho_2. \quad (3.31)$$

We can summarize the avalanche and no avalanche conditioning with

$$\rho_{0|d\mathcal{N}(t)} = [1 + \gamma_r P_1 dt - d\mathcal{N}(t)]\rho_0 \quad (3.32)$$

$$\rho_{1|d\mathcal{N}(t)} = (1 - \gamma_r dt + \gamma_r P_1 dt)\rho_1 + d\mathcal{N}(t) \left(\frac{1}{\text{Tr}[\rho_1]} - 1 \right) \rho_1 \quad (3.33)$$

$$\rho_{2|d\mathcal{N}(t)} = [1 + \gamma_r P_1 dt - d\mathcal{N}(t)]\rho_2. \quad (3.34)$$

Although we have already formed joint stochastic equations and averaged over unobserved processes, the above equations represent the step in Fig. 1 where we obtain the KSEs for classical probabilities of the APD states.

6. Superoperator Kushner-Stratonovich Equation

The expressions in Eqs. (3.32)–(3.34) can now be substituted into the RHS of Eqs. (3.14)–(3.16), using Eq. (3.9) for ρ . However, a simplifying deduction is that when the APD resets, the state of the quantum system must be equal to ρ_2 , which will already be normalized. We can therefore replace ρ by ρ_2 as the multiplier of the $d\mathcal{N}'(t - \tau_{\text{dd}})$ term. Similar reasoning allows us to ignore terms such as $-d\mathcal{N}(t)\rho_2$ and $\gamma_r \text{Tr}[\rho_1]\rho_2 dt$ that will appear in the equation for ρ_2 . The assumption of a known initial APD state allows us to make the following statement: *at all times either $P_0 + P_1 = 1$ or $P_2 = 1$* . After an avalanche, the detector is known to be in state 2 for

a time τ_{dd} . Once the APD has been reset a realistic observer will not know if the APD is in state 0 or 1 (but will know that it is not in state 2) until the time of the next avalanche.

Also note that $d\mathcal{N}(t)d\mathcal{N}'(t) = d\mathcal{N}(t) = d\mathcal{N}'(t)$ since here we are assuming that all avalanches are detected by the observer. This allows us to drop the prime on the non-Markovian resetting terms $d\mathcal{N}'(t - \tau_{\text{dd}})$, rather than perform a repetitive conditioning on the observation of the resetting process, which is actually just inferred from the observation of an avalanche.

Making these simplifications after substitution into Eqs. (3.14)–(3.16) gives

$$d\rho_0 = dt(\mathcal{L} - \gamma_{\text{dk}} - \eta\mathcal{J}[c + \mu] + \gamma_{\text{r}}\text{Tr}[\rho_1])\rho_0 - d\mathcal{N}(t)\rho_0 + d\mathcal{N}(t - \tau_{\text{dd}})\rho_2 \quad (3.35)$$

$$d\rho_1 = dt[(\mathcal{L} - \gamma_{\text{r}} + \gamma_{\text{r}}\text{Tr}[\rho_1])\rho_1 + (\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})\rho_0] - d\mathcal{N}(t)\rho_1 \quad (3.36)$$

$$d\rho_2 = dt\mathcal{L}\rho_2 + d\mathcal{N}(t)\frac{\rho_1}{\text{Tr}[\rho_1]} - d\mathcal{N}(t - \tau_{\text{dd}})\rho_2. \quad (3.37)$$

The above equations are the realistic quantum trajectories based on the observation of avalanches in the APD. We term the equations SKSEs as we have obtained a quantum analog of the KSE in that from measurement we are conditioning the state of a supersystem that contains a quantum system, as explained in Sec. IC. The conditioned system state is

$$\rho_{\mathcal{N}} = \rho_0 + \rho_1 + \rho_2. \quad (3.38)$$

Note that Eqs. (3.35)–(3.37) are nonlinear, so that the state remains normalized. This is in contrast to the corresponding equations in [14], which were linear and generated unnormalized states. For simulation purposes, it is more convenient to use the unnormalized forms.

C. Ignoring Imperfections

Depending on the experimental situation, one or more of the imperfections in the APD may be negligible. Easy limits to take are those of $\eta = 1$ and $\gamma_{\text{dk}} = 0$, as these values can be substituted into the already derived equations. More involved limits are those where γ_{r}^{-1} and τ_{dd} become small. These imply that one or two of the detector states are superfluous. The first limit can be treated by a process of adiabatic elimination.

When γ_{r} is much larger than the system rates (given by the eigenvalues of \mathcal{L}) we can adiabatically eliminate state 1 of the detector. When ρ_1 responds quickly to changes, $\dot{\rho}_1$ in Eq. (3.36) can be set to zero to give the slaved value

$$\rho_1 = \frac{1}{\gamma_{\text{r}}}(\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})\rho_0. \quad (3.39)$$

Note that $\mathcal{L}\rho_1$ and $\gamma_{\text{r}}\text{Tr}[\rho_1]\rho_1$ have been ignored as compared to $\gamma_{\text{r}}\rho_1$ and that Eq. (3.39) does not require the

inclusion of a jump term due to avalanches (as these are inherited from the dependence upon ρ_0). As ρ_1 can now be determined from Eq. (3.39), the number of differential equations for real variables that need to be evolved is greatly reduced

If $\gamma_{\text{r}} \rightarrow \infty$ then state 1 can be ignored completely, with Eqs. (3.35)–(3.37) reducing to

$$d\rho_0 = dt(\mathcal{L} - \gamma_{\text{dk}} - \eta\mathcal{J}[c + \mu] + \text{E}[d\mathcal{N}(t)])\rho_0 - d\mathcal{N}(t)\rho_0 + d\mathcal{N}(t - \tau_{\text{dd}})\rho_2, \quad (3.40)$$

$$d\rho_2 = dt\mathcal{L}\rho_2 - d\mathcal{N}(t - \tau_{\text{dd}})\rho_2 + d\mathcal{N}(t)\frac{(\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})\rho_0}{\text{Tr}[(\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})\rho_0]}. \quad (3.41)$$

These equations describe the detector jumping straight from the ready state to the dead state. Thus, the exact time at which a charged e^-h^+ pair is created is known, but creation by a dark count or a photon is still indistinguishable. The statistics of $d\mathcal{N}(t)$ are now defined by $\text{E}[d\mathcal{N}(t)] = dt\text{Tr}[(\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})\rho_0]$ and we have $\rho_{\mathcal{N}} = \rho_0 + \rho_2$. Note that only one of ρ_0 and ρ_2 will contribute to $\rho_{\mathcal{N}}$ at any particular time. If the expression for $\text{E}[d\mathcal{N}(t)]$ were included in Eq. (3.40) then the $-dt\gamma_{\text{dk}}\rho_0$ term would cancel due to it only having an effect on the transition rate out of state 0, which is now effectively being monitored.

If, on the other hand, the dead time goes to zero, then state 2 of the detector can be removed. Thus, when an avalanche occurs, the detector jumps straight to the ready state. The relevant equations are then (with state 1 not adiabatically eliminated)

$$d\rho_0 = dt(\mathcal{L} - \gamma_{\text{dk}} - \eta\mathcal{J}[c + \mu] + \gamma_{\text{r}}\text{Tr}[\rho_1])\rho_0 - d\mathcal{N}(t)\left(\rho_0 - \frac{\rho_1}{\text{Tr}[\rho_1]}\right), \quad (3.42)$$

$$d\rho_1 = dt[(\mathcal{L} - \gamma_{\text{r}} + \gamma_{\text{r}}\text{Tr}[\rho_1])\rho_1 + (\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})\rho_0] - d\mathcal{N}(t)\rho_1. \quad (3.43)$$

Here, the statistics of $d\mathcal{N}(t)$ are defined as for Eqs. (3.35)–(3.37) but we now have $\rho_{\mathcal{N}} = \rho_0 + \rho_1$.

Finally, we consider $\gamma_{\text{r}} \rightarrow \infty$ together with $\tau_{\text{dd}} = 0$. This leaves an equation for ρ_0 only. It is,

$$d\rho_0 = dt(\mathcal{L} - \gamma_{\text{dk}} - \eta\mathcal{J}[c + \mu] + \text{E}[d\mathcal{N}(t)])\rho_0 + d\mathcal{N}(t) \times \left(\frac{(\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})}{\text{Tr}[(\eta\mathcal{J}[c + \mu] + \gamma_{\text{dk}})\rho_0]} - 1\right)\rho_0, \quad (3.44)$$

with $\text{E}[d\mathcal{N}(t)]$ defined as for Eqs. (3.40)–(3.41). If $\gamma_{\text{dk}} = 0$ then Eq. (3.44) reduces to Eq. (2.17). Obviously, here we have $\rho_{\mathcal{N}} = \rho_0$.

IV. PHOTORECEIVER

When the incident photon flux is high, as in homodyne detection, an avalanche photodiode cannot be used because of the long dead time. Instead, a p-i-n photodiode

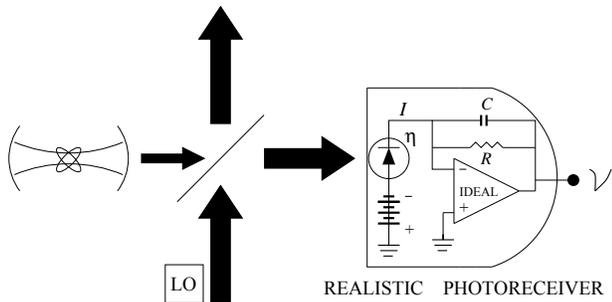


FIG. 4: Homodyne detection by a realistic photoreceiver of a TLA placed in an optical cavity. The photocurrent generated by a below breakdown p-i-n junction, of efficiency η , is input to a transimpedance amplifier. The voltage measured in the laboratory is \mathcal{V} .

is more appropriate. The inclusion of the high resistivity intrinsic region, i , allows the widening of the depletion region, which facilitates the absorption of light [30]. Because there is no avalanche, an external amplifier, such as a transimpedance amplifier (see Fig. 4) must be used [31]. We consider the whole system of photodiode plus amplifier to constitute the photoreceiver.

When a photon strikes the depletion region of the p-i-n junction, an electron-hole pair (that does not recombine) is produced with probability equal to the quantum efficiency η . The charge carriers drift under the influence of the below-breakdown reverse bias, and the resultant current I is fed into an operational amplifier (op-amp) set up as a transimpedance amplifier. This has a low effective input impedance, so that the diode acts as a current source, and I is converted into a voltage drop V across the feedback resistor, R . The capacitor C , in parallel with R , represents the total capacitance from the output of the op-amp back to its input, including capacitance added deliberately for the smoothing of noise and oscillations. If no electronic noise were present, the output voltage of the photoreceiver would be a filtered version of the input signal given, in the frequency domain, by

$$V(\omega) = \frac{-I(\omega)R}{1 + i\omega RC}, \quad (4.1)$$

where the negative is included because the output of the op-amp will be such that the input is kept at virtual ground.

It should be noted that if this were the case (that is, if there were no noise) then the input I could be perfectly reconstructed from the filtered signal V , since Eq. (4.1) is equivalent to

$$I(t) = -C\dot{V}(t) - \frac{V(t)}{R}. \quad (4.2)$$

Thus the resultant quantum trajectories would be no different from those of a photoreceiver with infinite bandwidth, since the input current to the amplifier is determinable. Everything of interest results therefore from

the presence of excess noise. We include only the Johnson noise V_J from the feedback resistor, which has a flat spectrum $S_J = 4k_B T R$. Here, k_B is the Boltzmann constant and T is the temperature. This simplification (neglecting contributions from voltage noise of the operational amplifier and dark counts from the photodiode) can be justified for practical receivers with $R \sim 10\text{k}\Omega$ [34].

The output voltage \mathcal{V} from the photoreceiver is given by sum of the filtered signal and the Johnson noise

$$\mathcal{V} = V + V_J. \quad (4.3)$$

Our aim is to find the quantum trajectory for the system, conditioned on continuously monitoring \mathcal{V} . Since the voltage V , which describes the detector state, is a continuous variable, in this case $\mathbb{S} = \mathbb{R}$, the real line, and the supersystem can be described by an operator function $\rho(V)$. Finding the stochastic equation of motion for $\rho(V)$ is quite involved.

A. Realistic Quantum Trajectories

In the following derivation we follow the order schematically illustrated in Fig. 1 more closely than in the case of the APD. The averaging over the unobserved processes is performed as a final step.

We begin by taking the output current I of the photodiode to be that from a perfect (apart from its efficiency η) unbalanced homodyne detection of the output field of the system. For a LO tuned to the resonant frequency of the system ω_0 , of power P , and phase Φ , the current is [5, 13]

$$I = e\sqrt{P/\hbar\omega_0} [\eta\langle e^{-i\Phi} c + e^{i\Phi} c^\dagger \rangle + \sqrt{\eta}\xi(t)], \quad (4.4)$$

where we have ignored the DC component due to the LO power and e is the electron charge. This current (in amps) is just the scaled current $J(t)$ introduced in Eq. (2.20), multiplied by $e\sqrt{P/\hbar\omega_0}$.

The Gaussian white noise [16] in Eq. (4.4) can be considered to arise from two independent sources: the Poisson statistics of the LO and the vacuum noise introduced by the inefficiency of the photodiode. The distinction between noise from these two sources will be illustrative when numeric simulations of the trajectories are performed (in a later paper). For now we need only consider $\xi(t)$.

1. Stochastic Master Equation

The evolution of the quantum system conditioned on I is given, in terms of the noise $\xi(t)$, by Eq. (2.21) and is restated here for the reader's convenience

$$d\rho_I = dt \{ \mathcal{L} + \sqrt{\eta}\xi(t)\mathcal{H}[e^{-i\Phi}c] \} \rho. \quad (4.5)$$

The subscript J in Eq. (2.21) has been changed to I . The fact that ρ on the RHS of Eq. (4.5) would usually be conditioned on previous values of $I(t)$ is to be understood. A realistic observer does not have direct access to $I(t)$, so later we will average over it.

2. Stochastic Differential Chapman-Kolmogorov Equation

Now, Eq. (4.1) is equivalent to the stochastic equation (4.2). Since the voltage V is not directly measured, we must consider a distribution $P(V)$ for it. Assuming that $C > 0$, and, for the moment, that I is known, Eq. (4.2) can be converted to an Itô stochastic Fokker-Planck equation for $P(V)$ conditioned on the photocurrent. This is done using the theory of Sect. II A, with the result being

$$dP_I(V) = \left(\frac{\partial}{\partial V} \frac{V + IR}{RC} + \frac{P\eta e^2}{2\hbar\omega_0 C^2} \frac{\partial^2}{\partial V^2} \right) P(V) dt. \quad (4.6)$$

A Fokker-Planck equation is one specific type of the broader class of differential Chapman-Kolmogorov equations. As in Eq. (4.5) we are using the convention that subscripts indicate that the increment is conditioned on that result. For example, the probability distribution for V given a current I is $P_I(V) \equiv P(V|I)$. It should be noted that Eq. (4.6) is different from a standard Fokker-Planck equation in that I still contains a white noise term. It is essential to retain this term as it is the same noise that appears in Eq. (4.4) - this will be used later in our derivation to link the evolution of the quantum system with that of the circuit of the photoreceiver.

3. Kushner-Stratonovich Equation

Next we need to determine the effect of the measurement of \mathcal{V} on $P(V)$. This can be calculated by using Bayes' theorem

$$P_{\mathcal{V}}(V) = \frac{P_{\mathcal{V}}(\mathcal{V})P(V)}{P(\mathcal{V})}. \quad (4.7)$$

Remembering that the Johnson noise is white, it follows from Eq. (4.3) that $P_{\mathcal{V}}(\mathcal{V})$ is a Gaussian with mean V and variance $4k_{\text{B}}TR/dt$. This enables us to calculate the properties of $P(\mathcal{V})$ from

$$P(\mathcal{V}) = \int dV P_{\mathcal{V}}(\mathcal{V})P(V) \quad (4.8)$$

$$= (2\pi\beta)^{-1/2} \int dV \exp(-(\mathcal{V} - V)^2/2\beta) P(V) \quad (4.9)$$

$$= \int dV \exp[(V - \langle V \rangle)(2\mathcal{V} - V - \langle V \rangle)/2\beta] P(V) \\ \times (2\pi\beta)^{-1/2} \exp(-(\mathcal{V} - \langle V \rangle)^2/2\beta). \quad (4.10)$$

For convenience, the variance $4k_{\text{B}}TR/dt$ in the above equations has been rendered by β . For an expansion

of the exponential inside the integral to be possible the factorization of Eq. (4.10) had to be made. The variance in V is much smaller than that of \mathcal{V} , which is of order $O(1/dt)$, due to the presence of the white Johnson noise.

Expanding terms in the above expression to leading order, we find eventually that

$$P(\mathcal{V}) = \frac{\exp(-(\mathcal{V} - \langle V \rangle)^2/2\beta)}{\sqrt{2\pi\beta}} \left[1 + O(dt^{3/2}) \right]. \quad (4.11)$$

It follows that we can write, correct to $O(1)$,

$$\mathcal{V} = \langle V \rangle + \sqrt{4k_{\text{B}}TR} \frac{d\mathcal{W}_J(t)}{dt}, \quad (4.12)$$

where $d\mathcal{W}_J(t)/dt$ is another Gaussian white noise source, independent of $\xi(t)$. Despite the fact that it arises from the Johnson noise, it is not the actual Johnson noise. It represents noise that the observer can determine from the measurement result and $\langle V \rangle$. Labeling the actual Johnson noise by $dW_J(t)$ we have

$$\mathcal{V} = V + \sqrt{4k_{\text{B}}TR} \frac{dW_J(t)}{dt}, \quad (4.13)$$

so that

$$d\mathcal{W}_J(t) = dW_J(t) + dt \frac{(V - \langle V \rangle)}{\sqrt{4k_{\text{B}}TR}}. \quad (4.14)$$

Substitution of $P_{\mathcal{V}}(\mathcal{V})$ and $P(\mathcal{V})$ into Eq. (4.7) gives to first order in dt

$$P_{\mathcal{V}}(V) = P(V) + d\mathcal{W}_J(t) \frac{V - \langle V \rangle}{\sqrt{4k_{\text{B}}TR}} P(V). \quad (4.15)$$

That is, we have derived the Kushner-Stratonovich equation

$$dP_{\mathcal{V}}(V) = d\mathcal{W}_J(t) \frac{(V - \langle V \rangle)}{\sqrt{4k_{\text{B}}TR}} P(V), \quad (4.16)$$

where the conditioning upon \mathcal{V} is contained in $d\mathcal{W}_J(t) = dt(\mathcal{V} - \langle V \rangle)/\sqrt{4k_{\text{B}}TR}$. This quantity is often referred to as the *residual* or *innovation* [18, 19].

4. Joint Stochastic Equations

To see how \mathcal{V} conditions the quantum system, we form the quantity $\rho(V) = \rho P(V)$, where ρ is here independent of $P(V)$ because we are still imagining I to be known at all times. The time evolution of $\rho(V)$, given that \mathcal{V} and I are known, is found from

$$\rho(V) + d\rho_{I,\mathcal{V}}(V) = [P(V) + dP_I(V) + dP_{\mathcal{V}}(V)] \\ \times (\rho + d\rho_I), \quad (4.17)$$

with the use of Eq. (4.6), Eq. (4.16) and Eq. (4.5).

Substitution into Eq. (4.17) gives

$$\rho_{I,\nu}(V) = \left[1 + d\mathcal{W}_J(t) \frac{(V - \langle V \rangle)}{\sqrt{4k_B T R}} + dt \left(\frac{\partial}{\partial V} \frac{V + IR}{RC} + \frac{P\eta e^2}{2\hbar\omega_0 C^2} \frac{\partial^2}{\partial V^2} \right) \right] P(V)$$

$$\times [1 + dt (\mathcal{L} + \sqrt{\eta}\xi(t)\mathcal{H}[e^{-i\Phi}c])] \rho. \quad (4.18)$$

The reader is reminded that $\xi(t)$ is uncorrelated with $d\mathcal{W}_J(t)$. Upon using I given in Eq. (4.4) and noting that it contains the white noise $\xi(t)$, the expansion of Eq. (4.18) to $O(dt)$ is

$$\begin{aligned} \rho_{I,\nu}(V) = & \left\{ 1 + d\mathcal{W}_J(t) \frac{(V - \langle V \rangle)}{\sqrt{4k_B T R}} + dt \left(\mathcal{L} + \frac{\partial}{\partial V} \frac{V}{RC} \right) \right. \\ & \left. + dt \frac{e}{C} \sqrt{\frac{P}{\hbar\omega_0}} \frac{\partial}{\partial V} [\eta \langle e^{-i\Phi}c + e^{i\Phi}c^\dagger \rangle + \sqrt{\eta}\xi(t)] \right\} \rho(V) \\ & + \left\{ dt \sqrt{\eta}\xi(t) + [dt\xi(t)]^2 \frac{e\eta}{C} \sqrt{\frac{P}{\hbar\omega_0}} \frac{\partial}{\partial V} \right\} \\ & \times [e^{-i\Phi}c\rho(V) + e^{i\Phi}\rho(V)c^\dagger - \langle e^{-i\Phi}c + e^{i\Phi}c^\dagger \rangle \rho(V)], \end{aligned} \quad (4.19)$$

where the form of the superoperator \mathcal{H} in Eq. (2.19) has been used.

5. Average Over Unobserved Processes

Finally, in reality \mathcal{V} is known but I is not. Therefore we should average over the noise $\xi(t)$, but keep the voltage noise $d\mathcal{W}_J(t)$. The relevant averages are Eq. (2.2), implying $dt^2 \mathbf{E}[\xi(t)\xi(t)] = dt$. An important consequence (necessary for consistency) is a cancellation of the terms involving $\langle e^{-i\Phi}c + e^{i\Phi}c^\dagger \rangle$. For convenience, we define a dimensionless voltage $v = V\sqrt{C/4k_B T}$, a rate $\gamma = 1/RC$ and a dimensionless noise power

$$N = \frac{4k_B T \hbar\omega_0}{\eta R P e^2}. \quad (4.20)$$

This last expression is the ratio of the Johnson noise power ($4k_B T R$) to the low-frequency power in \mathcal{V} due to the noise in the photocurrent ($R^2 e^2 P \eta / \hbar\omega_0$). The latter noise is that which would be present if a vacuum field was being combined with the LO field at the homodyne detection beam splitter.

Averaging over $\xi(t)$ and using v , γ and N we then obtain the following superoperator Kushner-Stratonovich equation (SKSE) for $\rho(v)$:

$$\begin{aligned} d\rho_{\mathcal{V}}(v) = & dt \left(\mathcal{L} + \frac{\gamma}{2N} \frac{\partial^2}{\partial v^2} + \gamma \frac{\partial}{\partial v} v \right) \rho(v) \\ & + dt \frac{\partial}{\partial v} \sqrt{\frac{\gamma\eta}{N}} [e^{-i\Phi}c\rho(v) + e^{i\Phi}\rho(v)c^\dagger] \\ & + \sqrt{\gamma} d\mathcal{W}_J(t) (v - \langle v \rangle) \rho(v). \end{aligned} \quad (4.21)$$

As noted below Eq. (4.16), the conditioning on the quantity that is measured in the laboratory, \mathcal{V} , is contained

in $d\mathcal{W}_J(t)$, with

$$\sqrt{\gamma} d\mathcal{W}_J(t) = dt\gamma \left(\sqrt{\frac{C}{4k_B T}} \mathcal{V} - \langle v \rangle \right). \quad (4.22)$$

Although the description of the state of the photoreceiver (by v) is essential in obtaining realistic trajectories, it is the evolution of the quantum system that we are most interested in. This state is given by

$$\rho_{\mathcal{V}} = \int \rho(v) dv. \quad (4.23)$$

The average $\langle v \rangle$, which appears in Eq. (4.21), is found from

$$\langle v \rangle = \int dv \text{Tr}[\rho(v)]v. \quad (4.24)$$

B. Effective Bandwidth

As mentioned earlier, if there were no excess (Johnson) noise added to the input signal, then this input signal could be perfectly reconstructed from the filtered output. However, in the case of realistic detection presented above there is noise, characterized by N , linearly added to the output of the filter. The extent to which information is lost due to this noise depends on the magnitude of the noise, the filter bandwidth, and the nature of the evolution of the monitored system. In this subsection, we argue that in the limit of small noise $N \ll 1$, the quality of the photoreceiver can be characterized by an *effective bandwidth* that depends upon the noise and the filter bandwidth. We expect the information loss to be small only if this effective bandwidth is large compared to the relevant system frequencies.

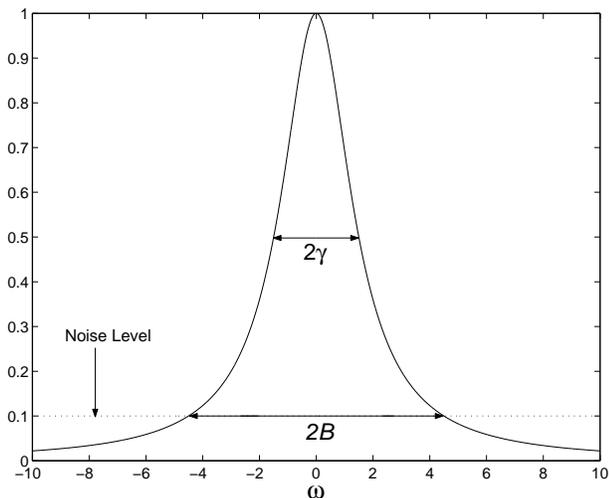


FIG. 5: The dimensionless quantities $\gamma^2/(\gamma^2 + \omega^2)$ and N are plotted, showing their frequency dependence (or lack of it). In this plot $\gamma = 1.5$ and $N = 0.1$. The frequency at which the vacuum signal drops in power to that of the noise is equal to the effective bandwidth B .

We identify the effective bandwidth, B , as being roughly the frequency at which the noise in a vacuum input signal is ‘lost’ in the Johnson noise. That is, the frequency at which the power of the vacuum input signal becomes equal to the Johnson noise power (which has a flat spectrum). Obviously the loss of very high frequency noise will invariably occur in any practical photoreceiver. The question of whether this loss is important or not is answered by looking at the eigenvalues of the Liouvillian. If the input noise is lost at frequencies that are well above the system rates then this noise would not significantly affect the evolution of the system. However, if noise is lost at frequencies at which the system can respond to, then the purity of the system state will decrease considerably. This is why we expect the effective bandwidth, as defined above, to be the relevant parameter for the photoreceiver.

Equating the vacuum signal and Johnson noise powers at the frequency $\omega = B$ gives

$$\frac{\gamma^2}{\gamma^2 + B^2} = \frac{4k_B T R \hbar \omega_0}{R^2 e^2 P \eta} = N \quad (4.25)$$

This is shown graphically in Fig. 5. Obviously if N is too large then this equation has no solution in the real numbers, and indeed the intuitive picture behind the argument fails. But for small N we have

$$B = \frac{\gamma \sqrt{1 - N}}{\sqrt{N}} \approx \frac{\gamma}{\sqrt{N}}. \quad (4.26)$$

This result will be investigated further for two different quantum systems in the following paper.

In terms of physical parameters

$$B = \frac{e}{2C} \sqrt{\frac{\eta P}{k_B T R \hbar \omega_0}}. \quad (4.27)$$

The photon flux $P/\hbar\omega_0$ of the LO cannot be increased arbitrarily as the regime in which the photodiode responds linearly to the field incident upon it must be adhered to. The resistor R cannot be reduced to zero as it is necessary for the reduction of the effective input impedance of the op-amp so that the photodiode can act as a current source. It also converts the current to a voltage, which presumably would be swamped by other, not considered, noise sources if reduced too far. Minimizing stray capacitances and the temperature of operation will improve the photoreceiver operation.

C. Adding Noise Only

The limit in which the power of the Johnson noise goes to zero has already been discussed, with the result being perfect detection. If the response time of the RC circuit vanishes then, provided Johnson noise is still added, the input signal is still obscured. For this to be a physically sensible consideration we assume that the capacitance, rather than the resistance, is zero.

The output voltage of the photoreceiver is then given by

$$\mathcal{V} = -IR + \sqrt{4k_B T R} \frac{dW_J(t)}{dt}, \quad (4.28)$$

as indicated by the form of Eq. (4.3). Here $dW_J(t)/dt$ represents the ‘real’ Johnson noise which is of a Gaussian nature. Using Eq. (4.4) and Eq. (4.20) gives (with $\Phi = 0$ for convenience),

$$\mathcal{V} \propto \frac{\eta}{1 + N} \langle c + c^\dagger \rangle + \sqrt{\frac{\eta}{1 + N}} \theta(t), \quad (4.29)$$

where $\theta(t)$ is normalized white noise consisting of a linear combination of the vacuum input noise and the Johnson noise. Comparison with Eq. (4.4) shows that the effect of adding noise but no filtering is to reduce the efficiency by a factor of $1 + N$. It can be shown using Bayesian analysis that this result holds for the form of the quantum trajectory as well.

V. CONCLUSION

The evolution of open quantum systems conditional upon detection results from *realistic* detectors cannot in general be generated by standard quantum trajectory theory (stochastic master equations). A method for generating this evolution was proposed in Ref. [14]. The crucial element is to include correlations between the system and classical detector states which cannot be observed in

practice. In this paper we have given a full description of this method, and shown how it can be applied in quantum optics. In particular, we derive realistic quantum trajectories for conditioning upon photon counting using an avalanche photodiode, and homodyne detection using a photoreceiver. These equations were presented in Ref. [14], but a full derivation was not given there.

In this paper we have not provided any solutions of the equations we have derived. To find and study these solutions is not a minor task, which is why it is reserved for the following paper [23]. There we use our equations to determine the evolution of a two-level cavity QED system, conditioned on four different types of detection (using the two detectors mentioned above). We also consider another system which can be treated analytically for realistic homodyne detection. Our study achieves five important aims, which are outlined in the introduction to that paper.

It is worth re-emphasizing the generality of our approach. It is applicable not just in quantum optics,

but in all areas where open quantum system theory is used. Prime examples are the mesoscopic measurement devices used in quantum electronics, such as single electron transistors. This device has received much attention as a possible measurement device for solid state qubits [35, 36, 37]. Conditional states may be used in quantum computation both for preparation, and for non-deterministic gate implementation [38]. In this context, the necessity of being able to relate realistically available results to the state of the qubit is obvious.

The greatest significance of our work is in the field of quantum control. The conditional state of a quantum system is synonymous with an observer's knowledge about that system. Thus (assuming all uncertainties are properly taken into account) it is by definition the optimum mathematical object to use to control that system. Properly taking into account detector imperfections is essential to building optimal control loops. We thus expect our theory to have broad applications in future quantum technology.

-
- [1] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Springer, Berlin, 1932); English translation (Princeton University Press, Princeton, 1955).
- [2] B. Misra and E. C. G. Sudarshan, *J. Math. Phys.* **18**, 756.
- [3] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).
- [4] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin, 1983).
- [5] H. J. Carmichael, *An Open Systems Approach to Quantum Optics* (Springer, Berlin, 1993).
- [6] J. Dalibard, Y. Castin and K. Mølmer, *Phys. Rev. Lett.* **68**, 580 (1992).
- [7] C. W. Gardiner, A. S. Parkins, and P. Zoller, *Phys. Rev. A* **46**, 4363 (1992).
- [8] V.P. Belavkin, "Nondemolition measurement and nonlinear filtering of quantum stochastic processes", pp. 245-66 of A. Blaquière (ed.), *Lecture Notes in Control and Information Sciences* **121** (Springer, Berlin, 1988).
- [9] V.P. Belavkin and P. Staszewski, *Phys. Rev. A* **45**, 1347 (1992).
- [10] A. Barchielli, *Quantum Opt.* **2**, 423 (1990).
- [11] A. Barchielli, *Int. J. Theor. Phys.* **32**, 2221 (1993).
- [12] H. M. Wiseman and G. J. Milburn, *Phys. Rev. A* **47**, 1652 (1993).
- [13] H. M. Wiseman and G. J. Milburn, *Phys. Rev. A* **47**, 642 (1993).
- [14] P. Warszawski, H. M. Wiseman and H. Mabuchi, *Phys. Rev. A* **65**, 023802 (2002).
- [15] C. W. Gardiner, *Quantum Noise* (Springer, Berlin, 1991).
- [16] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1985).
- [17] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis* (Addison-Wesley, Sydney, 1973).
- [18] T. P. McGarty, *Stochastic Systems and State Estimation* (Wiley-Interscience Publication, Sydney, 1974).
- [19] A. C. Doherty *et al.*, *Phys. Rev. A* **62**, 012105 (2000).
- [20] J. Gambetta and H. M. Wiseman, *Phys. Rev. A* **64**, 042105 (2001).
- [21] J.-I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, *Phys. Rev. Lett.* **78**, 3221 (1997).
- [22] A. C. Doherty and K. Jacobs, *Phys. Rev. A* **60**, 2700 (1999).
- [23] P. Warszawski and H.M. Wiseman, following paper.
- [24] H. M. Wiseman, *Phys. Rev. A* **49**, 2133 (1994).
- [25] H. M. Wiseman, *Mod. Phys. Lett. B* **9**, 629 (1995).
- [26] H. M. Wiseman and L. Diósi, *Chem. Phys.* **268**, 91 (2001); erratum **271**, 227 (2001).
- [27] L. Hardy, quant-ph/0101012; L. Hardy, "Why Quantum Theory?" in *Non-locality and modality*, edited by Tomasz Placek and Jeremy Butterfield (Kluwer, to be published), quant-ph/0111068.
- [28] G. Lindblad, *Commun. Math. Phys.* **48**, 199 (1976).
- [29] B. T. Debney and A. C. Carter, *Optical Fiber Sensors Components and Subsystems*, Ch. 4, Vol. 1 (Artech House, Boston 1996).
- [30] B. Garside, *Optical Fiber Sensors Components and Subsystems*, Ch. 5, Vol. 3 (Artech House, Boston 1996).
- [31] J. D. C. Jones, *Optical Fiber Sensor Technology*, Ch. 4 (Chapman and Hill, London 1995).
- [32] S. Takeuchi, J. Kim, Y. Yamamoto and H. H. Hogue, *Appl. Phys. Lett.* **74**, 1063 (1999).
- [33] A. Lacia *et al.*, *Appl. Phys. Lett.* **62**, 606 (1993).
- [34] H. Mabuchi, California Institute of Technology, private communication.
- [35] H. M. Wiseman *et al.*, *Phys. Rev. B* **63**, 235308 (2001).
- [36] M. H. Devoret and R. J. Schoelkopf, *Nature* **406**, 1039 (2000).
- [37] G. Schön, A. Shnirman and Y. Makhlin, cond-mat/9811029.
- [38] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).