INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF OPTICS B: QUANTUM AND SEMICLASSICAL OPTICS
J. Opt. B: Quantum Semiclass. Opt. **7** (2005) S179–S197 doi:10.1088/1464-4266/7/10/001

REVIEW ARTICLE

Modelling and feedback control design for quantum state preparation

Ramon van Handel, John K Stockton and Hideo Mabuchi

Physical Measurement and Control 266-33, California Institute of Technology, Pasadena, CA 91125, USA

E-mail: ramon@caltech.edu, jks@caltech.edu and hmabuchi@caltech.edu

Received 29 April 2005, accepted for publication 15 July 2005
Published 14 September 2005
Online at stacks.iop.org/JOptB/7/S179

Abstract
The goal of this article is to provide a largely self-contained introduction to the modelling of controlled quantum systems under continuous observation, and to the design of feedback controls that prepare particular quantum states. We describe a bottom-up approach, where a field-theoretic model is subjected to statistical inference and is ultimately controlled. As an example, the formalism is applied to a highly idealized interaction of an atomic ensemble with an optical field. Our aim is to provide a unified outline for the modelling, from first principles, of realistic experiments in quantum control.

Keywords: quantum feedback control, quantum state preparation, quantum probability, quantum filtering, continuous measurement

(Some figures in this article are in colour only in the electronic version)

1. Introduction

In recent years, advances in technology have enabled a proliferation of experiments where objects can be probed and manipulated near the fundamental quantum limits of performance. The manipulation and readout of single qubits with unprecedented coherence times both in condensed matter and in atomic setups, the reliable trapping, cooling and shot-noise limited continuous observation of single atoms in high-finesse optical cavities, and the production of various nonclassical states of light and of atomic ensembles is only a subset of recent achievements. The large degree of control that can be exerted at the quantum level suggests that classical engineering methodology can be fruitfully adapted to this new setting. In particular, it seems that the concept of feedback control should be of central importance in the engineering of reliable quantum technologies, as in the classical case.

This article is intended as an introduction to the theoretical description of quantum feedback control systems. We concentrate on a scenario that is common in quantum optical experiments, where the system to be controlled is brought in weak interaction with an external probe field which is subsequently detected. The detected signal can then be processed and fed back to the system through some actuator. There are various theoretical challenges in describing such a system:

- How does one model the system–probe interaction?
- How does one model a continuous measurement of the probe?
- How does one infer information on the system from the probe measurements?
- How does one design a feedback law that utilizes this information to achieve a particular control goal?

In the following we will address each of these questions in turn. Needless to say, it would be impossible to cover every aspect and intricacy of each of these questions within the scope of this article; rather, we aim to give a sufficiently detailed discussion to keep the article (mostly) self-contained, and refer to the bibliography for complete treatments.

As an example throughout the article, we discuss the preparation of entangled states of an atomic ensemble using feedback control. The model consists of an ensemble of atomic spins interacting dispersively with an optical probe, which is subjected to homodyne detection. Several recent experiments have exploited a similar setup to produce spin-squeezed states [28, 30, 47] which have applications in a variety

of metrology tasks, including magnetometers [29, 65] and atomic clocks [50, 59, 76].

We consider this model because it is illustrative in several respects. First, the model spans two quite different and interesting regimes. At short times the dynamics are approximately linear [65] and the model describes the production of spin squeezing. However, at long times the linear description is no longer valid, and we will show that then an eigenstate of the collective angular momentum of the ensemble (a Dicke state) is obtained. Although the long time limit described by this model is difficult to realize experimentally at this time, the consideration of the substantial differences between the regimes clearly demonstrates the challenges of quantum control. Second, the model is a convenient example to demonstrate the modelling of a quantum control system from first principles. Ultimately, by approaching the entire problem—from physical modelling to inference to control—in a systematic manner, we hope to provide a unified outline for future modelling efforts.

The article is roughly divided into two parts. The first part, consisting of sections 2–4, is somewhat technical in nature. Its goal is to obtain from first principles, using a simplified field-theoretical model of the interaction of an atomic ensemble with a probe field, the quantum filtering equation (52). To this end, we begin by reviewing in section 2 the statistical inference of quantum states. In section 3 we introduce a field-theoretical model of an atomic ensemble coupled to an electromagnetic probe field, and we discuss how it can be reduced to a stochastic equation. In section 4 we detail how to properly condition the ensemble state upon the results from continuous optical measurements of the field.

The second part, section 5, presents general principles of feedback control and demonstrates how they can be applied to enable quantum state preparation. This procedure is discussed in both the short time limit, where a linear approximation is valid, and in the long time limit, where a more complete description is required [66]. Section 5 is fairly independent from the first part of the article, and a reader who has some familiarity with the filtering equation, equation (52), could skip directly ahead to this section. We have attempted, however, to give in sections 2–5 a unified picture of quantum feedback control design, from the elementary physical interactions through feedback-enabled state preparation.

As we proceed, we attempt to review the literature concerning measurement and feedback control of atomic ensembles, while also putting into context related, but more mathematical, works concerning estimation and control. In the end, we hope to inspire further development in this field by highlighting the numerous connections between the problems of quantum control and problems considered in the culturally distinct context of the mathematics and control communities [4, 18, 71].

2. What is a quantum state?

Quantum mechanics describes the statistics of observable quantities, very much like classical probability theory. In fact, the foundation of quantum mechanics is just an extension of probability theory, as we will discuss in this section. Such a point of view allows us to apply classical constructions of probability theory directly to quantum models. Though this section contains no surprises, we aim to clarify the concepts and terminology used in the remainder of the article. We will pay particular attention to what is meant by a 'quantum state', an issue that must be resolved before we can discuss state preparation.

2.1. Classical probability

To set the stage for quantum probability we first discuss some of the elements of classical probability theory [75]. As an illustration, consider throwing two dice. The first ingredient we need in our theory is the *sample space*, usually denoted by Ω. This is just a set which describes all the 'microstates' of the system; in our case, it is the set of $6^2 = 36$ possible outcomes of a throw $11, 12, \ldots, 16, 21, 22, \ldots, 65, 66$. A *random variable* f is now a map $f : \Omega \to \mathbb{R}$. For example, we could define a random variable X that describes the sum of the two outcomes, i.e., $X(11) = 2$, $X(53) = 8$, etc.

To complete the picture we need to introduce an object that can provide answers to questions such as *what is the probability of having thrown 66?*, or *what is the probability of having thrown at least one three?* This is exactly provided by the notion of a *probability measure*. Note that we can represent any question as a subset of Ω; for example, our first question is represented by the set $\{66\}$, while the second is represented by $\{31, 32, \ldots, 36, 13, 23, \ldots, 63\}$. These sets (and the questions they represent) are called *events*. The probability measure $\mathbb{P}$ is a map that associates to every event a probability.

We can compose new events as follows. Given two events $A, B \subset \Omega$, the question *A or B?* is represented by $A \cup B$, whereas *A and B?* corresponds to $A \cap B$. In particular, the latter operation defines the *joint probability* $\mathbb{P}(A \cap B)$ of A and B. The probability measure needs to be consistently defined with respect to these operations in the sense that $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ if $A \cap B = \varnothing$, i.e., if A and B are mutually exclusive. Furthermore, $\mathbb{P}(\varnothing) = 0$ and $\mathbb{P}(\Omega) = 1$. In our example there is an equal probability of having thrown any combination; hence $\mathbb{P}(A) = \frac{1}{36}$ for any event A with a single element. Any other event can be constructed as a union of these 'elementary' events and its probability can be found using the formula for $\mathbb{P}(A \cup B)$.

Now suppose we wish to perform a particular observation on the system; we have already defined such observations (random variables) as maps on Ω. To obtain the probability of a particular observation, we simply invert the corresponding map. For example, the probability that we throw a combination that sums to 4 is $\mathbb{P}(X = 4) = \mathbb{P}(X^{-1}(4)) = \mathbb{P}(\{13, 22, 31\}) = \frac{1}{12}$. Hence the probability measure contains all the information available on the outcome of any observation, i.e., $\mathbb{P}$ represents the *state* of the system. The philosophy behind this choice of terminology is that physical theories exist to model the outcomes of observations; the 'state' is the object of the theory that gives rise to the statistics of any such observation.

Let us now consider classical state preparation. The physical mechanism that prepares the state of the dice, i.e., that causes every combination to have equal probability, is the throwing process. Suppose we want to prepare a different state, for example a state that has a high probability of obtaining two

sixes. We could obtain such a state by modifying the physical process that creates it. For example, we could engineer dice with an nonuniform mass distribution, so the sixth face is lighter than the other faces; then the rolling of the dice is more likely to terminate with the sixth face facing up.

There is a different way in which we can change the state. The *conditional probability* of event A given that we have measured event B is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{1}$$

Suppose we observe $X = 12$. Then the conditional probability of having thrown 66 is $\mathbb{P}(\{66\}|X = 12) = 1$, whereas without conditioning $\mathbb{P}(\{66\}) = \frac{1}{36}$. However, if we happened to measure $X \neq 12$ then $\mathbb{P}(\{66\}|X \neq 12) = 0$. This corresponds to the intuitive notion that if we *see* that we have thrown 66, then the probability that we have thrown 66 is one, no matter what its probability was before we had gained that information. However the probability that we would see 66 in the first place is only $\frac{1}{36}$. Hence we can create states by conditioning a 'prior' state on a measurement, but only very inefficiently: to prepare a state with high probability of obtaining 66, we have to keep throwing the dice until we happen to observe $X = 12$.

There is a final possibility which combines the two methods of state preparation. Suppose that we perform an observation not after the throw has completed, but while it is still in progress. Moreover, we allow ourselves to interfere with the dice: if the rolling dice threaten to terminate with a low value of X, we give them a shove so they keep rolling. This way the probability of throwing high numbers is elevated. In other words, we prepare the state of our choice by performing observation and applying *feedback* to the system dynamics. This crude example represents the type of state preparation that we consider in this article for quantum systems.

We conclude this section by introducing expectations and conditional expectations. If Ω is a finite countable set (which we have implicitly assumed in this section) then we may always decompose a random variable $f : \Omega \to \mathbb{R}$ as follows. The map f takes the values $f_i \in \mathbb{R}$ on disjoint subsets $S_i = f^{-1}(f_i) \subset \Omega$ such that $\bigcup_i S_i = \Omega$. Hence we can write

$$f(\omega) = \sum_i f_i \chi_{S_i}(\omega) \tag{2}$$

where χ_{S_i} is the indicator function of S_i, i.e., $\chi_{S_i}(\omega) = 1$ if $\omega \in S_i$, 0 otherwise. The *expectation* of f is given by

$$\mathbb{E}f = \sum_i f_i \mathbb{P}(S_i) \tag{3}$$

and represents the value that f takes on 'on average'. Note that the state $\mathbb{P}$ uniquely determines $\mathbb{E}$, but the converse is also true as by construction $\mathbb{E}\chi_S = \mathbb{P}(S)$ for any event $S \subset \Omega$. Hence we can equivalently define the state of the system by specifying the expectation of every system observable.

Similarly, we can define the *conditional expectation* of $f = \sum_i f_i \chi_{S_i}$ given that we have measured $g = \sum_i g_i \chi_{T_i}$:

$$\mathbb{E}(f|g)(\omega) = \sum_i \sum_j f_j \mathbb{P}(S_j|T_i)\chi_{T_i}(\omega). \tag{4}$$

Now $\mathbb{E}(\chi_S|g)(\omega) = \mathbb{P}(S|g = g(\omega))$. Hence we can equivalently define the state of the system, conditioned on a measurement of g, by specifying the conditional expectation of every system observable with respect to g.

Though entirely natural from a probabilist's point of view, it is not customary in physics to think of the conditional expectation as a random variable. One way to interpret equation (4) is that the random variable $\mathbb{E}(f|g)$ is the *best estimate* of f given g [60, 75]. To see this, first note that $\mathbb{E}(f|g)$ is by construction a function of g: $\mathbb{E}(f|g)(\omega) = \mathcal{X}(g(\omega))$ where we define $\mathcal{X} : g_i \mapsto \sum_j f_j \mathbb{P}(S_j|T_i)$. It is not difficult to show that, of all functions $\mathcal{X}'$, the one that minimizes the least-squares criterion $\mathbb{E}[(f - \mathcal{X}'(g))^2]$ is exactly $\mathcal{X}' = \mathcal{X}$. This is precisely what we mean by $\mathbb{E}(f|g)$ being the *best* estimate of f *given* g. Evidently this idea is equivalent, or in some sense dual, to the notion of a conditional state that we introduced earlier.

2.2. Quantum probability

We will now formulate quantum mechanics in the same language as the classical case [53, 68]. An observable (random variable) in quantum theory is given by a self-adjoint operator F on some complex Hilbert space $\mathcal{H}$. Assuming $\mathcal{H}$ is finite-dimensional, we always have the spectral decomposition

$$F = \sum_i f_i P_i \tag{5}$$

where $f_i \in \mathbb{R}$ are the eigenvalues of F and $P_i = P_i^2 = P_i^\dagger$ are projection operators onto the corresponding eigenspaces. The picture is completed by introducing a map $\mathbb{E} : \cdot \mapsto \mathrm{Tr}[\cdot\rho]$ with some $\rho = \rho^\dagger \geqslant 0$, $\mathrm{Tr}\,\rho = 1$. Then $\mathbb{E}F$ is the expectation of the observable F. In terms of the spectral decomposition

$$\mathbb{E}F = \sum_i f_i \mathbb{E}P_i. \tag{6}$$

Clearly the projectors P_i play the role of events χ_{S_i} in the classical theory. Indeed, a measurement of F yields the outcome f_i with probability $\mathbb{E}P_i$. Thus any quantum observable is identical to a classical random variable.

We can make the correspondence explicit in the following way. As we are free to choose any basis in the Hilbert space, we may always choose a basis in which F is diagonal. We can then interpret the diagonal elements of F as the values of the random variable f, where Ω is just the set of diagonal entries: $f : i \mapsto F_{ii}$. The P_i now correspond exactly to indicator functions on Ω and $\mathbb{P}(S) = \sum_i \chi_S(i)\rho_{ii}$. Note that the underlying Hilbert space plays a passive role in the theory, just like the sample space Ω in classical probability—the central element of the theory is the set of observables we are interested in. As long as we are interested in a set of observables that all commute with each other, then quantum and classical probability are identical theories: commuting observables can be simultaneously diagonalized, so we can follow the above 'recipe' to transform between the classical and quantum descriptions. In other words, classical probability theory is a special case of quantum probability theory.

The embedding of classical in quantum probability allows us to carry over directly concepts from classical probability to sets of commuting quantum observables. For example, in the classical case we defined the *joint probability* of two events A and B as $\mathbb{P}(A \cap B) = \mathbb{E}(\chi_A \chi_B)$. This carries

over directly to the quantum case for two quantum events P, Q as long as they commute: i.e., the joint probability of P and Q is $\mathbb{E}(PQ) = \mathrm{Tr}[PQ\rho]$. Similarly, we obtain an expression for conditional expectation for two commuting quantum observables $F = \sum_i f_i P_i$ and $G = \sum_i g_i Q_i$,

$$\mathbb{E}(F|G) = \sum_i \sum_j f_j \frac{\mathbb{E}(P_j Q_i)}{\mathbb{E}(Q_i)} Q_i \tag{7}$$

which is itself an observable as in the classical case, interpreted as the best estimate (in the least mean square sense) of F given G. Note that these are not even 'quantum analogues' of classical concepts—these are entirely *classical* operations. We can obtain these expressions by writing the commuting set of events in the diagonal basis, transforming to the classical picture, applying the classical operation, and transforming back to the quantum picture in the original basis.

What makes quantum probability different from classical probability is the existence of noncommuting observables. For events or observables that do not commute the classical probabilistic concepts do not make any sense: for example, the joint probability of P, Q with $[P, Q] \neq 0$ cannot be unambiguously defined as $\mathbb{E}(PQ) \neq \mathbb{E}(QP)$. Similarly, $\mathbb{E}(F|G)$ cannot be defined for $[F, G] \neq 0$. Hence we *do not allow* simultaneous measurement or statistical inference of noncommuting observables. The fact that noncommuting observables are inherent to quantum models restricts the amount of information that can be obtained from the system by measurement.

Once we have fixed a commuting family of observables to measure, however, the measurement process is reduced to straightforward application of classical probability theory. In particular, even if we are interested in modelling a pair of observables F and G that do not commute, we can still perform statistical inference as long as both observables commute with the observation H. After all, by equation (7), $\mathbb{E}(F|H)$ and $\mathbb{E}(G|H)$ commute and can hence be measured simultaneously, even though F and G do not commute. We will repeatedly exploit this fact throughout this article.

To illustrate these ideas, consider the example of a single spin-$\frac{1}{2}$ atom, and suppose we are interested in controlling the spin observables (Pauli matrices) $\sigma_x, \sigma_y, \sigma_z$. We run into problems if we try to directly measure σ_z, as this does not commute with σ_x and σ_y. Because the best estimate of σ_x or σ_y with respect to σ_z is undefined, it is unclear in what sense one could control σ_x and σ_y if we keep observing σ_z.

We have already hinted at the solution to this problem: we must observe a fourth observable X that commutes with $\sigma_x, \sigma_y, \sigma_z$. Then all three conditional expectations are well defined. A famous example of this procedure is the Stern–Gerlach apparatus: in this case the atom passes through a strong magnetic gradient which correlates the spin observables $\sigma_{x,y,z}$ with the spatial position X of the atom. By measuring X, which commutes with $\sigma_{x,y,z}$, we can form best estimates of the latter three observables, and thus at least conceptually these can be controlled.

In practice the Stern–Gerlach device is not a good system for controlling the spin, as the observable X is a different degree of freedom of the same atom that carries the spin. When the atom hits the screen, enacting a measurement of X, the atom is effectively destroyed and there is no point in updating the spin state for further control. The approach we take in this article is a realistic, though conceptually identical, version of this example. Instead of coupling the atomic spin to the atomic position, the spin interacts with an external electromagnetic field. Even though photodetection of the field is destructive this will not affect the atom itself.

The *quantum state* is an object that associates an expectation to the relevant set of observables. We refrain from defining the quantum state as the *density matrix* ρ. The properties that any expectation map must obey imply that we can always find a density matrix ρ such that the expectation can be expressed as $\mathbb{E}F = \mathrm{Tr}[\rho F]$ for the relevant set of observables F. What this relevant set is, however, depends on the context.

To illustrate this subtle distinction let us consider again the Stern–Gerlach example. Before conditioning we consider the four observables $\sigma_{x,y,z}$ and X. Hence we naturally express the state as a density matrix on $\mathcal{H}_s \otimes \mathcal{H}_q$, the tensor product of the atomic spin and position Hilbert spaces. However, we can only condition observables on X that commute with X. Hence after conditioning a spin-position density matrix is no longer meaningful, as many observables on the position Hilbert space (for example, momentum) will have an undefined conditional expectation[1].

To find the natural state after conditioning, recall that $\mathbb{E}(\sigma_{x,y,z}|X)$ all commute. Hence we can describe them as classical random variables $s_{x,y,z}(\omega)$ on some probability space Ω. To express the state as a density matrix, then, we must also make it random: we define $\rho(\omega)$ on $\mathcal{H}_s$ through $s_{x,y,z}(\omega) = \mathrm{Tr}[\rho(\omega)\sigma_{x,y,z}]$. This conforms to the intuitive idea that, after measurement, the conditional state is itself a classical random variable, where Ω is simply the set of possible outcomes of X. It also highlights, however, that in order to talk sensibly about state preparation we must carefully select which observables we wish to specify. Though the 'dual' description in terms of a density matrix is often more economical, we will often find it both conceptually and technically simpler to obtain results by considering conditional expectations to be observables on $\mathcal{H}_s \otimes \mathcal{H}_q$.

The three methods of state preparation discussed in the previous section carry over directly to the quantum case. All these methods have been discussed to various extent in the literature; references to their various experimental implementations will be given in section 5. The first method corresponds to designing a Hamiltonian whose time evolution generates the desired state. The drawback of this method is that such a Hamiltonian may be highly nonlinear and difficult to engineer in practice.

The second method corresponds to conditioning. As we saw in the example above, to do this we must 'open' the system by introducing another observable. We emphasize, however, that there is no physical 'collapse' associated to the actual measurement: we just use *classical* conditioning to update our state of knowledge. The drawback of this method is that

[1] Of course if we were interested in both position and momentum, we could couple to yet another observable that commutes with $\sigma_{x,y,z}$ as well as position and momentum. This way we move further and further down the 'Heisenberg chain'. Ultimately, however, we have to make an observation, which will rule out some incompatible observables.

the outcome of the measurement is random and will not always result in the desired state; particularly in cases where the state is prepared with low probability, this may not be a desirable option.

The third method, which is the main topic of this article, is that of conditioning with feedback. The advantage of such a method is that it can be implemented with simple Hamiltonians, while it does not suffer from the indeterminism of pure conditioning. The method can also be more robust than simple Hamiltonian evolution, as it is not as sensitive to, for example, timing errors or precise knowledge of experimental parameters [65, 70]. However, successful implementation of such a method requires sensitive, continuous-time quantum-limited measurements and fast in-line signal processing, techniques that have only recently become available.

We separate the development of quantum feedback control into three parts. In order to interpret the measurement current and feedback we must develop a physical model of the system and its interaction with the environment. This first step, the *physical modelling* step, embodies the 'physical content' of the problem. In the second step we condition the system dynamics based on an observation of the environment. This *statistical inference* step is, as we have discussed, entirely classical in nature. The third step is the *control problem*, finding a control law that will prepare the desired state. In the following sections we consider each of these problems separately.

Note that all the constructions in this section can be generalized to infinite-dimensional Hilbert spaces in the quantum case, and to infinite or continuous Ω in the classical case. However, a rigorous discussion of the associated mathematics is beyond the scope of this article. Though conceptually the finite and infinite cases are very similar, we will need to extend the finite techniques somewhat in section 4.2 in order to deal with continuous systems. For lucid introductions to the general theories of classical and quantum probability we refer to [75] and [53], respectively.

3. The physical model: from QED to stochastic equations

In this section we will describe a microscopic model for the class of systems we consider. The model consists of an atomic ensemble coupled weakly to an external electromagnetic field which is ultimately detected.

3.1. System model from quantum electrodynamics

It is well known from quantum electrodynamics [13, 54] that the observable for the free electric field is given by

$$\mathbf{E}(\mathbf{r},t)=\sqrt{\frac{\hbar}{(2\pi)^3\varepsilon_0}}\sum_s\int\sqrt{\frac{\omega}{2}}(\mathrm{i}a_{\mathbf{k},s}\boldsymbol{\varepsilon}_{\mathbf{k},s}\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)}+\mathrm{h.c.})\,\mathrm{d}^3k \tag{8}$$

where $\omega = c|\mathbf{k}|$, $\boldsymbol{\varepsilon}_{\mathbf{k},s}$ are polarization vectors and $a_{\mathbf{k},s}$ are plane wave (Fourier) mode annihilation operators that satisfy the commutation relations $[a_{\mathbf{k},s}, a^\dagger_{\mathbf{k}',s'}] = \delta^3(\mathbf{k}-\mathbf{k}')\delta_{ss'}$. We assume that the atomic ensemble (centred at the origin) interacts with the field predominantly through its collective dipole moment; i.e., the interaction Hamiltonian will be of the form $H_\mathrm{I}(t) = -\mathbf{d}(t)\cdot\mathbf{E}(0,t)$ where $\mathbf{d}(t)$ is the ensemble dipole operator. In practice there will be some ultraviolet cutoff, which we can obtain for example by averaging the electric field over the volume of the cloud of atoms instead of evaluating it at the origin. We will write

$$\mathbf{E}^{(+)}(\mathbf{r},t)=[\mathbf{E}^{(-)}(\mathbf{r},t)]^\dagger=\sum_s\int g(\mathbf{k})a_{\mathbf{k},s}\boldsymbol{\varepsilon}_{\mathbf{k},s}\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)}\,\mathrm{d}^3k \tag{9}$$

where $\mathbf{E} = \mathrm{i}(\mathbf{E}^{(+)} - \mathbf{E}^{(-)})$ and $g(\mathbf{k})$ is the mode function. For example, $g(\mathbf{k}) \propto \sqrt{\omega}\mathrm{e}^{-\mathrm{d}\omega^2}$ if we average $\mathbf{E}$ over a spatial Gaussian distribution.

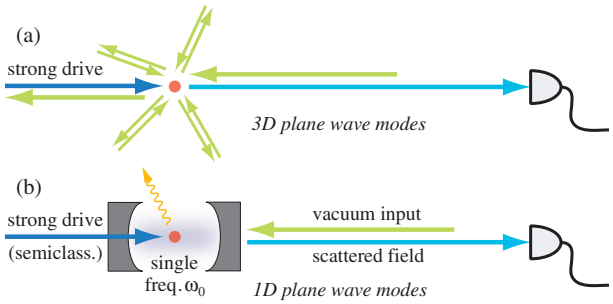

Figure 1. (a) Schematic diagram of the full interaction Hamiltonian, where all plane wave modes in three dimensions interact with an atomic ensemble. One of the incoming modes is coherently excited with frequency ω_0 and large amplitude; the coherent excitation scatters mainly in the forward direction. The remaining incoming modes are in the vacuum state and drive spontaneous emission. (b) Simplified one-dimensional model, where plane waves are scattered off a single-mode cavity in one direction only. The strong driving field is treated semiclassically and coherently excites the cavity mode, which has frequency ω_0. Spontaneous emission can be added to the model in a phenomenological manner.

The full interaction is sketched in cartoon form in figure 1(a). The atomic ensemble interacts through its dipole moment with all plane wave modes in three dimensions. A strong, focused laser beam at frequency ω_0 is modelled by bringing the corresponding modes into a large-amplitude coherent state. The drive is scattered predominantly in the forward direction, and is ultimately detected. The remaining modes are in the vacuum state and drive spontaneous emission of the ensemble in all directions. This essentially complete description of the interaction embodies all the physics of the problem, and thus allows one to predict quantities such as the spontaneous emission rate. The full picture is also very complicated, however, as it requires a detailed analysis of the atomic structure, a partitioning of the field into observed and unobserved modes, etc. Such an analysis is beyond the scope of this article.

Instead, we will investigate a highly simplified model (figure 1(b)) that is widely used in quantum optics [27, 78, 83]. To justify such a model, we claim that most of the interesting physics occurs in the direction of the driving laser, as most of the light is scattered forward and observed in this direction. Hence we can approximate the system by a one-dimensional model where only the forward modes are treated exactly and the strong drive is treated semiclassically. Spontaneous emission into the eliminated modes is unobserved, and hence we could include it phenomenologically by adding decoherence. Finally, to simplify the interaction with the ensemble, we place

the ensemble into a leaky single-mode cavity. This allows us to treat the interaction between the ensemble and the field for a single frequency only, that of the cavity mode, which is chosen to be at the laser driving frequency ω_0. The cavity dynamics is then adiabatically eliminated to give an effective interaction between the ensemble and the external field.

Let us systematically work out this simplified model. We begin by treating the one-dimensional external field that is ultimately detected. We can obtain an expression for the field by integrating equation (8) over a transverse area [27], or alternatively by directly quantizing the wave equation in one dimension [78]. We obtain

$$E(z,t) = \sqrt{\frac{\hbar}{2\pi\varepsilon_0 c}} \int_0^\infty \sqrt{\frac{\omega}{2}} (\mathrm{i}a_\omega \mathrm{e}^{-\mathrm{i}\omega(t-z/c)} + \mathrm{h.c.})\,\mathrm{d}\omega \quad (10)$$

for the electric field intensity in a single polarization state (we will assume polarized light), where $[a_\omega, a_{\omega'}^\dagger] = \delta(\omega - \omega')$. The annihilators a_ω correspond to plane wave modes in the z-direction, $\mathbf{k} = (\omega/c)\hat{\mathbf{z}}$, where positive z is defined to be on the left in figure 1. Thus the field for $z < 0$ is propagating towards the cavity mirror in time, whereas $z > 0$ propagates away from the mirror. In practice the cavity mirror will reverse the propagation direction, so we can reinterpret $z > 0$ as the component of the field propagating toward the detector, whereas $z < 0$ is the incident part of the field[2].

We now introduce the cavity mode with annihilation operator $b(t) = b\mathrm{e}^{-\mathrm{i}\omega_0 t}$. The interaction Hamiltonian between the cavity mode and the external field is given by

$$\begin{aligned} H_{\mathrm{CF}} &= \hbar \int_0^\infty \kappa(\omega)(\mathrm{i}a_\omega^\dagger b \mathrm{e}^{\mathrm{i}(\omega-\omega_0)t} + \mathrm{i}a_\omega^\dagger b^\dagger \mathrm{e}^{\mathrm{i}(\omega+\omega_0)t} + \mathrm{h.c.})\,\mathrm{d}\omega \\ &= \mathrm{i}\hbar(b(t) + b(t)^\dagger)(E^{(-)}(0,t) - E^{(+)}(0,t)) \end{aligned} \quad (11)$$

where we have used

$$E^{(+)}(z,t) = [E^{(-)}(z,t)]^\dagger = \int_0^\infty \kappa(\omega) a_\omega \mathrm{e}^{-\mathrm{i}\omega(t-z/c)}\,\mathrm{d}\omega. \quad (12)$$

Here $\kappa(\omega)$ does not only depend on the external field but also on the frequency-dependent transmission of the cavity mirror, and is unitless. An interaction Hamiltonian of this form can be obtained from the QED Hamiltonian by expanding it into 'quasi-modes' corresponding to either inside or outside the cavity; see [15].

We will briefly describe the remaining Hamiltonians. The interaction Hamiltonian of the cavity mode with a resonant classical drive is given by

$$H_{\mathrm{D}} = \hbar\mathcal{E}(b + b^\dagger) \quad (13)$$

where $\mathcal{E}$ is the drive amplitude. Spontaneous emission is treated by introducing another field $\tilde{E}$ with annihilation operators c_ω, called the side channel, and adding another Hamiltonian $H_\perp = -d(t)\tilde{E}(0,t)$ where $d(t)$ is an atomic dipole operator. Unlike the forward channel E, which we will ultimately observe, the side channel is left unobserved. This is a simple but effective way to model the partitioning of the full three-dimensional field $\mathbf{E}(\mathbf{r}, t)$ into observed and unobserved modes.

The atomic Hamiltonian H_{A} and the ensemble–cavity mode interaction H_{AC} are more variable, as they depend on the structure of the atoms in the ensemble. In particular, we get drastically different behaviour when the atoms have a transition that is resonant with the cavity mode than in the far detuned case. We will consider a specific example in section 3.3.

3.2. *Quantum noise and the Markov limit*

The discussion in the previous section was based entirely on 'mechanical' arguments; i.e., the electric field emerged naturally by quantization of Maxwell's equations and the dipole coupling to matter. Any physical model ultimately has its roots in this level of description. However, we have already discussed that the foundations of quantum theory are essentially a glorified probability theory, where any observable is equivalent to a random variable on some probability space. As we will be interested in observations of the field, it is essential to make the connection between the physical model and its manifestation as a (quantum) probabilistic dynamical system.

We can consider equation (12) as the Fourier transform of the operator distribution $\kappa(\omega)\theta(\omega)a_\omega$, where θ is the step function. We will always take the incoming field to be in the vacuum state as in figure 1. Each a_ω can be thought of as an independent quantum 'complex Gaussian' random variable, in the sense that its 'real' and 'imaginary' parts $a_\omega + a_\omega^\dagger$ and $\mathrm{i}a_\omega^\dagger - \mathrm{i}a_\omega$ are precisely Gaussian random variables. Hence $E^{(+)}(0,t)$ will be some sort of quantum complex Gaussian noise. Note that the two quadratures $E^{(+)} + E^{(-)}$ and $\mathrm{i}E^{(-)} - \mathrm{i}E^{(+)}$ do not commute, so we cannot interpret $E^{(+)}(0,t)$ as a classical complex noise. We would now like to consider the Heisenberg equation (in the interaction picture with respect to the field dynamics)

$$\dot{X}_t = \frac{\mathrm{i}}{\hbar}[H_\perp + H_{\mathrm{A}} + H_{\mathrm{AC}} + H_{\mathrm{D}} + H_{\mathrm{CF}}(t), X_t] \quad (14)$$

as being driven by the noise $\mathrm{i}E^{(-)} - \mathrm{i}E^{(+)}$, together with an observation of the field which need not commute with the driving noise. Then the statistical inference step can be formulated as finding the best estimate of the noisy time evolution of atomic observables given noisy observations of the field.

Similar problems have been studied in classical probability for about a century, and the main lessons learned there appear to carry over to the quantum case. In particular:

(i) Statistical inference of continuous-time processes is essentially intractable unless we approximate the noise process by a *white noise*. In this case, the time evolution of the system is Markovian [60] (i.e., the distribution of future system states depends only on the present state and not on past history) and statistical inference is described by the elegant theory of Markov nonlinear filtering [16, 51].

(ii) Dealing with white noise directly is possible, but the resulting theory is very technical due to the fact that white noise is an extremely singular object [35, 43, 45]. It is much easier to build a theory from a Wiener process, the integral of white noise, which is at least continuous [60].

[2] In a full three-dimensional description the cavity mirror would be modelled by an interaction Hamiltonian that scatters into the backward propagating modes with terms such as $a_{-\mathbf{k}}^\dagger a_{\mathbf{k}}$. In the one-dimensional case, however, we can simply absorb this reflection into the definition of the field.

We fill follow a similar program below for quantum systems; i.e., we will first find a Markov approximation of the full field-theoretic model described previously, then develop a theory of quantum Markov filtering.

Before we embark on this path, it should be mentioned that the problem with coloured noise takes on an even more severe form in the quantum case. In the classical case the problem is mainly technical; there is no conceptual problem associated to statistical inference with coloured noise, but it is not possible to obtain filtering equations in a recursive form [14]. In the quantum case, however, it is not even clear what we mean by an observation of coloured noise, let alone the associated statistical inference problem, as the field operators may not commute with themselves at different times or with the system [25]. There is as yet no satisfactory solution to this problem; in particular, a satisfactory theory of quantum non-Markovian continuous measurement has yet to be developed. As we will see, however, these problems do not appear in the Markov case.

3.2.1. Classical and quantum stochastic differential equations. Let us briefly review the classical concept of a stochastic differential equation (SDE) [60]. We denote by W_t a one-dimensional Wiener process. It is defined on a probability space Ω where each $\omega \in \Omega$ corresponds to a single sample path $\{W_t(\omega)\}$ of the Wiener process. Heuristically the time derivative $\dot{W}_t$ would be white noise, so we wish to give meaning to a differential equation of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}X_t(\omega) = f(X_t(\omega)) + \sigma(X_t(\omega))\frac{\mathrm{d}W_t(\omega)}{\mathrm{d}t}. \tag{15}$$

However, this equation makes no mathematical sense as W_t is differentiable with probability zero. The solution is to rewrite it as an integral equation

$$X_t = X_0 + \int_0^t f(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}W_s \tag{16}$$

and then to define the stochastic integral. As a notational analogy with ordinary differential equations we will also write

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t \tag{17}$$

which is equivalent to (16) by definition.

Itô defined a stochastic integral in the following way:

$$\int_{t_0}^{t_n} f_s\,\mathrm{d}W_s = \lim_{|t_{i+1}-t_i|\to 0}\sum_{k=0}^{n-1} f_{t_k}(W_{t_{k+1}} - W_{t_k}). \tag{18}$$

Precisely in which sense the limit is taken is a central construction in Itô's theory which we gloss over. A different definition, due to Stratonovich, is

$$\int_{t_0}^{t_n} f_s \circ \mathrm{d}W_s = \lim_{|t_{i+1}-t_i|\to 0}\sum_{k=0}^{n-1} \tfrac{1}{2}(f_{t_{k+1}} + f_{t_k})(W_{t_{k+1}} - W_{t_k}). \tag{19}$$

It is a signature of the singularity of the problem that these two integrals do not give the same answer; such integrals would necessarily be the same if we could interpret them in the Riemann–Stieltjes sense. It is now ambiguous, however, how we should interpret equation (16).

A major difference between the two integrals is their transformation property. Ordinary Riemann–Stieltjes integrals obey the Leibnitz rule $\mathrm{d}(X_tY_t) = Y_t\,\mathrm{d}X_t + X_t\,\mathrm{d}Y_t$ (we use the shorthand notation of equation (17)). It turns out that this property is also obeyed by the Stratonovich integral (19). The Itô integral, on the other hand, obeys the modified transformation property $\mathrm{d}(X_tY_t) = Y_t\,\mathrm{d}X_t + X_t\,\mathrm{d}Y_t + \mathrm{d}X_t\,\mathrm{d}Y_t$, where we use the Itô rules $\mathrm{d}W_t^2 = \mathrm{d}t$, $\mathrm{d}t^2 = \mathrm{d}W_t\,\mathrm{d}t = 0$ to evaluate the rightmost term. Similarly, the Itô transformation rule for arbitrary functions becomes

$$\mathrm{d}g(X_t) = g'(X_t)\,\mathrm{d}X_t + \tfrac{1}{2}g''(X_t)\,\mathrm{d}X_t^2. \tag{20}$$

Note how the shorthand notation of equation (17) allows us to express these deep results in a compact way. The power of the Itô calculus lies in the fact that complicated transformations can be performed using only simple symbolic manipulations.

The fact that the Stratonovich integral obeys the Leibnitz rule suggests that physical systems should be described by a Stratonovich SDE; after all, if we take a physical system with a smooth driving force, and add some noise to this force, we do not expect the transformation properties of the system to change. We will investigate this further in the next section. On the other hand, the Itô integral has the nice property that its expectation vanishes[3], which suggests that Itô SDEs are natural from the point of view of statistical inference. Fortunately we can have it both ways, as there is a conversion formula between Itô and Stratonovich SDEs: the solution of $\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + \sigma(X_t) \circ \mathrm{d}W_t$ is equivalent to the solution of

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + \tfrac{1}{2}\sigma(X_t)\cdot\nabla\sigma(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t. \tag{21}$$

We see that in the mean, the Stratonovich noise results in an effective drift. This additional term is known as the Itô correction.

Let us return to the quantum case. Define

$$a_t = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} a_\omega \mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega \tag{22}$$

where we have extended a_ω to negative frequencies. In the vacuum state, the two quadratures $x_t = a_t + a_t^\dagger$ and $y_t = \mathrm{i}a_t - \mathrm{i}a_t^\dagger$ have zero mean and delta-correlated covariance; for example, $\mathbb{E}x_t = 0$ and $\mathbb{E}[x_tx_s] = \mathbb{E}[a_ta_s^\dagger] = \delta(t-s)$. Moreover, it is easily verified that $[x_t, x_s] = 0$ for $t \neq s$ (and similarly for y_t); this is important, as it means that we can interpret x_t as a classical random process. Indeed, following the procedure of section 2.2, we can simultaneously diagonalize the operators x_t at all times and transform to a classical probability space. We find that both x_t and y_t are entirely identical to classical white noise. We will thus call the field a_t *quantum white noise*.

Note that the noise $E^{(+)}(0,t)$ that drives equation (14) is not white. However, if the system response has a sufficiently narrow bandwidth we would expect the noise to 'look' white on the slow timescale of the system, as $\kappa(\omega)$ is locally flat. Equation (22), and the associated introduction of negative frequencies, should be seen purely as a mathematical

[3] The Itô integral is only defined for *nonanticipative* integrands, i.e., f_t must be independent from any increment $W_{t_2} - W_{t_1}$ with $t_2 > t_1 \geqslant t$. It follows immediately from (18) that the integral has vanishing expectation.

construction that corresponds to noncommutative white noise. In the next section we will make these ideas more precise by showing in what sense the physical model (14) can be approximated using noises of this form.

We now proceed as in the classical case. Define the quantum Wiener process

$$A_t = \int_0^t a_t \, \mathrm{d}t. \tag{23}$$

We can now introduce the quantum Itô integral [10, 26, 37, 55]

$$\int_{t_0}^{t_n} X_s \, \mathrm{d}A_s = \lim_{|t_{i+1}-t_i|\to 0} \sum_{k=0}^{n-1} X_{t_k}(A_{t_{k+1}} - A_{t_k}) \tag{24}$$

for nonanticipative X_t (i.e., X_t is independent of any increment $A_u - A_v$, $u > v \geqslant t$.) It immediately follows that the integral has vanishing expectation in the vacuum state. Moreover, as X_{t_k} is independent from $A_{t_{k+1}} - A_{t_k}$, the process and increment commute: hence $X_t \, \mathrm{d}A_t = \mathrm{d}A_t \, X_t$. The quantum Itô rules are $\mathrm{d}A_t \, \mathrm{d}A_t^\dagger = \mathrm{d}t$, $\mathrm{d}A_t \, \mathrm{d}t = \mathrm{d}A_t^\dagger \, \mathrm{d}A_t = \mathrm{d}A_t^2 = 0$.

Similarly, we can define a quantum Stratonovich integral [26, 31, 32]

$$\int_{t_0}^{t_n} X_s \circ \mathrm{d}A_s = \lim_{|t_{i+1}-t_i|\to 0} \sum_{k=0}^{n-1} \tfrac{1}{2}(X_{t_{k+1}} + X_{t_k})(A_{t_{k+1}} - A_{t_k}) \tag{25}$$

which obeys the Leibnitz rule but does not have vanishing expectation. Additionally, in this case X_t does not commute with the noise increment, so $X_t \circ \mathrm{d}A_t$ and $\mathrm{d}A_t \circ X_t$ are two distinct forms of the Stratonovich integral.

Note that the above discussion is entirely heuristic; the mathematical objects we are using are extremely singular and require careful definition. The quantum Itô theory was introduced in a rigorous way by Hudson and Parthasarathy [37]; a more heuristic treatment can be found in Gardiner and Collett [26]. More recently the relations between the quantum Itô, Stratonovich and white noise formalisms were investigated by Gough [31, 32]. We refer to these references for a detailed treatment.

3.2.2. The Wong–Zakai theorem and the Markov limit. Mathematically, (quantum) stochastic differential equations are rather peculiar objects—strictly speaking they are not even differential equations, but integral equations. Nonetheless, SDEs are widely used to model physical phenomena. The reason that this is so successful stems from an important result, originally due to Wong and Zakai [82], which can be stated as follows. Suppose we have an ordinary differential equation of the form

$$\frac{\mathrm{d}x^\lambda(t)}{\mathrm{d}t} = f(x^\lambda(t)) + \sigma(x^\lambda(t))\xi^\lambda(t) \tag{26}$$

where $\xi^\lambda(t)$ is some piecewise smooth random process that converges to white noise in some appropriate sense as $\lambda \to 0$. Then the solution $x^\lambda(t)$ of equation (26) converges as $\lambda \to 0$ to the solution of

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + \sigma(X_t) \circ \mathrm{d}W_t. \tag{27}$$

This result tells us that the behaviour of a 'real' physical system is well approximated by the solution of an SDE as long as the noise is sufficiently wideband. Additionally our notion that physical systems are well described by Stratonovich equations is now rigorously justified. In the remainder of this section we will give a simple introduction to the quantum analogue of the Wong–Zakai procedure. For a rigorous treatment, we refer to [1, 33].

As a first step we partition our system into fast and slow timescales. The electromagnetic noise and the high-frequency oscillation of the cavity mode operate on the fast timescale, whereas the driving field and the coupling to the atoms operate on a much slower timescale. In order to study the Wong–Zakai limit we completely ignore the slow interactions by turning them off—a very good approximation if the correlation time of the noise is short. This is equivalent to the assumptions mentioned in the previous section: a short correlation time implies that $\kappa(\omega)$ is slowly varying, whereas ignoring the slow interactions assumes that these do not significantly shift the resonance frequency of the cavity.

What remains is the fast dynamics, which we write in propagator form

$$\frac{\mathrm{d}U_t}{\mathrm{d}t} = -\frac{\mathrm{i}}{\hbar} H_{\mathrm{CF}} U_t = (b(t) + b(t)^\dagger)(E^{(-)}(0,t) - E^{(+)}(0,t))U_t. \tag{28}$$

The key physical assumption we must make to obtain the white noise limit is that the cavity is weakly coupled to the external field. Naively one would expect that we could implement this limit by solving the equation $\mathrm{d}U_t/\mathrm{d}t = -\mathrm{i}\lambda H_{\mathrm{CF}} U_t/\hbar$ and then taking the limit $\lambda \to 0$. This clearly does not work, however, as this would just turn off the interaction between the cavity and the field. The problem is that λ not only changes the coupling strength, but also the timescale of the interaction dynamics.

The effect that we are trying to capture in the weak coupling limit is not a precise description of fast dynamics, but the effective contribution of the noise to the slow dynamics. We saw in the classical case, equation (21), that the noise causes an effective drift in the system dynamics. If we replace $\sigma(x) \mapsto \lambda\sigma(x)$, we infer from (21) that this drift occurs on a timescale t/λ^2. This suggests what we can make the substitution $H_{\mathrm{CF}} \mapsto \lambda H_{\mathrm{CF}}$ and let $\lambda \to 0$, but we will only obtain the weak coupling limit if we simultaneously rescale time as $t \mapsto t/\lambda^2$. This idea was originally suggested in the context of classical stochastic approximations by Stratonovich [67] and was introduced independently in the physics literature by Van Hove [72].

After performing these rescalings we obtain

$$\begin{aligned}\frac{\mathrm{d}U_t^\lambda}{\mathrm{d}t} &= \frac{1}{\lambda}\left[b\left(\frac{t}{\lambda^2}\right) + b^\dagger\left(\frac{t}{\lambda^2}\right)\right]\left[E^{(-)}\left(0,\frac{t}{\lambda^2}\right)\right. \\ &\quad \left. - E^{(+)}\left(0,\frac{t}{\lambda^2}\right)\right]U_t^\lambda \\ &= [a_\lambda^\dagger(t)b + \tilde{a}_\lambda^\dagger(t)b^\dagger - a_\lambda(t)b^\dagger - \tilde{a}_\lambda(t)b]U_t^\lambda.\end{aligned} \tag{29}$$

Let us investigate the behaviour of the rescaled noise

$$a_\lambda(t) = \frac{1}{\lambda}\int_0^\infty \kappa(\omega)\, a_\omega \mathrm{e}^{-\mathrm{i}(\omega-\omega_0)t/\lambda^2}\,\mathrm{d}\omega \tag{30}$$

as $\lambda \to 0$. In particular, we obtain for the correlation function

$$\begin{aligned}&\mathbb{E}[a_\lambda(t)a_\lambda^\dagger(s)] \\ &= \frac{1}{\lambda^2}\int_0^\infty \kappa(\omega)^2 \mathrm{e}^{-\mathrm{i}(\omega-\omega_0)(t-s)/\lambda^2}\,\mathrm{d}\omega \xrightarrow{\lambda\to 0} \gamma'\delta(t-s)\end{aligned} \tag{31}$$

with $\gamma' = 2\pi\kappa(\omega_0)^2$, where we have used $\lim_{\lambda\to 0} \mathrm{e}^{-\mathrm{i}\omega t/\lambda^2}/\lambda^2 = 2\pi\delta(\omega)\delta(t)$ (in the sense of Schwartz distributions). Hence in the weak coupling limit the resonant terms converge to white noise driving terms. However, for the rescaled noise

$$\tilde{a}_\lambda(t) = \frac{1}{\lambda}\int_0^\infty \kappa(\omega) a_\omega \mathrm{e}^{-\mathrm{i}(\omega+\omega_0)t/\lambda^2}\,\mathrm{d}\omega \tag{32}$$

we obtain

$$\mathbb{E}[\tilde{a}_\lambda(t)\tilde{a}_\lambda^\dagger(s)] = \frac{1}{\lambda^2}\int_0^\infty \kappa(\omega)^2 \mathrm{e}^{-\mathrm{i}(\omega+\omega_0)(t-s)/\lambda^2}\,\mathrm{d}\omega \xrightarrow{\lambda\to 0} 0. \tag{33}$$

Hence the nonresonant terms vanish in the weak coupling limit. We see that the weak coupling limit gives us the commonly used rotating wave approximation for free.

Studying the convergence of U_t^λ is more complicated, but can be performed by investigating the convergence of each term in the associated Dyson series [1, 33]. The result is, however, not surprising: equation (29) converges to the Stratonovich equation [31, 32]

$$\mathrm{d}U_t = \sqrt{\gamma'}\big[\mathrm{d}A_t^\dagger \circ bU_t - \mathrm{d}A_t \circ b^\dagger U_t\big] \tag{34}$$

which is essentially the quantum version of the Wong–Zakai theorem [33]. We can equivalently express the result in the Itô form as

$$\mathrm{d}U_t = \big[\sqrt{\gamma'}b\,\mathrm{d}A_t^\dagger - \sqrt{\gamma'}b^\dagger\,\mathrm{d}A_t - \tfrac{1}{2}\gamma' b^\dagger b\,\mathrm{d}t\big]U_t \tag{35}$$

where an Itô correction term emerges as in the classical case.

In addition to the emerging quantum stochastic equation, a detailed treatment of the quantum Wong–Zakai limit usually results in an additional small energy shift to the system Hamiltonian [1, 33]. This energy shift can be normalized away by a proper choice of the system Hamiltonian.

3.3. Example: spins with dispersive coupling

Now that we have made a Markovian approximation to the interaction, it remains to add the slow dynamics back in. We do this simply by adding the corresponding Hamiltonians. As discussed before, spontaneous emission is modelled by coupling the atoms directly to an unobserved field $\tilde{E}$ through their dipole moment $\mathrm{d}(t) = \sigma\mathrm{e}^{-\mathrm{i}\omega_\mathrm{d}t} + \sigma^\dagger\mathrm{e}^{\mathrm{i}\omega_\mathrm{d}t}$ (here σ is an atomic decay operator and ω_d is the dipole rotation frequency). Through a similar analysis as the one performed above, we obtain our complete physical model:

$$\begin{aligned}\mathrm{d}U_t = \Big[&\sqrt{\gamma'}b\,\mathrm{d}A_t^\dagger - \sqrt{\gamma'}b^\dagger\,\mathrm{d}A_t + \sqrt{\gamma_\perp}\sigma\,\mathrm{d}\tilde{A}_t^\dagger - \sqrt{\gamma_\perp}\sigma^\dagger\mathrm{d}\tilde{A}_t\\ &- \tfrac{1}{2}\gamma' b^\dagger b\,\mathrm{d}t - \tfrac{1}{2}\gamma_\perp\sigma^\dagger\sigma\,\mathrm{d}t\\ &- \frac{\mathrm{i}}{\hbar}(H_\mathrm{A}+H_\mathrm{AC}+H_\mathrm{D})\,\mathrm{d}t\Big]U_t.\end{aligned} \tag{36}$$

Before we specialize to the particular model that will be used in the remainder of the paper, let us digress for a moment and calculate the Heisenberg evolution $X_t = U_t^\dagger X U_t$ of an arbitrary observable X of the atom or cavity mode, as expressed symbolically in equation (14). Using the quantum Itô rules we easily obtain

$$\begin{aligned}\mathrm{d}X_t = &\frac{\mathrm{i}}{\hbar}[H_\mathrm{A}+H_\mathrm{AC}+H_\mathrm{D}, X_t]\,\mathrm{d}t\\ &+ \gamma'\mathcal{L}_{b_t}X_t\,\mathrm{d}t + \gamma_\perp\mathcal{L}_{\sigma_t}X_t\,\mathrm{d}t + \sqrt{\gamma'}[b_t^\dagger, X_t]\,\mathrm{d}A_t\\ &+ \sqrt{\gamma'}[X_t, b_t]\,\mathrm{d}A_t^\dagger + \sqrt{\gamma_\perp}[\sigma_t^\dagger, X_t]\,\mathrm{d}\tilde{A}_t\\ &+ \sqrt{\gamma_\perp}[X_t, \sigma_t]\,\mathrm{d}\tilde{A}_t^\dagger\end{aligned} \tag{37}$$

where $\mathcal{L}_c X = c^\dagger X c - \frac{1}{2}(c^\dagger c X + X c^\dagger c)$ is the well-known Lindblad term. As the expectations of Itô integrals vanish, clearly averaging away the noise terms ('tracing over the bath') results in a Lindblad-type master equation in the Heisenberg picture, which is ubiquitous in the description of quantum open systems (see for example [26]). In the language of quantum probability, the unitary solution U_t of the quantum Itô equation provides a *unitary dilation* of the associated Lindblad equation [36].

We now introduce a highly simplified model of an atomic ensemble interacting with an electromagnetic field [69, 70]. Consider an atomic ensemble consisting of a set of N atoms with a degenerate two-level ground state. We will assume that all atomic transitions are far detuned from the cavity resonance, so the interaction between the atoms and the cavity is well described by the dispersive Hamiltonian $H_\mathrm{AC} = \hbar\chi F_z b^\dagger b$ where F_z is the collective dipole moment of the ensemble, i.e., it is a spin-$N/2$ angular momentum operator, and χ determines the coupling strength. Such a Hamiltonian can be obtained, for example, by considering the full dipole coupling and then adiabatically eliminating all the excited states. We furthermore consider the atomic Hamiltonian $H_\mathrm{A} = \hbar\Delta F_z + \hbar h(t)F_y$, where Δ is the atomic detuning and $h(t)$ is the strength of a magnetic field in the y-direction. The latter will allow us to apply feedback to the system by varying the external magnetic field. We obtain

$$\begin{aligned}\mathrm{d}U_t = \Big[&\sqrt{\gamma'}b\,\mathrm{d}A_t^\dagger - \sqrt{\gamma'}b^\dagger\,\mathrm{d}A_t + \sqrt{\gamma_\perp}\sigma\,\mathrm{d}\tilde{A}_t^\dagger - \sqrt{\gamma_\perp}\sigma^\dagger\,\mathrm{d}\tilde{A}_t\\ &- \tfrac{1}{2}\gamma_\perp\sigma^\dagger\sigma\,\mathrm{d}t - \tfrac{1}{2}\gamma' b^\dagger b\,\mathrm{d}t - \mathrm{i}(\Delta F_z + h(t)F_y + \chi F_z b^\dagger b\\ &+ \mathcal{E}(b+b^\dagger))\,\mathrm{d}t\Big]U_t.\end{aligned} \tag{38}$$

Adiabatically eliminating the cavity [19, 24, 74], assuming that γ' and $\mathcal{E}$ are sufficiently large so this is a good approximation, yields

$$\begin{aligned}\mathrm{d}U_t = \Big[&\sqrt{\gamma_\perp}\sigma\,\mathrm{d}\tilde{A}_t^\dagger - \sqrt{\gamma_\perp}\sigma^\dagger\,\mathrm{d}\tilde{A}_t + \sqrt{M}F_z(\mathrm{d}A_t^\dagger - \mathrm{d}A_t)\\ &- \tfrac{1}{2}\gamma_\perp\sigma^\dagger\sigma\,\mathrm{d}t - \tfrac{1}{2}MF_z^2\,\mathrm{d}t\\ &- \mathrm{i}\left(\frac{4\chi\mathcal{E}^2}{(\gamma')^2} + \Delta\right)F_z\,\mathrm{d}t - \mathrm{i}h(t)F_y\,\mathrm{d}t\Big]U_t\end{aligned} \tag{39}$$

where $M = 16\chi^2\mathcal{E}^2/(\gamma')^3$ is the effective interaction strength. It is convenient to choose the atomic detuning $\Delta = -4\chi\mathcal{E}^2/(\gamma')^2$, and we will henceforth assume that this is the case (experimentally we can always fix the detuning by applying a magnetic field in the z-direction.)

Finally, we will for simplicity neglect spontaneous emission by setting $\gamma_\perp = 0$, a good approximation if $\gamma_\perp \ll M$ (in this case the interesting system dynamics takes place long before spontaneous emission sets in). This gives

$$\mathrm{d}U_t = \big[\sqrt{M}F_z(\mathrm{d}A_t^\dagger - \mathrm{d}A_t) - \tfrac{1}{2}MF_z^2\,\mathrm{d}t - \mathrm{i}h(t)F_y\,\mathrm{d}t\big]U_t. \tag{40}$$

We will use this highly simplified model as an example throughout the remainder of the article.

4. Conditioning: classical probability and quantum filtering

In the previous section we considered in detail the physical interactions between an atomic ensemble and the electromagnetic field, which, after many simplifications, were condensed into equation (40). This expression contains all the physical dynamics of our model. We now start the second step in our program, in which we perform statistical inference of the atomic dynamics based on an observation of the field. Our approach [71] is inspired by [5, 7].

4.1. Optical detection

Before we can derive a filtering equation we must specify what measurement is performed. We will consider the case of (balanced) homodyne detection, which measures a quadrature of the outgoing field. The principles of this method are discussed in many textbooks [62, 73] and a continuous time description in terms of quantum stochastic calculus can be found in [3]. Homodyne detection has the advantage that it gives rise to a continuous, Wiener process-type integrated photocurrent, which is particularly convenient for continuous time feedback control.

Other types of detection may be convenient in different situations depending on the experimental setup. For example, the spin squeezing experiment [30] makes use of polarimetry, which can be modelled in a very similar way as homodyne detection. Though photon counting detection also has a continuous time description in terms of quantum stochastic calculus, it gives rise to a discrete jump process which is much less convenient for the purpose of feedback control.

Heuristically, consider equation (40) as being driven by the white noise a_t, the 'derivative' of A_t. An ideal wide-band homodyne detector will measure the field observable $a_t + a_t^\dagger$ after the field has interacted with the ensemble; i.e., we observe the photocurrent $I(t) = U_t^\dagger(a_t + a_t^\dagger)U_t$. As usual, mathematically rigorous results are much more easily obtained in integrated form; hence we define as our observation the *integrated photocurrent*

$$Y_t = U_t^\dagger(A_t + A_t^\dagger)U_t \tag{41}$$

where $I(t)$ can be considered the 'derivative' of Y_t. For a rigorous treatment directly from the quantum stochastic description we refer to [3].

Finding an explicit expression for Y_t is a straightforward exercise in the use of the quantum Itô rules. From equations (40) and (41) we directly obtain

$$\mathrm{d}Y_t = 2\sqrt{M}U_t^\dagger F_z U_t\,\mathrm{d}t + \mathrm{d}A_t + \mathrm{d}A_t^\dagger. \tag{42}$$

Thus clearly homodyne detection of the field provides a measurement of the system observable $F_z(t)$ corrupted by the incident field noise.

We will extend our observation model a little further. We assumed in the above analysis that the detection efficiency is perfect. In practice there will always be some technical noise added to the signal, either due to the intrinsic loss mechanisms in the photodetectors or due to noise in the detection electronics (for example, amplifier noise). We will model these effects by the addition of an uncorrelated white noise term $\mathrm{d}W_t'$ to the observation current; i.e.,

$$\mathrm{d}Y_t = 2\sqrt{M\eta}U_t^\dagger F_z U_t\,\mathrm{d}t + \sqrt{\eta}(\mathrm{d}A_t + \mathrm{d}A_t^\dagger) + \sqrt{1-\eta}\,\mathrm{d}W_t' \tag{43}$$

where $\eta \in (0, 1]$ determines the relative strength of the technical noise ($\eta = 1$ is perfect detection). We can interpret the white noise $\mathrm{d}W_t'$ as an operator process by embedding it in a quantum probability space, for example, $\mathrm{d}W_t' = \mathrm{d}B_t + \mathrm{d}B_t^\dagger$ for some uncorrelated field B_t that does not interact with the system. Note that we have rescaled the current Y_t so that the total corrupting noise has unit variance, i.e., $\mathrm{d}Y_t^2 = \mathrm{d}t$; this gives a convenient normalization of the photocurrent. Experimentally the observed current will have some arbitrary amplification.

In order to make sense as an observed current Y_t must be a classical stochastic process, i.e., $[Y_t, Y_s] = 0\ \forall s \neq t$; clearly any sample path recorded in the laboratory is classical. From equation (43), however, it is not at all obvious that this is the case. Once again we resort to a heuristic argument which can be made rigorous in a detailed treatment of quantum stochastic calculus. Equation (40) implies that the observable a_t only interacts with the system at time t. As a_t is independent from a_s when $t \neq s$, it follows that $U_t^\dagger a_s U_t = U_s^\dagger a_s U_s \forall t \geqslant s$. But then $[I(t), I(s)] = U_t^\dagger[a_t + a_t^\dagger, a_s + a_s^\dagger]U_t = 0$, as we have already established that $a_t + a_t^\dagger$ is entirely classical white noise. Hence Y_t, the integral of $I(t)$ plus technical noise, is also a classical stochastic process.

There is another property of the observation, called the *nondemolition property* by Belavkin [6], that is essential in what follows. Let X be some observable of the atomic ensemble. Then it is easy to show, in exactly the same way we showed that Y_t is a classical process, that $[U_t^\dagger X U_t, Y_s] = 0\ \forall s \leqslant t$; i.e., any system observable at time t commutes with all prior observations. This means, as we saw in section 2.2, that finding the best estimate of a system observable given all prior observations is an entirely *classical* statistical inference problem. We will find the explicit solution to this problem, the quantum filtering equation, in the next section.

4.2. The quantum filter

Let us begin by establishing some notation. If X is an atomic ensemble observable, denote by $j_t(X) = U_t^\dagger X U_t$ its Heisenberg evolution at time t. Using equation (40) and the quantum Itô rules we easily obtain

$$\mathrm{d}j_t(X) = j_t(\mathcal{L}[X])\,\mathrm{d}t + \sqrt{M}\,j_t([X, F_z])(\mathrm{d}A_t^\dagger - \mathrm{d}A_t) \tag{44}$$

where $\mathcal{L}[X] = \mathrm{i}h(t)[F_y, X] + MF_zXF_z - \frac{1}{2}M(F_z^2X + XF_z^2)$. We have already established the observation equation

$$\mathrm{d}Y_t = 2\sqrt{M\eta}\,j_t(F_z)\,\mathrm{d}t + \sqrt{\eta}(\mathrm{d}A_t + \mathrm{d}A_t^\dagger) + \sqrt{1-\eta}\,\mathrm{d}W_t'. \tag{45}$$

Together, equations (44) and (45) form the system–observation pair of our model. Equation (44) describes the time evolution of any system observable, whereas equation (45) describes the observed current. The goal of the filtering problem is to find an expression for $\pi_t(X) = \mathbb{E}[j_t(X)|Y_{s\leqslant t}]$, the (least mean square) best estimate of the observable X given the prior observations $Y_{s\leqslant t}$. An essential point is that the conditional expectations

$\pi_t(X)$ are guaranteed to be well defined by the nondemolition property.

Due to the nondemolition property we could in principle simultaneously diagonalize $j_t(X)$ and Y_s, $s \leqslant t$ for every X, drop down to the associated classical probability space, and calculate the classical conditional expectation $\pi_t(X)$. This is not a very practical course of action, however, so we will need a shortcut. Moreover, our description of the conditional expectation in section 2.1 was rather limited: we only defined the conditional expectation with respect to one discrete random variable, whereas $Y_{s\leqslant t}$ is a continuous family of continuous random variables. To manipulate such continuous quantities one needs the mathematical machinery of real analysis.

We take the following approach. From the definition of conditional expectation is section 2.1 we can extract the following properties.

(i) $\mathbb{E}[X|Y]$ is a function on Y.
(ii) For any random variable Z that is a function of Y, we must have $\mathbb{E}[\mathbb{E}[X|Y]Z] = \mathbb{E}[XZ]$.

It is easy to see that the definition of section 2.1 implies these properties, and it is not hard to show that the converse is also true. In the continuous case we just take these properties as the definition of conditional expectation. This is precisely the real analytic definition, where the intuitive idea of being 'a function of Y' is replaced by the notion of measurability [75].

We are now ready to take our shortcut. By property (1), $\pi_t(X)$ must be a function of $Y_{s\leqslant t}$. Introduce the ansatz

$$\mathrm{d}\pi_t(X) = C_t(X)\,\mathrm{d}t + D_t(X)\,\mathrm{d}Y_t \tag{46}$$

where $C_t(X)$, $D_t(X)$ are functions of $Y_{s\leqslant t}$ to be determined. If we can determine C_t and D_t, the filtering problem has been solved.

To implement property (2) we use the following trick. We require that

$$\mathbb{E}\big[\pi_t(X)\mathrm{e}^{\int_0^t g(s)\,\mathrm{d}Y_s}\big] = \mathbb{E}\big[j_t(X)\mathrm{e}^{\int_0^t g(s)\,\mathrm{d}Y_s}\big] \tag{47}$$

for any function $g(t)$. The idea behind this is the same as that of a moment generating function: we can generate any (analytic) function of $Y_{s\leqslant t}$ by using an appropriate $g(t)$ and taking derivatives. Hence, if we have proved the relation (47) then we have essentially satisfied property (2).

What remains is mostly a direct application of the Itô rules. For convenience we multiply both sides of (47) by $\exp(-\frac{1}{2}\int_0^t g(s)^2\,\mathrm{d}s)$. Define

$$e_t^g = \mathrm{e}^{\int_0^t g(s)\,\mathrm{d}Y_s - \frac{1}{2}\int_0^t g(s)^2\,\mathrm{d}s} \qquad \mathrm{d}e_t^g = g(t)e_t^g\,\mathrm{d}Y_t. \tag{48}$$

It is now straightforward to evaluate

$$\frac{\mathrm{d}\mathbb{E}[e_t^g\pi_t(X)]}{\mathrm{d}t} = \mathbb{E}[e_t^g(C_t(X) + 2\sqrt{M\eta}\,j_t(F_z)D_t(X)) + g(t)e_t^g(D_t(X) + 2\sqrt{M\eta}\,j_t(F_z)\pi_t(X))] \tag{49}$$

$$\frac{\mathrm{d}\mathbb{E}[e_t^g\,j_t(X)]}{\mathrm{d}t} = \mathbb{E}[e_t^g\,j_t(\mathcal{L}[X]) + \sqrt{M\eta}g(t)e_t^g\,j_t(F_zX + XF_z)]. \tag{50}$$

We now invoke equation (47) and attempt to find $C_t(X)$, $D_t(X)$ by comparing (49) and (50) term by term. We run into a snag, however, as a naive comparison would yield C_t and D_t in terms of $j_t(F_z)$, etc, which are not functions of $Y_{s\leqslant t}$. Fortunately we can use property (2) of conditional expectations to change all the j_t terms in (49) and (50) to the corresponding π_t terms ($\mathbb{E}[j_t(\cdot)] = \mathbb{E}[\pi_t(\cdot)]$, etc). This gives immediately

$$\mathrm{d}\pi_t(X) = \pi_t(\mathcal{L}[X])\,\mathrm{d}t + \sqrt{M\eta}(\pi_t(F_zX + XF_z) - 2\pi_t(F_z)\pi_t(X))(\mathrm{d}Y_t - 2\sqrt{M\eta}\,\pi_t(F_z)\,\mathrm{d}t) \tag{51}$$

which is the quantum filtering equation for our model.

It is instructive to recall the example of section 2.2. There a simple filtering scenario was constructed by coupling a spin to a commuting observable, then conditioning the spin observables on the commuting observable. This gave rise to a set of classical random variables, representing the conditioned spin observables. Similarly, we have coupled an atomic ensemble to an optical mode and conditioned the atomic observables on a homodyne measurement in the field. This gave rise to a classical Itô equation (51) for the conditioned atomic observables, driven by the observations Y_t.

As in section 2.2, we will find it useful to represent the filter in its adjoint (density) form. To this end, we define the conditional atomic density matrix ρ_t as the random matrix that satisfies $\pi_t(X) = \mathrm{Tr}[\rho_t X]$. Equation (51) gives

$$\mathrm{d}\rho_t = -\mathrm{i}h(t)[F_y, \rho_t]\,\mathrm{d}t + M\mathcal{D}[F_z]\rho_t\,\mathrm{d}t + \sqrt{M\eta}\mathcal{H}[F_z]\rho_t\,\mathrm{d}W_t \tag{52}$$

where we have used the notation

$$\mathcal{D}[c]\rho \equiv c\rho c^\dagger - (c^\dagger c\rho + \rho c^\dagger c)/2 \tag{53}$$

$$\mathcal{H}[c]\rho \equiv c\rho + \rho c^\dagger - \mathrm{Tr}[(c + c^\dagger)\rho]\rho \tag{54}$$

and we have defined the *innovations process*

$$\mathrm{d}W_t = \mathrm{d}Y_t - 2\sqrt{M\eta}\,\mathrm{Tr}[\rho_t F_z]\,\mathrm{d}t. \tag{55}$$

An important result in filtering theory is that the innovations process W_t is in fact a Wiener process [6, 12]. Though we have not introduced sufficient technical machinery to prove this fact, we can give a simple interpretation. We can write W_t in the form

$$\mathrm{d}W_t = 2\sqrt{M\eta}(j_t(F_z) - \pi_t(F_z))\,\mathrm{d}t + \sqrt{\eta}(\mathrm{d}A_t + \mathrm{d}A_t^\dagger) + \sqrt{1-\eta}\,\mathrm{d}W_t'. \tag{56}$$

This expression consists of two parts: the last two terms are white noise terms, whereas the first term is the difference between an atomic observable and our best estimate of that observable, i.e., it represents the new information (the 'innovation') contained in the measurement.

4.3. Conditional spin dynamics

Before we add control to the picture it is interesting to take a look at the open-loop properties of the filtering equation (52), i.e., without feedback, by setting $h(t) = 0$. The equation propagates a density matrix, defined as the adjoint of a set of classical conditional expectations, which carries the interpretation of the 'statistically inferred' density matrix of the ensemble given the observations in the probe field. One might wonder how such a picture is related to the traditional picture of quantum measurements.

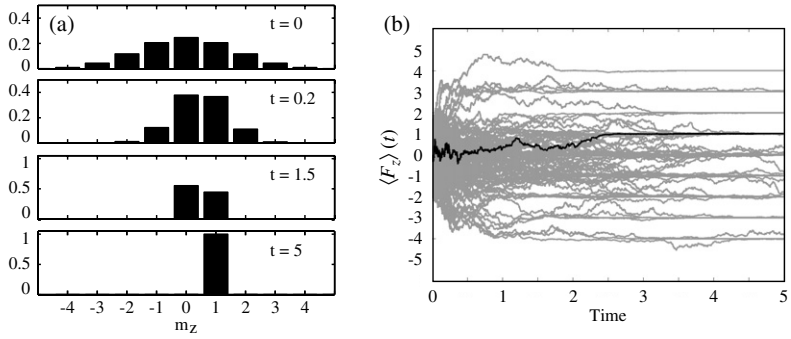

Figure 2. (a) Time evolution of the conditional state from a coherent spin state at $t = 0$ to an eigenstate at long times. The graph shows the population of each F_z eigenstate. (b) 100 sample paths of $\pi_t(F_z)$, with $M = \eta = 1$. The dark line is the sample path shown in (a), resulting in $m_z = 1$. Von Neumann projection is clearly visible at long times.

To illustrate the filtering process we have simulated equation (52) for a spin $F = 5$ ensemble (for example, 10 two-level atoms) [66]. Such simulations are highly simplified by the fact that the innovations process is a Wiener process. This means that we do not have to simulate the full quantum-mechanical model, equations (44) and (45), to obtain a photocurrent Y_t to drive (52). Instead, we just plug in a Wiener process for the innovations, for which straightforward numerical methods are available. The results are shown in figure 2.

At long times the conditional state is clearly driven to one of the eigenstates of F_z, i.e., Dicke states [17], just as predicted by the Von Neumann projection postulate. In fact, it can be rigorously proved that the $t \to \infty$ limit of equation (52) is *exactly* identical to the projection postulate, i.e., the probability of collapse onto each eigenstate is predicted correctly by the filtering equation [2, 66, 71].

Note that we have not previously mentioned the projection postulate in this article. As this result follows from our theory we do not need to postulate it: instead, we have 'derived' it using quantum dynamics and classical statistics[4]. In some sense the filtering process exposes the anatomy of a quantum measurement. We have explicitly modelled the coupling between the probe field and the system under measurement, equation (44), and we considered separately a step that involved purely the gain of information. Both processes conspire to bring about the traditional projection of the system state in the long time limit.

At intermediate times, $t < \infty$, the conditional state gradually collapses onto the F_z eigenstates. This process, for a single sample path, is shown in figure 2(a). Whereas a Von Neumann measurement would take the state discontinuously from the initial state to the final collapsed state, the filtering process continuously narrows the distribution over the eigenstates until only one remains. Aside from giving a more realistic description of continuous optical measurements, this description creates an opportunity that has no analogue with projective measurements: we can interfere with the collapse process while it is occurring by applying real-time feedback.

Finally, we should remark that not all filtering equations give rise to Von Neumann-type collapse. For example, homodyne detection of spontaneously emitted photons, or an atomic ensemble resonantly interacting with the probe field, will result in continuous decay of the conditional state into the ground state. Projective dynamics is obtained in our case because of the dispersive (off-resonant) interaction of the ensemble with the probe and the neglect of spontaneous emission. The latter can be justified, however, if there is a large separation of time scales between the time of collapse and the time at which the spontaneous emission sets in. In this case, the intermediate regime will be very similar to the long time limit of our model.

The range of dynamics emerging from filtering equations highlights the need for the separate modelling of the system–probe interation. Though we have only presented a very simple model, we have outlined a bottom-up approach in which the system–probe interaction is modelled from first principles using quantum electrodynamics. The detailed modelling of realistic experimental configurations will be invaluable for quantitative comparison of theoretical predictions and experimental data [28].

5. Feedback control and quantum state preparation

The intrinsic randomness of quantum measurement should not dissuade the capable observer from trying to control the dynamics of a system. In fact, it should do just the opposite. The inherent uncertainty in observation is the inspiration for the use of feedback control, and promotes it to the status of fundamental.

Although the physical constraints imposed by quantum mechanics are performance limiting, quantum feedback control problems are well defined and worth pursuing for all of the same reasons engineers use control on classical systems. Furthermore, quantum feedback control, while technically difficult, is simply a branch of traditional control and is amenable to the techniques developed therein [4, 18, 71].

[4] The reader should not get the impression, however, that we have now reduced all the peculiarities of quantum measurement to pure classical probability. In particular, we cannot derive why the measurement of an observable rules out the measurement of noncommuting observables, which has no counterpart in classical probability. Only the conditioning, which takes place after a measurement has been performed and the measurement result has been obtained, can be given a purely classical interpretation in this way as a statistical inference procedure. On the other hand, the 'back action' on the system is caused by the quantum dynamics of the interaction between the system and the probe, which we have explicitly modelled by a quantum stochastic differential equation.

Far from introducing an entirely new kind of problem, the challenges presented here highlight and motivate the extension of mathematical methods already in development elsewhere.

In this section we begin by discussing the types of problems and structure encountered in a typical quantum feedback control scenario, building upon the formalism developed above. Here we use language from classical control theory, and discuss the possible application of optimal and robust control theories to the quantum setting. We also emphasize experimental constraints which motivate simplifications of desired controls through model reduction. Next we demonstrate the utility of feedback in a review of applications to atomic ensemble experiments. We finish by focusing on the particular theoretical example of deterministically preparing a state with continuous measurement and control.

5.1. Defining feedback control

The term 'quantum feedback control' as used in this article refers to a particular class of problems that should be distinguished from other types of control with quantum systems. The class we consider involves the measurement of a quantum system by interaction with a quantum field. The field is destructively measured, resulting in a classical measurement record. That measurement record is then processed and fed back to Hamiltonian parameters affecting the same system.

The rest of this article is concerned with problems of this kind. However, it should be noted that there exist further types of control with this arrangement that we will not discuss. Aside from actuating Hamiltonian parameters of the system with feedback, the observer may possess the ability to adaptively change the measurement itself according to the measurement record. This leads to different 'unravellings' of the dynamics [79]. By the nature of the measurement, the ensemble average behaviour of the system will be the same for any chosen unravelling or adaptive measurement scheme. Of course, the same will not be true for the average trajectory behaviour of the system under different Hamiltonian control laws.

Additionally, there exist completely different types of control with quantum systems bearing little resemblance to the measurement techniques discussed here. For instance, one can imagine doing a type of feedback experiment where, instead of destructively measuring the ancilla system, it is returned to interact with the system of interest again, and possibly repeatedly. For the case of the usual optical ancilla system, this has been referred to as 'all-optical feedback' to distinguish it from the electrical measurement signal alternatively produced [81]. In certain cases this kind of 'coherent control' [52] can achieve state preparation goals with minimal processing overhead and delay. In the formalism presented here, one could describe such a process completely at the quantum stochastic level of section 3.

Finally, the term 'quantum control' is also used in the literature to refer to yet another scenario, with not one system, but an ensemble of identically prepared systems. Here a system is driven with a pulse, then the result is measured. Subsequently, another system is prepared, another pulse is used to drive it, the result is again measured, and so on. In between trials, the pulse shape is changed based on the previous measurements in some algorithmic way to optimize the effect of the pulse [61]. This procedure is a type of 'learning control' and, unlike in the examples we study, no feedback occurs during the lifetime of an individual system.

5.2. Separation structure

Generally speaking, the control problem consists of finding a mapping of the measurement record onto the actuation variables such that some predefined task is achieved. When stated in this way the problem is very difficult to solve; after all, when we allow any functional from the photocurrent history to the control variables, it is hard to know where to start.

Fortunately we can simplify the problem description considerably, using what is sometimes referred to as the *separation principle* or the *information state approach*, originally introduced in classical control theory by Mortensen [57]. The basic idea behind this approach is that we can never control the system more precisely than the precision with which the system state can be inferred from the observations. In many ways this is a statement of the obvious: for example, if we know that the system is controlled to within some bound, then clearly we can infer that the system state is within that bound. As a consequence, the best we can do is to control the best estimate of the system state, i.e., the conditional state.

The advantage of this approach is that we have converted the *output feedback* control problem into a *state feedback* control problem for the filter. Operationally, we then consider the filtering equation (52) as our new 'effective' dynamical equation to be controlled, where the feedback $h(t)$ can now be taken to be a function of the conditional state ρ_t as opposed to the measurement record. This is a less constrained problem than the output feedback problem and is hence often easier to solve. Control design is further simplified by the fact that the innovation, equation (55), is white. This means we can consider equation (52) as an ordinary Itô equation to be controlled, without separately modelling the statistics of the photocurrent driving noise.

The structure of the entire control setup, in the context of the model discussed in the previous sections, is shown in figure 3. The atomic ensemble and its interaction with the optical probe field and the magnetic control field was modelled in section 3. Homodyne detection was the subject of section 4.1. The photocurrent is processed by a digital control circuit which produces the feedback signal. Inside the controller, the 'whitened' photocurrent drives the quantum filter, as described in section 4.2. The control law is a function of the best estimate of the system state. To design the control law, however, we only need to consider the 'internal' feedback loop inside the digital circuit. From the controller's effective perspective, the only role of the physical experiment is to provide the innovation $\mathrm{d}W_t$, which is white by construction.

5.3. Defining an objective

We have separated the control design into an estimation problem, which was the subject of section 4.2, and a control problem. The control problem is undefined, however, until we state a goal that our controller should achieve.

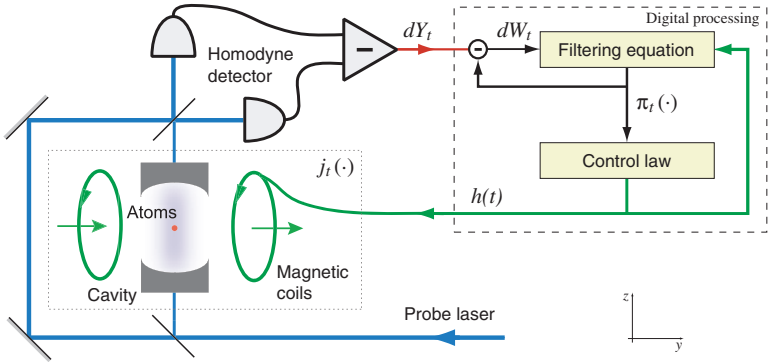

Figure 3. Schematic of the entire feedback control problem for an atomic ensemble. On the experimental level an ensemble interacts with a probe field, as described by equation (44). Homodyne detection gives rise to the photocurrent (45), which is processed by a digital controller. A magnetic field is used for feedback. On the controller level, the photocurrent drives the quantum filter (51) which updates recursively the best estimate of the atomic state. The control law is a functional of the current conditional state. The innovations structure (55) allows the control design to be based directly on the filtering dynamics.

As an example, an experimentalist may want to minimize some functional of the system and control variables, for example

$$C[h(t)] = \mathbb{E}\int_0^T (j_t(F_z^2) + \mu h(t)^2)\,\mathrm{d}t \tag{57}$$

where μ is a parameter that limits the degree that the control input is applied. To apply the separation principle to this case we must first convert the cost function C into a form which is only a function of the filter state. This is straightforward, however, due to the property $\mathbb{E}\,j_t(\cdot) = \mathbb{E}\pi_t(\cdot)$ of conditional expectations: we obtain

$$C[h(t)] = \mathbb{E}\int_0^T (\pi_t(F_z^2) + \mu h(t)^2)\,\mathrm{d}t. \tag{58}$$

As expected, the control goal depends only on the conditional state, i.e., the filter state is a *sufficient statistic* for this control problem.

In principle, the minimization of (58) using the dynamics of the filtering equation would produce a control law which is the optimal time-dependent mapping of the conditional state onto the control parameters. This type of problem is known as 'optimal control', and is one of the primary modes of thought in classical control theory [8, 20, 38]. As is apparent from equation (52), the general form of the filtering equation is nonlinear in the state and, as an unfortunate result, the optimal control solution is extremely difficult to find. Although both nonlinear and stochastic control theories are well developed fields classically, there is still much work to be done in their intersection.

Fortunately, there are alternative methods for gaining ground on the quantum feedback control problem. First, in some instances, it is possible to linearize the dynamics of the filtering equation via moment expansions. In this case, one can readily adopt 'LQG' techniques from classical control [8, 38] for linear systems (L), a cost function quadratic in linear observables and control variables as above (Q), and Gaussian dynamics (G), to solve the problem completely [4, 18, 19, 65]. In any given example, the needed linearization may only work for particular initial states and limited periods of time, but the LQG results can still be remarkably far reaching.

Second, we can choose to be less demanding of our controller, and instead formulate a non-optimal goal. For instance, suppose we are interested in preparing the quantum state ρ_c at long times. The control goal can then be formulated as *find a control law* $h(t)$ *so that* $\mathbb{E}[j_t(X)] \to \mathrm{Tr}[X\rho_c]$ *as* $t \to \infty$ *for any system observable* X. As above, it is easy to see that the filter state is a sufficient statistic, and hence we can directly apply the separation principle. In particular, if we can find a controller that makes ρ_c a global (stochastically) stable state for the filter dynamics, the eventual preparation of ρ_c is ensured. Although the state might not be prepared as quickly as is physically possible, it is an accomplishment to know that it will eventually be prepared with unit probability. Here there is much work to be done on constructively generating controllers and methods for proving the stability, but progress has been made for some simple problems [71].

5.4. Robustness and model reduction

If given the choice between a controller that works optimally under one set of ideal circumstances and a controller that works sub-optimally, but adequately, over a wide set of possible conditions, the wise experimentalist would always choose the latter. Due to unexpected modelling uncertainties and exogenous noise sources, the optimal control approach has the potential to fail catastrophically in realistic environments, a possibility that has motivated the development of 'robust control' for many years [84]. One could say the reason experiments are performed at all is to test the robustness of our model and control design.

The concept of robust control has been extensively studied in the classical deterministic setting, but the same logic holds true for quantum applications. Even as quantum technology reaches its limits, there will always be some degree of non-intrinsic system uncertainty to which the system should be robust. Not surprisingly, quantum feedback techniques have been shown to enable robustness to model uncertainty in metrology applications [65]. Of course one need not draw the line too sharply between optimal and robust control, as there exist types of risk-sensitive optimal quantum control that inherently consider certain kinds of robustness [39, 40].

Another practical reason why optimal control may not be ultimately relevant is that real-time information processing takes time. Even if the modelling is perfect and there are no excess noise sources, an optimal controller may not work due to processing constraints. If the processing delay of the actual controller is large compared to the relevant timescale of the filtering dynamics, then another approach will be needed. Despite the improving performance of programmable logic devices that might best implement the optimal control, there are few experiments with slow enough timescales that modern electronics can be optimally effective at real-time estimation and control [64].

Clearly it is of significant interest to be able to derive a controller that works without having to evolve the full filtering equation in real time. Recognizing this, physicists have proposed and used controllers for quantum feedback applications that use a simplified control law which bypasses the full state estimation. Even more easily, one can sometimes feed the measurement record directly back to the system with a gain tailored intelligently in time [69, 77, 80]. However, with any of these approaches, one must be cognizant of realistic gain and bandwidth constraints. For example, one cannot realistically feed pure white noise back into a system, as this would imply infinite sensor and detection bandwidths. Although most of the simplified quantum controllers suggested in the literature have been constructed through more or less heuristic means, we expect the continuing development of these techniques to resort to more mathematical notions of model reduction, where the degree of approximation and its effect on the feedback performance can be more explicitly quantified.

5.5. Measurement and feedback in atomic ensembles

When considering systems with the potential for interesting applications related to quantum information processing, there exists a natural tendency within many physicists to consider conceptually simplified systems, for example, a single atom or ion. While much progress has been made in trapping, measuring, and controlling single particles, it has also been realized for some time that the use of atomic ensembles does not preclude the observation of uniquely quantum effects nor a simple description. As compared to alternative systems, ensembles are experimentally convenient and, by the sheer number of participants, sufficient signal can be generated to make them powerful in quantum applications, with atomic clocks being just one prominent example.

Here we consider those experiments where continuous measurement and feedback have been used to generate entanglement either within or between atomic ensembles. We begin by discussing the use of dispersive measurement to produce a spin-squeezed state in a single ensemble in the short time limit, and how feedback can be used to make this process deterministic. We then focus on a particular theoretical limit where the linear approximation fails, but still highly entangled eigenstates of the measured F_z can be prepared by using the more complete filtering equation and an intuitive feedback law. Finally, we briefly discuss experiments and proposals involving the creation of entanglement between two ensembles with and without feedback.

5.5.1. One ensemble. For state preparation with an atomic ensemble, spin-squeezed state (SSS) [44] is a natural target states. This collective spin state is internally entangled, simply characterized by measured moments of the spin-operators, and useful in metrology tasks [29, 65, 76]. For an ensemble with N spin-f particles and collective angular momentum operators F_i, a state is defined as spin-squeezed, and entangled, if[5]

$$\frac{2fN\langle\Delta F_z^2\rangle}{\langle F_x\rangle^2} < 1 \tag{59}$$

where the spin-state is pointing along x so that $\langle F_x\rangle = F = Nf$ and $\langle F_y\rangle = \langle F_z\rangle = 0$ [63]. Methods to produce these states typically begin with an unentangled coherent spin state (CSS) with all spins exactly polarized along the x-direction and realizing the equality of the uncertainty relation

$$\langle\Delta F_y^2\rangle\langle\Delta F_z^2\rangle \geqslant \frac{\hbar^2\langle F_x\rangle^2}{4}. \tag{60}$$

For a SSS, the equality is roughly maintained with one component $\langle\Delta F_z^2\rangle$ squeezed smaller than the CSS value and the other $\langle\Delta F_y^2\rangle$ anti-squeezed.

There are many ways one can imagine producing the spin-correlations within the ensemble needed for the collective state to be squeezed. Examples include using direct Hamiltonian interactions [63] and also transferring correlations from an auxiliary system, for example, squeezed states of light [34, 49, 56]. We shall focus on the production of spin-squeezed states via dispersive measurement, the effects of which were originally discussed and demonstrated in [47, 48]. Subsequently, Thomsen, *et al* [69] proposed a feedback procedure, discussed below, that used a measurement-based field rotation to remove the randomness of the measurement while retaining the desired squeezing effect. Others have proposed using feedback to an optical pumping beam to achieve a similar result [56]. It has since been experimentally demonstrated that using a procedure similar to [69] feedback can enable the deterministic production of spin-squeezed states in cold atomic samples [28, 30]. Much work continues in this direction, in particular towards creating squeezed states with the Cesium clock transition, which would considerably improve current atomic clock performance [59].

To understand the conditional preparation of spin-squeezed states by dispersive measurement, consider the apparatus in figure 3. As shown above, the filtering equation is given by equation (52). This equation is only applicable at long times $t \gg 1/M$ if a sufficiently strong cavity is used to suppress the spontaneous emission to an insignificant level. Given existing experimental technology this is currently unrealistic; nevertheless, we consider the long time dynamics for purposes of demonstration.

The filtering equation was derived using a simplified one-dimensional model of the interaction. Although this model is often an adequate description of free-space experiments where a distribution of atoms interacts with a spatially extended probe beam, there is much interest in making the model more accurate by extending it to three dimensions. A complete

[5] We will denote by $\langle\cdot\rangle$ the expectation of an observable in a general sense. The associated state can be prepared either unconditionally or conditionally.

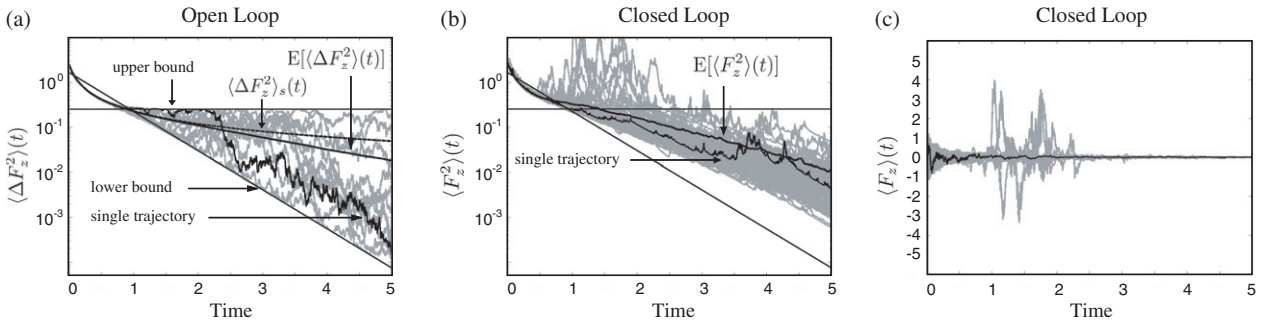

Figure 4. (a) $\langle\Delta F_z^2\rangle(t) = \pi_t(\Delta F_z^2)$ in open loop $h(t) = 0$. $\langle\Delta F_z^2\rangle_s$ is the approximate variance of equation (62). (b) $\langle F_z^2\rangle(t) = \pi_t(F_z^2)$ and (c) $\langle F_z\rangle(t) = \pi_t(F_z)$ with the control law (63) and $\lambda = 10$. Note that $\mathbb{E}\pi_t(F_z^2)$, the cost for preparation of the $m_z = 0$ eigenstate, decreases monotonically. All plots show 100 sample paths and $M = \eta = 1$.

model would consider the scattering process where all free-space field modes interact with the atomic distribution. Some of those channels would then be measured, and the results used to condition the atomic state. In this picture, the conditional entanglement results from the indistinguishability of the atoms in the measurement and 'spontaneous emission' is a term used to describe the effect of the remaining unobserved channels. A considerable amount of work remains to be done in describing collective scattering in terms of measurement theory, but much progress has been made [11, 21, 46, 58].

Returning to the one-dimensional model, we can extract the conditional evolution equations for the moments of any operator from the filtering equation. Under the approximation that there are many atoms and the initial collective state is nearly polarized along the x-direction, we can derive the closed set of equations

$$\mathrm{d}\pi_t(F_z) \approx F\exp[-Mt/2]h(t)\,\mathrm{d}t + 2\sqrt{M\eta}\pi_t(\Delta F_z^2)\,\mathrm{d}W_t \tag{61}$$

$$\mathrm{d}\pi_t(\Delta F_z^2) \approx -4M\eta\pi_t(\Delta F_z^2)^2\,\mathrm{d}t. \tag{62}$$

These equations are obtained by truncating the exact coupled expressions for $\pi_t(F_z^n)$, calculated from equation (51), at $n = 2$ [29, 65]. This reduced description is equivalent to a classical Kalman filter [8, 38] and corresponds to a local linearization of the spin dynamics.

Equations (61) and (62) are valid only in the short time limit $t \ll 1/M$, past which the full filtering equation is needed. At longer times terms neglected in the approximation grow to the point that the variance becomes stochastic [65], and the moment truncation is no longer a good description. This process can be seen in figure 4(a), where at small times the variance is deterministic, but then becomes random at longer times.

The deterministically shrinking variance of equation (62) at short times signifies that a spin-squeezed state is prepared with a random offset given by equation (61). The idea of [69] was to choose $h(t) \propto \dot{Y}_t$ with an intelligently chosen gain such that the first term effectively cancels the second term in equation (61), preparing the same SSS on every trial. Although this exact procedure cannot be implemented in practice due to the infinite detector and actuator bandwidths implied by the control law, it was essentially a similar, but filtered, current feedback law used in the experiment [30]. Because of the linearity of the dynamics in the short time limit, the simple current-based feedback law does not perform significantly worse than a law that changes $h(t)$ more optimally according to the state $\pi_t(F_z)$ [65].

Given these dynamics, another control strategy would be to separate the measurement and control in time: simply measuring for a finite amount of time, turning off the probe, and using the measurement result to rotate the spin-squeezed state to the desired location. However, as pointed out in [69], the continuous feedback approach is more robust than this procedure to, for example, uncertainty in the total atom number which is necessary to compute the size of the correcting rotation.

To further demonstrate the utility of continuous measurement and feedback, we now consider the long time behaviour of the filtering equation, past the point in time $t > 1/M$ when the linearized description fails. As discussed in section 4.3, the filtering equation stochastically prepares a random eigenstate of F_z asymptotically in time[6]. In [66] we investigated numerically the performance of particular controllers at producing one F_z eigenstate deterministically on every trial.

Here it is critical to point out that, unlike with the Gaussian spin-squeezed states, a post-measurement rotation strategy will not work in this regime. If the wrong eigenstate is randomly prepared in one measurement, it cannot be transformed into the correct eigenstate by a rotation alone. Furthermore, despite the adequacy of the direct current feedback law at short times, such a controller is less useful at longer times. As opposed to state-based control, this type of control will feed noise into the system even if the target state is reached, unless the gain is turned to zero. Although certain gain-tailored schemes can be made to optimize the feedback at small times [69], it is not at all obvious how such a procedure could be generalized to the long time case.

In contrast, if the control variable $h(t)$ is made a function of the conditional state, then it will naturally know when the goal has been achieved and no longer disturb the state unnecessarily. Numerically we were able to demonstrate [65] that with an initial x-polarized state, the control law

$$h(t) = -\lambda\pi_t(F_z) \tag{63}$$

[6] There are other schemes that produce superpositions of F_z eigenstates conditionally but without control, based on single photon detection of an ensemble in a cavity [22].

appeared to deterministically prepare the highly entangled state $m_z = 0$ on every trial, as seen in figure 4. Thus, continuous feedback, in addition to being robust, is also capable of preparing states on every trial that would be impossible to generate deterministically with measurement and control pulses separated in time.

Numerical evidence is encouraging, but more analytic statements about the performance of particular control laws are still desirable. Unfortunately, the more atoms the ensemble contains, the larger the Hilbert space becomes, and the more difficult it is to analytically prove that certain states are global attractors under particular feedback laws. However, as we have shown in [71], there exist methods adapted from nonlinear and stochastic control theory that can prove the global stability of F_z eigenstates for this problem. Although this has only been demonstrated for few atom systems, there is hope that the techniques can be extended to consider dynamics on larger Hilbert spaces. Much of the control design process remains guesswork, but ultimately we desire methodology that allows us to systematically construct both controllers and proofs that validate those controllers.

5.5.2. Two ensembles. The creation of a collective entanglement within a single atomic ensemble can be motivated with, for example, the need for noise reduction in metrology tasks, where the system is used as a relatively localized probe of some parameter of interest. In other practical applications, like quantum communication, it is desirable to have an entangled quantum state, but with constituents separated substantially in space [23]. Indeed it has been experimentally demonstrated that by detecting a single probe beam after it passes through two spatially separate atomic ensembles, the two ensembles can be made conditionally entangled [41].

Just as single-mode spin squeezing can be quantified with the collective variables for the one ensemble, here the 'two-mode' squeezing can be quantified with the joint collective operators describing both ensembles. Furthermore, in analogy to the work of [69], the random offset observed in the measurement process can in principle be eliminated with a suitable feedback law to deterministically produce the same two-mode spin-squeezed state on every trial [9].

In a related context, it has recently been experimentally demonstrated that the two-ensemble system may serve as an effective quantum memory for states of light [42]. This procedure differs from the deterministic state preparation discussed previously in that the state of light to be mapped onto the ensembles is not known beforehand. However, the procedure described in [42] is similar in that it does use feedback to rotate the Gaussian ensemble state in a way that maps one measured quadrature of the optical state onto the atoms, while the other unmeasured quadrature is mapped unconditionally by the interaction alone. Clearly, this process shares many of the same properties as the applications discussed previously and can similarly benefit from analyses with technical notions of robustness and optimality. Finally, this procedure becomes even more efficient if the input atomic state is a two-mode squeezed state, which highlights yet another practical application of deterministic entangled quantum state preparation.

6. Conclusion

In this article we have attempted to give a unified picture of a quantum feedback control setup. Starting from elementary physical interactions, as described by a field-theoretic model, we first performed statistical inference on this model, and then used this framework to develop feedback control strategies for state preparation in atomic ensembles. The latter is directly related to recent experimental work which we briefly summarized. It is our hope that such a unified picture will help linking the basic physics and experimental reality to a high-level, control-theoretic point of view.

Many open problems remain on both ends of the spectrum. On the physics side much work remains to be done on the realistic modelling of laboratory experiments. Ultimately a full three-dimensional field-theoretic model will be invaluable for quantitative comparison of theory and experiments. On the control-theoretic side many of the techniques that have been used are still heuristic in nature. Systematic, constructive design methods for nonlinear stochastic controllers, the incorporation of realistic robustness criteria, and efficient model reduction techniques with controllable approximation errors are some of the major outstanding issues. We believe that a fruitful interaction between the physics and mathematical control theory communities will open the road to significant advances in these directions.

Acknowledgments

The authors would like to thank Luc Bouten, J M Geremia and Howard Wiseman for many fruitful discussions. This work was supported by ARO and the Caltech MURI Center for Quantum Networks. JKS acknowledges a Hertz fellowship.

References

[1] Accardi L, Frigerio A and Lu Y G 1990 Weak coupling limit as a quantum functional central limit theorem *Commun. Math. Phys.* **131** 537–70

[2] Adler S L, Brody D C, Brun T A and Hughston L P 2001 Martingale models for quantum state reduction *J. Phys. A: Math. Gen.* **34** 8795–820

[3] Barchielli A 2003 Continual measurements in quantum mechanics *Lecture notes of the Summer School on Quantum Open Systems* (Grenoble: Institut Fourier)

[4] Belavkin V P 1988 Nondemolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes *Proc. Bellman Continuum, Sophia-Antipolis 1988* (*Springer Lecture Notes in Control and Information Sciences* vol 121) (Berlin: Springer) pp 245–65

[5] Belavkin V P 1992 Quantum continual measurements and a posteriori collapse on CCR *Commun. Math. Phys.* **146** 611–35

[6] Belavkin V P 1992 Quantum stochastic calculus and quantum nonlinear filtering *J. Multivariate Anal.* **42** 171–201

[7] Belavkin V P 1994 Quantum diffusion, measurement and filtering I *Theory Probab. Appl.* **38** 573–85

[8] Bensoussan A 1992 *Stochastic Control of Partially Observable Systems* (Cambridge: Cambridge University Press)

[9] Berry D W and Sanders B C 2002 Near-optimal two-mode spin squeezing via feedback *Phys. Rev.* A **66** 012313

[10] Biane P 1995 Calcul stochastique non-commutatif *Lectures on Probability Theory* (*Springer Lecture Notes in Mathematics* vol 1608) ed P Bernard (Berlin: Springer)

[11] Bouchoule I and Mølmer K 2002 Preparation of spin-squeezed atomic states by optical-phase-shift measurement *Phys. Rev.* A **66** 043811
[12] Bouten L, Guţă M and Maassen H 2004 Stochastic Schrödinger equations *J. Phys. A: Math. Gen.* **37** 3189–209
[13] Cohen-Tannoudji C, Dupont-Roc J and Grynberg G 1989 *Photons and Atoms: Introduction to Quantum Electrodynamics* (New York: Wiley)
[14] Coutin L and Decreusefond L 2000 Abstract non-linear filtering theory in presence of fractional Brownian motion *Ann. Appl. Prob.* **9** 1058–90
[15] Dalton B J, Barnett S M and Knight P L 1999 Quasi mode theory of macroscopic canonical quantization in quantum optics and cavity quantum electrodynamics *J. Mod. Opt.* **46** 1315–41
[16] Davis M H A and Marcus S I 1981 An introduction to nonlinear filtering *Stochastic Systems: The Mathematics of Filtering and Identification and Applications* ed M Hazewinkel and J C Willems (Dordrecht: Reidel) pp 53–75
[17] Dicke R H 1954 Coherence in spontaneous radiation processes *Phys. Rev.* **93** 99
[18] Doherty A C, Habib S, Jacobs K, Mabuchi H and Tan S M 2000 Quantum feedback control and classical control theory *Phys. Rev.* A **62** 012105
[19] Doherty A C and Jacobs K 1999 Feedback control of quantum systems using continuous state estimation *Phys. Rev.* A **60** 2700
[20] Doyle J, Francis B and Tannenbaum A 1990 *Feedback Control Theory* (London: Macmillan)
[21] Duan L M, Cirac J I and Zoller P 2002 Three-dimensional theory for interaction between atomic ensembles and free-space light *Phys. Rev.* A **66** 023818
[22] Duan L M and Kimble H J 2003 Efficient engineering of multiatom entanglement through single-photon detections *Phys. Rev. Lett.* **90** 253601
[23] Duan L M, Lukin M D, Cirac J I and Zoller P 2001 Long-distance quantum communication with atomic ensembles and linear optics *Nature* **414** 413–8
[24] Dunningham J A, Wiseman H M and Walls D F 1997 Manipulating the motion of a single atom in a standing wave via feedback *Phys. Rev.* A **55** 1398
[25] Gambetta J and Wiseman H M 2002 Non-Markovian stochastic Schrödinger equations: generalization to real-valued noise using quantum-measurement theory *Phys. Rev.* A **66** 012108
[26] Gardiner C W and Collett M J 1985 Input and output in damped quantum systems: quantum stochastic differential equations and the master equation *Phys. Rev.* A **31** 3761–74
[27] Gardiner C W and Zoller P 2004 *Quantum Noise* 3rd edn (Berlin: Springer)
[28] Geremia J, Stockton J K and Mabuchi H 2005 Continuous quantum measurement and conditional spin-squeezing in alkali atoms *Preprint* quant-ph/0501033
[29] Geremia J M, Stockton J K, Doherty A C and Mabuchi H 2003 Quantum Kalman filtering and the Heisenberg limit in atomic magnetometry *Phys. Rev. Lett.* **91** 250801
[30] Geremia J M, Stockton J K and Mabuchi H 2004 Real-time quantum feedback control of atomic spin-squeezing *Science* **304** 270
[31] Gough J 1999 Asymptotic stochastic transformations for nonlinear quantum dynamical systems *Rep. Math. Phys.* **44** 313–38
[32] Gough J 1999 The Stratonovich interpretation of quantum stochastic approximations *Potential Anal.* **11** 213–33
[33] Gough J 2005 Quantum flows as Markovian limit of emission, absorption and scattering interactions *Commun. Math. Phys.* **254** 489–512
[34] Hald J, Sørensen J L, Schori C and Polzik E S 1999 Spin squeezed atoms: a macroscopic entangled ensemble created by light *Phys. Rev. Lett.* **83** 1319–22
[35] Holden H, Øksendal B, Ubøe J and Zhang T 1996 *Stochastic Partial Differential Equations: A Modeling, White Noise Functional Approach* (Basle: Birkhäuser)
[36] Holevo A S 2001 *Statistical Structure of Quantum Theory* (Berlin: Springer)
[37] Hudson R L and Parthasarathy K R 1984 Quantum Itô's formula and stochastic evolutions *Commun. Math. Phys.* **93** 301–23
[38] Jacobs O L R 1996 *Introduction to Control Theory* 2nd edn (New York: Oxford University Press)
[39] James M R 2004 Risk-sensitive optimal control of quantum systems *Phys. Rev.* A **69** 032108
[40] James M R 2005 A quantum Langevin formulation of risk-sensitive optimal control *J. Opt. B: Quantum Semiclass. Opt.* **7** S198
[41] Julsgaard B, Kozhekin A and Polzik E S 2001 Experimental long-lived entanglement of two macroscopic objects *Nature* **413** 400
[42] Julsgaard B, Sherson J, Cirac J I, Fiurasek J and Polzik E S 2004 Experimental demonstration of quantum memory for light *Nature* **432** 482
[43] Kallianpur G and Karandikar R L 1985 White noise calculus and nonlinear filtering theory *Ann. Prob.* **13** 1033–107
[44] Kitagawa M and Ueda M 1993 Squeezed spin states *Phys. Rev.* A **47** 5138
[45] Kuo H H 1996 *White Noise Distribution Theory* (Boca Raton, FL: CRC Press)
[46] Kupriyanov D, Mishina O, Sokolov I, Julsgaard B and Polzik E S 2005 Multimode entanglement of light and atomic ensembles via off-resonant coherent forward scattering *Phys. Rev.* A **71** 032348
[47] Kuzmich A, Mandel L and Bigelow N P 2000 Generation of spin squeezing via continuous quantum nondemolition measurement *Phys. Rev. Lett.* **85** 1594–7
[48] Kuzmich A, Mandel L, Janis J, Young Y E, Ejnisman R and Bigelow N P 1999 Quantum nondemolition measurements of collective atomic spin *Phys. Rev.* A **60** 2346
[49] Kuzmich A, Mølmer K and Polzik E 1997 Spin squeezing in an ensemble of atoms illuminated with squeezed light *Phys. Rev. Lett.* **79** 4782
[50] Leibfried D, Barrett M D, Schaetz T, Britton J, Chiaverini J, Itano W M, Jost J D, Langer C and Wineland D J 2004 Toward Heisenberg-limited spectroscopy with multiparticle entangled states *Science* **304** 1476
[51] Liptser R S and Shiryaev A N 2001 *Statistics of Random Processes I: General Theory* (Berlin: Springer)
[52] Lloyd S 2000 Coherent quantum feedback *Phys. Rev.* A **62** 022108
[53] Maassen H 2003 Quantum probability applied to the damped harmonic oscillator *Quantum Probability Communications XII* ed S Attal and J M Lindsay (Singapore: World Scientific) pp 23–58
[54] Mandel L and Wolf E 1997 *Optical Coherence and Quantum Optics* (Cambridge: Cambridge University Press)
[55] Meyer P A 1995 *Quantum Probability for Probabilists* (*Springer Lecture Notes in Mathematics* vol 1538) (Berlin: Springer)
[56] Mølmer K 1999 Twin-correlations in atoms *Eur. Phys. J.* D **5** 301–5
[57] Mortensen R E 1966 Stochastic optimal control with noisy observations *Int. J. Control* **4** 455–64
[58] Mueller J H, Petrov P, Oblak D, Alzar C L G, de Echaniz S R and Polzik E S 2005 Diffraction effects on light-atomic ensemble quantum interface *Phys. Rev.* A **71** 033803
[59] Oblak D, Mikkelsen J K, Tittel W, Vershovski A K, Sorensen J L, Petrov P G, Alzar C L G and Polzik E S 2005 Quantum noise limited interferometric measurement of atomic noise: towards spin squeezing on the Cs clock transition *Phys. Rev.* A **71** 043807

[60] Øksendal B 1998 *Stochastic Differential Equations* 5th edn (Berlin: Springer)
[61] Rabitz H, de Vivie-Riedle R, Motzkus M and Kompa K 2000 Whither the future of controlling quantum phenomena? *Science* **288** 824
[62] Scully M O and Zubairy M S 1997 *Quantum Optics* (Cambridge: Cambridge University Press)
[63] Sorensen A, Duan L M, Cirac J I and Zoller P 2001 Many-particle entanglement with Bose–Einstein condensates *Nature* **409** 63
[64] Stockton J, Armen M and Mabuchi H 2002 Programmable logic devices in experimental quantum optics *J. Opt. Soc. Am.* B **19** 3019
[65] Stockton J K, Geremia J, Doherty A C and Mabuchi H 2004 Robust quantum parameter estimation: coherent magnetometry with feedback *Phys. Rev.* A **69** 032109
[66] Stockton J K, Van Handel R and Mabuchi H 2004 Deterministic Dicke-state preparation with continuous measurement and control *Phys. Rev.* A **70** 022106
[67] Stratonovich R L 1963 *Topics in the Theory of Random Noise* (London: Gordon and Breach)
[68] Streater R F 1995 *Statistical Dynamics* (London: Imperial College Press)
[69] Thomsen L K, Mancini S and Wiseman H M 2002 Continuous quantum nondemolition feedback and unconditional atomic spin squeezing *J. Phys. B: At. Mol. Opt. Phys.* **35** 4937
[70] Thomsen L K, Mancini S and Wiseman H M 2002 Spin squeezing via quantum feedback *Phys. Rev.* A **65** 061801(R)
[71] Van Handel R, Stockton J K and Mabuchi H 2005 Feedback control of quantum state reduction *IEEE Trans. Autom. Control* **50** 768–80
[72] Van Hove L 1955 Quantum mechanical perturbations giving rise to a statistical transport equation *Physica* **21** 617–40
[73] Walls D F and Milburn G J 1994 *Quantum Optics* (Berlin: Springer)
[74] Warszawski P and Wiseman H M 2000 Adiabatic elimination in compound quantum systems with feedback *Phys. Rev.* A **63** 013803
[75] Williams D 1991 *Probability with Martingales* (Cambridge: Cambridge University Press)
[76] Wineland D J, Bollinger J J, Itano W M and Heinzen D J 1994 Squeezed atomic states and projection noise in spectroscopy *Phys. Rev.* A **50** 67–88
[77] Wiseman H M 1994 Quantum theory of continuous feedback *Phys. Rev.* A **49** 2133
[78] Wiseman H M 1994 Quantum trajectories and feedback *PhD Thesis* University of Queensland
[79] Wiseman H M and Doherty A C 2005 Optimal unravellings for feedback control in linear quantum systems *Phys. Rev. Lett.* **94** 070405
[80] Wiseman H M, Mancini S and Wang J 2002 Bayesian feedback versus Markovian feedback in a two-level atom *Phys. Rev.* A **66** 013807
[81] Wiseman H M and Milburn G J 1994 All-optical versus electro-optical quantum-limited feedback *Phys. Rev.* A **49** 4110
[82] Wong E and Zakai M 1965 On the relationship between ordinary and stochastic differential equations *Int. J. Eng. Sci.* **3** 213–29
[83] Yuen H P and Shapiro J H 1978 Optical communication with two-photon coherent states. I—Quantum-state propagation and quantum-noise reduction *IEEE Trans. Inf. Theory* **24** 657–68
[84] Zhou K and Doyle J C 1997 *Essentials of Robust Control* 1st edn (Englewood Cliffs, NJ: Prentice-Hall)