PHYSICAL REVIEW A 89, 033850 (2014)

Ś

Three-dimensional light-matter interface for collective spin squeezing in atomic ensembles

Ben Q. Baragiola,^{1,*} Leigh M. Norris,¹ Enrique Montaño,² Pascal G. Mickelson,² Poul S. Jessen,² and Ivan H. Deutsch¹

¹Center for Quantum Information and Control, University of New Mexico, Albuquerque, New Mexico 87131, USA

²Center for Quantum Information and Control, University of Arizona, Tucson, Arizona 85721, USA

(Received 19 December 2013; published 26 March 2014; corrected 28 March 2014)

We study the three-dimensional nature of the quantum interface between an ensemble of cold, trapped atomic spins and a paraxial laser beam, coupled through a dispersive interaction. To achieve strong entanglement between the collective atomic spin and the photons, one must match the spatial mode of the collective radiation of the ensemble with the mode of the laser beam while minimizing the effects of decoherence due to optical pumping. For ensembles coupling to a probe field that varies over the extent of the cloud, the set of atoms that indistinguishably radiates into a desired mode of the field defines an inhomogeneous spin wave. Strong coupling of a spin wave to the probe mode is not characterized by a single parameter, the optical density, but by a collection of different effective atom numbers that characterize the coherence and decoherence of the system. To model the dynamics of the system, we develop a full stochastic master equation, including coherent collective scattering into paraxial modes, decoherence by local inhomogeneous diffuse scattering, and backaction due to continuous measurement of the light entangled with the spin waves. This formalism is used to study the squeezing of a spin wave via continuous quantum nondemolition measurement. We find that the greatest squeezing occurs in parameter regimes where spatial inhomogeneous model.

DOI: 10.1103/PhysRevA.89.033850

PACS number(s): 42.50.Lc, 03.67.Bg, 42.50.Dv

I. INTRODUCTION

Cold atomic ensembles interacting with electromagnetic fields are powerful tools in quantum information science with applications that include quantum memory [1-3], quantum communication [4,5], continuous variable quantum computing [6], and metrology [7,8]. At the heart of these protocols is the strong coupling between a quantum mode of the field and an effective collective spin of the ensemble. This coupling can generate entanglement between atoms and photons, such that measurement of the light yields strong quantum backaction on the atoms. Photons can also enable a quantum data bus for entangling atoms with one another. Enhancing the atom-light interface is thus essential for improving the performance of quantum technologies and for reaching new regimes where a quantum advantage becomes manifest. This can be achieved through confined modes such as in optical cavities [8-10] or wave guides in optical nanostructures [11-13].

Strong atom-photon coupling can also occur in free space in the interaction between light and an extended ensemble of atoms. This occurs when photons are indistinguishably scattered by the ensemble, and interference enhances the radiation into the probe mode relative to diffuse scattering into 4π sr [14,15]. Early experiments demonstrated such strong coupling and entanglement in high-pressure vapor cells where a one-dimensional description of plane-wave modes and uniform atomic density is applicable [16]. This theory accurately describes a variety of experiments including the entanglement of macroscopic ensembles in remote vapor cells [17] and quantum memory for continuous variables [2].

More recently, experiments have employed ensembles of ultracold atoms in pencil-shaped dipole traps probed by focused laser beams [18–20]. When the radiation pattern of the light scattered from the average atomic ensemble is well matched with the paraxial mode of the probe, the spatial mode of the scattered photons is effectively indistinguishable from the probe. In this case the probe mode becomes strongly entangled with a *collective* variable of the atomic ensemble. Such geometries have the potential to enhance the atom-photon quantum interface, but their description is more complex and requires a full treatment of radiation patterns, diffraction, inhomogeneous coupling, and decoherence. Harnessing the advantages of these atomic ensembles thus demands a threedimensional quantum theory of the underlying interaction, including both coherent coupling and quantum noise.

In the last decade there has been significant progress in developing a three-dimensional quantum description of the atom-light interface. Mode matching of the scattered light to the spatial mode of the probe laser, including the effects of diffraction, has been studied using a semiclassical scattering model [21]. A rigorous field-theoretic treatment separates the mean-field classical effects from the quantum fluctuations and noise, including the spatial inhomogeneities of the atomic and light modes [22]. Models that include spatial modes have been developed in a variety of contexts [23–26]. Applications include remote entanglement via collective Raman scattering in a DLCZ-type protocol [27,28] and for multimode quantum memories [29]. From such studies, it is clear that onedimensional models not only fail to describe the relevant coherent and incoherent effects, but they also do not take advantage of the resources associated with spatial modes [30,31].

In this paper we revisit the three-dimensional atom-light interface with particular emphasis on spin squeezing through quantum nondemolition (QND) measurement of the collective spin via the Faraday effect [32–34], shown schematically in Fig. 1. In this protocol, the key interaction is the off-resonant scattering of horizontally polarized photons into vertical polarization. Measurement in a balanced polarimeter corresponds to a homodyne measurement of these scattered photons. The

1050-2947/2014/89(3)/033850(22)

^{*}quinn.phys@gmail.com



FIG. 1. (Color online) Schematic of a linearly polarized laser probe with a Gaussian spatial mode (solid blue lines entering from the left of the cloud) interacting dispersively with a cold, trapped atomic ensemble. The light that is indistinguishably scattered by the average atomic density distribution defines the radiated paraxial mode (solid red arrows emanating from the cloud to the right). The interference of the radiated and probe modes leads to a rotation of the field polarization according to the Faraday effect. When measured in a polarimeter, this can be used to generate spin squeezing in the ensemble. The spatial overlap of the collectively scattered field and the probe, a measure of the strength of the atom-light coupling, depends highly on geometry. In addition, diffusely scattered photons due to density fluctuations in the ensemble lead to local decoherence in the collective spin variables.

degree of scattering into the local oscillator, defined by the paraxial laser mode, determines the measurement strength and the resulting backaction that generates squeezing.

Central to this problem are the spatial modes of the light and the collective spin waves of the atomic ensemble. In the one-dimensional model, a single collective parameter defines the strong-coupling regime of the atom-photon interface, the optical density on resonance, $OD = \eta \sigma_0 L = N \sigma_0 / A$, where η is the atomic density, σ_0 is the resonant scattering cross section, L is the length of the vapor cell, and N is the number of atoms in the volume V = AL for a uniform beam of area A. In contrast, in a fully three-dimensional model, where the atomic density, $\eta(\mathbf{r})$, and paraxial beam-intensity distribution, $\beta(\mathbf{r})$, are not uniform, there is a collection of parameters that dictates the strong-coupling regime. Different effective atom numbers, $N_{\text{eff}}^{(K)} = \int d^3 \mathbf{r} \, \eta(\mathbf{r}) \, [\beta(\mathbf{r})]^K$, govern different physical effects. For example, $N_{\text{eff}}^{(1)}$ determines the mean Faraday signal in the polarimeter, while $N_{\text{eff}}^{(2)}$ determines the size of the measurement uncertainty from spin projection noise.

The entangling strength of the atom-light interface is determined by the size of the spin projection uncertainty compared to the quantum uncertainty in the measured light quadratures (shot noise). This collective interaction is proportional to an effective optical density, $OD_{eff} = N_{eff}^{(2)} \sigma_0 / A$. In contrast, decoherence acts locally on the atoms in a noncollective manner, and the noise injected into the system due to optical pumping and spin flips is governed by other parameters. A proper accounting of the balance between the coherent coupling and decoherence is especially challenging given the tensor nature of the atom-photon interaction of real alkali-metal atoms. Previous treatments of quantum noise in a multimode Faraday-based atom-light interface have been carried out in a one-dimensional model [35,36]. Our goal is to extend this to the three-dimensional case.

In this work, we derive a stochastic master equation describing the dynamics of the collective atomic state conditioned on balanced polarimetry measurements, including the effects of measurement backaction, collective decoherence from unmeasured paraxial light, and local decoherence from diffuse photon scattering that gives rise to optical pumping. While we apply this to study conditional spin squeezing generated by a QND measurement, the formalism we develop is broadly applicable to other protocols where a strong, freespace, atom-light interface is essential and where measurement backaction may be a tool for induced atom-atom interactions.

The remainder of the article is organized as follows. We lay out the physical model for an ensemble of alkali-metal atoms dispersively interacting with a coherent probe laser in Sec. II. We begin with a semiclassical model that can be used to describe the scattered paraxial fields and to identify the collective spin wave that is coupled to the laser mode. To understand the entangling Faraday interaction in a multimode geometry, we then present a fully quantum mechanical model. This serves as the cornerstone for a complete description of QND squeezing and allows us to account for the damaging effects of decoherence. When the output light is measured continuously, the quantum dynamics, including the combined effects of measurement backaction and decoherence, are described by a stochastic master equation. We use this fully quantum mechanical atom-light description to study effects of spatial modes on the squeezing of spin waves in Sec. III. In particular, we use the multimode description to model the dynamics of spin squeezing and to analyze the dependence of peak squeezing on cloud and beam geometry. We use numerical simulations to help build physical intuition about the three-dimensional atom-light interface and to investigate how the model can be used to optimize an experimental design. Finally, we summarize our results and present future directions for this work in Sec. IV.

II. PARAXIAL ATOM-LIGHT INTERFACE

When driven by an off-resonant laser field such that the excited-state probability is small, atoms elastically scatter electromagnetic waves in a manner equivalent to a set of linearly polarizable particles. Thus, a great deal of qualitative and quantitative information can be obtained from classical radiation theory. In a rigorous field-theoretic analysis, Sørensen and Sørensen showed that the mean-field effect of the light interacting with an atomic ensemble gives rise to an index of refraction of the gas, while fluctuations are due to the random positions of the atoms and the vacuum noise of the light [22]. In particular, the index of refraction is due to the spatially averaged local density of the atoms, while the diffuse scattering into 4π arises from the random positions of the point atomic scatterers and is equivalent to decoherence by local spontaneous emission. This diffuse scattering, which leads to attenuation of the incident wave and optical pumping of the atomic state, is accounted for by an imaginary part of the polarizability according to the optical theorem.

We can thus break the problem into two pieces. First, the mean-field effect is described by classical scattering of a laser beam incident on a linearly polarizable dielectric whose shape is determined by the atomic density distribution. For a paraxial probe beam and an extended cloud, the scattered field is also paraxial, and the solution is easily found by Fraunhofer scattering theory [37]. Indeed, such a model was used in Ref. [21] to study the efficiency of forward scattering for various atomic and beam geometries. As we are interested in spin squeezing, within this semiclassical model we identify the mode-matched collective atomic spin wave arising from the Faraday effect, the scattering of an incident horizontal polarization to an orthogonal vertical polarization induced by the atoms' polarizability tensor. This is the entangling interaction of interest. Second, to properly account for quantum backaction on the atoms resulting from measurement and to describe the decoherence from diffuse scattering and optical pumping, we turn to the fully quantum theory. It is the trade-off between the effects of decoherence and the coherent coupling between the atomic ensemble and the probe that ultimately determines the strength of the atom-light interface.

A. Semiclassical theory

Consider the scattering of an incident paraxial laser beam with frequency ω_0 and complex amplitude, $\mathbf{E}_L(\mathbf{r}_{\perp},z) = \vec{\mathcal{E}}_L(\mathbf{r}_{\perp},z)e^{ik_0z}$, by a particle located at a position \mathbf{r}' with dynamical tensor electric-dipole polarizability $\vec{\alpha}$. The field envelope has the standard form, $\vec{\mathcal{E}}_L(\mathbf{r}_{\perp},z) = \vec{\epsilon}_L \, \mathcal{E}_0 u_{00}(\mathbf{r}_{\perp},z)$, where $\vec{\epsilon}_L$ is the laser polarization and $u_{00}(\mathbf{r}_{\perp},z)$ is chosen to be the Gaussian TEM₀₀ mode given by

$$u_{00}(\mathbf{r}_{\perp},z) = \frac{w_0}{w(z)} e^{-\frac{|\mathbf{r}_{\perp}|^2}{[w(z)]^2}} e^{\frac{ik_0|\mathbf{r}_{\perp}|^2}{2R(z)}} e^{-i\Phi(z)}.$$
 (1)

The *z*-dependent beam waist, the radius of curvature of the phase fronts, and the Guoy phase are given by

$$w(z) = w_0 \sqrt{1 + (z/z_R)^2},$$
 (2a)

$$R(z) = z[1 + (z_R/z)^2],$$
 (2b)

$$\Phi(z) = \tan^{-1}(z/z_R), \qquad (2c)$$

respectively, with beam waist w_0 and Rayleigh range $z_R \equiv k_0 w_0^2/2$. In the first Born approximation, the scattered field amplitude is that radiated by the induced dipole,

$$\mathbf{E}_{\text{scat}}(\mathbf{r}) = k_0^2 [\stackrel{\leftrightarrow}{\alpha} \cdot \vec{\mathcal{E}}_L(\mathbf{r}')]_\perp \frac{e^{ik_0|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}$$
$$\approx k_0^2 [\stackrel{\leftrightarrow}{\alpha} \cdot \vec{\mathcal{E}}_L(\mathbf{r}')]_\perp e^{ik_0(z-z')} \frac{e^{\frac{ik_0|\mathbf{r}_\perp - \mathbf{r}'_\perp|^2}{2(z-z')}}}{z-z'}, \qquad (3)$$

where the subscript \perp denotes the component of the dipole transverse to the direction of observation. The last approximation is valid for paraxial points of observation, $z \gg |\mathbf{r}_{\perp}|$. Gaussian-cgs units for the electromagnetic field equations are used throughout.

Because the dipole radiation is not mode matched with the Gaussian laser beam, the light is scattered into all paraxial modes as well as off-axis nonparaxial modes. In the far field, $z \gg z'$, the total field takes the form

$$\mathbf{E}_{\text{out}}(\mathbf{r}) = \mathbf{E}_{L}(\mathbf{r}) + \mathbf{E}_{\text{scat}}(\mathbf{r})$$

= $[\vec{\epsilon}_{L} + \vec{\Upsilon}]\mathcal{E}_{0}e^{ik_{0}z}u_{00}(\mathbf{r}_{\perp}, z) + \mathbf{E}'_{\text{scat}}(\mathbf{r}).$ (4)

Here we have separated $E_{scat}(\mathbf{r})$ into a portion forward-scattered into the spatial mode of the probe with amplitude

 $\vec{\Upsilon} \mathcal{E}_{0} e^{ik_0 z}$, and a portion scattered into all other spatial modes, $\mathbf{E}'_{scat}(\mathbf{r})$. As shown in Appendix A [Eq. (A10)],

$$\vec{\Upsilon} \mathcal{E}_0 e^{ik_0 z} \equiv \int \frac{d^2 \mathbf{r}_\perp}{A} u_{00}^*(\mathbf{r}_\perp, z) \mathbf{E}_{\text{scat}}(\mathbf{r})$$
$$= i \frac{2\pi k_0}{A} (\vec{\alpha} \cdot \vec{\epsilon}_L)_\perp |u_{00}(\mathbf{r}'_\perp, z')|^2 \mathcal{E}_0 e^{ik_0(z-z')}, \quad (5)$$

where $A = \int d^2 \mathbf{r}_{\perp} |u_{00}(\mathbf{r}_{\perp}, z)|^2 = \pi w_0^2 / 2$ is the effective beam area.

The key physical effects are seen in these equations. The component of $\tilde{\Upsilon}$ along the laser polarization $\vec{\epsilon}_L$ gives rise to the scalar index of refraction and attenuation. The component of $\tilde{\Upsilon}$ orthogonal to $\vec{\epsilon}_L$ gives rise to a rotation of the polarization on the Poincaré sphere: Faraday rotation and birefringence. For example, suppose the laser is linearly polarized along *x* ($\vec{\epsilon}_L = \mathbf{e}_x$). The total field thus can be written

$$\mathbf{E}_{\text{out}}(\mathbf{r}) = \mathbf{E}'_{\text{scat}}(\mathbf{r}) + \left[\left(1 + i\delta\phi - \frac{a}{2} \right) \mathbf{e}_x + \left(\frac{\chi + i\beta}{2} \right) \mathbf{e}_y \right] \\ \times \mathcal{E}_0 e^{ik_0 z} u_{00}(\mathbf{r}_{\perp}, z), \tag{6}$$

where

$$\begin{split} \delta \phi &= (2\pi k_0/A) |u_{00}(\mathbf{r}'_{\perp},z')|^2 \operatorname{Re} \{\alpha_{xx}\}, \\ a &= (4\pi k_0/A) |u_{00}(\mathbf{r}'_{\perp},z')|^2 \operatorname{Im} \{\alpha_{xx}\}, \\ \chi &= -(4\pi k_0/A) |u_{00}(\mathbf{r}'_{\perp},z')|^2 \operatorname{Im} \{\alpha_{yx}\}, \\ \beta &= (4\pi k_0/A) |u_{00}(\mathbf{r}'_{\perp},z')|^2 \operatorname{Re} \{\alpha_{yx}\}, \end{split}$$

where ϕ is the index of refraction phase shift, *a* is the Beer's law attenuation coefficient, χ is the rotation angle of the Stokes vector corresponding to the Faraday effect, and β is the corresponding angle for birefringence, with the polarizability matrix elements denoted as $\alpha_{ij} = \mathbf{e}_i \cdot \hat{\alpha} \cdot \mathbf{e}_j$ in the *x*-*y* basis.

The above description of the atom-field coupling is most easily generalized using the theory of scattering of paraxial waves [21]; details can be found in Appendix A. The mean field is described by the electric field envelope $\vec{\mathcal{E}}_L(\mathbf{r}_{\perp},z,t) = \mathcal{A}(t-z/c)\vec{\mathcal{U}}(\mathbf{r}_{\perp},z)$, where $\mathcal{A}(t)$ is the temporal pulse envelope and $\vec{\mathcal{U}}(\mathbf{r}_{\perp},z)$ is the spatial envelope satisfying the paraxial wave equation

$$\frac{\partial}{\partial z}\vec{\mathcal{U}}(\mathbf{r}_{\perp},z) = \frac{i}{2k_0}\nabla_{\perp}^2\vec{\mathcal{U}}(\mathbf{r}_{\perp},z) + i2\pi k_0 \overleftrightarrow{\chi}(\mathbf{r}_{\perp},z) \cdot \vec{\mathcal{U}}(\mathbf{r}_{\perp},z),$$
(7)

with spatially averaged dielectric susceptibility $\chi(\mathbf{r}_{\perp}, z)$. The scattering solution to this equation is well known [37]. In the first Born approximation, i.e., for dilute samples where multiple scattering is negligible, given an incident field $\tilde{\mathcal{U}}_{in}(\mathbf{r}_{\perp}, z)$, the total field is

$$\vec{\mathcal{U}}(\mathbf{r}_{\perp},z) = \vec{\mathcal{U}}_{in}(\mathbf{r}_{\perp},z) + i2\pi k_0 \int_{-\infty}^{z} dz' \int d^2 \mathbf{r}_{\perp}' K(\mathbf{r}_{\perp} - \mathbf{r}_{\perp}',z-z') \times \overleftrightarrow{\chi}(\mathbf{r}_{\perp}',z') \cdot \vec{\mathcal{U}}_{in}(\mathbf{r}_{\perp}',z'), \qquad (8)$$

where $K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z')$ is the paraxial propagator. This solution is the superposition of incident and radiated dipole fields. The solution for a paraxial field scattered from a point dipole at position \mathbf{r}' [Eq. (3)] is recovered by setting $\stackrel{\leftrightarrow}{\chi} \delta^{(3)}(\mathbf{r} - \mathbf{r}')$.

The diagonal matrix elements of the susceptibility give rise to the index of refraction and a slight distortion of the wave front of the beam. We can neglect this effect for dilute gases, though it is easily accounted for. The Faraday effect arises from the scattering of initially *x*-polarized light into orthogonal *y* polarization as discussed above, governed by the off-diagonal element of the dielectric susceptibility matrix, χ_{yx} . To measure Faraday rotation, one employs a balanced polarimeter at $\pm 45^{\circ}$, so that the signal \mathcal{M} is proportional to $\mathcal{U}_x^*\mathcal{U}_y + \mathcal{U}_y^*\mathcal{U}_x$, integrated across the detector surface at position z_D in the far field,

$$\mathcal{M} \propto \int d^2 \mathbf{r}_{\perp} \operatorname{Re} \{ \mathcal{U}_x^*(\mathbf{r}_{\perp}, z_D) \mathcal{U}_y(\mathbf{r}_{\perp}, z_D) \}.$$
(9)

The result is an effective homodyne detector for U_y , with U_x playing the role of the local oscillator. Using the solution for $U_y(\mathbf{r}_{\perp}, z_D)$ [Eq. (8)] and the properties of the propagator [Eq. (A6)],

$$\mathcal{M} \propto \int_{-\infty}^{z_D} dz' \int d^2 \mathbf{r}'_{\perp} \operatorname{Re}\{i \chi_{yx}(\mathbf{r}'_{\perp}, z')\} \mathcal{U}_x(\mathbf{r}'_{\perp}, z')$$
$$\times \left[\int d^2 \mathbf{r}_{\perp} \mathcal{U}_x^*(\mathbf{r}_{\perp}, z_D) K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z_D - z') \right]$$
$$= -\int d^3 \mathbf{r}' \operatorname{Im}\{\chi_{yx}(\mathbf{r}'_{\perp}, z')\} |\mathcal{U}_x(\mathbf{r}'_{\perp}, z')|^2.$$
(10)

The measured signal is thus proportional to the local value of the susceptibility component Im{ $\chi_{yx}(\mathbf{r}_{\perp}, z)$ } integrated over the dielectric, weighted by the local field intensity $|\mathcal{U}_x(\mathbf{r}_{\perp}, z)|^2$.

For a dilute ensemble of cold atoms at fixed positions \mathbf{r}_i , the dielectric susceptibility of the gas is

$$\hat{\vec{\chi}}(\mathbf{r}) = \sum_{i} \langle \hat{\vec{\alpha}}^{(i)} \rangle \, \delta^{(3)}(\mathbf{r} - \mathbf{r}_{i}), \tag{11}$$

where $\dot{\vec{\alpha}}^{(i)}$ is the the dynamic polarizability tensor operator for the *i*th atom. We consider here atoms restricted to a subspace defined by a total (hyperfine) angular momentum *f*. In terms of the hyperfine spin operator $\hat{\mathbf{f}}$, the polarizability operator can be decomposed into irreducible tensor components [38],

$$\hat{\alpha}_{ij} = \alpha_0 \bigg[C^{(0)} + i C^{(1)} \epsilon_{ijk} \hat{f}_k + C^{(2)} \bigg(\frac{\hat{f}_i \hat{f}_j + \hat{f}_j \hat{f}_i}{2} - \delta_{ij} \frac{\hat{\mathbf{f}}^2}{3} \bigg) \bigg],$$
(12)

where α_0 is the characteristic polarizability and $C^{(K)}$ is the coefficient of the irreducible rank-*K* tensor component. The rank-0 component is a scalar, which influences neither the atomic spin state nor the field's polarization dynamics. The vector (rank 1) component describes a fictitious magnetic field interaction that drives Larmor precession of the spin depending on the ellipticity of the field, as well as the Faraday effect that rotates the field polarization depending on the atomic magnetization. The irreducible tensor (rank 2) component generally induces nonlinear dynamics of the atomic spins [39], as well as birefringence on the polarization of the probe. For alkali-metal atoms driven on a fine-structure multiplet, α_0 and the $C^{(K)}$ coefficients are given in Ref. [38].

The effect of the tensor component complicates both the collective coupling of the atoms to the probe as well as the internal spin dynamics. In special cases, the deleterious effects of the rank-2 component of the tensor polarizability on the OND measurement can be removed via dynamical decoupling [19]. More generally, by applying a strong bias magnetic field in the direction of the probe's propagation along z, the birefringent effect on the probe arising from the coupling to the atoms via the rank-2 tensor term averages to zero [40]. The residual effect of the rank-2 component is nonlinear dynamics on the internal spin state of the atom [39], which does not effect the the QND measurement under consideration here. These nonlinear dynamics can affect the metrologically relevant spin squeezing by dephasing the collective mean spin, but this can be compensated through additional control techniques [41].

We thus retain only the vector component of the off-diagonal element of the dielectric susceptibility, χ_{yx} , which describes a pure Faraday interaction. Substituting $\operatorname{Im}\{\chi_{yx}(\mathbf{r}_{\perp},z)\} \propto \sum_{i} \langle \hat{f}_{z}^{(i)} \rangle \delta^{(3)}(\mathbf{r}-\mathbf{r}_{i})$ into Eq. (10) yields

$$\mathcal{M} \propto \sum_{i} |\mathcal{U}_{x}(\mathbf{r}_{\perp i}, z_{i})|^{2} \langle \hat{f}_{z}^{(i)} \rangle.$$
(13)

Equation (13) is the central result of the semiclassical model. In a plane-wave, homogeneous, one-dimensional description, the measured observable is $\mathcal{M} \propto \sum_i \langle \hat{f}_z^{(i)} \rangle = \langle \hat{F}_z \rangle$, the symmetric collective spin of the ensemble. For paraxial beams, the polarimeter measures an effective *spin wave* determined by the inhomogeneous weighting of the atomic spin operators by the local intensity of the beam. The spin wave is stationary because it is coupled to the *forward-scattered light*, where the absorbed and emitted modes are the same. Physically, this spin wave is the collective observable that radiates indistinguishably into the probe mode and is effectively selected by the homodyne measurement of the polarimeter.

Further intuition can be gained from the semiclassical model. We recover symmetric atom-light coupling when the field intensity $|\mathcal{U}_x(\mathbf{r}_{\perp},z)|^2$ is constant over the atomic ensemble. Geometrically, this is achieved when the beam waist, w_0 , is much larger than the transverse extent of the cloud and the length of the cloud is short compared to twice the Rayleigh range, $z_R = k_0 w_0^2/2$. The mean-field radiation pattern of such a cloud described by Eq. (8), however, has poor overlap with the probe as depicted in Fig. 2(a). The end result is that the polarimeter detects only a small fraction of the signal photons. On the other hand, perfect mode matching is achieved for atoms confined as a uniform dielectric sheet at a fixed zplane, similar to the geometry shown in Fig. 2(b). However, for a finite number of atoms, the realizable OD is low in this configuration. Indeed, a uniform dielectric slab of extent much larger than the beam waist achieves perfect mode matching, but one cannot achieve such an dielectric distribution with high OD using cold atomic gases. An intermediate pencil-shaped geometry is more realistic, allowing for reasonable mode



FIG. 2. (Color online) Diagrams of the scattered modes for various atomic cloud and beam geometries. The spatial profile of the probe mode is indicated with solid blue lines and that of the field scattered by a given dielectric distribution is indicated by solid red lines with arrows. (a) A pointlike atomic ensemble scatters light isotropically. (b) A pancake-shaped cloud at a fixed z plane radiates nearly perfectly into the probe mode. Extended pencil-shaped clouds can radiate into the probe mode well, as in (c), or poorly, as in (d), depending on the geometry of the probe.

matching while maintaining a high OD, as in Figs. 2(c) and 2(d).

In addition to maximizing the signal by mode matching the pattern of scattered light with that of the probe mode (a calculation the follows from the semiclassical theory), we must also minimize the sources of noise. There are two fundamental quantum effects: (i) the polarimeter has a finite shot-noise resolution; (ii) atoms scatter photons diffusely into all directions (spontaneous emission). The latter is accompanied by optical pumping that can both depolarize the spins and inject noise into the measured spin wave. To address these effects, we turn to the fully quantum theory.

B. Quantum theory

Following Refs. [36,42], we partition the quantized electric field into paraxial modes and nonparaxial, diffuse modes,

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) = \hat{\mathbf{E}}_{\text{para}}^{(+)}(\mathbf{r},t) + \hat{\mathbf{E}}_{\text{diff}}^{(+)}(\mathbf{r},t).$$
(14)

This decomposition is motivated by the geometry we consider: photon scattering of a paraxial laser beam by an extended atomic ensemble. The mean-field, spatially averaged atomic density, which plays the role of the index of refraction in the classical theory, appears as coherent radiation by a collective atomic observable in the quantum theory. The coupling of collective atomic observables to paraxial modes thus describes the coherent atom-photon light-shift interaction, mediated by the Hermitian part of the atomic polarizability operator.

The diffuse modes, in contrast, couple to the density fluctuations in the ensemble due to the discrete atomic positions and thus act locally on each atom [22]. In the usual Born-Markov approximation, tracing over these modes leads to decoherence and is described by the anti-Hermitian part of the atomic polarizability [38].

In this section we first derive a multimode generalization of the Faraday interaction that coherently entangles the atomic ensemble and the paraxial quantum field. Then we employ a master equation to account for the effects of local decoherence (optical pumping) driven by diffuse scattering. Finally, we present the stochastic master equation describing the conditional collective atomic state given polarimetry measurements of the paraxial field, which we use to analyze spin squeezing in Sec. III.

1. Paraxial multimode Faraday interaction

Quantization of paraxial electromagnetic fields was discussed in Ref. [43]; relevant extensions to the current problem are summarized in Appendix B. We decompose the paraxial field operator into an orthogonal set of transverse spatial modes, here the Laguerre-Gauss modes $u_{pl}(\mathbf{r}_{\perp}, z)$ given in Eq. (B8), which are convenient for cylindrical symmetry. The positive-frequency component of the electric field restricted to the paraxial subspace is

$$\hat{\mathbf{E}}_{\text{para}}^{(+)}(\mathbf{r},t) = \sum_{p,l,\alpha} \sqrt{\frac{2\pi\hbar\omega_0}{cA}} \, \mathbf{e}_{\alpha} \, \hat{a}_{pl,\alpha}(z,t) \, u_{pl}(\mathbf{r}_{\perp},z) e^{i(k_0 z - \omega_0 t)}, \quad (15)$$

where $\alpha = x, y$ labels transverse polarization, and the quantization area is chosen as the natural scale of the Gaussian beam, $A = \pi w_0^2/2$. The traveling wave creation and annihilation operators for each transverse mode freely propagate according the Hamiltonian

$$\hat{H}_{\text{free}} = \sum_{p,l,\alpha} \int dz \, \hat{a}_{pl,\alpha}^{\dagger}(z,t) \left(-i\hbar c \frac{\partial}{\partial z} \right) \hat{a}_{pl,\alpha}(z,t), \quad (16)$$

with solution $\hat{a}_{pl,\alpha}(z,t) = \hat{a}_{pl,\alpha}(z-ct,0) = \hat{a}_{pl,\alpha}(0,t-z/c)$ and free-field commutation relations,

$$[\hat{a}_{pl,\alpha}(z,t), \hat{a}_{p'l',\beta}^{\top}(z',t')] = \delta_{p,p'} \delta_{l,l'} \delta_{\alpha,\beta} \,\delta(t-t'-(z-z')/c).$$
(17)

We have normalized so that $\hat{a}_{pl,\alpha}^{\dagger}(z,t)\hat{a}_{pl,\alpha}(z,t)$ is the local photon flux in transverse mode pl with polarization α .

For weak excitation (linear atomic response), the interaction Hamiltonian governing the coupling of the quantized paraxial modes is

$$\hat{H}_{\text{int}} = -\sum_{i} \hat{\mathbf{E}}_{\text{para}}^{(-)}(\mathbf{r}_{i}, t) \cdot \hat{\vec{\alpha}}^{(i)} \cdot \hat{\mathbf{E}}_{\text{para}}^{(+)}(\mathbf{r}_{i}, t).$$
(18)

As before, the index *i* is summed over atoms in the ensemble at respective positions \mathbf{r}_i . Upon substituting the decomposition of $\hat{\alpha}_{ij}$ into its irreducible components given in Eq. (12), we find scalar (rank 0), vector (rank 1), and tensor (rank 2) contributions to the interaction. We retain only the vector contribution that leads to the Faraday effect, as discussed in the semiclassical case. The Faraday interaction is then

$$\hat{H}_{\text{int}} = -i\alpha_0 C^{(1)} \sum_i \hat{\mathbf{E}}_{\text{para}}^{(+)}(\mathbf{r}_i, t) \times \hat{\mathbf{E}}_{\text{para}}^{(-)}(\mathbf{r}_i, t) \cdot \hat{\mathbf{f}}^{(i)}$$

$$= -\frac{\hbar\chi}{2} \sum_{i, p, l, p', l'} [iu_{p'l'}^*(\mathbf{r}_{\perp i}, z_i)u_{pl}(\mathbf{r}_{\perp i}, z_i)$$

$$\times \hat{a}_{p'l', y}^{\dagger}(z_i, t) \hat{a}_{pl, x}(z_i, t) + \text{H.c.}]\hat{f}_z^{(i)}, \quad (19)$$

where

$$\chi = C^{(1)} \frac{4\pi \omega_0}{Ac} \alpha_0 = C^{(1)} \left(\frac{\sigma_0}{A}\right) \left(\frac{\Gamma}{2\Delta}\right)$$
(20)

is the Faraday rotation angle, $\sigma_0 = 3\lambda^2/(2\pi)$ is the resonant cross-section for unit oscillator strength, Γ is the atomic linewidth, and Δ is the detuning from resonance. For an $S_{1/2} \rightarrow P_J$ transition with Δ much larger than the excited-state hyperfine splitting, $|C^{(1)}| = g_f/3$, where g_f is the Landé g-factor. We can interpret Eq. (19) as a scattering process, whereby an x-polarized photon in a given transverse mode, pl, is absorbed and a y-polarized photon in the mode p'l' is emitted, and vice versa, as mediated by the collective atomic spin wave.

We consider an initial macroscopic occupation in the laser probe, again taken to be the fundamental Gaussian TEM₀₀ mode with x polarization. In that case the interaction can be linearized by substituting $\hat{a}_{pl,x}(z,t) \rightarrow \sqrt{\dot{N}_L} \delta_{p,0} \delta_{l,0}$, where $\dot{N}_L = AI_0/(\hbar\omega_0)$ is the photon flux of the laser with peak intensity I_0 . The quantum fluctuations in the field of interest are then represented by the y-polarized mode, $\hat{a}_{pl,y}(z,t)$, and the Faraday interaction takes the form

$$\hat{H}_{\text{int}} = -i \frac{\hbar \sqrt{\kappa}}{2} \sum_{i,p,l} [\beta_{pl}^*(\mathbf{r}_{\perp i}, z_i) \hat{a}_{pl,y}(z_i, t) - \text{H.c.}] \hat{f}_z^{(i)}, \quad (21)$$

where the local amplitude for scattering from the fundamental (laser) mode 00 into mode pl is given by

$$\beta_{pl}(\mathbf{r}_{\perp}, z) \equiv u_{pl}^*(\mathbf{r}_{\perp}, z) u_{00}(\mathbf{r}_{\perp}, z).$$
(22)

The interaction has been written in terms of the "measurement strength" per atom,

$$\kappa = \chi^2 \dot{N}_L = \frac{g_f^2}{9} \left(\frac{\sigma_0}{A}\right) \gamma_0, \tag{23}$$

which characterizes the rate at which photons are scattered into the paraxial modes, where $\gamma_0 = (\sigma_0 I_0/\hbar\omega_0)(\Gamma^2/4\Delta^2)$ is the unit-oscillator-strength photon scattering rate at the peak intensity.

The Heisenberg equation of motion for a *y*-polarized traveling wave mode interacting with the atomic media in the presence of the probe field is

$$\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial z}\right) \hat{a}_{pl,y}(z,t)$$

= $\frac{\sqrt{\kappa}}{2} \sum_{i} \beta_{pl}(\mathbf{r}_{\perp i}, z_i) \hat{f}_{z}^{(i)}(t) \delta(z - z_i),$ (24)

whose solution is

$$\hat{a}_{pl,y}(z,t) = \hat{a}_{pl,y}(0,t-z/c) + \frac{\sqrt{\kappa}}{2} \sum_{i} \beta_{pl}(\mathbf{r}_{\perp i}, z_{i}) \\ \times \hat{f}_{z}^{(i)}(t-(z-z_{i})/c)\Theta(z-z_{i}),$$
(25)

where $\Theta(z)$ is the Heaviside step function. Neglecting the time it takes light to propagate across the sample, the mode amplitude at the detector plane, z_D , in the far field is

$$\hat{a}_{pl,y}(z_D,t) = \hat{a}_{pl,y}(0,t-z_D/c) + \frac{\sqrt{\kappa}}{2}\hat{F}_z^{pl}, \qquad (26)$$

a form familiar from input-output theory [44]. The collective atomic spin wave that couples to paraxial mode pl is

$$\hat{F}_{z}^{pl} \equiv \sum_{i} \beta_{pl}(\mathbf{r}_{\perp i}, z_{i}) \hat{f}_{z}^{(i)}.$$
(27)

In the balanced polarimeter, the probe mode acts as a local oscillator so that one measures the Stokes vector component associated with the fundamental spatial mode defined by the laser beam, $(\hat{a}_{00,x}^{\dagger}\hat{a}_{00,y} + \hat{a}_{00,y}^{\dagger}\hat{a}_{00,x})/2 \approx \sqrt{N_L/2}\hat{X}_{00}$, where $\hat{X}_{00} = (\hat{a}_{00,y} + \hat{a}_{00,y}^{\dagger})/\sqrt{2}$ is the mode quadrature. The measured quadrature at the detector plane z_D ,

$$\hat{X}_{00}(z_D,t) = \hat{X}_{00}(0,t-z_D/c) + \sqrt{\frac{\kappa}{2}}\hat{F}_z^{00}, \qquad (28)$$

is proportional to the *fundamental spin wave*, \hat{F}_z^{00} . Note that this is exactly the spin wave found in the semiclassical calculation, Eq. (13). The total polarimeter signal, integrated over a time *T*, is determined by the output operator

$$\hat{\mathcal{M}} = \int_0^T dt \, \hat{X}_{00}(0, t - z_D/c) + T \sqrt{\frac{\kappa}{2}} \hat{F}_z^{00}.$$
 (29)

The fully quantum theory explicitly includes the additional vacuum noise entering the polarimeter, $\langle \Delta \hat{X}_{00}(0,t) \Delta \hat{X}_{00}(0,t') \rangle = \delta(t-t')/2$, that leads to a shot-noise (SN) variance of the polarimeter signal, in Eq. (29), $\Delta \mathcal{M}_{SN}^2 = T/2$.

Of particular interest here is the application to spin squeezing via QND measurement. In this case, the signal we seek to measure arises from different spin projections associated with the eigenstates of \hat{F}_z^{00} . Whereas in magnetometry these shot-to-shot variations are known as "projection noise" (PN), in the context of creating a spin-squeezed state, these variations from the mean value represent the "signal" one seeks to resolve over the laser shot noise. For the fundamental spin wave measured in the polarimeter, the projection noise variance is

$$\left(\Delta F_z^{00}\right)_{\rm PN}^2 = \sum_i \beta_{00}^2(\mathbf{r}_i) \left(\left(\Delta \hat{f}_z^{(i)} \right)^2 \right) + \sum_{i \neq j} \beta_{00}(\mathbf{r}_i) \beta_{00}(\mathbf{r}_j) \left(\Delta \hat{f}_z^{(i)} \Delta \hat{f}_z^{(j)} \right).$$
(30)

Given an initial spin coherent state polarized orthogonal to z, $\langle \Delta \hat{f}_z^{(i)} \Delta \hat{f}_z^{(j)} \rangle = (f/2)\delta_{ij}$, and thus,

$$\left(\Delta F_z^{00}\right)_{\rm PN}^2 = \frac{N_{\rm eff}^{(2)} f}{2}.$$
 (31)

Here we define a set of effective atom numbers,

$$N_{\rm eff}^{(K)} = \sum_{i} [\beta_{00}(\mathbf{r}_{i})]^{K} = \sum_{i} |u_{00}(\mathbf{r}_{i})|^{2K}$$
(32)

$$\rightarrow \int d^3 \mathbf{r} \, \eta(\mathbf{r}) |u_{00}(\mathbf{r}_i)|^{2K}, \qquad (33)$$

where the sum becomes an integral in the continuum limit. The atomic density distribution $\eta(\mathbf{r})$ is normalized so that $\int d^3 \mathbf{r} \, \eta(\mathbf{r}) = N$, the total atom number. The effective atom number $N_{\text{eff}}^{(2)}$ determines the projection-noise contribution to Eq. (29), $\Delta \mathcal{M}_{\text{PN}}^2 = \kappa T^2 (\Delta F_z^{00})_{\text{PN}}^2/2 = \kappa T^2 N_{\text{eff}}^{(2)} f/4$. The coupling strength ξ that sets the degree of entanglement one can attain between the atoms and photons is the ratio of the projection-noise variance to the shot-noise resolution [38]. Using Eqs. (23) and (31) we find

$$\xi = \frac{\Delta \mathcal{M}_{\rm PN}^2}{\Delta \mathcal{M}_{\rm SN}^2} = \left(\Delta F_z^{00}\right)^2 \kappa T = OD_{\rm eff} \frac{g_f^2 f}{18f} \gamma_0 T, \qquad (34)$$

where we have defined the effective optical density for the laser mode probing the spin wave on a unit-oscillator-strength transition,

$$OD_{eff} \equiv N_{eff}^{(2)} \frac{\sigma_0}{A}.$$
 (35)

The key to achieving a large OD_{eff} is choosing an atomic and beam geometry that addresses a large number of atoms and maximizes $N_{eff}^{(2)}$ while keeping the mode area A small. It should be noted that whereas in the one-dimensional case the optical density is associated with both the coupling strength and the Beer's law attenuation of the probe, in the three-dimensional case different parameters are associated with each of these effects. Because the attenuation coefficient in Eq. (6) is proportional to the local intensity of the field, the total attenuation depends upon the effective atom number $N_{eff}^{(1)}$.

While Eq. (34) implies an ever increasing coupling strength with integration time T, we have neglected so far the decoherence that limits the total useful integration time and the strength of the atom-light interface. In the following section we treat these effects from a first-principles master equation, including spatial variations in the scattering rate which drives local decoherence.

2. Local decoherence and optical pumping

The discrete random atomic positions are associated with the density fluctuations that give rise to diffuse scattering into 4π sr [22]. We consider light far detuned from any atomic resonance in a highly transparent regime, and thus we can safely neglect the small attenuation of the laser probe associated with this absorption. The scattering processes, however, cause decoherence of the spin wave due to optical pumping. This local decoherence breaks the collective symmetry of the problem and adds additional noise, which is detected in the polarimeter and competes with squeezing.

To treat the decoherence due to diffuse scattering, we employ a master equation,

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar} [\hat{H}_{\rm int}, \hat{\rho}] + \frac{d\hat{\rho}}{dt} \bigg|_{\rm diff}, \tag{36}$$

where \hat{H}_{int} is the multimode Faraday interaction given in Eq. (21). The key feature of this equation is that the paraxial modes couple to collective spin waves, while the diffuse scattering couples to localized atoms and induces optical pumping according to

$$\left. \frac{d\hat{\rho}}{dt} \right|_{\text{diff}} = \sum_{i} \gamma_s(\mathbf{r}_i) \mathcal{D}_i[\hat{\rho}]. \tag{37}$$

The map \mathcal{D}_i acts on the *i*th atom, proportional to the local scattering rate,

$$\gamma_s(\mathbf{r}_i) = I(\mathbf{r}_i) \frac{\sigma_0}{\hbar\omega} \frac{\Gamma^2}{4\Delta^2} = \gamma_0 \beta_{00}(\mathbf{r}_i).$$
(38)

Here $I(\mathbf{r}_i) = I_0 \beta_{00}(\mathbf{r}_i)$ is the local intensity at the position of the atom and γ_0 is the peak scattering rate. We consider here a probe driving an $S_{1/2} \rightarrow P_J$ transition in an alkalimetal atom, with a detuning that is small compared to the ground-state hyperfine splitting but large compared to any hyperfine splitting in the excited state. In this case, the light coherently couples substantially only to atoms in a given ground-electronic hyperfine manifold f and the master equation is restricted to this subspace. As shown in Appendix C, with an *x*-polarized probe and applying a large bias magnetic field along the z direction, the local decoherence in the master equation due to optical pumping is given by the map

$$\mathcal{D}_{i}[\hat{\rho}] = -\frac{2}{9}\hat{\rho} + \frac{g_{f}^{2}}{9} \bigg[\hat{f}_{z}^{(i)}\hat{\rho}\hat{f}_{z}^{(i)} + \frac{1}{2} \big(\hat{f}_{x}^{(i)}\hat{\rho}\hat{f}_{x}^{(i)} + \hat{f}_{y}^{(i)}\hat{\rho}\hat{f}_{y}^{(i)} \big) \bigg].$$
(39)

The first term on the right-hand side of Eq. (39) describes the decay of correlations due to optical pumping, while the second term (in brackets) represents a feeding due to "transfer of coherences" that can reduce this decay rate [45]. Note that for f > 1/2, this master equation is not trace preserving, since atoms can be optically pumped to the other ground hyperfine manifold, where they are lost to any further measurement.

Given the master equation, we can find the effect of diffuse scattering on atomic correlations. Consider an inhomogeneous collective operator of the form $\hat{X} = \sum_i \beta_{00}(\mathbf{r}_i)\hat{x}^{(i)}$. Because \hat{X} is a weighted sum over single atom operators, the equation of motion for its expectation value depends upon the evolution of the single atom density operator, $\hat{\rho}^{(i)}$. By summing over a single index *i* in Eq. (37), we obtain

$$\left. \frac{d\hat{\rho}^{(i)}}{dt} \right|_{\text{diff}} = \gamma_s(\mathbf{r}_i) \mathcal{D}_i[\hat{\rho}^{(i)}], \tag{40}$$

from which the evolution of $\langle \hat{X} \rangle$ is given by

$$\frac{d}{dt} \langle \hat{X} \rangle \bigg|_{\text{diff}} = \sum_{i} \beta_{00}(\mathbf{r}_{i}) \text{Tr} \bigg[\hat{x}^{(i)} \frac{d\hat{\rho}^{(i)}}{dt} \bigg|_{\text{diff}} \bigg]$$
$$= \sum_{i} \gamma_{s}(\mathbf{r}_{i}) \beta_{00}(\mathbf{r}_{i}) \langle \mathcal{D}_{i}[\hat{x}^{(i)}] \rangle.$$
(41)

For inhomogeneous collective operators that depend on pairs of atoms,

$$\hat{O} = \sum_{i \neq j} \beta_{00}(\mathbf{r}_i) \beta_{00}(\mathbf{r}_j) \hat{x}^{(i)} \hat{y}^{(j)}, \qquad (42)$$

we require the joint density operator of the *i*th and *j*th atoms, $\hat{\rho}^{(i,j)}$, with equation of motion

$$\left. \frac{d}{dt} \hat{\rho}^{(i,j)} \right|_{\text{diff}} = \gamma_s(\mathbf{r}_i) \mathcal{D}_i[\hat{\rho}^{(i,j)}] + \gamma_s(\mathbf{r}_j) \mathcal{D}_j[\hat{\rho}^{(i,j)}].$$
(43)

The evolution of $\langle \hat{O} \rangle$ due to diffuse scattering is then

$$\frac{d}{dt}\langle\hat{O}\rangle\Big|_{\text{diff}} = \sum_{i\neq j} \{\gamma_s(\mathbf{r}_i)\beta_{00}(\mathbf{r}_i)\beta_{00}(\mathbf{r}_j)\langle\mathcal{D}_i[\hat{x}^{(i)}]\hat{y}^{(j)}\rangle + \gamma_s(\mathbf{r}_j)\beta_{00}(\mathbf{r}_i)\beta_{00}(\mathbf{r}_j)\langle\hat{x}^{(i)}\mathcal{D}_j[\hat{y}^{(j)}]\rangle\}. \quad (44)$$

The degree of squeezing that one can ultimately produce is determined by a balance between QND measurement backaction on the spin wave mediated by the collective radiation and the damage to that observable caused by diffuse scattering. To properly treat this, we must include the effects of measurement on the atoms, as discussed in the next section.

3. The conditional stochastic master equation

The Faraday Hamiltonian [Eq. (21)] is an entangling interaction between the atomic spin waves and the paraxial modes of the field. When the light is measured in the polarimeter, quantum backaction leads to stochastic evolution of the atomic state, conditioned on the random measurement results. A complete description of the dynamics is then described by a stochastic master equation (SME), with decoherence from unmeasured light and squeezing due to information gained from the continuous measurement record. In a balanced polarimeter, the measurement signal is proportional to the interference of the probe and scattered fields integrated over the detector faces, as in Eq. (9). Due to the orthogonality of the spatial modes [Eq. (B9)], such a measurement selects only paraxial light that is scattered into the mode of the probe, u_{00} . The result is a continuous measurement of the quadrature \hat{X}_{00} .

We derive the SME for the atoms following [46,47], with details presented in Appendix D. Measurement of the quadrature \hat{X}_{00} by the homodyne polarimeter generates a differential stochastic measurement record,

$$dy_{00} = \left\langle \hat{F}_z^{00} \right\rangle dt + \frac{1}{\sqrt{\kappa}} dW, \tag{45}$$

where dW is a Weiner interval in the Itō calculus and κ is the measurement strength given in Eq. (23). Assuming unit measurement efficiency, the evolution of the ensemble conditioned upon the measurement record y_{00} is given by

$$d\hat{\rho} = \sqrt{\frac{\kappa}{4}} \mathcal{H}_{00}[\hat{\rho}] dW + \frac{\kappa}{4} \sum_{p,l} \mathcal{L}_{pl}[\hat{\rho}] dt.$$
(46)

The effects of measurement backaction on the ensemble are taken into account by the superoperator $\mathcal{H}_{00}[\hat{\rho}]$, where

$$\mathcal{H}_{pl}[\hat{\rho}] \equiv \hat{F}_{z}^{pl}\hat{\rho} + \hat{\rho}\hat{F}_{z}^{pl\dagger} - \mathrm{Tr}\big[\big(\hat{F}_{z}^{pl} + \hat{F}_{z}^{pl\dagger}\big)\hat{\rho}\big]\hat{\rho}.$$
 (47)

The Lindblad dissipator,

$$\mathcal{L}_{pl}[\hat{\rho}] \equiv \hat{F}_{z}^{pl} \hat{\rho} \hat{F}_{z}^{pl\dagger} - \frac{1}{2} \hat{F}_{z}^{pl\dagger} \hat{F}_{z}^{pl} \hat{\rho} - \frac{1}{2} \hat{\rho} \hat{F}_{z}^{pl\dagger} \hat{F}_{z}^{pl}, \quad (48)$$

describes the effect on the atomic ensemble arising from indistinguishable radiation into paraxial modes of the field.

Including local decoherence from diffuse scattering [Eq. (39)], the full SME for homodyne polarimetry measure-

ments of the 00 mode is

$$d\hat{\rho} = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}] dt + \sqrt{\frac{\kappa}{4}} \mathcal{H}_{00}[\hat{\rho}] dW + \frac{\kappa}{4} \sum_{p,l} \mathcal{L}_{pl}[\hat{\rho}] dt + \sum_{i} \gamma_{s}(\mathbf{r}_{i}) \mathcal{D}_{i}[\hat{\rho}] dt.$$
(49)

This SME is a complete description of the evolution of the collective atomic state, accounting for the three-dimensional nature of the atom-photon modes, decoherence, and measurement backaction. We see that through its interaction with the probe, the atomic ensemble undergoes an additional form of *collective* decoherence [Eq. (48)] corresponding to light radiated into paraxial modes $pl \neq 00$ that ultimately goes unmeasured. Thus, we have arrived at the same conclusion as in Ref. [27]. That is, decoherence arises through two distinct processes: first, the inherent mode mismatch that gives rise to collectively scattered light in spatial modes other than the probe mode, and, second, the diffuse scattering of photons that acts locally on atoms in the ensemble.

III. QND SQUEEZING OF SPIN WAVES

A. Quantifying squeezing of the spin waves

One typically quantifies the amount of squeezing created in a QND measurement according to the Wineland squeezing parameter [48],

$$\zeta = \left(\frac{\Delta\phi}{\Delta\phi_{\rm SCS}}\right)^2,\tag{50}$$

where $\Delta \phi$ is the projection-noise limited resolution when measuring an angle of rotation for a generic spin J of the given input state, and $\Delta \phi_{\text{SCS}}$ is the corresponding resolution when the input is a spin coherent state (SCS). For a mean value $J_{\parallel} \equiv$ $|\langle \hat{\mathbf{J}} \rangle|$ and variance ΔJ_{\perp}^2 orthogonal to the mean, the projectionnoise resolution is $\Delta \phi = \Delta J_{\perp}/J_{\parallel}$. With $\Delta \phi_{\text{SCS}} = 1/\sqrt{2J}$, the Wineland squeezing parameter is then

$$\zeta = 2J \frac{\Delta J_{\perp}^2}{J_{\parallel}^2}.$$
 (51)

For the spin waves of the inhomogeneous ensemble under consideration here, we must tie the squeezing parameter directly to the measured quantities. For an initial mean spin polarization along x and a small rotation around y, the polarimeter signal will be determined by the mean spin wave component $\langle \hat{F}_x^{00} \rangle = \sum_i \beta_{00}(\mathbf{r}_i) \langle \hat{f}_x^{(i)} \rangle$, and the projection noise contribution to the resolution of the measurement will be given by ΔF_z^{00} , defined in Eq. (30). The projection-noise limited resolution of this rotation is therefore $\Delta \phi^{00} = \Delta F_z^{00} / \langle \hat{F}_x^{00} \rangle$. Furthermore, given a SCS initially polarized along x, the mean spin of interest is $\langle \hat{F}_x^{00} \rangle_{\text{SCS}} = N_{\text{eff}}^{(1)} f$, where the effective atom number contributing to this signal is given in Eq. (32), while the projection-noise limited resolution for a SCS preparation is thus $(\Delta \phi_{\text{SCS}}^{00})^2 = N_{\text{eff}}^{(2)} / [2f(N_{\text{eff}}^{(1)})^2]$ and will depend on the shape of the atomic cloud and beam geometry. Putting this together, we define the squeezing parameter for the measured

spin wave to be

$$\zeta \equiv \left(\frac{\Delta\phi^{00}}{\Delta\phi_{\rm SCS}^{00}}\right)^2 = 2f \frac{\left(N_{\rm eff}^{(1)}\right)^2}{N_{\rm eff}^{(2)}} \frac{\left(\Delta F_z^{00}\right)^2}{\left(\hat{F}_x^{00}\right)^2}.$$
 (52)

This parameter quantifies the degree of "quantum backaction" on a SCS, accounting for the change in projection noise due to QND measurement as well as the damage done to both the mean spin polarization and variance due to optical pumping.

In a real-world metrological application such as an optically probed atomic magnetometer [49,50], spin rotations are measured by passing the probe through the atom sample and measuring the resulting Faraday rotation in a polarimeter. In addition to spin projection noise, the measurement resolution is then subject also to "technical noise," including probe shot noise, detector electronic noise, and atom number fluctuations. Under those circumstances, optimizing the squeezing parameter as defined in Eq. (52) is distinct from optimizing the magnetometer sensitivity.

B. The dynamical evolution of squeezing

To determine the squeezing as function of time, we employ the SME in Eq. (49) to track $(\Delta F_z^{00})^2$ and $\langle \hat{F}_x^{00} \rangle$. For ensembles with large numbers of atoms, we can work in the centrallimit approximation where fluctuations in the spin waves are treated as Gaussian random variables [36,51]. Following [46], the SME then couples solely means and covariances. The moments of the fundamental spin wave that characterize the spin-squeezing parameter then evolve according to

$$d(\Delta F_z^{00})^2 = -\kappa \left[\left(\Delta F_z^{00} \right)^2 \right]^2 dt + d\left(\Delta F_z^{00} \right)^2 \Big|_{\text{diff}}, \quad (53a)$$
$$d\langle \hat{F}_x^{00} \rangle = \sqrt{\frac{\kappa}{4}} \langle \mathcal{H}_{00} [\hat{F}_x^{00}] \rangle dW$$
$$+ \sum_{p,l} \frac{\kappa}{4} \langle \mathcal{L}_{pl} [\hat{F}_x^{00}] \rangle dt + d\langle \hat{F}_x^{00} \rangle \Big|_{\text{diff}}. \quad (53b)$$

Because we assume the fundamental mode is measured with unit efficiency, diffuse scattering by local spontaneous emission is the only process contributing to the decoherence of the variance $(\Delta F_z^{00})^2$. Collective radiation into other transverse modes commutes with \hat{F}_z^{00} and does not contribute to any decay or noise injection into the fundamental variance. In contrast, the mean spin $\langle \hat{F}_x^{00} \rangle$ decoheres due to both diffuse scattering and collective scattering into other unmeasured paraxial modes. It also evolves stochastically due to the continuous measurement of \hat{F}_z^{00} . However, the contributions to the dynamics from both collective scattering and continuous measurement are small in comparison to diffuse scattering and can be neglected when the radiation pattern of the cloud is well matched to that of the probe.

We consider the moment evolution [Eq. (53)] with the initial condition that the ensemble is in a SCS polarized along *x*. The initial mean spin and variance are $\langle \hat{F}_x^{00}(t_0) \rangle = N_{\text{eff}}^{(1)} f$ and $[\Delta F_z^{00}(t_0)]^2 = N_{\text{eff}}^{(2)} f/2$, respectively. Along with the cross-sectional area of the probe laser, $N_{\text{eff}}^{(2)}$ specifies the effective optical density, OD_{eff} defined in Eq. (35). The OD_{eff} is the critical geometric parameter for determining how the atomic density distribution influences collective scattering into the

probe mode and ultimately leads to spin squeezing. Both of these effective atom numbers are determined solely by the cloud shape and beam geometry and can be found from the semiclassical model in Sec. II A.

For times short compared to the photon scattering rate, where decoherence is negligible, the mean spin is essentially constant and the spin variance is affected only by measurement backaction. The solution to Eq. (53) takes the familiar form [51],

$$\left[\Delta F_z^{00}(t)\right]^2 = \frac{\left[\Delta F_z^{00}(t_0)\right]^2}{1 + \left[\Delta F_z^{00}(t_0)\right]^2 \kappa t} = \frac{\left[\Delta F_z^{00}(t_0)\right]^2}{1 + \xi}$$
(54)

$$\Rightarrow \zeta = \frac{1}{1+\xi},\tag{55}$$

where ξ is the integrated coupling strength in Eq. (34). In Eq. (55), the squeezing parameter decreases as OD_{eff}^{-1} for $\xi \gg 1$.

For longer times, decoherence due to diffuse photon scattering must be included. The mean spin will depolarize according to Eq. (41),

$$\frac{d}{dt} \langle \hat{F}_x^{00} \rangle = \sum_i \gamma_s(\mathbf{r}_i) \beta_{00}(\mathbf{r}_i) \langle \mathcal{D}_i [\hat{f}_x^{(i)}] \rangle.$$
(56)

The variance involves both single-atom and pairwise atomic correlations,

$$\frac{d}{dt} \left(\Delta F_z^{00} \right)^2 = \sum_i [\beta_{00}(\mathbf{r}_i)]^2 \frac{d}{dt} \left\langle \left(\Delta \hat{f}_z^{(i)} \right)^2 \right\rangle + \sum_{i \neq j} \beta_{00}(\mathbf{r}_i) \beta_{00}(\mathbf{r}_j) \frac{d}{dt} \left\langle \Delta \hat{f}_z^{(i)} \Delta \hat{f}_z^{(j)} \right\rangle, \quad (57)$$

where the first term is the spin projection noise of the uncorrelated spins and the second term contains the correlations that generate spin squeezing. Following Eqs. (41) and (44), these correlation functions decay due to diffuse scattering according to

$$\frac{d}{dt} \sum_{i} \langle (\Delta \hat{f}_{z}^{(i)})^{2} \rangle |_{\text{diff}}$$

$$= \sum_{i} \gamma_{s}(\mathbf{r}_{i}) \{ \langle \mathcal{D}_{i} [\hat{f}_{z}^{(i)2}] \rangle - 2 \langle \mathcal{D}_{i} [\hat{f}_{z}^{(i)}] \rangle \langle \hat{f}_{z}^{(i)} \rangle \}, \quad (58a)$$

$$\frac{d}{dt} \sum_{i \neq j} \langle \Delta \hat{f}_{z}^{(i)} \Delta \hat{f}_{z}^{(j)} \rangle |_{\text{diff}}$$

$$= \sum_{i \neq j} \{ \gamma_{s}(\mathbf{r}_{i}) \langle \Delta \mathcal{D}_{i} [\hat{f}_{z}^{(i)}] \Delta \hat{f}_{z}^{(j)} \rangle + \gamma_{s}(\mathbf{r}_{j}) \langle \Delta \hat{f}_{z}^{(i)} \Delta \mathcal{D}_{j} [\hat{f}_{z}^{(j)}] \rangle \}.$$

The local decoherence acts via the map D_i in Eq. (39).

1. Spin- $\frac{1}{2}$ ensembles

We first restrict our attention to ensembles of spin f = 1/2 atoms ($g_f = 2$) to focus on spatial effects without the

complications that arise for ensembles with larger spin. Using the fact that the local scattering rate is proportional to the probe intensity, $\gamma_s(\mathbf{r}) = \gamma_0 \beta_{00}(\mathbf{r})$, the mean spin evolution of Eq. (53b) is

$$\frac{d}{dt}\langle \hat{F}_x^{00} \rangle = -\frac{\gamma_0}{3} \sum_i [\beta_{00}(\mathbf{r}_i)]^2 \langle \hat{f}_x^{(i)} \rangle.$$
(59)

The local decoherence does not respect the orthogonality of the transverse paraxial modes and we will see that the diffuse scattering acts to couple the fundamental spin wave to higherorder spin waves.

Because the transverse modes are orthogonal in a plane at a fixed z, we can derive a set of coupled equations by decomposing products of the spatial weighting coefficients [Eq. (22)] in the basis of mode functions as

$$[\beta_{00}(\mathbf{r}_{\perp},z)]^{2} = |u_{00}(\mathbf{r}_{\perp},z)|^{4} = \sum_{p,l} c_{pl}^{00}(z)\beta_{pl}(\mathbf{r}_{\perp},z), \quad (60)$$

with z-dependent projection coefficients,

$$c_{pl}^{00}(z) \equiv \frac{1}{A} \int d^2 \mathbf{r}_{\perp} [u_{00}(\mathbf{r}_{\perp}, z)]^2 u_{00}^*(\mathbf{r}_{\perp}, z) u_{pl}(\mathbf{r}_{\perp}, z).$$
(61)

It follows that Eq. (59) can be expressed as

$$\frac{d}{dt}\langle \hat{F}_x^{00}\rangle = -\frac{\gamma_0}{3} \sum_{k,p,l} c_{pl}^{00}(z_k) \langle \hat{F}_x^{pl}(z_k) \rangle, \tag{62}$$

where we have defined spin waves in coarse-grained slices of thickness δz around z_k ,

$$\hat{F}_{z}^{pl}(z_{k}) \equiv \sum_{i_{k}} \beta_{pl} \left(\mathbf{r}_{\perp i_{k}}, z_{i_{k}} \right) \hat{f}_{z}^{(i_{k})}, \tag{63}$$

with the index i_k labeling atoms in the slice z_k . The total spin wave for a given transverse mode is $\hat{F}_z^{pl} = \sum_k \hat{F}_z^{pl}(z_k)$. The mean spin in the fundamental mode $\langle \hat{F}_x^{00} \rangle$ is thus coupled to other spin waves $\langle \hat{F}_x^{pl}(z_k) \rangle$ within each slice z_k . Generally, the expected value of the *pl*-spin wave in the slice z_k evolves according to

$$\frac{d}{dt} \langle \hat{F}_{x}^{pl}(z_k) \rangle = -\frac{\gamma_0}{3} \sum_{p',l'} c_{p'l'}^{pl}(z_k) \langle \hat{F}_{x}^{p'l'}(z_k) \rangle, \qquad (64)$$

with projection coefficients $c_{p'l'}^{pl}(z_k)$ given in Eq. (E3). Details of this derivation are found in Appendix E. The initial conditions [Eq. (E5)] account for the matching between the probe mode and cloud geometry. By projecting onto the spin waves, we obtain a hierarchy of coupled equations. Numerically, we truncate once the desired convergence is achieved.

The effect of diffuse scattering on the evolution of the collective spin variance follows in an analogous manner. For spin- $\frac{1}{2}$, $\langle \Delta \hat{f}_z^2 \rangle = 1/4$ for all atoms. The map for local decoherence, $\Delta D_i[\hat{f}_z^{(i)}] = -2\Delta \hat{f}_z^{(i)}/9$, corresponds to decay of spin-spin correlations with no feeding of coherences. The evolution of the fundamental spin wave variance [Eq. (53a)]

simplifies to

$$\frac{d}{dt} \left(\Delta F_z^{00} \right)^2 = -\kappa \left[\left(\Delta F_z^{00} \right)^2 \right]^2 - \frac{2\gamma_0}{9} \sum_{i,j} \left[\beta_{00}(\mathbf{r}_i) + \beta_{00}(\mathbf{r}_j) \right] \\
\times \beta_{00}(\mathbf{r}_i) \beta_{00}(\mathbf{r}_j) \left\langle \Delta \hat{f}_z^{(i)} \Delta \hat{f}_z^{(j)} \right\rangle + \frac{\gamma_0}{9} \sum_i \left[\beta_{00}(\mathbf{r}_i) \right]^3,$$
(65)

where again we have used Eq. (38). The first term describes squeezing of the variance due to measurement backaction, the second represents decay of correlations due to diffuse scattering, and the third is the noise injected into the variance from spin flips (optical pumping). Following the same procedure as above, the decay terms are projected onto higher-order spin waves,

$$\frac{d}{dt} \left(\Delta F_z^{00}\right)^2 = -\kappa \left[\left(\Delta F_z^{00}\right)^2 \right]^2 - \frac{4\gamma_0}{9} \sum_{p,l} \sum_{k,k'} c_{pl}^{00}(z_k) \\ \times \left\langle \Delta \hat{F}_z^{00}(z_k) \Delta \hat{F}_z^{pl}(z_{k'}) \right\rangle + \frac{\gamma_0}{9} N_{\text{eff}}^{(3)}.$$
(66)

Here $N_{\text{eff}}^{(3)}$ is the effective atom number governing the injection of noise through optical pumping, defined in Eq. (32). Equation (66) is a covariance description of the dynamics, similar to that commonly employed for spin squeezing [52], but which also accounts for local decoherence from first principles. To solve for the fundamental variance, we must track the evolution of the covariances between coarse-grained slices and between transverse modes, $\langle \Delta \hat{F}_z^{pl}(z_k) \Delta \hat{F}_z^{p'l'}(z_{k'}) \rangle =$ $\langle \hat{F}_z^{pl}(z_k) \hat{F}_z^{p'l'}(z_{k'}) \rangle - \langle \hat{F}_z^{pl}(z_k) \rangle \langle \hat{F}_z^{p'l'}(z_{k'}) \rangle$. Equations of motion for these covariances follow readily from the SME. A detailed derivation is given in Appendix E.

2. Spin-f alkali-metal atom ensembles

The constituent atoms in many spin-squeezing experiments are alkali-metal atoms whose ground-state structure is more complex than spin- $\frac{1}{2}$. For example, in ¹³³Cs, the groundelectronic subspace is defined by two hyperfine manifolds with total spin angular momentum $f = \{3,4\}$. Owing to the large ground-state hyperfine splitting (9.2 GHz in ¹³³Cs), a single hyperfine manifold f is addressed by the coherent interaction with the probe laser.

Though ensembles of higher spin atoms can be squeezed by the same QND measurement process, spin size affects both the coherent squeezing dynamics and decoherence. Recall that the strength of the Faraday interaction is quantified by the coupling strength ξ [Eq. (34)]. Because $\xi \propto 1/f^2$, the atom-light coupling decreases with increasing spin size. This decreased coupling strength is partially offset by an increased robustness to the effects of optical pumping. When f > 1/2, optical pumping events can be broadly divided into two categories: (i) "loss" that occurs when an atom is pumped from the *f* manifold into the other ground hyperfine manifold and (ii) "spin flips" that leave the atom in the *f* manifold. Because atoms lost to the other ground manifold are no longer resonant with the probe, loss events decrease the mean spin $\langle \hat{F}_x^{00} \rangle$, though they contribute no excess noise to $(\Delta F_z^{00})^2$. Spin flips are responsible for both a decrease in $\langle \hat{F}_x^{00} \rangle$ and a noise injection into $(\Delta F_z^{00})^2$. For the SCS preparation, the deleterious effects of spin flips are mitigated by "transfers of coherence" between pairs of magnetic sublevels that reduce the rate of decay of correlations [45]. While the interplay between these effects is complex, the rate of spin flips remains a good indicator of an ensemble's robustness to optical pumping. For an ensemble of spin-*f* alkali metals prepared PHYSICAL REVIEW A 89, 033850 (2014)

in a SCS, the spin flip rate is $\gamma_s(\mathbf{r})/(12f)$, thus decreasing for larger spin size.

Due to these decoherence processes, the dynamics of the squeezing parameter is substantially more complicated for larger spin atoms. For spin-f, we obtain the evolution of the mean value of a spin wave in slice z_k by projecting onto the different spin waves in a manner analogous to Eq. (64),

$$\frac{d}{dt} \langle \hat{F}_{x}^{pl}(z_{k}) \rangle = -\frac{2\gamma_{0}}{9} \sum_{p',l'} c_{p'l'}^{pl}(z_{k}) \langle \hat{F}_{x}^{p'l'}(z_{k}) \rangle + \frac{g_{f}^{2}\gamma_{0}}{9} \sum_{p',l'} \sum_{i_{k}} c_{p'l'}^{pl}(z_{k}) \beta_{p'l'} \big(\mathbf{r}_{\perp i_{k}}, z_{k} \big) \mathcal{C}_{i_{k}} \big[\hat{f}_{x}^{(i_{k})} \big].$$
(67)

Here we have defined a local superoperator that arises solely from the "feeding" terms in the master equation:

$$C_{i}[\hat{X}] \equiv \hat{f}_{z}^{(i)} \hat{X} \hat{f}_{z}^{(i)} + \frac{1}{2} \left(\hat{f}_{x}^{(i)} \hat{X} \hat{f}_{x}^{(i)} + \hat{f}_{y}^{(i)} \hat{X} \hat{f}_{y}^{(i)} \right).$$
(68)

Similarly, we find equations of motion for the fundamental spin wave variance,

$$\frac{d}{dt} \left(\Delta F_{z}^{00} \right)^{2} = -\kappa \left[\left(\Delta F_{z}^{00} \right)^{2} \right]^{2} - \frac{4\gamma_{0}}{9} \sum_{p,l} \sum_{k',k} c_{pl}^{00}(z_{k}) \left\{ \Delta \hat{F}_{z}^{00}(z_{k'}) \Delta \hat{F}_{z}^{pl}(z_{k}) \right\}
+ \frac{g_{f}^{2} \gamma_{0}}{9} \sum_{p,l} \sum_{k',k,i_{k}} c_{pl}^{00}(z_{k}) \beta_{pl}(\mathbf{r}_{i_{k}}) \left\{ \left\langle \Delta \hat{F}_{z}^{00}(z_{k'}) \Delta \mathcal{C}_{i_{k}}[\hat{f}_{z}^{(i_{k})}] \right\rangle + \left\langle \Delta \mathcal{C}_{i_{k}}[\hat{f}_{z}^{(i_{k})}] \Delta \hat{F}_{z}^{00}(z_{k'}) \right\rangle \right\}
+ \gamma_{0} \sum_{k,i_{k}} \left[\beta_{00}(\mathbf{r}_{i_{k}}) \right]^{3} \left\{ \frac{2}{9} \left\langle (\hat{f}_{z}^{(i_{k})})^{2} \right\rangle + \frac{g_{f}^{2}}{9} \left\langle \left\langle \mathcal{C}_{i_{k}}[\hat{f}_{z}^{(i_{k})2}] \right\rangle - \left\langle \{\hat{f}_{z}^{(i_{k})}, \mathcal{C}_{i_{k}}[\hat{f}_{z}^{(i_{k})}] \}_{+} \right\rangle \right) \right\},$$
(69)

where $\{\hat{X}, \hat{Y}\}_+$ denotes the anticommutator. As for the case of spin- $\frac{1}{2}$, we have an infinite hierarchy of equations that couple spin wave operators in the different z_k slices. In general, the feeding terms in Eq. (69) couple to covariances outside the set $\langle \Delta \hat{F}_z^{pl}(z_k) \Delta \hat{F}_z^{p'l'}(z_{k'}) \rangle$. This expands considerably the number of equations that must be solved to reach convergence. Solving these equations, furthermore, requires different methods than the spin- $\frac{1}{2}$ case. A detailed treatment of the spin-f case will be provided in future work.

C. Results

Using our formalism we can calculate the dynamics of QND measurement and the peak achievable squeezing in the presence of decoherence. We now consider the fundamental effects of geometry and the optimization of experimentally relevant quantities to maximize spin squeezing. Most of our results are shown for the simplest case of spin- $\frac{1}{2}$ atoms in order to focus on the effects of spatial modes and spin waves. We also consider some preliminary calculations for spin-f > 1/2; more complete studies will be presented elsewhere.

1. Geometric effects of local decoherence for a fixed rate of squeezing

The geometry of the atom-laser system plays two distinct roles in determining the amount of achievable squeezing. First, $OD_{eff} \propto N_{eff}^{(2)}$ [Eq. (35)] is a purely geometrical quantity, derivable from the semiclassical model (see also [21]). The

 OD_{eff} sets the measurement strength, ξ , that characterizes the amount of light that is collectively scattered into the spatial mode of the probe. Second, because of the inhomogeneous intensity profile of the laser mode, the rate of diffuse photon scattering that causes local decoherence and ultimately caps the amount of generated squeezing varies across the cloud. Further complications arise from the fact that optical pumping both injects noise into the spin wave variance and causes a decay of the mean spin.

To gain physical insight, in this section we fix the OD_{eff} as we vary the geometry in order to isolate the effects of local decoherence as they relate specifically to the squeezing parameter [Eq. (52)]. For simulations, we choose the ensemble to be a cylindrically symmetric Gaussian cloud with average density

$$\eta(\mathbf{r}) = \eta_0 \exp\left(-2\frac{\rho^2}{\sigma_\perp^2} - 2\frac{z^2}{\sigma_z^2}\right),\tag{70}$$

where σ_{\perp}^2 and σ_z^2 are the transverse and longitudinal $1/e^2$ variances, respectively, η_0 is the peak density, and the total atom number is found by integrating over the cloud, $N = \int d^3 \mathbf{r} \, \eta(\mathbf{r})$. To characterize the geometry of the atomic distribution we use the *aspect ratio*, defined as AR $\equiv \sigma_z/\sigma_{\perp}$. A longitudinally extended, pencil-shaped cloud commonly employed in cold, dipole-trapped atomic ensemble experiments has an AR $\gg 1$; a pancake-shaped cloud that is much wider than it is long has an AR $\ll 1$. Here we vary η_0 as a function of the cloud geometry such that OD_{eff} $\propto N_{\text{eff}}^{(2)}/A$ remains constant.



FIG. 3. (Color online) Squeezing dynamics for a fixed $OD_{eff} =$ 50 and different atomic cloud geometries. The laser probe is a TEM_{00} mode with beam waist $w_0 = 20 \ \mu m$. (a) Peak squeezing denoted as the inverse of the squeezing parameter, ζ^{-1} in dB, as a function of aspect ratio of the cloud. The inset shows effective atom numbers as a function of aspect ratio; $N_{\rm eff}^{(2)}$ is constant by design. (b) Comparison of squeezing dynamics for clouds with AR = 0.1 (solid green line) and AR = 316 (dashed red line). The behavior in the absence of decoherence [Eq. (55)] (dotted black line) is plotted at the same OD_{eff}, showing agreement for short times. (c) Dynamics of the spin wave variance for the two clouds, normalized by dividing each by its initial variance, $N_{\rm eff}^{(2)}/4$. (d) Dynamics of the mean spin for the two clouds, normalized by dividing each by its initial mean spin, $N_{\rm eff}^{(1)}/2$. For fixed OD_{eff}, the superior squeezing of the pencil-shaped cloud over the pancake-shaped cloud is attributed to slower decay of the mean spin.

To find the peak squeezing, we perform numerical simulations by integrating the evolution of the collective mean spin and variance [Eqs. (59) and (65)] and then calculating the spin squeezing as a function of time. Figure 3 shows the resulting spin squeezing for different cloud geometries for a fixed beam waist, $w_0 = 20 \ \mu m$. The effective optical density is held constant, $OD_{eff} = 50$, which guarantees identical squeezing in the absence of decoherence for any geometry. Figure 3(a)shows the peak squeezing as a function of the AR. An increase in peak squeezing accompanies an increasing aspect ratio, indicating that decoherence is less detrimental to longitudinally extended clouds. The dynamics of the squeezing parameter are plotted in Fig. 3(b) for the opposing cases of a pancake-shaped cloud with AR = 0.1 and a pencil-shaped cloud with AR = 316. For comparison, the short-time solution Eq. (55) is shown, which describes the squeezing for either cloud in absence of decoherence.

To understand these results we separately examine the dynamical evolution of the projection-noise variance and the mean spin in Figs. 3(c) and 3(d), both of which contribute to the squeezing parameter. The effects of decoherence lead to

different steady-state values of the fundamental spin wave variance in Fig. 3(c) because the noise injection due to optical-pumping-induced spin flips, set by $N_{\rm eff}^{(3)}$, is slightly smaller for the pencil than for the pancake [see subplot in Fig. 3(a)]. More importantly, the decay rate of the mean spin is a strong function of the AR, as seen in Fig. 3(d). For a fixed OD_{eff}, under consideration here, different cloud geometries correspond to different $N_{\rm eff}^{(1)}$, which determines the mean spin of the ensemble addressed by the beam. The pencil geometry addresses a larger $N_{\rm eff}^{(1)}$ when compared to the pancake geometry, as seen in the subplot of Fig. 3(a). In addition, for the pencil geometry $N_{\rm eff}^{(1)}$ also decays more favorably. This occurs because for a fixed OD_{eff}, in the pencil geometry a large fraction of the atoms are spread far from the beam waist where rates of optical pumping are lower. For the pancake geometry, to achieve the same OD_{eff}, more of the atoms the we address are concentrated in the high-intensity region and more quickly depolarize.

2. Optimizing geometry for fixed atom number

We gain further insight into the nature of the atom-light interface by keeping the atom number N fixed and optimizing the cloud dimensions for peak squeezing. We fix the peak density at $\eta_0 = 5 \times 10^{11}$ cm⁻³, typical of dipole-trapped atoms, and keep the total atom number constant, $N = 9.8 \times 10^6$, for a fixed cloud volume. In Fig. 4(a), we plot contours of peak squeezing as a function of aspect ratio and beam waist. The optimal peak squeezing, $\zeta_{opt}^{-1} = 10.0$ dB, is found for AR = 256 at a beam waist of $w_0^{opt} = 31 \ \mu$ m. At the optimal geometry, the cloud length extends over several Rayleigh ranges, $\sigma_z/z_R^{opt} = 2.42$, and the transverse width of the cloud is slightly larger than the beam waist, $\sigma_\perp/w_0^{opt} = 1.09$.

To further understand the region of peak squeezing, in Fig. 4(b), we plot contours of OD_{eff} . Comparison of Figs. 4(a) and 4(b) shows that the optimal peak squeezing occurs in a parameter region where OD_{eff} is high, as expected. However, the optimal peak squeezing arises from a balance between high ODeff with low noise injection into the spin wave variance and low decay of the mean spin. Figure 4(c) shows the fraction of total atoms contributing to the mean spin, $N_{\rm eff}^{(1)}/N$, to the effective optical density, $N_{\rm eff}^{(2)}/N$, and to the noise injection $N_{\rm eff}^{(3)}/N$. As the cloud becomes too long and narrow, there does not exist a beam waist that can address a sufficiently large number of atoms while keeping a high OD_{eff}. Said another way, when the cloud becomes too long, the diffraction of scattered light is too large to effectively mode match with the probe field, as seen in Fig. 2(c). Similarly, too small a waist leaves many atoms outside the Rayleigh range and too large a waist increases the beam area, thus decreasing OD_{eff} , both manifestations of poor mode matching of the probe and the scattered field from the atom cloud.

3. Optimizing the beam waist for a fixed atomic cloud geometry

With a better understanding of how cloud geometry influences decoherence, we study the optimization of squeezing in a situation typical of experiments with dipole-trapped cold atoms, where both the trap dimensions and beam waist can be tuned while the peak atomic density η_0 remains fixed. In



FIG. 4. (Color online) Squeezing for different cloud geometries with Gaussian atomic density distribution [Eq. (70)] and a fixed total atom number $N = 9.84 \times 10^6$. (a) Contours of peak squeezing, ζ^{-1} in dB, as a function of cloud aspect ratio and laser probe beam waist. (b) Contours of the coupling strength, OD_{eff} . The difference between the optimal coupling strength and the resulting squeezing depends on the balance between coherent interactions and decoherence, characterized by different effective atom numbers, $N_{eff}^{(1)}$, $N_{eff}^{(2)}$, and $N_{eff}^{(3)}$, shown in (c).

this situation, the total atom number N depends on the trap volume.

For each cloud geometry there exists a beam waist that maximizes OD_{eff} . This is seen in Fig. 5(b) where contours of OD_{eff} are shown for a cloud with a fixed transverse width of $\sigma_{\perp} =$ 100 μ m as the cloud length σ_z and beam waist w_0 are varied. Contours for peak squeezing are shown in Fig. 5(a). Comparison with Fig. 5(b) demonstrates that for a given cloud geometry, the peak squeezing is achieved with a smaller beam waist



FIG. 5. (Color online) Squeezing for a fixed peak density $\eta_0 = 5 \times 10^{11}$ cm⁻³ and variable atom number that fills a dipole trap for cold atoms. In (a)–(c) the transverse cloud width is fixed at $\sigma_{\perp} = 100$ μ m and cloud length is taken to be variable. (a) Contours of peak squeezing, ζ^{-1} in dB. (b) Contours of OD_{eff}. (c) Optimal beam waist for maximizing OD_{eff} (upper, red dots) and for maximizing peak squeezing (lower, blue dots). For a given atomic geometry, the beam waist that optimizes the OD_{eff} is not the same as that which optimizes peak squeezing. (d) Peak squeezing as a function of cloud size for the optimal beam waist at each point.

than that which optimizes OD_{eff}. This is seen most clearly in Fig. 5(c), where we compare the optimal beam waist for maximizing ODeff to the beam waist that maximizes peak squeezing. Optimal squeezing occurs at smaller beam waists, where the region of the beam with greatest intensity, the Rayleigh range, is smaller. Because the scattering rate $\gamma_s(\mathbf{r})$ is proportional to the local intensity, atoms outside the Rayleigh range experience a decreased rate of optical pumping. Although a smaller Rayleigh range implies a decreased OD_{eff} and $N_{eff}^{(1)}$ as well, the reduction of the decoherence rate dominates in this regime. This is a direct analogy to Sec. III C 1, in which pencilshaped clouds with higher mean spins were more robust to decay due to a large number of atoms farther away from the beam waist. Finally, in Fig. 5(d) we plot contours of peak squeezing for different geometries at the optimal beam waist for each point. Since larger clouds contain more atoms and, in general, for properly chosen probe geometry OD_{eff} and peak squeezing increase with more atoms, there is no local maximum in Figs. 5(a), 5(b), and 5(d).

4. Relation to the symmetric one-dimensional model

Spin squeezing by QND measurement is traditionally modeled using a one-dimensional description of the atom-light interface where the ensemble is symmetrically coupled to plane waves with no spatial variations [51]. When accounting only for squeezing due to collective scattering and QND measurement, the full three-dimensional system can be effectively described by such a model, with the symmetric OD replaced with OD_{eff}. When decoherence from local diffuse scattering is included, however, such models become insufficient. In addition, a symmetric description does not account for the difference between the effective atom number contributing to the spin wave variance, $N_{\rm eff}^{(2)}$, that contributing to the mean spin, $N_{\rm eff}^{(1)}$, and that contributing to noise injection by spin flips, $N_{\rm eff}^{(3)}$.

To better understand the limit in which we recover the simple symmetric description, consider a symmetric onedimensional model where an ensemble of spin- $\frac{1}{2}$ atoms is coupled to a uniform plane wave and scatters collectively into this mode and locally into diffuse modes. In this case a single atom number suffices; every atom contributes equally to the optical density, to the mean spin, and to the injection of noise, $N_{\rm eff}^{(1)} = N_{\rm eff}^{(2)} = N_{\rm eff}^{(3)} = N$. The equation of motion for the spin wave variance Eq. (65) becomes [38,40,52,53]

$$\frac{d}{dt}\Delta F_z^2 = -\kappa \left(\Delta F_z^2\right)^2 - \frac{4\gamma_0}{9}\Delta F_z^2 + \frac{\gamma_0}{9}N,\qquad(71)$$

where γ_0 is the scattering rate and $\kappa = (\sigma_0/A)(4\gamma_0/9)$ is the measurement strength corresponding to the rate of scattering into the probe mode per atom [Eq. (23)]. This equation can be solved analytically, yielding

$$\Delta F_z^2(t) = \frac{N}{4} \frac{\sqrt{\text{OD} + 1} + \tanh\left[\sqrt{\text{OD} + 1}\frac{2}{9}\gamma_0 t\right]}{\sqrt{\text{OD} + 1} + \left(\frac{\text{OD}}{2} + 1\right)\tanh\left[\sqrt{\text{OD} + 1}\frac{2}{9}\gamma_0 t\right]}.$$
(72)

In the limit of short times, $\gamma_0 t \ll 1$, we recover the expression for QND squeezing in the absence of decoherence [Eq. (55)], $\Delta F_z^2(t) \approx (N/4)[1 + \xi(t)]^{-1}$, where $\xi(t) = \text{OD } \gamma_0 t/9$. In the opposite limit of long times, $\gamma_0 t \to \infty$, and large optical density, $\text{OD} \gg 1$, we find the expected scaling $\Delta F_z^2(t \to \infty) \propto \text{OD}^{-1/2}[51]$.

We can compare the symmetric one-dimensional model to a limiting case of the full three-dimensional model developed here. When the transverse extent of the cloud is much smaller than the beam waist and the longitudinal extent is well within the Rayleigh range, then spatial variations of the field across the cloud are minimal and $\beta_{00}(\mathbf{r}_i) \rightarrow 1$. Although this limiting case replicates the squeezing expected from the symmetric one-dimensional model, it is in fact far from a single-mode description. As discussed in Sec. II B, this geometry radiates paraxial light into many of transverse modes defined relative to the beam, and the associated spin waves couple together through diffuse scattering [Eq. (66)]. After numerically solving these coupled equations according to the procedure outlined above, we recover the same results Eq. (71), as is manifest from Eq. (65) in the limit that $\beta_{00}(\mathbf{r}_i) \rightarrow 1$ over the extent of the atom cloud.

We investigate this limit numerically in Fig. 6 for a spherical cloud ($\sigma_{\perp}, \sigma_z = 100 \,\mu$ m) probed by beams of increasing waist w_0 . In each case $OD_{eff} = 50$, such that in the absence of decoherence the different geometries would achieve identical squeezing. In Fig. 6(a) we see that as the beam waist is increased, the peak squeezing approaches that of the symmetric one-dimensional model. The inset shows the convergence of the effective atom numbers as the beam waist increases. Figure 6(b) shows the dynamics of the squeezing parameter $\zeta^{-1}(t)$ for the spherical cloud at both extremes in Fig. 6(a). For comparison, the squeezing parameter for the symmetric one-dimensional model is plotted both with and without decoherence. The difference between the models is substantial; the optimal peak squeezing for the symmetric one-dimensional model and full model are $\zeta_{\text{peak}}^{-1} = \{3.52 \text{ dB}, 4.99 \text{ dB}\}$, respectively. This difference can be understood in terms of the effective atom numbers. The advantage for spin squeezing in the three-dimensional model comes from the fact that $N_{\rm eff}^{(1)} \ge N_{\rm eff}^{(2)} \ge N_{\rm eff}^{(3)}$ due to different dependence on the spatial weightings $\beta_{00}(\mathbf{r})$, while for the symmetric one-dimensional case they are equal. For the three-dimensional model, not only can the effective number of atoms contributing to the noise



FIG. 6. (Color online) Comparison between the symmetric onedimensional model and the three-dimensional spin wave model. (a) Peak squeezing, ζ^{-1} in dB, for a spherical cloud with OD_{eff} = 50 as the beam waist is increased. The inset shows the convergence of $N_{\rm eff}^{(1)}$ (upper, blue line), $N_{\rm eff}^{(2)}$ (middle, green line), and $N_{\rm eff}^{(3)}$ (lower, red line) as w_0 increases. (b) Comparison of squeezing dynamics for the extremal waists from (a): the smallest, $w_0 = 10 \ \mu$ m, and the largest, $w_0 = 10^4 \ \mu$ m. For comparison, the symmetric one-dimensional case using Eq. (71) is plotted with decoherence (dotted black line) and without (dashed black line).

injection be smaller than that contributing to the OD_{eff} , but the effective number of atoms contributing to the mean spin, and thus the signal, is larger than both. Inspecting Fig. 3(d) we see an additional advantage for the three-dimensional model; when geometry is properly chosen, the mean spin decays at a much reduced rate.

5. Spin f > 1/2 atoms

We present here some initial results that illustrate the differences between spin- $\frac{1}{2}$ and larger spin ensembles. Consider an ensemble of ¹³³C̃ s atoms prepared in the f = 4ground hyperfine manifold. Figure 7(a) shows contours of peak squeezing as a function of aspect ratio using the atomic density given by Eq. (70). Note that the peak squeezing is substantially smaller than the peak squeezing for spin- $\frac{1}{2}$ [compare to Fig. 4(a)] and the optimal aspect ratio and beam waist are different for spin-4 alkali-metal atoms than for spin- $\frac{1}{2}$ atoms. This can be attributed to a reduction of the coupling strength ξ that is not compensated by an equal reduction in the decoherence rate. In principle, the coupling strength can be increased by internal spin control of the f = 4 hyperfine spin [40]. Figure 7(b) compares the squeezing dynamics for spin-4 and spin- $\frac{1}{2}$ ensembles at the same geometry and illustrates not only the disparity in peak squeezing but also in the time at which it occurs. These effects arise from the subtle interplay



FIG. 7. (Color online) Squeezing of a spin-4 ensemble. (a) Contours of peak squeezing for an ensemble of spin-4 atoms for the same fixed atom number as in Fig. 4 as function of beam waist and cloud aspect ratio. The largest peak squeezing, $\zeta^{-1} = 7.8$ dB, is achieved for the optimal geometry (indicated by a white "x"): AR = 300, $w_0 =$ 28 μ m. (b) Dynamics of ζ^{-1} in dB for the optimal geometry, for spin-4 and for spin- $\frac{1}{2}$.

between the rates of depolarization and injected noise due to spin flips when applied to the spatial modes of larger spin ensembles.

IV. SUMMARY AND OUTLOOK

The entangling power of the quantum interface between photons and an ensemble of cold atoms is at the heart of a variety of quantum information processing tasks. When considering extended atomic clouds in dipole traps, one must consider a full three-dimensional description of the electromagnetic modes and atomic density distribution in order to optimize this entangling power. Inhomogeneous coupling between atoms and photons is essential to maximize the strength of the quantum interface, but this comes with substantial complexity in the theoretical description. The model presented in this work addressed these issues and examined regimes of optimal performance.

We have studied the strength of the atom-photon interface in a traveling wave configuration in the context of spin squeezing via QND measurement. We developed a description in terms of quantized paraxial modes of the field in order to model the inhomogeneous atom-light coupling across the atomic ensemble, which leads to two distinct effects. First, the collective coupling describes a generalization of the Faraday interaction that entangles the quantized Stokes vector of the laser field with a spin wave defined by the weighted ensemble of atoms that indistinguishably radiates into the mode of the probe. The spin wave that is squeezed is defined by the probe mode we measure in a balanced polarimeter. Second, diffusely scattered photons lead to optical pumping and decoherence across the ensemble at a rate proportional to the local probe intensity. The delicate balance of these two effects favors certain geometries for spin squeezing.

We have numerically investigated the ultimate limits to spin squeezing based on a SME, including the effects of QND measurement backaction and decoherence by photon scattering into unmeasured modes. Unlike the usual onedimensional description in which the amount of squeezing is set by a single parameter, the optical density, we find that due to inhomogeneous coupling, multiple parameters are required. Of particular importance in a metrological context are the mean collective spin and the projection-noise variance, determined by effective atom numbers $N_{\text{eff}}^{(1)}$ and $N_{\text{eff}}^{(2)}$, respectively. Optimal geometries maximize the effective optical density, OD_{eff} , proportional to $N_{\rm eff}^{(2)}$, while minimizing the depolarization of $N_{\rm eff}^{(1)}$ and injected noise into the spin wave by optical pumping. We found that optimal mode matching occurs for geometries where a large number of atoms are addressed by a beam with a small transverse area, yielding a high OD_{eff}, but also where the depolarization rate due to optical pumping is relatively small. This geometry corresponds to a longitudinally extended, pencil-shaped cloud, with a probe beam chosen to optimize the trade-offs between OD_{eff} and decoherence. Such a geometry is far from the regime describing squeezing of a symmetric atomic spin ensemble, as is typically assumed. One recovers the symmetric description only for ensembles confined with extents much smaller than the beam waist and Rayleigh range, which yield much smaller OD_{eff}.

We have studied here the case of interactions in a highly transparent regime where the optical density is very small at the detuning of the probe. All collective effects studied here arise solely from the fact that the each of the atoms that scatter photons into the same paraxial mode are indistinguishable. For much higher densities, and/or lower detunings, multiple scattering effects are non-negligible. Under these conditions, superradiant scattering [54,55] and stimulated Raman scattering [28,56] become important and can lead to addition collective effects. Additionally, at high atomic densities, one must also consider the effects of atomic collisions [20]. Finally, we consider ultracold ensembles where the atoms are well approximated as fixed, point scatterers. Effects of finite temperature can be included by averaging over the positions of the atoms [27].

While the three-dimensional model developed in this work was specifically tailored to study the problem of spin squeezing by QND measurement, it can be extended to other protocols involving the quantum interface between photons and free-space atomic ensembles. Mode-matching and spatial effects are important for other spin squeezing protocols including the double-pass countertwisting interaction [53,57] or the recently proposed planar squeezing protocol [58]. Understanding spatial effects in order to identify regimes of strong coupling is also essential for quantum memories and repeaters in free-space atomic ensembles. In addition, recent work has considered ensembles of higher-spin alkali-metal atoms, in which control over the rich internal hyperfine

structure can enhance the entangling strength of the atomlight interface [40]. Quantifying the gains achievable though such control techniques requires a realistic description of the inhomogeneous interaction between light and atoms. Finally, a multimode description of the entangling Hamiltonian offers the possibility to exploit spatial modes and their associated spin waves as a resource. The creation of entanglement between spin waves could lead to novel states with potential application in continuous variable quantum computation and communication [59].

ACKNOWLEDGMENTS

We thank Robert Cook for helpful discussions and insights. This work was supported by NSF Grants No. PHY-0969391, No. 0969997, No. 1212445, No. 1306171, and No. 1307520 and AFOSR Grant No. Y600242.

APPENDIX A: PARAXIAL SCATTERING—CLASSICAL THEORY

The complex electric field is paraxial and quasimonochromatic, $\vec{\mathcal{E}}(\mathbf{r}_{\perp}, z, t)e^{i(k_0z-\omega_0t)}$. Thus, the slowly varying envelope is governed by the time-dependent paraxial wave equation [60],

$$i\left(\frac{\partial}{\partial z} + \frac{1}{c}\frac{\partial}{\partial t}\right)\vec{\mathcal{E}}(\mathbf{r}_{\perp}, z, t)$$

= $-\frac{1}{2k_0}\nabla_{\perp}^2\vec{\mathcal{E}}(\mathbf{r}_{\perp}, z, t) - 2\pi k_0 \overleftrightarrow{\chi}(\mathbf{r}_{\perp}, z) \cdot \vec{\mathcal{E}}(\mathbf{r}_{\perp}, z, t), \quad (A1)$

where $\vec{\chi}(\mathbf{r}_{\perp}, z)$ is the spatially averaged dielectric susceptibility. Defining $\vec{\mathcal{E}}(\mathbf{r}_{\perp}, z, t) = \mathcal{A}(t - z/c)\vec{\mathcal{U}}(\mathbf{r}_{\perp}, z)$,

$$i\frac{\partial}{\partial z}\vec{\mathcal{U}}(\mathbf{r}_{\perp},z) = -\frac{1}{2k_0}\nabla_{\perp}^2\vec{\mathcal{U}}(\mathbf{r}_{\perp},z) -2\pi k_0\overleftrightarrow{\chi}(\mathbf{r}_{\perp},z)\cdot\vec{\mathcal{U}}(\mathbf{r}_{\perp},z).$$
(A2)

This equation is isomorphic to the time-dependent Schrödinger equation with the propagation distance z playing the role of time and the susceptibility playing the role of the potential.

We can define a Hilbert space of square-integrable functions on the transverse plane, and use Dirac notation to express the evolution of the transverse mode as a function of z, $\mathcal{U}(\mathbf{r}_{\perp}, z) = \langle \mathbf{r}_{\perp} | \mathcal{U}(z) \rangle$. The free-space propagator (*z*-evolution operator), satisfies the free-particle Schrödinger equation in two dimensions,

$$i\frac{\partial}{\partial z}\hat{K} = \frac{\hat{\mathbf{p}}_{\perp}^2}{2k_0}\hat{K},\tag{A3}$$

where $\hat{\mathbf{p}}_{\perp} = -i \nabla_{\perp}$ in the position representation. The solution, $\hat{K}(z - z') = \exp[-i \frac{\hat{\mathbf{p}}_{\perp}^2}{2k_0}(z - z')]$, has the familiar position-space representation for the spreading of a wave packet and Fraunhofer diffraction [37],

$$K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z') = \langle \mathbf{r}_{\perp} | K(z - z') | \mathbf{r}'_{\perp} \rangle$$

= $\frac{-ik_0}{2\pi(z - z')} \exp\left[\frac{ik_0 |\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}|^2}{2(z - z')}\right].$
(A4)

The z evolution of a freely propagating beam is given by

$$\mathcal{U}(\mathbf{r}_{\perp}, z) = \langle \mathbf{r}_{\perp} | \mathcal{U}(z) \rangle = \langle \mathbf{r}_{\perp} | \hat{K}(z - z') | \mathcal{U}(z') \rangle$$
$$= \int d^2 \mathbf{r}'_{\perp} K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z') \mathcal{U}(\mathbf{r}'_{\perp}, z').$$
(A5)

Other properties of the propagator follow from unitarity, $\hat{K}^{\dagger}(z - z') = \hat{K}(z' - z)$, and thus

$$\mathcal{U}^{*}(\mathbf{r}_{\perp}',z') = \langle \mathbf{r}_{\perp}' | \hat{K}(z'-z) | \mathcal{U}(z) \rangle^{*} = \langle \mathcal{U}(z) | \hat{K}(z-z') | \mathbf{r}_{\perp}' \rangle$$
$$= \int d^{2} \mathbf{r}_{\perp} \mathcal{U}^{*}(\mathbf{r}_{\perp},z) K(\mathbf{r}_{\perp}-\mathbf{r}_{\perp}',z-z').$$
(A6)

We define a complete orthogonal basis $\{|u_{pl}(z)\rangle\}$, normalized to a fixed transverse area *A*,

$$\langle u_{p'l'}(z)|u_{pl}(z)\rangle = \int d^2 \mathbf{r}_{\perp} u_{p'l'}^*(z) u_{pl}(z) = A \,\delta_{p,p'} \delta_{l,l'}.$$
 (A7)

The propagator can then be written as

$$\hat{K}(z-z') = \sum_{p,l} |u_{pl}(z)\rangle \langle u_{pl}(z')| \Rightarrow K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z')$$
$$= \frac{1}{A} \sum_{p,l} u^*_{pl}(\mathbf{r}'_{\perp}, z') u_{pl}(\mathbf{r}_{\perp}, z),$$
(A8)

with the boundary condition $K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, 0) = \delta^{(2)}(\mathbf{r}'_{\perp} - \mathbf{r}_{\perp})$ that follows from completeness [Eq. (B11)].

The scattering of paraxial fields thus follows in complete analogy to the scattering of nonrealistic Schrödinger waves [37], where the time-dependent formulation of scattering translates into the *z* dependence. The retarded Green's function for free propagation is defined as

$$G_{+}(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z') = -i\Theta(z - z')K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z'),$$
$$\left(i\frac{\partial}{\partial z} + \frac{1}{2k_{0}}\nabla_{\perp}^{2}\right)G_{+}(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z') \qquad (A9)$$
$$= \delta(z - z')\delta^{(2)}(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}),$$

where $\Theta(z)$ is the Heaviside step function. In the first Born approximation, i.e., for dilute samples where multiple scattering is negligible, given an incident field (free propagating solution) $\vec{U}_{in}(\mathbf{r}_{\perp}, z)$, the total scattering solution is

$$\vec{\mathcal{U}}(\mathbf{r}_{\perp},z) = \vec{\mathcal{U}}_{in}(\mathbf{r}_{\perp},z) + \int_{-\infty}^{\infty} dz' \int d^{2}\mathbf{r}_{\perp}' G_{+}(\mathbf{r}_{\perp} - \mathbf{r}_{\perp}',z - z')$$

$$\times (-2\pi k_{0}) \overleftrightarrow{\chi}(\mathbf{r}_{\perp},z) \cdot \vec{\mathcal{U}}_{in}(\mathbf{r}_{\perp}',z')$$

$$= \vec{\mathcal{U}}_{in}(\mathbf{r}_{\perp},z) + i2\pi k_{0} \int_{-\infty}^{z} dz' \int d^{2}\mathbf{r}_{\perp}'$$

$$\times K(\mathbf{r}_{\perp} - \mathbf{r}_{\perp}',z - z') \overleftrightarrow{\chi}(\mathbf{r}_{\perp},z') \cdot \vec{\mathcal{U}}_{in}(\mathbf{r}_{\perp}',z'), \qquad (A10)$$

corresponding to the superposition of incident and scattered fields.

APPENDIX B: QUANTIZATION OF THE PARAXIAL FIELD

Paraxial quantization follows from the slowly varying envelope approximation [43]. For these modes, we define the

positive-frequency component of the electric field analogous to a classical beam,

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) = \sqrt{2\pi\hbar\omega_0} \sum_{\alpha} \mathbf{e}_{\alpha} \hat{\Psi}_{\alpha}(\mathbf{r}_{\perp},z,t) e^{i(k_0 z - \omega_0 t)}, \quad (B1)$$

where α labels transverse polarization, and the slowly varying envelope satisfies the equal-time commutation relations of a nonrelativistic bosonic field,

$$[\hat{\Psi}_{\alpha}(\mathbf{r}_{\perp},z,t),\hat{\Psi}_{\beta}^{\dagger}(\mathbf{r}_{\perp}',z',t)] = \delta_{\alpha,\beta}\delta^{(2)}(\mathbf{r}_{\perp}-\mathbf{r}_{\perp}')\delta(z-z').$$
(B2)

The free field satisfies the paraxial wave equation

$$i\frac{\partial}{\partial t}\hat{\Psi}_{\alpha} = -ic\frac{\partial}{\partial z}\hat{\Psi}_{\alpha} - \frac{1}{2k_0}\nabla_{\perp}^2\hat{\Psi}_{\alpha}, \qquad (B3)$$

which is the Heisenberg equation of motion for an envelope governed by the free paraxial Hamiltonian,

$$\hat{H}_{\text{free}} = \hbar \sum_{\alpha} \int d^3 \mathbf{r} \, \hat{\Psi}^{\dagger}_{\alpha} \left(-ic \frac{\partial}{\partial z} - \frac{1}{2k_0} \nabla_{\perp}^2 \right) \hat{\Psi}_{\alpha}. \quad (B4)$$

The free-field solution is thus determined by the classical propagator,

$$\hat{\Psi}_{\alpha}(\mathbf{r}_{\perp},z,t) = \int d^2 \mathbf{r}_{\perp}' K(\mathbf{r}_{\perp} - \mathbf{r}_{\perp}',ct) \hat{\Psi}_{\alpha}(\mathbf{r}_{\perp}',z-ct,0).$$
(B5)

It then follows that the free field satisfies the general commutation relations,

$$\begin{aligned} \left[\hat{\Psi}_{\alpha}(\mathbf{r}_{\perp},z,t), \hat{\Psi}_{\beta}^{\dagger}(\mathbf{r}_{\perp}',z',t') \right] \\ &= K(\mathbf{r}_{\perp} - \mathbf{r}_{\perp}',z-z') \delta_{\alpha,\beta} \delta(z-z'-c(t-t')), \end{aligned} \tag{B6}$$

and thus equal-z, unequal-t commutation relations,

$$[\hat{\Psi}_{\alpha}(\mathbf{r}_{\perp},z,t),\hat{\Psi}_{\beta}^{\dagger}(\mathbf{r}_{\perp}',z,t')] = \frac{1}{c}\delta_{\alpha,\beta}\delta^{(2)}(\mathbf{r}_{\perp}-\mathbf{r}_{\perp}')\delta(t-t').$$
(B7)

The paraxial field is naturally decomposed into an orthonormal set of dimensionless transverse mode functions. Here we use the Laguerre-Gauss modes $\{u_{pl}(\mathbf{r}_{\perp}, z)\}$ in cylindrical coordinates,

$$u_{pl}(\mathbf{r}_{\perp},z) = \mathcal{N}_{pl} \frac{w_0}{w(z)} \left[\frac{\sqrt{2}\rho}{w(z)} \right]^{|l|} L_p^{|l|} \left\{ \frac{2\rho^2}{[w(z)]^2} \right\}$$
$$\times \exp\left\{ -\frac{\rho^2}{[w(z)]^2} \right\} \exp\left[\frac{ik_0\rho^2}{2R(z)} \right]$$
$$\times \exp[-i(2p+l+1)\Phi(z) - il\phi], \quad (B8)$$

where $\mathcal{N}_{pl} = \sqrt{p!/(|l|+p)!}$ is the normalization constant, $L_p^{|l|}(x)$ indicates an associated Laguerre polynomial, and parameters w(z), R(z), and $\Phi(z)$ are given in Eq. (2). These

modes satisfy

$$\int d^2 \mathbf{r}_{\perp} u_{pl}^*(\mathbf{r}_{\perp}, z) u_{p'l'}(\mathbf{r}_{\perp}, z) = A \,\delta_{p, p'} \delta_{l, l'},\tag{B9}$$

$$\sum_{p,l} u_{pl}(\mathbf{r}_{\perp}, z) u_{pl}^*(\mathbf{r}_{\perp}', z) = A \,\delta^{(2)}(\mathbf{r}_{\perp} - \mathbf{r}_{\perp}'), \tag{B10}$$

$$\sum_{p,l} u_{pl}(\mathbf{r}_{\perp}, z) u_{pl}^*(\mathbf{r}'_{\perp}, z') = A K(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z'), \quad (B11)$$

where we have defined a quantization area, $A = \pi w_0^2/2$, as the natural scale for Gaussian beams of waist w_0 . Using Eq. (B5) and the completeness relation [Eq. (B10)], we define local mode creation and annihilation operators,

$$\hat{a}_{pl,\alpha}(z,t) = \int d^2 \mathbf{r}_{\perp} \sqrt{\frac{c}{A}} \hat{\Psi}_{\alpha}(\mathbf{r}_{\perp},z,t) u^*_{pl}(\mathbf{r}_{\perp},z), \qquad (B12)$$

that evolve under the free-field Hamiltonian according to $\hat{a}_{pl,\alpha}(z,t) = \hat{a}_{pl,\alpha}(0,t-z/c) = \hat{a}_{pl,\alpha}(z-ct,0)$, and satisfy free-field commutation relations

$$[\hat{a}_{pl,\alpha}(z,t), \hat{a}_{p'l',\beta}^{\top}(z',t')] = \delta_{p,p'} \delta_{l,l'} \delta_{\alpha,\beta} \,\delta(t-t'-(z-z')/c).$$
(B13)

The positive-frequency component of the electric field expanded in these modes is

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r},t) = \sum_{p,l,\alpha} \sqrt{\frac{2\pi \hbar \omega_0}{cA}} \, \mathbf{e}_{\alpha} \, \hat{a}_{pl,\alpha}(z,t) \, u_{pl}(\mathbf{r}_{\perp},z) \, e^{i(k_0 z - \omega_0 t)}.$$
(B14)

APPENDIX C: MULTIMODE MASTER EQUATION

The joint dynamics of the collective atomic system and the paraxial field can be expressed using a master equation formalism [44],

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar} \left[\hat{H}_{\rm eff} \hat{\rho} - \hat{\rho} \hat{H}_{\rm eff}^{\dagger} \right] + \Gamma \sum_{i,q} \hat{W}_q^{(i)} \hat{\rho} \hat{W}_q^{(i)\dagger}, \qquad (C1)$$

where the index q runs over all spherical basis elements. The effective Hamiltonian has a real part that drives coherent dynamics and an imaginary part describing loss,

$$\hat{H}_{\rm eff} = \hat{H}_{\rm int} + \hat{H}_{\rm loss}.$$
 (C2)

For a probe laser with polarization $\vec{\epsilon}_L$, the jump operators are [38]

$$\hat{W}_{q}^{(i)} = \sum_{f'} \frac{\Omega(\mathbf{r}_{i})/2}{\Delta_{f'} + i\Gamma/2} \mathbf{e}_{q}^{*} \cdot \hat{\mathbf{d}}_{f'}^{(i)} \hat{\mathbf{d}}_{f'}^{(i)\dagger} \cdot \vec{\epsilon}_{L}, \qquad (C3)$$

with local Rabi frequency is $\Omega(\mathbf{r}_i) = \langle J' || d || J \rangle 2 \mathcal{E}_L(\mathbf{r}_i) / \hbar$.

For our spin-squeezing protocol, the probe is prepared with linear polarization along x and photon flux \dot{N}_L . The Hermitian part of the effective Hamiltonian, \hat{H}_{int} , is given by the multimode Faraday Hamiltonian [Eq. (21)]. For detuning large compared to the excited-state hyperfine splitting and neglecting terms that describe vacuum-vacuum scattering, the anti-Hermitian part is

$$\hat{H}_{\text{loss}} = -i\hbar \frac{C^{(0)}}{2} \sum_{i} \gamma_s(\mathbf{r}_i) \hat{\mathbb{1}}^{(i)}.$$
 (C4)

Using Eqs. (C4), (12), and (38) we can write the part of master equation from diffuse scattering,

$$\frac{d\hat{\rho}}{dt}\bigg|_{\text{diff}} = \sum_{i} \gamma_{s}(\mathbf{r}_{i}) \big[(|C^{(0)}|^{2} - C^{(0)})\hat{\rho} \\ + |C^{(1)}|^{2} \big(\hat{f}_{z}^{(i)} \hat{\rho} \hat{f}_{z}^{(i)} + \hat{f}_{y}^{(i)} \hat{\rho} \hat{f}_{y}^{(i)} \big) \big].$$
(C5)

For the case of an atom driven on an $S_{1/2} \rightarrow P_J$ transition, $|C^{(0)}|^2 - C^{(0)} = -2/9$ and $|C^{(1)}|^2 = g_f^2/9$. To define the quantization axis, a large bias magnetic field B_0 is applied along the *z* axis, causing rapid Larmor precession at frequency $\Omega_0 = g_f \mu_B B_0$. Transforming to the rotating frame, $\hat{f}_z^{(i)} \rightarrow \hat{f}_z^{(i)}$ and $\hat{f}_y^{(i)} \rightarrow \cos(\Omega_0 t) \hat{f}_y^{(i)} - \sin(\Omega_0 t) \hat{f}_x^{(i)}$. Substituting these relations into Eq. (C5) and time averaging according to the rotating wave approximation yields the map for local decoherence in Eq. (39).

APPENDIX D: DERIVATION OF THE MULTIMODE HOMODYNE POLARIMETRY STOCHASTIC MASTER EQUATION

The SME describes the evolution of the atomic ensemble as continuous homodyne polarimetry measurements are performed on the output light. Although we showed in Sec. II A that only light in the probe mode is measured, we present here the SME that results from independent measurements of the position quadratures \hat{X}_{pl} in each mode pl, following the standard prescription given in Refs. [46,47]. A more general SME arises from such a continuous polarimetry measurement than the case where solely the \hat{X}_{00} quadrature is measured and measurement records on all modes $pl \neq 00$ are discarded.

Prior to measurement, the time evolution operator $\hat{U}(\Delta t) = \exp(-i\Delta t \hat{H}/\hbar)$ describing the interaction of the light and spin waves over a time interval Δt is

$$\hat{U}(\Delta t) = \prod_{p,l} \hat{U}_{pl}(\Delta t).$$
(D1)

The interaction in each spatial mode is generated by the multimode Faraday Hamiltonian in (21). Written in terms of the traveling-wave quadratures that arrive at the detector plane, $\hat{X}_{pl}(z_D,t) \equiv (\hat{a}_{pl,y}(z_D,t) + \hat{a}_{pl,y}^{\dagger}(z_D,t))/\sqrt{2}$ and $\hat{P}_{pl}(z_D,t) \equiv -i(\hat{a}_{pl,y}(z_D,t) - \hat{a}_{pl,y}^{\dagger}(z_D,t))/\sqrt{2}$, the unitary interaction in each mode is

$$\hat{U}_{pl}(\Delta t) = \exp\left[-i\Delta t \sqrt{\frac{\kappa}{2}} \left(\operatorname{Re}\left\{\hat{F}_{z}^{pl}\right\}\hat{P}_{pl}(z_{D},t) - \operatorname{Im}\left\{\hat{F}_{z}^{pl}\right\}\hat{X}_{pl}(z_{D},t)\right)\right].$$
(D2)

After this interaction the light and spin waves are entangled so that a polarimetry measurement of \hat{X}_{pl} is correlated with quantum backaction on the atomic ensemble. The evolution of the system conditioned on independent measurements of each mode is determined by the Kraus operator,

$$\hat{A}(\Delta t) = \prod_{p,l} \hat{A}_{pl}(\Delta t).$$
(D3)

Here, $\hat{A}_{pl}(\Delta t)$ is the Kraus component for measurement outcome x_{pl} in the spatial mode pl:

$$\hat{A}_{pl}(\Delta t) \equiv \langle \hat{X}_{pl}(z_D, t) = x_{pl} | \hat{U}_{pl}(\Delta t) | 0 \rangle$$

$$= \exp\left[\Delta t \sqrt{\frac{\kappa}{2}} \hat{F}_z^{pl} x_{pl} - \frac{\kappa \Delta t}{4} \left(\operatorname{Re}\left\{ \hat{F}_z^{pl} \right\}^2 + i \operatorname{Im}\left\{ \hat{F}_z^{pl} \right\} \operatorname{Re}\left\{ \hat{F}_z^{pl} \right\} \right) \right].$$
(D4)

The measurement photocurrent is a Gaussian stochastic process with mean $\langle \operatorname{Re}\{\hat{F}_{z}^{pl}\}\rangle\Delta t$ and variance proportional to $1/\kappa$ [46],

$$\Delta y_{pl} = \left\langle \operatorname{Re}\{\hat{F}_{z}^{pl}\} \right\rangle \Delta t + \frac{\Delta W_{pl}}{\sqrt{\kappa}}, \tag{D5}$$

where ΔW_{pl} is a Wiener increment with zero mean and variance Δt . In the infinitesimal limit, $\Delta t \rightarrow dt$ and $\Delta W_{pl} \rightarrow dW_{pl}$, and we expand the Kraus component to first order in dt,

$$\hat{A}_{pl}(dt) = \hat{\mathbb{1}} + \frac{\kappa}{4} \hat{F}_z^{pl} \langle \hat{F}_z^{pl} + \hat{F}_z^{pl\dagger} \rangle dt - \frac{\kappa}{8} \hat{F}_z^{pl\dagger} \hat{F}_z^{pl} dt + \sqrt{\frac{\kappa}{4}} \hat{F}_z^{pl} dW_{pl}.$$
(D6)

We have used the statistical independence of the stochastic Wiener processes, $dW_{pl}dW_{p'l'} = \delta_{p,p'}\delta_{l,l'}dt$.

After the measurements are performed, the conditional collective atomic state is updated via the map

$$\hat{\rho}(t+dt) = \frac{\hat{A}(dt)\hat{\rho}(t)\hat{A}^{\dagger}(dt)}{\operatorname{Tr}[\hat{A}^{\dagger}(dt)\hat{A}(dt)\hat{\rho}(t)]}.$$
(D7)

Using Eqs. (D3) and (D6) with Eq. (D7), we derive the conditional atomic state. In differential form, the SME is

$$d\hat{\rho} = \sum_{p,l} \left(\sqrt{\frac{\kappa}{4}} \mathcal{H}_{pl}[\hat{\rho}] \, dW_{pl} + \frac{\kappa}{4} \mathcal{L}_{pl}[\hat{\rho}] \, dt \right), \qquad (D8)$$

where the measurement-update superoperator $\mathcal{H}_{pl}[\hat{\rho}]$ is defined in Eq. (47) and the Lindblad superoperator $\mathcal{L}_{pl}[\hat{\rho}]$ in Eq. (48).

For measurements of the fundamental mode only, we ignore the measurement records for $pl \neq 00$, which is equivalent to averaging over measurement records or tracing over these modes. The result is the SME

$$d\hat{\rho} = \sqrt{\frac{\kappa}{4}} \mathcal{H}_{00}[\hat{\rho}] \, dW + \frac{\kappa}{4} \sum_{p,l} \mathcal{L}_{pl}[\hat{\rho}] \, dt, \qquad (D9)$$

where $dW = dW_{00}$. Since the modes $pl \neq 00$ are unmeasured, information about the ensemble is lost and Eq. (D9) does not preserve purity.

APPENDIX E: DERIVATION OF THE MEAN SPIN AND COVARIANCE EQUATIONS OF MOTION

While the squeezing parameter [Eq. (52)] depends solely upon the mean and variance of the fundamental spin wave defined by the spatial mode of the laser probe, the diffuse scattering by individual atoms is not collective in nature and acts to couple the different spin waves to one another. In order to model the dynamical evolution of the squeezing, including decoherence, we must track the evolution of a hierarchy of differential equations coupling the means and covariances of spin waves in all spatial modes. This appendix provides a detailed derivation of these equations and the numerical methods used in their solution for the case of an ensemble of spin- $\frac{1}{2}$ atoms.

We first consider the evolution of $\langle \hat{F}_x^{pl} \rangle$, the mean of a spin wave in spatial mode pl, where $\hat{F}_x^{pl} = \sum_i \beta_{pl}(\mathbf{r}_i) \hat{f}_x^{(i)}$. Because collective scattering and measurement backaction negligibly affect the dynamics of the mean spin, to good approximation the evolution of $\langle \hat{F}_x^{pl} \rangle$ is dominated by diffuse scattering and is described by Eq. (41). For spin- $\frac{1}{2}$, the local map [Eq. (39)] simplifies to $\mathcal{D}_i[\hat{f}_x^{(i)}] = -\hat{f}_x^{(i)}/3$. Using Eq. (38), we get an equation of motion,

$$\frac{d}{dt}\langle \hat{F}_{x}^{pl}\rangle = -\frac{\gamma_{0}}{3}\sum_{i}\beta_{00}(\mathbf{r}_{i})\beta_{pl}(\mathbf{r}_{i})\langle \hat{f}_{x}^{(i)}\rangle.$$
 (E1)

By decomposing $\beta_{00}(\mathbf{r})\beta_{pl}(\mathbf{r})$ in terms of orthogonal mode functions, the right-hand side of Eq. (E1) can be expressed in terms of spin wave operators. In terms of the mode functions,

$$\beta_{00}(\mathbf{r}_{\perp},z)\beta_{pl}(\mathbf{r}_{\perp},z) = |u_{00}(\mathbf{r}_{\perp},z)|^2 u_{pl}^*(\mathbf{r}_{\perp},z)u_{00}(\mathbf{r}_{\perp},z)$$
$$= \sum_{p',l'} c_{p'l'}^{pl}(z)\beta_{p'l'}(\mathbf{r}_{\perp},z),$$
(E2)

where we have made use of orthogonality and completeness in Eqs. (B9) and (B10) to define projection coefficients,

$$c_{p'l'}^{pl}(z) \equiv \frac{1}{A} \int d^2 \mathbf{r}_{\perp} \left[u_{00}(\mathbf{r}_{\perp}, z) \right]^2 u_{pl}^*(\mathbf{r}_{\perp}, z) u_{p'l'}(\mathbf{r}_{\perp}, z).$$
(E3)

By restricting Eq. (E1) to a coarse-grained slice of thickness δz at longitudinal coordinate z_k and performing the projection in Eq. (E2), we obtain an infinite hierarchy of differential equations that couple mean spin waves in a given slice to one another,

$$\frac{d}{dt} \langle \hat{F}_x^{pl}(z_k) \rangle = -\frac{\gamma_0}{3} \sum_{p',l'} c_{p'l'}^{pl}(z_k) \langle \hat{F}_x^{p'l'}(z_k) \rangle.$$
(E4)

Solving the resulting system of coupled differential equations requires initial conditions of the mean spin waves in each slice. Using $\langle \hat{f}_x(t_0) \rangle = 1/2$ for the initial SCS state of the ensemble,

$$\left\langle \hat{F}_{x}^{pl}(z_{k},t_{0})\right\rangle = \sum_{i_{k}}\beta_{pl}\left(\mathbf{r}_{i_{k}}\right)\left\langle \hat{f}_{x}^{(i_{k})}(t_{0})\right\rangle = \frac{1}{2}\sum_{i_{k}}\beta_{pl}\left(\mathbf{r}_{i_{k}}\right), \quad (E5)$$

where i_k is an index over all atoms in slice z_k . For a average atomic density, $\eta(\mathbf{r})$, the sum becomes an integral,

$$\langle \hat{F}_x^{pl}(z_k, t_0) \rangle \rightarrow \frac{\delta z}{2} \int d^2 \mathbf{r}_\perp \eta(\mathbf{r}_\perp, z_k) \beta_{pl}(\mathbf{r}_\perp, z_k).$$
 (E6)

An approximate solution to Eq. (E4) is found for each slice by choosing δz and truncating the number of spin waves at some index p_{max} , l_{max} . Summing over the solutions at each slice gives the mean of the fundamental spin wave,

$$\langle \hat{F}_{x}^{00}(t) \rangle = \sum_{k} \langle \hat{F}_{x}^{00}(z_{k},t) \rangle.$$
 (E7)

Equation (E7) is the mean spin in the definition of the squeezing parameter.

To solve for the variance of the fundamental spin wave, we follow a similar procedure. As shown in Eq. (65), the fundamental variance couples through diffuse scattering to covariances between spin waves in slices z_k and $z_{k'}$:

$$\begin{split} \left\langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \right\rangle \\ &= \left\langle \hat{F}_{z}^{pl}(z_{k}) \hat{F}_{z}^{p'l'}(z_{k'}) \right\rangle - \left\langle \hat{F}_{z}^{pl}(z_{k}) \right\rangle \left\langle \hat{F}_{z}^{p'l'}(z_{k'}) \right\rangle. \end{split} \tag{E8}$$

From the SME in Eq. (49), we find the equations of motion for these covariances. Unlike the mean spin, the effects of continuous measurement must be included along with diffuse scattering. However, decoherence from collective scattering, described by the map \mathcal{L}_{pl} in Eq. (48), does not affect these covariances since the \hat{F}_z^{pl} commute with one another.

First, we examine the contribution of continuous measurement. From the SME in Eq. (46) and the rule of Itō calculus that differentials must be taken to second order [46], i.e., d(XY) = (dX)Y + X(dY) + (dX)(dY), we find

$$\begin{aligned} d \langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \rangle \Big|_{\text{meas}} \\ &= \sqrt{\frac{\kappa}{4}} \{ \langle \mathcal{H}_{00} [\hat{F}_{z}^{pl}(z_{k}) \hat{F}_{z}^{p'l'}(z_{k'})] \rangle - \langle \mathcal{H}_{00} [\hat{F}_{z}^{pl}(z_{k})] \rangle \langle \hat{F}_{z}^{p'l'}(z_{k'}) \rangle \\ &- \langle \hat{F}_{z}^{pl}(z_{k}) \rangle \langle \mathcal{H}_{00} [\hat{F}_{z}^{p'l'}(z_{k'})] \rangle \} dW \\ &- \frac{\kappa}{4} \langle \mathcal{H}_{00} [\hat{F}_{z}^{pl}(z_{k})] \rangle \langle \mathcal{H}_{00} [\hat{F}_{z}^{p'l'}(z_{k'})] \rangle dt. \end{aligned}$$
(E9)

The map \mathcal{H}_{00} , Eq. (47), couples the first- and second-order moments of the spin waves to higher-order moments. For the initial SCS along x and during its subsequent evolution, the spin waves \hat{F}_z^{pl} are Gaussian distributed, both over the entire cloud and within each coarse-grained slice z_k . Thus, third-order moments of commuting observables can be expressed in terms of first- and second-order moments with the relation $\langle XYZ \rangle =$ $\langle XY \rangle \langle Z \rangle + \langle XZ \rangle \langle Y \rangle + \langle YZ \rangle \langle X \rangle - 2 \langle X \rangle \langle Y \rangle \langle Z \rangle$ [46]. In this regime, all stochastic terms in Eq. (E9) cancel, leaving the deterministic equation:

$$\frac{d}{dt} \left\langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \right\rangle \Big|_{\text{meas}} = -\kappa \left\langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{00} \right\rangle \left\langle \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \Delta \hat{F}_{z}^{00} \right\rangle
= -\kappa \sum_{k'',k'''} \left\langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{00}(z_{k''}) \right\rangle \left\langle \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \Delta \hat{F}_{z}^{00}(z_{k'''}) \right\rangle.$$
(E10)

These dynamics, which arise from continuous polarimetry measurements, serve to generate the correlations that produce spin squeezing. Note that if we take l, l', p, p' = 0 and sum over all k and k', we recover the familiar case derived in Ref. [46].

We now turn our attention to diffuse scattering. The evolution of the first-order terms in the covariance [Eq. (E8)] is

$$\left. \frac{d}{dt} \left\langle \hat{F}_z^{pl}(z_k) \right\rangle \right|_{\text{diff}} = -\frac{2\gamma_0}{9} \sum_{p',l'} c_{p'l'}^{pl}(z_k) \left\langle \hat{F}_z^{p'l'}(z_{k'}) \right\rangle, \quad \text{(E11)}$$

where we have used the fact that $\mathcal{D}_i[\hat{f}_z^{(i)}] = -2\hat{f}_z^{(i)}/9$. The evolution of the second-order term in Eq. (E8) is governed

entirely by atomic pairwise correlations,

$$\frac{d}{dt} \left\langle \hat{F}_{z}^{pl}(z_{k}) \hat{F}_{z}^{p'l'}(z_{k'}) \right\rangle \Big|_{\text{diff}} = \sum_{i_{k} \neq j_{k'}} \beta_{pl}(\mathbf{r}_{i_{k}}) \beta_{p'l'}(\mathbf{r}_{j_{k'}}) \frac{d}{dt} \left\langle \hat{f}_{z}^{(i_{k})} \hat{f}_{z}^{(j_{k'})} \right\rangle \Big|_{\text{diff}}, \quad (E12)$$

where i_k and $j_{k'}$ label atoms in coarse-grained slices at z_k and $z_{k'}$, respectively. Using Eq. (44), the pairwise correlations evolve according to

$$\frac{d}{dt} \langle \hat{f}_{z}^{(i_{k})} \hat{f}_{z}^{(j_{k'})} \rangle \bigg|_{\text{diff}} = -\frac{2\gamma_{0}}{9} \big[\beta_{00}(\mathbf{r}_{i_{k}}) + \beta_{00}(\mathbf{r}_{j_{k'}}) \big] \langle \hat{f}_{z}^{(i_{k})} \hat{f}_{z}^{(j_{k'})} \rangle.$$
(E13)

Substituting this into Eq. (E12) yields

$$\frac{d}{dt} \langle \hat{F}_{z}^{pl}(z_{k}) \hat{F}_{z}^{p'l'}(z_{k'}) \rangle \Big|_{\text{diff}} = -\frac{2\gamma_{0}}{9} \sum_{i_{k} \neq j_{k'}} \beta_{pl}(\mathbf{r}_{i_{k}}) \beta_{p'l'}(\mathbf{r}_{j_{k'}}) [\beta_{00}(\mathbf{r}_{i_{k}}) + \beta_{00}(\mathbf{r}_{j_{k'}})] \langle \hat{f}_{z}^{(i_{k})} \hat{f}_{z}^{(j_{k'})} \rangle
= -\frac{2\gamma_{0}}{9} \left[\sum_{i_{k}} \beta_{00}(\mathbf{r}_{i_{k}}) \beta_{pl}(\mathbf{r}_{i_{k}}) \langle \hat{f}_{z}^{(i_{k})} \hat{F}_{z}^{p'l'}(z_{k'}) \rangle + \sum_{i_{k'}} \beta_{00}(\mathbf{r}_{i_{k'}}) \beta_{pl}(\mathbf{r}_{i_{k'}}) \langle \hat{F}_{z}^{pl}(z_{k}) \hat{f}_{z}^{(i_{k'})} \rangle \right]
+ \delta_{k,k'} \frac{\gamma_{0}}{9} \sum_{i_{k}} \beta_{00}(\mathbf{r}_{i_{k}}) \beta_{pl}(\mathbf{r}_{i_{k}}) \beta_{p'l'}(\mathbf{r}_{i_{k}}), \qquad (E14)$$

where the term in the last line comes from adding and subtracting the $i_k = j_k$ terms, which allowed us to perform the sum over one of the atom indices in the previous line. The sum in the final term can be expressed as an integral over the density of the atomic cloud,

$$N_{p'l'}^{pl}(z_k) = \delta z \int d^2 \mathbf{r}_{\perp} \eta(\mathbf{r}_{\perp}, z_k) \beta_{00}(\mathbf{r}_{\perp}, z_k) \beta_{pl}(\mathbf{r}_{\perp}, z_k) \beta_{p'l'}(\mathbf{r}_{\perp}, z_k).$$
(E15)

Note that for the fundamental mode, p, p', l, l' = 0, $N_{p'l'}^{pl}(z_k)$ is $N_{\text{eff}}^{(3)}$ at slice z_k . Combining Eqs. (E11) and (E14) and projecting the covariances into the spin waves using Eq. (E3), we arrive at a differential equation that couples spin wave covariances. Including the dynamics due to continuous measurement Eq. (E10), the full equation of motion for the covariances is

$$\frac{d}{dt} \langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \rangle = -\kappa \sum_{k'',k'''} \langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{00}(z_{k''}) \rangle \langle \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \Delta \hat{F}_{z}^{00}(z_{k'''}) \rangle
- \frac{2\gamma_{0}}{9} \sum_{p''l''} \left[c_{p''l''}^{pl}(z_{k}) \langle \Delta \hat{F}_{z}^{p''l''}(z_{k}) \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \rangle + c_{p''l''}^{p'l'}(z_{k'}) \langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{p''l''}(z_{k'}) \rangle \right] + \frac{\gamma_{0}}{9} N_{p'l'}^{pl}(z_{k}) \delta_{k,k'}.$$
(E16)

As in the case of the mean spin waves, this set of equations is solved by truncating Eq. (E16) at some p_{max} and l_{max} . Following Eq. (E12), using $\langle \hat{f}_z^{(i_k)} \hat{f}_z^{(j_{k'})} \rangle_{i_k \neq j_{k'}} = 0$ for the initial SCS, the initial spin wave covariances are

$$\left\langle \Delta \hat{F}_{z}^{pl}(z_{k}) \Delta \hat{F}_{z}^{p'l'}(z_{k'}) \right\rangle(t_{0}) = \delta_{k,k'} \frac{\delta z}{4} \int d^{2} \mathbf{r}_{\perp} \, \eta(\mathbf{r}_{\perp}, z_{k}) \beta_{pl}(\mathbf{r}_{\perp}, z_{k}) \beta_{p'l'}(\mathbf{r}_{\perp}, z_{k}). \tag{E17}$$

With these initial conditions and the equations of motion, we can solve for the evolution of all covariances in the presence of both QND measurement backaction and decoherence by optical pumping. Summing the covariances in the fundamental spin wave over all slices yields the variance in the fundamental spin wave:

$$\left[\Delta F_{z}^{00}(t)\right]^{2} = \sum_{k,k'} \left\langle \Delta \hat{F}_{z}^{00}(z_{k}) \Delta \hat{F}_{z}^{00}(z_{k'}) \right\rangle (t).$$
(E18)

From $[\Delta F_z^{00}(t)]^2$ and $\langle \hat{F}_x^{00}(t) \rangle$, we calculate the dynamics of the squeezing parameter [Eq. (52)].

For spin-f alkali-metal atoms, the derivation of the mean spin equations [Eq. (67)] and fundamental spin wave variance equation [Eq. (69)] follows a similar prescription as for spin- $\frac{1}{2}$, but with more general processes that include transfer of coherences, spontaneous emission, and hyperfine optical pumping. This makes the dynamics substantially more complex and will be treated in detail in a future publication.

- [1] M. Fleischhauer and M. D. Lukin, Phys. Rev. A 65, 022314 (2002).
- [2] B. Julsgaard, J. Sherson, J. I. Cirac, J. Fiurasek, and E. S. Polzik, Nature (London) 432, 482 (2004).
- [3] K. S. Choi, H. Deng, J. Laurat, and H. J. Kimble, Nature (London) 452, 67 (2008).
- [4] L. M. Duan, M. D. Lukin, J. Cirac, and P. Zoller, Nature (London) 414, 413 (2001).

- [5] D. N. Matsukevich and A. Kuzmich, Science 306, 663 (2004).
- [6] S. L. Braunstein and P. van Loock, Rev. Mod. Phys. 77, 513 (2005).
- [7] J. Appel, P. J. Windpassinger, D. Oblak, U. B. Hoff, N. Kjaergaard, and E. S. Polzik, Proc. Natl. Acad. Sci. USA 106, 10960 (2009).
- [8] I. D. Leroux, M. H. Schleier-Smith, and V. Vuletic, Phys. Rev. Lett. 104, 073602 (2010).
- [9] R. Miller, K. Northup, T. E. Birnbaum, A. Boca, A. Boozer, and H. Kimble, J. Phys. B: At. Mol. Opt. Phys. 38, S551 (2005).
- [10] Z. Chen, J. G. Bohnet, S. R. Sankar, J. Dai, and J. K. Thompson, Phys. Rev. Lett. **106**, 133601 (2011).
- [11] E. Vetsch, D. Reitz, G. Sagué, R. Schmidt, S. T. Dawkins, and A. Rauschenbeutel, Phys. Rev. Lett. 104, 203603 (2010).
- [12] R. Bose, D. Sridharan, H. Kim, G. S. Solomon, and E. Waks, Phys. Rev. Lett. **108**, 227402 (2012).
- [13] C.-L. Hung, S. Meenehan, D. Chang, O. Painter, and H. Kimble, New J. Phys. 15, 083026 (2013).
- [14] H. Tanji-Suzuki, I. D. Leroux, M. H. Schleier-Smith, M. Cetina, A. T. Grier, J. Simon, and V. Vuletić, in *Advances in Atomic, Molecular, and Optical Physics*, edited by P. B. E. Arimondo and C. Lin (Academic Press, San Diego, 2011), Vol. 60, pp. 201–237.
- [15] T. Bienaimé, R. Bachelard, N. Piovella, and R. Kaiser, Fortschr. Phys. 61, 377 (2013).
- [16] A. Kuzmich, L. Mandel, J. Janis, Y. E. Young, R. Ejnisman, and N. P. Bigelow, Phys. Rev. A 60, 2346 (1999).
- [17] B. Julsgaard, A. Kozhekin, and E. S. Polzik, Nature (London) 413, 400 (2001).
- [18] M. Kubasik, M. Koschorreck, M. Napolitano, S. R. de Echaniz, H. Crepaz, J. Eschner, E. S. Polzik, and M. W. Mitchell, Phys. Rev. A 79, 043815 (2009).
- [19] M. Koschorreck, M. Napolitano, B. Dubost, and M. W. Mitchell, Phys. Rev. Lett. 105, 093602 (2010).
- [20] F. Kaminski, N. S. Kampel, M. P. H. Steenstrup, A. Griesmaier, E. S. Polzik, and J. H. Müller, Eur. Phys. J. D 66, 227 (2012).
- [21] J. H. Müller, P. Petrov, D. Oblak, C. L. Garrido Alzar, S. R. de Echaniz, and E. S. Polzik, Phys. Rev. A 71, 033803 (2005).
- [22] M. W. Sørensen and A. S. Sørensen, Phys. Rev. A 77, 013826 (2008).
- [23] A. Kuzmich and T. A. B. Kennedy, Phys. Rev. Lett. 92, 030407 (2004).
- [24] P. J. Windpassinger, D. Oblak, U. B. Hoff, J. Appel, N. Kjærgaard, and E. S. Polzik, New J. Phys. 10, 053032 (2008).
- [25] M. Koschorreck and M. W. Mitchell, J. Phys. B 42, 195502 (2009).
- [26] J. D. Sau, S. R. Leslie, M. L. Cohen, and D. M. Stamper-Kurn, New J. Phys. 12, 085011 (2010).
- [27] L. M. Duan, J. I. Cirac, and P. Zoller, Phys. Rev. A 66, 023818 (2002).
- [28] M. W. Sørensen and A. S. Sørensen, Phys. Rev. A 80, 033804 (2009).
- [29] E. Zeuthen, A. Grodecka-Grad, and A. S. Sørensen, Phys. Rev. A 84, 043838 (2011).
- [30] A. Grodecka-Grad, E. Zeuthen, and A. S. Sørensen, Phys. Rev. Lett. 109, 133601 (2012).

- [31] D. B. Higginbottom, B. M. Sparkes, M. Rancic, O. Pinel, M. Hosseini, P. K. Lam, and B. C. Buchler, Phys. Rev. A 86, 023801 (2012).
- [32] A. Kuzmich, L. Mandel, and N. P. Bigelow, Phys. Rev. Lett. 85, 1594 (2000).
- [33] M. Koschorreck, M. Napolitano, B. Dubost, and M. W. Mitchell, Phys. Rev. Lett. **104**, 093602 (2010).
- [34] T. Takano, M. Fuyama, R. Namiki, and Y. Takahashi, Phys. Rev. Lett. 102, 033601 (2009).
- [35] D. V. Kupriyanov, O. S. Mishina, I. M. Sokolov, B. Julsgaard, and E. S. Polzik, Phys. Rev. A 71, 032348 (2005).
- [36] D. Vasilyev, K. Hammerer, N. Korolev, and A. S. Sørensen, J. Phys. B 45, 124007 (2012).
- [37] R. Newton, Scattering Theory of Waves and Particles (Dover, Mineola, NY, 1982).
- [38] I. H. Deutsch and P. S. Jessen, Opt. Comm. **283**, 681 (2010).
- [39] G. A. Smith, S. Chaudhury, and P. S. Jessen, J. Opt. B 5, 323 (2003).
- [40] L. M. Norris, C. M. Trail, P. S. Jessen, and I. H. Deutsch, Phys. Rev. Lett. 109, 173603 (2012).
- [41] By employing a two-color probe on the *D*1 and *D*2 resonance lines for alkali-metal atoms, one can cancel the residual nonlinear spin dynamics associated with the tensor light shift, while retaining the Faraday effect that is used in the QND measurement of the collective spin. This technique is currently under investigation and will be reported in a future publication.
- [42] K. Hammerer, Ph.D. thesis, Technische Universität München, 2006.
- [43] I. H. Deutsch and J. C. Garrison, Phys. Rev. A 43, 2498 (1991).
- [44] C. Gardiner and P. Zoller, *Quantum Noise* (Springer, Berlin, 2004).
- [45] C. Cohen-Tannoudji, in *Frontiers in Laser Spectroscopy*, edited by R. Balian, S. Haroche, and S. Liberman, Les Houches, Session XXVII, 1975 (North-Holland, Amsterdam, 1977), pp. 1–104.
- [46] K. Jacobs and D. A. Steck, Contemp. Phys. 47, 279 (2006).
- [47] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, UK, 2010).
- [48] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen, Phys. Rev. A 46, R6797 (1992).
- [49] D. Budker and M. Romalis, Nat. Phys. 3, 227 (2007).
- [50] R. J. Sewell, M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, and M. W. Mitchell, Phys. Rev. Lett. 109, 253605 (2012).
- [51] K. Hammerer, A. S. Sørensen, and E. S. Polzik, Rev. Mod. Phys. 82, 1041 (2010).
- [52] L. B. Madsen and K. Mølmer, Phys. Rev. A 70, 052324 (2004).
- [53] C. M. Trail, P. S. Jessen, and I. H. Deutsch, Phys. Rev. Lett. 105, 193602 (2010).
- [54] S. Inouye, A. P. Chikkatur, D. M. Stamper-Kurn, J. Stenger, D. E. Pritchard, and W. Ketterle, Science 285, 571 (1999), http://www.sciencemag.org/content/285/5427/571.full.pdf.
- [55] A. Hilliard, F. Kaminski, R. le Targat, C. Olausson, E. S. Polzik, and J. H. Müller, Phys. Rev. A 78, 051403 (2008).
- [56] M. G. Raymer and J. Mostowski, Phys. Rev. A 24, 1980 (1981).

- [57] M. Takeuchi, S. Ichihara, T. Takano, M. Kumakura, T. Yabuzaki, and Y. Takahashi, Phys. Rev. Lett. 94, 023003 (2005).
- [58] G. Puentes, G. Colangelo, R. J. Sewell, and M. W. Mitchell, New J. Phys. 15 103031 (2013).
- [59] N. Cerf, G. Leuchs, and E. Polzik (eds.), *Quantum Information with Continuous Variables of Atoms and Light* (World Scientific, Singapore, 2007).
- [60] J. C. Garrison and R. Y. Chiao, *Quantum Optics* (Oxford University Press, Oxford, UK, 2008).