I. INTRODUCTION

Two-mode spin squeezed (TMSS) states are important to consider for a number of reasons. First, they may be used to demonstrate entanglement experimentally using only spin measurements \(1\). Second, as they are entangled, they may be used as a resource for quantum information, for example, quantum teleportation \(2, 3\), quantum computing, superdense coding, etc. \(4\). In addition, TMSS states are analogous to two-mode squeezed states for light \(5\), which have proven to be useful for quantum information purposes \(6\).

There are several approaches to produce TMSS states. Recently they have been produced experimentally for samples of atoms by quantum nondemolition (QND) measurements of the spin components \(1\). In Ref. \(1\) it was proposed to produce TMSS states of light by combining the two Einstein-Podolsky-Rosen (EPR) output modes of an optical parametric oscillator with coherent light at polarizing beam splitters.

It is also possible to adapt proposals for generating single-mode spin squeezing to the two-mode case. For example, it has been proposed that single-mode spin squeezed states can be produced via an adiabatic approach \(6\). In this approach, one starts with a Hamiltonian for which the ground state is easily achieved, and slowly varies the Hamiltonian towards one that has as its ground state the optimal spin squeezed state. In principle, this procedure can be adapted to obtain optimal TMSS states. A drawback to this method is that, in practice, the appropriate Hamiltonian may be difficult to realize.

Another technique for generating TMSS states that may be generalized from the single-mode case is feedback. Feedback has proven to be a powerful technique for state preparation \(3, 4\) and measurement \(10, 11\). It has been shown that it is possible to produce near-optimal spin squeezing in the single-mode case by continual measurement of the spin operator \(J_z\), and using feedback to bring \(\langle J_z \rangle\) close to zero \(12\). Here we adapt this method to the QND measurement approach discussed in Ref. \(1\).

The method for production of TMSS states in Ref. \(1\) suffers the drawback that, although the variances in the sums of the spin components are small, their means are not close to zero. As is shown in Ref. \(3\), the optimal TMSS states have zero means. We would therefore wish to obtain states with zero means in order for them to be as close as possible to the optimal states. This is analogous to the problem with the production of single-mode spin squeezing as considered by Thomsen, Mancini, and Wiseman (TMW) \(5\).

II. TWO-MODE SPIN SQUEEZING

In Ref. \(1\), TMSS states (conditioned on the measurement record) are obtained via QND measurements of the spin of the atomic samples. In order to perform this QND measurement, an off-resonant pulse is transmitted through two cesium gas samples. The interaction Hamiltonian may be expressed as \(2, 12\)

\[
H_{int} = \hbar a \int S_z(z, t) J_z(z, t) dz,
\]  

(1)

where \(a\) is a coupling constant, \(S_z(z, t)\) is the \(z\) component of the instantaneous Stokes vector for the light, and \(J_z(z, t)\) is the \(z\) component of the continuous spin operator. These variables are explained further in Appendix \(A\).

In the Schrödinger picture \(J_z(z, t)\) is independent of \(t\), and the total Stokes vector \(S(z)\) is independent of \(z\). The total time evolution operator is, therefore,

\[
U = \exp \left( -\frac{i}{\hbar} \int H_{int} dt \right) \\
= \exp \left( -ia \int S_z(z, t) J_z(z) dz dt \right) \\
= \exp \left( -ia \int S_z J_z(z) dz \right) \\
= \exp(-ia S_z J_z). 
\]  

(2)

In Ref. \(1\) the light is initially linearly polarized, so the light is in the maximally weighted \(S_z\) eigenstate. In addition, the total spin \(s\) is large, so we may apply a
SU(2)→HW(2) contraction. That is, we make the substitutions \[ S_x \rightarrow s - b b^\dagger, \quad (3) \]
\[ S_y \rightarrow \sqrt{\frac{a}{2}} (b + b^\dagger), \quad (4) \]
\[ S_z \rightarrow \frac{i}{2} \sqrt{\frac{a}{2}} (b - b^\dagger), \quad (5) \]
where \( b b^\dagger \) is the number operator for photons subtracted from the \( x \) linear polarization mode (and added to the \( y \) linear polarization mode). We will assume that the probe beam is in a coherent spin state oriented in the \( x \) direction (i.e., \( x \) linearly polarized), and take the limit \( s \rightarrow \infty \) so that this contraction is exact.

Under this contraction, \( [S_y, S_z] \rightarrow is \). This means that, under the Heisenberg picture, the operator \( S_y \) transforms as
\[ S_y^{\text{out}} = S_y^{\text{in}} + iaS_zJ_z, \approx S_y^{\text{in}} + \frac{an}{2} J_z, \quad (6) \]
where \( n = 2s \) is the total number of photons in the pulse. Now we consider two samples, where we will use the notation \( J_k^{(1)} \) and \( J_k^{(2)} \), where \( k \in \{x, y, z\} \), for the spin operators for samples 1 and 2, respectively, and \( J_k = J_k^{(1)} + J_k^{(2)} \). We may replace \( J_z \) with \( J_z^{(+)} \), the total \( z \) component of spin for the two samples, giving
\[ S_y^{\text{out}} \approx S_y^{\text{in}} + \frac{an}{2} J_z^{(+)}. \quad (7) \]
Thus a measurement of \( S_y^{\text{out}} \) gives a QND measurement of \( J_z^{(+)} \). In considering detection, it is more convenient to consider the Stokes vector for the light integrated over time interval \( \delta t \). The transformation of this component is
\[ \delta S_y^{\text{out}} \approx \delta S_y^{\text{in}} + \frac{an}{2} J_z^{(+)}, \quad (8) \]
where \( \delta n \) is the total number of photons in time \( \delta t \). In Ref. \[1\], TMSS states are obtained by performing QND measurements on \( J_y^{(+)} \) as well. They achieve this by applying a magnetic field in the \( x \) direction, in order to induce Larmor precession of the \( y \) and \( z \) components of the spin. This means that Eq. \[8\] becomes
\[ \delta S_y^{\text{out}}(t) \approx \delta S_y^{\text{in}}(t) + \frac{an}{2} \left[ J_z^{(+)} \cos(\Omega t) + J_y^{(+)} \sin(\Omega t) \right], \quad (9) \]
where \( \Omega \) is the frequency of the precession. In this case measurement of \( \delta S_y^{\text{out}} \) alternately gives QND measurements of \( J_z^{(+)} \) and \( J_y^{(+)} \), thus giving reduced uncertainty in both \( J_z^{(+)} \) and \( J_y^{(+)} \).

For the TMSS states considered in Ref. \[1\] there is the requirement that \( (J_z^{(1)}) = -(J_z^{(2)}) \gg 1 \). Here we will instead take the convention that both \( J_z^{(1)} \) and \( J_z^{(2)} \) are large and close to the total spin \( J \), and replace \( J_y^{(+)} \) with \( J_y^{(-)} \). This is equivalent to a trivial rotation of coordinates for one of the modes and does not require a physical alteration to the experiment. If we now define
\[ J_z^{(0)}(t) \equiv J_z^{(+)} \cos(\Omega t) + J_y^{(-)} \sin(\Omega t), \quad (10) \]
then Eq. \[8\] becomes
\[ \delta S_y^{\text{out}}(t) \approx \delta S_y^{\text{in}}(t) + \frac{an}{2} J_z^{(0)}(t). \quad (11) \]

In order for this to be accurate, we require \( \delta n \) to be large, so that we may perform the SU(2)→HW(2) contraction. Under this contraction, \( \delta S_y = \sqrt{\delta s/2X} \), where \( X = b + b^\dagger \). Therefore the transformation in the quadrature \( X \) is
\[ X^{\text{out}}(t) = X^{\text{in}}(t) + a\sqrt{\delta n} J_z^{(0)}(t). \quad (12) \]
As the initial state is equivalent to the vacuum state \([\text{under the SU(2)→HW(2) contraction}] \), it has a variance of 1 in its measured value. The same is true for \( X^{\text{out}} \), as the extra term only changes the mean.

Measurements are made on the output beam by splitting the beam into +45° and −45° linearly polarized components at a polarizing beam splitter and measuring the intensities at the two outputs. We then define the photocurrent as
\[ I_c(t) = \lim_{\delta t \rightarrow 0} \frac{\delta N_+ - \delta N_-}{\sqrt{\nu \delta t}}, \quad (13) \]
where \( \delta N_+ \) and \( \delta N_- \) are the photon counts at the +45° and −45° linearly polarized outputs and \( \nu \) is the photon flux \( \delta n/\delta t \). We may interpret \( I_c(t) \) as the measured value of \( X \) divided by \( \sqrt{\delta t} \), so the photocurrent is given by
\[ I_c(t) = a\sqrt{\nu} \langle J_z^{(0)} \rangle_c + \xi(t), \quad (14) \]
where the subscript \( c \) indicates the average for the conditioned evolution, and \( \xi(t) \) is a real Gaussian white noise term such that \( \langle\xi(t)\xi(t')\rangle = \delta(t - t') \).

We take the limit of small time intervals \( \delta t \), though for finite photon flux this would mean that the SU(2)→HW(2) contraction is no longer valid, as \( \delta n \) would go to zero. In order for the limit to be rigorously correct, we need to take the limit of large \( \nu \) as well as small \( \delta t \). Nevertheless, this limit should be an accurate approximation for large photon flux. The conditioned master equation is then
\[ d\rho_c = \Gamma D[J_z^{(0)}]_c \rho_c dt + \sqrt{\Gamma} dW(t) \mathcal{H}[J_z^{(0)}]_c \rho_c, \quad (15) \]
where \( \Gamma = a^2 \nu/4 \), \( dW(t) = \xi(t) dt \) is an infinitesimal Wiener increment, \( \mathcal{D}[r] \rho = r \rho r^\dagger - (r^\dagger r \rho + r \rho^\dagger r)/2 \), and \( \mathcal{H}[r] \rho = r \rho + r^\dagger - Tr[(r + r^\dagger) \rho] \rho \).

### III. Feedback

The quadrature \[12\] is measured by detecting the photocurrent \[13\], thereby yielding TMSS states. The drawback is that this will yield nonzero values of \( \langle J_z^{(+)} \rangle \) and
\( \langle J_y^{-1} \rangle \), whereas for the optimal TMSS states \( \text{3} \), the expectation values are zero. In order to obtain states closer to the optimal TMSS states, we apply feedback to bring these expectation values closer to zero. For example, when \( \cos(\Omega t) = 1 \), we may apply a Hamiltonian proportional to \( J_y^{(1)} \) in order to bring \( \langle J_y^{(1)} \rangle \) towards zero, analogous to the case considered by TMW. When \( \sin(\Omega t) = 1 \), we wish to apply a Hamiltonian proportional to \( J_z^{-1} \) in order to bring \( \langle J_z^{(1)} \rangle \) towards zero. For arbitrary \( t \), we will apply the Hamiltonian

\[
H_{fb} = F(t)I_c(t),
\]

where

\[
F(t) = \frac{\lambda(t)}{\sqrt{T}} J_y^{(1)}(t),
\]

and

\[
J_y^{(1)}(t) \equiv J_y^{(+)} \cos(\Omega t) - J_z^{(-)} \sin(\Omega t).
\]

We choose this definition for \( J_y^{(1)} \) because \([J_y^{(1)}, J_y^{(0)}] = -iJ_y^{(+)} \) [we will omit the \( t \) from this point for brevity]. As \( \langle J_y^{(+)} \rangle \) is close to \( 2j \) for the states we are considering, this means that rotation about \( J_y^{(1)} \) will alter the expectation value of \( J_z^{(1)} \). This Hamiltonian may be implemented by a radio frequency magnetic field \([4] \). This generates a Hamiltonian proportional to \( J_y \), which becomes \( J_y^{(1)} \) due to the Larmor precession of the atomic spin.

Using this feedback, we obtain the unconditioned master equation \([3] \)

\[
\dot{\rho} = -i\left[ \frac{1}{2} c F(t) F(t)c, \rho \right] + g[\rho c - c\rho],
\]

with

\[
c = \sqrt{T} J_z^{(1)}.
\]

In the single-mode case TMW pointed out that the term \( c F(t) F(t)c \) is proportional to \( J_y J_x + J_x J_y \), which is the two-axis countertwisting Hamiltonian \([3] \). In a similar way we can derive a Hamiltonian for the two-mode case that produces TMSS states and is analogous to the two-axis countertwisting Hamiltonian. It is straightforward to show that

\[
c F(t) F(t)c = \lambda(t) \left\{ [J_y^{(1)}]^2 + (J_y^{(2)})^2 - (J_y^{(1)})^2 \right. \\
\left. - (J_z^{(2)})^2 \sin(2\Omega t) + (J_y^{(1)})^2 J_y^{(1)} J_z^{(1)} + J_z^{(2)} J_y^{(2)} + J_y^{(2)} J_z^{(2)}) \cos(2\Omega t) \\
+ 2(J_z^{(1)} J_y^{(2)} + J_y^{(1)} J_z^{(2)}) \right\}.
\]

In the limit of large \( \Omega \), the contribution to the evolution from the terms containing \( \sin(2\Omega t) \) and \( \cos(2\Omega t) \) is negligible. Omitting these terms yields

\[
c F(t) F(t)c \approx 2\lambda(t) (J_z^{(1)} J_y^{(2)} + J_y^{(1)} J_z^{(2)}),
\]

which indicates that it is possible to produce two-mode spin squeezing using the Hamiltonian

\[
H \propto J_z^{(1)} J_y^{(2)} + J_y^{(1)} J_z^{(2)},
\]

\[\text{A. Simple feedback}\]

Now we will consider the appropriate feedback strength, \( \lambda(t) \). It is easily shown that the variation in \( \langle J_y^{(0)} \rangle_c \) due to the conditioned evolution \([15] \) is

\[
\text{Tr}(J_z^{(0)} d\rho_c) = 2\sqrt{T} dW [(J_y^{(1)})^2_c - (J_y^{(1)})^2].
\]

The variation in \( \langle J_y^{(1)} \rangle_c \) due to the feedback is

\[
i[(H_{fb}, J_y^{(1)})] = i\frac{\lambda}{\sqrt{\Gamma}} I_c (|J_y^{(1)}, J_y^{(0)}\rangle_c)
\]

\[
= -\frac{\lambda}{\sqrt{\Gamma}} I_c (|J_y^{(1)}\rangle_c).
\]

If \( \langle J_y^{(1)} \rangle_c = 0 \), then the appropriate feedback to keep this equal to zero is

\[
\lambda(t) = \frac{2\Gamma \langle (J_y^{(1)})^2 \rangle_c}{\langle J_y^{(1)} \rangle_c}.
\]

If the unconditioned state is close to being pure, then we may also use this expression with the unconditioned averages without significant loss of accuracy:

\[
\lambda(t) = \frac{2\Gamma \langle (J_y^{(1)})^2 \rangle}{\langle J_y^{(1)} \rangle_c}.
\]

Note that this result is comparable to that found for the case of feedback for single-mode spin squeezing by TMW:

\[
\lambda(t) = \frac{2\Gamma \langle J_z^2 \rangle}{\langle J_z \rangle}.
\]

This is due to the fact that the commutation relations for \( J_x, J_y \), and \( J_z \) are identical to those for \( J_y^{(+)} \), \( J_y^{(1)} \), and \( J_z^{(1)} \).

\[\text{B. Analytic feedback}\]

Next we will consider the possibilities for using an analytic expression for the feedback strength \( \lambda(t) \). In Ref. \([2] \) the approximate relations

\[
\langle J_z^2 \rangle \approx (4\Gamma t + 2/j)^{-1},
\]

\[
\langle J_x \rangle \approx j \exp(-\Gamma t/2),
\]

are used to derive the analytic feedback scheme

\[
\lambda(t) = \frac{\Gamma \exp(\Gamma t/2)}{1 + 2j \Gamma t}.
\]

In the two-mode case, it is possible to obtain the corresponding relation for \( \langle J_y^{(1)} \rangle_c \), but not for \( \langle (J_z^{(1)})^2 \rangle \).
In these derivations it is convenient to define the scaled variables \( v = \Gamma t \) and \( \Lambda(t) = \lambda(t)/\Gamma \). In terms of these variables the master equation (19) becomes

\[
\dot{\rho} = -\frac{i\Lambda(t)}{2} \left[ (J^{(\Omega)}_x J^{(\Omega)}_y + J^{(\Omega)}_y J^{(\Omega)}_z), \rho \right] + \mathcal{D}[J^{(\Omega)}_z] \\
- i\Lambda(t) J^{(\Omega)}_y], \rho,
\]

where the dot indicates a derivative with respect to \( v \). Using these variables, the evolution is independent of \( \Gamma \). The value of \( \Gamma \) does not qualitatively affect the evolution and merely provides a scaling to the time and to \( \lambda(t) \).

It is straightforward to show that

\[
\frac{d}{dv} \langle J^{(+)}_x \rangle = \text{Tr} (J^{(+)}_x \dot{\rho}) \\
= -\frac{1}{2} \langle J^{(+)}_x \rangle + 2\Lambda \langle (J^{(\Omega)}_z)^2 \rangle - \frac{1}{2} \Lambda^2 \langle J^{(+)}_x \rangle \\
= -\frac{1}{2} \langle J^{(+)}_x \rangle + \frac{2 \langle (J^{(\Omega)}_z)^2 \rangle}{\langle J^{(+)}_x \rangle}.
\]

Initially this derivative is equal to zero. However, the value of \( \langle (J^{(\Omega)}_z)^2 \rangle \) falls rapidly and that of \( \langle J^{(+)}_x \rangle \) only falls slowly, so the second term quickly becomes negligible. If we neglect the second term, the evolution of \( \langle J^{(+)}_x \rangle \) is

\[
\langle J^{(+)}_x \rangle \approx 2j \exp(-v/2),
\]

which is equivalent to that in the single-mode case. This technique fails to find the corresponding expression for \( \langle (J^{(\Omega)}_z)^2 \rangle \), but we have found numerically that a good approximation is given by

\[
\langle (J^{(\Omega)}_z)^2 \rangle \approx \frac{e^{-v/4}}{2v+1/j}.
\]

Using these results in the expression for the feedback (27), we obtain

\[
\Lambda(t) = \frac{\Gamma \exp(\Gamma t/4)}{1 + 2j \Gamma t}.
\]

C. Optimal feedback

The third alternative that we will consider is optimal feedback. In order to derive this, it is convenient to define the variables

\[
\zeta = \frac{\langle (J^{(+)}_x)^2 + (J^{(-)}_y)^2 \rangle}{2j}, \quad \chi = \frac{\langle J^{(+)}_x \rangle}{2j}.
\]

The initial state used is that with individual coherent spin states in each arm, so both \( \zeta \) and \( \chi \) are equal to 1. As time progresses both \( \chi \) and \( \zeta \) decrease. In order for the feedback to be optimal, the value of \( \zeta \) should be as small as possible for each value of \( \chi \). Therefore, if \( \chi \) decreases by an amount \( \delta \chi \), then the amount by which \( \zeta \) decreases should be maximized. This means that the feedback that is optimum will be that which produces the maximum slope. Therefore, in order to determine the optimal feedback, we wish to solve

\[
\frac{d}{d\Lambda} \frac{\partial \zeta}{\partial \chi} = 0.
\]

This can be solved using

\[
\frac{d}{d\Lambda} \left[ \frac{d}{dv} (J^{(+)}_x)^2 + (J^{(-)}_y)^2 \right] = 0.
\]

Using the master equation (19), it is straightforward to show that

\[
\text{Tr} [(J^{(+)}_x) \dot{\rho}] = -\frac{1}{2} \left[ 1 + \Lambda^2 \right] \langle J^{(+)}_x \rangle + 2\Lambda \langle (J^{(\Omega)}_z)^2 \rangle - \Lambda \langle (J^{(\Omega)}_x)^2 \rangle.
\]

Determining the corresponding expression for the numerator in Eq. (39) is more difficult, and if it is performed directly it leads to an extremely complicated expression. However, there is a more convenient way of determining this expression. Note that

\[
\langle (J^{(\Omega)}_z)^2 \rangle = \langle (J^{(\Omega)}_x)^2 \rangle \cos^2(\Omega t) + \langle (J^{(\Omega)}_y)^2 \rangle \sin^2(\Omega t) + \langle (J^{(\Omega)}_z)^2 \rangle.
\]

For large \( \Omega \) we may average over \( t \), which gives

\[
\langle (J^{(\Omega)}_z)^2 \rangle \approx \frac{1}{2} \left[ \langle (J^{(+)}_x)^2 \rangle + (J^{(-)}_y)^2 \right].
\]

Numerically this approximation was found to be very accurate. Using this result we find that

\[
\frac{1}{2} \text{Tr} [(J^{(+)}_x)^2 + (J^{(-)}_y)^2] \approx \text{Tr} [(J^{(\Omega)}_z)^2] \\
= \Lambda^2 \langle (J^{(+)}_x)^2 \rangle - \Lambda \langle (J^{(\Omega)}_x)^2 \rangle - \Lambda \langle 4J^{(\Omega)}_x J^{(+)}_x J^{(+)}_z + J^{(+)}_y \rangle.
\]

Thus we find that Eq. (39) becomes

\[
\frac{d}{d\Lambda} \left\{ -\Lambda^2 d - \Lambda e \right\} = 0,
\]

where

\[
d = \langle (J^{(+)}_x)^2 \rangle - \langle (J^{(\Omega)}_z)^2 \rangle, \\
e = \langle 4J^{(\Omega)}_x J^{(+)}_x J^{(+)}_z + J^{(+)}_y \rangle, \\
f = \frac{1}{2} \langle J^{(+)}_x \rangle, \\
g = 2 \langle (J^{(\Omega)}_z)^2 \rangle.
\]

Solving this gives

\[
\Lambda = \frac{-fd \pm \sqrt{(fd)^2 + 4fe(fg - dg)}}{fe - dg}.
\]

Numerically it is found that the correct solution that maximizes the slope is that with the positive sign.

Note that this derivation of the optimum feedback does not rely on any commutation relations other than the usual \( su(2) \) commutation relations. It therefore is also applicable to the case of feedback for single-mode spin squeezing, with \( J^{(+)}_x \), \( J^{(\Omega)}_y \), and \( J^{(\Omega)}_z \) replaced by \( J_x \), \( J_y \), and \( J_z \).
for optimal TMSS states in Fig. 2. As time progresses, therefore a state with \( \zeta < \chi \). The plots of \( \zeta \) versus \( \chi \) for analytic feedback and optimal feedback are also shown in Fig. 4. The results for analytic feedback are very similar to those for simple feedback, with \( \zeta \) dramatically increasing at later times. In contrast, under optimal feedback the value of \( \zeta \) monotonically decreases to an asymptotic value for \( \chi = 0 \). For small times the three feedback schemes produce very similar results, and the states obtained using optimal feedback do not pass significantly closer to the optimal state for teleportation.

The last alternative that we consider in this section is the Hamiltonian (23), which is analogous to the two-axis countertwisting Hamiltonian in the single-mode case. The results for this Hamiltonian are also shown in Fig. 4. For this case there is significant spin squeezing, but the squeezing is generally poorer than for feedback. Nevertheless, at later times the value of \( \zeta \) does not rise dramat-
FIG. 3: The values of $\zeta$ plotted against $\chi$ for spin $j = 5$. The relation for optimal states is shown as the dashed line, the results for the feedback of Eq. (26) are shown as the continuous line, and without feedback as the dotted line. All results are for the conditioned master equation. The cross shows the position of the optimal state considered for teleportation in Ref. [3].

B. Conditioned master equation

Next we consider the results for the conditioned master equation. There are two examples where it is necessary to consider the conditioned master equation. In the case of simple feedback, at the time that $\zeta$ increases dramatically the purity of the state has greatly decreased. This means that the feedback based on the unconditioned averages, Eq. (27), is a poor approximation of the feedback based on conditioned averages, Eq. (26), and will not accurately keep the average $\langle J_z^{(1)} \rangle_c$ equal to zero. For this reason we may obtain better results using the conditioned master equation and the feedback (26).

The second case that we consider here is that without any feedback. In this case there is no reduction in the variances for the unconditioned equation. If we consider the conditioned equation, however, there is a reduction in the variances, though the means $\langle J_z^{(1)} \rangle_c$ and $\langle J_y^{(1)} \rangle_c$ are nonzero. For this reason we will consider the conditioned master equation for this case, with $\zeta$ defined by

$$\zeta = \frac{(\langle J_z^{(1)} \rangle_c^2 - \langle J_z^{(1)} \rangle_c^2)^2 + (\langle J_y^{(1)} \rangle_c^2 - \langle J_y^{(1)} \rangle_c^2)^2}{2j}. \quad (48)$$

To minimize the random differences between the two cases the same random numbers were used for each. The results for these two cases are plotted in Fig. 3.

The results for the feedback (26) are extremely close to the line for optimal states, with only temporary excursions from it. In particular, the value of $\zeta$ does not increase dramatically at large times. This indicates that improved results may be obtained using feedback based on the conditioned state.

On the other hand, implementing this feedback may be far more difficult. In the forms of feedback considered for the unconditioned equation, the feedback was proportional to the measurement record, with a proportionality constant $\lambda$ that is a function of time only. This feedback may be performed efficiently by calculating $\lambda(t) \text{ beforehand and programming it into a field programmable gate array}$ [17]. In contrast, the value of $\lambda$ given by Eq. (26) is dependent on the measurement record, and therefore would need to be calculated during the measurement, introducing a significant time delay.

The results without feedback quickly diverge from the line for optimal states, although they do not differ from it greatly. The major difference is at later times. While, with feedback, the value of $\zeta$ is reduced very close to zero, without feedback the value of $\zeta$ cannot be reduced as far. This means that, although a significant reduction in the variances can be achieved via feedback for strong QND measurements (where $\Gamma t > 1$), feedback does not give much improvement for weak measurements.

V. COMPARISON WITH SINGLE-MODE CASE

There is clearly a great similarity between the single-mode and two-mode cases. The operators $J_x^{(1)}$, $J_y^{(1)}$, and $J_z^{(1)}$ obey exactly the same commutation relations as $J_x$, $J_y$, and $J_z$ in the single-mode case. This means that, for example, the optimal feedback also applies to the single-mode case (as was mentioned in Sec. IV.C). As we may replace the sum of the variances $\langle (J_z^{(1)})^2 + (J_y^{(1)})^2 \rangle$ with $2\langle (J_z^{(1)})^2 \rangle^2$, it might appear that this case is identical to the single-mode case. Nevertheless, there is a subtle difference due to the fact that the operator $J_z^{(1)}$ is time dependent.

For the case of spin 1/2, there is complete equivalence between the single-mode case (for spin $j = 1$) and the two-mode case. We will take the Hilbert space for the single-mode case to be that for two spin 1/2 particles, so that it is the same as for the two-mode case. The optimal TMSS states for $j = 1/2$ are also optimal single-mode squeezed states in terms of the total spin. That is, they minimize $\langle (J_z^{(1)})^2 \rangle$ for a given $\langle J_z^{(1)} \rangle$. The converse is also true: the states optimized for single-mode spin squeezing are automatically optimized for two-mode spin squeezing.

In addition, both the single- and two-mode counter-twisting Hamiltonians produce optimal spin squeezed states. Not only this, but the feedback as given by Eq. (23) in the single-mode case, or Eq. (27) in the two-mode case, and the optimal feedback described in Sec. IV.C, give optimal spin squeezed states. The only feedback that does not give optimal states is the analytic feedback.
for total spin 1 we have the differential equations

\[ \frac{d}{dv}\langle J_z \rangle = 2\langle J_x^2 - J_y^2 \rangle, \]

\[ \frac{d}{dv}\langle J_x^2 \rangle = -\langle J_x \rangle, \]

\[ \frac{d}{dv}\langle J_y^2 \rangle = \langle J_x \rangle. \]

The solution of these equations is

\[ \langle J_x \rangle^2 = 4(\langle J_x^2 \rangle - \langle J_y^2 \rangle^2). \]  

This is the relation for optimal single-mode spin squeezed states as given by Sørensen and Mølmer. This shows that the states produced by the one- and two-mode countertwisting Hamiltonians and the one- and two-mode optimal states are all identical.

For the case of feedback, it is sufficient to show that the feedbacks of Eqs. (28) and (27) give optimal states, as this will imply that the optimal feedback gives optimal spin squeezed states. The derivation in this case is lengthy, and is given in Appendix B. It is interesting to note that in this case the optimal feedback is given by Eq. (B7), and differs dramatically from the analytic feedback used for higher spin.

The equivalence between the single- and two-mode cases does not hold for any total spin higher than 1. As shown in Fig. 5, there are significant differences between the optimal single- and two-mode spin squeezed states for total spin above 1. As can be seen, for a total spin of 2 or more the values of \( \xi \) are significantly less for the single-mode case than for the two-mode case. This reflects the fact that in the two-mode case we wish to minimize \( \langle J_y^{(-)} \rangle^2 \) as well as \( \langle J_x^{(+)} \rangle^2 \).

As there are such strong similarities between the single-mode and two-mode cases, it is reasonable to apply the squeezing parameter as considered in Refs. [8, 9, 19] to the two-mode case. In the single-mode case the squeezing parameter was defined by

\[ \xi_\epsilon^2 = 2j\langle J_y^2 \rangle / \langle J_x^2 \rangle. \]

With the definitions of \( \xi \) and \( \chi \) for the single-mode case above, this can be expressed as \( \xi_\epsilon^2 = \xi / \chi^2 \). It would seem
reasonably to use the same definition for the two-mode case.

For the two-mode case we find that the minimum value of \( \xi^2 \), namely \( \xi^2_{\text{min}} \), varies with spin \( j \) as shown in Fig. 6(a). In this figure the values of \( \xi^2_{\text{min}} \) have been multiplied by \( j + 1 \) in order to more easily compare the results for different spins. Very similar results are obtained for the simple feedback of Eq. (27) and the optimal feedback of Eq. (28). Using the analytic feedback of Eq. (29) gives results that are similar to, but slightly above, those for the other two feedback schemes. All of these feedback schemes give results that are significantly above those for feedback. In both cases there are only small differences between the results for the different types of feedback. The only qualitative difference between the scaling constants for the single- and two-mode cases is that in the single-mode case the analytic feedback gives slightly better results than the simple feedback of Eq. (27). In contrast, in the two-mode case the scaling constants are indistinguishable.

VI. EXPERIMENTAL PROSPECTS

There are some complications in applying this theory to the experiment of Ref. [1]. In this experiment, \( a \approx 5 \times 10^{-13} \), and \( \nu \approx 2 \times 10^{12} \text{s}^{-1} \). This means that \( \Gamma \approx 1.4 \times 10^{-9} \text{s}^{-1} \), which implies that the time required for maximal squeezing is on the order of \( 10^9 \text{s} \) (over 20 yr). Therefore a far more intense beam or a stronger interaction would be required to obtain the two-mode spin squeezing described here.

It must be emphasized that this long time is required for the measurements, not the feedback. Feedback may be applied to the states obtained in Ref. [1] with a negligible increase in the time required for the experiment. Nevertheless, for the conditions of this experiment, although the feedback will bring the means of \( J_z^+ \) and \( J_y^+ \) towards zero, it will not significantly reduce the variances. It is only when the QND measurement can be performed strongly enough (i.e., with a strong enough interaction over a long enough time) to obtain spin squeezing close to maximum that an improvement in the variance is obtained by using feedback.

Another problem is spontaneous losses due to absorption of QND probe light. The loss rate due to this is \( N^2 g^2 n/4 \Delta^2 \), where \( N = 4j \) is the total number of atoms and \( g \) is the one-photon Rabi frequency. As in the case of single-mode spin squeezing [1], this loss will be very large over the time period \( 1/\Gamma \) for free space. Both this problem and the problem of the long interaction time may be overcome by performing the experiment in a cavity in the strong-coupling regime.

Two other common experimental problems are inefficient detectors and time delays. As for the case of single-mode spin squeezing, inefficient detectors do not affect the scaling. The system has the potential to be far more sensitive to time delays than in the single-mode case, however. The problem is that, as the spin component that is being measured is rotating rapidly, the feedback may be correcting for measurements of a different spin component.
In order to correct the right spin component, the rotation of the spin component that is measured should have completed an integral number of rotations during the time delay. Experimentally, this means that the magnetic field should be adjusted such that the frequency Ω is an integral multiple of 2π/τ, where τ is the time delay. Provided that this is done, time delays should not be more of a problem than in the single-mode case.

VII. CONCLUSIONS

Two-mode spin squeezed states are important states to produce for quantum teleportation. Here we have shown that it is possible to produce states very close to the optimal TMSS states by adapting the feedback for single-mode spin squeezing considered by TMW. These states are not conditioned on the measurement record, in contrast to the conditional two-mode spin squeezing discussed in Ref. [1].

Using the simple feedback scheme (27) it is possible to obtain states that are quite close to optimal TMSS states for small times, but strongly diverge from these states at later times. In particular, for spins above about 5 it is possible to obtain states very close to the TMSS states considered for teleportation in Ref. [3]. An analytic feedback scheme (36) also gives similar results. This feedback scheme is more practical experimentally, as the appropriate feedback strength to be used is easily calculated.

We have derived the optimal feedback that produces the maximum possible spin squeezing. This feedback is also applicable to the case of feedback for single-mode spin squeezing. This feedback gives states even closer to the optimal TMSS states.

We have also derived a Hamiltonian for the two-mode case that is equivalent to the two-axis countertwisting Hamiltonian introduced in Ref. [10]. This Hamiltonian produces spin squeezing, but not as strongly as the measurements with feedback.

In the case of spin 1/2 both the feedback (except for the analytic feedback) and the countertwisting Hamiltonian produce optimal TMSS states. These states are equivalent to optimal single-mode spin squeezed states if the two modes are considered as a single spin-1 system. In the single-mode case optimal spin squeezed states are produced both by feedback and by the countertwisting Hamiltonian.

Acknowledgments

The authors acknowledge valuable discussions with Laura Thomsen and Howard Wiseman. We are also grateful for constructive criticism from Eugene Polzik. This project has been supported by an Australian Research Council Large Grant.

APPENDIX A: RELATION OF VARIABLES TO EXPERIMENT

Here we give further explanation of the variables introduced in Eq. (1). The coupling constant a is given by

\[ a = \frac{\sigma}{A(I + 1/2)} \frac{\gamma}{\Delta} \alpha_v, \]  

(A1)

where σ is the resonant absorption cross section, A is the area of the transverse cross section of the light beam, I is the nuclear spin, γ is the spontaneous emission rate of the upper atomic level, Δ is the detuning, and \( \alpha_v \) is the dynamic vector polarizability. We have omitted the bounds on the integral (1) for greater generality, so we may apply this expression to multiple samples. Explicit bounds are unnecessary, as there is no contribution to the integral from regions where there are no atoms.

The continuous spin operators for the sample are defined as

\[ J_k(z, t) = \lim_{\delta z \to 0} \frac{1}{\delta z} \sum_{\mu} \frac{1}{2} \sigma_k^\mu, \]  

(A2)

where \( k \in \{x, y, z\} \) and \( \sigma_k^\mu \) is the Pauli operator for particle \( \mu \). The sum is over all particles between \( z \) and \( z + \delta z \) over the cross section of the sample. The operator for the entire sample is obtained by integrating over \( z \).

The field is described by the continuous-mode annihilation operators \( a_k(z, t) \), where \( k = x \) and \( y \) for the \( x \) polarized and \( y \) polarized modes, respectively. The instantaneous Stokes parameters are

\[ S_x(z, t) = \frac{1}{2} [a_x^\dagger(z, t)a_x(z, t) - a_y^\dagger(z, t)a_y(z, t)], \]

\[ S_y(z, t) = \frac{1}{2} [a_x^\dagger(z, t)a_y(z, t) + a_y^\dagger(z, t)a_x(z, t)], \]

\[ S_z(z, t) = -\frac{1}{2} [a_y^\dagger(z, t)a_y(z, t) - a_x^\dagger(z, t)a_x(z, t)]. \]  

(A3)

The Stokes vector for the entire pulse at position \( z \) is given by \( \mathbf{S}(z) = \int \mathbf{S}(z, t) dt \).
APPENDIX B: FEEDBACK FOR TOTAL SPIN 1

Here we show that the feedback of either Eqs. (28) or (27) gives optimal spin squeezed states for a total spin of 1. To see this, note first that the first term in Eq. (19) is just the same as that produced by the counter twisting Hamiltonian, and so will produce optimal states. In order to show that the feedback produces optimal states, it remains to be shown that the second term, \( D \{ c - i F \} \rho \), does not alter the evolution. Specifically, in the single-mode case

\[
\text{Tr} \{ J_z D \{ J_z - i \Lambda J_y \} \rho \} = -\frac{1}{2}(1 + \Lambda^2) \langle J_z \rangle + \Lambda \langle J_x^2 \rangle + \Lambda \langle J_y^2 \rangle.
\]

(B1)

If the state \( \rho \) is an optimal state, then \( \langle J_x^2 \rangle = 1 \), so this simplifies to

\[
\text{Tr} \{ J_z D \{ J_z - i \Lambda J_y \} \rho \} = -\frac{1}{2}(1 + \Lambda^2) \langle J_z \rangle + \Lambda.
\]

(B2)

Similarly we can show that

\[
\text{Tr} \{ J_x^2 D \{ J_z - i \Lambda J_y \} \rho \} = -\frac{\Lambda}{2} \langle J_x \rangle + \Lambda^2 (\langle J_x^2 \rangle - \langle J_z^2 \rangle).
\]

(B3)

For optimal states, \( \langle J_x^2 \rangle = 1 \), so this becomes

\[
\text{Tr} \{ J_x^2 D \{ J_z - i \Lambda J_y \} \rho \} = -\frac{\Lambda}{2} \langle J_x \rangle + \Lambda^2 (1 - \langle J_z^2 \rangle).
\]

(B4)

It is simple to show that both Eqs. (B3) and (B4) will be zero if Eq. (54) is satisfied, and the feedback is given by Eq. (28).

Therefore, if the state is already in an optimal state, and the feedback as given by Eq. (28) is used, then the state will continue to be in an optimal state. On the other hand, if some other feedback is used, then optimal states will not be obtained. For example, the analytic feedback considered by LMW does not give optimal states. In order to determine analytic feedback that will give optimal states, note that the differential equations for \( \langle J_x \rangle \) and \( \langle J_x^2 \rangle \) are

\[
\frac{d}{dv} \langle J_x \rangle = \Lambda (2 \langle J_z^2 \rangle - 1),
\]

\[
\frac{d}{dv} \langle J_x^2 \rangle = -\frac{\Lambda}{2} \langle J_x \rangle.
\]

(B5)

Solving these using the feedback (28) and Eq. (54) gives

\[
\langle J_x \rangle = \sqrt{2e^{-v} - e^{-2v}},
\]

\[
\langle J_x^2 \rangle = \frac{\Lambda}{2} e^{-v}.
\]

(B6)

This implies that the value of \( \Lambda \) should change with time as

\[
\Lambda(v) = \frac{1}{\sqrt{2e^v - 1}}.
\]

(B7)

This is quite different from the analytic expression applied for larger spins.

The case for feedback for two-mode spin squeezing is analogous to this, except that \( J_z \), \( J_y \), and \( J_z \) are replaced with \( J_z^{(+)} \), \( J_y^{(1)} \), and \( J_z^{(2)} \). This therefore shows that optimal states are obtained in the one- and two-mode cases using the feedback given by Eqs. (28) and (27), respectively.

[1] B. Julsgaard, A. Kozhekin, and E. S. Polzik, Nature (London) 413, 400 (2001).
[2] A. Kuzmich and E. S. Polzik, Phys. Rev. Lett. 85, 5639 (2000).
[3] D. W. Berry and B. C. Sanders, New J. Phys. 4, 8 (2002).
[4] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, England, 2000).
[5] C. M. Caves and B. L. Schumaker, Phys. Rev. A 31, 3068 (1985).
[6] A. Furusawa, J. L. Sorensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik, Science 282, 706 (1998).
[7] A. S. Sorensen and K. Mølmer, Phys. Rev. Lett. 86, 4431 (2001).
[8] H. M. Wiseman and G. J. Milburn, Phys. Rev. A 49, 1350 (1994).
[9] L. K. Thomesen, S. Mancini, and H. M. Wiseman, Phys. Rev. A 65, 061801(R) (2002).
[10] H. M. Wiseman and R. B. Killip, Phys. Rev. A 57, 2169 (1998).
[11] D. W. Berry and H. M. Wiseman, Phys. Rev. A 63, 013813 (2001).
[12] A. Kuzmich, L. Mandel, J. Janis, Y. E. Young, R. Ejnissen, and N. P. Bigelow, Phys. Rev. A 60, 2346 (1999).
[13] T. Holstein and H. Primakoff, Phys. Rev. 58, 1098 (1940).
[14] R. T. Sang, G. S. Sumny, B. T. H. Varcoe, W. R. MacGillivray, and M. C. Standage, Phys. Rev. A 63, 023408 (2001).
[15] H. M. Wiseman, Phys. Rev. A 49, 2133 (1994).
[16] M. Kitagawa and M. Ueda, Phys. Rev. A 47, 5138 (1993).
[17] J. Stockton, M. Armen, and H. Mabuchi, quant-ph/0203143, 2002 (unpublished).
[18] A. Sørensen, L.-M. Duan, J. I. Cirac, and P. Zoller, Nature (London) 409, 63 (2001).
[19] X. Wang, J. Opt. B: Quantum Semiclassical Opt. 3, 93 (2001).