We study the effects of diffusion on a Λ-gradient echo memory, which is a coherent optical quantum memory using thermal gases. The efficiency of this memory is high for short storage time, but decreases exponentially due to decoherence as the storage time is increased. We study the effects of both longitudinal and transverse diffusion in this memory system, and give both analytical and numerical results that are in good agreement. Our results show that diffusion has a significant effect on the efficiency. Further, we suggest ways to reduce these effects to improve storage efficiency.

I. INTRODUCTION

Quantum memory is an important tool in many quantum information protocols, including quantum repeaters for long-distance quantum communication [1], and identity quantum gates in quantum computation [2]. Numerous optical quantum memories have been developed, including electromagnetically-induced transparency (EIT) based quantum memory [3, 4], far-detuned Raman process memory [5, 6], and photon-echo quantum memories: controlled reversible inhomogeneous broadening (CRIB) memory [7, 8], atomic frequency combs (AFC) memory [9], and gradient echo memory (GEM) [10–12]. A review of these schemes can be found in [13]. Of these schemes, the most impressive efficiency so far attained experimentally is 87% by Λ-GEM scheme [14] using warm rubidium vapor. In this paper, we will examine the effects of atomic diffusion on the Λ-GEM system, which may limit this efficiency for larger storage times.

Λ-GEM is a memory using a 3-level, Λ-type atom (Fig. 1). The input optical pulse couples the two metastable lower states through a control field. The excited state is coupled by a weak optical field, the positive frequency component of the electric field is described by the slowly varying operator

\[ \hat{E}(\mathbf{r}, t) = \sum_k \sqrt{\frac{1}{V}} a_k(t) e^{i k \cdot \mathbf{r}} e^{-i k_0 z} e^{i \omega_0 t} \]

with detuning \( \Delta \), where \( V \) is the quantization volume, \( \omega_0 \) is the carrier frequency of the quantum field and \( k_0 = \omega_0 / c \). The excited state |3⟩ is also coupled to the metastable state |2⟩ via a coherent control field with Rabi frequency \( \Omega_c \) and a two photon detuning \( \delta \). This two photon detuning is spatially varied \( \delta(z, t) = \eta(t) z \), with time dependent gradient \( \eta(t) \). Then the interaction Hamiltonian in the rotating frame with respect to the field

\[ gE(z, t) \]

We also suggest ways to reduce these effects to improve storage efficiency.

II. Λ GRADIENT ECHO MEMORY

We consider a medium consisting of Λ-type 3-level atoms with two metastable lower states as shown in Fig. 1. The ground state |1⟩ and the excited state |3⟩ are coupled by a weak optical field, the positive frequency component of the electric field is described by the slowly varying operator

\[ \hat{E}(\mathbf{r}, t) = \sum_k \sqrt{\frac{1}{V}} a_k(t) e^{i k \cdot \mathbf{r}} e^{-i k_0 z} e^{i \omega_0 t} \]
frequencies is
\[
\hat{H} = \sum_{n} \left[ \hbar \Delta \sigma_{33}^{(n)} + \hbar \delta(z_n, t) \sigma_{22}^{(n)} + \hbar g_{n} e^{i k \cdot r_n} \sigma_{31}^{(n)}(n) + \hbar \Omega_c (r_n) \sigma_{32}^{(n)}(n) + h.c. \right],
\]
where \(g_{n} = \gamma \sqrt{2 \omega_{0} / 2\pi} \) is the atom-field coupling constant with \(\gamma\) being the dipole moment of the 1-3 transition, and \(\sigma_{\mu\nu}^{(n)} = |\mu\rangle_i \langle \nu|\) is an operator acting on the \(n\)-th atom at \(r_n = (x_n, y_n, z_n)\). We assume that initially all atoms are in their ground state \(|1\rangle\). We transform to collective operators, which are averages over atomic operators over a small volume centered at \(r\) containing \(N_r \gg 1\) particles,
\[
\sigma_{\mu\nu}(r, t) = \frac{1}{N_r} \sum_{j=1}^{N_r} \sigma_{\mu\nu}^{(j)}(t)
\]
From the Heisenberg-Langevin equations in the weak probe region \((\sigma_{11} \simeq 1, \sigma_{22} \simeq \sigma_{33} \simeq 0)\), we get the Maxwell-Bloch equations\,[18],\]
\[
\dot{\sigma}_{13}^{(n)} = - (\gamma_{13} + i \Delta) \sigma_{13}^{(n)} + i g e^{i k \cdot z_n} E(r_n, t) + i \Omega_c e^{i k \cdot z_n} \sigma_{12}^{(n)},
\]
\[
\dot{\sigma}_{12}^{(n)} = - (\gamma_{12} + i \delta(z_n, t)) \sigma_{12}^{(n)} + i \Omega_c e^{-i k \cdot z_n} \sigma_{13}^{(n)},
\]
\[
\left( \frac{\partial}{\partial t} + c \frac{\partial}{\partial z} - \frac{i c \nabla^2 + c \nabla^2}{2 \omega_{0}} \right) E(r, t) = i g N e^{-i k_{0} z} \sigma_{13}(r, t),
\]
where \(\gamma_{\mu\nu}\) are the decay rates, \(g = \gamma \sqrt{2 \omega_{0} / 2\pi}\) and \(\Omega_c(r) = \Omega e^{ik_{0} z}\). We also omit the Langevin noise operators since we are more interested in the decoherence caused by diffusion. This is equivalent to making a semiclassical approximation for the electric field and the atomic coherences.

In Eq. (4), \(\nabla^2 \sum_{\nu} \sum_{\nu} \) is the diffraction term, and generally, the diffusion effects can be neglected\,[19]. Notice that we are here considering the regime \(\gamma_{13} \gg \gamma_{12}, \Omega_{c}\), we adiabatically eliminate the fast oscillations and set \(\sigma_{13}^{(n)} = 0\). Then we have \(\sigma_{13} = (g e^{i k_{0} z} E + \Omega e^{i k_{0} z} \sigma_{12}) / \Delta\), and we get the reduced Maxwell-Bloch equations,
\[
\dot{\sigma}_{12}^{(n)} = - (i \delta(z_n, t) - i \Omega_{c}^{2}) \sigma_{12}^{(n)} + i g \Omega_{c} e^{i(k_{0} - k) \cdot z} E(r_n, t),
\]
\[
\frac{\partial}{\partial z} E(r, t) = \frac{i g N \Omega_{c}}{c \Delta} e^{-i(k_{0} - k) \cdot z} \sigma_{12}(r, t) + \frac{i g^{2} N}{c \Delta} E(r, t).
\]
Here we neglect decay, i.e. \(\gamma_{12} \rightarrow 0\), since we consider the storage time much less than \(1/\gamma_{12}\).

III. DIFFUSION

We now consider the effects of diffusion on the atomic state. In order to isolate the motional effects of diffusion from collisional dephasing, we assume that the collisions between atoms do not change the state of the atom. Then we derive the diffusion equation for the atomic density matrix \(\rho\). Space is divided into volume elements with length \(\Delta r\) and center \(r\). We associate a density matrix \(\rho(r, t)\) with atoms in this volume element, given by
\[
\rho(r, t) = \frac{1}{N_r} \sum_{j=1}^{N_r} \rho^{(j)}(t),
\]
where \(N_r\) is the atom number in volume centered at \(r\). The total density matrix for the entire system is assumed to be the tensor product of these local density matrices.

Diffusion causes an exchange of atoms between adjacent volumes. During a short time \(\Delta t\), a fraction \(\epsilon\) of the atoms in slice \(r\) migrate into slice \(r \pm \Delta r\). There is also atomic flux back into slice \(r\) from \(r \pm \Delta r\). We assume that the total number density of the atoms is uniform, so the state at \(r\) and \(r \pm \Delta r\) is described by the new density matrix which is the average of the density matrix of atoms remaining in the volume and those that have migrated in to it. The diffusive component of the evolution is therefore
\[
\rho(r, t + \Delta t) = (1 - 2\epsilon)\rho(r, t) + \epsilon(\rho(r + \Delta r, t) + \rho(r - \Delta r, t))
\]
\[
\Rightarrow \partial_t \rho(r, t) = D \nabla^2 \rho(r, t)
\]
where \(D = \epsilon \Delta r^2 / \Delta t\) is the diffusion coefficient. With the same consideration, we get the diffusive component evolution for the atomic correlation functions
\[
\partial_{\mu\nu}(r, t) = D \nabla^2 \sigma_{\mu\nu}(r, t)
\]
Now we introduce the interaction with optical fields. Since diffusion is caused by Brownian motion, this will lead to Doppler shifts in the various detunings. We now consider the interaction between the optical field and a single atom, and quantify the effects of these Doppler shifts. The atom moves at some random velocity, and there will be a Doppler shift for both the signal and control fields. So the detunings in Eq. (4) become \(\Delta = \Delta_0 + \Delta_{Dopp}\), and \(\delta = \delta_0 + \delta_{Dopp}\), with \(\Delta_0, \delta_0\) the detunings for stationary atoms and \(\Delta_{Dopp}, \delta_{Dopp}\) the Doppler shifts. Typically, the one photon Doppler shift \(\Delta_{Dopp} \ll \Delta_0\), state \(|3\rangle\) is still far detuned. So the adiabatic elimination is still valid in the presence of the Brownian motion induced Doppler shift, and we can still reduce the 3-level atom to an effective 2-level atom. The Maxwell-Bloch equation will still reduce to Eq. (5), but
with one photon detuning \( \Delta = \Delta_0 + \Delta_{Dopp} \) and two photon detuning \( \delta = \delta_0 + \delta_{Dopp} \). So, for the reduced two level atomic system, the diffusive Maxwell-Bloch equation for the collective correlation \( \sigma_{12}(z,t) \) averaged over atoms in each volume is

\[
\dot{\sigma}_{12}(r,t) = i \frac{g\Omega_c}{\Delta} e^{i(k_0-k_z)z} E(r, t) - i\delta(z, t)\sigma_{12}(r, t) + D \nabla^2 \sigma_{12}(r, t),
\]

\[
\frac{\partial}{\partial z} E(r, t) = i \frac{gN\Omega_c e^{-i(k_0-k_z)z}}{c\Delta} \sigma_{12}(r, t) + i \frac{g^2N}{c\Delta} E(r, t). \tag{8}
\]

We have absorbed the Stark shift \( \frac{\Omega^2}{2} \) into the two-photon detuning. Here our diffusive Maxwell-Bloch equation is consistent with the result in the EIT system \[15, 17\].

Notice that the signal and control fields are co-propagating, so the Doppler broadening width for \( \delta \) is typically 1kHz, which is much smaller than the frequency width of the signal field (\( \sim 1\)MHz), so we neglect this two-photon Doppler broadening \( \delta_{Dopp} \) and replace \( \delta(z, t) \) by \( \delta_0(t), z \in [-L, L] \).

For the one photon detuning \( \Delta = \Delta_0 + \Delta_{Dopp} \), after we make the adiabatic elimination, it will appear in the denominator (see Eq. (8)), so

\[
\frac{1}{\Delta} \approx \frac{1}{\Delta_0} \left( 1 - \frac{\Delta_{Dopp}}{\Delta_0} + \left( \frac{\Delta_{Dopp}}{\Delta_0} \right)^2 \right).
\]

The term linear in \( \Delta_{Dopp} \) will vanish when we average over many atoms in a volume centred at \( r \), so we can replace \( \Delta \) by \( \Delta_0 \) in our Maxwell-Bloch equation, with second order accuracy [typically \( (\Delta_{Dopp}/\Delta_0)^2 \sim 10^{-3} \).

### IV. ANALYTIC CALCULATION AND NUMERICAL SIMULATION

To quantify the effects of diffusion, we solve for the atomic dynamics. There are three distinct phases during the storage: write-in \(-t_0 < t < 0\), during which the signal is absorbed by the memory; hold \( 0 < t < t_H\), during which the information is stored in the memory and the gradient is turned off; read-out \( t_H < t < t_{out} \), during which the signal is emitted by turning on the flipped gradient.

We quantify the effects of diffusion by the read-out efficiency \( \varepsilon \) defined to be

\[
\varepsilon = \frac{\int_{t_H}^{t_{out}} |f_{out}(t)|^2 dt}{\int_{-t_0}^{t_0} |f_{in}(t)|^2 dt} \tag{9}
\]

where \( f_{out}(t) = E(z = L, t > t_H) \) is the output field and \( f_{in}(t) = E(z = -L, t < 0) \) is the input field. We solve for \( f_{out}(t) \) both numerically and analytically, and consider the effects of diffusion in axial (longitude) and radial (transverse) directions separately.

### A. Longitudinal diffusion

For a uniform plane wave, transverse diffusion is irrelevant. We replace \( r \) by \( z \) in Eq. (8) and consider the longitude diffusion in a 1-dimensional model. Now the Maxwell-Bloch equation is

\[
\dot{\sigma}_{12}(z, t) = i \frac{g\Omega_c}{\Delta} e^{i(k_0-k_z)z} E(z, t) - i\delta(z, t)\sigma_{12}(z, t) + D \nabla^2 \sigma_{12}(z, t),
\]

\[
\frac{\partial}{\partial z} E(z, t) = i \frac{gN\Omega_c e^{-i(k_0-k_z)z}}{c\Delta} \sigma_{12}(z, t) + i \frac{g^2N}{c\Delta} E(z, t). \tag{10}
\]

We now investigate the longitude diffusion effects during the write-in process, the hold time and the read-out processes separately.

To compute \( f_{out}(t) \), we evolve Eq. (10) using \( \eta \) as in Fig. 3 (a). Following the method given in \[16\], we first propagate \( E(z = 0, t) \) and \( \sigma_{12}(z \leq L, t) \) forward with boundary condition \( E(z = -L, t < 0) = f_{in}(t) \) to find their values at time \( t_H \). Then we propagate \( E \) and \( \sigma_{12} \) backward to time \( t_H \), with final condition \( E(z = L, t > t_H) = f_{out}(t) \), and solve for \( f_{out}(t) \) by matching the two solutions at time \( t_H \).

Write: Consider the diffusion effects during the write-in process, we find that (see Appendix A)

\[
f_{out}(t_H + t) = e^{D\frac{\partial}{\partial z} e^{-2(k_0-k_z)t_H}} f_{in}(-t) \tilde{G} \tag{11}
\]

where \( k_1 = \frac{g^2N}{c\Delta} + k_0 - k_c - \frac{\beta}{\tau} \) is the initial spatial frequency of \( \sigma_{12}(z, t) \), and

\[
\tilde{G} = |\eta L^3 (t + \frac{\beta}{\eta L})|^{-\frac{1}{2}} e^{\frac{1}{2} \frac{\beta k_0^2 N}{\Delta}} e^{-\frac{1}{2} \frac{\beta k_c^2}{\eta L^3}} \left( \frac{1}{\Gamma(i\beta)}/\Gamma(-i\beta) \right).
\]

is a phase factor, with \( \eta_{eff} = g\Omega_c/\Delta \) and \( \beta = \frac{g^2N}{\eta L} \).

For a pulse with Gaussian temporal profile \( f_{in} = Ae^{-(t+t_H)^2/2T_H^2} \), we find

\[
\varepsilon_W = \frac{\int_{-t_0}^{t_H} dt e^{-2(k_0+k_c)t_H^2} e^{2D\frac{\partial}{\partial z} e^{-2(k_0+k_c)t_H^2}}}{\int_{-t_0}^{t_0} dt e^{-2(k_0+k_c)t_H^2}} \tag{12}
\]

Typically, \( D\eta^2T_H^2 \) is very small, then the efficiency is

\[
\varepsilon_W = \sqrt{\alpha_W} e^{-\tau_W} + O(D^2\eta^6T_H^6) \tag{13}
\]

where \( \alpha_W = \frac{1}{1 - D\eta^2T_H^2(k_0/k_1-\eta/t_H)} \) and \( \tau_W = \frac{2D^2\eta^2}{3} \left( \frac{k_0}{\eta} \right)^3 - \left( \frac{k_0}{\eta} - t_H \right)^3 \) are dimensionless parameters. For typical experimental parameters, \( \alpha_W \approx 1 \), then

\[
\varepsilon_W \approx e^{-\tau_W} \tag{14}
\]
We also numerically solve Eq. (10) with diffusion during the write-in process, using XMDS [20]. We calculate the efficiency for different values of the diffusion rate \( D \), input time \( t_{in} \) etc. The results are shown in Fig. 2 (points are numerical results, and the curve is Eq. (14)). We plot the efficiency \( \varepsilon_W \) with respect to the rescaled dimensionless parameter \( \tau_W \), so all the points with different parameters collapse on a single curve.

**Hold:** During the storage time \([0, t_H] \), we find (see Appendix A)

\[
f_{out}(t_H + t) = e^{-D t_H (k_i - \alpha)^2} f_{in}(-t) \bar{G}
\]

(15)

For the above Gaussian shape input, the efficiency is given by

\[
\varepsilon_H = \sqrt{\alpha_H} e^{-2\alpha_H \tau_H},
\]

(16)

where \( \alpha_H = \frac{k_i^2}{\tau_H^2} \left( \frac{1}{\tau_H^2} + D t_H \eta^2 \right) \) and \( \tau_H = D t_H k_i^2 \) are dimensionless parameters, with \( k_i = k_i - \eta t_{in} \). For typical experimental parameters, \( \alpha_H \approx 1 \) and we have

\[
\varepsilon_H \approx e^{-2\tau_H}.
\]

(17)

We also numerically solve Eq. (10) with diffusion during hold time, using XMDS. We calculate the efficiency for different values of the diffusion rate \( D \), storage time \( t_H \) etc. The results are shown in Fig. 3 (points are numerical results, and the curve is Eq. (14)). We plot the efficiency \( \varepsilon_H \) with respect to the rescaled dimensionless parameter \( \tau_H \), so all the points with different parameters collapse on a single curve.

**Read:** the diffusion effects during the read-out process are the same as the diffusion effects of the write-in process (see the appendix A), so we simply have

\[
\varepsilon_R = \varepsilon_W.
\]

**B. Transverse diffusion**

We now quantify the effects of diffusion for a beam with realistic transverse Gaussian profile. The efficiency for a 3-dimensional model is defined as

\[
\varepsilon = \frac{\int |f_{out}(x, y, t_H + t)|^2 dx dy dt}{\int |f_{in}(x, y, t)|^2 dx dy dt}
\]

(18)

Eq. (8) can be solved in Fourier space \( k_x, k_y \), and also notice that

\[
\varepsilon = \frac{\int |f_{out}(k_x, k_y, t_H + t)|^2 dk_x dk_y dt}{\int |f_{in}(k_x, k_y, t)|^2 dk_x dk_y dt}
\]

(19)

Eq. (8) can be reduced to a quasi-1D problem, and can be solved as before (see Appendix B). For transverse diffusion, the output pulse will be

\[
f_{out}(k_x, k_y, t_H + t) = e^{-2\gamma_k t} e^{-\gamma_k t_H} f_{in}(k_x, k_y, -t) \bar{G}.
\]

(20)

where \( \gamma_k = D(k_i^2 + k_k^2) \).

If the input pulse has both Gaussian temporal and transverse profile

\[
f_{in}(x, y, t) = A e^{-(x^2 + y^2)/a^2} e^{-(t + t_{in})^2)/(\tau_t^2) \]

then \( \gamma_k t_p \sim D t_p / a^2 \), which is typically small. Thus the memory efficiency is

\[
\varepsilon_{\perp} = \frac{1}{1 + \tau_{\perp}} + O(\gamma_k^2 t_p^2),
\]

(21)

where \( \tau_{\perp} = 4D(t_H + 2t_{in}) / a^2 \) is a dimensionless parameter.

We also numerically solve Eq. (8) with \( \nabla^2 = \nabla_x^2 + \nabla_y^2 \). We calculate the efficiency for different values of \( a, t_H \) etc. The results are shown in Fig. 4 (points are numerical results, and curve is Eq. (21)). We plot the efficiency \( \varepsilon_{\perp} \)
with respect to the rescaled dimensionless parameter $\tau_\perp$, so all the points with different parameters collapse on a single curve.

### C. Total diffusion

Experimentally, longitude and transverse diffusion co-exist during the whole process. Combining all the diffusive contributions mentioned above, we get the output field as (Appendix B)

$$f_{\text{out}}(k_x, k_y, t_H + t) = e^{-\frac{2D}{\hbar}(k_x^2 - (k_i - \eta t)^2)} e^{-D t_H (k_x - \eta t)^2} \times e^{\frac{1}{\alpha_H} \frac{1}{\alpha_W} - 2 \gamma t - \gamma t_H} f_{\text{in}}(k_x, k_y, -t) G.$$  

We consider input pulse with both Gaussian temporal and transverse profile as above, typically, $D\hbar^2/\alpha_p$, $\gamma\hbar t_p$ are very small. Then the total efficiency will be

$$\varepsilon_{\text{tot}} = \sqrt{\frac{1}{1/\alpha_H + 2/\alpha_W - 2}} e^{-2\gamma t} e^{-2\gamma t_H} \frac{1}{1 + \tau_\perp} + O[(D\hbar^2/\alpha_p, \gamma\hbar t_p)^2]$$

Typically, $\alpha_H \simeq 1$, $\alpha_W \simeq 1$, so we have

$$\varepsilon_{\text{tot}} \simeq \varepsilon_W \times \varepsilon_H \times \varepsilon_R \times \varepsilon_\perp.$$  

### D. Efficiency optimization and estimation

Our model did not examine other decoherence processes, such as control field-induced scattering and ground state decoherence. Our results simply quantify the effects of motional diffusion on GEM efficiency, and therefore the represent upper estimates for the performance of GEM.
From Eqs. (18, 20), we have higher order Hermite-Gaussian mode transverse profile. We then have $\varepsilon_H = \sqrt{\alpha}$ and $\varepsilon_W = e^{-\Delta t} t_{in}^3/\Delta$. Including transverse diffusion, we get the total efficiency for input field with transverse Gaussian profile

$$\varepsilon_{tot} = e^{-\Delta t} t_{in}^3/\Delta \sqrt{1 + 4D(t_H + 2t_{in})/a^2}$$  \hspace{1cm} (25)

The efficiency can be improved further by choosing a larger transverse width $a$, i.e., the effects of transverse diffusion will be reduced by using a smooth field in the transverse direction.

We note that the circumstances in which a GEM will be useful are those for which all dephasing, including that due to diffusion, is small. In this limit, a useful approximate expression for the GEM efficiency is given by

$$\varepsilon_{tot} \approx 1 - \frac{4Dk_t^2 t_{in}}{3} - \frac{DT_{\parallel}\eta^2 t_p^2}{2} - \frac{4D(t_H + 2t_{in})}{a^2}$$  \hspace{1cm} (26)

as the inefficiencies arising from each diffusive process considered above add together.

Experimental considerations give estimates of the achievable GEM efficiency. In particular, to ensure the bandwidth of the memory is large enough to absorb the achievable GEM efficiency. In particular, to ensure the pulse duration is large enough to absorb the achievable GEM efficiency. In particular, to ensure the pulse duration 2$t_p$ and the vapour length $L$ and beam width $a$. With a Gaussian transverse profile, we find that $k_H = 0, t_{in} > t_p$ to ensure that the whole pulse enters the medium during the write-in process, also $|k_i| > \frac{L}{a}$ is required to satisfy $k_H = 0$. So

$$\varepsilon_{tot} \approx 1 - \frac{4Dk_t^2 t_{in}}{3} - \frac{DT_{\parallel}\eta^2 t_p^2}{2} - \frac{4D(t_H + 2t_{in})}{a^2}$$  \hspace{1cm} (27)

This gives a reasonable upper bound on the GEM efficiency, given the pulse duration $2t_p$, the hold time $t_H$, and the vapour length $L$ and beam width $a$.

**Experimental considerations:** In experiments reported in [14, 21], Rb$^7$ atoms were used. Typical system parameters are $\omega_0 = 2\pi \cdot 377, 10746$ THz, $\omega_0 - \omega_p = 2\pi \cdot 6.8$ GHz, $\Delta = -2\pi \cdot 1.5$ GHz, $\Omega_r \approx 2\pi \cdot 20$ MHz, $\gamma \approx 2\pi \cdot 4.5$ Hz, $t_p = 1\mu s$, $a \approx 1.45$ mm, $2L = 0.2$ m, $\eta \approx -2\pi \cdot 0.1$ MHz/m, $N \approx 5 \times 10^{18}$ m$^{-3}$ [14, 21, 22]. The optical depth $|\beta|$ is sufficiently large. According to the formula in [23], we have $D \sim 0.004$ m$^2$/s for Rubidium atoms in buffer gas [14, 21].

With these parameters, the diffusive decay will be dominated by transverse diffusion. For example, for $t_{in} = 5\mu s$ and $t_H = 0$, the maximum achievable efficiency in $\varepsilon_{tot} \approx 93\%$ ($\varepsilon_{\parallel} \approx \varepsilon_{tot}, \varepsilon_W \approx 1, \varepsilon_H \approx 1$). We examine the input of the $(1,1)$ Hermite-Gaussian mode $f_{in}(x, y, t) \propto xy e^{-(x^2 + y^2)/a^2}$ as an example of a higher order Hermite-Gaussian mode transverse profile. From Eqs. (18, 20), we have $\varepsilon_{ij} = (1 + \mu)^3$, and the longitude diffusion effects are the same as the Gaussian profile (i.e. the $(0,0)$ Hermite-Gaussian mode). Thus, for diffusive decays, we have

$$\varepsilon_{ij} \propto \frac{1}{1 + \mu L}$$  \hspace{1cm} (28)

where $\varepsilon_{ij}$ is the read-out efficiency for $(ij)$ Hermite-Gaussian mode. We find that the efficiency decays faster for higher order modes, and the ratio Eq. (28) decreases when the storage time increases. This is in agreement with experimental investigations [21].

**E. Output beam width**

After some storage time, transverse diffusion will tend to smear the spin wave density in the radial direction. Intuitively, we would expect this to lead to a spatially wider output beam than would be the case in the absence of diffusion.

This is certainly the case when the control field is radially uniform. To see this, we define the intensity distribution for the read-out signal as

$$I(r_{\perp}) = \int |f_{out}(r_{\perp}, t_H + t)|^2 dt.$$  \hspace{1cm} (29)

We suppose that the control field is turned off during the hold time, $[0, t_H]$ to avoid control field-induced scattering, and that the gradient is always on and flipped at $t = 0.5t_H$. Also for typical experimental parameters, the effects of longitudinal diffusion is very weak, so we focus on transverse diffusion. We solve the Maxwell-Bloch equation using the same method as before. For a signal with a Gaussian transverse profile, we find that

$$I(r_{\perp}) \propto e^{r_{\perp}^2/[a^2 + 4D(2t_{in} + t_H)]},$$  \hspace{1cm} (30)

with $r_{\perp}^2 = x^2 + y^2$. Defining $w_{r_{\perp}}$ as the width of the output field, we have

$$w_{r_{\perp}}^2 = \frac{a^2}{4} + D(2t_{in} + t_H),$$  \hspace{1cm} (31)

which increases linearly with storage time (Fig. 6), at a rate determined by the diffusion coefficient.

Somewhat surprisingly, the experimentally measured rate of expansion of the read-out signal is smaller than that expected from atomic diffusion by a factor of 2 to 3 [21]. One possible explanation for this is the signal diffraction as suggested in [21], diffusion leads to a beam with reduced divergence and the measurement is taken downstream. However in this experiment the scale of experimental setup is much smaller than the Rayleigh range, so the diffraction effect is too small to explain the observed discrepancy.

Instead, we find that the anomalously narrow output beam width can be explained by considering the control field with realistic transverse Gaussian profile. This leads to a transverse variation in the phase of the spin
The extra phase \( \theta \) for control field with Gaussian profile. Points are numerical results using typical parameters given in the main text, and the curve is the approximate expression in Eq. (32).

The intensity distribution for the read-out signal with \( t_H = 16 \) \( \mu s \). To see the expansion clearly, we have renormalized the maximum of \( I(r_{\perp}) \) to 1, and \( I_0 \) is the renormalized intensity distribution. The black solid curve is input signal, the blue dotted one is the read-out signal for homogeneous control field, and the red dashed is read-out signal for control field with Gaussian profile.

The intensity distribution for the read-out signal. Circles are numerical results for homogeneous control field. Squares are numerical results for control field with Gaussian profile, the solid line is Eq. (31) for homogeneous control field.

give rise to a radially dependent phase on the spin wave, \( e^{i\theta(r_{\perp})} \), with the effect of the control field typically being dominant. We compare solutions for this inhomogeneous control field with solutions for the homogeneous control field to obtain the phase difference \( \theta(r_{\perp}) \) during hold time. Typically, the width of the control field, \( w_c \), is much larger than the width of the signal field, \( a \), so \( \theta(r_{\perp}) \) is approximately quadratic in \( r_{\perp}/w_c \):

\[
\theta(r_{\perp}) = \left[ -\frac{2\Omega^2 t_{in}}{\Delta} + 2\beta \ln \left( \frac{\eta L t_{in}}{\beta} + 1 \right) \right] + 2\beta \left( 1 - \frac{\Omega^2}{\Delta \eta L} \right) \left( \frac{\beta}{\eta L t_{in} + \beta} + \frac{r_{\perp}^2}{w_c^2} \right) \tag{32}
\]

where \( \beta \) is the optical depth corresponding to \( \Omega \). Because of this quadratic phase variation across the spin wave, diffusion acts to wash out the spin-wave coherence more quickly at larger radius, so the read-out efficiency is suppressed at larger \( r_{\perp} \). This will tend to reduce the apparent width of the emitted read-out signal.

**Experimental considerations**: In the experimental results reported in [21], \( w_c \simeq 3 \) mm, and \( t_{in} \simeq 2 \) \( \mu s \). Using these parameters, Fig. [3] shows the transverse variation in the phase of the spin-wave at \( (z = 0, t = 0.5 t_H) \) in the absence of diffusion. When diffusion is introduced, this transverse phase variation is smeared out, leading to reduced read-out efficiency in the wings of the spin-wave. Figure [4] compares the numerical results for the expansion of the read-out signal after a specific hold time, \( t_H = 16 \) \( \mu s \), with a homogeneous control field (dotted, blue) and with a spatially varying control field (dashed, red), assuming the diffusion rate \( D = 0.004 \) \( m^2/s \). Figure [5] shows the variation in the width of the output field as a function of hold time. We see that the expansion is slowed for a control field with Gaussian profile (squares), compared to the case of a uniform control field (circles).
Importantly, this corresponds to a reduction of the beam width expansion-rate by a factor of 2. The apparent diffusion rate extracted from this slower expansion rate is $D_{\text{eff}} \approx 0.002 \text{ m}^2/\text{s}$. This is quantitatively in agreement with the observations in [21].

V. SUMMARY

We have studied the effects of diffusion on the efficiency of the A-gradient echo memory, both numerically and analytically. We find that the efficiency is dependent on the spatial frequencies $k$ for both longitudinal diffusion and transverse diffusion: higher $k$ leads to more pronounced diffusive effects, and reduced efficiency, as expected. We show that the storage efficiency can be improved by appropriate choice of the gradient during the hold phase.

We established a mechanism by which the rate of expansion of the transverse width of the beam is reduced, compared to the naive expectation of diffusive effects. This mechanism arises from the effects of diffusion on the transverse variation in the spin wave phase. We showed that with an experimentally reasonable choice of parameters, the magnitude of this effect is the same as that observed in recent experiments. When the density of the buffer gas is increased, the collision rate increases, leading to a smaller diffusion rate. However, this will lead to collision-induced dephasing, which will dominate at sufficiently high buffer gas pressures. This implies a trade off between diffusion- and collision-induced dephasing. This will be the subject of future research.

ACKNOWLEDGMENTS

J. Hope and B. Hillman thank M. Hosseini, D. Higginbottom, O. Pinel and B. Buchler for helpful discussions of the modelling and experiments. J. Hope was supported by the ARC Future Fellowship Scheme. X.-W. Luo thanks G. J. Milburn for helpful discussions, and gratefully acknowledges the National Natural Science Foundation of China (Grants No. 11174270), the National Basic Research Program of China (Grants No. 2011CB921204), CAS for financial support, and The University of Queensland for kind hospitality.

[1] L.M. Duan, M. Lukin, I. Cirac, and P. Zoller. Long-distance quantum communication with atomic ensembles and linear optics. *Arxiv preprint quant-ph/0105105*, 2001.

[2] E. Knill, R. Laflamme, G.J. Milburn, et al. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46-52, 2001.

[3] M. Fleischhauer, A. Imamoglu, and J.P. Marangos. Electromagnetically induced transparency: Optics in coherent media. *Reviews of Modern Physics*, 77(2):633, 2005.

[4] M. Fleischhauer and MD Lukin. Dark-state polaritons in electromagnetically induced transparency. *Physical review letters*, 84(22):5094–5097, 2000.

[5] J. Nunn, IA Walmsey, MG Raymer, K. Surnacz, FC Waldermann, Z. Wang, and D. Jaksch. Modematching an optical quantum memory. *Arxiv preprint quant-ph/0503068*, 2005.

[6] J. Nunn, IA Walmsey, MG Raymer, K. Surnacz, FC Waldermann, Z. Wang, and D. Jaksch. Mapping broadband single-photon wave packets into an atomic memory. *Physical Review A*, 75(1):011401, 2007.

[7] B. Kraus, W. Tittel, N. Gisin, M. Nilsson, S. Kröll, and JI Cirac. Quantum memory for nonstationary light fields controlled on reversible inhomogeneous broadening. *Physical Review A*, 73(2):020302, 2006.

[8] N. Sangouard, C. Simon, M. Afzelius, and N. Gisin. Analysis of a quantum memory for photons based on controlled reversible inhomogeneous broadening. *Physical Review A*, 75(3):032327, 2007.

[9] M. Afzelius, C. Simon, H. De Riedmatten, and N. Gisin. Multimode quantum memory based on atomic frequency combs. *Physical Review A*, 79(5):052329, 2009.

[10] G. Hetet, JJ Longdell, AL Alexander, PK Lam, and MJ Sellars. Gradient echo quantum memory for light using two-level atoms. *Arxiv preprint quant-ph/0612169*, 2006.

[11] G. Hetet, JJ Longdell, AL Alexander, P.K. Lam, and MJ Sellars. Electro-optic quantum memory for light using two-level atoms. *Physical review letters*, 100(2):23601, 2008.

[12] G. Hétet. Quantum memories for continuous variable states of light in atomic ensembles. PhD thesis, Australian National University, 2008.

[13] A.I. Lvovsky, B.C. Sanders, and W. Tittel. Optical quantum memory. *Nature Photonics*, 3(12):706–714, 2009.

[14] M. Hosseini, BM Sparkes, G. Campbell, PK Lam, and BC Buchler. High efficiency coherent optical memory with warm rubidium vapour. *Nature communications*, 2:174, 2011.

[15] O. Firstenberg, M. Shuker, R. Pugatch, DF Redkin, N. Davidson, and A. Ron. Theory of thermal motion in electromagnetically induced transparency: Effects of diffusion, doppler broadening, and dicer and ramsey narrowing. *Physical Review A*, 77(4):043830, 2008.

[16] O. Firstenberg, M. Shuker, N. Davidson, and A. Ron. Elimination of the diffraction of arbitrary images imprint on slow light. *Physical review letters*, 102(4):43601, 2009.

[17] O. Firstenberg, P. London, D. Yankelev, R. Pugatch, M. Shuler, and N. Davidson. Self-similar modes of coherent diffusion. *Physical review letters*, 105(18):183602, 2010.

[18] Daniel F Walls and Gerard J Milburn. *Quantum optics*. Springer, 2008.

[19] TM Stace and AN Luiten. Theory of spectroscopy in an optically pumped effusive vapor. *Physical Review A*, 81(3):033848, 2010.
APPENDIX A

The Maxwell-Bloch equation for the 1-dimensional model is
\[
\dot{\sigma}_{12}(z,t) = i \frac{g \Omega_c}{\Delta} e^{i(k_0 - k_c)z} E(z,t) - i \alpha \dot{E}(z,t) \sigma_{12}(z,t) + D \nabla^2 \sigma_{12}(z,t),
\]
\[
\frac{\partial}{\partial z} E(z,t) = i \frac{g \Omega_c}{c \Delta} e^{-i(k_0 - k_c)z} \sigma_{12}(z,t) - \frac{i g^2 N}{c \Delta} E(z,t).
\]

To find the solution during \([-t_0,0]\), we first solve the equation without diffusion, then introduce the diffusion effects to our solutions.

When \(D = 0\), we can make transformation
\[
\tilde{\sigma}_{12}(z,t) = e^{-i \frac{2N}{L} z} e^{-i(k_0 - k_c)z} \sigma_{12}(z,t),
\]
\[
\tilde{E}(z,t) = e^{-i \frac{2N}{L} z} E(z,t),
\]
and get the new equations
\[
\partial_z \tilde{E}(z,t) = i \frac{g_{\text{eff}} N}{\eta} \tilde{\sigma}_{12}(z,t),
\]
\[
\partial_z \tilde{\sigma}_{12}(z,t) = - i \eta z \tilde{\sigma}_{12}(z,t) + i g_{\text{eff}} \tilde{E}(z,t)
\]
where \(g_{\text{eff}} = g \Omega_c / \Delta\). Following the method given in [10], and using the boundary conditions \(\tilde{\sigma}_{12}(z,t \to -\infty) = 0\) and \(\tilde{E}(z = -L, t < 0) = \tilde{f}_\text{in}(t)\), we integrate the first equation and substitute it in the second one. Making use of Fourier transformation, we find
\[
\tilde{E}(k,t) = \tilde{f}_\text{in}(t) \frac{k}{\eta} + \frac{\beta}{\eta L} t |k| [-i \beta - 1] G(\eta, \beta, L),
\]
\[
G(\eta, \beta, L) = \frac{1}{\eta \beta} e^{-\pi |\beta|^2/2 \sinh(\pi |\beta|)} \eta L |^{-i \beta} \Gamma(i \beta),
\]
where \(\tilde{E}(k,t) = \int \tilde{E}(z,t) e^{-ikz} dz\), \(\beta = g_{\text{eff}} N / \Delta\) is the optical depth and we assume \(\beta\) is sufficiently large, \(\Gamma(i \beta)\) is the Gamma Function, \(\tilde{f}_\text{in}(t) = \tilde{f}_\text{in}(t) e^{i \frac{2N}{L} k_0 t}\) is the input pulse. According to the Maxwell-Bloch equations, we have \(\tilde{\sigma}_{12}(k,t) = \frac{k}{g_{\text{eff}} N} \tilde{E}(k,t)\).

We transform \(\tilde{\sigma}_{12}(k,t)\) back to \(\sigma_{12}(k,t)\),
\[
\sigma_{12}(k,t) = f_{\text{in}}(t) \frac{k - k_i}{\eta} + t e^{\frac{2N}{L} k} L
\]
\[
\times \left| \frac{k - k_i}{\eta} - \frac{\beta}{\eta L} \right|^{-i \beta} \text{sgn} \left( \frac{k - k_i}{\eta} - \frac{\beta}{\eta L} \right) \frac{e^c}{g_{\text{eff}} N} G
\]
with \(k_i = \frac{2N}{\Delta} x + k_0 - k_c - \frac{\beta}{L}\).

Now we introduce the diffusion, for the short time interval \([t, t + \Delta t]\), diffusion will cause a decay \(e^{-D k^2 \Delta t}\) to
Thus, the solution for \( \sigma_{12} \) at \( t = 0 \) is

\[
\sigma_{12}(k,0) = e^{-\frac{\eta}{2}(k^3-k^3)} f_{\text{in}}(k) e^{\frac{\eta}{2} z^N L} \times \left( \frac{k-k_\eta}{\eta} - \frac{\beta}{\eta L} \right)^{-1} \frac{\eta}{\eta L} \sigma_{12}(k,0) \quad (38)
\]

where \( k = \frac{\eta}{\beta L} + k_\eta - k_c \). Notice that \( \sigma_{12}(k,t_H) \) get a phase \( e^{\frac{\eta}{2} z^N L} \frac{1}{k^3-k_c^3} \), so the group velocity for \( \sigma_{12}(z,t) \) is \( v_g(k) = \frac{\eta}{\beta L} \frac{1}{k^3-k_c^3} \). If the memory broadening \( \eta L \) is not much larger than the signal pulse bandwidth, the spin wave \( \sigma_{12}(z,t) \) will be nonzero near the ensemble boundary. Then the spin wave will propagate to the boundary and be reflected, this may ruin the spin wave coherence near the boundary and lower the memory efficiency. One way to avoid this effect is turning off the control field during storage, which makes the effective coupling \( g_{\text{eff}} = 0 \), and the group velocity \( v_g = 0 \).

To find the values for \( \sigma_{12} \) and \( E \) in the duration \([t_H, t_H + t_0]\), one needs to solve a modified version of Eq. (38), where the sign of \( \eta z \) is reversed. We follow the method given in [10], propagate this equation backwards with final conditions \( E(z = L, t > t_H) = f_{\text{out}}(t), \sigma_{12}(z,t \rightarrow \infty) = 0 \). Similar to the write-in process, at time \( t_H \), we have

\[
\sigma_{12}(k,t_H) = e^{-\frac{\eta}{2}(k^3-k^3)} f_{\text{out}}(t_H) + \frac{k-k_\eta}{\eta L} e^{-\frac{\eta}{2} z^N L} \times \left( \frac{k-k_\eta}{\eta} - \frac{\beta}{\eta L} \right)^{-1} \frac{\eta}{\eta L} \sigma_{12}(k,0) \quad (39)
\]

By matching the two solutions for \( \sigma_{12} \) at \( t_H \) Eqs. (39), (40), we get

\[
f_{\text{out}}(t_H + t) = d_W(t) d_H d_R(t) f_{\text{in}}(-t) \tilde{G} \quad (41)
\]

where

\[
\tilde{G} = \frac{\eta L \left( t + \frac{\beta}{\eta L} \right)}{1 + \frac{\beta}{\eta L}} e^{-\frac{\eta}{2} z^N L} e^{-\frac{\eta}{2} z^N L} \left( \frac{i\beta}{\eta L} \right)^{t_H} \Gamma(i\beta)/\Gamma(-i\beta)
\]

is a phase factor, \( d_W(t) = e^{-\frac{\eta}{2}(k^3-k^3)^2 t_H} \), \( d_H = e^{-\frac{\eta}{2}(k^3-k^3)^2 t_0} \) and \( d_R(t) = e^{-\frac{\eta}{2}(k^3-k^3)^2 t_H} \) are the diffusion decays for the write-in process \([-t_0,0]\), storage time \([0,t_H]\) and read-out process \([t_H,t_H + t_0]\) respectively.

**APPENDIX B**

The Maxwell-Bloch equation for the 3-dimensional model is

\[
\dot{\sigma}_{12}(r,t) = i \frac{g N_\Omega}{\Delta} e^{-i(k_0-k_c)z} E(r,t) - (i\eta z) \sigma_{12}(r,t) + \nabla^2 \sigma_{12}(r,t),
\]

\[
\frac{\partial}{\partial z} E(r,t) = i \frac{\gamma N_\Omega c}{\Delta} e^{-i(k_0-k_c)z} \sigma_{12}(r,t) + i \frac{g^2 N_c}{\Delta} E(r,t).
\]

To solve these equations, we first transform transverse coordinates \( x, y \) to Fourier space \( k_x, k_y \),

\[
\dot{\sigma}_{12}(k_x,k_y,z,t) = - (i\eta z + \gamma_k) \sigma_{12}(k_x,k_y,z,t)
\]

\[
+ i \frac{g N_\Omega}{\Delta} e^{-i(k_0-k_c)z} E(k_x,k_y,z,t)
\]

\[
+ \nabla^2 \sigma_{12}(k_x,k_y,z,t),
\]

\[
\frac{\partial}{\partial z} E(k_x,k_y,z,t) = i \frac{\gamma N_\Omega c}{\Delta} e^{-i(k_0-k_c)z} \sigma_{12}(k_x,k_y,z,t)
\]

\[
+ i \frac{g^2 N_c}{\Delta} E(k_x,k_y,z,t),
\]

where \( \gamma_k = D(k_x^2 + k_y^2) \). Now we make the following transformation:

\[
\dot{\sigma}_{12}(k_x,k_y,z,t) = e^{\gamma t} \sigma_{12}(k_x,k_y,z,t),
\]

\[
\dot{E}(k_x,k_y,z,t) = e^{\gamma t} E(k_x,k_y,z,t),
\]

then we have

\[
\dot{\sigma}_{12}(k_x,k_y,z,t) = - (i\eta z + \gamma_k) \dot{\sigma}_{12}(k_x,k_y,z,t)
\]

\[
+ i \frac{g N_\Omega}{\Delta} e^{-i(k_0-k_c)z} \dot{E}(k_x,k_y,z,t)
\]

\[
+ \nabla^2 \dot{\sigma}_{12}(k_x,k_y,z,t),
\]

\[
\frac{\partial}{\partial z} \dot{E}(k_x,k_y,z,t) = i \frac{\gamma N_\Omega c}{\Delta} e^{-i(k_0-k_c)z} \dot{\sigma}_{12}(k_x,k_y,z,t)
\]

\[
+ i \frac{g^2 N_c}{\Delta} \dot{E}(k_x,k_y,z,t).
\]

These are actually quasi-1D equations, so we can solve these equations by the method we used before, and the output field is:

\[
\dot{f}_{\text{out}}(k_x,k_y,t_H + t) = d_W(t) d_H d_R(t) \tilde{f}_{\text{in}}(k_x,k_y,-t) \tilde{G}.
\]
We transform back to $f_{\text{out}}(k_x, k_y, t_H + t)$, and get

$$f_{\text{out}}(k_x, k_y, t_H + t) = dW(t) dH dR(t) \times d_{\perp}(t) f_{\text{in}}(k_x, k_y, -t) \bar{G}, \quad (46)$$

where $d_{\perp}(t) = e^{-2\gamma t} e^{-\gamma t_H}$ is the transverse diffusion decay.