Spin-mediated Mott excitons

T.-S. Huang,1, 2 C. L. Baldwin,1, 3 M. Hafezi,1 and V. Galitski1, 2

1Joint Quantum Institute, University of Maryland, College Park, MD 20742, USA
2Condensed Matter Theory Center, University of Maryland, College Park, MD 20742, USA
3National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

(Dated: April 24, 2020)

Motivated by recent experiments on Mott insulators, in both iridates and ultracold atoms, we theoretically study the effects of magnetic order on the Mott-Hubbard excitons. In particular, we focus on spin-mediated doublon-holon pairing in Hubbard materials. We use several complementary theoretical techniques: mean-field theory to describe the spin degrees of freedom, the self-consistent Born approximation to characterize individual charge excitations across the Hubbard gap, and the Bethe-Salpeter equation to identify bound states of doublons and holons. The binding energy of the Hubbard exciton is found to increase with increasing the Néel order parameter, while the exciton mass decreases. We observe that these trends rely significantly on the retardation of the effective interaction, and require consideration of multiple effects from changing the magnetic order. Our results are consistent with the key qualitative trends observed in recent experiments on iridates. Moreover, the findings could have direct implications on ultracold atom Mott insulators, where the Hubbard model is the exact description of the system and the microscopic degrees of freedom can be directly accessed.

I. INTRODUCTION

The physics of excitons in semiconductors, i.e., bound states of electrons and holes, is by now well-established [1, 2]. Excitons play an essential role in technologies such as light-emitting diodes [3], organic solar cells [4], and photodetectors [5], among others. Furthermore, there has recently been tremendous interest in hybridizing exciton states with photon modes in optical cavities [6]. Such exciton polaritons can form (non-equilibrium) Bose-Einstein condensates at remarkably high temperatures, even room temperature [7–11].

Given these applications, it is important to study the properties of excitons in systems other than conventional semiconductors. It has been convincingly established that excitonic states do exist in strongly correlated materials such as Mott insulators [12–21], yet essential aspects of Mott excitons remain poorly understood.

For example, it is known that Mott insulators are often antiferromagnetic at low temperature, but very little work has been done to understand how and to what extent the presence of such order affects exciton properties. The qualitative role of magnetization is sketched in Fig. 1 – charges remain bound so as to minimize the number of spins disrupted by their motion – but a quantitative description has been lacking.

Recent experiments have begun to investigate this question. In Refs. [22, 23], pump-probe experiments were performed on the Mott insulator Na$_2$IrO$_3$ both with and without magnetic order (controlled by varying temperature or applying an intermediate pulse). The authors concluded that the binding energy and exciton mass are both enhanced by the presence of magnetization. Ref. [24] similarly observed that the binding energy increases with the spin-spin interaction strength in cuprates. See also Ref. [25], which found that the relaxation time in Mott insulators decreases with increasing spin correlations.

The same question can apply to Mott insulators in synthetic quantum systems, such as ultracold gases. By loading fermionic atoms into an optical lattice and tuning their interactions, the Fermi-Hubbard model can be synthesized experimentally [26–30]. Unlike condensed matter systems, such as the iridates, neutral fermionic atoms in an optical lattice are genuinely described by the Hubbard Hamiltonian, without any additional effects arising from longer-range Coulomb interactions, phonons, etc. Researchers have quite recently begun investigating the interplay of spin and charge degrees of freedom in this setting [31–37] (note that here the “charge” excitations are not actually charged).

In this paper, we perform a theoretical study of the role of magnetic order in Mott excitons, the first such to our knowledge. As depicted in Fig. 1 in an antiferromagnetic background, a hole and doubly occupied site can bind through a string of flipped spins. Such Mott excitons differ from conventional excitons formed...
by Coulomb interaction in two aspects. First, the spin-mediated interaction is far from instantaneous, and second, the individual charges are themselves renormalized by spin fluctuations. We shall demonstrate that both effects are necessary ingredients in the trends reported here.

Given the complexity of the problem, our analysis requires multiple stages. We first use slave particles to isolate spin and charge degrees of freedom, then describe the spin dynamics by mean-field theory, calculate the dispersion of charges self-consistently, and finally characterize excitonic states via the Bethe-Salpeter equation. Many of the steps in this program are analogous to those in Ref. [38], which studied charge dynamics in the Hubbard model. The good agreement between the results of Ref. [38] and alternate numerical methods lends support to the present approach.

Our key finding is that larger magnetization leads to an increased binding energy of the Hubbard exciton but a decreased mass. This observation is in some tension with interpretations of recent experiments [23]. It also stands in contrast to conventional Coulomb-mediated excitons, where the binding energy and mass are proportional to an increased binding energy of the Hubbard exciton but a decreased mass. This observation is in some tension with our knowledge of how the bound state properties change as a function of magnetization has not yet been carried out.

In the following Sec. [11] we describe the steps of our analysis in detail. Results are presented in Sec. [11] and Sec. [11] concludes.

II. FORMALISM & METHODS

Our starting point is the 2D Fermi-Hubbard model, which by now needs no introduction:

$$H_{\text{Hub}} = -t \sum_{\langle ij \rangle, \sigma} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}, \quad (1)$$

where $\sigma \in \{\uparrow, \downarrow\}$ and $\langle ij \rangle$ denotes nearest-neighbor sites on a square lattice. $c_{i\sigma}$ is the usual electron annihilation operator and $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$. We shall consider the system at half filling in the $U \gg t$ limit.

It is well-known that in this limit, the Hubbard model features two types of excitations, associated with the transport of charge and spin respectively [14]-[17]. Furthermore, the charge excitations can be either positive or negative, corresponding to sites with zero or two electrons, and their creation comes with a large energy cost of order $U$. By analogy with conventional semiconductors, we thus expect this system to support well-defined excitons in the dilute-charge limit. However, long-wavelength spin excitations do not come with an energy cost, and their presence plays a significant role in determining the exciton properties.

There are many formalisms with which to study the Hubbard model [38]-[51]. Since our focus is on the motion of only a few charges within a background of spin excitations, the slave-particle formalism is particularly well-suited [38]-[52]-[53]. The steps of our calculation are as follows:

i) Express the Hamiltonian in terms of slave particles – doublons, holons, & spinons – and reduce to the t-J model following the standard procedure [54].

ii) Make a mean-field approximation to the Heisenberg interaction, thus neglecting the back action of doublons and holons on the magnetic order.

iii) Calculate the dispersion of individual doublons and holons in the magnetic background via the self-consistent Born approximation.

iv) Calculate exciton properties using the Bethe-Salpeter equation.

The major limitation of this program is the use of mean field theory to describe the magnetic order. Thus we do not claim to have quantitatively accurate results for small magnetization. That said, we do expect that the qualitative trends seen here are accurate, especially near the equilibrium value of magnetization, for which mean field theory is known to work reasonably well (see Ref. [38] and references therein).

A. Slave particles

In the slave-particle formalism, we express the electron operator as ($\sigma = \pm 1$)

$$c_{i\sigma} = s_{i\sigma}^\dagger d_i + \sigma e_i^\dagger s_{i\sigma}, \quad (2)$$

where $d_i$ and $e_i$ are fermionic operators and $s_{i\sigma}$ is bosonic. One can confirm that Eq. (2) is consistent with
the commutation relations. A site with a $d$ particle is to be interpreted as a site with two electrons (a “doublon”), a site with an $e$ particle is to be interpreted as an empty site (a “holon”), and a site with an $s$ particle is one with a single electron having spin $\sigma$ (a “spinon”). See Fig. 2. The physical content of Eq. (2) is then clear: removing an electron of given spin is equivalent to replacing the doublon with the opposite spinon if the site is doubly-occupied and replacing the spinon with a holon if the site is singly-occupied (otherwise the state is annihilated). Note that since every site is in one of the four states – empty, spin-up, spin-down, doubly-occupied – there must be exactly one of the fictitious particles on each site:

$$d_i^\dagger d_i + e_i^\dagger e_i + s_i^\dagger s_i^\uparrow + s_i^\dagger s_i^\downarrow = 1, \forall i.$$  (3)

The original Hamiltonian clearly preserves this relationship.

Substituting Eq. (2) into Eq. (1), we have that

$$H_{\text{Hub}} = -t \sum_{\langle ij \rangle, \sigma} (d_i^\dagger d_j - e_i^\dagger e_j)s_i^\sigma s_j^\sigma + U \sum_i d_i^\dagger d_i$$

$$- t \sum_{\langle ij \rangle, \sigma} \sigma(d_i^\dagger e_j^\dagger s_i^\sigma s_j^\sigma + e_i^\dagger s_j^\sigma s_i^\sigma).$$  (4)

Note that the first line preserves the number of doublons and holons, whereas the second line does not.

At large $U$, the second line of Eq. (4) can be treated by perturbation theory in $t/U$. The method as applied here is standard, and can be found in, e.g., Ref. [51]. We obtain the t-J model:

$$H_{\text{t-J}} = -t \sum_{\langle ij \rangle, \sigma} (d_i^\dagger d_j - e_i^\dagger e_j)s_i^\sigma s_j^\sigma + U \sum_i d_i^\dagger d_i$$

$$- J \sum_{\langle ij \rangle} \left(s_i^\dagger s_j^\sigma + s_j^\dagger s_i^\sigma\right)(s_j^\downarrow s_i^\uparrow - s_j^\uparrow s_i^\downarrow),$$  (5)

where $J \equiv 4t^2/U$. Strictly speaking, Eq. (5) should include additional next-nearest-neighbor terms, as well as a direct interaction between nearest-neighbor doublons and holons, but these are commonly neglected.

### B. Magnetic ordering

The second line of Eq. (5) is precisely the antiferromagnetic Heisenberg Hamiltonian, expressed in terms of Schwinger bosons (here the spinons $s_{i\sigma}$) [51]. We treat that term via mean-field theory. Since we are interested in the dilute-charge limit, we neglect the charge-spinon interaction (first line of Eq. (5)) when determining the mean-field properties.

Assume Néel order on the $A$ and $B$ sublattices ($\hat{\sigma}_i^z$ is the Pauli operator on site $i$):

$$\langle \hat{\sigma}_i^z \rangle = \begin{cases} m, & i \in A \\ -m, & i \in B \end{cases}.$$  (6)

where $0 < m < 1$. In the spinon language, this corresponds to Bose condensation: we replace the operators $s_i^\uparrow$ on sublattice $A$ and $s_i^\downarrow$ on sublattice $B$ by the coefficient $b$. This leaves a single spinon operator on each site $(s_i^\uparrow$ on $A$, $s_i^\downarrow$ on $B)$, which we will denote simply by $s_i$ and keep only to second order:

$$H_{\text{t-J}} \approx -tb \sum_{\langle ij \rangle} (d_i^\dagger d_j - e_i^\dagger e_j)(s_i + s_j^\dagger) + U \sum_i d_i^\dagger d_i$$

$$-JB^2 \sum_{\langle ij \rangle} (b^2 - s_i^\dagger s_j^\dagger - s_i s_j).$$  (7)

Note that Eq. (7) no longer respects the constraint in Eq. (3), thus we must include a Lagrange multiplier to satisfy the constraint on average. After performing a Bogoliubov transformation and switching to momentum space, one finds the Hamiltonian

$$H_{\text{t-J}} \approx -tb \sum_{kq} d_{k+q}^\dagger d_k (M_{kq}b_{q} + M_{k+q,-q}b_{-q}^\dagger)$$

$$+ \frac{tb}{N} \sum_{kq} e_k^\dagger e_k (M_{kq}b_{q} + M_{k+q,-q}b_{-q}^\dagger)$$

$$+ U \sum_q d_q^\dagger d_q + \sum_{q} \omega_q b_{q}^\dagger b_{q},$$

where the sum is over the 2D Brillouin zone and $\beta_q$ is the transformed spinon operator. $N$ is the number of sites in the lattice. The frequencies $\omega_q$ and vertices $M_{kq}$ are

$$\omega_q = 8Jb^2 \sqrt{1 - \gamma_q^2},$$  (9)

$$M_{kq} = 4\gamma_k u_q - 4\gamma_{k+q} v_q,$$  (10)

where

$$\gamma_q = \frac{1}{2} \left( \cos q_x + \cos q_y \right),$$  (11)

$$u_q \equiv \sqrt{\frac{1}{2} \left( 1 + \frac{1}{\sqrt{1 - \gamma_q^2}} \right)},$$  (12)

$$v_q \equiv \text{sgn}[\gamma_q] \sqrt{\frac{1}{2} \left( 1 + \frac{1}{\sqrt{1 - \gamma_q^2}} - 1 \right)}.$$  (13)

Normally one would determine $b$ so as to minimize the ground state energy. This is known to give $b^2 \approx 0.8$ for a 2D square lattice [22]. We shall instead treat $b$ as an independent parameter and calculate exciton properties as a function of magnetization (using that $m = 2b^2 - 1$).
The charge dynamics is strongly affected by spinons but the doublon-doublon and holon-holon propagators by the integral equation in Fig. 3. The same equation holds the doublon-doublon and holon-holon propagators by the vertex, corresponding to the Hamiltonian in Eq. (8).

FIG. 3. The self-consistent Born approximation (SCBA) for the single-particle propagator (either doublon-doublon or holon-holon). The solid single line is the free propagator, in this case simply \( G^0_k(\epsilon) = 1/\epsilon \), and the solid double line is the full propagator. The dashed line is the spinon propagator and the black dot is the vertex, corresponding to the Hamiltonian in Eq. (8).

C. Self-consistent Born approximation

The self-consistent Born approximation (SCBA) gives the doublon-doublon and holon-holon propagators by the integral equation in Fig. 3. The same equation holds for each propagator separately. This approximation is expected to be accurate in the dilute-charge limit, where the charge dynamics is strongly affected by spinons but not vice-versa.

In terms of the doublon/holon self-energy \( \Sigma_k(\epsilon) \), Fig. 3 translates to (after a frequency integration)

\[
\Sigma_k(\epsilon) = \frac{t^2 b^2}{N} \sum_q \frac{M_{kq}^2}{\epsilon - \omega_q - \Sigma_{k-q}(\epsilon - \omega_q)}. \tag{14}
\]

The quasiparticle spectrum \( \epsilon_k \) is given by the solution to \( \Sigma_k(\epsilon_k) = \epsilon_k \).

Eq. (14) can be solved quite efficiently. Note that all \( \omega_q \) are positive [55], thus Eq. (14) in fact expresses \( \Sigma_k(\epsilon) \) in terms of the self-energy at lower frequencies. We start at sufficiently negative \( \epsilon \), below which we approximate \( \Sigma_k(\epsilon) \approx (t^2 b^2 / N) \sum_q M_{kq}^2 / (\epsilon - \omega_q) \), and then compute the self-energy at incrementally higher frequencies in terms of the previous values. To help avoid numerical errors, we add a small imaginary part \( 0.2 i J b^2 \) to \( \epsilon \).

While one could proceed using the full \( \Sigma_k(\epsilon) \), it has been found that the quasiparticle dispersion can be well-approximated by the form [56]

\[
\epsilon_k = - 2 t_1 (\cos(k_x + k_y) + \cos(k_x - k_y)) \\
- 2 t_2 (\cos(2k_x) + \cos(2k_y) + 2) \tag{15}
\]

This expression has a clear physical interpretation: \( t_1 \) is the amplitude for performing a two-step hop along the diagonals of the lattice, and \( t_2 \) is the amplitude for a two-step hop along the principal axes (see Fig. 4). Thus in what follows, we shall use for the single-particle propagators the simpler expression

\[
G_k(\epsilon) = \frac{1}{\epsilon - (1 - i 0)\epsilon_k}, \tag{16}
\]

with \( \epsilon_k \) given by Eq. (15).

D. Bethe-Salpeter equation

We next consider the two-particle Green’s function \( T \) denotes time ordering

\[
\mathcal{G}_{j_d k, j_d' k'} (t_d, t_e; t_d', t_e') = - \langle T d_{j_d}(t_d) e_{j_e}(t_e) e_{j_e'}(t_e') \rangle d_{j_d'}(t_d') \tag{17}
\]

and its Fourier transform \( \mathcal{G}_{k_d k, k'_d k'} (\epsilon_d, \epsilon_e; \epsilon'_d, \epsilon'_e) \). Due to translational invariance, \( \mathcal{G} \) depends only on differences in position and time, which we choose to parametrize by the relative coordinates

\[
j \equiv j_d - j_e,
\]

\[
j' \equiv j_d' - j_e',
\]

\[
r \equiv \frac{j_d + j_e}{2} - \frac{j_d' + j_e'}{2}, \tag{18}
\]

with relative times defined analogously. The corresponding momenta are

\[
k = \frac{k_d - k_e}{2},
\]

\[
k' = \frac{k_d' - k_e'}{2}, \tag{19}
\]

We will use absolute and relative momenta interchangeably, depending on notational convenience, with...
Eq. (19) always giving the relationship between the two. \( G_{k_k k'_k, \ddot{q}_k \ddot{q}'_k; \epsilon_d, \epsilon; \epsilon'_d, \epsilon'_e} \) will often be written as \( G_{k_k' k; \epsilon, \epsilon'; E} \).

Within the ladder approximation, \( G \) is determined by the integral equation of Fig. 5. Written out,

\[
G_{k_k' k; \epsilon, \epsilon'; E} = G_{k_d, \epsilon_d} G_{k_e, \epsilon_e} \left[ \delta_{k_k, k_k'} + \frac{i^2 \hbar^2}{N} \sum_q \int \frac{d \omega}{2 \pi i} \left( \frac{M_{k_d - q, q} M_{k_e, q}}{\omega - (1 - i) \omega_q} - \frac{M_{k_d - q, q} M_{k_e + q, -q}}{\omega + (1 - i) \omega_q} \right) G_{k - q, k'; \epsilon - \omega, \epsilon'; E} \right],
\]

where \( G_{k} \) is given by Eq. (16) and the vertices \( M_{kq} \) are as in Eq. (8).

Since our goal is to identify bound states, we reduce Eq. (20) to the Bethe-Salpeter equation. The details of this approach can be found in Ref. [57]. We assume that \( G \) has an isolated pole in the total energy \( E \), near which it has the form

\[
G_{k_k' k; \epsilon, \epsilon'; E} \sim -i \frac{\psi_k(\epsilon) \bar{\psi}_{k'}(\epsilon')}{E - (1 - i)E_b},
\]

where the “wavefunction” \( \psi_k(\epsilon) \), its time-reversed partner \( \bar{\psi}_k(\epsilon) \), and the bound state energy \( E_b \) remain to be determined. Inserting this ansatz into both sides of Eq. (20) and equating the residues at \( E_b \) on each side, we obtain a non-linear eigenvalue problem (the Bethe-Salpeter equation):

\[
\psi_k(\epsilon) = G_{k_d} \left( \frac{E_b}{2} + \epsilon \right) G_{k_e} \left( \frac{E_b}{2} - \epsilon \right) \frac{i^2 \hbar^2}{N} \sum_q \int \frac{d \omega}{2 \pi i} \left( \frac{M_{k_d - q, q} M_{k_e, q}}{\omega - (1 - i) \omega_q} - \frac{M_{k_d - q, q} M_{k_e + q, -q}}{\omega + (1 - i) \omega_q} \right) \psi_{k - q}(\epsilon - \omega).
\]

\( E_b \) and \( \psi_k(\epsilon) \) are given by the solution to Eq. (22). Note that they will depend on the center-of-mass momentum \( K \).

\( \psi_k(\epsilon) \) is the bound state wavefunction in a quite literal sense: it is the Fourier transform of

\[
\psi_j(t) = \langle 0 | T d_j(t) e_0(0) | b \rangle,
\]

where \( |b\) denotes the bound state and \( |0\) denotes the ground state. Note that \( t = 0 \) is of particular interest, since it gives the amplitude for simultaneously observing the holon at site 0 and the doublon at site \( j \). Thus to simplify the problem, we integrate Eq. (22) over \( \epsilon \), and furthermore, make the ansatz

\[
\psi_k(\epsilon) = -G_{k_d} \left( \frac{E_b}{2} + \epsilon \right) G_{k_e} \left( \frac{E_b}{2} - \epsilon \right) (E_b - \epsilon_{k_d} - \epsilon_{k_e}) \Psi_k,
\]

with \( \Psi_k \) independent of \( \epsilon \). The explicit factor of \( E_b - \epsilon_{k_d} - \epsilon_{k_e} \) is included so that \( \Psi_k \) is the equal-time wavefunction, i.e., \( \Psi_k = \psi_k(t = 0) \). This ansatz allows us to perform the \( \epsilon \) integral straightforwardly, giving a closed equation for \( E_b \) and \( \Psi_k \):

\[
(E_b - \epsilon_{k_d} - \epsilon_{k_e}) \Psi_k = -\frac{i^2 \hbar^2}{N} \sum_q \left( \frac{M_{k_d - q, q} M_{k_e, q}}{E_b - \epsilon_{k_d} - \omega_q - \epsilon_{k_e}} + \frac{M_{k_d - q, q} M_{k_e + q, -q}}{E_b - \epsilon_{k_d} - \epsilon_{k_e} + \omega_q} \right) \Psi_{k - q}.
\]

Eq. (25) is the two-particle Schrödinger equation albeit with an energy-dependent potential. We find the values of \( E_b \) at which it has a non-zero solution, and record the corresponding eigenvector.

Strictly speaking, Eq. (24) is not a valid ansatz for \( \psi_k(\epsilon) \), i.e., it does not solve the frequency-dependent Eq. (22). However, it has a clear physical interpretation. The Fourier transform \( \psi_k(t) \) gives the wavefunction for inserting the doublon and holon separated by time \( t \) (see Eq. (23)). The poles coming from the single-particle propagators in Eq. (24) correspond to the phase factor acquired by the remaining particle during that interval, and our ansatz amounts to neglecting any other time dependence. This approximation has been applied previously to study holon-holon binding [42], and we expect it to be qualitatively accurate for our purposes.

Eq. (25) and those preceding it differ from the equations for holon-holon binding in two respects. First, the holon-holon equations must include exchange terms not found here. Second, due to the relative phase between the doublon-spinon and holon-spinon vertices, the effective potential in Eq. (25) would have the opposite sign for the holon-holon problem.
FIG. 6. Dispersion of individual quasiparticles (both doublons and holons) within the SCBA, for a lattice of size $32 \times 32$.

III. RESULTS

A. Single-particle properties

We first review the behavior of individual quasiparticles, determined within the SCBA as described above. Although these calculations have been reported previously, e.g., in Refs. [52, 56], it will be useful to reproduce them here.

Fig. 6 shows the quasiparticle dispersion throughout the Brillouin zone, with the magnetization set to the equilibrium value for concreteness. As noted above, it can be well-approximated by a next-nearest-neighbor hopping model with amplitude $t_1$ for moving along the diagonals of the lattice and amplitude $t_2$ for moving along the principal axes (cf. Fig. 4). The form of the dispersion is not sensitive to the value of magnetization. However, the effective hopping amplitudes, which we determine empirically by fitting the computed spectrum to Eq. (15), do depend on $m$ as shown in Fig. 7.

Some features of the dispersion can be explained by a simple Hartree-Fock approximation to the original Hamiltonian, in which the Hubbard interaction is replaced by $n_{i\uparrow} \langle n_{i\downarrow} \rangle + \langle n_{i\uparrow} \rangle n_{i\downarrow}$. Assuming Néel order for $\langle n_{i\sigma} \rangle$, the Hamiltonian becomes a tight-binding model on a bipartite lattice with dispersion

$$\epsilon_k^{(HF)} = \sqrt{U^2 + 4t^2 \left( \cos k_x + \cos k_y \right)^2} \sim U + \frac{2t^2}{U} \left( \cos k_x + \cos k_y \right)^2,$$

using that $t \ll U$. Up to a constant shift, the second line is equivalent to Eq. (15) for the special case $t_1 = 2t_2$. Note in particular that $t_1, t_2 < 0$. Thus Hartree-Fock correctly predicts that the band minimum is within the lines $k_x + k_y = \pm \pi$. However, it incorrectly claims that the dispersion is degenerate along the entire lines. The more sophisticated SCBA resolves this degeneracy, identifying four minima at $(k_x, k_y) = (\pm \pi/2, \pm \pi/2)$.

Returning to Fig. 7, we see that the bandwidth $W$ increases significantly as the magnetization increases. Equivalently, the single-particle mass decreases. Within the framework of our calculation, the explanation is clear: a doublon/holon can move only if a spinon takes its place (see Eq. (5)), and since the Néel order ensures that every spinon hop is either into or out of the condensate, the doublon/holon hopping term is proportional to $b$.

B. Exciton properties

We now turn to the exciton properties as functions of magnetization, using the Bethe-Salpeter equation. All of the quantities presented here are straightforward to compute from the energy $E_b$ and wavefunction $\psi_k$ given by Eq. (22).

Fig. 8 shows the energy of the lowest internal state as a function of the center-of-mass momentum $K$. As was the case for the single-particle dispersion, the shape of the exciton dispersion is not particularly sensitive to the magnetization. Note that the bottom of the band is not at the origin but rather at $(K_x, K_y) = (\pi, \pi)$. The wavefunction of the $(\pi, \pi)$ state is shown in Fig. 9.

We
FIG. 8. Dispersion of the exciton center-of-mass motion, for a lattice of size 32 × 32.

FIG. 9. Exciton wavefunction in (relative) momentum space, for the same lattice as in Fig. 8.

see that it has p-wave symmetry, unlike what one would expect for a two-holon bound state.

The binding energy, mass, and radius of the exciton are plotted versus magnetization in Fig. 10. We see that as one increases the magnetization \( m \), the mass decreases while the binding energy and size increase. It is interesting to compare these trends with what one would expect for a conventional exciton formed via Coulomb attraction. In that situation, a decrease in mass is associated with an increase in radius and a decrease in binding energy. Here, we find a similar relationship between radius and mass, but the binding energy instead scales inversely with mass.

Eq. (25) can be simplified further in the large-\( t/J \) limit. We will see that \( E_b \) scales as \( t \), whereas \( \epsilon_k \) and \( \omega_q \) are asymptotically smaller [56]. Thus we can neglect the single-particle and spinon dispersions, leaving the equation

\[
E_b^2 \Psi_k = -\frac{t^2 \hbar^2}{N} \sum_q (M_{k_+q,k_-q}M_{k_-q,k_+q} + M_{k_+q,k_-q}M_{k_-q,k_+q}) \Psi_{k-q}.
\]

Although still not of the Schrödinger form, Eq. (27) is much simpler to solve than Eq. (25): the kernel on the right-hand side no longer depends self-consistently on the energy (and as claimed, \( E_b \sim t \)). The results obtained from the large-\( t/J \) equation are plotted alongside the others in Fig. 10.

As is clear from Eq. (25), the spinon-mediated interaction between charges is not instantaneous. To assess
FIG. 11. Comparison between the full results and the two approximations considered in the text, for the same properties as in Fig. 10. Vertical dashed lines indicate the equilibrium value of $m$.

the importance of this retardation, we have compared the results in Fig. 10 to what would be obtained through the static approximation (setting $\omega = 0$ in the kernel of Eq. (22)). The static approximation would predict significantly different results, as seen in Fig. 11: the binding energy would instead decrease slightly with magnetization and the mass would increase slightly. Thus the retardation of the effective interaction is an essential ingredient to the behavior seen here.

Similarly, one can ask whether the trends observed in Fig. 10 are due primarily to changes in the spinon behavior or rather due to the single-particle mass, which itself decreases with magnetization. We have repeated the above calculations under “fixed-mass” conditions, in which the single-particle parameters $t_1$ and $t_2$ are kept fixed (to their values at $m = 0.5$) as we vary the magnetization. Fig. 11 shows that each of the three observables responds differently. The binding energy becomes more sensitive to magnetization, indicating that the quasiparticle and spinon properties play antagonistic roles. On the other hand, the exciton mass becomes less sensitive—the change to the effective interaction suppresses the mass by itself. Finally, the exciton radius shows the reverse behavior to before, instead decreasing with magnetization (although the size remains quite small in absolute terms).

The recent pump-probe experiments in Refs. 22, 23 have investigated how excitons are influenced by magnetic order in the Mott insulator Na$_2$IrO$_3$. Our results support their interpretation in some aspects but not in others. In Ref. 22, the authors observe an increase in the fraction of bound excitations when below the Néel temperature, which they attribute to an increase in the exciton binding energy. Fig. 11 shows that magnetic order does indeed increase the binding energy. On the other hand, Ref. 23 demonstrates that the relaxational dynamics following a pump are slower in the presence of magnetic order. This is attributed to the mass increasing with magnetization, yet we have observed the opposite (consistent with past works calculating the dependence on $J/t$ 53, 56, 58). Given the highly non-equilibrium nature of the experiments, as well as the approximations inherent in an analytical approach, further investigation is clearly needed.

Finally, let us compare the present calculation of doublon-holon binding to that of holon-holon binding, which is obviously of significant interest in its own right 51, 59. Clearly the two have much in common, yet there are two important differences. First, the integral equation which determines the two-particle Green’s function (Fig. 5) has an additional exchange term due to the indistinguishability of the holons. Second, even the direct term comes with an extra minus sign, i.e., the effective interaction is of opposite sign. The sign can be removed by redefining the hole operator on one sublattice, but the additional phase may modify further results depending on the application. It is important to keep these distinctions in mind when relating the present results to the high-$T_c$ literature.

IV. CONCLUSION

We have studied the role that magnetic order plays in the formation of excitons within Mott insulators, using the Hubbard model as a concrete Hamiltonian. The binding energy increases in the presence of (antiferromagnetic) magnetization, whereas the exciton mass decreases. The size of the exciton increases slightly, yet the radius is never more than a lattice spacing. Using the standard classification, these are Frenkel excitons regardless of magnetic order.

In addition, we have established that the trends ob-
served here require a detailed understanding of the many-body dynamics in these systems. Retardation effects in the effective spinon-mediated interaction are essential. Furthermore, the constituent charge and spin excitations are each affected separately by the background magnetic order, in ways cooperative for some exciton properties but antagonistic for others.

It must be noted that despite the complexity, there are significant limitations to our approach. In particular, since we have made a mean-field approximation to describe the antiferromagnetic order, we can only hope to capture the effects of small fluctuations (i.e., low density of spin waves). Thus we do not expect our results to be quantitatively accurate in the small-$m$ regime, but at the same time, we do expect to have captured the overall qualitative trends.

As an outlook, the results presented here will be important when analyzing recent and future experiments on the optical properties of strongly correlated electronic materials. The existing experiments are quite complex, and require interpretations of their own. Our results agree with those interpretations in some respects but disagree in others. A complete understanding of the systems will require numerous approaches, both experimental and theoretical, including but not limited to the one described here.

Particularly promising are the recent experiments on fermionic atoms in optical lattices [34] [35] [37]. Since ultracold gases do not have many of the complicating features found in condensed matter systems, we expect that this will be a valuable direction to explore further. It is also likely that our conclusions, being based on the single-band nearest-neighbor Hubbard model, are more applicable to those systems than to materials such as the iridates. Importantly, current quantum gas microscopes allow one to directly create localized doublons and holons via optical tweezers, and reliably measure the spin correlation functions [32]. Such an unprecedented direct access to the system microscopics will provide a powerful way of investigating many-body excitons.

V. ACKNOWLEDGEMENTS

We would like to thank Eugene Demler, Immanuel Bloch, Andrew Allocca, and Zachary Raines for stimulating discussions and suggestions. This work was supported by the NSF Physics Frontier Center at the Joint Quantum Institute, the NSF DMR-1613029 and US-ARO contracts W911NF1310172 (T.-S.H), NRC Research Associateship award at the National Institute of Standards and Technology (CLB), AFOSR-MURI FA9550-16-1-0323 (M.H.), DOE-BES award DESC0001911 and the Simons Foundation (V.G.). M.H. & V.G. acknowledge the hospitality of KITP-UCSB, which is supported in part by Grant No. NSF PHY-1748958.


[55] More precisely, all \( \omega \) are non-negative, but the vertex vanishes on those momenta such that \( \omega = 0 \).


