Phase diagram of the bosonic Kondo-Hubbard model

Michael Foss-Feig and Ana Maria Rey

Phys. Rev. A **84**, 053619 — Published 16 November 2011

DOI: 10.1103/PhysRevA.84.053619

Phase diagram of the Bose Kondo-Hubbard model

Michael Foss-Feig and Ana Maria Rey
JILA, National Institute of Standards and Technology, and University of Colorado, Boulder, Colorado 80309, USA

We study a bosonic version of the Kondo lattice model with an on-site repulsion in the conduction band, implemented with alkali atoms in two bands of an optical lattice. Using both weak and strong-coupling perturbation theory, we find that at unit filling of the conduction bosons the superfluid to Mott insulator transition should be accompanied by a magnetic transition from a ferromagnet (in the superfluid) to a paramagnet (in the Mott insulator). Furthermore, an analytic treatment of Gutzwiller mean-field theory reveals that quantum spin fluctuations induced by the Kondo exchange cause the otherwise continuous superfluid to Mott-insulator phase transition to be first order. We show that lattice separability imposes a serious constraint on proposals to exploit excited bands for quantum simulations, and discuss a way to overcome this constraint in the context of our model by using an experimentally realized non-separable lattice. A method to probe the first-order nature of the transition based on collapses and revivals of the matter-wave field is also discussed.

PACS numbers: 03.75.Ss, 37.10.Jk, 67.85.–d, 71.27.+a, 75.30.Hx

I. INTRODUCTION

Ultracold atoms in optical lattices have been used to simulate a variety of condensed matter Hamiltonians [1], with eminent successes including the simulation of both the Bose [2] and Fermi [3, 4] *single-band* Hubbard models. More recently, progress in controlling and stabilizing atoms in the excited bands of an optical lattice [5–8] has given rise to the exciting possibility of simulating *multi-band* condensed matter Hamiltonians, which involve a nontrivial interplay of spin, charge, and orbital degrees of freedom. These achievements have precipitated a variety of theoretical investigations into the new physics made possible by the orbital degrees of freedom in an optical lattice [9–12].

The Kondo lattice model (KLM), in which tightly bound electrons act as spinful scattering centers for electrons in a conduction band [13], is a typical example of the type of model one would like to simulate. In the KLM, the orbital degree of freedom gives rise to a rich phase diagram that includes, e.g., magnetically ordered states, a heavy Fermi liquid, and unconventional superconductors. In this manuscript we will revisit a version of the KLM first proposed in Ref. [14], in which the electrons are replaced by spin-$\frac{1}{2}$ bosons, with the spin degree of freedom encoded in two hyperfine states of an alkali atom. Our primary new finding is that, for any small but nonzero Kondo coupling, the typically continuous superfluid (SF) to Mott insulator (MI) phase transition becomes first-order. The qualitative difference between the pure Hubbard and Kondo-Hubbard model, even at arbitrarily weak Kondo coupling, is reminiscent of similar results for the Fermi Kondo-Hubbard model obtained in Ref. [15]. That the inclusion of small inter-band interactions (which are often relevant in real materials) can have such a dramatic effect on the Bose Hubbard phase diagram underscores the importance of generalizing optical lattice simulations to include orbital degrees of freedom.

The structure of the manuscript is as follows. The experimental implementation of the Bose Kondo-Hubbard model is discussed in Sec. II. This section deviates from the original proposal [14] in that we emphasize the essential role of lattice non-separability in obtaining the model in more than one spatial dimension. We begin our investigation of the model's phase diagram in Sec. III, where we derive effective spin Hamiltonians valid in the weak coupling (Sec. III A) and strong coupling (Sec. III B) limits in order to understand the magnetic properties of the SF and MI phases. We then employ an analytic treatment of Gutzwiller mean-field theory (MFT) in Sec. IV to map out the ground state phase diagram. At mean-field level, one can observe the interplay of two competing tendencies: Superfluidity of the conduction bosons tends to spontaneously break SU(2) symmetry, whereas the Kondo interaction prefers a rotationally symmetric ground state composed of localized singlets. While the rigidity of the superfluid protects it from magnetic fluctuations, we will see that these fluctuations give rise to a metastable MI of Kondo singlets, causing the MI to SF transition to become first-order. Such a first-order transition has been realized—and confirmed by Quantum Monte Carlo—in the spin-$\frac{1}{2}$ boson model of Ref. [16]. In Sec. V A we consider experimental details related to dynamically maintaining a two-band model. In particular we discuss the implementation of the model, and the relevant parameter regimes, using the non-separable lattice of Ref. [17]. In Sec. V B we discuss the preparation of the unit-filled MI phase. We then suggest the possibility of experimentally observing the first-order transition by ramping down the lattice to enter the SF regime; due to the discontinuous nature of the phase transition, even an arbitrarily slow ramp should excite collapses and revivals of the superfluid order parameter. In Sec. VI we summarize our findings and their relevance.

II. THE MODEL

Everything that follows assumes a three dimensional (3D) optical lattice $\mathcal{I}(\boldsymbol{r}) + \mathcal{I}_z(z)$ populated by bosonic alkali atoms with mass m and s-wave scattering length a_s. Here $\boldsymbol{r} = \{x, y\}$ is a vector living in the x-y plane. The model itself will be two dimensional (2D), which can be achieved by choosing the lattice along z to be sufficiently deep, and loading every atom into a Wannier orbital $w_0^z(z)$ of the lowest band of $\mathcal{I}_z$. With the exception of the integral $\chi \equiv \int dz\, w_0^z(z)^4$, which will appear in the interaction matrix elements below, we will from now on ignore the existence of the z-direction completely. We assume that all atoms are in the $F = 1$ hyperfine manifold with $m_F = \pm 1$. In ^{87}Rb, the similarity of scattering lengths for total spin $F = 0$ and $F = 2$ collisions, together with the ability to offset the $m_F = 0$ state via the quadratic Zeeman effect [7, 18], strongly suppresses spin changing collisions into the $m_F = 0$ state. This justifies considering only two internal states for times long after the initial preparation, and we label these two internal (spin) degrees of freedom by $\sigma = \uparrow, \downarrow$. At sufficiently low temperatures, the full interatomic potential can be replaced with an s-wave pseudopotential [19, 20] $V_s(\boldsymbol{r}_{\rm rel}) = \delta(\boldsymbol{r}_{\rm rel}) 4\pi a_s \hbar^2/m$, where $\delta(\boldsymbol{r}_{\rm rel})$ is the 3D delta function and $\boldsymbol{r}_{\rm rel}$ is the (3D) relative vector between two colliding atoms [32]. Neglecting nearest neighbor interactions, the Hamiltonian describing this system is

$$\mathcal{H} = \sum_{ij\alpha\sigma} J_{\alpha ij} \alpha^\dagger_{i\sigma} \alpha_{j\sigma} + \sum_{j\alpha} \frac{U_\alpha}{2} \hat{n}_{j\alpha}(\hat{n}_{j\alpha} - 1) + \sum_{j,\alpha>\beta} V_{\alpha\beta} \hat{n}_{j\alpha} \hat{n}_{j\beta} + \sum_{j,\alpha>\beta,\sigma\sigma'} V_{\alpha\beta} \alpha^\dagger_{j\sigma} \beta^\dagger_{j\sigma'} \alpha_{j\sigma'} \beta_{j\sigma} + \sum_{j,\{\alpha,\beta\}\neq\{\gamma,\delta\}} W_{\alpha\beta\gamma\delta} \alpha^\dagger_{j\sigma} \beta^\dagger_{j\sigma'} \gamma_{j\sigma'} \delta_{j\sigma}. \quad (1)$$

In general there should also be a term accounting for the mean band-energies. However, in consideration of the forthcoming approximation that there are no band changing collisions, which will render such terms conserved quantities, we ignore them from the outset. In Eq. (1), $\alpha^\dagger_{j\sigma}$ creates a spin-σ boson in a Wannier orbital $w_\alpha(\boldsymbol{r} - \boldsymbol{r}_j)$ belonging to the $\alpha^{\rm th}$ band of $\mathcal{I}(\boldsymbol{r})$, located on the lattice site centered at $\boldsymbol{r}_j$. The density operator for bosons on site j in the $\alpha^{\rm th}$ band is $\hat{n}_{j\alpha} = \sum_\sigma \alpha^\dagger_{j\sigma} \alpha_{j\sigma}$, and the various parameters are given by

$$\begin{aligned} J_{\alpha ij} &= \int d^2\boldsymbol{r} w_\alpha(\boldsymbol{r} - \boldsymbol{r}_i) \left[\mathcal{I}(\boldsymbol{r}) - \frac{\hbar^2 \nabla^2}{2m} \right] w_\alpha(\boldsymbol{r} - \boldsymbol{r}_j) \\ U_\alpha &= \frac{4\pi\hbar^2 a_s \chi}{m} \int d^2\boldsymbol{r} w_\alpha(\boldsymbol{r})^4 \\ V_{\alpha\beta} &= \frac{4\pi\hbar^2 a_s \chi}{m} \int d^2\boldsymbol{r} w_\alpha(\boldsymbol{r})^2 w_\beta(\boldsymbol{r})^2 \\ W_{\alpha\beta\gamma\delta} &= \frac{2\pi\hbar^2 a_s \chi}{m} \int d^2\boldsymbol{r} w_\alpha(\boldsymbol{r}) w_\beta(\boldsymbol{r}) w_\gamma(\boldsymbol{r}) w_\delta(\boldsymbol{r}). \quad (2) \end{aligned}$$

The notation $\{\alpha, \beta\} \neq \{\gamma, \delta\}$ means that these two sets of indices cannot contain identical elements, and is introduced so that the final term in $\mathcal{H}$ contains all (and only) scattering processes that change the band populations on a given site. The hoppings $J_{\alpha ij}$ are so far unspecified; in order to obtain a 2D model, we only require that, for a given α, all such hoppings emanating from one site to its nearest neighbors are similar in magnitude. To set the overall scale of kinetic energy we define the effective hopping strength $J_\alpha = \frac{1}{4}\sum_i |J_{\alpha ij}|$ [33], which does not depend on the site index j.

We further restrict our attention to an initial state with one atom per site in the lowest vibrational band (b band, or localized band) of $\mathcal{I}(\boldsymbol{r})$, and an average of n atoms per site in a single excited band (a band, or conduction band). The assumption that these conditions will be maintained dynamically, a necessary condition for obtaining a bosonic Kondo model, amounts to disregarding all of the $W_{\alpha\beta\gamma\delta}$ terms in Eq. (1). In previous works [5, 8, 21], such an approximation has been justified largely by the restricted density of states in an optical lattice. Very roughly, the argument is that when interactions are small compared to typical band gaps, and if the bands themselves are narrow, then band changing collisions tend to be off resonant and strongly suppressed. Here we point out a notable exception to the validity of this reasoning, which to our knowledge has not been previously described in the literature. This exception affects the implementation of the Kondo-Hubbard model, and also of the multi-flavor models described in Ref. [21].

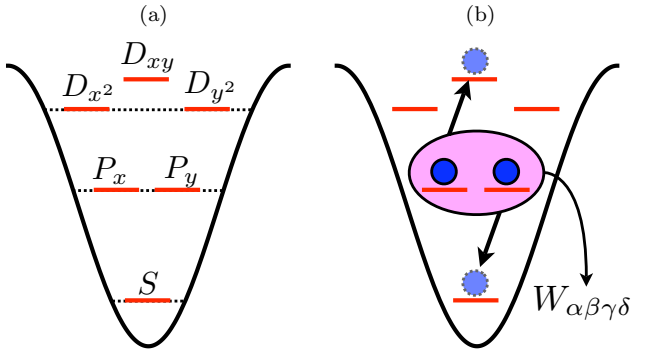

FIG. 1: (Color online) In (a) we plot an energy level diagram for states $|00\rangle$ (S band), $|10\rangle$ (P_x band), $|01\rangle$ (P_y band), $|20\rangle$ (D_{x^2} band), $|02\rangle$ (D_{y^2} band), and $|11\rangle$ (D_{xy} band). The resonant scattering process (due to the $W_{\alpha\beta\gamma\delta}$ terms in $\mathcal{H}$) that transfers atoms from the initially populated S and D_{xy} bands into the P_x and P_y bands is shown in (b).

For a lattice that can be written $\mathcal{I}(\boldsymbol{r}) = \mathcal{I}_x(x) + \mathcal{I}_y(y)$, a Wannier orbital in the a band can be written $w_a(\boldsymbol{r}) = w^x_{a_x}(x) w^y_{a_y}(y)$, whereas a Wannier orbital in the b band can be written $w_b(\boldsymbol{r}) = w^x_0(x) w^y_0(y)$. To have a 2D model, it must be true that hopping in both the x and y directions is greater in the conduction band than in the localized band, and hence $a_x, a_y > 0$. Because of the separability of the non-interacting part of the Hamiltonian in the x-y basis, the state $|00\rangle \otimes |a_x a_y\rangle$ [describing a single site with one atom in $w^x_0(x) w^y_0(y)$ and one atom in $w^x_{a_x}(x) w^y_{a_y}(y)$] is, even in the presence of interactions, *exactly* degenerate with the state $|0 a_y\rangle \otimes |a_x 0\rangle$. Furthermore, these states are connected by the interaction term W with a matrix element equal to twice the exchange integral V between the conduction and localized orbitals (see Appendix C). The situation is depicted graphically in Fig. 1 for the case where $a_x = a_y = 1$, and the states $|00\rangle$, $|10\rangle$, $|01\rangle$, and $|11\rangle$ belong to the S, P_x, P_y, and D_{xy} bands respectively. The P_x and P_y bands will be populated via collisions on a timescale $\tau = \pi/V$; this is unacceptable, since τ will turn out to be the timescale for singlet formation, and one can't expect to see Kondo like physics if the approximations yielding the model break down so quickly. The way around this problem is to use an optical lattice potential which cannot be separated in cartesian coordinates. In section V B we will give an example of an existing non-separable lattice which is favorable for avoiding this problem, but for the time being we continue with the approximation that the $W_{\alpha\beta\gamma\delta}$ can be ignored.

If the atoms in the b band are deep in the unit-filled MI regime, and we drop terms that are therefore constant, then $\mathcal{H}$ can be reduced to a bosonic Kondo-Hubbard model [14]

$$\begin{aligned} \mathcal{H}_{\mathrm{K}} &= -\sum_{ij\sigma} J_{ij} a^\dagger_{i\sigma} a_{j\sigma} + \frac{U}{2} \sum_j \hat{n}_{ja}(\hat{n}_{ja} - 1) \\ &+ 2V \sum_j \boldsymbol{S}_{ja} \cdot \boldsymbol{S}_{jb} - \mu \sum_j \hat{n}_{ja}. \end{aligned} \tag{3}$$

In Eq. (3) a chemical potential has been included to facilitate the forthcoming mean-field treatment, and the spin operators are defined by $\boldsymbol{S}_{j\alpha} \equiv \frac{1}{2} \sum_{\sigma\sigma'} \alpha^\dagger_{j\sigma} \boldsymbol{\tau}_{\sigma\sigma'} \alpha_{j\sigma'}$, with $\boldsymbol{\tau}$ being a vector whose components are the Pauli matrices. In order to avoid a clutter of indices, we have defined $V \equiv V_{ab}$, $J_{ij} \equiv J_{ija}$ (and $J \equiv J_a$), and $U \equiv U_a$. Because V has the same sign as a_s, the Kondo interaction is antiferromagnetic (AFM) for repulsive interactions, which is a manifestation of Hund's rule adapted for bosons: Antisymmetrization of the spin wave function for two identical bosons requires the antisymmetrization of their spatial wave function, thus keeping them apart and lowering the energy cost of repulsive interactions.

III. EFFECTIVE SPIN MODELS

In the presence of the Kondo term, one still expects the conduction bosons to undergo a MI to SF phase transition at unit filling ($n = 1$) as the ratio U/J is varied. However, the magnetic properties of these phases will be heavily influenced. The goal of this section is to study magnetic ordering in the SF and MI phases by deriving effective spin Hamiltonians that are accurate in either the weak coupling or strong coupling limits. The results presented are meant to reinforce and complement the mean-field picture that will subsequently be developed.

A. Weak coupling

In the fermionic KLM at $V = 0$, the localized spins are decoupled from the conduction electrons, giving rise to a large spin degeneracy in the ground state manifold. The lifting of this degeneracy is the result of virtual excitations of the conduction electrons from below to above the Fermi-surface, which mediate long ranged couplings between the localized spins known as the Rudderman-Kittel-Kasuya-Yosida (RKKY) interaction [22–24]. In dimensions greater than one, the RKKY interaction is believed to stabilize long-range magnetic order in the KLM ground state [23]. The primary difference in the bosonic model is that the $V = 0$ spin degeneracy extends to the conduction band, where the ferromagnetic superfluid can point in any direction. Surprisingly, this additional freedom actually simplifies the situation; the degeneracy will now be lifted at first (rather than second) order in V, and at this order the effective Hamiltonian can be solved exactly.

Although we are not considering the presence of a confining potential in this manuscript, in order to preserve the greatest generality in the ensuing discussion we will state our results in terms of arbitrary single particle eigenstates $\psi_r(j)$ and eigenvalues ϵ_r for a lattice plus trap. At $V, U = 0$, the ground state is formed by putting all conduction bosons in the single particle wavefunction $\psi_0(j)$ and has a degeneracy of $(N_a + 1) \times 2^{\mathcal{N}}$, where $\mathcal{N}$ is the number of lattice sites (also the number of b atoms), and N_a is the number of a atoms. It is useful to define the functions $\mathcal{G}^r_{jl} = \psi^*_r(j) \psi_r(l)$, in terms of which the spin density operator for a conduction boson in the single particle groundstate is

$$\boldsymbol{S}^0_a = \frac{1}{2} \sum_{jl} \mathcal{G}^0_{jl} \sum_{\sigma\sigma'} a^\dagger_{j\sigma} \boldsymbol{\tau}_{\sigma\sigma'} a_{l\sigma'}. \tag{4}$$

Projection of $\mathcal{H}_{\mathrm{K}}$ onto the degenerate groundstate manifold yields an effective weak coupling Hamiltonian that is first order in the Kondo coupling V

$$\mathcal{H}^{(1)}_{\mathrm{wc}} = 2V \boldsymbol{S}^0_a \cdot \sum_j \mathcal{G}^0_{jj} \boldsymbol{S}_{jb}. \tag{5}$$

The first order energy shift due to U has been dropped because it does not lift the spin degeneracy. Equation (5)

describes the so-called *central spin model*, which is familiar in the context of electron spin decoherence in quantum dots [25]. In the application to quantum dots, the model describes coupling between the spin of an electron in the dot and the nuclear spins of the atomic lattice in which the dot sits, with a coupling function determined by the square of the electron's wavefunction. In the present case, the condensate atoms are Schwinger bosons representing the central spin $\boldsymbol{S}_a^0$, which couples to the mutually non-interacting localized spins with a coupling function given by the square of the condensate wavefunction. The central spin model has been studied extensively in the literature, and exact solutions have been obtained by Bethe ansatz [26]. However, for a translationally invariant system $\mathcal{H}_{\rm wc}^{(1)}$ can easily be diagonalized by rewriting it in terms of the conserved quantities s_b, s_a, and s, where s_α is the total spin quantum number of $\boldsymbol{S}_\alpha = \sum_j \boldsymbol{S}_{j\alpha}$, and s is the total spin quantum number of the combine spin $\boldsymbol{S}_a + \boldsymbol{S}_b$. The ground state is formed when the localized spins align fully ferromagnetically ($s_b = \mathcal{N}/2$), and then their total spin $\boldsymbol{S}_b$ couples as antiferromagnetically as possible to the condensate spin $\boldsymbol{S}_a$ ($s = \frac{\mathcal{N}}{2}|n-1|$).

Second-order perturbation theory in the Kondo exchange generates a correction to the weak coupling Hamiltonian

$$\begin{aligned}\mathcal{H}_{\rm wc}^{(2)} &= -n\mathcal{N}V^2 \sum_{j,l} \mathcal{R}_{jl} \boldsymbol{S}_{jb}\cdot\boldsymbol{S}_{lb} \\ &+ 2V^2 \boldsymbol{S}_a \cdot \sum_j \mathcal{R}_{jj}\boldsymbol{S}_{jb}, \end{aligned} \tag{6}$$

where $\mathcal{R}_{jl} = \sum_r \mathcal{G}^r_{jl}\mathcal{G}^0_{lj}(\varepsilon_r)^{-1}$ [see Appendix A for a derivation of Eq. (6)]. The first term in Eq. 6 is the bosonic analog of the RKKY interaction for fermions, and has the same physical origin; a spin at site j scatters an a atom out of the condensate, which then re-enters the condensate upon scattering off the spin at site l, and in this way the two spins talk to each other. The second term, which renormalizes the spin couplings of $\mathcal{H}_{\rm wc}^{(1)}$, reflects the ability of the scattered boson to return to the condensate with its spin flipped. Such a term is absent for fermions because Pauli exclusion principle prevents a conduction fermion from returning to the Fermi sea with its spin flipped[34]. Unlike the RKKY term this correction is local in space, a property that can be attributed to the cancellation of time reversed scattering processes involving more than one localized spin (the sign for such a scattering process depends on whether the spin flip occurs when the conduction atom is exiting or returning to the condensate).

Although the first term has the same physical origin as RKKY for fermions, an important difference arrises due to the structure of the coupling function. In the fermionic case, the Fermi surface introduces a length scale k_F^{-1}, at which the long-range coupling oscillates. For noninteracting bosons there is no such length scale, the coupling function $\mathcal{R}$ does not oscillate, and one can show that it is strictly positive at any finite separation for a translationally invariant system in the thermodynamic limit. Thus any groundstate of $\mathcal{H}_{\rm wc}^{(1)}$ automatically minimizes $\mathcal{H}_{\rm wc}^{(2)}$, and the polarization of the localized spins persists to second order in V (note also that the renormalization of the central spin coupling constants by the second term in $\mathcal{H}_{\rm wc}^{(2)}$ is toward larger positive values). It should be noted, however, that $\mathcal{N}\mathcal{R}_{jj}$ diverges logarithmically with the system size in 2D, suggesting that the energy cannot actually be expanded in powers of V. Nevertheless, the existence of exclusively ferromagnetic (FM) terms in the first two orders of perturbation theory strongly suggests a FM ground state at weak coupling. We note that at unit filling, the true groundstate of $\mathcal{H}_{\rm wc}$ is a singlet ($s = 0$), but nevertheless the superfluid and the localized spins are aligned ferromagnetically within themselves ($s_a = s_b = \frac{\mathcal{N}}{2}$). That $s_b = \frac{\mathcal{N}}{2}$ means that all off-diagonal elements of the correlation function

$$\chi_{jl} = \langle \boldsymbol{S}_{jb}\cdot\boldsymbol{S}_{lb}\rangle \tag{7}$$

obtain the maximum value of $\frac{1}{4}$. Hence when we say that the $n = 1$ superfluid phase is FM, we mean that ferromagnetism exists independently within the superfluid and within the localized spins.

B. Strong coupling

We will take the strong coupling limit to be defined by $U \gg J$, for any V, and for simplicity we will restrict our discussion to the case of commensurate filling in the conduction band. However, we will see that a similar limit arrises for $V \gg J$ and any U when the a atoms are at unit filling. For integer $n \geq 1$ the ground state is a MI with n conduction bosons per site. The eigenstates on a single site follow from the addition of angular momenta $\boldsymbol{S}_{ja} + \boldsymbol{S}_{jb} \equiv \boldsymbol{S}_j$, and therefore have total spin quantum number $s_j^{\pm} = (n \pm 1)/2$. Because the interaction is AFM, the eigenstate with lower total spin is the ground state, and the MI phase with density n must have total spin $(n-1)/2$ at each site. If on a single site we label the state by its total spin quantum number s_j and the z projection of total spin s_j^z, then the n filling MI at $J = 0$ is given by

$$|{\rm MI}_n\rangle = \bigotimes_j |(n-1)/2, s_j^z\rangle, \tag{8}$$

and it has energy per site $E_n = \frac{U}{2}n(n-1) - \frac{V}{2}(2+n)$. For $n = 1$ each site contains a singlet $|0,0\rangle_j$, spin excitations are gapped ($\Delta_{\rm s} = 2V$), and the ground state is PM. We also note that for unit filling the charge gap is $\Delta_{\rm c} = U + V$, and does not vanish as $U \to 0$. Hence the unit-filled MI exists whenever either U or V is large compared to J.

States with $n \geq 2$ can be obtained from the singlet by repeated application of the a atom creation operators:

$$|s_j^-, s_j^z\rangle \propto (a_{j\uparrow}^\dagger)^{s_j^- + s_j^z}(a_{j\downarrow}^\dagger)^{s_j^- - s_j^z}|0,0\rangle_j. \tag{9}$$

It is interesting to note that for a MI phase with $n \geq 2$ the charge gap is given by $\Delta_c = U$, and has no dependence on V. Therefore the limit $V \gg J$ alone does not yield a MI in this case. Because the states $|s_j^-, s_j^z\rangle$ are degenerate ($s_j^z = -s_j^-, ..., s_j^-$), it is possible to derive an effective super-exchange Hamiltonian between the total spins on neighboring sites by perturbation theory in the hoppings J_{ij}. The calculation—though complicated by the existence of virtual excited states with two possible total spin quantum numbers—is straightforward, and yields an effective Hamiltonian that only contains the *total spin operators* $\boldsymbol{S}_j$

$$\mathcal{H}_{\rm sc} = -g(n) \sum_{ij} \frac{J_{ij}^2}{U} \boldsymbol{S}_i \cdot \boldsymbol{S}_j. \tag{10}$$

In Eq. (10),

$$g(n) = \frac{4n+8}{(n+1)^2} \left(\frac{V(n+1)^2 + U(n+2)^2}{Vn+U} \right) \tag{11}$$

is a strictly positive, density dependent coupling constant, and we have dropped an overall density dependent energy shift. The ground state is therefore FM with total spin $\frac{\mathcal{N}}{2}(n-1)$, and, unlike for the $n = 1$ MI, there is no spin gap. Notice that this result for the total spin also holds for the weak coupling case, although the nature of the FM state is completely different in the two limits. While the inter-site spin correlations between the localized spins are maximized at weak coupling, in the strong-coupling limit we find the diminished correlation $\chi_{jl} = \frac{1}{4}(\frac{n-1}{n+1})^2$, which vanishes when $n = 1$.

IV. MEAN FIELD THEORY

As was shown by a numerical analysis of the Gutzwiller variational ansatz in Ref. [14], in the presence of the localized spins the a atoms continue to exhibit a MI to SF phase transition at commensurate filling. Here we adopt an alternate but equivalent description of the Gutzwiller variational ansatz as a site-decoupled mean-field theory (MFT) [27], and obtain an analytical description of the phase transition. The motivation for this mean-field treatment is that (1) it semi-quantitatively describes the MI to SF phase transition of bosons in $D > 1$ and (2) the new features of the present model—the Kondo term—are entirely local, and thus are included without further approximation.

The starting point of Gutzwiller mean-field theory is the decoupling of the kinetic energy term

$$a_{i\sigma}^\dagger a_{j\sigma} \to \langle a_{i\sigma}^\dagger \rangle a_{j\sigma} + a_{i\sigma}^\dagger \langle a_{j\sigma} \rangle - \langle a_{i\sigma}^\dagger \rangle \langle a_{j\sigma} \rangle \tag{12}$$

Assuming that the mean-field order parameter $\psi_\sigma = \langle a_{j\sigma} \rangle$ is translationally invariant, and defining $\mathcal{J} = 4J/U$, $\mathcal{V} = V/U$, and $\mathcal{M} = \mu/U$, we can rewrite the exact Hamiltonian $\mathcal{H}_{\rm K}$ as a sum over identical Hamiltonians defined at each site (any one of which we call the mean-field Hamiltonian):

$$\begin{aligned} \mathcal{H}_{\rm GA} &= -\mathcal{J} \sum_\sigma (\psi_\sigma a_\sigma^\dagger + \psi_\sigma^* a_\sigma) + \mathcal{J} \sum_\sigma \psi_\sigma^* \psi_\sigma \\ &+ \frac{1}{2} \hat{n}_a (\hat{n}_a - 1) + 2\mathcal{V} \boldsymbol{S}_a \cdot \boldsymbol{S}_b - \mathcal{M} \hat{n}_a. \end{aligned} \tag{13}$$

The mean-field ground state is obtained by minimizing $E_{\rm GA}(\psi_\sigma) \equiv \langle \mathcal{H}_{\rm GA} \rangle$ with respect to ψ_σ. The order parameter describes the formation of superfluid coherence; if ψ_σ is finite, then by virtue of having to point somewhere it describes a state that spontaneously breaks SU(2) symmetry, as the superfluid should. On the other hand, $\psi_\sigma = 0$ is the signature of a Mott insulator phase, in which number fluctuations are suppressed and phase coherence vanishes. Because $\mathcal{H}_{\rm GA}$ is invariant under the U(1) gauge transformations $a_\sigma, \psi_\sigma \to e^{i\theta_\sigma} a_\sigma, e^{i\theta_\sigma} \psi_\sigma$, we are justified to choose both components of our spinor order parameter to be real. Physically, this choice amounts to confining any superfluid that arrises to live in the $x-z$ plane, since the y component of the (vector) superfluid density

$$\boldsymbol{\rho} = \sum_{\sigma\sigma'} \psi_\sigma^* \boldsymbol{\tau}_{\sigma\sigma'} \psi_{\sigma'} \tag{14}$$

vanishes when $\psi_\sigma^* = \psi_\sigma$. It should be understood in what follows that this apparent broken symmetry is an artifact maintained for simplicity.

The mean-field Hamiltonian can easily be solved numerically by truncating the single-site Hilbert space to contain no more than some finite number of a atoms. However, much insight can be gained by proceeding as far as possible analytically. If one makes the assumption that the SF to MI transition is continuous, then the mean-field phase boundary is described exactly by perturbation theory in ψ_σ. We want to emphasize from the beginning that this assumption will later be shown to break down in certain parameter regimes. Nevertheless, with this assumption in mind one writes the mean-field energy as

$$E_{\rm GA} = \mathcal{A}(\mathcal{J}, \mathcal{V}, \mathcal{M}) + \mathcal{B}(\mathcal{J}, \mathcal{V}, \mathcal{M}) \psi^2 + \mathcal{O}(\psi^4), \tag{15}$$

where $\psi^2 \equiv \sum_\sigma \psi_\sigma^2$ is the superfluid density, and the phase boundary is determined by the condition $\mathcal{B} = 0$. That $E_{\rm GA}$ can be written as an even function of ψ follows from the symmetries of $\mathcal{H}_{\rm GA}$. For a Mott lobe of filling n, the groundstate at finite $\mathcal{V}$ is in general degenerate (see Sec. III B), and $\mathcal{B}$ must be found by calculating the lowest eigenvalue of the effective Hamiltonian

$$\mathcal{H}_{\rm eff} = \mathcal{J}^2 \sum_{n\sigma\sigma'} \psi_\sigma \psi_{\sigma'} \frac{(a_\sigma^\dagger + a_\sigma)|n\rangle\langle n|(a_{\sigma'}^\dagger + a_{\sigma'})}{E_0 - E_n}. \tag{16}$$

When $n = 1$ the MI ground state is unique (the singlet), and while degenerate perturbation theory is unnecessary it is not wrong (we simply end up finding the

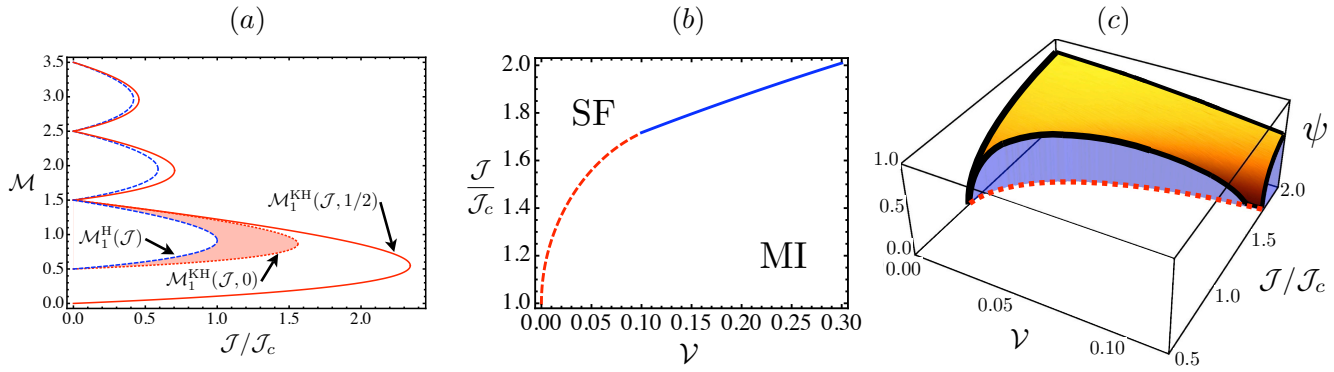

FIG. 2: (Color online) (a) Mott lobes in the $\mathcal{M}-\mathcal{J}$ plane, with $\mathcal{J}$ plotted in units of $\mathcal{J}_c = (3+\sqrt{8})^{-1}$, which is where the single band mean-field transition occurs at unit filling. The red solid lines depict the boundaries of Mott lobes for $V = U/2$, given by $\mathcal{M}_n^{\rm KH}(\mathcal{J}, 1/2)$. The blue dashed lines are the boundaries of Mott lobes for the single band boson Hubbard model, denoted $\mathcal{M}_n^{\rm H}(\mathcal{J})$. These have been shifted up by $\mathcal{V} = 1/2$ along the $\mathcal{M}$-axis in order to more easily compare their widths with those of the Kondo-Hubbard lobes. The red dotted line is the boundary $\mathcal{M}_1^{\rm KH}(\mathcal{J}, 0)$, which does not agree with $\mathcal{M}_n^{\rm H}(\mathcal{J})$. In Fig. 2(b) we plot the phase boundary of the $n = 1$ MI in the $\mathcal{J}$-$\mathcal{V}$ plane. The solid blue line demarks a continuos transition, while the red dotted line demarks a first-order transition. Fig. 2(c) shows the superfluid order parameter ψ in the $\mathcal{J}$-$\mathcal{V}$ plane, showing the discontinuity along the first-order portion of the phase transition.

eigenvalue of a 1×1 matrix, which is of course its only entry). Some algebra reveals that the above effective Hamiltonian can be rewritten in the following sensible form

$$\mathcal{H}_{\rm eff} = \mathcal{J}^2 c_1(\mathcal{V}, \mathcal{M}, n)\rho - \mathcal{J}^2 c_2(\mathcal{V}, \mathcal{M}, n)\boldsymbol{\rho}\cdot\boldsymbol{S}, \qquad (17)$$

where $\boldsymbol{S} = \boldsymbol{S}_a + \boldsymbol{S}_b$, the coefficients c_1 and c_2 are presented in Appendix B, and the vector superfluid density $\boldsymbol{\rho}$ acts as a magnetic field. It is not hard to show that c_2 is strictly positive [35], and as a result minimization of $\mathcal{H}_{\rm eff}$ is achieved when $\boldsymbol{S}$ points along the magnetic field. This result makes perfect sense, since it amounts to self consistency in the direction of the order parameter. Because the quantum number of total spin in the unperturbed groundstate is $s = (n-1)/2$, we find

$$\mathcal{B}(n) = \mathcal{J}^2\left(\frac{1}{\mathcal{J}} + c_1(\mathcal{V}, \mathcal{M}, n) - c_2(\mathcal{V}, \mathcal{M}, n)\frac{n-1}{2}\right). \qquad (18)$$

$\mathcal{B}(n) = 0$ determines the boundary of the n-filling Mott lobe, which we will denote by $\mathcal{M}_n^{\rm KH}(\mathcal{J}, \mathcal{V})$. The first three lobes are shown in Fig. 2(a). Notice that the width of the $n = 1$ MI is increased (from 1 to $1+\mathcal{V}$ along the $\mathcal{M}$ axis), while the widths of the higher filling lobes are unaffected. The reason for this goes back to the discussion in Sec. III B, where we pointed out that the charge gap is given by $\Delta_{\rm c} = U(1+\mathcal{V})$ for the $n = 1$ MI and by $\Delta_{\rm c} = U$ for an $n \geq 2$ MI.

At this point we are in a position to see that something is wrong with the phase boundaries as they have been presented so far [the solid red lines in Fig. 2(a)]. Explicit calculation yields the boundaries at zero Kondo coupling

$$2\mathcal{M}_n^{\rm KH}(\mathcal{J}, 0) = 2n - 1 - \mathcal{J} \pm \sqrt{(\mathcal{J}-1)^2 - 4\mathcal{J}n - \frac{4J}{n+1}} \qquad (19)$$

whereas the mean-field phase boundaries for the single-band Bose Hubbard model are given by

$$2\mathcal{M}_n^{\rm H}(\mathcal{J}) = 2n - 1 - \mathcal{J} \pm \sqrt{(\mathcal{J}-1)^2 - 4\mathcal{J}n}, \qquad (20)$$

which only agree at $\mathcal{J} = 0$. But the conduction bosons of the Kondo-Hubbard model certainly *are* governed by the single band Bose Hubbard model at $\mathcal{V} = 0$, since they don't talk to the localized spins. What has gone wrong? The problem is that Eq. (19) was derived by assuming that the mean-field ground state in the MI phase has $s = (n-1)/2$, which is true whenever $\mathcal{V} > 0$. However, there are also the excited spin states with $s = (n+1)/2$, separated from the ground state manifold by $\Delta_n = \mathcal{V}(n+1)$, which have been ignored. Formally, such a procedure is correct for any finite $\mathcal{V}$, so long as $\mathcal{J}\psi \ll \Delta_n$ is satisfied. As $\mathcal{V} \to 0$, the range of validity shrinks, until eventually at $\mathcal{V} = 0$ the perturbative results break down for any finite ψ.

Once this issue is understood, it becomes clear that the phase transition must in fact be first-order, and the argument is as follows. If we sit somewhere in the red shaded region of Fig. 2(a), at $\mathcal{V} = 0$ the system is a superfluid (since we are outside of the dashed blue line, which *is* correct at $\mathcal{V} = 0$), and so $E_{\rm GA}$ looks like it does in Fig. 3(b). Now as we turn on small but finite $\mathcal{V}$, a local minimum of the energy *must* immediately develop at $\psi = 0$ [see Fig. 3(c)], since perturbation theory has a finite range of validity in ψ and predicts $\mathcal{B} > 0$ [a Mott insulator, see the red dotted curve in Fig. 2(a)]. As $\mathcal{V}$ becomes larger, the range of validity for the perturbative result Eq. (19) increases, until eventually the $\psi = 0$ minimum must be the global minimum; at this point there is a first-order transition into a Mott insulator of Kondo singlets. Notice that the first-order transition is not tied to the development of a cubic term in the energy, as is often the case, but rather results from the negative contributions

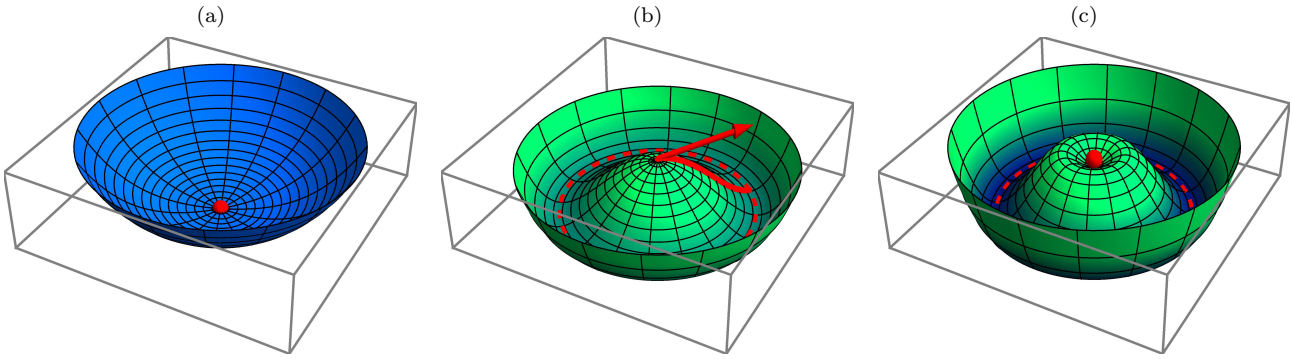

FIG. 3: (Color online) Schematic illustration of the mean-field energy, demonstrating the mechanism that drives the SF to MI phase transition first-order. (a) and (b) show the two forms that $E_{\rm GA}$ can take when $\mathcal{V}=0$, (a) in the MI phase and (b) in the SF phase, with the order parameter spontaneously breaking SU(2) symmetry by choosing a particular direction. In (c) we show a possible scenario for a small nonzero $\mathcal{V}$. The superfluid is weakly affected (being rigid against spin fluctuations), but when the order parameter is small magnetic fluctuations are allowed, reducing the energy and causing the formation of a metastable MI phase.

of even powers of ψ beyond ψ^2. By numerically solving the mean-field Hamiltonian we can obtain the first-order phase boundaries, defined by where the metastable MI and the superfluid become degenerate. The boundary for the first Mott lobe is shown in Fig. 2(b), whereas the discontinuous jump in the order parameter across this boundary is shown in Fig. 2(c).

While the above argument is quite rigorous, it offers little physical insight into what is really going on; there is a more intuitive argument that explains the existence of the metastable MI phase and pinpoints the general features of our model that give rise to it. The superfluid ground state has, in a sense to be made precise shortly, a certain rigidity against spin fluctuations. Denoting the $\mathcal{V}=0$ mean-field ground state $|\langle \boldsymbol{S}_b\rangle, \boldsymbol{\rho}\rangle$, and letting $\hat{\boldsymbol{\rho}}=\boldsymbol{\rho}/\psi^2$ be a unit vector in the direction of the superfluid, one might guess that turning on a small $\mathcal{V}$ causes the groundstate to be the singlet-like $|-\frac{1}{2}\hat{\boldsymbol{\rho}},\boldsymbol{\rho}\rangle - |\frac{1}{2}\hat{\boldsymbol{\rho}},-\boldsymbol{\rho}\rangle$, gaining $3\mathcal{V}/2$ from the Kondo term. However, it can be seen that such a state actually has $\sim \mathcal{J}\psi^2$ more kinetic energy than the mean-field ground state at $\mathcal{V}=0$. Hence when $\mathcal{J}\psi^2 \gtrsim \mathcal{V}$, rather than fluctuating its spin the superfluid will just anti-align with the localized spin, gaining only $\mathcal{V}/2$ from the Kondo term. Notice that this rigidity is purely kinetic in origin, and that the mean-field ground state for small $\mathcal{V}$ agrees, both in form and energy, with the lowest order results of Sec. III B. On the other hand, when $\mathcal{J}\psi^2 \lesssim \mathcal{V}$, the superfluid gains more energy from the Kondo term by fluctuating its spin than it looses in kinetic energy, and hence its energy decreases by $3\mathcal{V}/2$. The tendency of the mean-field energy to be lowered more for small ψ than for large ψ is what enables a local minimum at $\psi=0$ to arise [36].

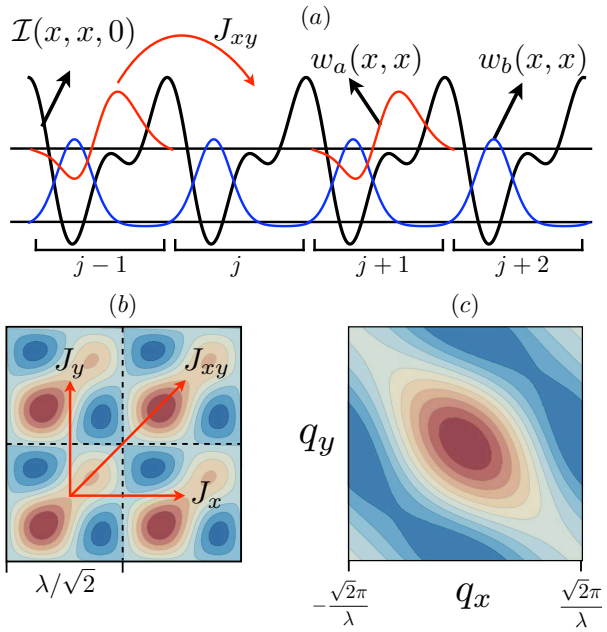

FIG. 4: (Color online). (a) Plot of $\mathcal{I}(x,x)$, with the wave functions w_a and w_b shown schematically. (b) A contour plot of $\mathcal{I}(x,y)$, showing the primary hoppings. The resulting spectrum in the a band is plotted as a function of quasi-momenta q_x and q_y (c), and can be fit to a tight-binding model with hoppings J_x, J_y, and J_{xy} [see (b)].

V. EXPERIMENTAL DETAILS

A. The optical lattice

As a specific example of how the approximations and parameters considered in this manuscript can be obtained in an optical lattice, we consider the 2D potential of Ref.

[17]

$$\mathcal{I}(x,y) = -4\mathcal{I}_o[\cos kx + \cos ky]^2 - \mathcal{I}_i[2\cos(2kx - 2\varphi) + 2\cos 2ky], \quad (21)$$

and a deep transverse confining lattice with potential $\mathcal{I}_z(z) = \mathcal{I}_\perp \cos^2(k_z z)$. Each unit cell consists of a biased double well [Fig. 4(a)], and control of the bias allows for an adjustment of the overlap integral $\int d^2\boldsymbol{r}\ |w_a(\boldsymbol{r})|^2|w_b(\boldsymbol{r})|^2$ and hence of V. This lattice offers a large parameter space to play with, and here we give just one example of a lattice configuration that could facilitate our model. Defining $E_R = \frac{h^2}{2m\lambda^2}$, we choose the parameters $\{\mathcal{I}_o, \mathcal{I}_i, \varphi\} = \{0.9E_R, 2.52E_R, 0.3\pi\}$, for which the deep well has a depth of $\sim 21E_R$ and the shallow well has a depth of $\sim 11E_R$.

Using a transverse lattice with $\mathcal{I}_\perp = 40E_R$ and $\lambda = 2\pi k_z^{-1} = 780$nm, and the scattering length of ^{87}Rb, we find $U_b/4J_b \approx 2\times 10^3$, $U_a/4J \approx 10$, and $V/U_a \approx 0.1$. Here J_b is the nearest neighbor hopping matrix element for the b band. This is just inside the $n = 1$ mean-field Mott lobe, with V/U_a in the correct range to observe the first-order phase transition, and these parameters can be adjusted to exit the Mott insulator in the a band, or to change V/U_a. One consequence of this geometry is that the bands are not isotropic. The effect is very small for the b band, but for the a band, by fitting the numerically calculated dispersion [Fig. 4(c)] to a tight binding model, we find primary hoppings $J_x = J_y = 0.002E_R$ and $J_{xy} = 0.0035E_R$ [Fig. 4(b)].

Other important energy scales are the band gaps, and for the parameters given above the first 4 gaps (all those between the b band and the a band and the gap above the a band) are all at least $1E_R$. This should be compared to the largest relevant interaction energy U_b, and for the parameters given above we find $U_b \approx 0.28E_R$. It should be noted that this comparison is an extremely conservative metric of how the interaction energies compare to the band gaps. After all, U_b is the largest interaction energy in the model, and the band gap separating the b band from the band directly above it is larger than $5E_R$.

B. Observation of the first-order transition

In order to probe the SF to MI phase transition, an ideal starting point would be a 2D ($x-y$ plane) MI of spin triplet pairs in the lowest vibrational level of the 2D lattice with $\mathcal{I}_i = 0$. The PM MI could then be achieved with high fidelity by ramping up $\mathcal{I}_i$ to establish an array of double wells in the $x-y$ plane, and then using either Raman pulses [5] or the population swapping techniques of Refs. [6, 7] to populate the a band. Standard time-of-flight imaging, combined with band-mapping techniques [7] and spin selective imaging, could be used to resolve both the superfluid coherence and the magnetic ordering in either band.

As we have discussed in Sec. V A, $\mathcal{V}$ can be tuned in a double well lattice. Therefore the fsirst-order phase transition should be observable by sitting just outside the unit-filled Mott insulator lobe and increasing $\mathcal{V}$ from 0 to some value $\mathcal{V}_f$ large enough to support a Mott insulating ground state. The discontinuous nature of the phase transition will cause a failure of adiabaticity for even an arbitrarily slow sweep of $\mathcal{V}$. At the mean-field level, the effect of ramping $\mathcal{V}$ can be explored with no further approximation by solving the time-dependent equations of motion that result from minimization of the Lagrangian [28]

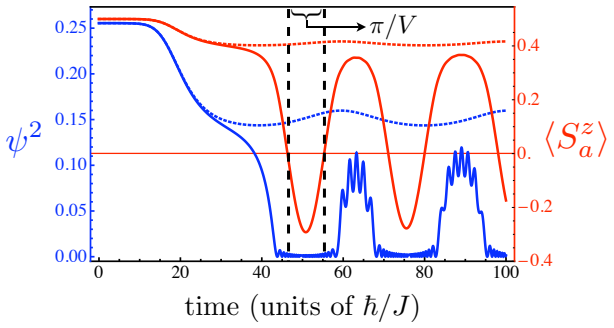

FIG. 5: (Color online) Dynamics during a slow ramp of $\mathcal{V}$ from the SF to the MI at fixed conduction band filling $n = 1$ (through a first-order phase transition). The ramping function is $\mathcal{V}(t) = \mathcal{V}_f(\tanh(t/t_0) + 1)/2$, with $t_0 \approx 5.7\hbar/J$. For a ramp to $\mathcal{V}_f = \mathcal{V}_c + \delta$ ($\mathcal{V}_f = \mathcal{V}_c - \delta$), the lower blue solid (dotted) line is ψ^2 and the upper red solid (dotted) line is $\langle S_a^z \rangle$. The fast oscillations of ψ^2 in the non-adiabatic case occur on a time scale of order $\frac{2\pi}{U}$.

$$\mathcal{L} = \langle i\frac{d}{dt} - \mathcal{H}_{\rm GA}\rangle, \quad (22)$$

where the expectation value is taken in the single site Hilbert space (for practical calculations, this space must be truncated by cutting off the maximum number of allowed bosons). Defining $\mathcal{V}_c$ to be the value of $\mathcal{V}$ at which the metastable SF solution disappears (note that this is *not* where we define the phase boundary in Sec. IV), we compare a slow ramp of $\mathcal{V}$ from 0 to $\mathcal{V}_f = \mathcal{V}_c \pm \delta$. For $\mathcal{V}_f < \mathcal{V}_c$, we observe a nearly adiabatic reduction of the SF component, whereas for $\mathcal{V} > \mathcal{V}_c$ we observe collapses and revivals of the SF component. This is reminiscent of the behavior seen in Refs. [29, 30], where a fast quench was studied in the single-band Bose Hubbard model, but here the collapses occur even for very slow lattice ramps. Since the collapses are pinned to a rotation of the magnetization [Fig. 5], it is not surprising that they repeat on a time scale $\sim \frac{2\pi}{V}$.

Regarding the experimental feasibility of the proposed dynamics around the MI to SF transition, the total time elapsed in Fig. 5 is $100\hbar/J$, which is comparable to the longest excited band decay times measured in Ref. [5] for the $n = 1$ situation. This suggests that dynamical evidence of the first-order transition, e.g. loss of adiabaticity or hysteresis, should be within reach of current experiments. Hysteretic behavior will be limited to temperatures that are comparable to or smaller than the energy

barrier between the two metastable phases. This height is set by $V \lesssim 0.1U$, suggesting that the required temperature is cold, but much more favorable than that required for super-exchange physics. A more quantitative determination of the temperature requirements necessitates a more rigorous (numerically exact, rather than mean-field) study, and will be addressed in future research. We also note that the first-order phase transition could be explored by measuring local and static observables in the trap, for instance it could manifest as a discontinuity in the density profile.

VI. SUMMARY AND CONCLUSIONS

In real metals, it is known that when conduction electrons interact with magnetic impurities their behavior is drastically altered; in order to describe such systems, it is necessary to study many-body Hamiltonians that include spin, charge, and orbital degrees of freedom, such as the KLM. The bosonic analogues of such systems are experimentally accessible using ultracold alkali atoms in *non-separable* optical lattices, motivating this fairly in depth study of the Bose Kondo-Hubbard model.

Our primary new finding is that the SF to MI phase transition of the conduction bosons is qualitatively modified by the existence of an arbitrarily small Kondo coupling to a band of localized spins. When approaching the MI, as the superfluid density of the conduction bosons is reduced, magnetic fluctuations driven by the Kondo exchange induce a metastable MI of Kondo singlets, and as a result the associated phase transition becomes first-order. We expect the first-order phase transition to be observable in experiment, making this model a leading candidate for observing entirely new, strongly correlated, multi-band phenomena in optical lattices.

Acknowledgments

We would like to thank S. Fölling, I. Bloch and L.-M. Duan for helpful discussions. This work was supported by grants from the NSF (PFC and Grant No. PIF-0904017), the AFOSR, and a grant from the ARO with funding from the DARPA-OLE.

Appendix A: Derivation of the weak coupling Hamiltonian

Here we derive the second order weak-coupling effective Hamiltonian via perturbation theory in the Kondo term $\mathcal{H}_V = 2V\sum_j \boldsymbol{S}_{ja}\cdot\boldsymbol{S}_{jb}$

$$\mathcal{H}^{(2)}_{\rm wc} = \sum_{n,\Sigma}\frac{\mathcal{H}_V|n,\Sigma\rangle\langle n,\Sigma|\mathcal{H}_V}{\varepsilon_0-\varepsilon_n}. \tag{A1}$$

The state $|n,\Sigma\rangle$ above has one conduction boson in the single particle state $\psi_n(j)$, the rest of the conduction bosons in $\psi_0(j)$, and a spin configuration labeled by the index Σ; only states of this form contribute at this order. Because the energy denominators have no dependence on Σ, the sum over Σ is a completeness identity in spin space and can be dropped. Using the basis transformation $a^\dagger_{j\sigma} = \sum_m a^\dagger_{m\sigma}\psi_m(j)$ we can rewrite the Kondo coupling as

$$\mathcal{H}_V = 2V\sum_{jmn}\boldsymbol{S}^b_j\cdot\boldsymbol{S}^a_{mn}\psi_m(j)\psi_n(j), \tag{A2}$$

where $\boldsymbol{S}^a_{mn} = \sum_{\sigma\sigma'}a^\dagger_{m\sigma}\boldsymbol{\tau}_{\sigma\sigma'}a_{n\sigma'}$. For expectation values within the degenerate ground state manifold we have the equivalence

$$\mathcal{H}^{(2)}_{\rm wc} = 4V^2\sum_{mjl}\frac{\left(\boldsymbol{S}^b_j\cdot\boldsymbol{S}^a_{0m}\right)\left(\boldsymbol{S}^b_l\cdot\boldsymbol{S}^a_{m0}\right)\mathcal{G}^0_{jl}\mathcal{G}^m_{jl}}{\varepsilon_0-\varepsilon_m}. \tag{A3}$$

Taking advantage of common identities for pauli matrices, within the ground state manifold we can rewrite

$$\left(\boldsymbol{S}^b_j\cdot\boldsymbol{S}^a_{0m}\right)\left(\boldsymbol{S}^b_l\cdot\boldsymbol{S}^a_{m0}\right) = \frac{\mathcal{N}_a}{4}\boldsymbol{S}^b_j\cdot\boldsymbol{S}^b_l - \frac{\delta_{jl}}{2}\boldsymbol{S}^a_{00}\cdot\boldsymbol{S}^b_j, \tag{A4}$$

which leads to

$$\begin{aligned}\mathcal{H}^{(2)}_{\rm wc} &= -n\mathcal{N}V^2\sum_{j,l}\mathcal{R}_{jl}\boldsymbol{S}_{jb}\cdot\boldsymbol{S}_{lb}\\ &+ 2V^2\boldsymbol{S}_a\cdot\sum_j\mathcal{R}_{jj}\boldsymbol{S}_{jb}.\end{aligned} \tag{A5}$$

Appendix B: Parameters for $\mathcal{H}_{\rm eff}$.

The effective Hamiltonian that yields the Mott lobe boundaries [Eq. 17] follows from tedious but straightforward manipulation of Eq. 16. The coefficients c_1 and c_2 are found to be

$$\begin{aligned}c_1 = \frac{1}{2n}\Bigg(&\frac{n^2+2n}{\mathcal{M}-\mathcal{V}-n} - \frac{n^2-1}{1+\mathcal{M}-\mathcal{V}-n}\\ &- \frac{1}{1+\mathcal{M}-\mathcal{V}+\mathcal{V}n-n}\Bigg)\end{aligned} \tag{B1}$$

and

$$\begin{aligned}c_2 = \frac{-1}{n^2+n}\Bigg(&\frac{n^2+2n}{\mathcal{M}-\mathcal{V}-n} - \frac{(n+1)^2}{1+\mathcal{M}-\mathcal{V}-n}\\ &+ \frac{1}{1+\mathcal{M}-\mathcal{V}+\mathcal{V}n-n}\Bigg).\end{aligned} \tag{B2}$$

Appendix C: Scattering between P and D bands in a separable optical lattice

Here we explain in more detail the problem of having 2D Kondo-like physics in a separable lattice $\mathcal{I}(\boldsymbol{r}) = \mathcal{I}_x(x) + \mathcal{I}_y(y)$; these considerations also apply to the

D > 1 multi-flavor models discussed in Ref. [21]. Though we are interested in spin-$\frac{1}{2}$ bosons, for simplicity we will ignore the spin degree of freedom (or alternatively restrict ourselves to the fully polarized states in our spin-$\frac{1}{2}$ model). Projecting the spinless version of the Hamiltonian in Eq. (1) onto the Hilbert space of a single site, we obtain

$$\tilde{\mathcal{H}} = \sum_{\alpha} E_{\alpha}\hat{n}_{\alpha} + \sum_{\alpha}\frac{U_{\alpha}}{2}\hat{n}_{\alpha}(\hat{n}_{\alpha}-1) + 2\sum_{\alpha>\beta} V_{\alpha\beta}\hat{n}_{\alpha}\hat{n}_{\beta} + \sum_{\{\alpha,\beta\}\neq\{\gamma,\delta\}} W_{\alpha\beta\gamma\delta}\alpha^{\dagger}\beta^{\dagger}\gamma\delta, \tag{C1}$$

where the site index has been dropped and the mean band energies E_{α} have been accounted for.

Because the noninteracting part of the Hamiltonian is separable in x-y coordinates, a generic Wannier function can be labeled by independent band indices for the x and y directions, and so we introduce a vector band index $\boldsymbol{\alpha} = \{\alpha_x, \alpha_y\}$. In this notation, a Wannier function is given by

$$w_{\boldsymbol{\alpha}}(\boldsymbol{r}) = w^x_{\alpha_x}(x)w^y_{\alpha_y}(y), \tag{C2}$$

and has single-particle energy $E_{\boldsymbol{\alpha}} = \epsilon^x_{\alpha_x} + \epsilon^y_{\alpha_y}$, where

$$\epsilon^x_{\alpha_x} = \int dx\; w^x_{\alpha_x}(x)\left[\mathcal{I}_x(x) - \frac{\hbar^2\partial_x^2}{2m}\right] w^x_{\alpha_x}(x) \tag{C3}$$

(and likewise for $\epsilon^y_{a_y}$). We assume all of these Wannier functions to be real. The lowest energy single particle state $|00\rangle$, with $\boldsymbol{\alpha} = \{0,0\}$, is both the most stable (against vibrational decay) and the least mobile, and hence is the obvious candidate for the localized spins in the Kondo-Hubbard model. The lowest energy state that has an enhanced hopping in *both* spatial directions is $|11\rangle$, with $\boldsymbol{\alpha} = \{1,1\}$, and we choose it for the conduction atoms.

The two-particle state with one boson in $|00\rangle$ and one boson in $|11\rangle$ is, at the single particle level, degenerate with the two-particle state in which one boson is in $|10\rangle$ ($\boldsymbol{\alpha} = \{1,0\}$) and one boson is in $|01\rangle$ ($\boldsymbol{\alpha} = \{0,1\}$), both states having energy $E_0 = \epsilon^x_0 + \epsilon^x_1 + \epsilon^y_0 + \epsilon^y_1$. Therefore, the minimal Hilbert space for an interacting model that contains both $|00\rangle$ and $|11\rangle$ must also contain $|10\rangle$ and $|01\rangle$. Defining the two particle states $|1\rangle = |00\rangle \otimes |11\rangle$ and $|2\rangle = |10\rangle \otimes |01\rangle$, it is straightforward to check that $\langle 1|\tilde{\mathcal{H}}|1\rangle = \langle 2|\tilde{\mathcal{H}}|2\rangle = 2V$, where

$$V = \frac{4\pi\hbar^2 a_s\chi}{m}\int d^2\boldsymbol{r} w^x_0(x)^2 w^y_0(y)^2 w^x_1(x)^2 w^y_1(y)^2 \tag{C4}$$

is the same Kondo exchange defined in the manuscript. In other words, in addition to having the same single particle energy, the states $|1\rangle$ and $|2\rangle$ have the same interaction energy, and remain degenerate at the *many-body* level. Furthermore, $|1\rangle$ and $|2\rangle$ are connected by the W terms in $\tilde{\mathcal{H}}$, with matrix element $\langle 1|\mathcal{H}|2\rangle = 2V$. Because $|1\rangle$ and $|2\rangle$ are degenerate at the many-body level, and are connected by an interaction with the same strength as the Kondo exchange, a site initially in state $|1\rangle$ will scatter into the state $|2\rangle$ in a characteristic time $\tau = \pi/V$.

For implementation of the multi-flavor models of Ref. [21] in D > 1, some non-separability of the lattice would need to be present in order to prevent the above mentioned collision channels from populating states outside of the Hilbert space considered (which includes the state $|2\rangle$ but excludes the state $|1\rangle$).

[1] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).
[2] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, Nature **415**, 39 (2002).
[3] U. Schneider, L. Hackermüller, S. Will, T. Best, I. Block, T. A. Costi, R. W. Helmes, D. Rasch, and A. Rosch, Science **322**, 5907 (2008).
[4] R. Jördens, N. Strohmaier, K. Günter, H. Moritz, and T. Esslinger, Nature **455**, 204 (2008).
[5] T. Müller, S. Fölling, A. Widera, and I. Bloch, Phys. Rev. Lett. **99**, 200405 (2007).
[6] G. Wirth, M. Olschlager, and A. Hemmerich, Nat. Phys. (Advanced online publication) (2010).
[7] M. Anderlini, P. J. Lee, B. L. Brown, J. Sebby-Strabley, W. D. Phillips, and J. V. Porto, Nature **448**, 452 (2007).
[8] I. B. Spielman, P. R. Johnson, J. H. Huckans, C. D. Fertig, S. L. Rolston, W. D. Phillips, and J. V. Porto, Phys. Rev. A **73**, 020702 (2006).
[9] J. Larson, A. Collin, and J.-P. Martikainen, Phys. Rev. A **79**, 033603 (2009).
[10] Q. Zhou, J. V. Porto, and S. D. Sarma, arXiv[cond-mat.quant-gas] **1010.1534** (2010).
[11] V. M. Stojanović, C. Wu, W. V. Liu, and S. Das Sarma, Phys. Rev. Lett. **101**, 125301 (2008).
[12] C. Wu, Phys. Rev. Lett. **101**, 186807 (2008).
[13] H. Tsunetsugu, M. Sigrist, and K. Ueda, Rev. Mod. Phys. **69**, 809 (1997).
[14] L. Duan, Europhys. Lett. **67**, 721 (2004).

[15] M. Feldbacher, C. Jurecka, F. F. Assaad, and W. Brenig, Phys. Rev. B **66**, 045103 (2002).
[16] L. de Forger de Parny, F. Hebert, V. G. Rousseau, R. T. Scalettar, and G. G. Batrouni, arXiv[cond-mat.quant-gas] **1105.6141** (2011).
[17] J. Sebby-Strabley, M. Anderlini, P. S. Jessen, and J. V. Porto, Phys. Rev. A **73**, 033605 (2006).
[18] T. Ho and S. Zhang, arXiv[cond-mat.quant-gas] **1007.0650** (2010).
[19] E. Fermi, Nuovo Cimento **11**, 157 (1934).
[20] K. Huang and C. N. Yang, Phys. Rev. **105**, 767 (1957).
[21] A. Isacsson and S. M. Girvin, Phys. Rev. A **72**, 053604 (2005).
[22] M. A. Ruderman and C. Kittel, Phys. Rev. **78**, 275 (1950).
[23] T. Kasuya, Prog. Theo. Phys. **16**, 45 (1956).
[24] K. Yosida, Phys. Rev. **106**, 896 (1957).
[25] A. V. Khaetskii, D. Loss, and L. Glazman, Phys. Rev. Lett. **88**, 186802 (2002).
[26] M. Gaudin, J. Physique **37**, 1087 (1976).
[27] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Phys. Rev. B **40**, 546 (1989).
[28] D. Jaksch, V. Venturi, J. I. Cirac, C. J. Williams, and P. Zoller, Phys. Rev. Lett. **89**, 040402 (2002).
[29] M. Greiner, O. Mandel, T. W. Hänsch, and I. Bloch, Nature **419**, 51 (2002).
[30] S. Will, T. Best, U. Schneider, L. Hackermüller, D.-S. Lühmann, and I. Bloch, Nature **465**, 197 (2010).
[31] D. Blume and C. H. Greene, Phys. Rev. A **65**, 043613 (2002).
[32] In general, the pseudopotential should be regularized and the scattering length should depend on momentum. The regularization can be disregarded when treating the interaction in first-order perturbation theory, which is valid for any lattice model truncated to bands in which the single-particle wave functions oscillate on a length scale $l_{\rm sp}$ that is large compared to the scattering length a_s (which is certainly the case for ^{87}Rb in the first several bands of an optical lattice). The momentum dependence can be ignored whenever the effective range of the true interatomic potential is much smaller than $l_{\rm sp}$ [31], which is also an excellent approximation.
[33] The factor of $\frac{1}{4}$ is so that J_α agrees with the usual hopping integral in the case of nearest neighbor hopping on an isotropic square lattice.
[34] This is strictly true in the thermodynamic limit. For a finite fermionic system an intensive central spin like term exists whenever the number of conduction fermions is odd.
[35] For a given $n > 1$, the Mott lobe is bounded below by $\mathcal{V} + n - 1$ and above by $\mathcal{V} + n$, and it is within these constraints that the coefficient c_2 is strictly positive.
[36] It is also essential that for a superfluid density ψ^2, the depth of the superfluid minimum can become much smaller than $\mathcal{J}\psi^2$ as $\mathcal{B} \to 0$ (i.e., as the Mott transition is approached). If this were not the case, the larger energy gain from the Kondo term at small ψ might not induce a local minimum at $\psi = 0$.