Correlated hopping in the Falicov-Kimball model: A large-dimensions study

Avraham Schiller
Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel
(March 24, 2022)

The Falicov-Kimball model with a correlated-hopping interaction is solved using an extended dynamical mean-field theory that becomes exact in the limit of large dimensions. The effect of correlated hopping is to introduce nonlocal self-energy components that retain full dynamics as $D \to \infty$, thus introducing an explicit $k$-dependence to the single-particle self-energy. An explicit solution for the homogeneous phase at $D = 2$ reveals significant nonlocal dynamical contributions in the physically relevant regime of moderately large correlated-hopping amplitude, indicating that important nonlocal correlations are omitted in Hubbard-like models upon neglecting the correlated-hopping interaction.

PACS numbers: 71.10.Fd, 71.27.+a, 71.30.+h

The development of the limit of infinite dimensions in the Dynamical Mean-Field Theory (DMFT), has opened exciting new possibilities in the study of correlated electron systems. One notable example is the Mott-Hubbard metal-insulator transition realized in $\text{V}_2\text{O}_3$ and $\text{Ca}_1_{-x}\text{Sr}_x\text{VO}_3$, where the main features of the optical and photoemission data are successfully accounted for within the DMFT treatment of the single-band Hubbard model. This approach, however, fails to explain the moderate mass enhancement and the apparent $k$-dependence of the single-particle self-energy close to the transition in these materials, indicating that important nonlocal interactions are either omitted in the Hubbard model, or mistreated in the DMFT.

One nonlocal interaction term clearly absent in the Hubbard model is that of correlated hopping (CH), in which hopping between neighboring lattice sites depends on the occupancy of the opposite spin orientation. Such an interaction term is an integral part of the Coulomb repulsion between electrons, and with estimates ranging from about 0.5eV in transition metals to 0.8eV in the cuprates and up to 3.3eV in benzene, it is certainly comparable to, if not larger than, the corresponding amplitude for single-particle hopping in real systems. Despite this fact and in spite of Hirsch’s suggestion of a new mechanism for superconductivity, there is still no qualititative understanding of the effects that a CH term might have. While some rigorous arguments can be made, these are restricted to certain models in one spatial dimension, and to a few very special cases in higher dimensions. There is thus an obvious need for a careful investigation of CH within a reliable non-perturbative approach, applicable to the entire range of interactions from weak to strong coupling.

In this paper, we generalize the limit of infinite dimensions, $D \to \infty$, as to include the CH interaction. Focusing on the Falicov-Kimball model — a simplified Hubbard model with one spin species tied down — the effect of CH is to introduce nonlocal self-energy components that retain full dynamics as $D \to \infty$. This marks departure from the local approximation. Similar to existing formulations of the infinite-dimensional limit, though, both the local and nonlocal self-energy components are extracted from an effective single-site action, which can still be treated analytically for the Falicov-Kimball model. Solving the resulting DMFT for the homogeneous phase at $D = 2$ and in the physically relevant regime of a moderately large CH amplitude, the dynamical contribution of the nearest-neighbor self-energy is found to be comparable to that of the local self-energy, indicating that important nonlocal correlations are omitted in Hubbard-like models upon neglecting the CH interaction. This not only suggests that CH may be essential for the complete understanding of the metal-insulator transition in $\text{V}_2\text{O}_3$ and $\text{Ca}_1_{-x}\text{Sr}_x\text{VO}_3$, but further calls into question its omission in any Hubbard-like modeling of real systems.

We begin our discussion with the Hamiltonian

\[
\mathcal{H} = \frac{t_1}{\sqrt{2D}} \sum_{<i,j>} \left\{ d_i^\dagger d_j + d_j^\dagger d_i \right\} + U \sum_i d_i^\dagger d_i f_i^\dagger f_i \\
+ \frac{t_2}{\sqrt{2D}} \sum_{<i,j>} \left\{ d_i^\dagger d_j + d_j^\dagger d_i \right\} \left( f_i^\dagger f_i + f_j^\dagger f_j \right) \\
+ (E_f - \mu) \sum_i f_i^\dagger f_i - \mu \sum_i d_i^\dagger d_i,
\]

in which itinerant $d$ fermions interact via an on-site Coulomb repulsion $U$ with localized $f$ fermions on a $D$-dimensional hypercubic lattice. Motion of the $d$ fermions has two components: a single-particle, nearest-neighbor hopping term $t_1$, and a correlated-hopping term $t_2$, in which the hopping amplitude for the $d$ fermions depends on the occupancy of the corresponding $f$ fermions. The two fermion species share a joint chemical potential, $\mu$.

For $t_2 = 0$, Eq. (1) reduces to the Falicov-Kimball model which is perhaps the simplest many-body Hamiltonian to exhibit both a Mott-Hubbard metal-insulator transition and long-range ordered phases. A nonzero $t_2$ modifies the single-particle dynamics of the itinerant fermions in two ways. First, there is a straightforward Hartree renormalization of the static single-particle hopping amplitude according to $t_1 \to t = t_1 + ...$
2nfft_2, where n_f is the f occupation number (assuming a homogeneous phase). In addition, the d and f fermions acquire new self-energies with full dynamics, which survive also in the absence of an on-site U. To investigate the effect of CH we employ the limit of large dimensions, which has proven useful in studying the Falicov-Kimball model. This to end, both the single-particle and CH terms in Eq. (3) have been properly rescaled with D.

For conventional Hubbard-like models, the self-energy is purely local at D = ∞, rendering the local approximation exact. This paradigm breaks down as soon as CH is considered, as exemplified in Fig. 1 for the Hamiltonian of Eq. (3). In Fig. 1, sites j and k are distinct nearest neighbors of site i. Thus, after summation over all equivalent k, the diagram of Fig. 1(a) gives a contribution of order 1/√D to the nearest-neighbor d self-energy Σ_n(i,j), while the diagram of Fig. 1(b) (in which both j and k are fixed) gives a contribution of order D^{-1} to Σ_kj. Both terms are relevant in the limit D → ∞.

A systematic formulation of the D → ∞ limit for the Hamiltonian of Eq. (3) employs the Luttinger-Ward functional Φ — the sum of all vacuum-to-vacuum skeleton diagrams. Consider a generic Φ diagram. Labeling each vertex according to the site index of its incoming and outgoing f lines, we attach the 1/√D factor associated with each CH vertex to its incoming or outgoing nearest-neighbor d line. With this convention we notice that, even after summation over all intermediate sites, any path connecting two vertices at sites i ≠ j gives a contribution of order D^{-∥i−j∥/2} or higher to the diagram at hand (∥i−j∥ is the minimal number of nearest-neighbor steps leading from site i to site j). Given that (i) any two vertices within a Φ diagram can be connected by four independent paths having no lines in common, and (ii) the total number of equivalent sites j scales as D^{∥i−j∥}, the overall contribution of such a diagram (after summation over all equivalent j’s) is O(D^{-∥i−j∥}). Thus, to order D^{-n}, only skeleton diagrams with a maximal inter-vertex distance ∥i−j∥_{max} ≤ n need to be considered. Specifically, only purely local vertices are left at D = ∞.

While the above classification appears identical to that for conventional Hubbard-like models, it should be emphasized that purely local vertices in our convention do not imply purely local diagrams. To the contrary, a CH vertex at site i couples the operators d_i and f_i to the nearest-neighbor shell operator

\[ ψ_i = \frac{1}{\sqrt{2D}} \sum_{\langle j,i \rangle} d_j, \]

in which j is summed over the 2D nearest neighbors of site i. Thus, a CH vertex at site i necessarily introduces at least one propagator that is not confined to site i.

Despite the above nonlocality, the D → ∞ limit remains tractable since Φ decouples in this limit into

\[ Φ = \sum_i Φ_{loc}[G^{dd}_{ii}, G^{\psi d}_{ii}, G^{d\psi}_{ii}, G^{ff}_{ii}], \]

where Φ_{loc} is the functional generated by the action

\[ S_{eff} = -\int_0^\beta d\tau \int_0^\beta d\tau' \sum_{\alpha,\beta=d,\psi} \alpha^\dagger(\tau) [G^{-1}(\tau−\tau')]_{\alpha\beta}(\tau') \]

\[ + \int_0^\beta d\tau \int_0^\beta d\tau' f(\tau') fj \{ f(\tau') + d(\tau') \psi(\tau) + \psi^\dagger(\tau)d(\tau) \} f(\tau) \]

\[ + U \int_0^\beta d\tau \int_0^\beta d\tau' d(\tau)d(\tau') fj. \]

Here G(\tau−\tau') is a 2 × 2 matrix propagator, chosen such that, when dressed with the effective-action self-energy, it coincides with the corresponding lattice propagator:

\[ G^{-1}(i\omega_n) = G^{-1}(i\omega_n) + Σ^D(i\omega_n) \]

with

\[ G(i\omega_n) = \begin{bmatrix} G^{dd}_{ii}(i\omega_n) & G^{d\psi}_{ii}(i\omega_n) \\ G^{\psi d}_{ii}(i\omega_n) & G^{\psi\psi}_{ii}(i\omega_n) \end{bmatrix} \]

[ω_n = π(2n + 1)T are the Matsubara frequencies]. Here, in using the 2 × 2 matrix notation of Eqs. (4)–(6), we have identified the indices α = d, ψ with α = 1, 2, respectively. Contrary to G, which requires knowledge of the fully dressed lattice d propagator, the bare f propagator in Eq. (6) is simply G_f^{-1}(i\omega_n) = i\omega_n + μ − E_f, which stems from the localized nature of the f fermions.

From Eq. (6) it is apparent that Φ at D = ∞ is a functional of the local d and f propagators, as well as the nearest-neighbor and next-nearest-neighbor d Green functions which enter via G_{ii}^{dd}, G_{ii}^{d\psi}, and G_{ii}^{\psi\psi}. Thus, in

![Diagram](image-url)
addition to the local $d$ and $f$ self-energies, one must also consider the nearest- and next-nearest-neighbor $d$ self-energies. From a functional derivative of Eq. (3) one has

$$
\Sigma_{ii}(i\omega_n) = \Sigma_{dd}^S(i\omega_n) + \Sigma_{\psi\psi}^S(i\omega_n), \quad (7a)
$$

$$
\Sigma_{<i,j>}(i\omega_n) = \frac{1}{2D} N_{<i,j>}(\Sigma_{\psi\psi}^d(i\omega_n)), \quad (7b)
$$

$$
\Sigma_{<i,j>}(i\omega_n) = \frac{1}{2D} N_{<i,j>}(\Sigma_{\psi\psi}^d(i\omega_n)), \quad (7c)
$$

where $<i,j>$ denotes distinct lattice sites with a common nearest neighbor. In Eq. (7c), $N_{<i,j>}$ is equal to one for $i$ and $j$ on the same axis; otherwise it is equal to two. Accordingly, the $\mathbf{k}$-dependent $d$ self-energy reads

$$
\Sigma_{\mathbf{k}}(z) = \Sigma_{dd}^S(z) + \epsilon_{\mathbf{k}} \left[ \Sigma_{\psi\psi}^S(z) + \Sigma_{\psi\psi}^d(z) \right] + \epsilon_{\mathbf{k}}^2 \Sigma_{\psi\psi}^S(z),
$$

where $\epsilon_{\mathbf{k}} = \sqrt{2/D} \sum_{i} \cos(k_i)$. Notice that, contrary to conventional formulations of the limit of large dimensions, $\Sigma_{\mathbf{k}}$ has an explicit $\mathbf{k}$ dependence which enters solely through the dimensionless energy $\epsilon_{\mathbf{k}}$. $\Sigma_{\mathbf{k}}(z) \equiv \Sigma_{\mathbf{k}}(z)$.

A self-consistency loop can finally be closed by expressing the lattice Green function of Eq. (4) in terms of $\Sigma_{\mathbf{k}}$. For the spatially homogeneous phase this gives

$$
G(z) = \int_{-\infty}^{\infty} d\epsilon \frac{\rho(\epsilon)}{z + \mu - t_1 \epsilon - \Sigma_{\mathbf{k}}(z)} \begin{bmatrix} 1, \epsilon & \epsilon^2 \end{bmatrix}, \quad (9)
$$

where $\rho(\epsilon)$ is the density of states for the energy $\epsilon_{\mathbf{k}}$, i.e., a Gaussian for $D = \infty$. Equation (4) can further be written in closed form in terms of the bare $(U, t_2 = 0)$ tight-binding $d$ propagator. We also note that Eqs. (3)–(4) were derived under the explicit assumption of a hypercubic lattice with only nearest-neighbor single-particle hopping. Different self-consistency equations will generally apply to other lattices and other tight-binding models.

Equations (3)–(4) are an exact formulation of the limit $D \to \infty$ for the Hamiltonian of Eq. (1). In particular, for $t_2 = 0$, when only $\Sigma_{dd}(i\omega_n)$ is nonzero, they properly reduce to the corresponding equations for the Falicov-Kimball model at $D = \infty$. In the spirit of the DMFT, the same set of equations can also be applied to any finite dimension $D$, provided the actual $D$-dimensional density of states $\rho(\epsilon)$ is used in Eq. (3). Below we present results for $D = 2$. Although designed at present to describe the homogeneous phase, this formulation is easily extended to phases with long-range order, as well as to other lattice models with CH. For example, in the Hubbard model one has to introduce separate $\psi_{\sigma}$ fields for each spin orientation, but the basic formulation remains the same.

A crucial advantage of the Hamiltonian of Eq. (1) over its Hubbard counterpart is that the action of Eq. (4) can be treated analytically. Integrating out the $f$ particles, the lattice Green function of Eq. (3) is expressed as

![Figure 2](image_url)

**FIG. 2.** The $d$ spectral function as a function of $t_2/t_1$, for $n_f + n_d = 1$, $U/t_1 = 4$, $T/t_1 = 1$, and $D = 2$. For both $t_2/t_1 = 0$ and $t_2/t_1 = -1$, the half-filled model is particle-hole symmetric, with a Mott gap for $U/t_1 = 4$. This gap gradually closes upon increasing $t_2/t_1 > 0$ or decreasing $t_2/t_1 < -1$, until the lower and upper Hubbard bands are merged into one broad band. Note the general resemblance between curves with values of $t_2/t_1$ that are symmetric about $t_2/t_1 = -1/2$.

$$
G(i\omega_n) = (1-n_f)G(i\omega_n) + n_f \left[ G^{-1}(i\omega_n) - \begin{pmatrix} U & t_2 \\ t_2 & 0 \end{pmatrix} \right]^{-1},
$$

where $n_f$ is the $f$ occupation number: $n_f = 1/(e^{\beta \epsilon} + 1)$,

$$
\epsilon = E_f - \mu + U/2 - \frac{1}{\beta} \sum_{n=-\infty}^{\infty} \ln \det \left[ 1 - G(i\omega_n) \begin{pmatrix} U & t_2 \\ t_2 & 0 \end{pmatrix} \right]. \quad (11)
$$

Once Eqs. (3), (4), (10), (11) have been iterated and the $f$ occupation number determined, the same equations can be used to calculate $G(z)$ for an arbitrary complex $z$.

Restricting attention to $D = 2$ and $E_f = 0$, we focus hereafter on the homogeneous phase of the half-filled case: $n_f + n_d = 1$, where $n_d$ is the $d$ occupation number. Setting $U/t_1 = 4$ and $T/t_1 = 1$, Fig. 2 shows the evolution of the $d$ spectral function as a function of $t_2/t_1$. For the Falicov-Kimball model, $t_2 = 0$, the $d$ spectrum has a Mott gap. This gap persists throughout the range $-1.4 \lesssim t_2/t_1 \lesssim 1$, although the chemical potential does not always fall inside the gap. For either $t_2/t_1 \gtrsim 1$ or $t_2/t_1 \lesssim -1.4$, the lower and upper Hubbard bands merge into one broad band, reflecting the increase in kinetic energy of the itinerant fermions. Due to the lack of particle-hole symmetry, the $d$ spectrum is generally asymmetric. Indeed, only for $t_2 = 0$ and $t_2/t_1 = -1$ is the half-filled model particle-hole symmetric, in which case the resulting spectrum is temperature independent.

As shown in Fig. 3, one cannot simply absorb $t_2$ into an effective Hartree renormalization of the single-particle hopping within the simpler Falicov-Kimball
model. Specifically, tuning $\mu$ and $E_f$ in the latter model as to maintain the same $n_f$ and $n_d$, there are substantial deviations between the $d$ spectral function for the Hamiltonian of Eq. (1) and that obtained within the DMFT for the Falicov-Kimball model with the renormalized hopping, $t_{FKM} = t_1 + 2t_2 n_f$. Here $n_f$ and $n_d = 1 - n_f$ are kept the same in both models, as are $U/t_1$ and $T/t_1$.

The effect of CH is best seen, though, in the nonlocal dynamical contributions through $\Sigma_{<i,j>$} and $\Sigma_{ii}$, defined from the expansion $\Sigma_{\mathbf{k}} = \Sigma_0 + \Sigma_1 \epsilon_{\mathbf{k}} + \Sigma_2 \epsilon_{\mathbf{k}}^2$ [see Eq. (5)], and related to the local, nearest-neighbor, and next-nearest-neighbor self-energies through $\Sigma_{ii} = \Sigma_0 + \Sigma_2$, $\Sigma_{<i,j>}$, $\Sigma_{\mathbf{k}} = \Sigma_1 / \sqrt{2D}$, and $\Sigma_{<\mathbf{k},\mathbf{p}>} = N \Sigma_{<\mathbf{k},\mathbf{p}>} \Sigma_2 / 2D$ [see Eqs. (6)]. Here $n_f + n_d = 1$, $U/t_1 = 4$, $T/t_1 = 1$, and $D = 2$. In the physically relevant regime, $|t_2/t_1| \gtrsim 1$, the imaginary parts of $\Sigma_1$ and $\Sigma_0$ are comparable in size, indicating that important nonlocal dynamical contributions are omitted in the DMFT upon neglecting $t_2$.

Anticipating a qualitatively similar result for the Hubbard model, CH thus provides a natural mechanism for a strong $k$-dependence in the single-particle self-energy, including for $D \to \infty$. This may prove important, e.g., for the understanding of the apparent $k$-dependence and the moderate mass enhancement close to the metal-insulator transition in $V_{2-x}O_3$ and $Ca_{1-x}Sr_xVO_3$. Study of the Hubbard model along these lines is currently underway.

FIG. 3. The $d$ spectrum for the Hamiltonian of Eq. (1) at $D = 2$ versus the one obtained within the DMFT for the Falicov-Kimball model with the renormalized single-particle hopping, $t_{FKM} = t_1 + 2t_2 n_f$. Here $n_f$ and $n_d = 1 - n_f$ are kept the same in both models, as are $U/t_1$ and $T/t_1$.

FIG. 4. Imaginary part of the $d$ self-energy components $\Sigma_0$, $\Sigma_1$, and $\Sigma_2$, defined from the expansion $\Sigma_{\mathbf{k}} = \Sigma_0 + \Sigma_1 \epsilon_{\mathbf{k}} + \Sigma_2 \epsilon_{\mathbf{k}}^2$ [see Eq. (5)], and related to the local, nearest-neighbor, and next-nearest-neighbor self-energies through $\Sigma_{ii} = \Sigma_0 + \Sigma_2$, $\Sigma_{<i,j>}$, $\Sigma_{\mathbf{k}} = \Sigma_1 / \sqrt{2D}$, and $\Sigma_{<\mathbf{k},\mathbf{p}>} = N \Sigma_{<\mathbf{k},\mathbf{p}>} \Sigma_2 / 2D$ [see Eqs. (6)]. Here $n_f + n_d = 1$, $U/t_1 = 4$, $T/t_1 = 1$, and $D = 2$.

1 W. Metzner and D. Vollhardt, Phys. Rev. Lett. 62, 324 (1989).
2 See, e.g., A. Georges et al., Rev. Mod. Phys. 68, 13 (1996).
3 M. J. Rozenberg et al., Phys. Rev. Lett. 75, 105 (1995).
4 M. J. Rozenberg et al., Phys. Rev. Lett. 76, 4781 (1996).
5 H.-D. Kim et al., Phys. Rev. B 57, 1316 (1998).
6 I. H. Inoue et al., Phys. Rev. Lett. 74 2539 (1995); I. H. Inoue et al., Phys. Rev. B 58 4372 (1998); H. Makino et al., Phys. Rev. B 58 4384 (1998).
7 J. Hubbard, Proc. Roy. Soc. London A 276, 238 (1963).
8 J. Appel et al., Phys. Rev. B 47, 2812 (1993).
9 R. G. Parr et al., J. Chem. Phys. 18, 1561 (1950).
10 J. E. Hirsch, Phys. Lett. A 134, 451 (1989); J. E. Hirsch and F. Marsiglio, Phys. Rev. B 39, 11515 (1989).
11 L. Arrachea and A. A. Aligia, Phys. Rev. Lett. 73, 2240 (1994); L. Arrachea et al., Phys. Rev. Lett. 76, 4396 (1996).
12 I. N. Karnaukhov, Phys. Rev. Lett. 73, 1130 (1994); R. Bariev et al., Europhys. Lett. 32, 85 (1995).
13 A. Schadschneider, G. Su, and J. Zittartz, Z. Phys. B 102, 393 (1997), and references therein.
14 J. de Boer et al., Phys. Rev. Lett. 74, 789 (1995); A. Schadschneider, Phys. Rev. B 51, 10386 (1995).
15 R. Strack and D. Vollhardt, Phys. Rev. Lett. 70, 2637 (1993).
16 A. A. Ovchinnikov, Mod. Phys. Lett. B 7, 1397 (1993).
17 L. M. Falicov and J. C. Kimball, Phys. Rev. Lett. 22, 997 (1969).
18 U. Brandt and C. Mielich, Z. Phys. B. 75, 365 (1989); ibid. 79, 295 (1990); ibid. 82, 37 (1991); U. Brandt and M. P. Urbaneck, ibid. 89, 297 (1992).
19 J. K. Freericks, Phys. Rev. B 47, 9263 (1993); ibid. 48, 14797 (1993).
This stems from the fact that $T$ enters $G(z)$ in our formulation only through $n_f$ and $\mu$, which are fixed at $n_f = 1/2$ and $\mu = U/2$ for both $t_2 = 0$ and $t_2/t_1 = -1$. 

\bibitem{20} E. Müller-Hartmann, Z. Phys. B \textbf{74}, 507 (1989).