Preemptive nematic order, pseudogap, and orbital order in the iron pnictides

R. M. Fernandes, A. V. Chubukov, J. Knolle, I. Eremin, and J. Schmalian

Phys. Rev. B **85**, 024534 — Published 25 January 2012

DOI: 10.1103/PhysRevB.85.024534

Preemptive nematic order, pseudogap, and orbital order in the iron pnictides

R. M. Fernandes,[1] A. V. Chubukov,[2] J. Knolle,[3] I. Eremin,[4] and J. Schmalian[5]

[1]*Department of Physics, Columbia University, New York, New York 10027, USA*
[2]*Department of Physics, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA*
[3]*Max-Planck-Institut für Physik komplexer Systeme, D-01187 Dresden, Germany*
[4]*Institut für Theoretische Physik III, Ruhr-Universität Bochum, D-44801 Bochum, Germany*
[5]*Institut für Theorie der Kondensierten Materie, Karlsruher Institut für Technologie, D-76131 Karlsruhe, Germany*

(Dated: January 12, 2012)

Starting from a microscopic itinerant model, we derive and analyze the effective low-energy model for collective magnetic excitations in the iron pnictides. We show that the stripe magnetic order is generally preempted by an Ising-nematic order which breaks C_4 lattice symmetry but preserves $O(3)$ spin-rotational symmetry. This leads to a rich phase diagram as function of doping, pressure, and elastic moduli, displaying split magnetic and nematic tri-critical points. The nematic transition may instantly bring the system to the verge of a magnetic transition, or it may occur first, being followed by a magnetic transition at a lower temperature. In the latter case, the preemptive nematic transition is accompanied by either a jump or a rapid increase of the magnetic correlation length, triggering a pseudogap behavior associated with magnetic precursors. Furthermore, due to the distinct orbital character of each Fermi pocket, the nematic transition also induces orbital order. We compare our results to various experiments, showing that they correctly address the changes in the character of the magneto-structural transition across the phase diagrams of different compounds, as well as the relationship between the orthorhombic and magnetic order parameters.

I. INTRODUCTION

The current interest in iron-based superconductors (FeSCs) lies not only in the superconducting pairing mechanism, but also in the peculiar normal-state properties of these materials (for reviews, see[1]). Of particular interest is the fact that in weakly doped iron pnictides the stripe spin-density wave order at $T < T_N$, with ordering vectors $(0,\pi)$ or $(\pi,0)$ in the 1-Fe Brillouin zone[2,3], is often preceded by a "nematic" phase with broken C_4 tetragonal symmetry but unbroken $O(3)$ rotational symmetry. The emergence of such a phase is not only manifested by a tetragonal to orthorhombic transition at $T_s \geq T_N$, but also by the onset of significant anisotropies in several quantities[4], such as dc resistivity[5,6], optical conductivity[7,8], local density of states[9], orbital occupancy[10], uniform susceptibility[11], and the vortex core in the mixed superconducting state[12]. The fact that the spin-density-wave and structural transition lines follow each other across all the phase diagrams of 1111 materials (RFeAsO, with rare-earth R) and 122 materials (AeFe$_2$As$_2$,with alkaline-earth Ae), even inside the superconducting dome[13,14], prompted researchers to propose that these phases are intimately connected. The interplay between magnetic and structural transitions in FeSCs is also quite rich: while in 1111 materials the two transitions are second-order and split ($T_s > T_N$), in most of the 122 materials they seem to occur simultaneously or near-simultaneously at small dopings, but clearly split above some critical doping - $x \approx 0.022$ in $Ba\left(Fe_{1-x}Co_x\right)_2 As_2$, see[15,16], and $x \approx 0.039$ in $Ca\left(Fe_{1-x}Co_x\right)_2 As_2$, see[17].

Early theoretical proposals explored two alternative scenarios for the nematic order. One scenario, which borrows concepts from the manganites, is that the driving force is orbital ordering - it induces the structural transition and triggers the magnetic transition at a lower temperature by renormalizing the exchange constants[18–27]. An alternative, magnetic scenario, which borrows concepts from the studies of the $J_1 - J_2$ model of localized spins[28], is build upon the fact that the order parameter manifold for the stripe magnetic order is $O(3) \times Z_2$, with the Z_2 Ising symmetry differentiating between the two possible ordering vectors $(0,\pi)$ and $(\pi,0)$. This scenario explores the possibility that the Z_2 Ising variable orders before the $O(3)$ rotational symmetry is broken, leading to an intermediate phase with a broken tetragonal symmetry but no long-range magnetic order. The Ising order makes the Fermi pockets at $(0,\pi)$ and $(\pi,0)$ non-equivalent and induces orbital order, since the two electron pockets are constituted of different orbitals[29,30]. Furthermore, the breaking of the Z_2 symmetry also makes the x and y directions inside the unit cell inequivalent, inducing a structural instability via the bilinear coupling between the Z_2 order parameter and the orthorhombic distortion (see below). For this reason, the state with broken Ising symmetry is often called an Ising-nematic phase. The magnetic scenario was first applied to localized or nearly-localized spin models for Fe-pnictides[31–33], but was later extended to itinerant systems in a phenomenological way[34,35].

Since orbital order and Ising-nematic order break the same symmetry, one order generates the other, making the experimental distinction between the two scenarios rather subtle. The same is true if one wants to distinguish between itinerant and localized magnetic models, which we will address as well. In both situations, the only real way to distinguish between different approaches

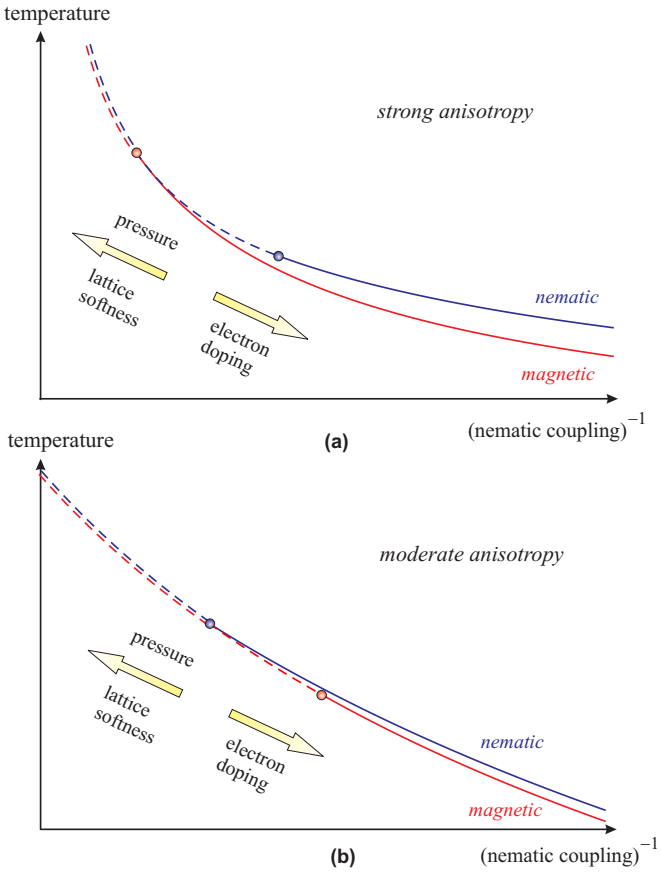

Figure 1: (Color online) Two characteristic phase diagrams of the system behavior as function of temperature and nematic coupling. Panel (a) is for strongly anisotropic, quasi-two dimensional systems, while panel (b) is for moderately anisotropic, less quasi-two dimensional systems. Red (light gray) and blue (dark gray) curves represent magnetic and Ising-nematic transitions, respectively. The arrows show how the inverse nematic coupling changes as function of doping, pressure, and lattice stiffness. Here and in the other figures, a solid (dashed) line denotes a second-order (first-order) transition, and a double-dashed line indicates a simultaneous first-order transition. The two solid points mark the positions of the nematic and magnetic tri-critical points. The difference between the two phase diagrams is in the system behavior at intermediate nematic coupling. In (a), the system displays a first-order nematic transition followed by a second-order magnetic transition at lower temperatures. In (b), the system first undergoes a second-order nematic transition, and then, at smaller temperatures, a meta-nematic transition which triggers a first-order magnetic transition. At the meta-nematic transition, the nematic order parameter jumps between two finite values.

is by explicit calculations followed by comparison to the available experimental results. In this regard, the richness of the phase diagrams of the iron pnictides is an important cornerstone, because one has to explain not only the very existence of the nematic phase, but also how this transition changes as a function of doping, pressure, and material. In addition, one has to address the feedback effect from the nematic order on fermions. In particular, there is growing volume of experimental evidence of pseudogap-like electronic behavior in the iron pnictides, which in some cases seems to emerge at the same temperature at which nematic order sets in[36–39].

In this paper, we argue that the changes in the character of the nematic transition with doping, pressure, and alkaline-earth substitutions, as well as the development of the pseudogap and other experimentally detectable features, can be understood within a magnetic scenario for an itinerant fermionic model. We depart from a multi-band model of interacting fermions and *derive* (instead of assuming) the Ginzburg-Landau (GL) effective action for the two low-energy collective $O(3)$ magnetic degrees of freedom $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ associated with the ordering vectors $(\pi, 0)$ and $(0, \pi)$, respectively. This action contains, besides other terms, the term $g(\mathbf{\Delta}_X^2 - \mathbf{\Delta}_Y^2)^2$ which describes fluctuations in the Ising-nematic channel. We extend the original $O(3)$ model to N field components and study the limit $N = \infty$ within mean-field theory, and arbitrary N by the renormalization group (RG) technique. We find different types of system behavior depending on the strength of the dimensionless parameter $\alpha \propto g^{-1}$ and on the degree of magnetic anisotropy. In all cases, however, we find two distinct multi-critical points in the phase diagram, namely a *magnetic tri-critical point* and a *nematic tri-critical point*.

In the case of strongly anisotropic, quasi-2D systems, the nematic tri-critical point precedes the magnetic one, resulting in three types of behavior - see Fig. 1(a). (i) For small α (largest nematic coupling g) we find a strong first-order Ising-nematic transition at T_s, accompanied by a discontinuous jump of the magnetic correlation length ξ and a jump of the magnetic order parameter to a finite value. In this case, the stripe magnetic order emerges simultaneously with the Ising-nematic order, via a first-order magnetic transition ($T_N = T_s$), but because ξ jumps, there are no critical magnetic fluctuations above the transition. (ii) For intermediate α we find a first-order preemptive Ising-nematic transition accompanied by a discontinuous increase of the magnetic correlation length to a larger but still finite value, such that the stripe magnetic order does not develop at T_s and emerges only at a smaller T. In this case, the magnetic transition is second-order. (iii) For large α (small nematic coupling g) we find a second-order preemptive Ising-nematic transition followed by a second-order magnetic stripe transition at a smaller T. Near the nematic tri-critical point we find that the Ising-nematic order emerges very rapidly such that, in practice, the second-order Ising transition is hard to distinguish from a first-order transition. The behavior of the nematic and magnetic order parameters for the phase diagram in Fig. 1(a) is shown in Fig. 9(d), (c), and (b) for regimes (i), (ii), and (iii), respectively.

In the case of moderately anisotropic, less quasi-2D systems, the positions of the nematic and magnetic tri-critical points can be reversed, resulting in a new behavior in the regime of intermediate α, see Fig.1(b).

The regimes (i) and (iii) are still present, when both the nematic and magnetic transitions are first-order and second-order, respectively. However, in regime (ii) the upper Ising-nematic transition is second-order and at a lower T there is a meta-nematic transition (i.e. the nematic order parameter undergoes a finite jump from one non-zero value to the other), which induces a first-order magnetic transition. The behavior of the nematic and magnetic order parameters for the phase diagram in Fig. 1(b) is shown in Fig. 15(c), (b), and (a) for regimes (i), (ii), and (iii), respectively.

As shown in Figs. 1, we find from our microscopic model that α decreases with pressure but increases with increasing lattice stiffness and electron doping, i.e. larger electron doping results in a larger value of α. Our results then predict that the Ising-nematic and magnetic transitions split upon electron doping, but tend to remain simultaneous and first-order upon pressure, in agreement with the experimental data for FeSCs. We also show that the nematic order parameter couples linearly to the orbital polarization and lattice distortion, and hence nematic order generates orbital and structural order. We argue, however, that the sign of the orbital order may differ between hole and electron-doped materials.

We also argue that, for the phase diagram in Fig. 1a, the discontinuous (or nearly-discontinuous) increase of the magnetic correlation length at T_s in the regimes (ii) and (iii) greatly increases the strength of thermal magnetic fluctuations. Once enhanced, these fluctuations account for spectral weight redistribution (magnetic precursors), what gives rise to pseudogap behavior in the fermionic spectral function and other observables. We present more detailed comparison with the data later in the paper and also compare our results with earlier studies of Ising-nematic order.

In the itinerant picture which we adopt here, the nature of the Ising-nematic phase has a clear interpretation in terms of magnetic fluctuations. The system has two degenerate stripe magnetic ground states with ordering vectors $(\pi, 0)$ and $(0, \pi)$, described by the two order parameters $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$. At high temperatures $\langle \mathbf{\Delta}_X \rangle = \langle \mathbf{\Delta}_Y \rangle = 0$, and the fluctuations of each order parameter have equal strength, i.e., $\langle \mathbf{\Delta}_X^2 \rangle = \langle \mathbf{\Delta}_Y^2 \rangle$. The Ising-nematic phase emerges when fluctuations associated with one of the ordering vectors become stronger than the other, $\langle \mathbf{\Delta}_X^2 \rangle \neq \langle \mathbf{\Delta}_Y^2 \rangle$, while still $\langle \mathbf{\Delta}_X \rangle = \langle \mathbf{\Delta}_Y \rangle = 0$ (see Fig. 2). Because there are two possible choices, $\langle \mathbf{\Delta}_X^2 \rangle > \langle \mathbf{\Delta}_Y^2 \rangle$ or $\langle \mathbf{\Delta}_X^2 \rangle < \langle \mathbf{\Delta}_Y^2 \rangle$, the symmetry which breaks down at the Ising-nematic phase transition is a Z_2 symmetry. Once magnetic fluctuations around $(\pi, 0)$ become stronger (or weaker) than the fluctuations around $(0, \pi)$, the equivalence between x and y directions inside the unit cell breaks down. The Ising-nematic transition then triggers the transition from a tetragonal to an orthorhombic phase, and also imposes orbital order because the Fermi pockets centered at $(\pi, 0)$ and $(0, \pi)$ have different orbital character. Furthermore, this anisotropy in the spectrum of magnetic fluctuations leads to anisotropic scattering of electrons, resulting in different in-plane resistivities along x and y directions[40].

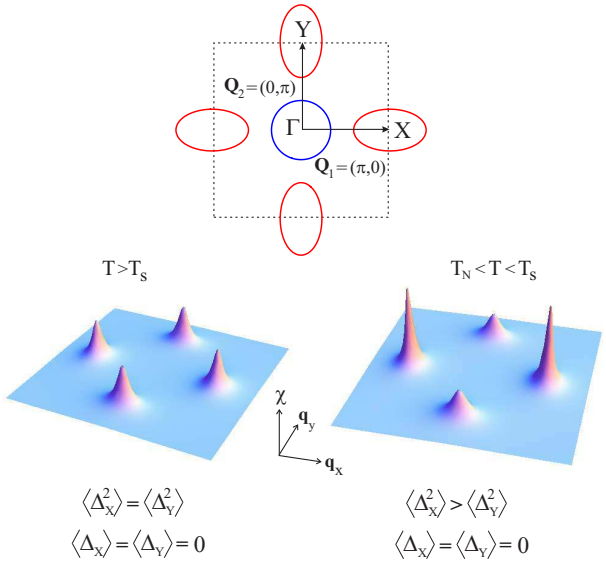

Figure 2: (Color online) (*upper panel*) The band-structure with a circular hole pocket at Γ and two electron pockets at X and Y. The Brillouin zone contains one Fe atom. (*lower panels*) Static magnetic susceptibility $\chi_{\mathbf{q}}$ across the Brillouin zone for different temperatures. At high temperatures, fluctuations near the two stripe magnetic ordering vectors are equally strong, $\langle \Delta_X^2 \rangle = \langle \Delta_Y^2 \rangle$. Above the magnetic ordering temperature T_N but below the Ising-nematic ordering temperature T_s, fluctuations associated with one of the stripe states become stronger (in the figure, $\langle \Delta_X^2 \rangle > \langle \Delta_Y^2 \rangle$) and the tetragonal symmetry is broken inside the unit cell. Stronger fluctuations around one ordering vector yield stronger intensity and narrower peaks.

The structure of the paper is the following. In Section II we derive the effective action for the nematic order parameter using a two-stage Hubbard-Stratonovich transformation. In Sec. III we extend the effective action to arbitrary N, where N is the number of components of the magnetic order parameter ($N = 3$ in the original $O(3)$-isotropic model). We solve for the nematic order in the mean-field approximation, justified in the $N \to \infty$ limit, obtain different regimes of system behavior and discuss each regime in some detail. In Sec.IV we study the effective action at a finite N using the renormalization group (RG) technique. We find that the system behavior for $N = 3$ is the same as for large N, but there is a change in the RG flow structure at $N \leq 2$. In Sec. V we discuss the interplay between Ising-nematic, orbital, and structural order, and show how the nematic order gives rise to pseudogap behavior despite the fact that the Z_2 order parameter has zero momentum and does not by itself reconstructs low-energy fermionic states. In Sec. VI we compare our theory with the experiments and with other theoretical works on the nematic order. We present our conclusions in Sec.VII.

II. MICROSCOPIC MODEL AND THE EFFECTIVE ACTION

We start from the minimal four-band model with two hole pockets Γ at the center of the Brillouin zone and two electron pockets X and Y at $\mathbf{Q}_1 = (\pi, 0)$ and $\mathbf{Q}_2 = (0, \pi)$, respectively (see Fig. 2). We follow Ref.[41] and consider that only one of the two hole pockets participate in the SDW state. The extension to the case when both Γ pockets are involved in the SDW reconstruction is straightforward and does not change the analysis below except for the renormalization of the couplings.

For simplicity, we consider parabolic dispersions with $\varepsilon_{\Gamma,\mathbf{k}} = \varepsilon_0 - \frac{k^2}{2m} - \mu$, $\varepsilon_{X,\mathbf{k}+\mathbf{Q_1}} = -\varepsilon_0 + \frac{k_x^2}{2m_x} + \frac{k_y^2}{2m_y} - \mu$, and $\varepsilon_{Y,\mathbf{k}+\mathbf{Q_2}} = -\varepsilon_0 + \frac{k_x^2}{2m_y} + \frac{k_y^2}{2m_x} - \mu$, where m_i denotes the band masses, ε_0 is the offset energy, and μ is the chemical potential. Near the Fermi energy and for small ellipticity, the dispersions can be approximated by $\varepsilon_{\Gamma,\mathbf{k}} = -\varepsilon$, $\varepsilon_{X,\mathbf{k}+\mathbf{Q_1}} = \varepsilon - \delta_0 + \delta_2 \cos 2\theta$, $\varepsilon_{Y,\mathbf{k}+\mathbf{Q_2}} = \varepsilon - \delta_0 - \delta_2 \cos 2\theta$, with $\delta_0 = 2\mu$, $\delta_2 = \varepsilon_0 m(m_x - m_y)/(2m_x m_y)$, and $\theta = \tan^{-1} k_y/k_x$[42].

Electrons with spin α of the band i are created by the operators $c^\dagger_{i,\mathbf{k}\alpha}$, yielding the non-interacting Hamiltonian:

$$\mathcal{H}_0 = \sum_{i,\mathbf{k}} \varepsilon_{i,\mathbf{k}} c^\dagger_{i,\mathbf{k}\alpha} c_{i,\mathbf{k}\alpha} \tag{1}$$

Here and for the rest of the paper the summation over repeated spin indices is assumed, and we shift the momenta of the fermions near the X and Y Fermi pockets by $\mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively, i.e. $\varepsilon_{X,\mathbf{k}+\mathbf{Q_1}} \to \varepsilon_{X,\mathbf{k}}$, $\varepsilon_{Y,\mathbf{k}+\mathbf{Q_2}} \to \varepsilon_{Y,\mathbf{k}}$.

As discussed in Ref.[43], this model has eight fermionic interactions u_n which can be decomposed into the spin density-wave (SDW), the charge density-wave (CDW) and the pairing channels. Since our goal is to study stripe magnetism and the accompanied Ising-nematic order, we keep only the interactions in the spin channel with momenta near $\mathbf{Q}_1$ and $\mathbf{Q}_2$, restricting the interacting Hamiltonian to

$$\mathcal{H}_{\rm int} = -\frac{1}{2} u_{\rm spin} \sum_{i,\mathbf{q}} \mathbf{s}_{i,\mathbf{q}} \cdot \mathbf{s}_{i,-\mathbf{q}} \tag{2}$$

where $\mathbf{s}_{i,\mathbf{q}} = \sum_k c^\dagger_{\Gamma,\mathbf{k}+\mathbf{q}\alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{i,\mathbf{k}\beta}$ is the electronic spin operator, with Pauli matrices $\boldsymbol{\sigma}_{\alpha\beta}$. The coupling $u_{\rm spin}$ is the combination of density-density and pair-hopping interactions between hole and electron states (u_1 and u_3 terms in the notation of Ref.[44]):

$$\begin{aligned} u_1 c^\dagger_{\Gamma,\alpha} c_{\Gamma,\alpha} c^\dagger_{X,\beta} c_{X,\beta} &= -\frac{u_1}{2} c^\dagger_{\Gamma,\alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{X,\beta} \cdot c^\dagger_{X,\gamma} \boldsymbol{\sigma}_{\gamma\delta} c_{\Gamma,\delta} \\ &\quad + (\cdots) \\ u_3 c^\dagger_{\Gamma,\alpha} c_{X,\alpha} c^\dagger_{\Gamma,\beta} c_{X,\beta} &= -\frac{u_3}{2} c^\dagger_{\Gamma,\alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{X,\beta} \cdot c^\dagger_{X,\gamma} \boldsymbol{\sigma}_{\gamma\delta} c_{\Gamma,\delta} \\ &\quad + (\cdots) \end{aligned} \tag{3}$$

where the dots stand for the terms with $\delta_{\alpha,\beta}\delta_{\gamma,\delta}$, which only contribute to the CDW channel. Combining the two contributions for the SDW channel, we find $u_{\rm spin} = u_1 + u_3$. Once $u_{\rm spin}$ exceeds some critical value (which gets smaller when δ_0 and δ_2 decrease), static magnetic susceptibility diverges at $(0, \pi)$ and $(\pi, 0)$, and the system develops long-range magnetic order. An excitonic-type SDW instability in Fe-pnictides, resulting from the interaction between hole and electron pockets, has been considered by several authors[41,45–53].

Our calculations are done in two steps. In the first step, we introduce the bosonic fields $\boldsymbol{\Delta}_{(X,Y)} \propto \sum_{\mathbf{k}} c^\dagger_{\Gamma,\mathbf{k}\alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{(X,Y),\mathbf{k}\beta}$ for the collective magnetic degrees of freedom, integrate out the fermions, and obtain a Ginzburg-Landau (GL) action for $\boldsymbol{\Delta}_X$ and $\boldsymbol{\Delta}_Y$. We show, in agreement with earlier results[41], that in a mean-field approximation only one of the magnetic order parameters - either $\langle \boldsymbol{\Delta}_X \rangle$ or $\langle \boldsymbol{\Delta}_Y \rangle$ - becomes non-zero in the magnetically ordered state. This leads to stripe-type SDW order in which spins are ordered ferromagnetically in one direction and antiferromagnetically in the other, i.e. the ordering momentum is either $(\pi, 0)$ or $(0, \pi)$. In the second step, we include fluctuations of the $\boldsymbol{\Delta}_{X,Y}$ fields, introduce the collective Ising-nematic bosonic variable $\phi \propto \Delta_X^2 - \Delta_Y^2$ together with $\psi \propto \boldsymbol{\Delta}_X^2 + \boldsymbol{\Delta}_Y^2$, integrate over $\boldsymbol{\Delta}_X$ and $\boldsymbol{\Delta}_Y$, and obtain an effective action in terms of ϕ and ψ. We analyze this action and check whether the system develops an instability towards $\langle \phi \rangle \neq 0$ before $\langle \boldsymbol{\Delta}_X \rangle$ or $\langle \boldsymbol{\Delta}_Y \rangle$ becomes non-zero.

A. The action in terms of $\boldsymbol{\Delta}_X$ and $\boldsymbol{\Delta}_Y$

A straightforward way to obtain the action in terms of $\langle \boldsymbol{\Delta}_X \rangle$ and $\langle \boldsymbol{\Delta}_Y \rangle$ is to start with the fermionic Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm int}$ in Eqs. (1) and (2), write the partition function as the integral over Grassmann variables:

$$Z \propto \int dc_{i,\mathbf{k}} dc^\dagger_{i,\mathbf{k}} \mathrm{e}^{-\beta \mathcal{H}} \tag{4}$$

and then decouple the quartic term in fermionic operators using the Hubbard-Stratonovich transformation:

$$\mathrm{e}^{\frac{ax^2}{2}} = \frac{1}{\sqrt{2\pi a}} \int dy\, \mathrm{e}^{\left(-\frac{y^2}{2a} + yx\right)} \tag{5}$$

where, in our case, $x = \mathbf{s}_{i,\mathbf{0}} = \sum_k c^\dagger_{\Gamma,\mathbf{k}\alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{(X,Y),\mathbf{k}\beta}$ and $y = \boldsymbol{\Delta}_{(X,Y)}$. We then integrate Eq. (4) over fermionic variables using the fact that after the Hubbard-Stratonovich transformation the effective action becomes quadratic with respect to the fermionic operators. The result of the integration is recast back into the exponent and the partition function is expressed as:

$$Z \propto \int d\boldsymbol{\Delta}_X d\boldsymbol{\Delta}_Y \mathrm{e}^{-S_{\rm eff}[\boldsymbol{\Delta}_X, \boldsymbol{\Delta}_Y]} \tag{6}$$

If the relevant $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ are small, which we assume to hold even if the magnetic transition is first-order (we present the conditions on the parameters below), one can expand $S_{\rm eff}\left[\mathbf{\Delta}_X, \mathbf{\Delta}_Y\right]$ in powers of $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ and obtain the Ginzburg-Landau type of action for the order parameters $\mathbf{\Delta}_X, \mathbf{\Delta}_Y$. For uniform $\mathbf{\Delta}_i$, the most generic form of $S_{\rm eff}\left[\mathbf{\Delta}_X, \mathbf{\Delta}_Y\right]$ is

$$\begin{aligned} S_{\rm eff}\left[\mathbf{\Delta}_X, \mathbf{\Delta}_Y\right] &= r_0\left(\mathbf{\Delta}_X^2+\mathbf{\Delta}_Y^2\right)+\frac{u}{2}\left(\mathbf{\Delta}_X^2+\mathbf{\Delta}_Y^2\right)^2 \\ &-\frac{g}{2}\left(\mathbf{\Delta}_X^2-\mathbf{\Delta}_Y^2\right)^2+v\left(\mathbf{\Delta}_X\cdot\mathbf{\Delta}_Y\right)^2 \end{aligned} \quad (7)$$

Carrying out this procedure, we obtain the coefficients r_0, u, g, and v in terms of the non-interacting fermionic propagators convoluted with Pauli matrices (details can be found in Appendix A). The coefficient v vanishes in our model because of the anti-commutation property of the Pauli matrices: $\sigma^i\sigma^j+\sigma^j\sigma^i=0$ for $i\neq j$. To get a non-zero v, one needs to include direct interactions between the two electron pockets[41]. The other three prefactors are expressed via fermionic propagators $G^{-1}_{j,\mathbf{k}}=i\omega_n-\xi_{j,\mathbf{k}}$ as

$$\begin{aligned} r_0 &= \frac{2}{u_{\rm spin}}+2\int_k G_{\Gamma,k}G_{X,k} \\ u &= \frac{1}{2}\int_k G^2_{\Gamma,k}\left(G_{X,k}+G_{Y,k}\right)^2 \\ g &= -\frac{1}{2}\int_k G^2_{\Gamma,k}\left(G_{X,k}-G_{Y,k}\right)^2 \end{aligned} \quad (8)$$

where $\int_k = T\sum_n \int \frac{d^dk}{(2\pi)^d}$ and $k=(\mathbf{k},\omega_n)$, with momentum $\mathbf{k}$ and Matsubara frequency $\omega_n=(2n+1)\pi T$. Similar coefficients were found in Ref.[54], which focused on the magnetic instabilities in a two-band model.

Evaluating the momentum integrals and summing over Matsubara frequencies we obtain that $\int_k G_{\Gamma,k}G_{X,k}$ is negative and at perfect nesting diverges as $N_F\log\Lambda/T$, where $N_F\approx m/(2\pi)$ is the density of states of the hole pocket and Λ is the upper cutoff for the low-energy theory - i.e. the scale at which corrections to the parabolic dispersion become of the order one. Away from perfect nesting, the logarithm saturates at $T=0$ at some finite but still large value[55]. As a result, r_0 decreases with decreasing T and, if the amplitude of the interaction $u_{\rm spin}$ is not too small, changes sign at some temperature $T=T_{N,0}$. This is the mean-field transition temperature below which magnetic order appears. Near $T_{N,0}$ one can expand r_0 as $r_0=a(T-T_{N,0})$, with $a>0$.

The GL action can be straightforwardly extended to include the momentum and frequency dependence of the $\mathbf{\Delta}$ fields, $\mathbf{\Delta}_i\to\mathbf{\Delta}_i(q)$. This extension does not modify in any relevant way the pre-factors of the quartic terms, but changes the pre-factor for the quadratic term at small q to:

$$\begin{aligned} r_0\to\chi^{-1}_{i,q} &= \frac{2}{u_{\rm spin}}+2\int_k G_{\Gamma,k}G_{i,k+q} \\ &= r_0+\gamma\left|\nu_n\right|+f_{\mathbf{q}} \end{aligned} \quad (9)$$

Here $i=X,Y$, $\nu_n=2\pi Tn$ is the bosonic Matsubara frequency, γ is the Landau damping coefficient, and $f_{\mathbf{q}}$ is in general an anisotropic function of $\mathbf{q}$, $f_{\mathbf{q}}=q_x^2(1\pm\eta)+q_y^2(1\mp\eta)+\eta_z q_z^2$, with $-1<\eta<1$ and upper (lower) signs referring to Δ_X (Δ_Y). We verified that all results which we obtain below do not depend on whether η is finite or zero (see Appendix B). To simplify the presentation, we then set $\eta=0$ and approximate $\chi^{-1}_{i,q}$ by $r_0+\gamma|\nu_n|+q_\parallel^2+\eta_z q_z^2$, the same for $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$. Below we consider this anisotropic 3D dispersion and also an isotropic dispersion in dimensions $2\leq d\leq 4$, in which case $q_\parallel^2+\eta_z q_z^2$ is replaced by the d-dimensional momentum amplitude q^2.

For the two fourth-order terms in (8) we obtain at a finite T, expanding around perfect nesting to leading order in the chemical potential μ and in the mass anisotropy $\delta m=m(m_x-m_y)/(2m_xm_y)$:

$$\begin{aligned} u &\approx \frac{7\zeta(3)N_F}{4\pi^2T^2} \\ g &\approx 0.024u\left(\frac{\varepsilon_0\delta m}{T}\right)^2 \end{aligned} \quad (10)$$

for $\delta m\ll T/\varepsilon_0\ll 1$. At $T=0$, we obtain

$$\begin{aligned} u &\approx \frac{N_F}{4\mu^2} \\ g &\approx u\left(\frac{\varepsilon_0\delta m}{2\mu}\right)^2 \end{aligned} \quad (11)$$

for $\delta m\ll|\mu|/\varepsilon_0\ll 1$. In general, $u>0$ and $u>g$, i.e. the action (7) increases when either Δ_X or Δ_Y becomes large. If this condition was not satisfied, the expansion of $S_{\rm eff}$ in powers of $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ would only make sense if higher order terms were considered. The crucial result for our consideration is that g vanishes for circular electron Fermi surfaces ($\delta m=0$), but is positive for any non-zero ellipticity, independent on the sign of ϵ_0 and on whether m_x is larger or smaller than m_y.

The action $S_{\rm eff}$ is exact and includes all fluctuations of the two bosonic fields. Before we consider these fluctuations, let us analyze Eq. (7) in the mean-field approximation. The conventional way to justify mean-field theory is to extend the original one-flavor fermionic model to L flavors, such that $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ couple to L clones of $\mathbf{s}_{i,\mathbf{q}=\mathbf{0}}$, and take the limit $L\to\infty$. To do this properly, one has to rescale simultaneously $u_{\rm spin}\to u_{\rm spin}/L$. After this rescaling, the effective action can be written as $S_{\rm eff}=L\tilde{S}_{\rm eff}$, where $\tilde{S}_{\rm eff}$ is the same as Eq. (7), but with rescaled $u_{\rm spin}$. Because of the overall factor L the action $\tilde{S}_{\rm eff}$ can be approximated by its value at the minimum, as corrections to the partition function from fluctuations of $\mathbf{\Delta}_i$ are small in $(\log L)/L$.

After solving for the minimum of $S_{\rm eff}\left[\mathbf{\Delta}_X, \mathbf{\Delta}_Y\right]$ in Eq. (7), we find that, when $g=0$, the ground state has a huge degeneracy because any configuration $\mathbf{\Delta}=\langle\mathbf{\Delta}_X\rangle e^{i\mathbf{Q}_1\cdot\mathbf{r}}+\langle\mathbf{\Delta}_Y\rangle e^{i\mathbf{Q}_1\cdot\mathbf{r}}$ with $\langle\mathbf{\Delta}_X\rangle^2+\langle\mathbf{\Delta}_Y\rangle^2=-r_0/u$ minimizes $\tilde{S}_{\rm eff}$. A non-zero g gives rise to the additional coupling

$2g\Delta_X^2\Delta_Y^2$, which breaks this degeneracy. For a positive g, this term favors the states in which only one order parameter has a nonzero value, i.e. configurations with either $\langle \mathbf{\Delta}_X \rangle \neq 0$ or $\langle \mathbf{\Delta}_X \rangle \neq 0$, but not both.

To relate $\langle \mathbf{\Delta}_X \rangle$ or $\langle \mathbf{\Delta}_Y \rangle$ to the magnetic ordering, we return to the effective action written in terms of the double functional integral over the fermionic and Δ fields. Assuming that the electronic spin $\mathbf{s}_{i,\mathbf{q}=0} = \sum_{\mathbf{k}} c^\dagger_{\Gamma,\mathbf{k}\alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{i,\mathbf{k}\beta}$ acquires a non-zero magnitude $\langle s_{i,\mathbf{0}} \rangle \neq 0$, we again minimize S_{eff}, but this time with $\langle s_{i,\mathbf{0}} \rangle$ as a parameter, yielding $\langle s_{i,\mathbf{0}} \rangle = \langle \Delta_i \rangle / u_{\text{spin}}$. Since a non-zero $s_{X,\mathbf{0}}$ ($s_{Y,0}$) implies magnetic order with the momentum $\mathbf{Q}_1 = (\pi, 0)$ ($\mathbf{Q}_2 = (0, \pi)$), the fact that only one of the $\mathbf{\Delta}_i$ orders means that the magnetic ordering has a particular momentum. One can easily verify that in such a state spins order ferromagnetically along one direction and antiferromagnetically along the other one.

B. The action in terms of the Ising-nematic order parameter

Since the action (7) is invariant with respect to the interchange between $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$, the onset of either $(\pi, 0)$ or $(0, \pi)$ SDW state breaks not only the conventional $O(3)$ spin-rotational symmetry, but also the additional Z_2 (Ising) symmetry associated with choosing between $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ (Refs.[28,32,33]). The issue we now consider is whether the Z_2 symmetry breaking preempts the $O(3)$ symmetry breaking, i.e. it happens before the Ginzburg-Landau parameter r_0 changes sign and the magnetic order sets in.

Such a Z_2 symmetry breaking without magnetic order would imply that fluctuations associated with one of the bosonic fields are larger than the fluctuations associated with the other one, e.g., $\langle \Delta_X^2 \rangle > \langle \Delta_Y^2 \rangle$ while $\langle \Delta_X \rangle = \langle \Delta_Y \rangle = 0$. A direct experimental detection of this state could be done by performing inelastic neutron scattering in detwinned samples and measuring the spectrum at $(\pi, 0)$ and $(0, \pi)$.

That the action (7) can potentially lead to a preemptive Ising-nematic instability is evident from the presence of the term $g\left(\Delta_X^2 - \Delta_Y^2\right)^2$, which can give rise to an ordered state with $\langle \Delta_X^2 \rangle - \langle \Delta_Y^2 \rangle \neq 0$ in a way similar to how the $\boldsymbol{s}_{i,\mathbf{q}} \boldsymbol{s}_{i,-\mathbf{q}}$ term in the Hamiltonian (2) gives rise to a state with non-zero $\langle \boldsymbol{s}_{i,0} \rangle \neq 0$. The preemptive Ising-nematic instability, however, does not appear in the mean-field limit of $L \to \infty$ fermionic flavors simply because magnetic fluctuations are absent at $L = \infty$, and a non-zero $\langle \Delta_i^2 \rangle \neq 0$ appears simultaneously to $\langle \Delta_i \rangle \neq 0$, once r_0 changes sign. However, it may well happen once we return to the original model with $L = 1$ fermionic flavor and include magnetic fluctuations.

To study a potential preemptive Z_2 transition, we need to introduce collective variables of the fields $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$. Let us introduce auxiliary scalar fields ϕ for $\mathbf{\Delta}_X^2 - \mathbf{\Delta}_Y^2$ and ψ for $\mathbf{\Delta}_X^2 + \mathbf{\Delta}_Y^2$. The field ψ always has a non-zero expectation value $\langle \psi \rangle \neq 0$, which describes Gaussian corrections to the magnetic susceptibility $\chi_{i,q}^{-1}$ in Eq. 9. Meanwhile, the field ϕ may or may not have a non-zero expectation value. If it does, it generates a non-zero value of $\left\langle \mathbf{\Delta}_X^2 - \mathbf{\Delta}_Y^2 \right\rangle$ and the system develops an Ising-nematic order.

The effective action in terms of ϕ and ψ is obtained by using again the Hubbard-Stratonovich transformation of Eq. (4), but this time the variable x is either $\Delta_X^2 + \Delta_Y^2$ or $\Delta_X^2 - \Delta_Y^2$. After applying this transformation, we express the partition function in terms of double integrals over the fields $(\mathbf{\Delta}_X, \mathbf{\Delta}_Y)$ and (ϕ, ψ):

$$Z \propto \int d\mathbf{\Delta}_X d\mathbf{\Delta}_Y d\phi \, d\psi \, \mathrm{e}^{-S_{\text{eff}}[\mathbf{\Delta}_i, \phi, \psi]} \tag{12}$$

where:

$$\begin{aligned} S_{\text{eff}}\left[\mathbf{\Delta}_i, \phi, \psi\right] \quad &= \int_q \chi_q^{-1} \left(\Delta_X^2 + \Delta_Y^2\right) + \int_x \left(\frac{\phi^2}{2g} - \frac{\psi^2}{2u}\right) + \\ &\int_x \psi \left(\Delta_X^2 + \Delta_Y^2\right) + \int_x \phi \left(\Delta_X^2 - \Delta_Y^2\right) \end{aligned} \tag{13}$$

Once ϕ becomes non-zero, we have from (13)

$$\left\langle \Delta_X^2 \right\rangle - \left\langle \Delta_Y^2 \right\rangle = \frac{\phi}{g} \tag{14}$$

If the magnetic long-range order is not developed, i.e. $\langle \Delta_i \rangle = 0$, then it is straightforward to integrate over the fields $(\mathbf{\Delta}_X, \mathbf{\Delta}_Y)$. Carrying out the integration we obtain the effective action in terms on ϕ and ψ in the form

$$S_{\text{eff}}\left[\phi, \psi\right] = \int_q \left\{ \frac{\phi^2}{2g} - \frac{\psi^2}{2u} + \frac{3}{2} \log\left[\left(\chi_q^{-1} + \psi\right)^2 - \phi^2\right] \right\} \tag{15}$$

We later modify the derivation of $S_{\text{eff}}\left[\phi, \psi\right]$ to the case where the system has magnetic order. In the next two sections we will study the effective action (15) in its own mean-field theory and then using a renormalization-group (RG) formalism.

III. MEAN-FIELD THEORY

To justify a mean-field treatment of S_{eff} from (15) we do a trick similar to the one before, but with the bosonic rather than the fermionic variables. Namely, we extend the number of components of the fields $\mathbf{\Delta}_{X,Y}$ from the original value $N = 3$ to arbitrary N, assume that the ϕ and ψ fields interact equally with all components of $\mathbf{\Delta}_{X,Y}$, and rescale the coupling constants to $g \to g/N$ and $u \to u/N$. The effective action (15) then has an overall pre-factor N implying that for large N it can be analyzed by just taking its value at the extremum.

We emphasize that the extension to $N \to \infty$ components is a different approximation than the previous one, in which we made $L \to \infty$ copies of the fermionic fields.

For $L \to \infty$, the terms of the effective action containing the field ϕ have an overall pre-factor of $1/L \to 0$. This means that fluctuations of ϕ are large, preventing a preemptive Ising-nematic order. On the other hand, by abandoning the large L limit and taking instead the $N \to \infty$ limit, we allow strong fluctuations of both Δ_X^2 and Δ_Y^2, but eliminate fluctuations of the scalar fields ϕ and ψ which only account for a small correction, of order $(\log N)/N$, to the partition function. Rewriting the action of Eq. (15) for N components of the Δ_i field gives:

$$S_{\text{eff}}[\phi,\psi] = N\int_q \left\{ \frac{\phi^2}{2g} - \frac{\psi^2}{2u} + \frac{1}{2}\log\left[\left(\chi_q^{-1}+\psi\right)^2 - \phi^2\right]\right\} \tag{16}$$

The mean-field theory for this action is the saddle-point solution of (16), i.e. the minimum of $S_{\text{eff}}[\phi,\psi]$ with respect to ϕ, and the maximum with respect to ψ. That $e^{-S_{\text{eff}}[\phi,\psi]}$ increases when ψ gets larger is related to the fact that Gaussian corrections to r_0 coming from the quartic term $(\Delta_X^2+\Delta_Y^2)^2$ are confined to the upper cutoff of the theory. These corrections just renormalize r_0 by a constant, independent on ϕ, and play no role in our analysis. What matters to us is how ψ is affected by ϕ and vise versa. These mutually influencing terms are independent on the upper cutoff and are therefore well described within the mean-field theory. Differentiating Eq. (16) with respect to ϕ and ψ and taking $\partial S_{\text{eff}}[\phi,\psi]/\partial\phi = \partial S_{\text{eff}}[\phi,\psi]/\partial\psi = 0$, we obtain two non-linear coupled equations for ϕ and ψ:

$$\begin{aligned} \frac{\psi}{u} &= \int_q \frac{r_0+\psi+q^2+\gamma|\nu_n|}{\left(r_0+\psi+q^2+\gamma|\nu_n|\right)^2-\phi^2} \\ \frac{\phi}{g} &= \int_q \frac{\phi}{\left(r_0+\psi+q^2+\gamma|\nu_n|\right)^2-\phi^2} \end{aligned} \tag{17}$$

It is convenient to re-express the first formula as a self-consistent equation for the renormalized mass of the bosonic field, $r = r_0+\psi$. In the paramagnetic phase, $r \propto \xi^{-2}$, where ξ is the magnetic correlation length. It is also useful to remove the high-energy contribution to ψ (i.e. the contribution coming from the upper cutoff), by incorporating it into the renormalization of r_0. Specifically, we rewrite the first equation as:

$$\begin{aligned} r &= r_0+\psi = r_0 + u\int_q \frac{r+q^2+\gamma|\nu_n|}{\left(r+q^2+\gamma|\nu_n|\right)^2-\phi^2} \\ &= \bar{r}_0 + u\int_q\left[\frac{r+q^2+\gamma|\nu_n|}{\left(r+q^2+\gamma|\nu_n|\right)^2-\phi^2} - \frac{1}{q^2+\gamma|\nu_n|}\right] \end{aligned} \tag{18}$$

where $\bar{r}_0 = r_0 + u\int_q \frac{1}{q^2+\gamma|\nu_n|}$. For classical systems, only the $\nu_n = 0$ term matters. The remaining momentum integral in the last line in (18) is infrared and ultraviolet convergent for $2 < d < 4$, such that the upper limit of integration can be safely extended to infinity. Since r_0 and $\bar{r}_0$ differ only by a constant, $\bar{r}_0$ is also a monotonic increasing function of T and can be expressed as $\bar{r}_0 = \bar{a}\left(T - \bar{T}_{N,0}\right)$. For quantum systems such a renormalization of r_0 is not enough in $d > 2$, and the $d+1$ dimensional integral over momenta and frequency in (18) is still generally confined to the upper cutoff. In this situation, we use additional renormalizations (see below) to restrict the consideration to small energies, at which the effective action describes universal low energy behavior.

For $\phi = 0$, magnetic order emerges when $r = 0$, i.e. when the static susceptibility for the fields Δ_X and Δ_Y, $\chi_0^{-1} = r$, diverges. The relationship between r and $\bar{r}_0$ follows from (18):

$$r = \bar{r}_0 + u\int_q\left[\frac{1}{r+q^2+\gamma|\nu_n|} - \frac{1}{q^2+\gamma|\nu_n|}\right] \tag{19}$$

For classical systems the integrals are infrared divergent for $d \le 2$, meaning that r never reaches zero - this is nothing but the Mermin-Wagner theorem. For $d > 2$ we immediately find from (19) that r and $\bar{r}_0$ vanish simultaneously, i.e. in the absence of a preemptive Ising instability long-range magnetic order appears at $r = \bar{r}_0 = 0$.

Suppose now that a non-zero solution for ϕ appears already at $r > 0$, i.e. when the system is still in the paramagnetic phase. Once ϕ becomes non-zero, the static magnetic susceptibilities for the fields Δ_X and Δ_Y become non-equivalent:

$$\begin{aligned} \chi_X(q=0) &= \frac{1}{r-\phi} \\ \chi_Y(q=0) &= \frac{1}{r+\phi} \end{aligned} \tag{20}$$

Now the magnetic transition occurs when $r = |\phi|$, i.e. at a temperature larger than without Ising-nematic order. In other words, a preemptive Ising-nematic order increases the magnetic transition temperature.

In what follows, we analyze the phase diagrams resulting from the set of non-linear equations (17) in three different regimes: the classical regime, where thermal fluctuations dominate, $T\gamma \gg \bar{r}_0$ (Section III-A); the quantum regime, $T = 0$ (Section III-B); and the regime of intermediate temperature $T\gamma \sim \bar{r}_0$, where thermal and quantum fluctuations are equally important (Section III-C). In all cases, the key parameter that controls the characters of the magnetic and nematic transitions is the dimensionless ratio:

$$\alpha \equiv \frac{u}{g} \tag{21}$$

which measures the strength of the nematic coupling g in units of the magnetic coupling constant u. As it is evident from Eqs. (10) and (11), α depends on the parameters describing the band structure dispersions. It is therefore affected by changes in the chemical potential and in the ellipticity - see Section VI for a systematic analysis of α as function of doping, pressure, and lattice stiffness.

In the classical regime (Section III-A) and in the intermediate regime (Section III-C), the other independent

variable in the phase diagrams (besides α) is the temperature. In the former case, the temperature dependence appears only via the difference from the mean-field Neel transition temperature, $T - T_{N,0}$. In the latter, we consider the explicit dependency on T. On the other hand, in the $T = 0$ quantum regime (Section III-B), the independent variable is the distance to the mean-field critical point at $r_0 = 0$.

Our primary interest is to obtain the phase diagram in the most general case of a quasi-two-dimensional system with anisotropic magnetic interactions. To better understand the results in this case, we first consider a few limiting cases:

- We first analyze the classical $d = 2$ case in Section III-A-1. In this situation, there is no finite temperature magnetic transition, but there is always an intermediary paramagnetic phase with nematic order. At small α (large nematic coupling g), the nematic transition is first-order, whereas at larger α (smaller g), it is second-order. The $d = 2$ case is investigated also in the quantum limit (Section III-B-1) and in the intermediate-temperature regime (Section III-C-1). The behavior in the intermediate regime is the same as in the classical regime. At $T = 0$, we find instead that the magnetic and the nematic transitions are simultaneous and first-order.

- Next we analyze the classical $d = 3$ case with isotropic magnetic interactions in Section III-A-2. We find that, regardless of the value of the dimensionless coupling α, the magnetic and nematic transitions are always simultaneous and first-order. We obtain a similar result in the quantum case (Section III-B-2) and in the intermediary temperature regime (Section III-C-2).

- To model the more general case of anisotropic quasi-2D systems, we first consider in Section III-A-3 the classical system with isotropic dispersion in arbitrary dimensionality $2 < d < 3$. We find that, for large values of α (small g), the system behavior is the same as in the $d = 2$ case, with split second-order magnetic and nematic transitions. On the other hand, for small α (large g), the behavior is similar to the $d = 3$ case, with simultaneous first-order transitions. The two regimes are separated by two tri-critical points and an intermediary regime with a first-order nematic transition split from a lower-temperature second-order magnetic transition. The quantum and intermediary-temperature regimes are analyzed in Sections III-B-3 and III-C-3, respectively.

- In Section III-A-4, we obtain the classical phase diagram of a quasi-two-dimensional system with anisotropic out-of-plane magnetic dispersion, characterized by the parameter η_z. For small η_z (weak anisotropy), the phase diagram is the same as in the previous case of intermediary dimension $2 < d < 3$ and isotropic interactions. For larger η_z (moderate anisotropy), we still obtain a regime of split second-order nematic and magnetic transitions at large α and a regime of simultaneous first-order transitions at small α. However, the two regimes are now separated by an intermediary regime with a higher-temperature second-order nematic transition split from a first-order magnetic transition. In Sections III-B-4 and III-C-4, we present the results for the system behavior in the quantum and intermediary-temperature regimes, respectively.

- Finally, in Section III-A-5, we analyze in more details the intermediary regime of split second-order nematic and first-order magnetic transitions, which appear in the classical phase diagram of the quasi-two-dimensional system with moderate out-of-plane anisotropy. In particular, we show that the first-order magnetic transition is simultaneous to a meta-nematic transition, where the nematic order parameter jumps between two finite values.

A. Classical phase diagram as a function of u/g

When the temperature is high enough ($T\gamma \gg \bar{r}_0$), the dominant contribution to the sum over Matsubara frequencies, $\int_q = T\sum_{\nu_n}\int d^dq/(2\pi)^d$ in Eqs. (17), comes from the term with zero Matsubara frequency. This approximation substantially simplifies Eqs. (17) as the remaining momentum integrals can be evaluated exactly. We begin with the $d = 2$ case.

1. *The case $d = 2$*

For $d = 2$, the integration over momentum in Eqs. (17) yields:

$$\begin{aligned} r &= \bar{r}_0 - \frac{\bar{u}}{4}\log\left(r^2 - \phi^2\right) \\ r &= \phi\coth\left(\frac{2\phi}{\bar{g}}\right) \end{aligned} \tag{22}$$

where we defined the renormalized parameters $\bar{r}_0 = r_0 + \bar{u}\log\Lambda$, $\bar{u} = uT/(2\pi)$, and $\bar{g} = gT/(2\pi)$, with Λ denoting the upper momentum cutoff. The second equation implies that a solution with $\phi \neq 0$ is only possible when $r > 0$. Eliminating r from these equations, we obtain after further rescaling:

$$\phi^*\coth\phi^* + \alpha\log\left(\frac{\phi^*}{\sinh\phi^*}\right) = \bar{\bar{r}}_0 \tag{23}$$

where $\phi^* = 2\phi/\bar{g}$, $\bar{\bar{r}}_0 = 2\bar{r}_0/\bar{g} - \alpha\log(\bar{g}/2)$, and $\alpha = u/g$ (see Eq. 21).

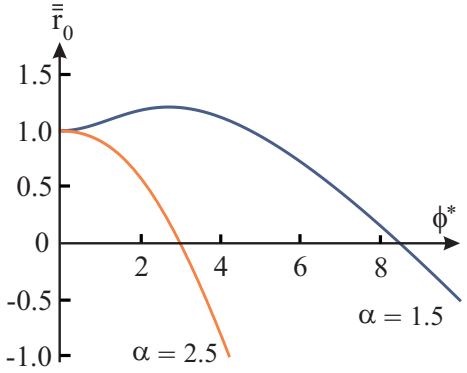

Figure 3: (Color online) $\bar{\bar{r}}_0$ as the function of the Ising-nematic order parameter $\phi^* = 2\phi/\bar{g}$ for two representative values $1 < \alpha < 2$ (blue/dark line) and $\alpha > 2$ (orange/light line). The value of ϕ^* at which the solution first appears upon decreasing $\bar{\bar{r}}_0$ sets the type of the nematic transition. For $\alpha > 2$, the solution first emerges at $\phi^* = 0$, and the transition is second-order. For $1 < \alpha < 2$, $\bar{\bar{r}}_0$ is non-monotonic function of ϕ^* and the solution first emerges at a non-zero ϕ^*. The nematic transition then becomes first-order.

Recall that the original model is constrained to $u > g$, i.e. $\alpha > 1$. The variable $\bar{\bar{r}}_0$ decreases with decreasing T, since it only differs by a constant and by an overall factor from r_0. The leading instability of the system upon decreasing T is into a state with ϕ^* corresponding to the maximum of the left-hand side of Eq. (23). A simple analysis shows that the maximum is at $\phi^* = 0$ for $\alpha \geq 2$ and at a finite ϕ^* for $1 < \alpha < 2$ (see Fig. 3). The implication is that, for $\alpha \geq 2$, the system undergoes a second-order Ising-nematic transition at $\bar{\bar{r}}_0^{cr} = 1$ ($r = \bar{g}/2$), while for $1 < \alpha < 2$ the Ising-nematic transition becomes first-order and the solution for ϕ^* first appears at a larger $\bar{\bar{r}}_0 = \bar{\bar{r}}_0^{max}$ (see Figs. 3 and 4). The value ϕ^*_{cr} at which the left-hand side of Eq. (23) has a maximum gradually increases as α decreases, approaching infinity as $\alpha \to 1$.

The actual $\bar{\bar{r}}_0^{cr}$ at which the first-order Ising-nematic transition occurs is somewhat smaller than $\bar{\bar{r}}_0^{max}$ because at $\bar{\bar{r}}_0^{max}$ the effective action $S_{\rm eff}$ only develops a local minimum at a non-zero ϕ^*_{cr}. The actual transition occurs when the value of the effective action at the local minimum becomes equal to $S_{\rm eff}(\phi^* = 0)$. To obtain $\bar{\bar{r}}_0^{cr}$ we then need to evaluate the effective action at both minima $\phi^* \neq 0$ and $\phi^* = 0$, and find $\bar{\bar{r}}_0^{cr}$ at which the two terms become equal. For better clarity, we compute $S_{\rm eff}$ not only at the extrema (when the self-consistent equations (22) are valid), but for arbitrary ϕ^* at a given $\bar{\bar{r}}_0$. To do this, we solve the equation $\partial S_{\rm eff}/\partial\psi = 0$ to obtain $r\,(\alpha,\, \phi^*, \bar{\bar{r}}_0)$, substitute it back into the effective action and obtain $S_{\rm eff}\,[\alpha,\, \phi^*, \bar{\bar{r}}_0]$. Carrying out the calculations, we obtain, neglecting a constant term,

$$S_{\rm eff}\,[\alpha,\, \phi^*, \bar{\bar{r}}_0] = \frac{\bar{g}}{8}\,\tilde{S}_{\rm eff}\,[\alpha,\, \phi^*, \bar{\bar{r}}_0]\,, \qquad (24)$$

where

$$\tilde{S}_{\rm eff}\,[\alpha,\, \phi^*, \bar{\bar{r}}_0] = (\phi^*)^2 + r^*\left(2 - \log\left[(r^*)^2 - (\phi^*)^2\right]\right)$$
$$-\frac{\alpha}{4}\log^2\left[(r^*)^2 - (\phi^*)^2\right] - \phi^*\log\left(\frac{r^* + \phi^*}{r^* - \phi^*}\right) \qquad (25)$$

and $r^* = r^*\,(\phi^*, \alpha, \bar{\bar{r}}_0)$ is the solution of the equation

$$r^* + \frac{\alpha}{2}\log\left[(r^*)^2 - (\phi^*)^2\right] = \bar{\bar{r}}_0 \qquad (26)$$

In Fig. 4 we plot $\tilde{S}_{\rm eff}\,[\alpha,\, \phi^*, \bar{\bar{r}}_0]$ for $\alpha = 1.5$ as a function of ϕ^* for several values of $\bar{\bar{r}}_0$. We see that for $\bar{\bar{r}}_0 > 1.21$, $S_{\rm eff}$ has a minimum only at $\phi^* = 0$, and $S_{\rm eff}\,[1.5,\, \phi^*, \bar{\bar{r}}_0]$ monotonically increases with increasing ϕ^*. However, once $\bar{\bar{r}}_0$ becomes smaller than $\bar{\bar{r}}_0^{max}$, which for this value of α is $\bar{\bar{r}}_0^{max} = 1.21$, the function $S_{\rm eff}\,[1.5,\, \phi^*, \bar{\bar{r}}_0]$ develops an inflection point at $\phi^*_{cr} \approx 2.72$. At smaller $\bar{\bar{r}}_0$, this inflection point gradually splits into a maximum at $\phi < \phi^*_{cr}$ and a minimum at $\phi > \phi^*_{cr}$. At $\bar{\bar{r}}_0^{cr} = 1.15$ the values of $\tilde{S}_{\rm eff}$ at $\phi^* = 0$ and at the local minimum $\phi^* \neq 0$ become equal, and the system undergoes a first-order Ising-nematic transition. The local minimum of $S_{\rm eff}$ at $\phi^* = 0$ survives down to $\bar{\bar{r}}_0^{min} = 1$. Below this temperature, the effective action has only one minimum at a finite ϕ^*.

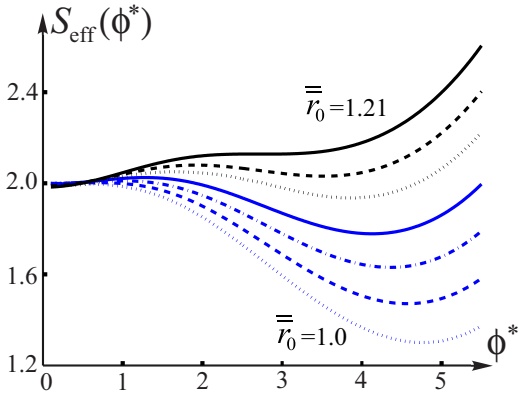

Figure 4: (Color online) The effective action $\tilde{S}_{\rm eff}\,[\alpha,\, \phi^*, \bar{\bar{r}}_0]$ from Eq. 25 as a function of ϕ^* for $\alpha = 1.5$ and several $\bar{\bar{r}}_0$. The evolution of $\tilde{S}_{\rm eff}$ is typical of a first-order phase transition. At large $\bar{\bar{r}}_0$, $\tilde{S}_{\rm eff}\,[1.5,\, \phi^*, \bar{\bar{r}}_0]$ monotonically increases with increasing ϕ^*. At $\bar{\bar{r}}_0^{max} = 1.21$, the effective action develops an inflection point (black/dark solid curve) which at smaller $\bar{\bar{r}}_0$ splits into a maximum and a minimum. At $\bar{\bar{r}}_0^{cr} = 1.15$, the values of $\tilde{S}_{\rm eff}$ at $\phi^* = 0$ and at the local minimum at finite ϕ^* become equal (black dotted curve), and the system undergoes a first-order Ising-nematic transition. The local minimum in $S_{\rm eff}$ at $\phi^* = 0$ survives down to $\bar{\bar{r}}_0^{min} = 1$ (blue/light gray dotted curve).

The jump in ϕ at the first-order transition affects the susceptibilities associated with the two magnetic order parameters Δ_X and Δ_Y, which become non-equivalent once ϕ becomes finite, see Eq. (20). This implies that the static susceptibility and the magnetic correlation length

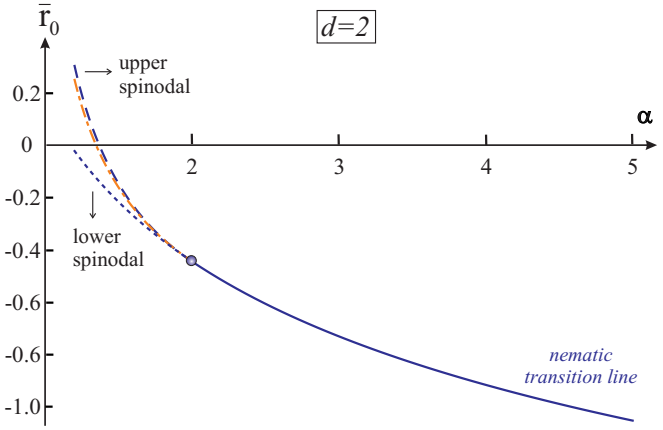

Figure 5: (Color online) Calculated phase diagram in the $(\alpha, \bar{r}_0)$ plane for the $d = 2$ case and high enough temperatures, where classical fluctuations dominate. In the region $\alpha < 2$, the nematic transition is first-order. The dashed blue/dark line represents the upper spinodal (where the ordered state becomes metastable), while the dotted blue/dark line denotes the lower spinodal (where the disordered state ceases to be metastable), and the dashed-dotted orange/light line is the actual transition line where the global minimum of $S_{\rm eff}$ shifts from $\phi = 0$ to $\phi \neq 0$. We recall there is no magnetic order at a finite temperature in $d = 2$. In this figure, we fixed $\bar{u} = 1$ and changed $\bar{g}$.

change discontinuously at the first-order Ising transition, even though there is no magnetic instability (the value to which ϕ jumps is always smaller than r, see the second equation in (22)). Actually, as we already mentioned, magnetic order never emerges for $d = 2$ at a finite temperature, so Z_2 is the only symmetry that gets broken.

To obtain the phase diagram in the $(\alpha, \bar{\bar{r}}_0)$ plane, we need to analyze the behavior of the system once the Ising order sets in. We found that the Ising-nematic order parameter continuously increases with decreasing $\bar{\bar{r}}_0$ for all $\alpha > 1$, implying that there is no other first-order transition line in the phase diagram besides the one at which the Ising-nematic order develops. The phase diagram is presented in Fig. 5. The upper spinodal in this figure corresponds to $\bar{r}_0^{max}$ where a local minimum of $S_{\rm eff}$ appears at $\phi^* \neq 0$, and the lower spinodal refers to $\bar{r}_0^{min}$ where $\phi^* = 0$ ceases to be a local minimum of $S_{\rm eff}$. A first-order transition happens in between the upper and lower spinodal lines.

2. *The case $d = 3$*

A very different phase diagram emerges in three dimensions. The momentum integration in Eqs. (17) now yields

$$r = \bar{r}_0 - \frac{\bar{u}}{4}\left(\sqrt{r+\phi} + \sqrt{r-\phi}\right)$$

$$\phi = \frac{\bar{g}}{4}\left(\sqrt{r+\phi} - \sqrt{r-\phi}\right) \quad (27)$$

where $\bar{r}_0 = r_0 - 2\Lambda\bar{u}/\pi$, and, as before, $\bar{u} = uT/(2\pi)$, and $\bar{g} = gT/(2\pi)$. It is convenient to re-express these equations in terms of r and $z \equiv \phi/r$ $(0 \leq z \leq 1)$. Eliminating r from (27) we then obtain:

$$\bar{r}_0 = \frac{\bar{g}^2}{8}\left(\alpha + \frac{1}{1+\sqrt{1-z^2}}\right) \quad (28)$$

The right-hand side of (28) is an increasing function of z for all values of α. Therefore, the first instability of the system is at $\bar{r}_0 = \bar{g}^2(1+\alpha)/8$ towards the state with the largest possible value of z, namely, $z = 1$. As a result, the order parameter ϕ jumps at the Ising-nematic transition to $\phi = r$. For such ϕ, the static susceptibility associated with one of the magnetic order parameters diverges, see Eqs. (20), implying that the first-order Ising-nematic transition instantaneously brings the system to the verge of magnetic order. The divergence of the susceptibility reflects the fact that the Goldstone modes of the magnetically ordered state are gapless. In the large-N limit, the impact of longitudinal fluctuations, where the susceptibility remains finite, is negligible.

In this situation, we have to extend our analysis and investigate the possibility that Δ_X jumps to a non-zero value at the Ising transition. To do this, we go back to the effective action (12) written in terms of double integrals over both $\Delta_{X,Y}$ and their collective variables ϕ and ψ, and allow one of components of Δ_X to have a non-zero mean value $M = \langle \Delta_X^i \rangle$. Expressing $\Delta_{X,\mathbf{q}}^j = M\delta_{\mathbf{q},0}\delta^{i,j} + \tilde{\Delta}_{X,\mathbf{q}}^j$ and integrating over $\tilde{\mathbf{\Delta}}_X$ and $\mathbf{\Delta}_Y$ we obtain:

$$S_{\rm eff}[\phi,\psi,M] = S_{\rm eff}[\phi,\psi] + M^2(r-\phi)\left[1 + (r-\phi)^2\int_q \frac{1}{q^2(r-\phi+q^2)}\right] \quad (29)$$

In the spirit of the $N \to \infty$ approximation, we rescale $M^2 \to M^2 N$ and differentiate $S_{\rm eff}[\phi,\psi,M]$ over all three variables. This yields the set of three self-consistent equations for ϕ, r, and M. The equation for M is:

$$M(r-\phi)\left(1 + (r-\phi)^2\int_q \frac{1}{q^2(r-\phi+q^2)}\right) = 0. \quad (30)$$

The solution is either $M = 0$ or $r = \phi$. We take $M \neq 0$ and $r = \phi$. The other two equations for $r = \phi$ are

$$\phi = \bar{r}_0 - \frac{\bar{u}}{4}\sqrt{2\phi} + \bar{u}\bar{M}^2,$$
$$\phi = \frac{\bar{g}}{4}\sqrt{2\phi} + \bar{g}\bar{M}^2, \quad (31)$$

where $\bar{M} = M\sqrt{2\pi/T}$. Solving the first equation for ϕ we obtain:

$$\sqrt{\phi} = \frac{\bar{g}}{4\sqrt{2}}\left(1+\sqrt{1+\frac{32\bar{M}^2}{\bar{g}}}\right). \quad (32)$$

Substituting this solution into the second equation yields $\bar{r}_0$ as a function of $\bar{M}$:

$$\bar{r}_0 = \frac{\bar{g}^2}{16}\left[-16\frac{\bar{M}^2}{\bar{g}}(\alpha-1)+ (1+\alpha)\left(1+\sqrt{1+\frac{32\bar{M}^2}{\bar{g}}}\right)\right] \quad (33)$$

Following the same strategy as before, we look for the value of $\bar{M}$ correspondent to the largest $\bar{r}_0$, i.e. we determine the value of $\bar{M}$ that emerges at the largest temperature. A straightforward calculations shows that it is non-zero and equals to

$$\bar{M} = \frac{\sqrt{2\bar{g}\alpha}}{4(\alpha-1)}. \quad (34)$$

This implies that the first-order Ising-nematic transition at which ϕ reaches its maximum value ($= r$) triggers a first-order magnetic transition into a state with a finite $M = \langle \Delta_X^i \rangle$. We emphasize, however, that in this and similar cases which we consider below, the magnitude of the nematic order parameter ϕ is larger than $g\bar{M}^2$, i.e. the magnetic transition is secondary to the nematic transition. Indeed, substituting $\bar{M}$ from (34) into (32) we obtain

$$\phi = \alpha\left(\bar{g}\bar{M}^2\right) > \bar{g}\bar{M}^2. \quad (35)$$

We emphasize that, without nematic instability, a "pure" magnetic transition would take place at a smaller temperature, when the condition $r = 0$ is satisfied, instead of $r = \phi$. Therefore, even though both transitions are simultaneous and first-order, the nematic one is the primary transition, and the magnetic transition is induced by the feedback from the nematic order.

Note also that in the mean-field theory the susceptibility associated with the field $\mathbf{\Delta}_X$ remains massless ($\chi_X^{-1} = r - \phi = 0$) despite that magnetic order develops and enforces a gap in the spectrum of longitudinal fluctuations. This is a consequence of taking the $N \to \infty$ limit, in which one longitudinal mode is negligible compared to $N-1$ gapless transverse modes.

Like in the $d = 2$ case, the value of $\bar{r}_0$, at which the effective action develops an inflection point and the solution with $\phi = r$ and $M \neq 0$ first appears, is the upper spinodal $\bar{r}_0^{max}$. The actual first-order transition occurs at a smaller $\bar{r}_0^{cr}$, at which the values of S_{eff} at $\phi = M = 0$ and at the local minimum at $\phi = r$, $M \neq 0$ become equal (for $r = \phi$ the local minimum is with respect to variations of M).

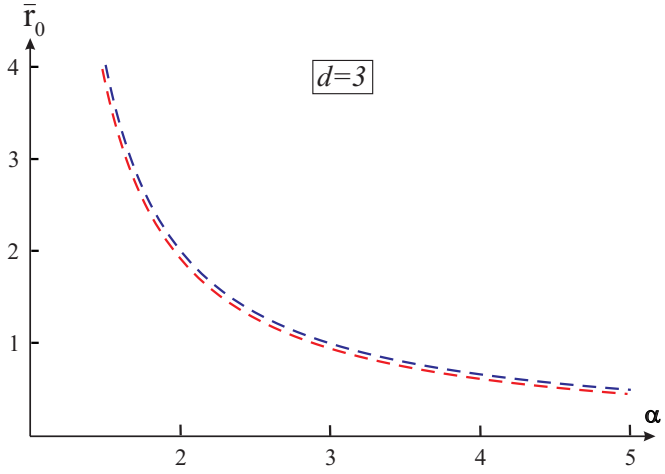

Figure 6: (Color online) Classical phase diagram in the $(\alpha, \bar{r}_0)$ plane for the $d = 3$ case. For any value of α, the nematic and magnetic transition are simultaneous and first-order.

The phase diagram for $d = 3$ is presented in Fig. 6. As in the $d = 2$ case, there are no additional transition lines, i.e. after jumping to finite values at the first-order transition, both ϕ and M continuously increase with decreasing $\bar{r}_0$. The similarities between the $d = 2$ and $d = 3$ cases, however, stop here because the phases below the critical line are very different. In $d = 3$, Ising-nematic order immediately triggers a magnetic order, such that there is no regime where the Z_2 symmetry is broken but the $O(3)$ symmetry is unbroken. In $d = 2$, there is no magnetic order below the Ising-nematic transition, which can be either first-order or second-order, depending on α.

We now consider what happens in anisotropic three dimensional systems. There are several ways to model the anisotropy: one can either consider the dimension d to be an arbitrary number between 2 and 3 or one can keep $d = 3$ but consider an anisotropic magnetic dispersion with different stiffness along q_z and along q_x, q_y. It turns out that the system behavior is universal at small and large α, but at intermediate α it depends on the choice of the model. This will lead us to the two phase diagrams shown in Fig. 1. We first consider arbitrary $2 < d < 3$ and then an anisotropic $d = 3$ dispersion.

3. *The case* $2 < d < 3$.

We assume first that $M = 0$ and later extend the formalism to include a non-zero magnetization. For $2 < d < 3$, the integration over momentum in Eqs. (17) yields

$$r = \bar{r}_0 - \frac{\bar{u}}{4}\left[(r+\phi)^{\frac{d-2}{2}} + (r-\phi)^{\frac{d-2}{2}}\right]$$
$$\phi = \frac{\bar{g}}{4}\left[(r+\phi)^{\frac{d-2}{2}} - (r-\phi)^{\frac{d-2}{2}}\right], \quad (36)$$

where:

$$\frac{\bar{u}}{u} = \frac{\bar{g}}{g} = \frac{2TS_d}{(2\pi)^d}\int_0^\infty dx\,\frac{x^{d-3}}{1+x^2} = \frac{T}{2^{d-1}\pi^{\frac{d-2}{2}}\Gamma\left(\frac{d}{2}\right)\sin\left(\frac{(d-2)\pi}{2}\right)} \quad (37)$$

and $S_d = 2\pi^{d/2}/\Gamma(d/2)$ is the area of a d-dimensional sphere with unit radius. Introducing, as before, $z \equiv \phi/r$ $(0 \le z \le 1)$, we solve the second equation for $r\,(z)$, substitute the solution into the first equation, and obtain:

$$\bar{r}_0 = \left(\frac{\bar{g}}{4}\right)^{\frac{2}{4-d}} Q\,(z,\alpha)\,, \quad (38)$$

with:

$$Q\,(z,\alpha) = \left[\frac{(1+z)^{\frac{d-2}{2}} - (1-z)^{\frac{d-2}{2}}}{z}\right]^{\frac{d-2}{4-d}} \times \left[(1+z)^{\frac{d-2}{2}}\left(\alpha + \frac{1}{z}\right) + (1-z)^{\frac{d-2}{2}}\left(\alpha - \frac{1}{z}\right)\right] \quad (39)$$

We determine the leading instability of the system by again looking at the location of the maximum of $Q\,(z,\alpha)$ at a given α. In Fig. 7, we plot $Q(z,\alpha)$ for $d = 2.5$ and three different values of α. We find three different regimes for the behavior of the system: For $1 < \alpha < \alpha_{c1} = 2$, the first instability is at $z = 1$; for $\alpha_{c1} < \alpha < \alpha_{c2} = 3.5$, the first instability is at an intermediate value $0 < z < 1$; for $\alpha > \alpha_{c2}$, the first instability is at $z = 0$. Expanding $Q\,(z,\alpha)$ near $z = 0$ and $z = 1$ and analyzing the sign of the slope, we find the expressions for α_{c1} and α_{c2} in an arbitrary dimension:

$$\alpha_{c1} = \frac{1}{3-d}\,,\ \alpha_{c2} = \frac{6-d}{6-2d}. \quad (40)$$

For $d = 2$ we have $\alpha_{c1} = 1$ and $\alpha_{c2} = 2$, i.e. the regime $1 < \alpha < \alpha_{c1}$ disappears, in agreement with Fig. 5. On the other hand, for $d = 3$, both α_1 and α_2 tend to infinity, and the region $1 < \alpha < \alpha_{c1}$ extends to all values of α, in agreement with Fig. 6. For d between 2 and 3 all three regions are present, as shown in the phase diagram of Fig. 8.

In the first region $1 < \alpha < \alpha_{c1}$ (region I in Fig. 8), the behavior of the system is the same as we found in $d = 3$: ϕ jumps at the Ising-nematic transition to the largest possible value $\phi = r$, triggering a simultaneous magnetic transition. To determine whether the latter is also first order, we extend the analysis of the effective action in the same way as we did for $d = 3$: we introduce magnetic long-range order via the order parameter $M = \langle \Delta_X^i \rangle$ and obtain the set of three coupled equations for M, ϕ, and ψ. The equation for M again gives either $M = 0$ or $\phi = r$.

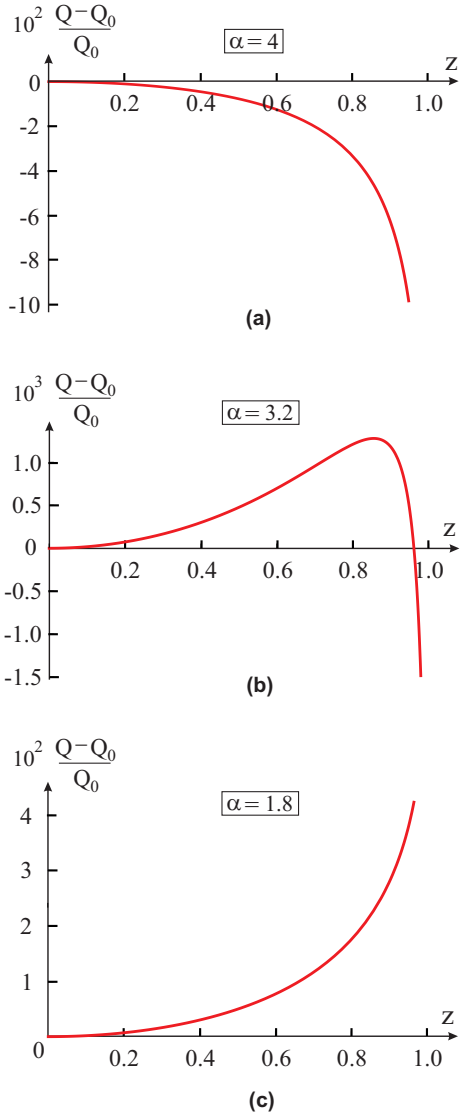

Figure 7: (Color online) $Q\,(z,\alpha) \propto \bar{r}_0$, as defined in Eqs. (38) and (39), plotted as function of $z \equiv \phi/r$ for three representative values of α at $d = 2.5$: (a) $\alpha > \alpha_{c2} = 3.5$, (b) $2 = \alpha_{c1} < \alpha < \alpha_{c2}$, and (c) $\alpha < \alpha_{c1}$. The maximum of Q shifts from $z = 0$ at large α to intermediate z at intermediate α, and to $z = 1$ at small α.

We choose $r = \phi$ and re-express the other two equations as

$$\phi = \bar{r}_0 - \frac{\bar{u}}{4}(2\phi)^{\frac{d-2}{2}} + \bar{u}\bar{M}^2$$
$$\phi = \frac{\bar{g}}{4}(2\phi)^{\frac{d-2}{2}} + \bar{g}\bar{M}^2 \quad (41)$$

with $\bar{M}/M = \sqrt{u/\bar{u}}$ (see Eq. 37). For $\alpha \le \alpha_{c1}$, we expand in $\bar{M}$ and obtain an explicit equation relating $\bar{r}_0$ to the magnetization $\bar{M}$:

$$\bar{r}_0 = \left(\frac{\bar{g}}{2}\right)^{\frac{2}{4-d}}\left(\frac{1+\alpha}{2}\right) + \frac{\bar{g}}{2(4-d)}\bar{Q}\,(\bar{M},\alpha)\,, \quad (42)$$

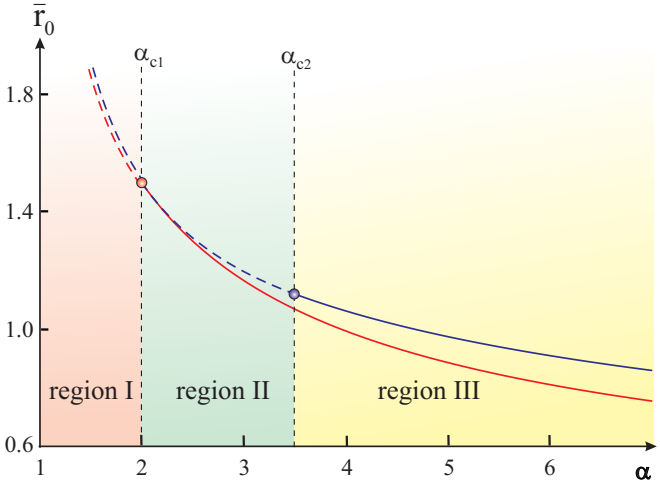

Figure 8: (Color online) The calculated phase diagram in the $(\alpha, \bar{r}_0)$ plane for the representative case $d = 2.5$. In region I, there is a simultaneous nematic and magnetic first-order transition. In region II, the two transitions split, but the nematic transition remains first-order, while the magnetic transition becomes second-order. In region III, the nematic transition also becomes second-order. The three regions are separated by two tri-critical points. The temperature behavior of the nematic and magnetic order parameter in each region is shown in Fig. 9 below.

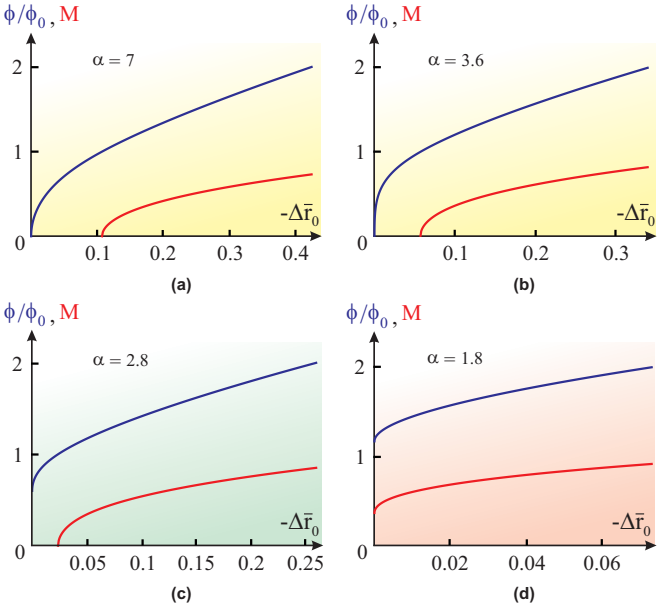

Figure 9: (Color online) Nematic order parameter ϕ (blue/dark gray curves) and magnetic order parameter M (red/light gray curves) as function of $-\Delta\bar{r}_0 = r_{0,cr} - r_0 \propto (T_s - T)$ for four different values of α at $d = 2.5$. In (a) and (b) we use $\alpha > \alpha_{c2} = 3.5$ (region III of the phase diagram), in (c) we use $\alpha_{c2} > \alpha > \alpha_{c1} = 2$ (region II of the phase diagram), and in (d) we use $\alpha < \alpha_{c1}$ (region I of the phase diagram). ϕ_0 is the value of ϕ corresponding to $\phi = r$.

with:

$$\bar{Q}\left(\bar{M},\alpha\right) \approx 4\left(3-d\right)\left(\alpha_{c1}-\alpha\right)\bar{M}^2 - 8\left(\frac{\bar{g}}{2}\right)^{\frac{2-d}{4-d}}\left(1+\alpha\right)\left(\frac{d-2}{4-d}\right)\bar{M}^4. \quad (43)$$

For $\alpha \lesssim \alpha_{c_1}$ the maximum of $\bar{Q}\left(\bar{M},\alpha\right)$ is at a finite magnetization:

$$\bar{M} = \frac{1}{2}\left[\frac{\left(4-d\right)\left(3-d\right)\left(2\bar{g}\right)^{\frac{d-2}{4-d}}}{\left(1+\alpha\right)\left(d-2\right)}\right]^{1/2}\sqrt{\alpha_{c1}-\alpha}. \quad (44)$$

Therefore, in region I, the first-order Ising transition triggers the first-order magnetic transition. We again emphasize, however, that the Ising-nematic order parameter is *larger* than the square of the magnetic order parameter, i.e. the magnetic transition is the secondary transition, triggered by the preemptive Ising transition. This is most clearly seen from Eq. (44): the jump in the magnetization approaches zero as α approaches α_{c1}, while the jump in the Ising-nematic order parameter remains finite and reaches $\phi = (1/2)(\bar{g}/2)^{2/(4-d)}$. Thus, α_{c1} corresponds to a *magnetic tri-critical point,* while a preemptive first-order Ising transition exists on both sides of it. In Fig. 9 d, we present numerical results for ϕ and M as functions of $\bar{r}_0$ in region I.

In the second region $\alpha_{c1} < \alpha < \alpha_{c2}$ (region II in Fig. 8), the Ising-nematic transition is still first-order, but the magnitude of the jump of ϕ is smaller than the value required to trigger a magnetic transition, i.e. $\phi < r$. As a result, in this region the magnetic and Ising transitions are split, with the former occurring at a smaller $\bar{r}_0$ (i.e. at a smaller temperature) than the latter. The magnetic transition then becomes second-order, as the maximum of $\bar{Q}\left(M,\alpha\right)$ remains at $M = 0$ for $\alpha \gtrsim \alpha_{c1}$, see Eq. (43). In Fig. 9 c, we present the numerical solution for ϕ and M in this region, as function of temperature.

To determine what happens as the system approaches α_{c2}, we expand $Q\left(z,\alpha\right)$ in Eq. (39) around $z = 0$ and obtain:

$$\frac{Q(z,\alpha)}{(d-2)^{\frac{2}{4-d}}} \approx \left(\frac{2\alpha}{d-2}+1\right) + \frac{(3-d)}{6}\times \left[\left(\alpha_{c2}-\alpha\right)z^2 + \frac{(15+2d)\left(6-d\right)}{120}\left(\kappa\alpha_{c2}-\alpha\right)z^4\right] \quad (45)$$

with a constant $\kappa \equiv \frac{90-7d-d^2}{(6-d)(15+2d)} < 1$. Clearly, the jump of the Ising-nematic order parameter across the transition approaches zero as α increases towards α_{c2}, $z_{max} \sim \sqrt{\alpha_{c2}-\alpha}$. Therefore, α_{c2} is a *nematic tri-critical point,* beyond which the Ising-nematic transition becomes second-order (region III in Fig. 8). In this region, a non-zero ϕ gradually develops below the transition line, and the magnetic transition splits even further from the Ising-nematic transition. We present the numerical solution for ϕ and M in region III in Fig. 9 b,a.

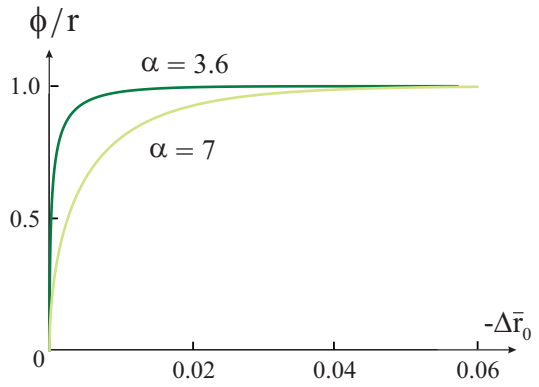

Figure 10: (Color online) Nematic order parameter ϕ, in units of the renormalized magnetic mass term r, as function of the reduced temperature $-\Delta\bar{r}_0 = r_{0,cr} - r_0 \propto (T_s - T)$ at $d = 2.5$ for two different values of α : $\alpha = 7$, which is far from the nematic tri-critical point $\alpha_{c2} = 3.5$ (light-green/ light gray curve), and $\alpha = 3.6$, which is very close to the nematic tri-critical point (dark-green/dark gray curve). In both cases, ϕ evolves continuously from zero, but the slope is very large for α only slightly above α_{c2}.

Note that the slope with which ϕ increases below the instability remains very high over some range of α in region III, as illustrated in Fig. 9(b) and, more transparently, in Fig. 10. From the practical point of view, the Ising transition can then still be viewed as almost first-order. This can also be seen from Eq. 45 because in coefficient of the z^4 term, κ remains very close to one for all $2 < d < 3$ ($0.048 < 1 - \kappa < 0.054$). As a result, if α is not too far from α_{c2}, ϕ increases very rapidly. For instance, the value $\bar{r}_0/\bar{r}_0^{max}$ for which ϕ increases to $\phi = r/2$ is $|\bar{r}_0/\bar{r}_0^{max} - 1| \lesssim 10^{-3}\,(\alpha - \alpha_{c2})$, almost independent on the dimensionality.

We emphasize once again that for $\alpha < \alpha_{c2}$, when the nematic transition is first order, the value of $\bar{r}_0$ at which the solution of Eq. (38) first appears is the upper spinodal of the system, $\bar{r}_0^{max}$. The lower spinodal $\bar{r}_0^{min}$ is the value of $\bar{r}_0$ below which $\phi = 0$ is no longer the local minimum of $S_{\rm eff}$. The actual phase transition in the presence of fluctuations is between $\bar{r}_0^{max}$ and $\bar{r}_0^{min}$, and is close to $\bar{r}_0^{min}$ if fluctuations are weak. If we use $\bar{r}_0^{min}$ instead of $\bar{r}_0^{max}$ for the value of $\bar{r}_0^{cr}$ at which the first-order transition occurs, we find that α_{c2} remains intact, but α_{c1} moves to a larger value α'_{c1} given by:

$$\alpha'_{c1} = \frac{\left(\frac{2}{d-2}\right)^{\frac{d-2}{4-d}} - (d-2)}{2 - \left(\frac{2}{d-2}\right)^{\frac{d-2}{4-d}}} \qquad (46)$$

One can easily check that for $2 < d < 3$, $\alpha_{c1} < \alpha'_{c1} < \alpha_{c2}$. Thus, the magnetic tri-critical point does not merge with the nematic tri-critical point, even if the Ising-nematic transition occurs at the lowest possible $\bar{r}_0^{min}$.

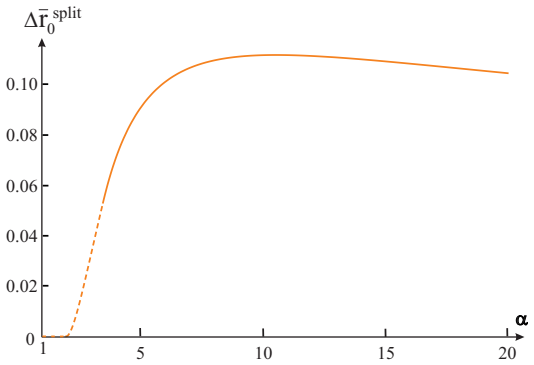

Figure 11: (Color online) Splitting between the nematic and magnetic transitions, $\Delta\bar{r}_0^{\rm split} \equiv \bar{r}_0^{\rm nem} - \bar{r}_0^{\rm mag}$, as function of $\alpha \equiv u/g$, for the case $d = 2.5$. The dashed line refers to regions I and II of the phase diagram of Fig. 8, whereas the solid line refers to region III.

Notice that in the phase diagram of Fig. 8 the splitting between the magnetic and nematic transitions initially increases with increasing α, once the system crosses the nematic tri-critical point. However, for small values of the nematic coupling g (large values of α), this splitting is expected to decrease, since it should tend to zero for infinitesimal g. This non-monotonic behavior is a consequence of the fact that, at smaller α, the strong first-order nematic transition brings in a simultaneous magnetic transition, despite the fact that the magnetic correlation length is rather small immediately above the nematic transition. This correlation length, meanwhile, increases monotonically as α increases. We show the non-monotonic behavior of the splitting between the magnetic and nematic transitions as function of α in Fig. 11 for the case $d = 2.5$. Note that the splitting begins to slowly decrease only at $\alpha \geq 10$.

4. *Anisotropic $d = 3$ case*

We now show that the phase diagram obtained for the case $2 < d < 3$ is qualitatively the same as the phase diagram of the $d = 3$ model with strongly anisotropic, quasi-2D magnetic dispersion. However, for less anisotropic systems, we find a different phase diagram, with new behavior at intermediate α.

To model the anisotropy, we consider the system composed of stacked two-dimensional layers. In each layer n, we define the magnetic order parameters $\mathbf{\Delta}_i^n$, with $i = X, Y$. The coupling between different layers adds the term $-\frac{\eta_z}{2}\mathbf{\Delta}_i^n \cdot \mathbf{\Delta}_i^{n+1}$ to the action with $\eta_z < 1$. As a result, the magnetic susceptibility in Eq. (9) becomes:

$$\chi_{i,q}^{-1} = r_0 + q_\parallel^2 + \eta_z \sin^2 \frac{q_z}{2} \qquad (47)$$

with $0 < q_z < 2\pi$ and $\mathbf{q}_\parallel = (q_x, q_y)$. An alternative possibility is to consider anisotropic but still quadratic

dispersion

$$\chi_{i,q}^{-1} = r_0 + q_\parallel^2 + \beta^2 q_z^2 \tag{48}$$

with $0 < \beta < 1$, and set the same cutoff Λ for all three components of $\mathbf{q}$. We found that the phase diagram is *the same* regardless of whether we use Eq. (47) or (48) for the bosonic susceptibility. For definiteness, here we present the results for χ_q given by Eq. (47) and consider the case of Eq. (48) in Appendix C.

We plug Eq. (47) into the self-consistent equations (17) and evaluate the three-dimensional integrals. Absorbing the cutoff Λ into r_0, $\bar{r}_0 = r_0 + 2\bar{u}\log(2\Lambda)$, we obtain in the paramagnetic phase:

$$\begin{aligned} r &= \bar{r}_0 - \bar{u}\log\left(\sqrt{r+\phi} + \sqrt{\eta_z + r + \phi}\right) \\ &\quad -\bar{u}\log\left(\sqrt{r-\phi} + \sqrt{\eta_z + r - \phi}\right) \\ \phi &= \bar{g}\log\left(\frac{\sqrt{r+\phi} + \sqrt{\eta_z + r + \phi}}{\sqrt{r-\phi} + \sqrt{\eta_z + r - \phi}}\right) \end{aligned} \tag{49}$$

where, as before, $\bar{u} = uT/(2\pi)$ and $\bar{g} = gT/(2\pi)$. >From the second equation, we obtain r as function of ϕ:

$$r(\phi) = -\frac{\eta_z}{2} + \cosh\left(\frac{\phi}{g}\right)\sqrt{\frac{\eta_z^2}{4} + \frac{\phi^2}{\sinh^2\left(\frac{\phi}{g}\right)}} \tag{50}$$

Substituting (50) into the first equation in (49) we obtain $\bar{r}_0(\phi)$, whose maximum determines the first instability of the system. At a non-zero $\bar{M}$ we find, as before, $r = \phi$ and

$$\begin{aligned} &\bar{r}_0(\phi) = \\ &\phi + \bar{u}\left[\log\left(\sqrt{2\phi\eta_z} + \sqrt{\eta_z^2 + 2\phi\eta_z}\right) - \bar{M}^2\right] \end{aligned} \tag{51}$$

where the magnetization is given by:

$$\bar{M}^2 = \frac{\phi}{g} - \log\left(\frac{\sqrt{2\phi} + \sqrt{\eta_z + 2\phi}}{\sqrt{\eta_z}}\right) \tag{52}$$

Notice the characteristic logarithmic dependence on the anisotropy, $\log\eta_z$. Using these expressions, we obtain the phase diagram of the anisotropic system by analyzing the first instability upon decreasing $\bar{r}_0$.

We show our results for various η_z in Fig.12, where we present four possible phase diagrams together with their respective $\bar{r}_0(z,\alpha)$ profiles for different values of α (recall that $z = \phi/r$). The phase diagrams shown in panels (a) and (d) exist over sizable ranges of small and moderate η_z, respectively. On the other hand, the phase diagrams in panels (b) and (c) exist only over rather narrow ranges of η_z and just show how the system actually evolves from the regime (a) into the regime (d).

The behavior at small η_z shown in Fig.12(a) is the same as for a fractional dimension $2 < d < 3$, Fig. 8. Namely,

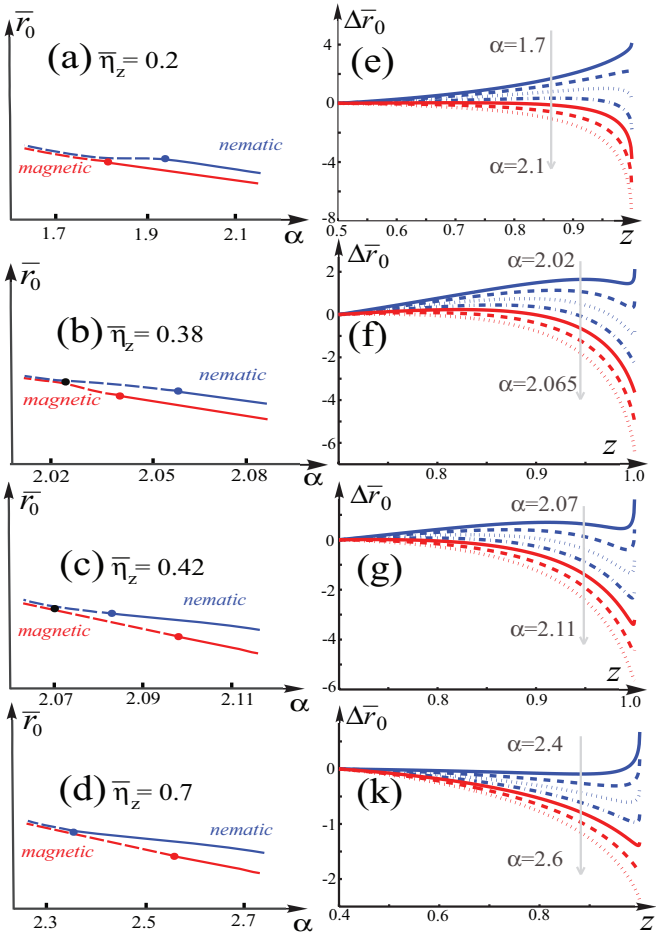

Figure 12: (Color online) Left panels: the four different phase diagrams for the anisotropic 3D model with the magnetic susceptibility given by Eq. (47). The behavior found in isotropic systems for a fractional $2 < d < 3$ only holds at small $\bar{\eta}_z \equiv \eta_z/\bar{g}$. At larger $\bar{\eta}_z$, the magnetic and nematic tri-critical points α_{c1} and α_{c2} interchange and a new behavior emerges at intermediate α. The behavior presented in panels (a) and (d) takes place over sizable ranges of small and moderate $\bar{\eta}_z$, respectively. The behavior presented in panels (b) and (c) takes place only over a narrow range of $\bar{\eta}_z$, around the $\bar{\eta}_z$ value for which α_{c1} and α_{c2} cross. Right panels: the behavior of $\bar{r}_0(z)$ for several different α for each phase diagram ($z = \phi/r$). The key difference between the regimes (a) and (d) is that in (a) $\bar{r}_0(z)$ has only one maximum at every α, while in (d) $\bar{r}_0(z)$ has two distinct maxima and a minimum in between.

there is a magnetic tri-critical point at α_{c1} and a nematic tri-critical point at $\alpha_{c2} > \alpha_{c1}$, with the regime of split first-order nematic and second-order magnetic transition in between. The behavior at larger η_z is, however, different. We see from Fig.12(d) that, in the regime of intermediate α, there is a second-order nematic followed by a first-order magnetic transition.

This change in the system behavior upon increasing η_z can be better understood by considering how the ratio

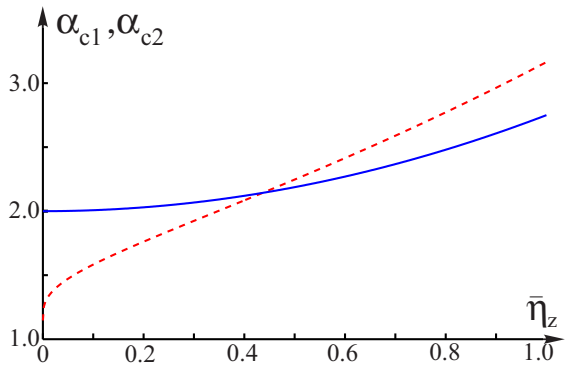

Figure 13: (Color online) Magnetic and nematic tri-critical points α_{c1} (red/light gray dotted line) and α_{c2} (blue/dark gray solid line) as functions of $\bar{\eta}_z$ for the 3D anisotropic model with the magnetic susceptibility given by Eq. (47). Notice that α_{c1} and α_{c2} cross at $\bar{\eta}_z \sim 0.43$.

of the two tri-critical points α_{c1}/α_{c2} evolves with η_z. In 2D, when $\eta_z = 0$, we have $\alpha_{c2}/\alpha_{c1} = 2$. For the isotropic system in $2 < d < 3$, we have from (40) that $\alpha_{c2}/\alpha_{c1} = 3 - d/2$, implying that α_{c2} remains larger than α_{c1} for all $2 < d < 3$, what gives rise to the phase diagram of Fig. 8. However, there is no requirement that α_{c2} must remain larger than α_{c1} as the system approaches the isotropic 3D regime - in fact, the only requirement is that in this limiting case both α_{c1} and α_{c2} diverge. In Fig. 13 we plot α_{c1} and α_{c2} as a function of η_z for the 3D system with anisotropic dispersion. We see that α_{c1} and α_{c2} cross at a certain η_z, beyond which the nematic tri-critical point occurs at a smaller α than the magnetic tri-critical point. This leads to the phase diagram of Fig. 12(d).

We can show analytically that α_{c1} and α_{c2} cross upon increasing η_z. To obtain α_{c2}, we substitute $r(\phi)$ from (50) in (49) and expand the right-hand side in powers of ϕ. Since the quadratic ϕ^2 term vanishes at α_{c2}, we obtain:

$$\alpha_{c2} = 2 + \frac{3\bar{\eta}_z^2}{4} \tag{53}$$

where $\bar{\eta}_z \equiv \eta_z/\bar{g}$. To obtain α_{c1}, we express ϕ in terms of $\bar{m}$ using Eq. (52), substitute the result in Eq. (51) and expand the right-hand side in powers of $\bar{M}$. Since the quadratic $\bar{M}^2$ term vanishes at α_{c1}, we find:

$$\alpha_{c1} = \left(1 - \frac{2}{\sqrt{2\bar{\phi}_{c1}}\sqrt{\bar{\eta}_z + 2\bar{\phi}_{c1}}}\right)^{-1} \tag{54}$$

where $\bar{\phi}_{c1}$ is the solution of

$$\bar{\phi}_{c1} = \log\left(\frac{\sqrt{2\bar{\phi}_{c1}} + \sqrt{\bar{\eta}_z + 2\bar{\phi}_{c1}}}{\sqrt{\bar{\eta}_z}}\right) \tag{55}$$

A simple analysis then shows that α_{c2} becomes larger than α_{c1} at $\bar{\eta}_z^c = 0.43$. One has to be careful to properly determine α_{c2} for $\bar{\eta}_z > \bar{\eta}_z^c$, since the actual nematic tri-critical point takes place when $z = 0$ is the global maximum of $\bar{r}_0(z)$, and not only a local maximum. This subtlety does not affect the result that α_{c1} and α_{c2} cross, nor the regime $\bar{\eta}_z < \bar{\eta}_z^c$. In the phase diagram shown in Fig. 14, α_{c2} is the actual nematic tri-critical point.

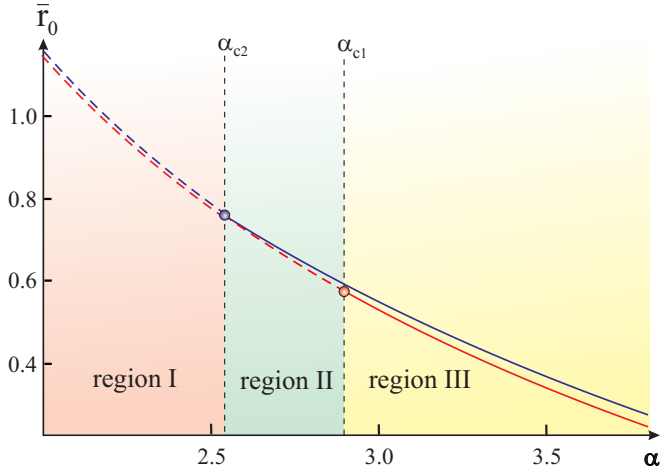

Figure 14: (Color online) Calculated phase diagram in the $(\alpha, \bar{r}_0)$ plane for moderately anisotropic $d = 3$ system with $\eta_z = 0.3\bar{u}$. Three different regions of system behavior are marked in the same way as in Fig. 8. Notice that the positions of the tri-critical points are reversed with respect to the case of strongly anisotropic systems.

The phase diagram shown in Fig. 12(d) contains a new type of a phase transition not seen in quasi-two dimensional systems, namely, a meta-nematic transition. We explore it in more detail in the next subsection.

5. *Meta-nematic transition*

In Fig. 14 we explicitly show the three different regimes that appear in Fig. 12(d), as we did in Fig. 8. The behavior of the nematic and the magnetic order parameters as functions of the distance to the transitions in all three regimes is shown in Fig. 15. For these particular figures, for convenience, we considered $\eta_z = 0.3\bar{u}$ instead of $\eta_z = 0.7\bar{g}$, which does not change the properties of the phase diagram.

A careful analysis of the profile of $\bar{r}_0(z, \alpha)$ as a function of z for the phase diagram in Fig. 12 (d) shows that, at small α, the maximum of $\bar{r}_0(z, \alpha)$ is at $z = 1$ (i.e. at $\phi = r$). This implies that upon decreasing $\bar{r}_0$ the system undergoes a first-order nematic transition, which triggers a simultaneous first-order magnetic transition. This system behavior takes place in region I of Fig. 14, and is displayed in Fig. 15 (c).

However, as α increases, $\bar{r}_0(z, \alpha)$ develops a local maximum at $z = 0$ as well. At the nematic tri-critical point α_{c2}, $\bar{r}_0(0, \alpha)$ becomes equal to $\bar{r}_0(1, \alpha)$, and $\bar{r}_0(z, \alpha)$ has a minimum for $0 < z < 1$. Once α becomes larger than α_{c2}, the absolute maximum of $\bar{r}_0(z, \alpha)$ shifts to $z = 0$,

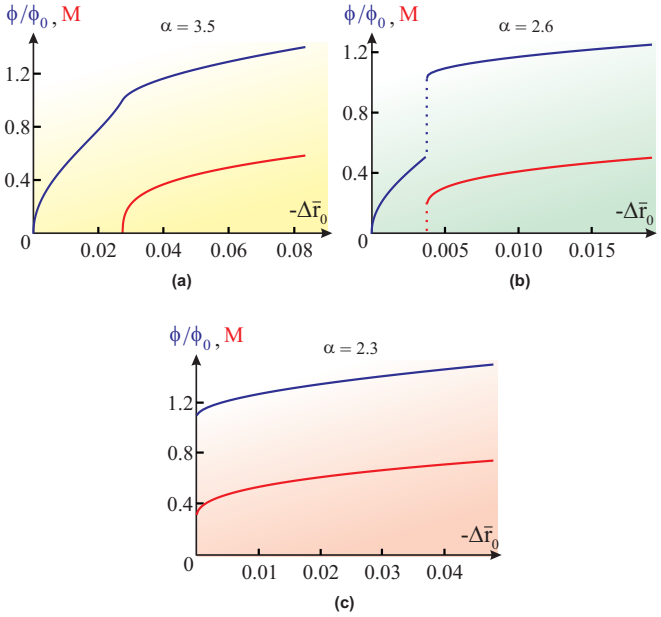

Figure 15: (Color online) Nematic order parameter ϕ (blue/dark gray curves) and magnetic order parameter m (red/light gray curves) as function of $-\Delta\bar{r}_0 = r_{0,cr} - r_0 \propto (T_s - T)$ for three different values of α in the case of a moderately anisotropic system with $\eta_z = 0.3\bar{u}$ (see Fig. 14). Panel (a) is for $\alpha > \alpha_{c1} \approx 2.91$ (region III of the phase diagram), panel (b) for $\alpha_{c2} \approx 2.54 < \alpha < \alpha_{c1}$ (region II of the phase diagram), and panel (c) is for $\alpha < \alpha_{c2}$ (region I of the phase diagram). In panel (b), the dotted line marks the meta-nematic transition, coincident with the magnetic transition. ϕ_0 is the value of ϕ corresponding to $\phi = r$.

i.e. the nematic transition becomes second-order. As $\bar{r}_0$ decreases, a solution with small and finite $z \neq 0$ develops. But because the profile of $\bar{r}_0(z, \alpha)$ has two maxima, another solution with $z = 1$ necessarily appears once $\bar{r}_0$ gets smaller than some critical value. We checked that the value of S_{eff} at $z = 1$ eventually becomes larger than the value corresponding to the solution at small z. As a result, at some $\bar{r}_0$, the nematic order parameter undergoes a *meta-nematic transition*, where it jumps from some small value $\phi \neq 0$ to $\phi = r$ ($z = 1$). In accordance to what we found earlier, we explicitly confirmed that such a transition triggers a first-order magnetic transition. Therefore, for $\alpha > \alpha_{c2}$, the second-order nematic and meta-nematic transitions split. The second-order transition occurs first, and the meta-nematic transition occurs at a smaller $\bar{r}_0$, triggering a first-order magnetic transition. This system behavior takes place in region II of the phase diagram of Fig. 14, as shown in Fig. 15 (b).

As α increases further, the magnetic transition remains first-order up to $\alpha = \alpha_{c1} > \alpha_{c2}$. At this point, the position of the maximum of $\bar{r}_0(M, z = 1, \alpha)$ shifts to $M = 0$, i.e the magnetic transition becomes second order. We found both analytically and numerically that at exactly the same $\alpha = \alpha_{c1}$, the local minimum of $\bar{r}_0(z, \alpha)$ moves to $z = 1$, meaning that $\bar{r}_0(z, \alpha)$ becomes a monotonically decreasing function of z for all $0 \leq z \leq 1$. As a result, for $\alpha > \alpha_{c1}$, the nematic order parameter ϕ monotonically increases with decreasing $\bar{r}_0$ until it reaches the value $\phi = r$, where the system undergoes a second-order magnetic transition. This system behavior corresponds to region III in Fig. 14, and is the same as region III in Fig. 8. Notice from Fig. 15 (a), however, that for $\alpha \gtrapprox \alpha_{c1}$, the proximity to the meta-nematic transition line produces a kink in the temperature dependence of the nematic order parameter.

We acknowledge that, for values of η_z that are not very small - such as the one considered here - the magnetic susceptibility of the anisotropic 3D system may have extra terms not contemplated by our formalism, such as corrections to the continuous in-plane dispersion. However, as we show in details in Appendix C, the same behavior is obtained for an anisotropic quadratic dispersion with equal momentum cutoff along all three momentum directions. This gives extra confidence that the phase diagram of Fig. 14 may be realized in at least some moderately anisotropic systems.

We point out that a phase diagram with the intermediate regime of a second-order nematic transition and first-order magnetic transition was earlier obtained in a semi-phenomenological model for the interaction between the structural and magnetic degrees of freedom[56]. Kim *et al* considered a microscopic version of that model[15], showing that anharmonic elastic terms may bring the system into the regime of split second-order nematic and first-order magnetic transitions (we discuss the coupling between the lattice and the nematic degrees of freedom in Section V below). Our results show that such a behavior can be obtained in a purely magnetic model, even if the coupling to structural degrees of freedom is negligibly small.

B. Quantum phase diagram as a function of u/g

To complement our analysis of thermal fluctuations, we now consider the opposite limit of $T = 0$, when the Matsubara frequency become a continuous variable and $\int_q = \int d^d q d\nu_n/(2\pi)^{(d+1)}$. Now r_0 is a function of some control parameter, e.g. doping, pressure, or applied field. We show that no new phases appear in the $T = 0$ limit, compared to the three phases that we found previously in the classical limit. As before, we consider first $d = 2$, then $d = 3$, and then arbitrary d between 2 and 3 and anisotropic 3D systems.

1. $d = 2$

Integrating the self-consistent equations for ϕ and r in (17) over both ν_n and $\mathbf{q}$, we obtain for $d = 2$

$$1 = \tilde{g}\left(\log\frac{\Lambda^2}{\sqrt{r^2-\phi^2}} + 1 - \frac{r}{\phi}\tanh^{-1}\frac{\phi}{r}\right)$$

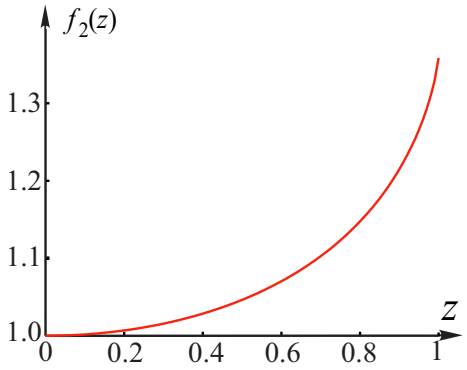

Figure 16: (Color online) The function $f_2(z)$ from Eq.(58).

$$r = \bar{r}_0 - \tilde{u}\left(r\log\frac{\Lambda^2}{\sqrt{r^2-\phi^2}} + r - \phi\tanh^{-1}\frac{\phi}{r}\right) \quad (56)$$

where $\tilde{g} = g/\left(4\pi^2\gamma\right)$, $\tilde{u} = u/\left(4\pi^2\gamma\right)$ are dimensionless couplings, and Λ is the upper limit of the integral over momentum. The parameter $\bar{r}_0$ again decreases as the system approaches the magnetic transition. Note that the dependence on Λ is still present even after we absorbed the $r-$ and ϕ-independent contribution of the right-hand side of Eq. (18) into $\bar{r}_0$. This dependence is eliminated only after we rescale all variables by Λ^2. The implication is that the relevant $\bar{r}_0$, r, and ϕ are all of the order of Λ^2 if the dimensionless $\tilde{g}$ and $\tilde{u}$ are of order one. Thus, unlike the classical case, where the dependence on Λ was fully absorbed into $\bar{r}_0$, here the low-energy behavior becomes non-universal. To proceed - and to make later comparison with the RG results - here we consider the case of $\bar{r}_0$ much smaller than Λ. This is achieved by taking the dimensionless couplings $\tilde{g}$ and $\tilde{u}$ to be small, which also implies that r and ϕ are small compared to Λ^2.

In the absence of the preemptive Ising-nematic instability, the magnetic instability would again occur at $\bar{r}_0 = 0$. Searching for the solution with $\phi \neq 0$, we introduce as before $z = \phi/r$, solve the first equation for $r(z)$ and substitute the result into the second equation. This gives the relation between $\bar{r}_0$ and z. For small $\tilde{g}$ this relation takes the form

$$\bar{r}_0 = \Lambda^2 \mathrm{e}^{-1/\tilde{g}}\,(1+\alpha)\,f_2(z) \quad (57)$$

where

$$f_2(z) = \frac{\mathrm{e}}{\sqrt{1-z^2}} \times \left(\frac{1-z}{1+z}\right)^{\frac{1}{2z}}. \quad (58)$$

Note that $\bar{r}_0 \ll \Lambda^2$ for $\tilde{g} \ll 1$, as we anticipated. The dependence on z comes through $f_2(z)$. This function, which we plot in Fig. 16, monotonically increases with increasing z, implying that the first instability occurs at $z = 1$. Because $f(z)$ does not depend on α, this result obviously holds for any α.

When $\tilde{g}$ and $\tilde{u}$ are not small, the relation between r_0 and z is more complex, but the result is the same - the first instability upon decreasing $\bar{r}_0$ is into the state with the largest $\phi = r$. As in earlier cases, an instability with $\phi = r$ implies that the Ising and magnetic transitions occur simultaneously and that both are first-order transitions. We analyzed the whole phase diagram and again found that there is only one transition line at which both ϕ and M jump to finite values. At smaller $\bar{r}_0$, ϕ and M monotonically increase.

2. $d = 3$

For $d = 3$, the dependence on the upper cutoff is more severe - it is power-law rather than logarithmic. The set of two equations on r and ϕ becomes, after we integrate over frequencies

$$\begin{aligned} \phi &= \frac{\tilde{g}}{2\pi}\int_0^{\Lambda^2} dy\sqrt{y}\tanh^{-1}\frac{\phi}{y+r} \\ r &= \bar{r}_0 + \frac{\tilde{u}}{4\pi}\int_0^{\Lambda^2} dy\sqrt{y}\log\frac{y^2}{(y+r)^2-\phi^2} \end{aligned} \quad (59)$$

The dependence on Λ can be eliminated by rescaling $(r,\phi,\bar{r}_0) \to \Lambda^2(r,\phi,\bar{r}_0)$ and $(\tilde{g},\tilde{u}) \to (\tilde{g},\tilde{u})/(2\pi\Lambda)$. Introducing, as before $z = \phi/r$, we rewrite Eqs. (59) in rescaled variables as

$$\begin{aligned} z &= \tilde{g}\sqrt{r}\int_0^{\frac{1}{r}} du\sqrt{u}\tanh^{-1}\frac{z}{u+1} \\ r &= \bar{r}_0 + \frac{\tilde{u}}{2}r\sqrt{r}\int_0^{\frac{1}{r}} du\sqrt{u}\log\frac{u^2}{(u+1)^2-z^2} \end{aligned} \quad (60)$$

One can easily make sure that the the first equation in (60) has a solution only when the rescaled $\tilde{g}$ is above the threshold $\tilde{g}_{cr} = 1/2$. Once $\tilde{g}$ is above $\tilde{g}_{cr}$, the relevant rescaled $\bar{r}_0$ is of order one, i.e. the actual $\bar{r}_0$ is of order Λ^2. While the model remains perfectly well defined, universal predictions with respect to the low energy behavior cannot be made. Still, like in $d = 2$, one can make the relevant $\bar{r}_0$ to be much smaller than Λ^2. For this, one has to place $\tilde{g}$ close to the critical value, $\tilde{g} = 1/2 + \epsilon$, and consider $\epsilon \ll 1/\Lambda$. Expanding in ϵ in (60) and relating r to z we obtain, in the original variables,

$$\bar{r}_0 = \frac{2}{\pi}\Lambda^3\,(\tilde{g}-\tilde{g}_{cr})\,(1+\alpha)\,f_3(z) \quad (61)$$

where

$$f_3(z) = \left[\pi - \frac{1}{z}\int_0^{\infty} du\sqrt{u}\left(-\frac{z}{u+1} + \tanh^{-1}\frac{z}{u+1}\right)\right]^{-1} \quad (62)$$

This function, which we plot in Fig.17, increases with z such that the first instability occurs into the state with $z = 1$, i.e. $\phi = r$. This implies that the Ising-nematic and magnetic orders appear simultaneously, via a first-order transition. We see therefore that at $T = 0$ there is no difference between $d = 3$ and $d = 2$ - a first-order simultaneous Ising/magnetic transition occurs in both cases.

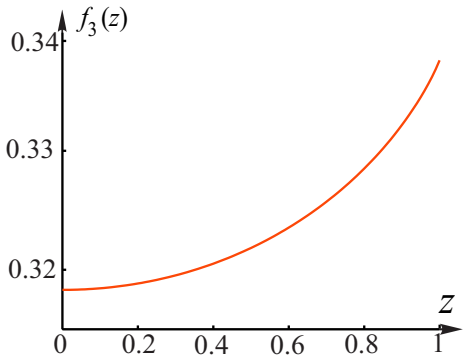

Figure 17: (Color online) The function $f_3(z)$ from Eq.(62).

3. $2 < d < 3$

We also analyzed the case of arbitrary d between 2 and 3 and found the same result as for $d = 2$ and $d = 3$: there is a first-order Ising-nematic transition which triggers a simultaneous first-order magnetic transition. One can easily show that this result holds for all $d > 1$ in the quantum limit, and for all $d > 3$ in the classical limit. The analogy between quantum systems in $d > 1$ and classical systems in $d > 3$ is not surprising because the dynamical exponent is $z_{dyn} = 2$, meaning that the behavior of a quantum system in d dimensions is the same as that of a classical system with $d+2$ dimensions.

4. *Anisotropic $d = 3$ case*

The same result - a simultaneous first-order nematic and magnetic transition, holds also for anisotropic $d = 3$ systems with the susceptibility χ_q given by Eq. (47). In this respect, at $T = 0$ there is no difference between the system behavior in $2 < d < 3$ and for anisotropic 3D dispersion, no matter what the degree of anisotropy is. We stress again that this conclusion is not universally true and that it could be possible to construct models with same order parameter symmetry that display second order transitions (see, for instance, Ref.[57]).

C. Phase diagram at arbitrary T

We now combine the quantum and classical analysis and consider the phase diagram at a finite $T\gamma \sim \bar{r}_0$ when both quantum and classical fluctuations are equally relevant.

1. $d = 2$

For $d = 2$, we have at high temperatures a second order Ising transition at $\alpha > 2$ and a first-order Ising transition into $\phi < r$ at $1 < \alpha < 2$. At $T = 0$, we have instead a first order transition into $\phi = r$ and $\bar{m} \neq 0$ for all α. A simple analysis shows that the behavior at any finite T remains the same as at high temperatures simply because at any $T > 0$ classical fluctuations do not allow a non-zero magnetic order. The value of α at which the first-order Ising transition becomes second order changes with T, but the phase diagram at any T still consists of a single line along which the system undergoes either first-order or second-order Ising transition.

2. $d = 3$

At $d = 3$, the behavior at large T and at $T = 0$ are identical - in both cases the first instability is into a state with $\phi = r$ (a simultaneous first-order Ising/magnetic transition). A simple analysis shows that this behavior holds for any T, no matter how small or large.

3. $2 < d < 3$

This case is the most interesting one. At high temperatures, all three types of transitions are realized, depending on α, while at $T = 0$ the system only undergoes a phase transition into a state with $\phi = r$. As a result, at a given d between 2 and 3, the character of the transition changes as a function of α at a fixed T, and as a function of T at a fixed α.

We verified that the phase diagrams do not change if we impose upper cutoff on the frequency summation rather than on the integration over momentum. The former is more convenient for numerical calculations, and below we use frequency rather than momentum cutoff. Physically, the frequency cutoff Λ_ν becomes more important than the momentum cutoff $\Lambda = \Lambda_q$ if the frequency dependence of the bosonic $\chi(\mathbf{q}, \nu_n)$ becomes stronger than $\gamma|\nu_n|$ at energies smaller than Λ_q.

For a generic $2 < d < 3$ and an arbitrary T, the equations for r and ϕ become, after integrating over momentum

$$\begin{aligned} r &= \bar{r}_0 - \frac{\bar{u}}{4}\sum_{\nu_n}\Big[-2\left(\gamma|\nu_n|\right)^{\frac{d-2}{2}} + \left(r+\phi+\gamma|\nu_n|\right)^{\frac{d-2}{2}} \\ &\quad + \left(r-\phi+\gamma|\nu_n|\right)^{\frac{d-2}{2}}\Big] \\ \phi &= \frac{\bar{g}}{4}\sum_{\nu_n}\Big[\left(r+\phi+\gamma|\nu_n|\right)^{\frac{d-2}{2}} \\ &\quad - \left(r-\phi+\gamma|\nu_n|\right)^{\frac{d-2}{2}}\Big] \end{aligned} \tag{63}$$

where $\bar{u}$ and $\bar{g}$ are defined in (37). The frequency summation extends up to $n = n_{max} = \Lambda_\nu/(2\pi T)$. Introducing as before $z = \phi/r$ and rescaling in addition the temperature $\bar{T} = \gamma T/\bar{r}_0$, we numerically extract r as a function of z from the second equation, substitute into the first equation, and obtain $\bar{r}_0$ as a function of z. In Fig. 18 we

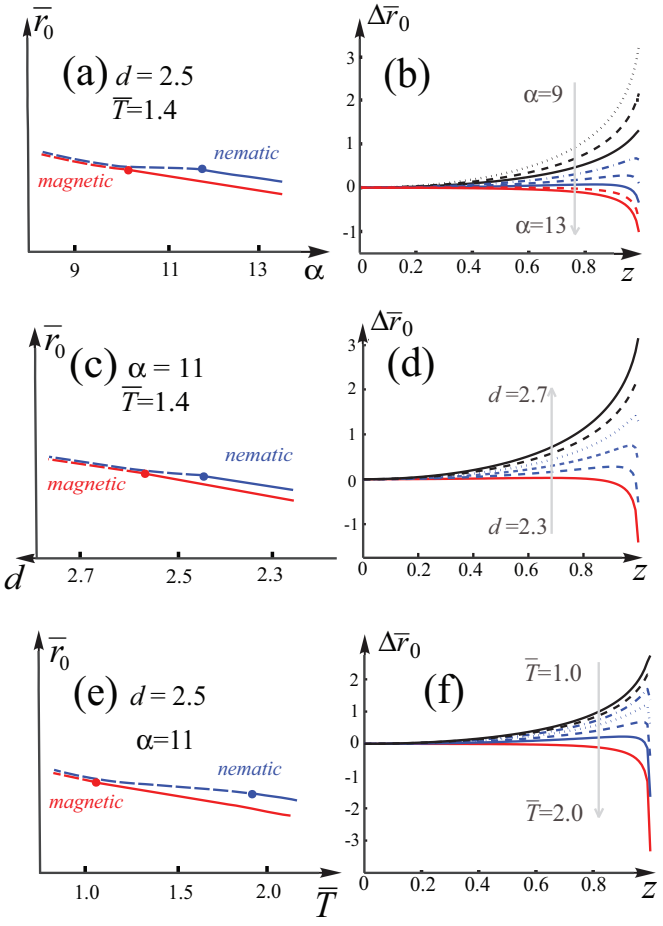

Figure 18: (Color online) The phase diagram as a function of the ratio of the couplings α at fixed reduced temperature $\bar{T}$ and dimensionality d (panels (a)-(b)), as a function of d at fixed α and $\bar{T}$ (panels (c)-(d)), and as a function of $\bar{T}$ at fixed α and d (panels (e)-(f)). The parameters α and $\bar{T}$ are defined in the text. We set $n_{max} = 100$.

plot $\bar{r}_0(z)$ and the phase diagrams upon varying α, d and $\bar{T}$. We see the same trend as in the classical phase diagram, namely, as α gets larger at some fixed $\bar{T}$ and d, the Ising and magnetic transitions split, and at even larger α the Ising-nematic transition eventually becomes second order (panels (a)-(b)). The same trend holds upon the reduction of dimensionality (i.e., make the system more two-dimensional) at a fixed α and $\bar{T}$ (panels (c)-(d)) and also upon decreasing $\bar{T}$ at a fixed α and d (panels (e)-(f)).

Overall, we see that the phase diagram as a function of α does not change qualitatively between high temperatures, when the classical approximation is valid, and $\bar{T} = O(1)$, when both classical and quantum fluctuations are relevant. The only real difference is the change in the critical values α_{c1} at which the Ising-nematic and magnetic transitions split, and α_{c2} at which the Ising-nematic transition becomes second-order. As $\bar{T}$ gets smaller, both critical α become larger. Note that the rescaling of T to $\bar{T} = \gamma T/\bar{r}_0$ is important for obtaining the correct temperature dependence of the transitions due to hidden T dependence in $z = \phi/r$ via $r = r(T)$.

4. *Anisotropic $d = 3$ case*

We analyzed the two cases of anisotropic magnetic dispersion, Eqs. (47) and (48), at a finite T. In each case we found the same behavior as at large T, with different phase diagrams depending on the degree of anisotropy, as discussed in Fig. 12. For strongly anisotropic, quasi-2D spin susceptibility χ_q, the results are similar to the case $2 < d < 3$, i.e the magnetic tri-critical point is located on the left of the nematic tri-critical point, and in the intermediate region the nematic transition is first-order, while the magnetic transition is second-order (recall Fig. 18). For weaker anisotropy, however, the two tri-critical points interchange, as we observed in the purely classical analysis, see Fig. 12. Once this happens, in the intermediate region between the two tri-critical points the nematic transition is second-order, while the magnetic transition is first-order. The only difference with respect to the purely classical case is that both α_{c1} and α_{c2} shift to larger values at a smaller T.

IV. RG ANALYSIS AT A FINITE N

The mean-field analysis is quite straightforward, but it is rigorously justified only in the artificial limit of large N, where, we remind, N is the number of components of the $\mathbf{\Delta}$ fields. The actual number of spin components is $N = 3$, and it is by no means guaranteed that the behavior at $N = 3$ is the same as at large N. To verify this, we need to return back to the effective action $S_{\text{eff}}[\mathbf{\Delta}_X, \mathbf{\Delta}_Y]$, Eq. (7), and use a complementary approach which is not restricted to large N. One such approach, commonly used to study phase transitions, is the renormalization group (RG) technique. In RG, one progressively integrates out contributions from high energies down to E and analyzes how the parameters of the effective model vary with $L \equiv \log W/E$, where W is the bandwidth (the highest energy scale in the problem). In our case, the parameters are u and g, and in the RG approach one studies the flow of the running couplings u_L and g_L and of any other coupling generated by the RG flow. Alternatively, one can vary the distance to a transition, i.e. vary $\bar{r}_0$, in which case $L = \log mW/\bar{r}_0$.

Quite generally, the RG flow may lead to three types of behavior depending on the structure of the RG equations and on the bare values u and g. One possibility is that the couplings u_L and g_L flow to zero, which implies that there is no preemptive Ising-nematic transition. Another possibility is that g_L and u_L flow to infinity (more accurately, to strong coupling) in such a way that the stability condition for the effective model $u_L > g_L$ is not broken. In this situation, the system undergoes a second-order

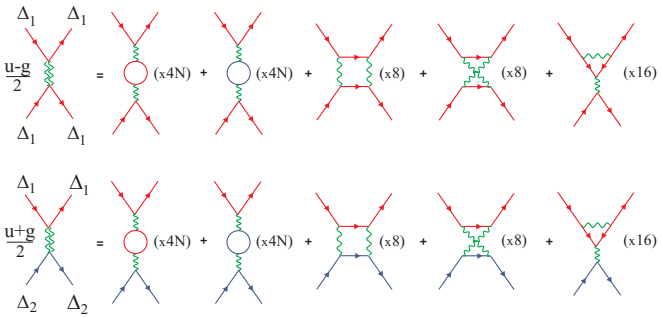

Figure 19: (Color online) The one-loop diagrams responsible for the renormalization of the two bosonic vertices $u_{11} = (u-g)/2$ and $u_{12} = (u+g)/2$, with their respective combinatorial factors. The wavy lines correspond to the interactions, whereas the continuous lines correspond to Δ_1 and Δ_2 (red/light gray and blue/dark gray lines, respectively). Notice that the two first diagrams in each line contain a closed bosonic loop and, therefore, an overall pre-factor of N.

Ising-nematic transition at the scale $L_{cr} = \log mW/\bar{r}_0^{cr}$ at which g_L diverges. The third possibility is that the stability condition $u_L > g_L$ gets broken at some $L^* < L_{cr}$. In this case, the effective action becomes unstable with respect to a discontinuous variation of $\mathbf{\Delta}^2_{X,Y}$, and the system undergoes a first-order transition.

The RG approach is still a weak-coupling approach in the sense that the bare couplings u and g are assumed to be much smaller than the bandwidth. The advantage of the RG technique is that it can be applied to any N, and, from this perspective, it goes beyond mean-field approximation. However, the RG approach has its own limitations - it can be rigorously justified only in the marginal dimension $d_{eff} = d + z_{dyn} = 4$, when the renormalizations are logarithmic. In our case $z_{dyn} = 2$, meaning that the marginal behavior takes place at $T = 0$ and $d = 2$. Still, we obtained the $N = \infty$ phase diagram at $T = 0$ and $d = 2$ in the previous section, and it is instructive to compare that phase diagram with the RG phase diagram at arbitrary N to verify whether or not the behavior at $N = 3$ is the same as at $N = \infty$, at least in this particular case. We remind that our $N = \infty$ quantum analysis in $d = 2$ shows that the system undergoes a first-order transition into the magnetic state with $\phi = r$ for any $\alpha > 1$. We now analyze how the phase diagram looks like for arbitrary N.

We derived the one-loop RG equations for the flow of the running couplings u_L ad g_L by the momentum-shell method (see, for instance, Ref.[58]) and also derived them by analyzing the parquet diagrams for the renormalization of the four-boson vertices, which are presented in Fig. 19 together with their respective combinatorial prefactors. Notice the special role of the diagrams which contain a closed bosonic loop. Summation over the internal bosonic indices yields a factor N, which does not appear in the diagrams without closed loops. The mean-field results are reproduced if we consider only these diagrams.

For further convenience, we rescale the coupling constants to $u, g \to 2N\,(u, g)$. Combining logarithmic contributions from all parquet diagrams we can cast the renormalizations of u and g into the form of differential RG equations

$$\begin{aligned} \dot{u}_L &= -2\left(1+\frac{4}{N}\right)u_L^2 - \frac{4}{N}g_L^2 + \frac{4}{N}u_L g_L \\ \dot{g}_L &= 2\left(1+\frac{2}{N}\right)g_L^2 - \frac{12}{N}u_L g_L \end{aligned} \tag{64}$$

where $\dot{X}_L = dX_L/dL$.

For the particular cases $N = 3$ and $N = 6$, these equations reduce to those obtained in Refs.[34] and[57] for the similar model. We checked that no other couplings allowed by symmetry are generated by the RG flow if only u and g have non-zero initial values. In particular, if the third coupling v in Eq. (7) is initially zero, it is not generated by RG.

We recall that both u_L and g_L are initially positive (bare values u and g are given by Eqs. 8), and that the bosonic action (7) is stable as long as $u_L/g_L > 1$. If this condition breaks down at some L, one of the coefficients of the quartic terms becomes negative and the system undergoes a first-order transition into a state with a non-zero ϕ.

For $N \to \infty$ the equations for u_L and g_L decouple and we can easily solve them and obtain

$$\begin{aligned} u_L &= \frac{u}{1+2uL} \\ g_L &= \frac{g}{1-2gL} \end{aligned} \tag{65}$$

We see that u_L flows to zero, while g_L increases under the RG flow, diverging at $L_{cr} = \frac{1}{2g_0}$. If g_L would be the only parameter in the problem, this divergence would indicate a preemptive second-order Ising-nematic instability, since the susceptibility of the Ising-nematic order parameter diverges at L_{cr}. However, in our case there are two couplings, and the action is stable only as long as $u_L > g_L$. This condition breaks down at a smaller $L^* = \frac{1}{4g_0} - \frac{1}{4u_0}$, before g_L diverges. The outcome is that for $N = \infty$ and $d = 2$ the system undergoes a first-order Ising-nematic transition at $T = 0$. This is in agreement with the mean-field analysis.

At a finite N, the two equations are coupled and both u_L and g_L can diverge. To understand what happens in this case, it is convenient to define the ratio $p_L = u_L/g_L$ and re-express the flow equations in terms of g_L and p_L:

$$\dot{p}_L = 2g_L\left[\left(-1+\frac{2}{N}\right)p_L^2 - p_L - \frac{2}{N}\right] \tag{66}$$

$$\dot{g}_L = 2g_L^2\left(1+\frac{2}{N}-\frac{6}{N}p_L\right) \tag{67}$$

It is straightforward to verify that this set of RG equations has several fixed trajectories along which p_L is a

constant and g_L evolves. The fixed trajectories are obtained by setting $\dot{p}_L = 0$ in Eq. (66). Solving the quadratic equation $(-1+2/N)\,p_L^2 - p_L - 2/N = 0$ we find two fixed values $p_1 = -1$ and $p_2 = -2/(N-2)$. The coupling g_L diverges along the fixed trajectory with $p_L = p_1$ as $\dot{g}_L = 2g_L^2\,(1+8/N)$. Along the second fixed trajectory p_2, g_L evolves according to $\dot{g}_L = 2g_L^2\left(N^2+8\right)/\left(N^2-2N\right)$, i.e. it diverges for $N > 2$ and tends to zero for $N < 2$.

To understand which trajectory is stable and which is not, we consider small deviations from a fixed trajectory, $p_L = p_i + \delta p_{iL}$, and expand the flow equations to lowest order in δp_{iL}. We obtain

$$\begin{aligned} \delta\dot{p}_{1L} &= 2g_L\delta p_{1L}\left(\frac{N-4}{N}\right) \\ \delta\dot{p}_{2L} &= -2g_L\delta p_{2L}\left(\frac{N-4}{N}\right) \end{aligned} \tag{68}$$

We see that p_1 is a stable fixed trajectory for $N < 4$ and unstable for $N > 4$, while p_2 is a stable fixed trajectory for $N > 4$ and unstable for $N < 4$.

There is also the third fixed trajectory $g_L = 0$ ($p_3 = \infty$). Expanding near $g_L = 0$ we obtain from Eq. (64)

$$\begin{aligned} \dot{u}_L &= -2\left(1+\frac{4}{N}\right)u_L^2 \\ \dot{g}_L &= -\frac{12}{N}u_L g_L \end{aligned} \tag{69}$$

whose solution is

$$\begin{aligned} u_L &= \frac{u}{1+2\left(1+\frac{4}{N}\right)uL} \\ g_L &= g\left(\frac{u_L}{u}\right)^{\frac{6}{N+4}} \end{aligned} \tag{70}$$

The fixed trajectory $g_L = 0$ is stable as long as g_L remains small compared to u_L. Evaluating $p_L = u_L/g_L$ from the solutions of (70) we find

$$p_L = \frac{u}{g}\left(\frac{u_L}{u}\right)^{\frac{N-2}{N+4}} = \frac{u}{g}\left[1+2\left(1+\frac{4}{N}\right)uL\right]^{\frac{2-N}{N+4}} \tag{71}$$

We see that, for $N > 2$, p_L decreases under the RG flow such that eventually g_L exceeds u_L. This implies that the fixed trajectory $g_L = 0$ is unstable. For $N < 2$, however, p_L increases under the RG, and the trajectory $g_L = 0$ is stable.

Combining this analysis with the numerical solution of Eqs. (66) and (67) at intermediate energies, we obtain three different regimes of system behavior depending on the value of N.

For $N > 4$, the RG trajectory is as shown in Fig. 20 (a)-(b). For arbitrary $\alpha = u/g$, the system approaches the stable fixed trajectory $p_L = p_2 = -2/(N-2)$. Since $p_2 < 0$, and the bare value of p is positive and larger than one, the running coupling p_L necessarily becomes

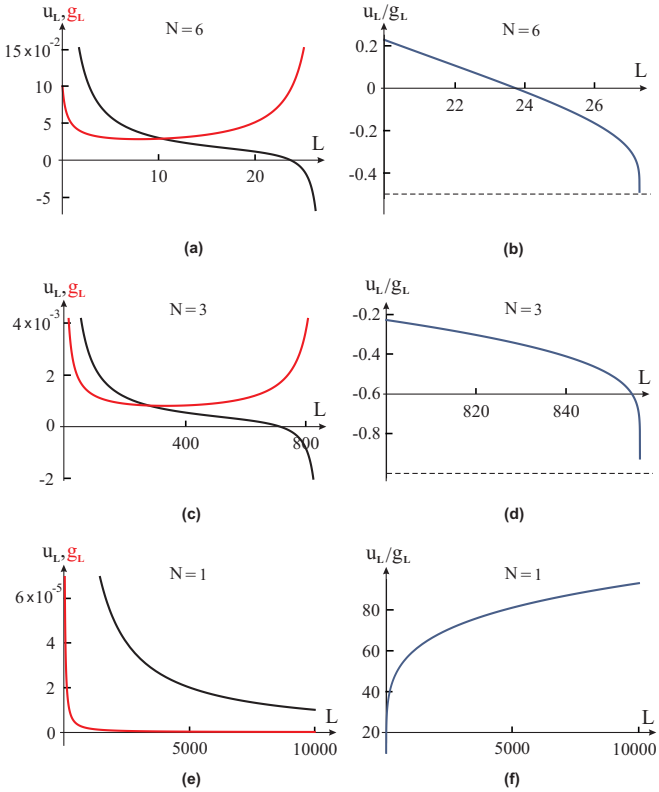

Figure 20: (Color online) RG flow of the running coupling constants u (black line) and g (red/light gray line), as well as their ratio u/g (blue/dark gray line), as function of L. In (a)-(b), we present the results for $N = 6$; in (c)-(d), we display the results for $N = 3$; in (e)-(f) the results shown are for $N = 1$. In all cases, the initial values were $u = 1$, $g = 0.1$. The dashed lines in (b) and (d) refer to the stable fixed trajectories $p_2 = -2/(N-2)$ and $p_1 = -1$, respectively.

one at some $L = L^*$ along the RG flow. At this L^* the action becomes unstable and the system undergoes a first-order transition into a state with a non-zero Ising-nematic order parameter. The only difference between finite $N > 4$ and $N = \infty$ is that, for a finite N, u_L eventually flows to infinity while for $N \to \infty$ it flows to zero. This difference, however, does not play any role in our consideration as the RG flow makes sense only as long as u_L/g_L remains larger than one.

For $2 < N < 4$, the fixed trajectory p_2 becomes unstable and cannot be reached if the RG flow starts with $u > g > 0$. The stable trajectory to which the system flows is now $p_1 = -1$, as shown in Fig. 20 (c)-(d). Near this fixed trajectory we find from Eqs. (68) and (67) that g_L increases and diverges at some L_{cr}, while $p_L + 1 = \delta p_{1L} \propto (1/g_L)^{(4-N)/(N+8)} \to 0$.

While the fixed trajectory is different for $2 < N < 4$ and for $N > 4$, the behavior relevant to our purposes remains the same for all $N > 2$, namely in the process of RG flow towards a fixed trajectory, the ratio u_L/g_L reaches one at some $L = L^*$. At this point, the action becomes unbounded and the system undergoes a first-

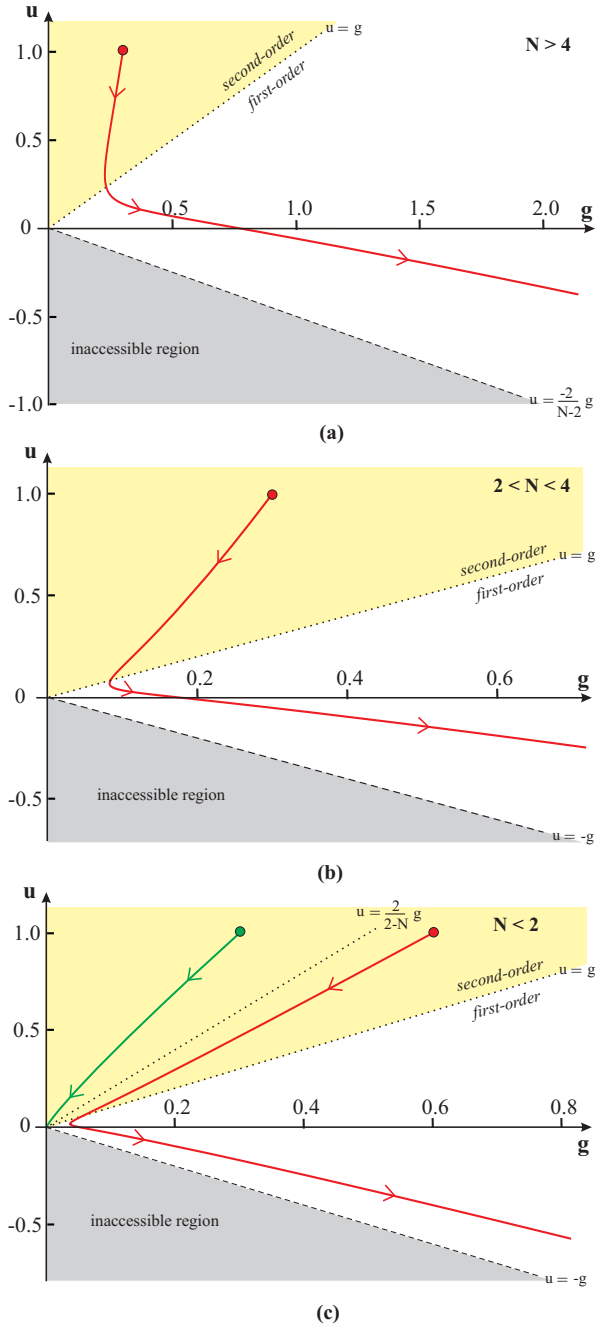

Figure 21: (Color online) Structure of the RG flow in the (g, u) plane for the representative values (a) $N = 6$, (b) $N = 3$, and (c) $N = 1$. The fixed trajectories $u = -p_i g$ are shown as dashed lines in the lower-half planes ($p_1 = -1$ for (b) and (c) and $p_2 = -2/(N-2)$ in (a)), and are approached only at very large L. In the upper-half planes, the dotted line $u = g$ separates the second-order and first-order regimes. In (c), the dotted line $u = 2g/(2-N)$ separates initial conditions that fall in the basis of attraction of the $g = 0$ fixed point ($p_3 = \infty$, green/light solid line) from those that flow to the $p_1 = -1$ fixed trajectory (red/dark solid line).

order phase transition into a state with a finite Ising-nematic order parameter. The dependence on N is then only quantitative - as N decreases, the scale L^* gets progressively larger. To see this, we note that for any finite N, the couplings u_L and g_L first decrease under the RG flow and only at larger L reverse the trend and approach the fixed trajectory at $g_L \to \infty$ and $u_L \to -\infty$. To get an estimate of how L^* evolves with N, we use the approximate Eq.(71) and identify L^* with the RG scale at which p_L becomes one. We obtain

$$L^* \propto \left(\frac{u}{g}\right)^{\frac{N+4}{N-2}} \tag{72}$$

For a given $u/g \gg 1$, the value L^* rapidly increases with decreasing N; for instance, for $u/g = 10$, the ratio of L^* for $N = 3$ and for $N = 5$ is 10^4. This means that at smaller N the system behaves over a wide range of energies as if magnetic fluctuations were absent, and only very near $T_{N,0}$ it recognizes that it actually undergoes a first-order nematic transition.

For $N < 2$, a new behavior becomes possible, as shown in Fig. 20 (e)-(f). Now the fixed trajectory $p_2 = 2/(2-N)$ crosses the region $u > g > 0$ from where the RG flow begins. Once the bare u and g are such that $u/g > 2/(2-N)$, the RG flow is sandwiched between the fixed trajectories p_2 and p_3 which are respectively unstable and stable for $N < 2$. The RG flow then moves both g_L and u_L towards $g_L, u_L = 0$, keeping $g_L < u_L$, i.e. without crossing the first-order instability line. In this situation, no preemptive nematic instability develops, and the system only undergoes a mean-field magnetic transition at $T_{N,0}$. If, however, the initial $u/g < 2/(2-N)$, the system behavior is the same as before, with the couplings evolving towards the fixed trajectory $p_1 = -1$, and p_L becoming equal to one at some scale L^*, at which the system undergoes a first-order nematic transition. In the formal limit $N = 0$, the whole region $u > g$ falls into the basis of attraction of the $g_L = u_L = 0$ fixed point, i.e. there is no preemptive nematic instability for any $u/g > 1$. The general structure of the RG flow equations in the (g, u) plane is shown in Fig. 21.

For $d = 2$ and $T = 0$ this scenario of no preemptive nematic instability holds only for $N < 2$ and does not affect our actual case of $N = 3$, for which the system behavior under the RG flow is qualitatively the same as in the mean-field, $N = \infty$ analysis. What happens for $d > 2$ and/or a finite T is unclear because, in the absence of the logarithmic terms, the approximations leading to the RG equations are not justified. It is possible in principle that a preemptive nematic transition does not occur for some large enough $\alpha = u/g$. If this is the case, then there must be a reentrant behavior for large α in Fig. 8, i.e. the magnetic transition temperature T_N must reverse trend and come closer to the nematic instability. However, the more likely scenario is that the phase diagram which we obtained in the mean-field approximation survives for the actual $N = 3$ component bosonic field for all α.

V. CONSEQUENCES OF THE ISING-NEMATIC ORDER

A. Orbital Order

ARPES measurements on detwinned samples have found that the onset of resistivity anisotropy is accompanied by the onset of orbital order in the paramagnetic phase, with different occupations for the d_{xz} and d_{yz} orbitals[10]. One possibility, explored by several authors in different contexts, is that this orbital ordering is an intrinsic instability of the system[18–24,26,27]. In line with the theme of this work, we explore another possibility, namely that the orbital order is induced by the Ising-nematic order. This scenario is generally consistent with the small value of the measured orbital polarization. To investigate this scenario quantitatively, we consider a simplified two-orbital model in which the entire X pocket has d_{yz} character, while the entire Y pocket has d_{xz} character[30], and assume that there is a splitting $\Delta_{\rm orb}$ between the on-site energies of the d_{xz} and d_{yz} orbitals. In the presence of such splitting, the fermionic dispersion becomes anisotropic, see Fig. 22, and the Hamiltonian acquires the additional terms

$$\mathcal{H}_{\rm orb} = -\sum_{\mathbf{k}} \Delta_{\rm orb} \left(c^{\dagger}_{X,\mathbf{k}\alpha} c_{X,\mathbf{k}\alpha} - c^{\dagger}_{Y,\mathbf{k}\alpha} c_{Y,\mathbf{k}\alpha} \right) + \frac{a_0}{2} \Delta^2_{\rm orb} \tag{73}$$

Including these two terms into the Hubbard-Stratonovich procedure and expanding the effective action in powers of $\Delta_{\rm orb}$ we obtain

$$\begin{aligned} S\left[\mathbf{\Delta}_X, \mathbf{\Delta}_Y, \Delta_{\rm orb}\right] &= S_{\rm eff}\left[\mathbf{\Delta}_X, \mathbf{\Delta}_Y\right] + \frac{a}{2}\Delta^2_{\rm orb} \\ &\quad - w\left(\Delta^2_X - \Delta^2_Y\right)\Delta_{\rm orb} \end{aligned} \tag{74}$$

with $S_{\rm eff}\left[\mathbf{\Delta}_X, \mathbf{\Delta}_Y\right]$ given by Eq. (7) and:

$$\begin{aligned} a &= a_0 + 4\int_k G^2_{X,k} \\ w &= 2\int_k G_{\Gamma,k} G^2_{X,k} \end{aligned} \tag{75}$$

Evaluating the integrals and expanding around perfect nesting we find $a > 0$ and $w = -c\mu$ with $c > 0$. The w term describes the linear coupling between the orbital and Ising-nematic order parameters, i.e. the development of one order triggers the development of the other. Differentiating Eq. (74) with respect to $\Delta_{\rm orb}$ we obtain

$$\langle \Delta_{\rm orb} \rangle = \frac{w}{a}\left\langle \Delta^2_X - \Delta^2_Y \right\rangle \tag{76}$$

Since $w \propto -\mu$ and μ scales with doping, the induced orbital order is expected to be small at small doping. This is in accordance with the experimental data. The sign of the orbital splitting also agrees with the data[10]: polarized ARPES measurements on detwinned electron-doped $Ba\left(Fe_{1-x}Co_x\right)_2 As_2$ samples have shown that when $\left\langle \Delta^2_X \right\rangle > \left\langle \Delta^2_Y \right\rangle$, $\langle \Delta_{\rm orb} \rangle$ is negative. This is consistent with (76) since for electron-doped materials $\mu > 0$ and $w < 0$. To the best of our knowledge, similar measurements have not been carried out for hole-doped samples. Interestingly, Eq. (76) shows that the sign of $\langle \Delta_{\rm orb} \rangle$ should change for hole-doped materials, which are described in our model by $\mu < 0$. Note that long-range magnetic order can also induce orbital polarization[22].

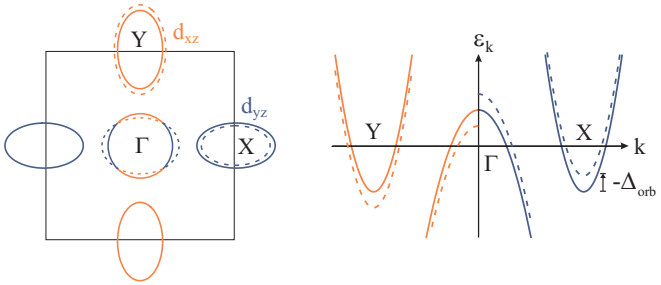

Figure 22: (Color online) Schematic representation of the effect of orbital ordering on the Fermi surface (left panel) and on the band dispersions (right panel). The light/orange curves have d_{xz} orbital character, while the dark/blue curves have d_{yz} orbital character. The solid lines correspond to the Fermi surface and band dispersions without orbital order ($T > T_s$), whereas the dashed lines refer to temperatures below the onset of orbital order ($T < T_s$). In this figure we considered $\Delta_{\rm orb} < 0$.

B. Structural order

The same reasoning also applies to the interplay between the Ising-nematic and the orthorhombic orders. The structural order is detected experimentally as the difference between the lattice constants a and b along the x and y directions of the Fe-plane, respectively. In the ideal situation, structural order appears only below a particular structural transition temperature T_s. In reality, however, some orthorhombic distortion can be present at any T due to internal strains. In the case of detwinned samples, a small in-plane strain is applied along one of the orientations[4,6,59,60]. Then, strictly speaking, $\varepsilon_s = a - b$ is never zero, i.e. there is no well-defined T_s for finite strain. Still, experimentally one can identify the crossover temperature below which ε_s sharply increases.

The relationship between Ising-nematic and structural orders can be obtained in the same way as in the previous subsection. Introducing the orthorhombic order parameter in a way similar to Eq. (73) yields

$$\mathcal{H}_{\rm str} = -\lambda \sum_{\mathbf{k}} \varepsilon_s \left(c^{\dagger}_{X,\mathbf{k}\alpha} c_{X,\mathbf{k}\alpha} - c^{\dagger}_{Y,\mathbf{k}\alpha} c_{Y,\mathbf{k}\alpha} \right) + \frac{C_s}{2}\epsilon^2_s \tag{77}$$

where λ is a coupling constant and C_s is the shear modulus. Including these two terms into the Hubbard-

Stratonovich procedure we obtain, in the mean-field approximation, the same effective action as Eq. (74), but with ε_s instead of $\Delta_{\rm orb}$ and with the renormalized shear modulus C_s instead of a. Accordingly, the orthorhombic and Ising-nematic order parameters are linearly related:

$$\langle \varepsilon_s \rangle = \frac{\lambda w}{C_s} \left\langle \Delta_X^2 - \Delta_Y^2 \right\rangle \tag{78}$$

i.e. one order immediately triggers the other. Notice, however, that in distinction to the case of orbital order (Eq. 76), the proportionality constant between $\langle \varepsilon_s \rangle$ and the nematic order parameter contains not only the parameter w, but also the magneto-elastic coupling λ, which can itself depend on the chemical potential and on additional details of the band structure.

The linear relation between the two orders has been discussed in a number of earlier papers[32,35,61], and is not surprising because both orders break the same Z_2 symmetry between the x and y directions (the orbital order does the same). It also implies that, in detwinned samples, the applied strain plays the role of a conjugate field to the Ising-nematic order parameter[40]. We recall that the proportionality coefficient w scales with μ and is generally small. As a result, if the orthorhombic order is induced by the Ising-nematic order at $T = T_s$, the orthorhombic order parameter is initially small and may become visible only at some distance below the Ising-nematic transition. This may explain why recent magnetic torque experiments observed nematic order up to higher temperatures than the structural order[11]. We emphasize, however, that there is only one well-defined transition temperature below which the tetragonal symmetry is broken.

Note that the coupling to structural degrees of freedom also renormalizes the nematic coupling constant as[35]

$$\tilde{g} = g + \frac{\lambda^2 w^2}{C_s} \tag{79}$$

Therefore, even if initially $g = 0$, the coupling to the lattice generates a non-zero nematic coupling and, consequently, nematic order[61].

C. Pseudogap behavior

Recent ARPES data on NaFeAs, whose strongly anisotropic Fermi surface is very similar to that of the 1111 compounds[62], show that the reconstruction of the fermionic spectrum begins at temperatures around $T_s \approx 54$ K, where the nematic order sets in, rather than at $T_N \approx 39$ K, where the stripe magnetic order develops[36]. Such a reconstruction increases below T_N, which is a good indication that the effect likely has a magnetic origin.

Thermal magnetic fluctuations do give rise to a pseudogap behavior (often termed as magnetic precursors) by transferring spectral weight from small frequencies to a frequency comparable to the spectral gap developed in the magnetically ordered state[63]. The only precondition for magnetic precursors is that the magnetic correlation length ξ must be large enough.

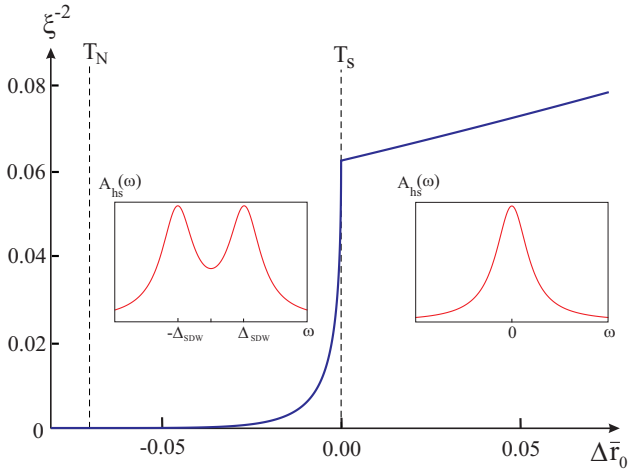

Figure 23: (Color online) Inverse squared magnetic correlation length $\xi^{-2} \propto r$ (in arbitrary units), as function of the reduced temperature $\Delta\bar{r}_0 \propto T - T_s$, for the case $\alpha = 4$ in the phase diagram of Fig. **8**. The dashed lines denote the split second-order nematic (T_s) and magnetic (T_N) transitions. Note the sharp increase and the discontinuity in the derivative of ξ at the nematic transition. The insets show schematically, at different temperatures, the frequency-dependence of the fermionic spectral function $A_{\rm hs}(\omega)$ at the hot spots, i.e. the Bragg points of the band dispersion $\varepsilon_{\Gamma,\mathbf{k}} = \varepsilon_{(X,Y),\mathbf{k}+\mathbf{Q}_i}$. Notice the development of peaks at $\omega = \pm\Delta_{\rm SDW}$ below the nematic transition, where $\Delta_{\rm SDW}$ denotes the value of gap that is opened due to SDW order at zero temperature.

How can the pseudogap develop at the nematic transition? At first glance, this seems a mere coincidence because the nematic order is a collective instability in the particle-hole channel with momentum $q = 0$, and a $q = 0$ boson cannot reconstruct the fermionic spectrum. However, one of our key results is that at the nematic transition the magnetic correlation length increases either discontinuously or very sharply. Even away from the nematic tri-critical point, where the Ising-nematic transition is second-order, the magnetic correlation length has a discontinuous derivative at T_s, since $\xi^{-2} \to \xi^{-2} - \phi$, making it increase faster. This behavior is shown in Fig. 23, where we plot the inverse magnetic correlation length $r \propto \xi^{-2}$ as a function of temperature for $\alpha = 4$ for the phase diagram of Fig. 8**.** Interestingly, in NaFeAs, as well as in other iron pnictides[64], a significant enhancement of magnetic fluctuations was observed[65,66] just below T_s by nuclear magnetic resonance (NMR). Once the correlation length jumps (or sharply increases) to a larger value, the strength of thermal magnetic fluctuations rapidly increases. As a result, the fermionic spectral function $A_{\mathbf{k}}(\omega)$ develops a magnetic pseudogap via the transfer of spectral weight from zero to finite frequencies at the

hot spots $\varepsilon_{\Gamma,\mathbf{k}} = \varepsilon_{(X,Y),\mathbf{k}+\mathbf{Q}_i}$ (see Fig. 23), what leads to the reconstruction of the fermionic spectrum, although zero-frequency states appear only below T_N. We argue therefore that the jump (or sharp increase) in ξ at the nematic transition is the "glue" that links the nematic instability and the development of the pseudogap in the fermionic spectral function.

VI. COMPARISON WITH EXPERIMENTS AND EARLIER THEORIES

A. Experimental phase diagrams

We now compare our results of Section III for anisotropic systems to the experimental phase diagrams of the 122 and 1111 compounds. Our theoretical phase diagram is plotted in Fig. 1 (a) and (b). Our goal is to relate the changes of the system behavior as function of the parameter $\alpha = u/g$ to the measured changes imposed by doping, pressure, and alkaline-earth substitution. For this we need to (i) properly place the parent compounds onto our phase diagram, (ii) decide which of the two similar but not identical phase diagrams in Fig. 1 is more appropriate, and (iii) understand how α varies with the experimental parameters. We first consider 122 materials and then briefly discuss 1111 materials.

We begin with the parent 122 materials. There are three well studied types of 122 systems, namely, $CaFe_2As_2$, $SrFe_2As_2$, and $BaFe_2As_2$. The first two materials undergo a strong simultaneous magneto-structural first-order transition, as evidenced from several thermodynamic measurements[67–69]. This clearly places both materials well into the region of small α, to the left of the first tri-critical point - either α_{c1} in Fig. 1(a) or α_{c2} in Fig. 1(b). For $BaFe_2As_2$, X-ray diffraction[15,70] as well as high-accuracy magnetization measurements[16] find a very small splitting between the structural and magnetic transitions ($T_s \sim 141$ K, $T_N \sim 140$ K). This places $BaFe_2As_2$ in the region of larger α (i.e. smaller g), which can be either slightly to the right of the magnetic tri-critical point α_{c1} in Fig. 1(a), or slightly to the right of the nematic tri-critical point α_{c2} in Fig. 1(b). A smaller α for the parent compounds $CaFe_2As_2$, $SrFe_2As_2$ compared to the parent compound $BaFe_2As_2$ can be explained by the difference in the values of the shear modulus C_s. The analysis of their phase diagrams under pressure[71–73], as well as of their mechanical properties[74], shows that $CaFe_2As_2$ and $SrFe_2As_2$ are softer than $BaFe_2As_2$. According to our Eq. (79), softer systems with smaller shear modulus C_s (and possibly larger magneto-elastic coupling λ) have larger g, and hence smaller α.

Which of the two phase diagrams in Fig. 1 is more appropriate for $BaFe_2As_2$ is a more subtle issue. The x-ray data[15,70] show that the orthorhombic order parameter evolves continuously immediately below T_s and then jumps at the same temperature where the magnetic transition takes place, as evidenced by the magnetization data[16]. Neutron diffraction data[75] do not detect critical magnetic fluctuations above T_N, consistent with the idea that the magnetic transition is first-order and simultaneous to the meta-nematic transition. This favors the phase diagram of Fig. 1(b). However, the same neutron data do not detect a clear jump of the magnetic order parameter at the magnetic transition, as expected at a first-order transition. So, it is possible in principle that the continuous evolution of the nematic/orthorhombic order parameter immediately below T_s is a secondary effect - due to internal strain, for instance. Then, the true nematic transition would be first-order, while the magnetic transition occurring at $T_N \leq T_s$ would be second order. This would be consistent with the phase diagram of Fig. 1(a). In any case, however, the key observation is that the parent compound $BaFe_2As_2$ is in the region where the structural and magnetic transitions are quite close to each other and at least one of them is first order.

To investigate the effect of doping and pressure, we need to understand how $\alpha = u/g$ changes with the chemical potential μ and the mass anisotropy of the electron pockets $\delta m = m\left(m_x - m_y\right)/2m_x m_y$. We use Eqs. (8) and evaluate u and g for small chemical potential μ/ε_0 and small δm. For the ratio u/g we obtain

$$\alpha = \frac{u}{g} \approx 42 \left(\frac{T}{\varepsilon_0 \delta m}\right)^2 \left[1 + 0.9 \left(\frac{\mu}{T}\right)^2 + 0.01 \left(\frac{\varepsilon_0 \delta m}{T}\right)^2\right] \tag{80}$$

Electron doping adds carriers to the electron pockets and increases the magnitude of the chemical potential μ. This, according to (80), *increases* α, i.e., under electron doping the system should move to the regime of split second-order transitions (see Fig. 1). This is the main prediction of our theory.

This prediction generally agrees with the experimental phase diagrams of electron-doped 122 materials, where Fe is substituted by Co, Ni, Cu, Pd, or Rh[76]. In particular, for $Ba\left(Fe_{1-x}Co_x\right)_2 As_2$, magnetic and x-ray measurements demonstrated[15,16] that the magnetic and structural transitions rapidly split and both become second-order above $x \approx 0.022$. This is particularly evidenced by the fact that the peak in the derivative of the magnetic susceptibility χ (indicative of a first order magnetic transition) is strongly suppressed beyond this doping value[16]. The splitting of both transitions upon doping has also been observed in $Ca\left(Fe_{1-x}Co_x\right)_2 As_2$ at $x \approx 0.039$ by neutron diffraction data[17]. In these systems, the structural transition remains first-order at least for some doping range after the splitting, what is consistent with the phase diagram in Fig. 1(a) if we identify $x \approx 0.039$ with α_{c1}. The behavior of the isovalently doped compound $Ba\left(Fe_{1-x}Ru_x\right)_2 As_2$ is also consistent with our theory. In this material, the chemical potential does not change with doping[77], hence the structural and magnetic transitions should remain very close for all x. Thermodynamic measurements did indeed find that the two transitions do not split upon increasing x[78].

The interplay between structural and magnetic tran-

sitions in the hole-doped $(Ba_{1-x}K_x)Fe_2As_2$ compounds is not so well established, with conflicting reports of either simultaneous first-order transitions[79], or split first-order structural transition and second-order magnetic transition[80]. In the context of our model, hole doping adds an extra complication since at least some effects of doping are absorbed into the changes of an additional hole pocket at (π,π)[81]. This suggests that the chemical potential in our effective 4-band model changes at a slower rate than in the case of electron doping, since extra holes do not necessarily go to the central pockets. If this is the case, then a first-order structural transition extends to larger dopings, in agreement with the data.

Consider now the evolution of α with pressure. Band structure calculations show that pressure reduces the nesting features of the Fermi surface[82,83], what in our model implies that δm increases. According to Eq. (80), $\alpha \propto 1/(\delta m)^2$ then *decreases*, bringing the system deeper into the regime of simultaneous first-order structural and magnetic transitions (see Fig. 1). This agrees with the experimental results that structural and magnetic transitions do not split under pressure in $AeFe_2As_2$ (Ae = Ba, Co, Sr)[67,73].

We now briefly consider the 1111 materials, whose Fermi surfaces are also similar to the previously discussed NaFeAs compound. In these systems, the nematic and magnetic transitions are split and second-order already in the parent compounds. In our theory, these compounds should then be placed to the right of α_{c2} in Fig. 1(a) or to the right of α_{c1} in Fig. 1(b). The reason why α is larger in the 1111 materials is the significantly larger degree of out-of-plane anisotropy in the 1111 materials compared to the 122 materials, what brings the former closer to the $d = 2$ limit. In our modeling, these systems are then described by a smaller effective d. According to our theory, as d gets smaller, both tri-critical points $\alpha_{c1,2}$ shift towards smaller values (see Fig. 18 and Eq. (40)), extending the regime where the magnetic and structural transitions are split and second-order. A similar argument was given in Ref.[34].

Finally, for the iron chalcogenides $FeTe_{1-x}Se_x$, our model is applicable in the regime of intermediary Se doping, near the superconducting dome of the (x,T) phase diagram. In this region, ARPES measurements[84] reveal that the electronic structure is similar to the one considered in our model (see Fig. 2), and neutron scattering shows[85] that magnetic fluctuations are peaked at $\mathbf{Q}_1 = (\pi,0)$ and $\mathbf{Q}_2 = (0,\pi)$. On the other hand, our model is not suitable for the undoped FeTe sample, where the same ARPES data show the absence of electron pockets centered at $\mathbf{Q}_1$ and $\mathbf{Q}_2$[84].

B. Linear relation between the magnetic and nematic order parameters

Several experimental groups showed that in some 122 materials, most notably $SrFe_2As_2$, the magnetic and nematic order parameters have very similar temperature dependencies below the simultaneous first-order magneto-nematic transition[79,86,87]. In our analysis, the relationship between M and ϕ is given by Eq. (41) and can be readily seen in Fig. 9d for a simultaneous first-order transition. Re-plotting in Fig. 24 the nematic order parameter ϕ and the properly re-scaled magnetic order parameter M as functions of temperature, we see that both follow the same trend, indicating that the relationship between the two order parameters is nearly linear, in agreement with the experimental data. Even better linear relation is obtained for α closer to one, as we show in Fig. 25, where $\phi/\Delta\phi$ is plotted explicitly as function of $M/\Delta M$ ($\Delta\phi$ and ΔM denote the values of the jumps across the first-order magneto-nematic transition).

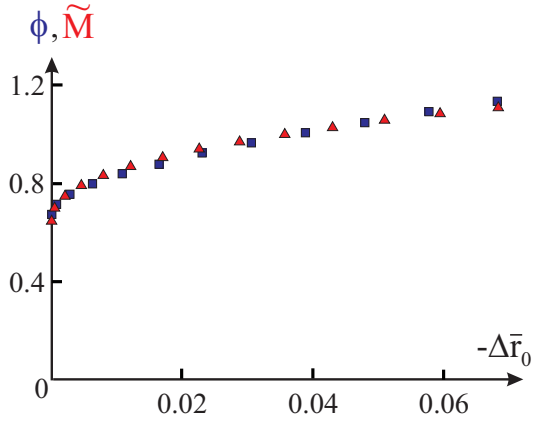

Figure 24: (Color online) Nematic order parameter ϕ (square/blue symbols) and re-scaled magnetic order parameter $\tilde{M}$ (triangles/red symbols) for $d = 2.5$ and $\alpha = 1.8$ as function of $-\Delta\bar{r}_0 = r_{0,cr} - r_0 \propto (T_s - T)$ (the same as in Fig. 9d). Here, $\tilde{M} = 0.87M + 0.33$.

At first glance, this near-linear relation seems non-trivial, because for split second-order magnetic and nematic transitions, a straightforward expansion leads to $\phi - \phi\,(m = 0) \propto M^2$. It can, however, be easily understood analytically by expanding Eq. (41) around the jumps $\Delta\phi$ and ΔM. We obtain

$$\frac{\phi - \Delta\phi}{\Delta\phi} = \sum_{j=1}^{\infty} a_j \left(\frac{M - \Delta M}{\Delta M}\right)^j \qquad (81)$$

For small $\phi - \Delta\phi$ and $M - \Delta M$, the relationship is indeed linear, since a_1 is non-zero for all $2 < d \leq 3$ (see Appendix D). In fact, in $SrFe_2As_2$, where the linear relation was experimentally observed[87], the measured temperature dependent orthorhombic and magnetic order parameters are rather small compared to the magnitude of the jumps. Interestingly, when $\alpha \to 1$ and the first-order transition gets stronger, the coefficients of Eq. (81) satisfy $a_1/a_2 = 2$ and $a_{j+2} = 0$ for any $2 < d \leq 3$, i.e., the dependence has only linear and quadratic terms. In the same limit, $\Delta\phi$ and ΔM become large, since both scale as $1/(\alpha - 1)$. Thus, a plot of ϕ versus M in absolute units

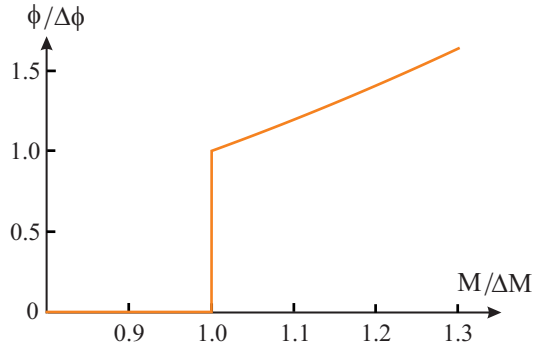

Figure 25: (Color online) Nematic order parameter ϕ as function of the magnetic order parameter M for $d = 2.5$ and $\alpha = 1.05$. Note the near linear behavior between them. Here $\Delta\phi$ and ΔM refer to the jumps at each transition.

would show a strictly linear dependence for $\alpha \to 1$.

C. $J_1 - J_2$ and phenomenological models

The possibility of an Ising-nematic order induced by magnetic fluctuations was first proposed for the iron pnictides within the localized-moment scenario[31–33], built upon earlier results by Chandra, Coleman, and Larkin[28]. In this localized-moment approach, one considers spins on a square lattice interacting via a nearest-neighbor antiferromagnetic exchange J_1 and a next-nearest neighbor antiferromagnetic exchange J_2. It was argued[31] that the hybridization between the Fe and As orbitals in the iron pnictides gives rise to a rather large J_2, which can exceed $J_1/2$. Once this happens, the magnetic ground state develops the stripe order with $O(3) \times Z_2$ order parameter manifold.

As shown in Ref.[28], one can interpret the stripe order as composed of two inter-penetrating Neel sublattices, with staggered magnetization $\mathbf{M}_1$ and $\mathbf{M}_2$. The configuration with $\mathbf{M}_1$ parallel (anti-parallel) to $\mathbf{M}_2$ corresponds to the $(\pi, 0)$ $((0, \pi))$ state. At the mean-field level, the two sublattices are uncoupled. Thermal and quantum fluctuations, however, induce a coupling between the two sublattices that favors the collinear configurations. Once this coupling is included, the effective action takes the form

$$\begin{aligned} S_{J_1-J_2}\left[\mathbf{M}_i\right] \;=\; & J_2 \int_q q^2 \left(\mathbf{M}_{1,q} \cdot \mathbf{M}_{1,-q} + \mathbf{M}_{2,q} \cdot \mathbf{M}_{2,-q}\right) \\ & + J_1 \int_q q_x q_y \left(\mathbf{M}_{1,q} \cdot \mathbf{M}_{2,-q}\right) \\ & - \zeta \frac{J_1^2}{J_2} \int_x \left(\mathbf{M}_1 \cdot \mathbf{M}_2\right)^2 \end{aligned} \tag{82}$$

where the magnitudes of $\mathbf{M}_1$ and $\mathbf{M}_2$ are fixed, and ζ is a dimensionless constant which is non-zero because of thermal and/or quantum fluctuations. This constant is small in a $1/S$ expansion, where S denotes spin, and remains small numerically even for $S = 1/2$ (see, for instance[88]).

It is straightforward to make connection between our itinerant model, Eq. (7), and the $J_1 - J_2$ model. First, the relationship between the real-space order parameters $\mathbf{M}_1$ and $\mathbf{M}_2$ and the momentum-space order parameters $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ is $\mathbf{\Delta}_X = \mathbf{M}_1 + \mathbf{M}_2$ and $\mathbf{\Delta}_Y = \mathbf{M}_1 - \mathbf{M}_2$. The scalar product $\mathbf{M}_1 \cdot \mathbf{M}_2$ is then the same as $\mathbf{\Delta}_X^2 - \mathbf{\Delta}_Y^2$ and the nematic coupling is given by $g = \zeta J_1^2/J_2$. Second, in the itinerant approach, the hard constraint $\mathbf{M}_1^2 = \mathbf{M}_2^2 = 1$ is replaced by the quartic terms, which play the role of soft constraints. We see that the itinerant and the $J_1 - J_2$ models are indeed quite similar, and in both models the Ising-nematic order results from the Z_2 degeneracy of the stripe magnetic ground state.

Whether the two models have identical phase diagrams and show the same behavior upon doping, pressure, and alkaline-earth substitution is a more subtle issue. In the $J_1 - J_2$ model, the analog of α is $J_2^2/(\zeta J_1^2)$. Taken at face value, this quantity is large (because ζ is small) and depends only weakly on doping and pressure. For instance, the parameters used in Ref.[32] yield $\alpha \sim 100$, which places both undoped and doped iron pnictides in the regime of split second-order transitions, i.e. region III of the phase diagrams of Figs. 8 and 14**.** On the other hand, in the itinerant model, α is generally of order one and changes with doping and pressure due to the changes in the Fermi surface.

It is also unclear (chiefly due to the lack of results) whether the phase diagram of the $J_1 - J_2$ model contains an intermediate phase in which one transition is first-order and the other is second-order, as in Fig. 1. The mean-field analysis of Eq. (82) for a quasi two-dimensional system did find split second-order nematic and magnetic transitions[32] for small ζ and a simultaneous first-order transition[57] for large enough ζ. A similar result was obtained by Monte Carlo simulations of Eq. (82) in Refs.[89,90], which treated ζ as a phenomenological input parameter. It was also shown by mean-field[91], RG[34], and Monte Carlo[92] calculations that in three dimensional systems with anisotropic magnetic dispersion the degree of anisotropy tunes the system between the regimes of split second-order transitions and simultaneous first-order transitions. Making the spin interaction anisotropic in spin space has the same effect, i.e. it gives rise to a transformation from split second-order transitions to simultaneous first-order transitions[93]. However, in all cases, the intermediate phase either was not discussed, or was assumed to be absent[34]. If only the regimes of split second-order transitions and simultaneous first-order transitions occur in the $J_1 - J_2$ model, it will be difficult to explain the different characters of the transitions observed experimentally upon doping, pressure, or alkaline-earth substitution. Also, to the best of our knowledge, there have been no studies within the $J_1 - J_2$ model of whether the magnetic correlation length jumps at the nematic transition. As we showed in Section V C, such a jump (or a sharp increase), obtained

with the itinerant approach, is fundamental to explain the pseudogap behavior above T_N.

Several research groups, using $J_1 - J_2$ based models, also put forward various arguments to relate the transformation from simultaneous to split nematic and magnetic transitions to doping and pressure[34,94,95]. Ref.[34] suggested that the change from simultaneous first-order to split second-order transitions with doping is due to doping-induced change of the spin dynamics from Landau-overdamped ($z_{dyn} = 2$) to propagating ($z_{dyn} = 1$). However, ARPES measurements later showed that hot spots - and therefore Landau damping - are present even at optimal doping[96]. The authors of Ref.[95] assumed that doping increases the disorder concentration and decreases the coupling to the lattice, what leads to a decrease in g. However, recent data on isovalent doped pnictides[77,78] show that, when the chemical potential remains unchanged, the character of the transitions remain the same as in the parent compound, even for large doping concentrations and large in-plane disorder. Ref.[94] suggested that the phase diagram of the iron pnictides is close to a magnetic quantum tri-critical point, but did not analyze the character of the structural transition.

There have also been studies of nematic and magnetic transitions within the itinerant approach to the iron-pnictides. Ref.[15] considered a model very similar to the one presented here and argued that the phase diagram of Fig. 1(b) can be obtained even for quasi-two dimensional systems if there is a strong enough coupling to anharmonic elastic terms.

A phase diagram similar to our Fig. 1(b), with the intermediate regime, was also obtained in Ref.[56]. These authors assumed that the structural and magnetic transitions occur independent of each other, and that the two order parameters are linearly coupled, as in Eq. (78). The intermediate regime then emerges in some range of parameters, primarily due to the presence of a quartic term in the elastic free-energy. The authors of[56], however, did not argue why the intrinsic magnetic and structural transitions would occur at about the same temperature.

One argument of why this may be the case, without invoking nematic degrees of freedom, was presented in Ref.[97]. The authors of[97] argued that, under special conditions (which may or may not be satisfied in the iron pnictides), a SDW instability and a CDW instability in the form of orbital currents occur at almost the same temperature, with the orbital-current instability occurring first. They further argued that orbital-current order and magnetic order are orthogonal to each other in momentum space and coexist.

As we already mentioned, some elements of the physics that leads to our phase diagrams in Fig. 1 are similar to the physics of the $J_1 - J_2$ and phenomenological models. The key elements that distinguish our study from previous works are (i) that the whole phase diagram is entirely driven by magnetic degrees of freedom, and (ii) that the parameters of the effective bosonic model are derived from the original itinerant model of interacting fermions, such that the evolution of the system behavior with doping, pressure, and alkaline-earth substitution is fully described within the model itself.

Finally, we point out that an RG analysis similar to the one presented in Section III was performed in Ref.[34] for $N = 3$ components of the magnetic order parameter, and in Ref.[57] for $N = 6$, which refers to the case of incommensurate magnetic order parameter. Interestingly, although both works obtained runaway flows, the former associated the first-order instability to the divergence of the coupling constants, while the latter pointed out that it happens much earlier, when the condition $u_L < g_L$ is first satisfied. Our reasoning is similar to the one in Ref.[57].

VII. CONCLUDING REMARKS

We argued in this work that the development of the preemptive nematic order and its interplay with the stripe magnetic order can be fully understood within an itinerant magnetic scenario for the iron pnictides. We obtained (rather than assumed) the dependence of the nematic coupling upon doping, pressure, and alkaline-earth substitution, what enabled us to consistently explain the character of the magneto-structural transitions for a variety of iron pnictides. We also demonstrated how the nematic transition induces orbital and structural order, and triggers the pseudogap behavior observed in the paramagnetic phase of these materials.

The authors thank E. Abrahams, J. Analytis, E. Bascones, J. van den Brink, P. Brydon, S. Bud'ko, P. Canfield, P. Chandra, M. Daghofer, L. Degiorgi, I. Fisher, A. Goldman, A. Kaminski, A. Kreyssig, W. Ku, J. Lorenzana, S. Maiti, R. McQueeney, Y. Matsuda, A. Millis, R. Moessner, R. Prozorov, M. Takigawa, M. Tanatar, Z. Tesanovic, and M. Vavilov for useful discussions. R. M. F. acknowledges the support from ICAM and NSF-DMR 0645461, as well as the valuable support from the NSF Partnerships for International Research and Education (PIRE) program. A.V.C. acknowledges the support from NSF-DMR 0906953. R. M. F. and A.V.C. thank the hospitality of the Aspen Center for Physics, where part of this work has been done, and the support from NSF grant number 1066293. I. E. and J.K. acknowledges the financial support from DAAD (PPP USA No. 50750339).

Appendix A: Microscopic calculation of the Ginzburg-Landau coefficients

In this Appendix, we show explicitly how to calculate the Ginzburg-Landau coefficients of Eq. (7) starting with the interacting Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm int}$ of Eqs. (1)

and (2). First, we introduce the 6-dimensional creation Nambu operator:

$$\Psi_{\mathbf{k}}^{\dagger} = \left(\begin{array}{cccccc} c_{\Gamma,\mathbf{k}\uparrow}^{\dagger} & c_{\Gamma,\mathbf{k}\downarrow}^{\dagger} & c_{X,\mathbf{k}\uparrow}^{\dagger} & c_{X,\mathbf{k}\downarrow}^{\dagger} & c_{Y,\mathbf{k}\uparrow}^{\dagger} & c_{Y,\mathbf{k}\downarrow}^{\dagger} \end{array} \right) \quad \text{(A1)}$$

After introducing the bosonic fields $\mathbf{\Delta}_X$ and $\mathbf{\Delta}_Y$ via the Hubbard-Stratonovich transformation and evaluating the products of Pauli matrices, we can write the partition function as:

$$Z = \int d\Delta_i d\Psi \mathrm{e}^{-S[\Psi,\Delta_i]} \quad \text{(A2)}$$

with the action written in compact form:

$$S\left[\Psi,\Delta_i\right] = -\int_k \Psi_k^{\dagger}\mathcal{G}_k^{-1}\Psi_k + \frac{2}{u_{\mathrm{spin}}}\int_x \left(\Delta_X^2+\Delta_Y^2\right) \quad \text{(A3)}$$

Here, $\Delta_i = |\mathbf{\Delta}_i|$ and the Green's function $\mathcal{G}_k^{-1}$ is given by:

$$\mathcal{G}_k^{-1} = \mathcal{G}_{0,k}^{-1} - \mathcal{V} \quad \text{(A4)}$$

with the bare term:

$$\mathcal{G}_{0,k} = \left(\begin{array}{ccc} \hat{G}_{\Gamma,k} & 0 & 0 \\ 0 & \hat{G}_{X,k} & 0 \\ 0 & 0 & \hat{G}_{Y,k} \end{array} \right) \quad \text{(A5)}$$

and the interacting term:

$$\mathcal{V} = \left(\begin{array}{ccc} 0 & -\hat{\Delta}_X & -\hat{\Delta}_Y \\ -\hat{\Delta}_X & 0 & 0 \\ -\hat{\Delta}_Y & 0 & 0 \end{array} \right) \quad \text{(A6)}$$

Here, we defined the 2×2 matrices $\hat{G}_{i,k} = G_{i,k}\mathbb{I}$ and $\hat{\Delta}_i = \mathbf{\Delta}_i\cdot\boldsymbol{\sigma}$, with identity matrix $\mathbb{I}$, Pauli matrices σ^j, and $G_{i,k}^{-1} = i\omega_n - \xi_{i,\mathbf{k}}$ the non-interacting single-particle Green's functions.

It is now straightforward to integrate out the fermions, since the action is quadratic in them, and obtain the effective magnetic action:

$$S_{\mathrm{eff}}\left[\mathbf{\Delta}_X,\mathbf{\Delta}_Y\right] = -\mathrm{Tr}\ln\left(1-\mathcal{G}_{0,k}\mathcal{V}\right) + \frac{2}{u_{\mathrm{spin}}}\int_x \left(\Delta_X^2+\Delta_Y^2\right) \quad \text{(A7)}$$

Here, $\mathrm{Tr}\left(\cdots\right)$ refers to sum over momentum, frequency and Nambu indices. A series expansion in powers of Δ_i^2 then gives:

$$\begin{aligned} S_{\mathrm{eff}}\left[\mathbf{\Delta}_X,\mathbf{\Delta}_Y\right] &= \frac{1}{2}\mathrm{Tr}\left(\mathcal{G}_{0,k}\mathcal{V}\right)^2 + \frac{1}{4}\mathrm{Tr}\left(\mathcal{G}_{0,k}\mathcal{V}\right)^4 \\ &\quad \frac{2}{u_{\mathrm{spin}}}\int_x \left(\Delta_X^2+\Delta_Y^2\right) \end{aligned} \quad \text{(A8)}$$

Evaluation of the traces yields:

$$S_{\mathrm{eff}}\left[\mathbf{\Delta}_X,\mathbf{\Delta}_Y\right] = \sum_i r_{0,i}\Delta_i^2 + \sum_{i,j} u_{ij}\Delta_i^2\Delta_j^2 \quad \text{(A9)}$$

with the coefficients:

$$\begin{aligned} r_{0,i} &= \frac{2}{u_{\mathrm{spin}}} + 2\int_k G_{\Gamma,k}G_{i,k} \\ u_{ij} &= \int_k G_{\Gamma,k}^2 G_{i,k}G_{j,k} \end{aligned} \quad \text{(A10)}$$

Due to the $\pi/2$ rotation symmetry relating the X and Y bands, it follows that $r_{0,1} = r_{0,2} \equiv r_0$ and $u_{11} = u_{22}$. After re-arranging the terms, we obtain:

$$\begin{aligned} S_{\mathrm{eff}}\left[\mathbf{\Delta}_X,\mathbf{\Delta}_Y\right] &= r_0\left(\Delta_X^2+\Delta_Y^2\right) \\ &\quad + \left(\frac{u_{12}+u_{11}}{2}\right)\left(\Delta_X^2+\Delta_Y^2\right)^2 \\ &\quad - \left(\frac{u_{12}-u_{11}}{2}\right)\left(\Delta_X^2-\Delta_Y^2\right)^2 \end{aligned} \quad \text{(A11)}$$

with the coefficients $u = u_{12}+u_{11}$ and $g = u_{12}-u_{11}$ given in Eq. (8).

Appendix B: Effects of the anisotropic momentum dispersion

In this Appendix we briefly show that the inclusion of an anisotropic momentum dispersion in the bare dynamic susceptibilities $\chi_{i,q}$ do not change our main results. After denoting the renormalized susceptibilities by $\tilde{\chi}_{X,q}^{-1} = \chi_{X,q}^{-1} + \psi - \phi$ and $\tilde{\chi}_{X,q}^{-1} = \chi_{Y,q}^{-1} + \psi + \phi$, with $\chi_{i,q}^{-1} = r_0 + f_{i,\mathbf{q}} + \gamma\left|\nu_n\right|$, we can rewrite the mean-field equations (17) as:

$$\begin{aligned} \psi &= \frac{u}{2}\int_q \left(\tilde{\chi}_{X,q}+\tilde{\chi}_{Y,q}\right) \\ \phi &= \frac{g}{2}\int_q \left(\tilde{\chi}_{X,q}-\tilde{\chi}_{Y,q}\right) \end{aligned} \quad \text{(B1)}$$

In the main text, we considered the case of an isotropic momentum dispersion $f_{i,\mathbf{q}} = q^2$. Most generally, $f_{i,\mathbf{q}}$ will have an anisotropic form preserving the tetragonal symmetry of the system:

$$f_{i,\mathbf{q}} = q_x^2\left(1\pm\eta\right) + q_y^2\left(1\mp\eta\right) \quad \text{(B2)}$$

where $-1<\eta<1$ and the upper (lower) sign refers to band X (Y). Indeed, inelastic neutron scattering measurements find this form for the dynamic susceptibility in several iron pnictide compounds[98,99].

Substituting it in the self-consistent equations (B1), we can re-scale the momentum by $\tilde{q}_x = q_x\sqrt{1\pm\eta}$ and $\tilde{q}_y = q_y\sqrt{1\mp\eta}$ depending on whether the integral involves $\tilde{\chi}_{X,q}$ (upper signs) or $\tilde{\chi}_{Y,q}$ (lower signs). In either case, the Jacobian of the transformation is the same, yielding:

$$\begin{aligned}\psi &= \frac{u}{\sqrt{1-\eta^2}}\int_{\tilde{q}}\frac{r_0+\psi+\tilde{q}^2+\gamma|\nu_m|}{\left(r_0+\psi+\tilde{q}^2+\gamma|\nu_m|\right)^2-\phi^2}\\ \phi &= \frac{g}{\sqrt{1-\eta^2}}\int_{\tilde{q}}\frac{\phi}{\left(r_0+\psi+\tilde{q}^2+\gamma|\nu_m|\right)^2-\phi^2}\end{aligned} \quad \text{(B3)}$$

Thus, comparing the previous equations with the original mean-field expressions (17), we conclude that the only effect of the anisotropic dispersion is to renormalize the coupling constants u and g in the same way, yielding:

$$\begin{aligned}\bar{\bar{u}} &= \frac{\bar{u}}{\sqrt{1-\eta^2}}\\ \bar{\bar{g}} &= \frac{\bar{g}}{\sqrt{1-\eta^2}}\end{aligned} \quad \text{(B4)}$$

This does not change the value of the ratio $\alpha = u/g = \bar{\bar{u}}/\bar{\bar{g}}$, implying that the phase diagrams discussed in the main text remain valid, with the same values for the tri-critical points α_{c1} and α_{c2}. The only modification will be in the absolute value of the temperature, since $\bar{r}_0$ is proportional to g (see Eq. 38).

Appendix C: Anisotropic 3D model with quadratic dispersion

In this Appendix we consider the behavior of the magnetic and nematic tri-critical points in the anisotropic 3D model with the bosonic susceptibility

$$\chi_{i,q} = r_0 + q_{||}^2 + \beta^2 q_z^2 \quad \text{(C1)}$$

where $0 \le \beta \le 1$ and the same momentum cutoff Λ is taken for all three momentum components. This model naturally interpolates between the 2D case ($\beta = 0$) and the isotropic 3D case ($\beta = 1$). We show that the behavior of α_{c1} and α_{c2} as function of β is very similar, although not identical, to the behavior of the two tri-critical points as functions on η_z displayed in Fig.13.

The self-consistent equations for ϕ and r become:

$$\begin{aligned} r &= r_0 + \bar{u}\int_0^{\Lambda_{||}}\frac{dq_{||}}{2\pi}q_{||}\int_0^{\Lambda_z}dq_z\times\\ &\left(\frac{1}{r+q_{||}^2+\beta^2q_z^2-\phi}+\frac{1}{r+q_{||}^2+\beta^2q_z^2+\phi}\right)\\ \phi &= \bar{g}\int_0^{\Lambda_{||}}\frac{dq_{||}}{2\pi}q_{||}\int_0^{\Lambda_z}dq_z\times\\ &\left(\frac{1}{r+q_{||}^2+\beta^2q_z^2-\phi}-\frac{1}{r+q_{||}^2+\beta^2q_z^2+\phi}\right)\end{aligned} \quad \text{(C2)}$$

For simplicity, we set $\bar{g} = 1$ below, but the results can be easily generalized for arbitrary $\bar{g}$ after rescaling $\beta^2 \to \beta^2/\bar{g}$. We first do the 2D integral over $q_{||}$ and then evaluate the one-dimensional integral in the q_z direction. We define

$$I(c) = \int_0^{\Lambda}\frac{dq_{||}}{2\pi}q_{||}\int_0^{\Lambda}dq_z\frac{1}{q_{||}^2+\beta^2q_z^2+c} \quad \text{(C3)}$$

with $c = r \pm \phi$. Evaluating the momentum integrals we obtain

$$\begin{aligned}4\pi I(c) &= \frac{2\sqrt{\Lambda^2+c}}{\beta}\arctan\left(\frac{\beta\Lambda}{\sqrt{\Lambda^2+c}}\right)\\ &+\Lambda\log\left(\Lambda^2+\beta^2\Lambda^2+c\right)-\frac{2\sqrt{c}}{\beta}\arctan\left(\frac{\beta\Lambda}{\sqrt{c}}\right)\\ &-\Lambda\log\left(\beta^2\Lambda^2+c\right)\end{aligned} \quad \text{(C4)}$$

As before, we assume that Λ is large compared to both r and ϕ. Then the first line does not depend on c and can be absorbed into the renormalizion of r_0, i.e. into $\bar{r}_0$. The 2D and 3D results are indeed reproduced: in the 2D limit $\beta \to 0$, and we reproduce the logarithmic behavior and Eqs. (22); in the 3D limit $\beta \to 1$ and $\Lambda \to \infty$ such that $\arctan \to \pi/2$ and we reproduce the characteristic square root behavior as well as Eqs. 27.

A simple analysis shows that the crossover from 2D to 3D behavior occurs at a rather small $\beta^2 \sim 1/\Lambda$. If α_{c1} and α_{c2} cross, they must cross in this regime. We rescale β^2 by Λ ($\beta = \frac{\tilde{\beta}}{\sqrt{\Lambda}}$) and also rescale r and ϕ by Λ ($r = \Lambda\tilde{r}$, $\phi = \Lambda\tilde{\phi}$), obtaining a cut-off independent equation in terms of $\tilde{\beta}$

$$\frac{4\pi I(\tilde{c})}{\Lambda} = -\frac{2\sqrt{\tilde{c}}}{\tilde{\beta}}\arctan\left(\frac{\tilde{\beta}}{\sqrt{\tilde{c}}}\right)-\log\left(\tilde{\beta}^2+\tilde{c}\right) \quad \text{(C5)}$$

The integral can be plugged into Eqs.(C2) to obtain the self-consistent equations for $\tilde{r}$ and $\tilde{\phi}$. In order to obtain α_{c2} corresponding to the nematic tri-critical point, we expand the self-consistent equations to second order in $\tilde{\phi}$ and look for the value of α when the coefficient of the ϕ^2-term changes sign. A straightforward analysis yields:

$$\alpha_{c2} = \frac{\tilde{\beta}\sqrt{\tilde{r}_{c2}}\left[\tilde{\beta}\sqrt{\tilde{r}_{c2}}\left(3\tilde{\beta}^2+5\tilde{r}_{c2}\right)+3\left(\tilde{\beta}^2+\tilde{r}_{c2}\right)^2\arctan\left(\frac{\tilde{\beta}}{\sqrt{\tilde{r}_{c2}}}\right)\right]}{\tilde{\beta}^2\tilde{r}_{c2}\left(6-3\tilde{\beta}^2-5\tilde{r}_{c2}\right)+3\left(\tilde{\beta}^2+\tilde{r}_{c2}\right)\arctan\left(\frac{\tilde{\beta}}{\sqrt{\tilde{r}_{c2}}}\right)\left[\tilde{\beta}\sqrt{\tilde{r}_{c2}}\left(4-\tilde{\beta}^2-\tilde{r}_{c2}\right)+2\left(\tilde{\beta}^2+\tilde{r}_{c2}\right)\arctan\left(\frac{\tilde{\beta}}{\sqrt{\tilde{r}_{c2}}}\right)\right]} \tag{C6}$$

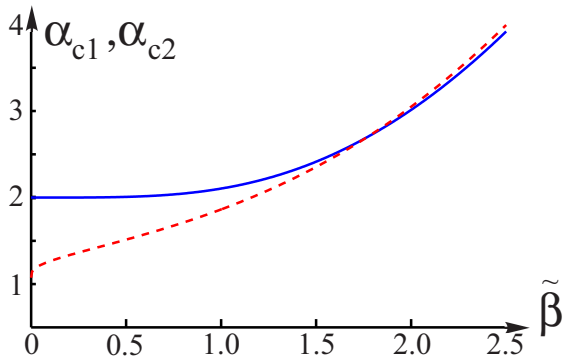

Figure 26: (Color online) The behavior of the magnetic and nematic tri-critical points α_{c1} (dashed red/light gray line) and α_{c2} (solid blue/dark gray line) as functions of the rescaled $\tilde{\beta}$ in the 3D anisotropic model with the bosonic susceptibility given by Eq. (48). Similar to the other anisotropic 3D model, α_{c1} and α_{c2} cross at $\tilde{\beta} \approx 1.7$, and at larger $\tilde{\beta}$, α_{c2} becomes smaller than α_{c1}.

where $\tilde{r}_{c2}$ is the critical value for the onset of Ising-nematic order:

$$1 = \frac{2\arctan\left(\frac{\tilde{\beta}}{\sqrt{\tilde{r}_{c2}}}\right)}{\tilde{\beta}\sqrt{\tilde{r}_{c2}}} \tag{C7}$$

To calculate α_{c1} corresponding to the magnetic tri-critical point, we extend the self-consistent equations to include M in the standard manner and set $r = \phi$. The equation for the critical $\tilde{\phi}_{c1}$ at which $M = 0$ is

$$\tilde{\phi}_{c1} = \frac{2\sqrt{2}\arctan\left(\frac{\tilde{\beta}}{\sqrt{2\tilde{\phi}_{c1}}}\right)}{\tilde{\beta}} - \log\left(\tilde{\beta}^2\right) + \log\left(\tilde{\beta}^2+2\tilde{\phi}_{c1}\right). \tag{C8}$$

Expanding the self-consistent equations in powers of M, it is straightforward to obtain the value of $\alpha = \alpha_{c1}$ for which the coefficient of the quadratic term M^2 changes sign:

$$\alpha_{c1} = \frac{\tilde{\beta}}{\tilde{\beta} - 2\sqrt{2}\arctan\left(\frac{\tilde{\beta}}{\sqrt{2\tilde{\phi}_{c1}}}\right)/\sqrt{\tilde{\phi}_{c1}}} \tag{C9}$$

We plot α_{c1} and α_{c2} as function of $\tilde{\beta}$ in Fig. 26. We see that the two tri-critical points cross at a certain $\tilde{\beta}$, beyond which α_{c2} becomes smaller than α_{c1}. As expected, as $\tilde{\beta}$ increases, both α_{c1} and α_{c2} increase, and their ratio approaches 1 from below. The difference between α_{c1} and α_{c2}, however, stays finite.

We also solved numerically the self-consistent equations (C2) for r and ϕ and obtained the phase diagram in the variables $\bar{r}_0$ and α for various β. We found the same four phase diagrams as in Fig. 12. Namely, for small $\tilde{\beta}$ we recover the behavior of Fig. 12 (a), while for $\tilde{\beta} \gg 1$ (but with $\beta \leq 1$), we recover the behavior of Fig. 12 (d). The two phase diagrams at intermediate $\tilde{\beta}$ are also the same as those in Figs. 12 (b), (c).

Appendix D: Linear relationship between the nematic and magnetic order parameters

In this Appendix, we show how near a simultaneous first-order magneto-structural transition, the nematic and magnetic order parameters obey an approximately linear relationship. We start with the self-consistent equations (41) in the magnetically ordered phase for arbitrary dimension $2 < d \leq 3$. A straightforward manipulation leads to the equivalent equations:

$$\begin{aligned} \bar{r}_0 &= \frac{\bar{u}}{4}(2\phi)^{\frac{d-2}{2}} - \phi(\alpha-1) \\ \phi &= \frac{\bar{g}}{4}(2\phi)^{\frac{d-2}{2}} + \bar{g}\bar{M}^2 \end{aligned} \tag{D1}$$

with $\phi = r$. The condition $d\bar{r}_0/d\phi = 0$ gives the value of $\Delta\phi$ for which $\bar{r}_0$ is maximum and, therefore, the first instability of the system. Substitution of $\Delta\phi$ in the second equation then gives $\Delta\bar{M}$. Evaluating the algebraic equations yields:

$$\begin{aligned} \Delta\phi &= \frac{(g/2)^{\frac{2}{4-d}}}{2}\left[\frac{\alpha(d-2)}{\alpha-1}\right]^{\frac{2}{4-d}} \\ \Delta\bar{M} &= 2(g/2)^{\frac{d-2}{2(4-d)}}\left[\frac{\alpha(d-2)}{\alpha-1}\right]^{\frac{d-2}{2(4-d)}}\sqrt{\frac{1-\alpha(3-d)}{\alpha-1}} \end{aligned} \tag{D2}$$

These are the values of the nematic and magnetic jumps for $\alpha < \alpha_{c1} = 1/(d-3)$. Notice that $\Delta\bar{M} \to 0$ for $\alpha \to \alpha_{c1}$ and that the $d = 3$ result of Eq. (34) is reobtained from the second equation.

Using the second equation of (D1), we can expand it for ϕ close to $\Delta\phi$ and $\bar{M}$ close to $\Delta\bar{M}$, obtaining:

$$\frac{\phi-\Delta\phi}{\Delta\phi} = \frac{4[1-\alpha(3-d)]}{(d-2)(\alpha+1)} \times$$

$$\left[\left(\frac{\bar{M}-\Delta\bar{M}}{\Delta\bar{M}}\right)+\frac{a_2}{a_1}\left(\frac{\bar{M}-\Delta\bar{M}}{\Delta\bar{M}}\right)^2\right] + (\alpha-1)\,\mathcal{O}\left[\left(\frac{M-\Delta M}{\Delta M}\right)^3\right] \tag{D3}$$

with:

$$\frac{a_2}{a_1}=\frac{(6-d)\,(1-2\alpha\,(3-d))+\alpha^2\left(22-13d+2d^2\right)}{2\,(d-2)\,(\alpha+1)^2} \tag{D4}$$

Thus, the linear relationship between ϕ and M is always present for small enough deviations from the jump. For $\alpha \to 1$, this linear relationship dominates and extends to larger values of N. Indeed, as $\alpha \to 1$, the jumps $\Delta\phi$ and $\Delta\bar{M}$ of Eq. (D2) become larger and, consequently, $\left(\frac{\bar{M}-\Delta\bar{M}}{\Delta M}\right)$ becomes smaller for a fixed $\bar{M}$. Furthermore, all the coefficients of the series expansion (D3) of order higher than quadratic go to zero, and the ratio a_1/a_2 (D4) between the linear and the quadratic coefficients tends to 2 for any dimension.

[1] K. Ishida, Y. Nakai and H. Hosono, J. Phys. Soc. Japan **78**, 062001 (2009); J. Paglione and R.L. Greene, Nature Phys. **6**, 645 (2010); D. C. Johnston, Adv. Phys. **59**, 803 (2010).

[2] T. Yildirim, Phys. Rev. Lett. **101**, 057010 (2008).

[3] F. Ma, Z.-Y. Lu, and T. Xiang, Phys. Rev. B **78**, 224517 (2008).

[4] I. R. Fisher, L. Degiorgi, and Z. X. Shen, arXiv:1106.1675

[5] J.-H. Chu, J. G. Analytis, K. De Greve, P. L. McMahon, Z. Islam, Y. Yamamoto, and I. R. Fisher, Science **329**, 824 (2010).

[6] M. A. Tanatar, E. C. Blomberg, A. Kreyssig, M. G. Kim, N. Ni, A. Thaler, S. L. Bud'ko, P. C. Canfield, A. I. Goldman, I. I. Mazin, and R. Prozorov, Phys. Rev. B **81**, 184508 (2010).

[7] A. Dusza, A. Lucarelli, F. Pfuner, J.-H. Chu, I. R. Fisher, and L. Degiorgi, EPL **93**, 37002 (2011); A. Lucarelli, A. Dusza, A. Sanna, S. Massidda, J.-H. Chu, I.R. Fisher, and L. Degiorgi, arXiv:1107.0670

[8] M. Nakajima, T. Liang, S. Ishida, Y. Tomioka, K. Kihou, C. H. Lee, A. Iyo, H. Eisaki, T. Kakeshita, T. Ito, and S. Uchida, PNAS **108**, 12238 (2011).

[9] T.-M. Chuang, M. P. Allan, J. Lee, Y. Xie, N. Ni, S. L. Bud'ko, G. S. Boebinger, P. C. Canfield, and J. C. Davis, Science **327**, 181 (2010).

[10] M. Yi, D. Lu, J.-H Chu, J. G. Analytis, A. P. Sorini, A. F. Kemper, B. Moritz, S.-K. Mo, R. G. Moore, M. Hashimoto, W.-S. Lee, Z. Hussain, T. P. Devereaux, I. R. Fisher, and Z.-X. Shen, PNAS **108**, 6878 (2011).

[11] Y. Matsuda *et al.*, in preparation.

[12] C.-L. Song, Y.-L. Wang, P. Cheng, Y.-P. Jiang, W. Li, T. Zhang, Z. Li, K. He, L. Wang, J.-F. Jia, H.-H. Hung, C. Wu, X. Ma, X. Chen, and Q.-K. Xue, Science **332**, 1410 (2011).

[13] R. M. Fernandes, D. K. Pratt, W. Tian, J. Zarestky, A. Kreyssig, S. Nandi, M. G. Kim, A. Thaler, N. Ni, P. C. Canfield, R. J. McQueeney, J. Schmalian, and A. I. Goldman, Phys. Rev. B **81**, 140501(R) (2010).

[14] S. Nandi, M. G. Kim, A. Kreyssig, R. M. Fernandes, D. K. Pratt, A. Thaler, N. Ni, S. L. Bud'ko, P. C. Canfield, J. Schmalian, R. J. McQueeney, and A. I. Goldman, Phys. Rev. Lett. **104**, 057006 (2010).

[15] M. G. Kim, R. M. Fernandes, A. Kreyssig, J. W. Kim, A. Thaler, S. L. Bud'ko, P. C. Canfield, R. J. McQueeney, J. Schmalian, and A. I. Goldman, Phys. Rev. B **83**, 134522 (2011).

[16] C. R. Rotundu and R. J. Birgeneau, arXiv:1106.5761

[17] K. Prokeš, S. Mat'aš, L. Harnagea, S. Singh, S. Wurmehl, D. N. Argyriou, and B. Büchner, Phys. Rev. B **83**, 104414 (2011).

[18] F. Krüger, S. Kumar, J. Zaanen, J. van den Brink, Phys. Rev. B **79**, 054504 (2009).

[19] R. R. P. Singh, arXiv:0903.4408

[20] W. Lv, F. Krüger, and P. Phillips, Phys. Rev. B **82**, 045125 (2010); W. Lv and P. Phillips, arXiv:1105.4630

[21] W. G. Yin, C. C. Lee, and W. Ku, Phys. Rev. Lett. **105**, 107004 (2010).

[22] M. J. Calderon, B. Valenzuela, and E. Bascones, Phys. Rev. B **80**, 094531 (2009); E. Bascones, M. J. Calderon, and B. Valenzuela, Phys. Rev. Lett. **104**, 227201 (2010); M. J. Calderon, G. Leon, B. Valenzuela, and E. Bascones, arXiv:1107.2279

[23] P. M. R. Brydon, M. Daghofer, and C. Timm, J. Phys.: Condens. Matter **23**, 246001 (2011); A. Nicholson, Q. Luo, W. Ge, J. Riera, M. Daghofer, G. B. Martins, A. Moreo, and E. Dagotto, Phys. Rev. B **84**, 094519 (2011).

[24] C.-C. Chen, J. Maciejko, A. P. Sorini, B. Moritz, R. R. P. Singh, and T. P. Devereaux, Phys. Rev. B **82**, 100504 (2010).

[25] M. S. Laad and L. Craco, arXiv:1010.2940

[26] D.-Y. Liu, Y.-M. Quan, D.-M. Chen, L.-J. Zou1, and H.-Q. Lin, Phys. Rev. B **84**, 064435 (2011).

[27] A. H. Nevidomskyy, arXiv:1104.1747

[28] P. Chandra, P. Coleman, and A.I. Larkin, Phys. Rev. Lett. **64**, 88 (1990).

[29] S. Graser, T. A. Maier, P. J. Hirschfeld, and D. J. Scalapino, New J. Phys. **11**, 025016 (2009).

[30] J. Zhang, R. Sknepnek, R. M. Fernandes, and J. Schmalian, Phys. Rev. B **79**, 220502(R) (2009).

[31] Q. Si and E. Abrahams, Phys. Rev. Lett. **101**, 076401 (2008); E. Abrahams and Q. Si, J. Phys.: Condens. Matter **23**, 223201 (2011).

[32] C. Fang, H. Yao, W.-F. Tsai, J. Hu, and S. A. Kivelson, Phys. Rev. B **77,** 224509 (2008).

[33] C. Xu, M. Müller, and S. Sachdev, Phys. Rev. B **78**, 020501(R) (2008).

[34] Y. Qi and C. Xu, Phys. Rev. B **80**, 094402 (2009).

[35] R. M. Fernandes, L. H. VanBebber, S. Bhattacharya, P. Chandra, V. Keppens, D. Mandrus, M. A. McGuire, B. C. Sales, A. S. Sefat, and J. Schmalian, Phys. Rev. Lett. **105**, 157003 (2010).

[36] C. He, Y. Zhang, B. P. Xie, X. F. Wang, L. X. Yang, B. Zhou, F. Chen, M. Arita, K. Shimada, H. Namatame, M. Taniguchi, X. H. Chen, J. P. Hu, and D. L. Feng, Phys. Rev. Lett. **105**, 117002 (2010).

[37] Y.-M. Xu, P. Richard, K. Nakayama, T. Kawahara, Y. Sekiba, T. Qian, M. Neupane, S. Souma, T. Sato, T. Takahashi, H.-Q. Luo, H.-H. Wen, G.-F. Chen, N.-L. Wang, Z. Wang, Z. Fang, X. Dai, and H. Ding, Nature Comm. **2**, 392 (2011).

[38] M. A. Tanatar, N. Ni, A. Thaler, S. L. Bud'ko, P. C. Canfield, and R. Prozorov, Phys. Rev. B **82**, 134528 (2010).

[39] D. N. Basov *et al.*, in preparation.

[40] R. M. Fernandes, E. Abrahams, and J. Schmalian, Phys. Rev. Lett. **107**, 217002 (2011).

[41] I. Eremin and A. V. Chubukov, Phys. Rev. B **81**, 024511 (2010).

[42] A. B. Vorontsov, M. G. Vavilov, and A. V. Chubukov, Phys. Rev. B **81**, 174538 (2010).

[43] S. Maiti and A. V. Chubukov, Phys. Rev. B **82**, 214515 (2010).

[44] A. V. Chubukov, D. V. Efremov, and I. Eremin. Phys. Rev. B, **78**, 134512 (2008).

[45] V. Cvetkovic and Z. Tesanovic, Phys. Rev. B **80**, 024512 (2009).

[46] V. Barzykin and L. P. Gor'kov, JETP Lett. **88**, 131 (2008).

[47] P. M. R. Brydon and C. Timm, Phys. Rev. B **79**, 180504(R) (2009); B. Zocher, C. Timm, and P. M. R. Brydon, arXiv:1106.5315

[48] K. Kuroki, S. Onari, R. Arita, H. Usui, Y. Tanaka, H. Kontani, and H. Aoki, Phys. Rev. Lett. **101**, 087004 (2008).

[49] F. Wang, H. Zhai, Y. Ran, A. Vishwanath, and D.-H. Lee,

Phys. Rev. Lett. **102**, 047005 (2009).
[50] C. Platt, C. Honerkamp, and W. Hanke, New J. Phys. **11**, 055058 (2009).
[51] A. B. Vorontsov, M. G. Vavilov, and A. V. Chubukov, Phys. Rev. B **79**, 060508(R) (2009).
[52] R. M. Fernandes and J. Schmalian, Phys. Rev. B **82**, 014521 (2010).
[53] J. Knolle, I. Eremin, A.V. Chubukov, and R. Moessner, Phys. Rev. B **81**, 140506(R) (2010).
[54] P. M. R. Brydon, J. Schmiedt, and C. Timm, arXiv:1109.2071
[55] V. Cvetkovic and Z. Tesanovic, EPL **85**, 37002 (2009).
[56] A. Cano, M. Civelli, I. Eremin, and I. Paul, Phys. Rev. B **82**, 020408(R) (2010).
[57] A. J. Millis, Phys. Rev. B **81**, 035117 (2010).
[58] S. Ma, *Modern Theory of Critical Phenomena,* Westview Press (2000).
[59] E. C. Blomberg, M. A. Tanatar, A. Kreyssig, N. Ni, A. Thaler, Rongwei Hu, S. L. Bud'ko, P. C. Canfield, A. I. Goldman, and R. Prozorov, Phys. Rev. B **83**, 134505 (2011).
[60] H.-H. Kuo, J.-H. Chu, S. C. Riggs, L. Yu, P. L. McMahon, K. De Greve, Y. Yamamoto, J. G. Analytis, and I. R. Fisher, Phys. Rev. B **84**, 054540 (2011).
[61] I. Paul, Phys. Rev. Lett. **107**, 047004 (2011).
[62] K. Kusakabe and A. Nakanishi, J. Phys. Soc. Jpn. **78**, 124712 (2009).
[63] M. Vilk and A.-M. S. Tremblay, J. Phys. I France **7**, 1309-1368 (1997); J. Schmalian, D. Pines, and B. Stojković, Phys. Rev. Lett. **80**, 3839 (1998); Phys. Rev. B **60**, 667 (1999); E.Z. Kuchinskii and M. V. Sadovskii, JETP **88**, 968 (1999); T. Sedrakyan and A.V. Chubukov, Phys. Rev B **81**, 174536 (2010).
[64] Y. Nakai, K. Ishida, Y. Kamihara, M. Hirano, and H. Hosono, J. Phys. Soc. Jpn. **77**, 073701 (2008).
[65] K. Kitagawa, Y. Mezaki, K. Matsubayashi, Y. Uwatoko, and M. Takigawa, J. Phys. Soc. Jpn. **80**, 033705 (2011).
[66] L. Ma, G. F. Chen, D.-X. Yao, J. Zhang, S. Zhang, T.-L. Xia, and W. Yu, Phys. Rev. B **83**, 132501 (2011).
[67] E. Colombier, S. L. Bud'ko, N. Ni, and P. C. Canfield, Phys. Rev. B **79**, 224518 (2009).
[68] M. S. Torikachvili, S. L. Bud'ko, N. Ni, and P. C. Canfield, Phys. Rev. Lett. **101**, 057006 (2008).
[69] N. Ni, S. Nandi, A. Kreyssig, A. I. Goldman, E. D. Mun, S. L. Bud'ko, and P. C. Canfield, Phys. Rev. B **78**, 014523 (2008).
[70] C. R. Rotundu, B. Freelon, T. R. Forrest, S. D. Wilson, P. N. Valdivia, G. Pinuellas, A. Kim, J.-W. Kim, Z. Islam, E. Bourret-Courchesne, N. E. Phillips, and R. J. Birgeneau, Phys. Rev. B **82**, 144525 (2010).
[71] W. Uhoya, A. Stemshorn, G. Tsoi, Y. K. Vohra, A. S. Sefat, B. C. Sales, K. M. Hope, and S. T. Weir, Phys. Rev. B **82**, 144118 (2010).
[72] W. O. Uhoya, J. M. Montgomery, G. M. Tsoi, Y. K. Vohra, M. A. McGuire, A. S. Sefat, B. C. Sales, and S. T. Weir , J. Phys.: Condens. Matter **23**, 122201 (2011).
[73] A. I. Goldman, A. Kreyssig, K. Prokes, D. K. Pratt, D. N. Argyriou, J. W. Lynn, S. Nandi, S. A. J. Kimber, Y. Chen, Y. B. Lee, G. Samolyuk, J. B. Leao, S. J. Poulton, S. L. Bud'ko, N. Ni, P. C. Canfield, B. N. Harmon, and R. J. McQueeney, Phys. Rev. B **79**, 024513 (2009).
[74] P. C. Canfield, private communication.
[75] K. Matan, R. Morinaga, K. Iida, and T. J. Sato, Phys. Rev. B **79**, 054526 (2009).
[76] P. C. Canfield, S. L. Bud'ko, Ni Ni, J. Q. Yan, and A. Kracher, Phys. Rev. B **80**, 060501(R) (2009); N. Ni, A. Thaler, A. Kracher, J. Q. Yan, S. L. Bud'ko, and P. C. Canfield, Phys. Rev. B **80**, 024511 (2009); P. C. Canfield and S. L. Bud'ko, Annu. Rev. Condens. Matter Phys. **1**, 27 (2010).
[77] R. S. Dhaka, C. Liu, R. M. Fernandes, R. Jiang, C. P. Strehlow, T. Kondo, A. Thaler, J. Schmalian, S. L. Bud'ko, P. C. Canfield, and A. Kaminski, Phys. Rev. Lett. **107**, 267002 (2011).
[78] A. Thaler, N. Ni, A. Kracher, J. Q. Yan, S. L. Bud'ko, and P. C. Canfield, Phys. Rev. B **82**, 014534 (2010).
[79] S. Avci, O. Chmaissem, E. A. Goremychkin, S. Rosenkranz, J.-P. Castellan, D.-Y. Chung, I. S. Todorov, J. A. Schlueter, H. Claus, M. G. Kanatzidis, A. Daoud-Aladine, D. Khalyavin, and R. Osborn, Phys. Rev. B **83**, 172503 (2011).
[80] R. R. Urbano, E. L. Green, W. G. Moulton, A. P. Reyes, P. L. Kuhns, E. M. Bittar, C. Adriano, T. M. Garitezi, L. Bufaiçal, and P. G. Pagliuso, Phys. Rev. Lett. **105**, 107001 (2010).
[81] A. F. Kemper, T. A. Maier, S. Graser, H.-P. Cheng, P. J. Hirschfeld, and D. J. Scalapino, New J. Phys. **12**, 073030 (2010).
[82] S. A. J. Kimber, A. Kreyssig, Y.-Z. Zhang, H. O. Jeschke, R. Valentí, F. Yokaichiya, E. Colombier, J. Yan, T. C. Hansen, T. Chatterji, R. J. McQueeney, P. C. Canfield, A. I. Goldman, and D. N. Argyriou, Nature Mater. **8**, 471 (2009).
[83] Y.-Z. Zhang, H. C. Kandpal, I. Opahle, H. O. Jeschke, and R. Valentí, Phys. Rev. B **80**, 094530 (2009).
[84] Z. X. Shen, private communication.
[85] T. J. Liu, J. Hu, B. Qian, D. Fobes, Z. Q. Mao, W. Bao, M. Reehuis, S. A. J. Kimber, K. Prokeš, S. Matas, D. N. Argyriou, A. Hiess, A. Rotaru, H. Pham, L. Spinu, Y. Qiu, V. Thampy, A. T. Savici, J. A. Rodriguez, and C. Broholm, Nature Materials **9**, 718 (2010)
[86] S. D. Wilson, C. R. Rotundu, Z. Y., P. N. Valdivia, B. Freelon., E. Bourret-Courchesne, and R. J. Birgeneau, Phys. Rev. B **81**, 014501 (2010).
[87] A. Jesche, N. Caroca-Canales, H. Rosner, H. Borrmann, A. Ormeci, D. Kasinathan, H. H. Klauss, H. Luetkens, R. Khasanov, A. Amato, A. Hoser, K. Kaneko, C. Krellner, and C. Geibel, Phys. Rev. B **78**, 180504(R) (2008).
[88] R. R. P. Singh, W. Zheng, J. Oitmaa, O. P. Sushkov, and C. J. Hamer, Phys. Rev. Lett. **91**, 017201 (2003).
[89] A. L. Wysocki, K. D. Belashchenko, and V. P. Antropov, Nat. Phys. **7**, 485 (2011).
[90] J. Hu, B. Xu, W. Liu, N. Hao, and Y. Wang, arXiv:1106.5169
[91] C. Fang, J. Hu, S. Kivelson, and S. Brown, Phys. Rev. B **74**, 094508 (2006).
[92] Y. Kamiya, N. Kawashima, and C. D. Batista, arXiv:1108.1599
[93] R. Applegate, R. R. P. Singh, C.-C. Chen, and T. P. Devereaux, arXiv:1110.0434
[94] G. Giovannetti, C. Ortix, M. Marsman, M. Capone, J. van den Brink, and J. Lorenzana, Nature Comm. **2**, 398 (2011).
[95] M. Capati, M. Grilli, and J. Lorenzana, arXiv:1106.3238
[96] C. Liu, T. Kondo, R. M. Fernandes, A. D. Palczewski, E. D. Mun, N. Ni, A. N. Thaler, A. Bostwick, E. Rotenberg, J. Schmalian, S. L. Bud'ko, P. C. Canfield, and A. Kaminski, Nature Phys. **6**, 419 (2010).
[97] J. Kang and Z. Tesanovic, Phys. Rev. B **83**, 020505 (2011).

[98] S. O. Diallo, D. K. Pratt, R.M. Fernandes, W. Tian, J. L. Zarestky, M. Lumsden, T. G. Perring, C. L. Broholm, N. Ni, S. L. Bud'ko, P. C. Canfield, H.-F. Li, D. Vaknin, A. Kreyssig, A .I. Goldman, and R. J. McQueeney, Phys. Rev. B **81**, 214407 (2010).

[99] H.-F. Li, C. Broholm, D. Vaknin, R. M. Fernandes, D. L. Abernathy, M. B. Stone, D. K. Pratt, W. Tian, Y. Qiu, N. Ni, S. O. Diallo, J. L. Zarestky, S. L. Bud'ko, P. C. Canfield, and R. J. McQueeney, Phys. Rev. B **82**, 140503(R) (2010).