Ising-nematic order in the bilinear-biquadratic model for the iron pnictides

Patricia Bilbao Ergueta and Andriy H. Nevidomskyy
Department of Physics and Astronomy, Rice University, Houston, TX 77005

Motivated by the recent inelastic neutron scattering (INS) measurements in the iron pnictides which show a strong anisotropy of spin excitations in directions perpendicular and parallel to the ordering wave-vector even above the magnetic transition temperature $T_N$, we study the frustrated Heisenberg model with a biquadratic spin-spin exchange interaction. Using the Dyson-Maleev (DM) representation, which proves appropriate for all temperature regimes, we find that the spin-spin dynamical structure factors are in excellent agreement with experiment, exhibiting breaking of the $C_4$ symmetry even into the paramagnetic region $T_N < T < T_s$ which we refer to as the Ising-nematic phase. In addition to the Heisenberg spin interaction, we include the biquadratic coupling $K(S_i \cdot S_j)^2$ and study its effect on the dynamical temperature range $T_o - T_N$ of the Ising-nematic phase. We find that this range reduces dramatically when even small values of the interlayer exchange $J_c$ and biquadratic coupling $K$ are included. To supplement our analysis, we benchmark the results obtained using the DM method against those from different non-linear spin-wave theories, including the recently developed generalized spin-wave theory (GSWT), and find good qualitative agreement among the different theoretical approaches as well as experiment for both the spin-wave dispersions and the dynamical structure factors.

PACS numbers: 75.10.Jm, 74.70.Xa, 75.40.Gb

I. INTRODUCTION

The recent discovery of iron-based high-$T_c$ superconductor\cite{liu95} and the antiferromagnetically ordered nature of their parent compounds\cite{yoshida05} sheds more light on the importance of understanding the electronic correlations and the magnetic excitations in these materials, especially due to the similarities between their phase diagram to that of the cuprates. Like the cuprates, the parent compounds of the iron pnictides exhibit an antiferromagnetic ground state below the Néel ordering temperature $T_N$. However instead of the regular Néel phase, pnictides order into a columnar antiferromagnet (CAF) with the ordering wave-vector $Q = (\pi,0)$ or $(0,\pi)$\cite{martirosyan08}. The transition to the long-range magnetic order in the pnictides takes place in a very close proximity to a structural transition from tetragonal to orthorhombic phase below temperature $T_s \geq T_N$. Therefore, it is not a priori clear whether the electronic degrees of freedom or purely elastic lattice response play the primary role.

Recent resistivity $\rho$ measurements under fixed strain $\delta = (a-b)/(a+b)$ ($a$ and $b$ are the in-plane lattice constants) have detected divergent nematic susceptibility $d\rho/d\delta$, proving that the nematicity is of electronic origin rather than due to an elastic lattice instability\cite{wollersheim09}.

Two different mechanisms have been proposed for the electronic nematic ordering: the spin-nematic scenario\cite{sun07} and the orbital ordering with unequal population of iron $d_{xz}$ and $d_{yz}$ orbitals\cite{baoguo07,sun09}. Because the corresponding order parameters are coupled linearly in the Landau free energy, the appearance of one will cause a non-zero expectation value of the other, and disentangling which is the cause and which is the consequence is very difficult\cite{sun07}. In this work, we will not attempt to answer this question. Rather, we are interested in the physical signatures of electronic nematic order in the electron spin response.

In particular, we are motivated by the recent inelastic neutron scattering (INS) data on detwinned BaFe$_2$As$_2$ that exhibit a two-fold anisotropy in the spin excitations even above the structural transition temperature $T_s$ in the nominally tetragonal phase\cite{yamashita08}.

To address this problem theoretically, we chose an approach in terms of quasi-local moments on Fe sites, following earlier works by other authors\cite{fujita08,baoguo09,li09,xue09,lv09,lv10}. This strong coupling perspective is motivated by the “bad metal” nature of the parent compounds and the proposed proximity to the Mott localization transition\cite{li09,lv09,lv10}. Indeed, superconductivity was found to border a Mott insulating phase in alkaline iron selenides $A_1-x$Fe$_2$Se$_2$ (the “245” family, with $A = K$, Rb, Cs, or Tl)\cite{fujita08,baoguo09,li09,xue09,lv09,lv10}. The Mott insulating ground state has also been identified in the iron oxycanalogues La$_2$O$_3$Fe$_2$(Se,S)$_2$ $\text{O}$\cite{fujita08,baoguo09,li09,xue09,lv09,lv10}, $R_2$O$_3$Fe$_2$Se$_2$ (here $R = \text{Ce,Pr, Nd, Sm}$)\cite{fujita08,baoguo09,li09,xue09,lv09,lv10}, and Sr$_2$Fe$_2$O$_5$\cite{li09}. Further evidence in favor of proximity to the incipient Mott phase comes from the suppression of the Drude peak in optical conductivity measurements on iron pnictides\cite{yu10} and the spectral weight transfer induced by temperature\cite{sun10}. Theoretically, it was proposed that Hund’s coupling plays a crucial role in enhancing strong electron correlations in the iron pnictides\cite{sun07}, due to the fact that it decouples charge fluctuations in individual orbitals, leading to the orbital-dependent mass renormalization\cite{sun07}. These predictions have been confirmed experimentally in e.g. KFe$_2$As$_2$\cite{sun09}. As the Coulomb repulsion strength grows, this has been suggested to eventually result in an orbital-selective Mott transition\cite{sun09} observed in alkaline iron selenides\cite{sun10}.

While some aspects of magnetism can be understood from a weak coupling approach of itinerant electrons with the Fermi surface nesting\cite{sun07}, the aforementioned studies justify the use of an effective model of localized spins $S_i$ on iron sites with nearest $(<i,j>)$ and next-nearest
neighbor ($\ll i,k \gg$) interactions:

$$
H_{SD} = J_1 \sum_{\langle i,j \rangle} S_i \cdot S_j - K \sum_{\langle i,j \rangle} (S_i \cdot S_j)^2 + J_2 \sum_{\langle i,k \rangle} S_i \cdot S_k + J_c \sum_{i} S_i \cdot S_{i+\hat{z}}
$$

(1)

The size of the effective spin $S$ is dictated by the strength of the Hund’s coupling relative to the Fe $d$-electron bandwidth and crystal-field splittings. In the parent compounds of iron pnictides, $S = 1$ agrees with the integrated spin spectral weight of $\sim 3\mu_B$ per Fe from INS measurements.\textsuperscript{52} In the iron chalcogenides the electron bandwidth is narrower, so that even though the absolute value of the Hund’s coupling is similar to that in the iron pnictides ($J_H \sim 0.7$ eV), its role is more pronounced, resulting in a larger spin $S = 2$.

In addition to the Heisenberg spin interaction, the effective Hamiltonian\textsuperscript{1} also contains the biquadratic spin-spin interactions. The latter are important to correctly capture the dispersion of the spin excitations near the Brillouin zone boundary,\textsuperscript{53,54} observed via the INS.\textsuperscript{21,22} Indeed, without the biquadratic $K$-term, one is forced to adopt an anisotropic nearest neighbor (NN) coupling constant $J_{\perp} \neq J_{||}$ even when modeling the neutron spectra above the structural transition,\textsuperscript{22} which is unphysical. Moreover, such analysis would predict $J_1$ to be wildly different in two crystallographic directions, with antiferromagnetic $J_{\perp}$ and ferromagnetic $J_{||}$ which are impossible to reconcile even when the small ($\delta \lesssim 1\%$) orthorhombic lattice distortion is taken into account. Instead, the inclusion of a biquadratic term in Eq. (1) dynamically generates the anisotropy in the effective NN Heisenberg couplings, in agreement with the experimental results.\textsuperscript{48–50}

In this work, we show that the spin wave dispersion and dynamical spin structure factor from INS measurements\textsuperscript{16,52} can be modeled semi-quantitatively using the effective spin model Eq. (1). Moreover, we find that there is a temperature range $T_N < T < T_N$ in the paramagnetic phase with nematic anisotropy of the spin excitations, similar to the recent INS data on BaFe$_2$As$_2$.\textsuperscript{16} On a technical level, this work improves significantly upon the earlier work by one of the co-authors,\textsuperscript{52} who used a simple mean-field decoupling of the biquadratic spin-spin interaction in the spirit of a linear spin-wave approximation. By contrast, here we employ several more exacting methods to treat the model Eq. (1), including a non-linear spin-wave theory, Dyson–Maleev spin representation,\textsuperscript{55,56} and the recently developed so-called generalized spin-wave theory (GSWT).\textsuperscript{57} Last but not the least, we demonstrate that the biquadratic spin interaction $-K(S_i \cdot S_j)^2$ can be derived from a theory of coupled spin and orbital degrees of freedom by integrating out the latter, thereby establishing a connection between the orbital mechanism of nematicity and the spin response.

The paper is organized as follows. First, we provide the aforementioned derivation of the effective spin Hamiltonian in section \textsuperscript{[D]} Then, in section \textsuperscript{[E]} we summarize our main results and compare them to experiment. We then go on to show the spin-wave dispersions obtained with the different methods in section \textsuperscript{[K]} In section \textsuperscript{[L]} we include plots for the low temperature dynamical structure factors and comment on their agreement with experimental results. We analyze the evolution of both the staggered magnetization and the nematic order parameter with temperature in detail in section \textsuperscript{[M]} and finally summarize our conclusions in section \textsuperscript{[N]}. The details of the different methods used to calculate spin waves are given in the Appendices for convenience.

II. EFFECTIVE SPIN MODEL

In the original Refs.\textsuperscript{48–50}, the biquadratic spin-spin interaction in Eq. (1) was introduced heuristically as a higher order spin exchange process, derived for instance as a fourth order perturbation in the Schrieffer–Wolff projection of the Hubbard model.\textsuperscript{61} In this section we show that alternatively, the microscopic origin of the $K$-term can be traced to orbital ordering within the theory of coupled spin and orbital degrees of freedom. The advantage of such an interpretation is that it allows to incorporate the orbital physics into the spin response, since ultimately both orbital and spin degrees of freedom are involved in the electronic nematic phase.\textsuperscript{15,61}

To see this explicitly, let us consider as a starting point the Kugel–Khomskii model\textsuperscript{62} formulated for the $d_{xz}$ and $d_{yz}$ orbitals of the iron pnictides following Refs.\textsuperscript{10,11,13} and \textsuperscript{62}.

$$
H = J \sum_i \left[ (S_i \cdot S_{i+\hat{x}} + 1) \tau_i^a \tau_{i+\hat{x}}^a + (S_i \cdot S_{i+\hat{y}} + 1) \tau_i^b \tau_{i+\hat{y}}^b \right]
$$

(2)

where $S_i = 1$ is the magnetic moment originating from the Hund’s coupled electron spins in Fe $d_{xz}$ and $d_{yz}$ orbitals, and $\{\tau^a, \tau^b\}$ are the pseudospin operators that act in the orbital subspace of $|xz\rangle$ and $|yz\rangle$ states and depend on the directionality of the Fe-Fe bond. This model is complicated to deal with, but for our purposes, a simple mean-field decoupling will suffice, denoting $\Gamma_x \equiv \langle S_i \cdot S_{i+\hat{x}} \rangle$ and $\Gamma_y \equiv \langle S_i \cdot S_{i+\hat{y}} \rangle$ for brevity:

$$
E_{MF} = J \left[ (\Gamma_x + 1) \langle \tau_i^a \tau_{i+\hat{x}}^a \rangle + (\Gamma_y + 1) \langle \tau_i^b \tau_{i+\hat{y}}^b \rangle \right]
$$

(3)

To proceed, we relate the average of the orbital pseudospin operators to the occupation number of the corresponding orbital: $\langle \tau^a \rangle = n_{xz}$ and $\langle \tau^b \rangle = n_{yz}$. This is similar to earlier studies in Refs.\textsuperscript{11,13} and \textsuperscript{62} who treated $\tau^a$ as Ising degrees of freedom. We point out that $n_{xz}$ and $n_{yz}$ orbitals are occupied by one electron each in the tetragonal phase of the parent compound, so that total occupancy $n = n_{xz} + n_{yz} = 2$. Then, the correlators
\langle \tau_i \tau_{i+\gamma} \rangle \) can be written as follows:

\[
\langle \tau_i^a \tau_{i+\gamma}^a \rangle = \left( n_{xz} + \frac{\delta n}{2} \right)^2 = \frac{n + P + \delta n}{2}
\]

\[
\langle \tau_i^b \tau_{i+\gamma}^b \rangle = \left( n_{yz} + \frac{\delta n}{2} \right)^2 = \frac{n - P + \delta n}{2}
\]

where we have introduced the orbital polarization \( P = n_{x} - n_{yz} \) and \( \delta n \) is a phenomenological parameter that accounts for orbital fluctuations. Indeed, \( \delta n \) is a measure of the variance in the orbital occupation number \( \langle \tau_i^a \tau_{i+\gamma}^a \rangle \) = \( n_{x} + O(\delta n^2) \) and similarly for the \( yz \)-orbital.

Minimizing the mean-field energy Eq. (3) with respect to \( P \), we find the expectation value of the orbital polarization

\[
P = \frac{n + \delta n}{2 + \Gamma_x + \Gamma_y} (\Gamma_y - \Gamma_x) \approx (\Gamma_y - \Gamma_x)
\]

where the last equality is obtained by noting that \( \Gamma_x = -\Gamma_y \) in the columnar antiferromagnetic phase and \( n + \delta n \approx 2 \). We see that the orbital polarization is linearly proportional to the Ising-nematic order parameter \( \langle \tau_i^a \tau_{i+\gamma}^a \rangle \) as it should be based on the Landau theory of coupled order parameters, since the bilinear combination \( P \cdot (\Gamma_y - \Gamma_x) \) is allowed by symmetry. Using Eqs. (4) and (5), we can now decouple the orbital degrees of freedom in the Kugel–Khomskii model at the mean-field level, resulting in an effective spin Hamiltonian to linear order in \( P \):

\[
H_{\text{MF}} \sim \left( \frac{n + \delta n}{2} \right)^2 J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - \left( \frac{n + \delta n}{2} \right)^2 J \sum_{\langle i,j \rangle} (\Gamma_y - \Gamma_x)(\mathbf{S}_i \cdot \mathbf{S}_{i+y} - \mathbf{S}_i \cdot \mathbf{S}_{i+x})
\]

If we now forget that this mean-field Hamiltonian originated from orbital physics, it is tempting to interpret it as a mean-field approximation to the following spin Hamiltonian:

\[
H_{\text{spin}} = J_1 \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - \frac{J_1}{2} \sum_{\langle i,j \rangle} (\mathbf{S}_i \cdot \mathbf{S}_j)^2 + J_1 \sum_i \mathbf{S}_i \cdot \mathbf{S}_{i+x} \]

with \( J_1 = J(n + \delta n)^2/4 \). The first two terms of this effective Hamiltonian are the same as in the Eq. (1), with the biquadratic term appearing naturally as a result of integrating out the orbital degrees of freedom. The last term in Eq. (7) involves three spin interaction and similar to the \( J_0 \) Heisenberg term, favors columnar antiferromagnetic order with \( \langle S_i \cdot S_{i+x} \rangle = -\langle S_i \cdot S_{i+y} \rangle \). Note also that within this derivation, one obtains a value of \( K = J_1/2 \), explaining the relatively large value of \( K \approx 0.6J_1 \) necessary to fit the spin wave dispersion from inelastic neutron scattering. We note that the possibility of generating a significantly large coupling \( |K| \) provided the system has quasi-degenerate orbitals has also been pointed out by Mila and Zhang who used higher order perturbation theory to derive the biquadratic spin exchange from the Hubbard model. For completeness, we mention that the biquadratic spin-spin interaction can also be the result of the magnetoelastic coupling provided the lattice has suitable phonon modes, as was proposed for Fe\(_{1+y}\)Te in Ref. [53].

### III. MAIN RESULTS

At low enough temperatures \( T < T_N \) the iron pnictides exhibit a columnar antiferromagnetic spin-stripe ground state with two degenerate in energy wave-vectors Q = (π, 0) or (0, π). Within this region, the staggered magnetization \( m_s \) has a finite value. The magnetic transition temperature is always equal or lower than the structural transition temperature \( T_s \) and the INS experiment have found anisotropy of the spin-wave dispersions above \( T_N \). Therefore, there is a finite range of temperatures \( T_N < T < T_s \) with a nematic anisotropy in the spin response, prompting the researchers call this a spin Ising-nematic phase following early ideas of spon-
taneous $Z_2$ symmetry breaking in the frustrated $J_1 - J_2$ model on a square lattice.\cite{takahashi2010}

Our theoretical calculations based on the effective spin Hamiltonian Eq. (1) confirm the above picture. Namely, we find a paramagnetic nematic temperature region $T_N < T < T_\sigma$ where the staggered magnetization vanishes but the correlations along the $x$-direction and $y$-direction are different. Therefore, we take the quantity

$$\Gamma_y - \Gamma_x \equiv \langle \mathbf{S}_x \cdot \mathbf{S}_{x+y} \rangle - \langle \mathbf{S}_x \cdot \mathbf{S}_{x+y} \rangle$$

(8)
as a measure of electron spin nematicity, which is plotted as a function of temperature in Fig. 1.

The nematic phase is most pronounced in two spatial dimensions ($J_c = 0$), where the true long-range magnetic order cannot exist by virtue of the Mermin–Wagner theorem\cite{mermin1966} whereas the discrete $Z_2$ symmetry can still be broken. However, we find that inclusion of a very small interplanar coupling $J_\sigma$ is sufficient to make $T_N$ approach $T_\sigma$, and the dynamic temperature range $T_\sigma - T_N$ shrinks rapidly as a function of $J_c$, as shown in Fig. 2a). In the absence of a biquadratic spin coupling, this result has already been anticipated in Ref.\cite{lee2010} using a large-$N$ approach and in Ref.\cite{lee2011} using Dyson–Maleev large-$S$ spin-wave theory (see Appendix\cite{appendix} for more details of the method).

Here, we are interested in how the biquadratic spin coupling $-K(\mathbf{S}_i \cdot \mathbf{S}_j)^2$ affects the above result. What we found is that the dynamic temperature range of the nematic phase shrinks considerably upon including even a small biquadratic term $K = 0.03 J_2$, see Fig. 2b). In fact, $T_\sigma$ practically coincides with $T_N$ for sufficiently large $J_c/J_2 > 0.05$. We conclude that considerable fine-tuning is needed in order to achieve a purely Ising-nematic phase within the model in Eq. (1) and only a very narrow nematic region is observed, which completely disappears for appreciable values of $K$ such as $K \sim 0.6 J_1$ required to fit the INS data.\cite{lee2011} This indicates that while the effective spin model Eq. (1) is very successful in modeling the spin-wave dispersions (Section IV) and INS spin structure factor (Section IV), an effectively single orbital spin physics may be insufficient to explain the considerable dynamical range $(T_\sigma - T_N)$ observed experimentally in the iron pnictides. Even more striking, the apparent absence of a long-range magnetic order in stoichiometric FeS,\cite{70,71} while $T_\sigma \sim 90 \text{K}$ remains large, indicates a failure of a pure spin approach and highlights the importance of multi-orbital physics that likely plays an important role in FeSe and in the iron pnictides. This conclusion is corroborated by the recent angle-resolved photoemission spectroscopy (ARPES)\cite{72,73} and nuclear magnetic resonance (NMR) studies\cite{74,75} on FeSe.

We now proceed to address the issue of recent INS measurements on BaFe$_{2-x}$Ni$_x$As$_2$ which show that the low-energy spin-wave excitations break the $C_4$ symmetry and remain anisotropic even above the structural transition temperature $T_s$, in the nominally tetragonal phase.\cite{70,71} We have calculated the dynamical structure factor $S(q, \omega)$ in the paramagnetic phase using the Dyson–Maleev spin representation\cite{76,77} and we indeed find the $C_2$ anisotropy of the intensity in the temperature range $T_N < T < T_\sigma$, similar to the experimental results (see Fig. 3). It has to be pointed out however that normally, one would associate the temperature $T_\sigma$ with the structural transition temperature $T_s$ because it marks the breaking of the $C_4$ lattice symmetry. Therefore, it is puzzling that the experimentally observed anisotropy persists above $T_s$.\cite{70,71}
It has to be remembered however that the INS experiment in Ref. [16] was done on detwinned crystals, i.e. in the presence of a non-zero uniaxial strain which itself breaks the $C_4$ lattice symmetry. Then, the notion of a spontaneous symmetry breaking no longer applies and instead becomes a crossover. Given the very large nematic susceptibility near $T_s$, as inferred from the resistivity measurements, a natural explanation of the neutron scattering data would be that the applied strain triggers a nematic response whose tail is seen at elevated temperatures $T > T_s$. The spin (and coupled orbital) fluctuations in the nematic channel thus contribute to the observed signal.

**IV. SPIN-WAVE DISPERSIONS FOR THE MAGNETIC GROUND-STATE**

The theory of the Dyson–Maleev bosons (sometimes also referred to as *modified* spin-wave theory) has already succeeded in giving a good qualitative picture of experimental data in various studies [49,50,66]. However, we now put our method of choice to further test by benchmarking its results against those obtained with several other spin-wave theories (see Appendices for more details). For this purpose, we choose to focus on the spin-wave dispersions in the magnetically ordered ground state, and throughout this and next section, we shall be using the two-dimensional version of our model Hamiltonian, with $J_c = 0$. In addition to giving us more insight into the validity of the methods typically used for these studies,
this analysis serves two other purposes.

First, we note that Stanek et al. in Ref. [49] have found several discrepancies between the results of the Schwinger...
bosl (SB) representation and those obtained with our method of choice (DM), with the latter approach producing more accurate results when a quadratic spin-spin coupling $K \neq 0$ is included in Eq. (1). Here, we would like to investigate whether this accuracy of the SB method pertains to any of the other spin-wave theories, as well as to try to get some insight into its origin. Second, this comparison offers an ideal opportunity to put the recently developed generalized spin-wave theory [58] to the test.

We start with the study of non-linear spin-wave (NLSW) theories, which use a semi-classical approach to the Holstein-Primakoff spin representation by expanding into powers of the small parameter $1/S$ (see Appendix A for more details). However, this procedure does not always guarantee the mandatory existence of the Goldstone modes in the long-range ordered magnetic ground state, so we only include the orders of expansion at which these are conserved, namely up to order $O(S^1)$ (low order) and up to order $O(S^0)$ (high order). Full decoupling of the biquadratic spin term turns out to also destroy the Goldstone modes, so we stick instead to the Hubbard-Stratonovich decoupling to treat the terms of higher order in the spins:

$$\langle S_r \cdot S_{r'} \rangle^2 \approx 2 \langle S_r \cdot S_{r'} \rangle S_r \cdot S_{r'} - \langle S_r \cdot S_{r'} \rangle^2 \quad (9)$$

before introducing the following order parameters of interest:

$$n = \langle a_r^{\dagger} a_r \rangle$$
$$g_x = \langle a_r a_{r+\hat{x}} \rangle = \langle a_r^{\dagger} a_{r+\hat{x}}^{\dagger} \rangle$$
$$f_y = \langle a_r^{\dagger} a_{r+\hat{y}} \rangle = \langle a_r a_{r+\hat{y}} \rangle$$
$$g_{xy} = \langle a_r a_{r+\hat{x}\pm\hat{y}} \rangle = \langle a_r^{\dagger} a_{r+\hat{x}\pm\hat{y}}^{\dagger} \rangle \quad (10)$$

As usual, $n$ represents the on-site average number of bosons that decrease the value of the sublattice magnetization $m_x$ from the classical value of $m_{\text{class}} = S$. The remaining three order parameters correspond to the correlations between neighboring spins in the $\hat{x}$, $\hat{y}$ and $\hat{x}\pm\hat{y}$ directions. It is worth noting that both $g_x$ and $g_{xy}$ are anomalous averages in the sense that the operators involved do not conserve the particle number. This is, however, an artifact of the method of choice, which requires a $\pi$-rotation around, say, $S_z$ direction for one of the sublattices so that the classical spin orientation now effectively become ferromagnetic instead of antiferromagnetic one. Finally, the averages not included must vanish in order for the total $z$-component of the spin $S_{\text{tot}}^z = \sum_i S_i^z$ to be conserved.

Once the Hamiltonian is decoupled and diagonalized, we are left with typical Bogoliubov dispersions (see the Appendix A for more details). As mentioned earlier, we restrict our analysis up to the order $O(S^0)$ in the expansion. Since the classical limit of these theories corresponds to the case where $S \to \infty$, we expect that going up to $O(S^0)$ will give a more reliable description than keeping only the terms up to order $O(S)$. Indeed, if we compare the obtained spin-wave dispersions in the magnetic ground state (see Fig.[8] we observe that, although both orders give similar results for the purely bilinear case ($K = 0$), the discrepancies increase dramatically with the value of the biquadratic coupling, especially around the zone boundary or $M = (\pi, \pi)$ point. More specifically, the region around the $M$ point remains a local high for the spin-wave dispersion in all cases for the higher order results, while it progressively becomes a local maximum for higher values of $K$ in the lower order $O(S^1)$ approximation. This difference becomes less pronounced as we increase the value of spin (see Fig.[5] for the case of $S = 2$), which is readily understood since quantum fluctuations are diminished as $S$ grows and becomes more classical. For reference, we include the spin-wave dispersions that we obtain from a simple linear spin-wave theory (LSWT) by keeping only the quadratic terms in the Hamiltonian, without performing any decoupling. The low order non-linear results $O(S^1)$ almost overlap with those of the LSWT, reinforcing the conclusion that the higher order expansion $O(S^0)$ constitutes a significant improvement over the former method.

Another method used for benchmarking is the recently developed generalized spin-wave theory [58]. This approach is also semi-classical in the sense that it still involves an expansion about a small parameter which measures the deviations from a purely classical ground state. However, it has the advantage of using the fundamental representation of the $SU(N)$ group instead of $SU(2)$ (in our case, $N = 2S + 1 = 3$ for $S = 1$), designed to capture spin-quadrupolar order in addition to the dipolar magnetic order. As we show below, this approach also allows for a more accurate treatment of the biquadratic spin-spin interaction. We note that unlike the usual spin-wave theory, the GSWT introduces several bosonic modes, so we always have $m = 2S$ different dispersions instead of a single one. We only plot the lowest-energy mode, since this is the one that describes the low-lying spin-wave excitations.

When plotting the results of all these alternate approaches side-by-side with the DM results [Fig. 6(a)], we can readily check that the high order NLST, GSWT and our main choice DM all reflect accurately the effects of quantum fluctuations. In the case of the GSWT, this is achieved by the inclusion of the most general order parameter, largely improving over the usual LSWT which only accounts for fluctuations around the classical vector field. For the DM approach, it is essential to include all the possible spin correlations, which is achieved by performing a full decoupling of the biquadratic term (see Appendix C), rather than the mean-field decoupling used in Eq. (10). It is worth noting that such a full decoupling is possible in the DM approach while still preserving the Goldstone modes because no $1/S$ terms are necessary when using the DM bosons. This is precisely what also makes it our method of choice for the analyses at higher temperatures.
FIG. 6. (Color online) Discrepancies among the results from the different spin-wave theories. In a) we plot the spin-wave dispersions of the magnetically ordered ground-state of the $J_1 - J_2 - K$ model for the case on $S = 1$, for a value of $K = 0.8 J_2$, using the results from all the methods we tried. Low order NLSWT is almost as ineffective as simple LSWT at capturing the effect of the quantum fluctuations while its high order counterpart, GSWT and the DM boson method all show a similar level of accuracy. In b) we plot the evolution of the normalized maximum of the ground-state dispersion [at point $(\pi, \pi)$] for the $J - K$ model, comparing the results from all our methods to those of Ref. [49]. Again, results from low order NLSWT overlap with those obtained via SB (known to also underestimate fluctuations) while the rest of the methods account for all the possible correlations to a similar extent.

For completeness, we also apply all of the above techniques to the simpler $J - K$ model with a Néel ground state ($J_2 = 0$), in order to further the comparison between different methods in Ref. [49]. We obtain the maxima of the dispersions at the Brillouin zone edge which are shown in Fig. 6 b) for several different value of the biquadratic coupling $K$ (the values are normalized with respect to the case of $K = 0$ for ease of comparison). As expected, the results for the low order $O(S^1)$ NLSWT overlap with the curve for the SB results. As explained above, the overestimation of the dispersion maxima originates from the inability of these approaches to correctly capture the effect of quantum fluctuations. By contrast, higher order NLSWT $O(S^0)$, the GSWT and the DM methods all give the slopes of the spin-wave dispersion that are nearly identical to those found by exact diagonalization in Ref. [49]. We must note that we are using a self-consistent set of Euler–Lagrange equations to solve for the expectation values of the variables in Eq. (10), whereas the authors of Ref. [49] directly minimize the free energy. While in principle, the two methods are equivalent, there are small numerical discrepancies between our results and those of Stanek et al. [49], which however do not affect any of our conclusions. It turns out that the latter method is much harder to implement in the case of the frustrated Hamiltonian, which is why we stick with the self-consistent Euler–Lagrange approach in this study.

V. DYNAMICAL STRUCTURE FACTORS

The differential cross section of the inelastic neutron scattering is proportional to the dynamical spin structure factor $S(q, \omega) = \int dt e^{i\omega t} \langle S_q(t) \cdot S_{-q}(0) \rangle$ which measures spin-spin correlations. At zero temperature, only the transverse components of $S^{xx}(q, \omega) = S^{yy}(q, \omega)$ contribute, which can be expressed as follows [49]:

$$S^{xx}(q, \omega) = N_s \pi S^{eff} \left( \frac{\mu q_0 - \Delta q_0}{\omega q_0} \right) \delta(\omega - \omega q_0),$$

where the coefficients $\mu q_0$ and $\Delta q_0$ are given by the expressions in Eqs. [10] and [11].

At finite temperatures, longitudinal correlations ($S^{zz}$) must also be taken into account. This gives a somewhat more complex formula [49]:

$$S(q, \omega) = \frac{2\pi}{N_s} \sum_k \sum_{s,s'} [\cosh(2\theta_k + q) - 2\delta_{k+1,0}] \times \delta(\omega - s\epsilon_k - s'\epsilon - q) n_k^{s} n_{k+q}^{s'},$$

The momentum-dependent angle $\theta_k$ is obtained from the particular Bogoliubov transformation and is given by $\tan(2\theta_k) = \frac{\Delta_k}{\epsilon_k}$. We use the notation $n_k^{s}$ to refer to $n_{k}^{s} = n_{k}$ and $n_{k}^{\pm} = n_{k} \pm 1$, respectively, where $n_{k}$ is the Bose distribution function evaluated at the spin-wave frequency $\omega_k = 2\sqrt{A_k^2 - |B_k|^2}$. 

For $J_2 = 0$, the NLSWT is almost as ineffective as simple LSWT at capturing the effect of the quantum fluctuations while its high order counterpart, GSWT and the DM method all show a similar level of accuracy. In b) we plot the evolution of the normalized maximum of the ground-state dispersion [at point $(\pi, \pi)$] for the $J - K$ model, comparing the results from all our methods to those of Ref. [49]. Again, results from low order NLSWT overlap with those obtained via SB (known to also underestimate fluctuations) while the rest of the methods account for all the possible correlations to a similar extent.
FIG. 7. (Color online) Comparison of the calculated (left panels) and experimentally measured (right panels) dynamical structure factors at low temperature (T = 7 K in Ref. [52]). Calculations were performed using GSWT at K = 0 for energy cuts of a) ω = 2J₂, b) ω = 3J₂, c) ω = 5J₂, and d) ω = 6J₂. A value of J₂ = 25 meV was used with broadening γ = 0.5J₂ in Eq. (11).

To obtain finite results, we substitute the δ-function in Eqs. (11) and (12) by a Lorentzian broadening:

\[ \delta(\omega - \Delta \epsilon) \rightarrow \frac{1}{\pi} \frac{\gamma}{(\omega - \Delta \epsilon)^2 + \gamma^2} \] (13)

The width γ includes the instrumental broadening used to mimic the finite experimental resolution and, more importantly, it also incorporates the Landau damping effect due to coupling of spin waves to itinerant electrons. Calculating the magnitude of γ would require a detailed microscopic theory that is beyond the scope of this article. Instead, we use γ = 0.5J₂, somewhat smaller than the value deduced from INS data on CaFe₂As₂ [66] and used in previous theoretical works [50]. We find that including larger values of γ ≥ J₂ renders the transverse contribution in Eq. (11) nearly featureless, and we find that the main non-trivial effect of broadening is on the longitudinal component in Eq. (12).

We have used both the DM method (see Fig. 3) and the generalized spin-wave theory (see Fig. 7) to compute the dynamical spin structure factor and compared it with the INS experiments on BaFe₂As₂ from Refs. 16 and 52. At high temperatures T > T_N we used the DM method to compare with the recent experiments [15], since the DM bosons faithfully capture large deviations from the classical ground state even in the disordered phase [50]. The results of this comparison are plotted in Fig. 3 and have been discussed already in Section III.

To compute the spin structure factor at low temperatures, we chose the GSWT method due to its simplicity, having already checked that its accuracy is comparable to that of the high order NLSWT and the DM method in the previous section. Indeed, both the DM and GSWT produce nearly identical results when using Eq. (11) at T = 0, but the GSWT is much more straightforward to implement (see Appendix B for more details). To compare with the experimental results, we chose to plot four different energy cuts in Fig. 7 and we have rotationally symmetrized the results so that they can be directly compared with the C₄-symmetric results observed in the twinned samples of BaFe₂As₂ in Ref. [52].

The positions of the peaks in S(q, ω) are determined by the condition of energy conservation in the δ-function in Eqs. (11) and (12), which dictates that at each frequency, maxima will appear at ω ≈ ω_q. For the lower energy cuts in Fig. 7(a) and (7)b, we observe rings that appear centered at q₁ = (π, 0) and q₂ = (0, π), the two possible degenerate ordering wave-vectors. The ellipticity of the rings is an indicator of the anisotropy of the system. As we increase the energy of the cuts, the rings expand towards the magnetic zone boundary eventually shifting their peaks from q₁ and q₂ to q = (π/2, π/2), as Figs. 7c) and 7d) demonstrate. Our results are in semi-qualitative agreement with experimental data [22] (shown in the right panels of Fig. 7 for comparison). The main discrepancies are most likely due to a larger effective broadening in the real compounds, produced by the Landau damping as mentioned earlier.
VI. EVOLUTION OF NEMATICITY

In recent neutron scattering experiments, anisotropies have been measured in the iron pnictides, even in the tetragonal, paramagnetic phase above $T_s$. This suggests that the inhomogeneities are not linked to the structural changes, but rather to the magnetic fluctuations. This idea is further supported by the recent nematic susceptibility measurements inferred from electric resistivity in detwinned samples of BaFe$_2$As$_2$. Therefore, the structural lattice instability would not be the cause of nematicity. On the contrary, incipient magnetic fluctuations in the nematic regime appear to be responsible for the structural changes.

In two-dimensional spin systems studied previously, the Ising-nematic transition is possible even at finite temperature since it does not break any continuous symmetry and is not subject to the Mermin–Wagner theorem. Here we study the nematic order in three spatial dimensions, by analyzing the temperature evolution of both the staggered magnetization $m_s$ and the electronic nematicity ($\Gamma_y - \Gamma_x$) in Eq. (8). For reasons already mentioned in section IV, namely the lack of a small expansion parameter, the DM boson representation is our method of choice due to its reliability even deep into the paramagnetic phase. We use the full decoupling of the biquadratic term outlined in Ref. 73 and contrast them with the result of the simple mean-field decoupling [see Eq. (9)] used in the earlier work by one of the authors.

We define $T_\sigma$ as the temperature at which the spin-spin correlations $\Gamma_x$ and $\Gamma_y$ in Eq. (8) become equal in both crystallographic directions so that the system covers the full $C_4$ symmetry of the tetragonal lattice. As we anticipated in section III, we find that the staggered magnetization $m_s$ always vanishes at a lower or equal temperature than that at which the nematicity does, that is $T_N \leq T_\sigma$. We refer to the region $T_N \leq T \leq T_\sigma$ as the pure Ising-nematic phase. However, we find that for typical values of $J_c$ and $K$ [see Fig. 1] the temperature range of this phase is either very narrow or non-existent. In order to find out the origin of this behavior, we perform an exhaustive analysis of the $T_\sigma - T_N$ range as a function of the two tunable parameters $J_c$ and $K$.

The evolutions of both $T_N$ and $T_\sigma$ have been previously studied for the case with no biquadratic term ($K = 0$) in Ref. 60. As expected, we obtain the same qualitative behavior [Fig. 2b] as other authors, where $T_\sigma$ increases slowly and linearly with the inter-planar coupling, whereas $T_N$ is zero in two spatial dimensions but quickly approaches $T_\sigma$ as soon as even a small value for $J_c$ is allowed. This is the expected behavior since tuning $J_c$ amounts to varying the dimensionality of the system and thus as soon as we enter the three-dimensional regime the magnetic long-range order stabilizes fast, while $T_\sigma$ still varies weakly. In this work, we further investigate how these results are affected by the inclusion of non-zero biquadratic coupling $K$.

What we find is that $K$ has a dramatic effect and leads to the very quick narrowing of the dynamical temperature range $T_\sigma - T_N$ even for small values of $K$ and realistic $J_c$ [Fig. 2b]. Similarly, one can fix $J_c$ and study the evolution of the range $T_\sigma - T_N$ as function of biquadratic spin coupling $K$. The results are plotted in Fig. 8 which show that $T_N$ quickly approaches $T_\sigma$ as $K$ is increased. This behavior contrasts with the result in Ref. 50 which find an increase of both transition temperatures with $K$.  

![FIG. 8. (Color online) $T$ vs. $K$ phase diagrams for the case of $S = 1$ with values of the interlayer coupling of a) $J_c = 0.1 J_2$ and b) $J_c = 0.01 J_2$, respectively. The colored area corresponds to the Ising-nematic region. Both temperatures decrease rapidly with $K$ and coincide fast, with wider nematic regions still appreciable for smaller interlayer couplings.](image)
Because a simple mean-field decoupling Eq. (1) was used in Ref. [50, they results are easily explained in terms of effective couplings $J_{1}^{x,y} = J_1 - 2K \Gamma_x(\gamma)$, where the antiferromagnetic character in the $x$-direction in enhanced due to the larger effective value of the coupling in this direction (note that $\Gamma_x < 0$), while the coupling in the $y$ direction becomes smaller. Thus, within this mean-field picture, $K$ is clearly responsible for enhancing the anisotropy, so that the $C_2$ symmetry breaking becomes more stable for larger $K$. In our Dyson–Maleev treatment, we use a non-mean field decoupling method [53] which more accurately accounts for quantum fluctuations. Ironically, this results in the diminished regime of stability of the Ising nematic phase as $K$ increases (see Fig. 5).

We conclude that the pure nematic regime disappears for realistic values of $J_i$ and $K$, so that considerable fine-tuning of these parameters is necessary for Ising nematicity to occur in an appreciable temperature range above $T_N$.

**VII. CONCLUSIONS**

We studied the frustrated bilinear-biquadratic spin model applied to the iron pnictides, both in the magnetically ordered phase where the compounds form columnar antiferromagnets, as well as above the magnetic transition temperature, in the paramagnetic regime. We found signatures of a pure Ising-nematic phase in the form of a $C_2$ anisotropy in the spin-spin dynamical structure factors, one which persists even for $T > T_N$. We identify the temperature ($T_{\sigma}$) at which the system recovers the full $C_4$ symmetry of the spin response with the physical structural transition temperature $T_s$.

However, recent experiments on Ni-doped BaFe$_2$As$_2$ have found anisotropies in the spin excitations even beyond this point, up to some temperature $T^\ast > T_{\sigma}$. The lack of ability of our theory to capture these anisotropic features above $T_{\sigma}$ can be explained in one of the two ways. The first scenario, is that the spin response anisotropy is not static but rather originates from nematic fluctuations alone. Since we use static order parameters for the decoupling of the Hamiltonian in the spin-wave theories [see e.g. Eq. (1)], this may be the reason why these features are not captured in our theoretical results above $T_s$. The second scenario is that in the absence of applied uniaxial strain, the $C_4$ symmetry is indeed restored above the temperature $T_{\sigma}$ which we identify with $T_{s}$ in our theory. However, the effect of uniaxial strain used to detwin the samples is to smear the transition, making it a smooth crossover with $C_2$ anisotropies that persist up to some higher temperature $T^\ast$. Given the very large nematic susceptibility near $T_s$, as inferred from the resistivity measurement, this would be a natural explanation since the applied strain would be expected to trigger a nematic response whose tail is seen at elevated temperatures $T > T_s$. If this is indeed the case, one would expect the crossover temperature $T^\ast$ to be strain-dependent. Future experiments under variable strain would therefore be very desirable to help clarify this issue.

The present study improves upon previous work by other authors [50,52] by including the effect of the biquadratic coupling $-K(S_i \cdot S_j)^2$ to study the evolution of the dynamical temperature range $T_{\sigma} - T_N$ of the Ising nematic phase. We confirmed that this range decreases rapidly with the inclusion of even a small finite interlayer coupling $J_c$, since it stabilizes the magnetic long range order. Furthermore, we found that the inclusion of $K$ has a similar effect, requiring precise fine-tuning in order to get a pure nematic phase over an appreciable temperature range. Unlike the previous work in Ref. [50] which adopted a simple mean-field treatment of the biquadratic term [see Eq. (9)], we have used a more accurate scheme that includes fluctuations beyond mean-field using a variety of theoretical techniques (non-linear spin-wave theory, Dyson–Maleev method, and GSWT). In all cases, we found that the present approach results in a drastically different spin-wave dispersion near the Brillouin zone boundary compared to that from mean-field treatment in Ref. [50]. Similarly, the evolution of the transition temperatures ($T_N$ and $T_{\sigma}$) with increasing $K$ is also very different in the present work, reflecting the higher accuracy of the non-mean-field decoupling that we employed.

On a more technical level, we have benchmarked several spin-wave theories by comparing the spin-wave dispersions and the dynamical spin structure factors to the inelastic neutron scattering experiments. We found our main method of choice to be the Dyson–Maleev modified spin-wave theory, which accurately captures the effect of quantum and thermal fluctuations and produces a good semi-quantitative agreement with INS experiment, especially at higher temperatures $T \gtrsim T_N$. Finally, the recently developed generalized spin wave theory [53] deserves a special mention due to its elegance and simplicity, while providing results comparable in accuracy to those obtained with the DM and high order NLSWT methods. However, the GSWT fails closer to $T_N$ when the thermal fluctuations significantly reduce the ordered moment.

**Appendix A: Non-linear spin-wave theory**

As mentioned before, our starting point is the frustrated Heisenberg Hamiltonian with additional biquadratic coupling.

$$\mathcal{H}_{2D} = J_1 \sum_{r} \mathbf{S}_r \cdot \mathbf{S}_{r'} + J_2 \sum_{<r,r'>} \mathbf{S}_r \cdot \mathbf{S}_{r'} - K \sum_{<r,r'>} (\mathbf{S}_r \cdot \mathbf{S}_{r'})^2$$

(A1)

We will concentrate on the regime of parameters $J_2/J_1 > 1/2$ (particularly, we choose $J_1 = J_2$ through-
out the entire paper) so that the lattice is in the column-
lar AFM phase with ordering wave-vectors $Q = (\pi, 0)$ or $Q = (0, \pi)$, evidenced by neutron scattering experiments. We can then consider two interpenetrating sublattices $A$ and $B$, and sum over all points $r \in A$ and $r' \in B$. Since the spins in sublattice $B$ are aligned antiferromagneti-
cally with respect to those in sublattice $A$, we perform a rotation by $\pi$ in the former, sending $S^x_r \to S^x_r = S^y_r$, $S^y_r \to -S^y_r$ and $S^z_r \to -S^z_r$.

In our first approach, we write the spin operators via
the well established Holstein-Primakoff representation, in
terms of the bosonic operators $a^\dagger_r, a_r$.

$$S^+_r = \sqrt{2S - a^\dagger_r a_r} a_r$$
$$S^-_r = a^\dagger_r \sqrt{2S - a^\dagger_r a_r}$$
$$S^z_r = S - a^\dagger_r a_r$$

(A2)

After performing the rotation of the $B$ sublattice, the
Heisenberg terms in the hamiltonian take the following
form, for the antiferromagnetically and ferromagnetically
aligned directions, respectively.

$$(S_r \cdot S_{r'})_{AFM} = -S^2 + S \left( a^\dagger_r a_r + a^\dagger_{r'} a_{r'} + a_r a_{r'} + a^\dagger_r a^\dagger_{r'} a_r a_{r'} - \frac{1}{4} \left( a^\dagger_r a_r a^\dagger_{r'} a_{r'} + a^\dagger_r a^\dagger_{r'} a_r a_{r'} + a^\dagger_r a_{r'} a^\dagger_{r'} a_r + a^\dagger_r a_{r'} a_r a^\dagger_{r'} + a^\dagger_{r'} a_r a_r a^\dagger_{r'} + a^\dagger_{r'} a^\dagger_r a_r a_r + a^\dagger_{r'} a_r a_r a^\dagger_r + a^\dagger_{r'} a_r a^\dagger_r a_r \right) \right) - \sum_k \omega_k \left( a^\dagger_k a_k + \frac{1}{2} \right)$$

(A4)

Where we expanded the square roots by taking
$a^\dagger_r a_r / 2S$ as our small parameter. For orders of $O(S^0)$
and higher in $1/S$, this yields terms with more than 2
bosonic operators. To make them solvable, we decou-
ple them by using Wick’s theorem and consider all the
possible decouplings. We take all averages to be real for
convenience, without loss of generality.

$$n = \langle a^\dagger_r a_r \rangle$$
$$g_x = \langle a_r a_{r+\hat{x}} \rangle = \langle a^\dagger_r a^\dagger_{r+\hat{x}} \rangle$$
$$f_y = \langle a_r a_{r+\hat{y}} \rangle = \langle a^\dagger_r a^\dagger_{r+\hat{y}} \rangle$$
$$g_{xy} = \langle a_r a_{r+\hat{x}+\hat{y}} \rangle = \langle a^\dagger_r a^\dagger_{r+\hat{x}+\hat{y}} \rangle$$

(A5)

We assume the rest of the averages to be zero by virtue of
the conservation of the total z-component of the spin
$(S^2 = \sum_i S^i)$ in each direction. In principle, both bilin-
ear and biquadratic terms can be treated in this manner.
However, our calculations show that this approach over-
estimates the fluctuations and can destroy long-range or-
der and biquadratic terms can be treated in this manner.
Thus, we stick to a more usual Hubbard-Stratonovich decoupling of the terms of higher
order in the spins.

$$(S_r \cdot S_{r'})^2 \approx 2 \langle S_r \cdot S_{r'} \rangle S_r \cdot S_{r'} - \langle S_r \cdot S_{r'} \rangle^2, \quad (A6)$$

where the staggered magnetizations in each direction
can be expressed in terms of the averages mentioned
above as follows [note that we only take the terms up
to order $O(S)$ for the lower order NLSWT].

$$\Gamma_x = \langle S_r \cdot S_{r+\hat{x}} \rangle = -\left( S - n - g_x \right)^2$$
$$\Gamma_y = \langle S_r \cdot S_{r+\hat{y}} \rangle = (S - n + f_y)^2 \quad (A7)$$

After decoupling the higher order terms and applying
the Hubbard-Stratonovich transformation to the bi-
quadratic term, we obtain a hamiltonian of the form (up
to constant terms):

$$H_{NLSW} = \sum_k \left[ A_k \left( a^\dagger_k a_{-k} + a^\dagger_{-k} a_k \right) + B_k \left( a^\dagger_k a_{-k} + a^\dagger_{-k} a_k \right) \right] = \sum_k \omega_k \left( a^\dagger_k a_k + \frac{1}{2} \right)$$

(A8)

Which, after proper diagonalization has the following
Bogoliubov dispersion:

$$\omega_k = 2\sqrt{A_k^2 - B_k^2} \quad (A9)$$

where the coefficients $A_k$ and $B_k$ are given by the fol-
lowing expressions. Once again, we only consider terms
up to order $O(S)$ in the lower order case.

$$A_k = (J_1 - 2K_x)(S - n - g_x) + (J_1 - 2K_x)(S - n + f_y)(\cos k_y - 1) + 2J_2(S - n - g_{xy})$$
$$B_k = (J_1 - 2K_x)(S - n - g_x) \cos k_x + 2J_2(S - n - g_{xy}) \cos k_x \cos k_y \quad (A10)$$

Using the coefficients of the transformation, we can
finally solve the system self-consistently.
\[ \alpha_x = \frac{1}{N_s} \sum_k \left( \langle a_k^+ a_k \rangle + \langle a_k a_{-k} \rangle \cos k_x \right) \]
\[ = -\frac{1}{2} + \frac{1}{N_s} \sum_k (1 + 2n_k) \frac{A_k - B_k \cos k_x}{\omega_k} \]
\[ \beta_y = \frac{1}{N_s} \sum_k \langle a_k^+ a_k \rangle (1 - \cos k_y) = \]
\[ = -\frac{1}{2} + \frac{1}{N_s} \sum_k (1 + 2n_k) \frac{A_k (1 - \cos k_y)}{\omega_k} \]
\[ \alpha_{xy} = \frac{1}{N_s} \sum_k \left( \langle a_k^+ a_k \rangle + \langle a_k a_{-k} \rangle \cos k_x \cos k_y \right) = \]
\[ = -\frac{1}{2} + \frac{1}{N_s} \sum_k (1 + 2n_k) \frac{A_k - B_k \cos k_x \cos k_y}{\omega_k} \]

Where for convenience we’re using: \( \alpha_i = n + g_i \) for \( i = x, xy \) and \( \beta_y = n - f_y \) (taken from Ref. 58).

**Appendix B: Generalized spin-wave theory**

This new approach, proposed recently in Ref. 58 is derived starting from an \( N \) number of Schwinger bosons instead of the usual two modes. While the usual linear spin-wave theory (LSWT) describes only fluctuations around the classical vector field, by considering the most general \( SU(N) \) order parameter, we obtain as a result waves that describe any underlying ground state multipolar order. They satisfy the local \( SU(N) \) constraint on the local number of bosons:

\[ \sum_{m=0}^{N-1} b_{km}^m b_{km} = NS \]  

(B1)

In the fundamental representation of \( SU(N) \) where \( NS = 1 \), all spin operators can be expressed as a bilinear combination in the bosonic modes. In particular, for the local spin operators where \( N = 2S + 1 \), the matrix elements are given by the following expressions.

\[ S_{mm'}^{x} = \delta_{m(m'-1)} \frac{\sqrt{(m+1)(2S-m)}}{2} + \]
\[ + \delta_{m(m+1)m'} \frac{\sqrt{(m'+1)(2S-m')}}{2} \]
\[ S_{mm'}^{y} = \delta_{m(m'-1)} \frac{\sqrt{(m+1)(2S-m)}}{2i} - \]
\[ - \delta_{m(m+1)m'} \frac{\sqrt{(m'+1)(2S-m')}}{2i} \]
\[ S_{mm'}^{z} = \delta_{m(m'-1)} (S-m) \]

And the elements of the matrix associated with the bilinear operator (for details, see Ref. 58) \( S_{\mu}^{\nu} S_{\gamma}^{\nu} \) are:

\[ S_{mm'}^{\nu \mu} = \sum_{m''} S_{mm''}^{\nu} S_{m''m'}^{\mu} \]

(B3)

Finally, following the spirit of the Holstein-Primakoff representation, we impose the constraint (6) by requiring that the condensed fraction satisfies.

\[ b_{r0}^0 = b_{r0} = \sqrt{1 - \sum_{m=1}^{N-1} b_{rm}^m b_{rm} \simeq 1 - \frac{1}{2} \sum_{m=1}^{N-1} b_{rm}^m b_{rm} } \]  

(B4)

Keeping only the quadratic terms will result in an effective Hamiltonian, which we can diagonalize, obtaining \( (N - 1) \) Bogoliubov dispersions, one for each bosonic mode. Using these representations for the spin operators, we obtain the following Hamiltonian (again, up to constant terms):

\[ H_{GW} = \sum_{k,m} \mu_{km} \left( b_{km}^+ b_{km} + b_{km}^+ b_{km}^-- \right) + \]
\[ + \Delta_{km} \left( b_{km}^+ b_{km}^-- + b_{km}^+ b_{km}^- \right) \]

(B5)

With coefficients (in the two-dimensional case) given by:

\[ \mu_{k1} = S \left[ 2J_2 + K(2S-1)^2 \right] + \]
\[ + S \left[ J_1 - 2KS(S-1) \right] \cos k_y \]
\[ \mu_{k2} = 4S \left[ J_2 + K(S-1)(2S-1) \right] - \]
\[ - KS(2S-1) \cos k_y \]
\[ \mu_{km} = mS \left[ 2J_2 - K(m-2S)(2S-1) \right] \]

(B6)

\[ \Delta_{k1} = \left\{ J_1S + KS \left[ 1 + 2S(S-1) \right] \right\} \cos k_x + \]
\[ + 2J_2 S \cos k_x \cos k_y \]
\[ \Delta_{k2} = - KS(2S-1) \cos k_x \]
\[ \Delta_{km} = 0 \]

(B7)

For \( m > 1 \). Finally, the diagonalized hamiltonian takes the following form.

\[ H_{GW} = \sum_k \left( \epsilon_k \beta_k^1 \beta_k^1 + \epsilon_k \beta_k^2 \beta_k^2 \right) + \]
\[ + \sum_{m=3}^{2S} \epsilon_{km} \beta_{km}^1 \beta_{km} \]

(B8)

With Bogoliubov dispersions:

\[ \epsilon_k = 2 \sqrt{\mu_{k1} - \Delta_{k1}^2} \]
\[ \epsilon_{k2} = 2 \sqrt{\mu_{k2} - \Delta_{k2}^2} \]

(B9)
And flat dispersions: $\omega_{kn} = 2\mu_{kn}$ for $m > 2$. For convenience, we identify the modes with their value for the $z$-component of the spin, for the particular case of $S = 1$ we’ll be using. Thus the first mode $m = 1$ (with $m_z = 0$) becomes $m = 0$ and the second mode (with $m_z = -1$) $m = 2$ is now represented by $m = \downarrow$.

Appendix C: Dyson-Maleev bosons

In the methods explained above the spin operators are expanded around the classical ground state configuration. This approach, however, is not a valid one in the paramagnetic regime, or even in the ordered phase, when the sublattice magnetization becomes too small. To study the temperature evolution of the system is thus more appropriate to use an alternate approach, which doesn’t rely on any small parameter. In this representation, the spin operators are expressed in terms of Dyson-Maleev bosons.

\[
S^+_r = \sqrt{2S} \left(1 - \frac{a^+_r a_r}{2S}\right) \, a_r
\]

\[
S^-_r = \sqrt{2S} a^+_r
\]

\[
S^z_r = S - a^+_r a_r
\]

We note that $S^+_r$ and $S^-_r$ are no longer complex conjugates. However, the resulting Hamiltonian is Hermitian so we proceed anyways.

At this point we can treat the biquadratic term using two different approaches. The first one uses the mean-field procedure outlined above, where the higher order term is substituted by the following decoupling: $(S_r \cdot S_r)^2 \simeq 2 (S_r \cdot S_r) S_r \cdot S_r - (S_r \cdot S_r)^2$. The results obtained by this method have already been studied for the finite temperature case in\(^\text{[55]}\). An alternate approach is to decouple the entire biquadratic term using all possible decouplings via Wick’s theorem and the averages specified before. This method was used by\(^\text{[23]}\) in the $T = 0$ case and we expand it to include the finite temperature case. Once again, we obtain dispersions of the form: $\omega_k = 2\sqrt{A_k^2 - B_k^2}$ for the two- and three-dimensional cases, respectively. Following again the notation from Ref.\(^\text{[58]}\)

\[
A_{k(2D)} = \lambda + 2J_2(S - \alpha_{xy}) + J_1 [r_x(S - \alpha_x) + r_y(S - \beta_y)(\cos k_y - 1)]
\]

\[
B_{k(2D)} = J_1 r_x(S - \alpha_x) \cos k_x + 2J_2(S - \alpha_{xy}) \cos k_x \cos k_y
\]

\[
A_{k(3D)} = A_{k(2D)} + J_6 (S - \alpha_z)
\]

\[
B_{k(3D)} = B_{k(2D)} + J_6 (S - \alpha_z) \cos k_z
\]

Where $r_x$, $r_y$ stand for the following expressions:

\[
r_x = 1 + \frac{K}{S - \alpha_x} \left[2S^3 - 2S^2(1 + 5\alpha_x) + S(18\alpha_x^2 + 8\alpha_x + 1) - 12\alpha_x^3 - 9\alpha_x^2 - 2\alpha_x\right]
\]

\[
r_y = 1 - \frac{K}{S - \beta_y} \left[2S^3 - 2S^2(1 + 5\beta_y) + S(18\beta_y^2 + 8\beta_y - 12\beta_y^3 - 9\beta_y^2 - \beta_y\right]
\]

Finally, we introduced the chemical potential $\lambda$ to enforce the constraint $(S_z) = 0$ in the paramagnetic regime. Thus, $\lambda = 0$ in the magnetically ordered phase. The resulting set of self-consistent equations has the same form as [A11], with the addition of the following equation in the three-dimensional case.

\[
\alpha_z = \frac{1}{N_z} \sum_k \left(\langle a^+_k a_k \rangle + \langle a^+_k a_{-k} \rangle \cos k_z \right) = -\frac{1}{2} + \frac{1}{N_z} \sum_k (1 + 2\eta_k) \frac{A_k - B_k \cos k_z}{\omega_k}
\]

We differentiate the two domains, $T < T_N$ (where $n < S$ and $\lambda = 0$) and $T > T_N$ (where $n = S$ and $\lambda \neq 0$).

ACKNOWLEDGMENTS

The authors are grateful to Pengcheng Dai for providing Fig. 3c, d) and to Zhentao Wang and Yang-Zhi Chou for helpful discussion. This work was performed with partial support from the Welch Foundation grant no. C-1818, and the NSF CAREER award no. DMR-1350237.

1. Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono, J. Am. Chem. Soc. 130, 3296 (2008).
2. Z.-A. Ren et al., Chin. Phys. Lett. 25, 2215 (2008).
3. C. de la Cruz et al., et al., Nature 453, 899 (2008).
4. J.-H. Chu, H.-H. Kuo, J. G. Analytis, and I. R. Fisher, Science 337, 710 (2012).
5. C. Fang, H. Yao, W.-F. Tsai, J. P. Hu, and S. A. Kivelson, Phys. Rev. B 77, 224509 (2008).
6. C. Xu, M. Müller, and S. Sachdev, Phys. Rev. B 78, 020501 (2008).
7. J. Dai, Q. Si, J.-X. Zhu, and E. Abrahams, Proc. Natl. Acad. Sci. USA 106, 20550 (2009).
8. R. M. Fernandes, E. Abrahams, and J. Schmalian, Phys. Rev. Lett. 107, 217002 (2011).
