Three-band Hubbard model for \( \text{Na}_2\text{IrO}_3 \):
Topological insulator, zigzag antiferromagnet, and Kitaev-Heisenberg material

Manuel Laubach,1, Johannes Reuther,2,3 Ronny Thomale,4 and Stephan Rachel1,5,6
1Institut für Theoretische Physik, Technische Universität Dresden, 01062 Dresden, Germany
2Dahlem Center for Complex Quantum Systems and Fachbereich Physik, Freie Universität Berlin, 14195 Berlin, Germany
3Helmholtz-Zentrum Berlin für Materialien und Energie, D-14109 Berlin, Germany
4Institut für Theoretische Physik, Universität Würzburg, 97074 Würzburg, Germany
5Department of Physics, Princeton University, Princeton, New Jersey 08544, USA
6School of Physics, University of Melbourne, Parkville, VIC 3010, Australia

\( \text{Na}_2\text{IrO}_3 \) was one of the first materials proposed to feature the Kane-Mele type topological insulator phase. Contemporaneously it was claimed that the very same material is in a Mott insulating phase which is described by the Kitaev-Heisenberg (KH) model. First experiments indeed revealed Mott insulating behavior in conjunction with antiferromagnetic long-range order. Further refined experiments established antiferromagnetic order of zigzag type which is not captured by the KH model. Since then several extensions and modifications of the KH model were proposed in order to describe the experimental findings. Here we suggest that adding charge fluctuations to the KH model represents an alternative explanation of zigzag antiferromagnetism. Moreover, a phenomenological three-band Hubbard model unifies all the pieces of the puzzle: topological insulator physics for weak and KH model for strong electron–electron interactions as well as a zigzag antiferromagnet at intermediate interaction strength.

Introduction. The past decade of condensed matter physics has been strongly influenced by the rise of spin orbit coupling. Although known from the early days of quantum mechanics, only recently it started to unfold its complexity and potential. Most notably, topological insulators \cite{1,2} are spin orbit dominated band insulators with exotic surface properties. Amongst strongly correlated systems, spin-orbital liquids \cite{3} and spin–orbit–assisted Mott insulators \cite{4} are fascinating states of matter which were proposed in the past years.

Soon after the theoretical prediction of the quantum spin Hall (QSH) \cite{5} effect it was realized that graphene has negligibly small spin orbit coupling such that the topological insulator phase is not observable. Ever since, the field has been searching for other honeycomb lattice materials with possibly heavier elements and alternative mechanisms to accomplish large gap QSH states in honeycomb monolayers \cite{6}. One of the first proposals to feature Kane-Mele-type QSH physics focussed on \( \text{Na}_2\text{IrO}_3 \) where the Ir atoms form an effective 2D honeycomb net \cite{7}. Shitade et al. argued that if Coulomb interactions are not too strong \( \text{Na}_2\text{IrO}_3 \) might exhibit a correlated topological insulator (TI) ground state. Around the same time, Jackeli and Khaliullin proposed that \( \text{Na}_2\text{IrO}_3 \) might be a Mott insulator where the spin physics is dominated by both Heisenberg and exotic Kitaev spin exchange \cite{8}, referred to as Kitaev-Heisenberg (KH) model. Since the pure Kitaev model \cite{9} is a rare example of an exactly soluble quantum spin liquid system, the prospect of approximately realizing it in an actual material has attracted enormous attention. The phase diagram of the KH model contains a Néel ordered antiferromagnet, the Kitaev spin liquid, and a striped antiferromagnet as an intermediate phase \cite{10}. The successful growth of single crystals, however, established a zigzag antiferromagnetic ground state \cite{11,12} – as suggested earlier by \textit{ab initio} calculations \cite{13} – which cannot be explained by the original KH model. Several works proposed extensions of this model \cite{14,15} to account for the experimental findings, ranging from longer-ranged Heisenberg exchange \cite{14,15,26} and “negative” KH interactions \cite{17} to off-diagonal Γ exchange \cite{21,22} and longer-ranged Kitaev interactions \cite{19,23,25}. Furthermore, the formation of molecular orbital crystals has been proposed as an explanation for the phenomenology of \( \text{Na}_2\text{IrO}_3 \) \cite{16}. Despite the variety of different models, it remains elusive until today which might be the spin Hamiltonian to most accurately describe the phenomenology of \( \text{Na}_2\text{IrO}_3 \). Moreover, it has even been questioned whether Coulomb interactions are as strong as believed and whether the Mott limit is justified: \textit{ab initio} results rather suggest a moderate interaction strength \cite{15,13} or even a strong 3D TI phase \cite{27}; likewise, recent ARPES experiments \cite{28} report the observation of metallic surface states which are not compatible with the picture of an antiferromagnet in the deep Mott limit.

Given the enormous interest in the KH model and its extensions, one might wonder how a corresponding Hubbard model could be reconstructed of which the KH model is its strong coupling limit. A first guess might be that the non-interacting TI Hamiltonian considered by Shitade \textit{et al.} \cite{7} leads in the strong coupling limit to the KH model. A simple analysis reveals, however, that in this case Kitaev spin exchange is generated on second neighbor links, while nearest neighbor exchange is of pure Heisenberg type \cite{29}. Likewise, one can construct a single-orbital band structure which indeed leads to the KH Hamiltonian for large \( U \) \cite{30}; it turns out, however,
that such a band structure explicitly breaks time-reversal symmetry since real spin-dependent hoppings on nearest neighbor bonds are involved.

In the first part of this paper, we introduce a three-band Hubbard model with a time-reversal invariant band structure which is constructed such that its strong coupling limit is the pure KH spin model. The minimal model we identified has three orbital degrees of freedom which can be interpreted as the $t_{2g}$ manifold. Thus the phenomenological model introduced here has a strong similarity to the model studied earlier by Rau et al. [21]. The analysis of the non-interacting band structure reveals an extended topological insulator phase. Since there are only two types of time-reversal invariant insulating bandstructures in 2D [33], the weak-coupling limit of our three-band model is topologically equivalent to the one proposed by Shitade et al. [7]. As we analyze the model in the presence of increasing interaction strength we find that it hosts a zigzag ordered phase at intermediate interaction strength $U$ in consistency with experiments for Na$_2$IrO$_3$. In this sense, our phenomenological model unifies the work by Shitade et al. [7] at weak $U$, the work by Jackeli and Khaliullin [8] at strong $U$, and the experimental results [11, 12] at intermediate interaction strength $U$.

In the second part, we choose the model’s parameters to better describe Na$_2$IrO$_3$. We show that the zigzag phase at intermediate interaction strengths persists and we compute single particle spectral functions which can be compared to ARPES spectra. We demonstrate that topologically trivial edge states as remnants of the nearby topological insulator phase are present in the band gap of this three-band model.

**Phenomenological model.** The spinfull three-band model investigated in this work consists of a kinetic, spin-orbit and interaction term

$$H = H_{\text{kin}} + H_{SO} + H_{I}$$

with 1/6 electron filling. The kinetic part takes the form

$$H_{\text{kin}} = \sum_{(i,j)} \sum_{n,n',\alpha} d_{in\alpha}^{\dagger} T^{\gamma(i,j)}_{nn'} d_{in'\alpha} + \text{h.c.},$$

where $d_{in\alpha}^{\dagger}$ denotes a fermionic annihilation (creation) operator on site $i$, with orbital $n \in \{yz, xz, xy\}$ and spin $\alpha \in \{\uparrow, \downarrow\}$. The orbital indices are reminiscent of the common labelling of a $t_{2g}$ manifold which our model is closely related to. The sum is over pairs of nearest neighbor sites $\langle i,j \rangle$ on the honeycomb lattice and $\gamma(i,j) \in \{x, y, z\}$ indicates the type of inequivalent Kitaev bond the pair $\langle i,j \rangle$ belongs to. For a $z$-bond the hopping matrix $T^z_{nn'}$ is given by

$$T^z = \begin{pmatrix} t_1 & t_2 & 0 \\ t_2 & t_1 & 0 \\ 0 & 0 & t_1' \end{pmatrix},$$

and the matrices $T^x_{nn'}$ and $T^y_{nn'}$ follow by cyclic permutations of the rows and columns. Unless explicitly stated otherwise we will assume $t_1 = t_1'$ in the following. The spin-orbit coupling has the usual form

$$H_{SO} = \frac{\lambda}{2} \sum_{i} \sum_{n,n',\alpha,\alpha'} d_{in\alpha}^{\dagger} L_{nn'} \cdot \sigma_{\alpha\alpha'} d_{in'\alpha'},$$

where $L$ denotes an angular momentum operator in $3 \times 3$ matrix representation with $L^2 = l(l+1) = 2$ and $\sigma$ is the Pauli vector. When $\lambda > 0$ the spin-orbit coupling generates a low energy $J = 1/2$ doublet, representing the subspace in which the spin physics of the KH model takes place.

The simplest type of interaction term is an onsite Hubbard repulsion $H_1 = U \sum_i (N_i - 1)^2$ with $N_i = \sum_{n,\alpha} d_{in\alpha}^{\dagger} d_{in\alpha}$ which energetically favors single site occupancy. Performing a strong coupling expansion in second order in $t_{1/2}/U$ as described in Ref. [21] and projecting the result on the low energy $J = 1/2$ subspace (which implies $U \gg t_{1/2}$ and $\lambda \gg t_1'^2/U$), yield an usual spin-isotropic Heisenberg coupling $H = J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j$ with $J = \frac{4t_1'^2}{U}$ where $\vec{S}_i$ are local spin-1/2 operators acting in the $J = 1/2$ doublet. We find that the minimal SU(2) symmetric extension of the plain Hubbard repulsion which generates anisotropic spin interactions in this subspace has the form

$$H_1 = \frac{U - 3J_H}{2} \sum_i (N_i - 1)^2 - 2J_H \sum_i S_i^2,$$

with

$$S_i = \frac{1}{2} \sum_{n,\alpha,\alpha'} d_{in\alpha}^{\dagger} \sigma_{\alpha\alpha'} d_{in\alpha}.$$
and the Hund’s coupling \( J_H \). The prefactors have been chosen to resemble the Kanamori Hamiltonian which has been used to describe multiplet interactions in \( t_{2g} \) shells \[21\] \[22\] \[23\]. (Note that the full Kanamori Hamiltonian also contains additional \( \mathbf{L}^2 \) terms.) Together with Eqs. \[3\] and \[4\] this type of minimal interaction will be investigated for different coupling strengths below.

At strong coupling (assuming \( U \gg t_{1/2} \) and \( \lambda \gg t_{1/2}^2/U \) the Hamiltonian \[[1]\]) reduces to the KH model

\[
H_{\text{KH}} = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - K \sum_{\langle i,j \rangle} \mathbf{S}_i^{\gamma(i,j)} \cdot \mathbf{S}_j^{\gamma(i,j)}
\]

with parameters

\[
J = \frac{4t_2^2(3U - 4J_H)}{3U(U - 4J_H)} , \quad K = \frac{16J_H t_2^2}{3U(U - 4J_H)} .
\]

When \( U > 4J_H \) the model features antiferromagnetic Heisenberg and ferromagnetic Kitaev interactions (i.e., \( J, K > 0 \)) as has been originally proposed by Chaloupka et al. \[10\]. Casting this Hamiltonian into the convenient form where the exchange interactions are written as \( J = 1 - \alpha \), \( K = 2\alpha \), Eq. \[5\] yields the hopping amplitudes

\[
t_1 = C \sqrt{\frac{1 - \alpha}{3UJ_H - 4}} , \quad t_2 = C \sqrt{\frac{\alpha}{2}}
\]

where the parameter \( C \) sets the energy scale of the kinetic Hamiltonian. In total, \( \alpha, U, J_H, \lambda, C \) define a five parameter model, for which we will study the effects of charge fluctuations away from the strong coupling limit (note that one parameter can be absorbed into an overall energy scale). This three band system can be considered as a simplified (and particle-hole transformed) model for \( \text{Na}_2\text{IrO}_3 \) as it has been studied in Ref. \[21\]. A significant simplification comes from setting \( t_1 = t_1' \) in Eq. \[3\].

Given the orbital structure of the \( t_{2g} \) states, \( t_1 < t_1' \) would certainly better account for the microscopic situation in \( \text{Na}_2\text{IrO}_3 \). However, this generates additional off-diagonal and symmetric \( \Gamma \)-exchange, which complicates the effective spin model and may drive spiral types of magnetic order. We, hence, mostly consider \( t_1 = t_1' \) but also discuss the qualitative changes of our results when \( t_1 < t_1' \).

We first focus on the properties of the non-interacting model, \( U = J_H = 0 \). Instead of \( t_{1/2} \) we use \( \alpha \) and the spin orbit amplitude \( \lambda \) as parameters (here we keep \( U/J_H = 5 \) in order to be consistent with the discussion of the interacting case below). Calculating the band structure, we identify (semi-)metallic and insulating phases where the \( \mathbb{Z}_2 \) topological invariant reveals both topologically trivial and non-trivial insulating regimes. These findings are summarized in the non-interacting phase diagram in Fig. \[1\] where we also show the single-particle gap for the TI and normal insulator (NI) phases. Since there are only two types of time-reversal invariant insulating band structures in 2D (trivial and topological) \[31\] the TI phase is topologically equivalent to the one proposed by Shitade et al. Hence, together with the above results from a strong coupling expansion, our three-band model naturally unifies the proposals of Shitade et al. and Jackeli and Khaliullin.

**Zigzag antiferromagnetic phase.** In general, Hubbard models in \( d = 2 \) are not exactly soluble and one is restricted to approximations. Below, we will present numerical evidence that there is a phase transition from the topological insulator phase into a zigzag antiferromagnet (AFM) at intermediate interaction strength based on the Variational Cluster Approach (VCA). VCA is a quantum cluster approach very similar to Cluster Perturbation Theory \[34\] \[35\] where the interacting Green’s function is computed on a small cluster which is then used to construct the full Green’s function for the infinitely large system within a self-consistent, variational scheme. Here we use a \( C \)-shaped four-site cluster with three orbitals per site corresponding to an effective 12-site cluster. For details about the method we refer the interested reader to Refs. \[36\] \[39\]. The VCA method \[36\] \[40\] has been successfully applied to study interacting topological phases \[39\] \[37\] \[38\] \[41\] \[42\] \[44\], as well as magnetically ordered systems \[37\] \[38\] \[42\] \[44\].

Using VCA, we obtain the \( U-\alpha \) phase diagram for \( 0 < U \leq 100 \) and \( \lambda = C = 1 \), see Fig. \[2\]. To illustrate different coupling regimes, we divided the range of
interactions into two parts: weak to intermediate interactions \(0 < U < 3\) and intermediate to strong interactions \(3 < U < 100\). In addition, we show the result in the strong coupling limit \((U \to \infty)\) obtained from exact diagonalization (ED) for a finite system with 24 lattice sites which agrees with previous studies such as Ref. [21]. As a central VCA result, we find that within wide ranges of \(\alpha\), the topological insulator phase is stable up to \(U \approx 1\) followed by a second order phase transition without gap closing into a zigzag AFM regime.

Mapping out the detailed interplay of different phases, we first observe a TI, NI and two semi-metallic points at \(U = 0\), see Fig. 2. Up to \(U \approx 1\) the phase diagram remains largely unchanged except that the semi-metallic point between the TI and NI phases splits into two semi-metallic lines. For \(U \geq 1\) the TI phase, evolving over a wide range \(0 < \alpha < 0.8\), undergoes a phase transition into the zigzag AFM phase. Note that the interaction driven phase transition from the TI phase into a collinear magnet is in general expected for non-frustrated lattices [32, 33]. For \(U \approx 1.6\) to 1.8 a first-order phase transition from zigzag AFM to a Néel ordered AFM takes place. While the Néel phase extends up to \(\alpha \approx 0.8\) for small interactions, with increasing \(U\) it slowly shrinks down to \(\alpha \approx 0.4\) for \(U = 50\) which coincides with the phase transition in the ED phase diagram. For large \(\alpha\) we detect a non-magnetic insulator (NMI) phase which persists up to \(U \approx 10\) and down to very small \(U\) at \(\alpha \approx 0.85\). The origin of this phase is unclear: candidate states are quantum paramagnets including spin liquids and valence bond solids but also ordered states which are incommensurate with respect to the cluster size such as spiral magnets or large-unit cell magnets. At \(U \approx 5\) a stripy AFM becomes stable for large \(\alpha\) and starts to spread over an extended \(\alpha\)-range for increasing \(U\). The reader should notice that VCA is unable to identify the Kitaev spin liquid phase, which is clearly resolved by ED for \(\alpha > 0.8\). Due to the small cluster size and the suppression of long-range entanglement (which is an inescapable consequence of quantum cluster approaches) this is by no means surprising. For large \(U\) and \(\alpha \approx 1\), however, the VCA condensation energies of possible magnetic orders become almost indistinguishable. Such a nearly degenerate ground state scenario might indicate the proximity to a spin-liquid phase. Given the overall phase diagram in Fig. 2, a rough experimental estimate of \(J_f \approx 4\) (or even larger) as found in Ref. [40] would locate Na\(_2\)IrO\(_3\) near the right boundary of the zigzag phase.

On a conceptually more simple level, the presence of a zigzag AFM at intermediate interaction strength can be further substantiated by a Hartree-Fock (HF) mean-field analysis. Recently, the HF approach has been applied to a three-band Hubbard model which is similar to Eq. (1) but differs in details about the interaction terms and the signs of the parameters [47, 48]. In these works, a zigzag AFM phase was likewise identified at intermediate interaction strength.

**Discussion.** So far we considered \(\lambda = C = 1\) and \(t_1 = t'_1\) where the latter condition led to the KH model in the strong coupling limit. The appropriate choice of effective parameters to account for the actual scenario found in experiment is a contentious issue and still under debate [49]. From our findings, we can still make the general statement that zig zag magnetic order generically occurs in a large range of parameter space for intermediate Hubbard strength. Given the microscopic situation of Na\(_2\)IrO\(_3\), the parameter choice \(|\lambda| \ll C\) and \(t_1 < t'_1\) certainly better accounts for the orbital structure and orientation of the \(t_{2g}\) states as well as for \textit{ab initio} results obtained earlier [16, 18]. To make the model more realistic, we consider \(t'_1 = 1.5t_1\), \(\alpha = 0.4\), and \(0.8 \leq C \leq 6\) in the remainder of the paper [note that \(C\) represents the overall energy scale, see Eq. (1)]. The consequences of this generalization are summarized as follows: (i) As mentioned before, due to \(t'_1 \neq t_1\) symmetric off-diagonal \(\Gamma\) exchange is generated in the spin Hamiltonian [7]. (ii) As a result of these couplings, the ED phase diagram at \(U \to \infty\) shows an additional small 120° Néel regime between the stripy AFM phase and the Kitaev spin liquid [21]. (iii) The non-interacting phase diagram only undergoes minor shifts of the phase boundaries. (iv) In the interacting VCA phase diagram, the zigzag phase persists in the whole range of investigated \(C\) parameters.
but is shifted to larger values of $U/C$, see Fig. 3(a).

In the literature, there has been an intensive debate about what are the “correct” spin Hamiltonian and exchange couplings $J$, $K$, and $\Gamma$ for Na$_2$IrO$_3$ [8] 10,13,15,18,23,25,26. The couplings given in Eq. (5) will be modified for $t_1 \neq t'_1$ as follows,

$$J = \frac{4}{9} \left( \frac{2t_1^2 + t_1^2}{U} + \frac{2t_1(t_1 + 2t'_1)}{U - 4J_H} \right)$$ (10) 

$$K = -\frac{16J}{9U} \frac{3t_1^2 + (t_1 - t'_1)^2}{4J_H - U}$$ (11) 

$$\Gamma = \frac{32J}{9U} \frac{t_2(t_1 - t'_1)}{U - 4J_H}$$ (12)

These expressions can be used to obtain sign and magnitude of the magnetic couplings depending on the input tight-binding parameters as well as $U$ and $J_H$. In principle, it can even be applied to the analogous discussions for other “Kitaev-Heisenberg materials” [30,32].

We finally focus on the regime near the transition between the TI and zigzag AFM phase. Assuming that the experimentally observed zigzag phase is indeed not deep in the Mott phase but much closer to the weak-coupling regime [16,27], it is interesting to investigate whether remnants of the TI phase can still be detected in the zigzag phase. We compute the single-particle spectral function $A(k_x, \omega)$ on a ribbon geometry which allows to track the edge states even in the presence of interactions. Using the above parameter setting $t'_1 = 1.5t_1$, $\alpha = 0.4$ and fixing $C = 5$, Fig. 3(b-f) shows the evolution of surface spectral functions for varying $U/C$. In the TI phase for $U/C \leq 3.6$ we clearly observe helical edge states traversing the bulk gap while the magnetization remains zero. At $U = 3.7$ the second order phase transition occurs marking the smooth onset of magnetization. Once magnetic order sets in, we immediately observe a small gap in the edge states at $k_x = \pi$, which is a consequence of the lost topological protection. Even around $U/C = 4$, we still observe the gapped edge states, see Fig. 3(f). Interestingly, in a recent angle-resolved photo emission spectroscopy (ARPES) measurement on single crystals of Na$_2$IrO$_3$, metallic surface states with nearly linear dispersion have been revealed [28,53]. It remains as an exciting task to verify whether these metallic surface states are the topologically trivial remnants of the TI phase as suggested by our VCA analysis.

Conclusion. We introduced a three-band Hubbard model which unifies these two influential works in the context of Na$_2$IrO$_3$: the correlated TI phase proposed by Shitade et al. [7], the KH model advocated by Jackeli and Khaliullin [8], and the inelastic neutron scattering results of Choi et al. [11] indicating a zigzag antiferromagnetic ground state. Our three-band Hubbard model hosts for weak $U$ an extended TI phase, at strong $U$ the KH model, and at intermediate $U$ a zigzag antiferromagnetic phase.

Moreover, our analysis predicts topologically trivial edge states in the zigzag phase as a remnant of the TI phase which might have been observed in recent ARPES measurements [25].

Acknowledgements.—We acknowledge discussions with R. Valenti, N. Perkins, G. Jackeli, I. Mazin, and P. Gegenwart. We acknowledge financial support by the DFG through SFB 1143 (ML, SR) and SFB 1170 (RT), by the ERC through the starting grant TOPOLECTRICS (ERC-StG-Thomale-336012, RT), and by the Freie Universität Berlin within the Excellence Initiative of the German Research Foundation (JR). We thank the Center for Information Services and High Performance Computing (ZIH) at TU Dresden for generous allocations of computer time.