
Editorial Recommendation

Dear Authors, the first round of review has yielded two reports, that are available to you for your perusal. It is felt that important revisions may be needed, in order for the editorial college to proceed with recommending publication. Upon resubmission, the Authors are kindly requested to offer a pointed response to all the comments made by the reviewers and to provide a detailed list of all the changes implemented. Sincerely, Massimo Boninsegni Editor-in-charge

Our reply:

We thank the editor and reviewers for their handling of our manuscript. Below we address all points raised by the two reviewers, and we believe our paper is now ready for publication. We are a bit surprised about the editor's judgement that a "major revision" may be needed: Reviewer 1 has not made any recommendation (neither for or against publication), whereas Reviewer 2 has explicitly stated that they "*strongly recommend publication of the paper in its present form*".

With our detailed replies at hand, we hope that our paper can quickly be published in its present form. A complete list of all changes to the manuscript can be found at the end of this document.

Anonymous Report 1 on 2022-10-13

The authors reconsider a long-standing problem, the properties of two propagating impurities (holes, doubly occupied sites) in a Neél background. The authors propose an approximate theory based on a simplified string potential. The following remarks are not necessarily in order of relevance.

Our reply:

We thank the referee for their quick evaluation of our work in record time (<3 days!), and for their sharing their insightful remarks below. We are glad that the referee agrees that we are addressing a long-standing problem in the field.

Referee:

(a)
Strings of flipped spins break into two parts when self-intersecting. Self-intersection is neglected by the authors on the basis that the relative number of self-intersecting strings is small. This is true, the argument is nevertheless a fallacy, at least on a formal level.

Relative to its last direction of movement, a hole can move right/straight/left (R/S/L). Strings are hence sequences of R/S/L tokens, such as ...RSSLLS...

A string will self-intersect if the patterns
..RRR.. or ..LLL..
occur. The probability for this to happen is
2/3**3 = 2/27 = 7.4%, as listed in Table 1.

The probability that strings do not self-interact decays exponentially with the length of the string, roughly as (1-0.074)**(length)

This is relevant because R/S/L hoppings are associated with distinct spin configurations, which makes the propagation of holes equivalent to that of a classical particle, which is also stated by the authors. Strings ceases to exist on a formal basis once they self-intersects for the first time.

Neglecting self-intersection is hence OK for short, but not for long strings. The authors work in the limit t >> J, which implies a weak confining potential and hence long strings. This consideration suggests therefore that the approach presented is void, on a formal basis, being based on long, but not self-intersecting strings. The effect of string intersections on an effective confining

potential needs to be worked out (see next point). It is
a bit disturbing that the authors did not discuss this issue.

Our reply:

The referee is correct in pointing out that the effective string model we work with cannot be exactly mapped to the problem of two holes moving in an antiferromagnet, since the Hilbertspaces of the two problems are not isomorphic.

We emphasize this point early on and very explicitly in our manuscript: Directly at the beginning of Sec. II (Microscopic Models) we state:

"*In this article we introduce and solve an effective theory describing bound states of holes. As described in detail in Sec. III, we will make approximations on the level of both the Hilbert space and the Hamiltonian. Nevertheless, our starting point are microscopic models of doped AFM Mott insulators to which, we argue, our results apply within some approximations.* ***Critical minds should simply view these models as motivating our effective theory, although our numerical analysis in Ref. [41] indicates remarkable similarities with the semi-analytical predictions derived here.***"

In the beginning of Sec. III (Effective String Model) we re-iterate that we perform approximations on the level of the effective Hamiltonian as well as the Hilbertspace and clearly indicate that the model we analyze is an effective model.

That said, we also agree with he referee that it is important to understand when and how our approximations fail; and we agree that self-crossings of the string do lead to effects not captured by our effective model. However, we disagree with the statement that "*Strings ceases to exist on a formal basis once they self-intersects for the first time*": A string (inside a classical Néel state) with a RRR configuration in it still corresponds to a particular spin configuration which, due to the RRR loop, has a defect inside the surrounding Néel pattern. Hence this configuration corresponds to a unique state, orthogonal to other string states, to which the hole-tunneling Hamiltonian couples by a few hops. And therefore this state *has to be kept* as a unique state in the approximate basis we construct (and it does not cease to exist, if by that the reviewer means that the state has already been counted as a different state beforehand). That said, the contribution to the potential energy corresponding to the described loop does not increase linearly with the string length anymore, a further assumption we make in our model. How good this simplification is can be debated — we prefer to make it in order to make our model analytically accessible. Ultimately only a 1:1 comparison with unbiased large-scale numerics, as we performed recently in [Bohrdt et al., arXiv:2210.02322] or an experiment can clarify whether our approximations merely lead to quantitative differences or striking qualitative deviations. Our comparison to large-scale DMRG studies on 4-leg cylinders suggested the former — see [Bohrdt et al., arXiv:2210.02322], of which we reproduced Fig. 6 in Reply-Fig.1 below for the reviewers convenience.

It remains true, however, that longer loops exist for which a new spin configuration is constructed by following the string, that is already counted in the string bases. These cases correspond to Trugman loops: they require at least 6 moves of one hole, not just 3 consecutive moves as discussed by the referee. For the string lengths relevant at low energies (up to a length around l~8, see Fig. 7 in our manuscript), the fraction of such Trugman loops is indeed small, justifying our over-counting of these states. Longer strings correspond to higher energies (in the form of localization energy to special loop states in string-space, or through more ferromagnetically aligned spins, see [Grusdt et al., PRX 8, 011046 (2018)] for a more in-depth discussion of this issue). To address the comment by the referee, we have clarified in our

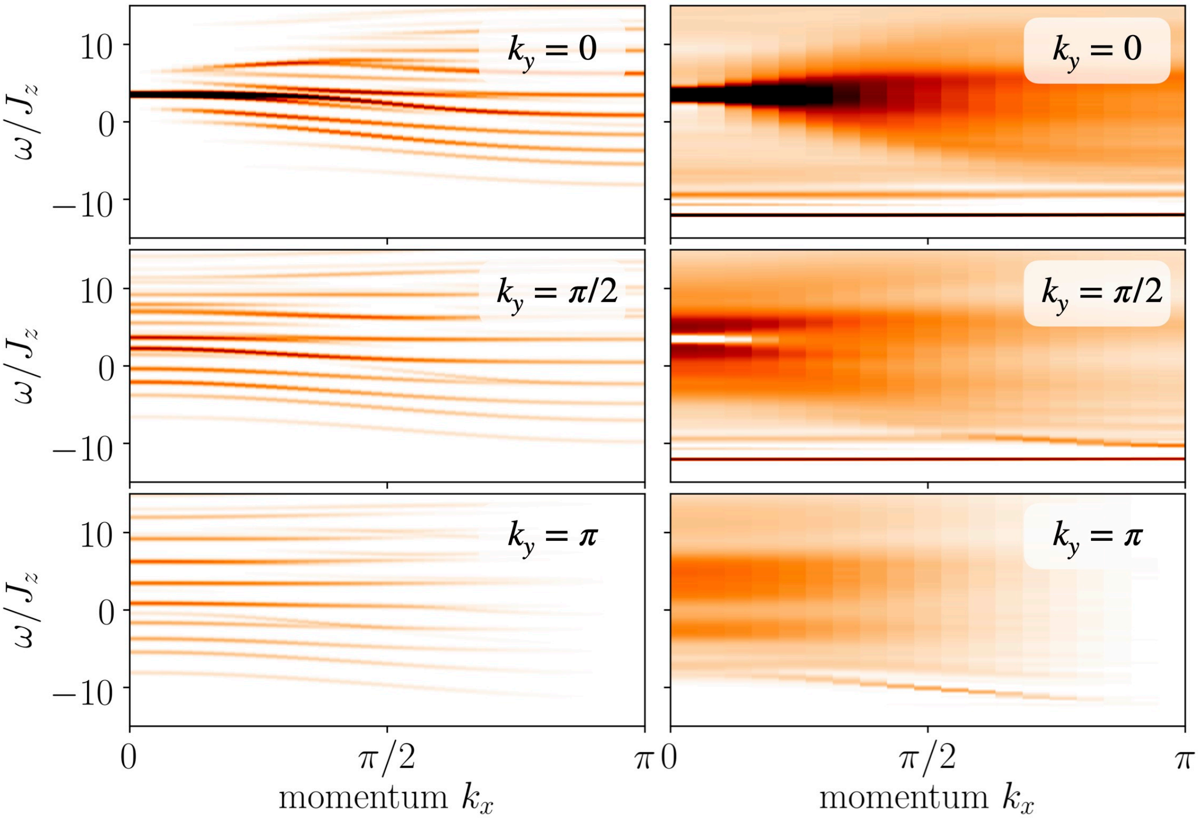


FIG. 6. **Comparison of s-wave spectra to geometric string theory** prediction, [15], left, to numerically calculated spectra (right), for the $t - J_z$ model with $t/J = 3$, $\chi = 600$, time evolution up to $T_{max}/J_z = 10$, on a 40×4 cylinder. Top, middle, bottom plots correspond to momentum $k_y = 0, \pi/2, \pi$, respectively.

Reply-Fig. 1: Re-print of Fig. 6 in [Bohrdt et al., arXiv:2210.02322], showing predictions by the effective string theory introduced in the present manuscript (left column; Ref. [15] mentioned in the caption is our present manuscript) and numerical DMRG results on 4-leg cylinders (right column). Overall, good qualitative agreement is obtained, suggesting that our approximations are accurate enough to capture the essential structure of the observed 2-hole bound states.

manuscript that Trugman loops can be ignored "*in the range of string lengths relevant to low-energy states*". In practice this means typical string lengths remain < 10.

Our discussion above highlights two important points: Firstly, the number of problematic configurations scales less unfavorably with the string length than suggested by the reviewer; Secondly, even for relatively large values of *t/J ~ 3* which we consider, the relevant string lengths at small energies remain small enough to reasonably neglect effects of loops and self-intersections of the strings.

However, we do generally agree with the reviewer that the limit *J/t* ➔ *0* when the string length in the effective model diverges, is problematic: Indeed, when strings of length *l >> 10* are considered, the string basis becomes over-complete. To understand the correct scaling of how over-complete the effective basis is, the following argument is useful: The number of string states scales as 3^ℓ when ℓ is the maximum string length; These string states reshuffle spins on an disc of radius scaling as $r \propto \sqrt{\ell}$ since the direction of individual string segments can be considered as being random; Within this disc $\sim r^2$ spins are reshuffled, whose Hilbertspace dimension thus scales as $2^{r^2} \sim 2^\ell$. Since $2^\ell \ll 3^\ell$ for long strings ℓ, we indeed expect over-completeness of the string Hilbertspace to eventually become a problem — but only beyond the regime of J/t values we are interested in (see also the discussion below).

Referee:

(b)
An equivalent problem concerns the confining potential.
The patterns ..LL.. or ..RR..
lead to adjacent string elements and hence to a correction to the confining potential, here with respect to the linear approximation used by the authors. The probability for this to occur for every two steps is substantial, 2/9 = 22.2%

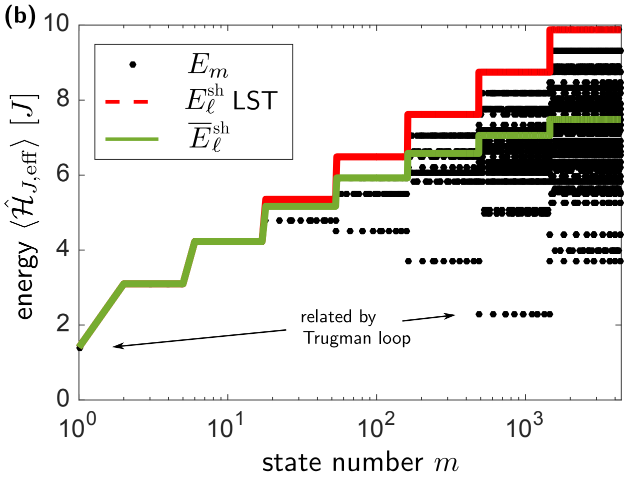
Given that string self-interactions are attractive, LL and RR token patterns will occur dynamically even with a larger probability. The same holds for the arguments in (a). The linear simplification used by the authors is hence questionable. It is a bit strange that this issue is not mentioned in the manuscript.

Our reply:

We agree with the referee that self-interactions of the strings are attractive, and that we do not capture such effects within our effective model. We politely disagree with the referee, however, that "this issue is not mentioned in the manuscript". In Sec. III.A.2. where we introduce the effective model, we explicitly state that: "*fluctuations of* V_Σ *from string to string (or site to site on the Bethe lattice) result from string-string interactions*"; Just below we point out that "*we can estimate the string-length potential* $V(l) \approx V_{LST}(l)$ *by considering only straight strings in Eq. (12). Since string-string interactions are always attractive, this linear string theory (LST) estimate also defines an upper bound for the averaged potential*".

We emphasize once more that we study an effective model, which is related but not 1:1 identical to microscopic *t-Jz* or *t-J* models, as clearly pointed out at several points, including early on, in our manuscript.

Moreover, we would like to emphasize that the linear string approximation is not expected to cause major qualitative changes (as supported by our numerical studies in [Bohrdt et al., arXiv:2210.02322], see above). To get an idea, we present here in Reply-Fig. 2 the string potentials calculated when string-states of a single hole distort the spin-background, starting from



Reply-Fig. 2: Variational energies of string states of a single hole, created by removing a spin from a 4x4 periodic Heisenberg AFM ground state and moving the hole along different string trajectories (m labels the string states, going through all string-length 0, then 1, then 2, … states as m increases). The green solid line denotes the average energy at a given string length, red corresponds to the linear string potential.

the ground state of a quantum Heisenberg AFM on a 4x4 periodic lattice (we expect similar results when a classical product Néel state is considered, and for two-hole cases). We find that the average string potential continues to increase with the string length, falling just somewhat below the linear string prediction (which, as mentioned in our manuscript, represents an upper limit). The most pronounced effect is the existence of a few very-low energy states that correspond to Trugman-loop trajectories. But as discussed above (and in detail in [Grusdt et al., PRX 8, 011046 (2018)]), occupying Trugman loops requires significant localization energy on the Bethe lattice, allowing to incorporate them perturbatively in a tight-binding calculation even if $t>J$. Therefore, and given that our goal is to obtain analytical insights into a simplified effective model, we believe our introduction of a linearized string potential in the theory does not invalidate our approach.

Referee:

(c)
The authors claim that their approach works both for the t-J_z and the t-J model, as long as t>>J_z, respectively t>>J. This claim seems blatantly wrong. The probability that J_xy fluctuations disrupt the string increases exponentially with string length, the presumed regime of validity of the proposed approach. Strings cease to exist altogether, at least in their naive form. It is worrisome that the authors did not mention this well-known problem.

Our reply:

First of all, we agree with the reviewer that our effective model cannot accurately describe the underlying physics of pairs of holes in the extreme limit $t \ggg J$, but for a different reason than mentioned by reviewer: In this limit, the spin environment becomes polarized when the Nagaoka effect starts to set in, and the total spin of the ferromagnetic dressing cloud surrounding the hole(s) grows with increasing t/J. However, the transition to this Nagaoka regime sets in only for $t/J \gtrsim 25$, according to large-scale DMRG simulations by White & Affleck [PRB 64, 024411 (2001)]. Following the remark by the reviewer, we have now added a clarifying comment in our manuscript stating that "*we do assume throughout, however, that* $t/J \lesssim 25$*before the Nagaoka ferromagnetism arises for even larger values of t/J*".

Second, we respectfully disagree with the statement by the referee that transverse spin fluctuations J_xy become detrimental in the regime of long strings, or "cease to exist altogether". One can take two different perspectives on this problem, as detailed next. The first applies for higher-energy string states, which can decay through string breaking in a mechanism akin to the celebrated Schwinger mechanism. The second is focusing on low-energy states, where dressing with virtual magnons can occur beyond the physics of strings; Let us provide more details:

1) High-energy perspective: If one starts from an excited string state, where the string is longer than in the ground state, a natural process that can occur is string breaking: The spin-flip processes can cut the string in two, and two holes connected by a long string can decay into two separate shorter strings, each connecting one hole to a spinon-type defect. This effect is akin to the celebrated Schwinger mechanism occurring when two quarks are separated sufficiently far from one another. We agree that, intuitively, the matrix elements for such string decays increase with increasing string lengths, and it is indeed expected that for high-energy string states this is a

dominant decay mechanism. However, for the lowest energy string states, the situation is different: If there are no lower-energy final states that the system can decay to, the string remains stable — or if the available phase space of possible final eigenstates is small, which is the case at the lowest energies, the string states can be stable of at least meta-stable excitations of the system, much like many elementary particles in nature have a finite life-time.

Our numerical DMRG simulations in [Bohrdt et al., arXiv:2210.02322] show well-defined quasiparticle peaks in the spectrum at low-to-intermediate energies, which we take as a strong sign that the described meta-stable string states actually exist. In any case, it is meaningful to calculate the structure of the (meta-) stable states, as we do in our paper, which can be used as a valuable starting point later on to estimate their life-times from the above-described decay processes. Such detailed calculations will be subject of future work and would go significantly beyond the scope of our current paper.

2) Low-energy perspective: In an intuitive semi-classical picture it appears logical that longer strings give more space for spin-flip terms to 'disrupt' the string, and we expect that the reviewer is arguing that every given bond along the string contributes a certain probability p for a string-breaking to occur — hence the probability for the string to survive would be scaling as $(1-p)^{\ell}$, i.e. exponential in ℓ.

However, this logic is too simple: The point is that the mobile dopant is constantly creating *and retracing* the string. This means, following creation of any given string segment, there is only limited time for a spin-flip process to occur before the charge returns, restoring the original undisturbed spin background. The question we need to ask is whether the fast hole motion can adiabatically follow its slowly reacting spin-background. Putting numbers on this shows that — counterintuitively — longer strings become *increasingly more robust* against string-breaking:

The time it takes for a given bond to flip is given by $\tau_J = 1/J$; the time it takes the mobile charge to create and retrace the string up to that bond scales as $\tau_\Sigma = \ell_\Sigma \tau_t$, where ℓ_Σ is the typical length of the string and $\tau_t = 1/t$ is the characteristic hopping time. Since the characteristic length of the string in the (approximately) linear string potential scales as $\ell_\Sigma \simeq (t/J)^{1/3}$, we arrive at: $\tau_J/\tau_\Sigma \simeq \left(\frac{t}{J}\right)^{2/3}$. Hence for $t \gg J$ we expect $\tau_J > \tau_\Sigma$ — i.e. it would take longer for the spin-flip to occur than healing the considered part of the string. I.e., longer strings should be *more robust* against string-breaking, because of their strong fluctuations (for a static string, we completely agree with the reviewer's intuition however: longer strings become exponentially more short-lived in this case).

What the above argument suggests is that coherent superpositions of string states should become well-defined when $t \gg J$: namely, *the hole motion can adiabatically adjust to it spin-background*. This is not to say that spin-flip processes as described by the referee are completely inconsequential! Rather, they can be included on top of the fast hole motion that creates the string: We can first calculate the string, which describes a coherent superposition of various hole-trajectories. By the separation of time scales, we can then average over these different hole-trajectories in the string basis to calculate how the spin-environment will be affected. This, in turn, would slightly modify the string potential, and we could go and perform a refined calculation of the best effective string Hamiltonian. This procedure could be repeated to reach convergence. So far, in our manuscript, we only performed the first step in this calculation, which is closest to the t-Jz model. However, we are currently developing a formalism to perform the refined calculation and describe additional modifications of the spin background quantitatively, using the so-called generalized 1/S expansion, the basic idea of which we introduced already in [Grusdt et al., PRX 8, 011046 (2018)]. As long as the matrix elements leading to this additional dressing of the string with low-energy magnon excitations remains small, the string picture should provide an accurate representation of the essential physics.

So to summarize, we fully agree with the reviewer that studies of the spin-flip terms are interesting and important. Our above scaling analysis is reassuring, demonstrating that the fast hole motion can adiabatically adjust to a slowly varying spin background — i.e. string states remain well defined objects. Going into further detail would go beyond the scope of this reply and/or our manuscript, and will be subject of extensive future research.

Finally, we note that restricting the values of t/J to below 25 also limits the relevant string lengths. As mentioned above, in the (approximately) linear string potential the average string length scales as $\ell_\Sigma \simeq (t/J)^{1/3}$. Since we obtain $\ell_\Sigma \simeq 3$ when t/J =3 (see Fig. 7 in our manuscript), by this scaling we expect that a typically value of only $\ell_\Sigma \simeq 6$ would be reached for t/J =25. This reassures us that the use of an approximate string basis in this regime (which is not the physically most relevant) may still be reasonable on a qualitative level.

Referee:

(d)
The transformation from the Hubbard to the t-J model leads to an intra-sublattice correlated hopping term, let's call it here T_2. The prefactor is J. It allows for coherent intra-sublattice propagation, which contrasts qualitatively with the bare NN hopping. No strings are generated by T_2. Without a word, the authors disregard T_2, keeping only the bare J term. This can be valid for suitable ion-trap experiments, but most probably not for solid-state applications. In view of the amusing phrasing:

"The system most closely related to our effective theory is constituted by the 2D t-J_z model on a square lattice."

the authors may argue that their theory is any case only somewhat remotely related to real-world physics. But leaving out a term that would add qualitative new features, and possibly invalidate the entire formalism, needs supporting arguments.

Our reply:

As the reviewer already points out, we are studying an effective model with the goal to capture the key qualitative aspects of tightly bound pairs of holes connected by a string. We completely agree that it is an interesting question, which needs to be addressed in future work, how a next-nearest neighbor (NNN) hole-hopping term affects this picture: Such NNN terms can arise directly on the level of the microscopic one-band or three-band Hamiltonian, or as the referee describes when going from a one-band Hubbard model to a simpler t-J model.

In our manuscript, we choose not to discuss such terms for several reasons: (1) We want to first understand the qualitative physics of the ideal case with strings, which is why we do not claim that our effective model is 1:1 related to the various microscopic models discussed in the context of high-Tc; (2) We want to avoid any discussion which model captures high-Tc most accurately — this is an interesting and important question, which needs far more space to discuss properly (1-band Hubbard model vs. 3-band Hubbard model vs. models with phonons vs. t-J model vs. even higher-order ring-exchange terms, etc etc); (3) As long as the NNN hopping terms are small compared to the NN tunneling t, we do not even expect strong qualitative modifications. This can be argued either just perturbatively (where the energy gaps between different bound states of holes connected by a string provide the relevant energy scale to compare to), or by again evoking additional magnon dressing: in the generalized 1/S expansion which we are currently implementing, the NNN hopping process can be captured by formally keeping the string but dressing it with additional magnon excitations describing the modified spin configuration around the string. Such additional magnon dressing will surely lead to corrections of the energy, but these are again expected to be small as long as the NNN hopping terms do not become dominant.

A detailed treatment and discussion of these additional terms will be subject of future research but goes beyond the scope of the present manuscript.

Referee:

(e)
The two end-impurities of a string are governed by identical dynamics. It would have hence been intuitive to use double-sided stacks for the encoding of strings. The authors opted instead for an asymmetric formulation that needs to be symmetrized in a second step. Is there a rational for the choice of an asymmetric encoding?

Our reply:

This is a very interesting question raised by the referee, and indeed Shraiman & Siggia have chosen a double-sided encoding of strings in [PRL 60, 740 (1988)]. Our one-sided stacking (with subsequent (anti-)symmetrization) has the advantages that (1) the discrete C4 and C3 angular momenta could be more straightforwardly included in the formalism (at least for us, having worked with spinon-chargon strings before), which we use to truncate our effective string basis; and (2) we found it more intuitive since there is no need in our formalism to single out any particular ‘center’ of the strongly fluctuating string around which to define the stacks of strings.

That said, we agree with the reviewer that a double-sided stacking can also be used. From such an approach we expect different results, since basis truncation would need to be performed differently, but we don’t want to speculate which approach is more accurate in the end.

Anonymous Report 2 on 2022-11-27

In the present manuscript the authors attempt to tackle a (at least) three decades old problem of hole pairing in quantum antiferromagnets. They consider the $t - J$ and $t - J_z$ models and study the role of Brinkman-Rice strings in the pairing mechanism. The idea that holes in quantum antiferromagnets bind through the strings has been examined by a few authors in the past where it was found that $d_{x^2-y^2}$ pairing is favored in this case. It is clear that the present paper is solid work and sheds additional light in this problem. Using an effective model of partons connected by a confining string they calculate the spectral properties of bound states.

Our reply:

We thank the reviewer for their very positive evaluation of our manuscript, and for concluding that it is solid work which sheds additional light on a long-standing problem.

Referee:

I have only one comment for the authors' consideration which is not a condition for publication. This work as some others (especially within the $t - J_z$ model) predict heavy states of immobile pairs with flat-band dispersions. However, it is well-known that quantum spin-fluctuations erase the origin of the string to which each hole is attached, through a pair of hole-hoppings with an additional spin-pair flip. This mechanism is the one which gives rise to a finite bandwidth to the single-hole dispersion, i.e., of the order of t^2/J in the $t - J$ model. This mechanism allows the holes to keep their kinetic energy low while bound to one-another which opens the door for bound-states of mobile pairs.

Our reply:

We thank the reviewer for this comment. We agree that quantum spin-fluctuations lead to a non-vanishing dispersion of what used to be completely flat bands of pairs of holes. We have numerically analyzed and confirmed this effect using large-scale DMRG simulations on 4-leg cylinders in our accompanying paper [Bohrdt et al., arXiv:2210.02322]. Since we focus on the effective toy model in the present manuscript (where such spin-flip processes are not yet included), we decided not to add a more detailed discussion to the text for now.

But let us add two more comments: First, the reviewer mentions a situation where each hole is attached to a string with a given origin. This corresponds to a scenario where one starts from two magnetic or spin polarons, each with its own spin, which can then form a bound state (indeed such bound states of weakly interacting magnetic or spin polarons where described in [Brügger et al., PRB 74, 224432 (2006)], to which we now added a reference in the outlook of our paper). These pairs of polarons have a very different character, however, than the tightly-bound two hole states we discuss here: In our case, the string connects one hole to the other hole, and no additional origin is left. Nevertheless we fully agree that spin fluctuations can lead to fluctuations of the string itself, which leads exactly to the scenario described by the reviewer, namely that the flat band becomes weakly dispersive.

Second, the reviewer makes an argument that the bandwidth of the so-obtained weakly dispersing state should scale as t^2/J. We agree with this scaling only in the perturbative limit when $t < J$, but note that our paper focuses on the strong-coupling regime where roughly $1 < t/J < 25$. Hence we expect a different scaling of the weakly dispersive band, to leading order independent of t: In order to move a segment in the middle of a string, first a transverse spin-fluctuation

$J_\perp(S^xS^x + S^yS^y)$ is required to break up the string. The intermediate state has an energy $\propto J_z$; repairing the string requires another spin-flip $\propto J_\perp$. Hence, perturbatively in $J_\perp/J_z$ we expect the bandwidth to scale as $W \propto J_\perp^2/J_z$ — which is exactly what we numerically found in [Bohrdt et al., arXiv:2210.02322] for the t-XXZ model. When $J_\perp = J_z = J$, we numerically found $W \simeq J$ on the order of *J*.

Finally, in the limit $t < J$ (which we do not consider), the distinction between pairs of magnetic or spin-polarons on one hand, and the tightly bound pairs of holes with just one string, becomes ill-defined. In that limit we fully agree with the arguments by the reviewer.

Referee:

In conclusion, I believe that the present work is a significant contribution to this important outstanding problem and I strongly recommend publication of the paper in its present form.

Our reply:

We thank the reviewer again for their time, and we are in particular delighted about their enthusiastic recommendation for publication of our work in its present form.

List of changes:

- In the introduction we have added a reference to [Mezzacaop et al., PRB 94, 155120 (2016)];
- In the outlook we have added references to [Brügger et al., PRB 74, 224432 (2006)] and [Danilov et al., npc Quantum Materials 7, 50 (2022)];
- Following the comment by Reviewer 1, we have added a clarifying statement that loop configurations of strings can reasonably be neglected in the relevant range of string lengths considered (below 10)
- Following the comment by Reviewer 1, we have added a clarifying statement that we only consider values of $t/J < 25$, before the Nagaoka ferromagnetism sets in (as demonstrated by White & Affleck in PRB 64,024411 (2001).)