

Controlled expansion for certain non-Fermi-liquid metals

David F. Mross, John McGreevy, Hong Liu, and T. Senthil

Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(Received 9 March 2010; revised manuscript received 26 May 2010; published 26 July 2010)

The destruction of Fermi-liquid behavior when a gapless Fermi surface is coupled to a fluctuating gapless boson field is studied theoretically. This problem arises in a number of different contexts in quantum many-body physics. Examples include fermions coupled to a fluctuating transverse gauge field pertinent to quantum spin-liquid Mott insulators, and quantum critical metals near a Pomeranchuk transition. We develop a controlled theoretical approach to determine the low-energy physics. Our approach relies on combining an expansion in the inverse number (N) of fermion species with a further expansion in the parameter $\epsilon = z_b - 2$, where z_b is the dynamical critical exponent of the boson field. We show how this limit allows a systematic calculation of the universal low-energy physics of these problems. The method is illustrated by studying spinon Fermi-surface spin liquids, and a quantum critical metal at a second-order electronic nematic phase transition. We calculate the low-energy single-particle spectra, and various interesting two-particle correlation functions. In some cases, deviations from the popular random-phase approximation results are found. Some of the same universal singularities are also calculated to leading nonvanishing order using a perturbative renormalization-group calculation at small N extending previous results of Nayak and Wilczek. Implications for quantum spin liquids and for Pomeranchuk transitions are discussed. For quantum critical metals at a nematic transition, we show that the tunneling density of states has a power-law suppression at low energies.

DOI: [10.1103/PhysRevB.82.045121](https://doi.org/10.1103/PhysRevB.82.045121)

PACS number(s): 71.27.+a, 71.10.Hf

I. INTRODUCTION

This paper is concerned with the destruction of Fermi-liquid behavior in two-dimensional systems where a gapless Fermi surface is coupled to a fluctuating gapless boson field. This problem arises in a number of different contexts in quantum many-body physics. A well-known example is where the fermions are coupled to a gapless transverse gauge boson. This describes the low-energy effective theory of certain quantum spin-liquid phases,^{1,2} the theory of the half-filled Landau level,³ and various non-Fermi-liquid metallic phases.⁴⁻⁸ A different and equally well-known example is as a description of quantum critical metals at a ‘‘Pomeranchuk’’ instability.⁹ The classic example is the Stoner transition associated with the onset of ferromagnetism in a metal. In recent years, attention has focused on a different example of a Pomeranchuk transition that associated with the onset of electronic nematic order¹⁰⁻¹⁹ from a Fermi-liquid metal. Here electronic nematic order means a phase where the lattice point-group symmetry but not translation symmetry is broken. Such order has been observed with increasing frequency in a number of different correlated metals²⁰⁻²⁵ giving rise to an interest in the associated quantum phase transition. At such a quantum phase transition, the nematic order parameter is described as a gapless fluctuating Bose field, and its coupling to the gapless Fermi surface destroys Fermi-liquid behavior.^{11,13,14}

The purpose of this paper is to formulate a controlled theoretical approach to this class of problem, the need for which has been emphasized recently.^{26,27} The low-energy physics of the resulting non-Fermi-liquid metal is characterized by universal scale invariant behavior. Our approach provides a systematic method of calculating the exponents and other universal properties associated with this scale invariant behavior. We illustrate this by studying many physical prop-

erties of the gauge-field model and of the nematic quantum critical metal in detail.

Quite generally the low-energy physics of problems of this sort is conveniently described by restricting attention to fermionic modes in the immediate vicinity of the Fermi surface, and the long-wavelength, low-frequency modes of the fluctuating boson field. The model is described by the Euclidean action,

$$S = S_f + S_{int} + S_a, \quad (1)$$

$$S_f = \int_{\vec{k}, \omega} \bar{f}_{k\alpha} (-i\omega - \mu_f + \epsilon_{\vec{k}}) f_{k\alpha}, \quad (2)$$

$$S_{int} = \int_{\vec{k}, \omega} a(\vec{k}, \omega) O(-\vec{k}, -\omega), \quad (3)$$

$$S_a = \int_{\vec{k}, \omega} \frac{1}{e^2} k^2 |a(k, \omega)|^2. \quad (4)$$

Here $f_{k\alpha}$, $\alpha = 1, \dots, N$, is a fermion field with N possible flavors and a is the boson field. In the gauge model a is the transverse component of a $U(1)$ gauge field, and $O(\vec{x}, \tau)$ is the transverse component of the current density of fermions. At a nematic quantum critical point, a will be taken to be the nematic order-parameter field, and $O(\vec{k}, \omega)$ is the fermion bilinear with the same symmetry. For instance, on a two-dimensional square lattice with lattice constant ℓ a uniform nematic order parameter couples to $N^{-1/2} \sum_{\vec{k}} [\cos(k_x \ell) - \cos(k_y \ell)] \bar{f}_{k\alpha} f_{k\alpha}$.

Much prior work of course exists on this problem. In a number of early papers,^{3,4,28-30} the problem was analyzed in a random-phase approximation (RPA) and various related approaches. This showed that the fermions and gauge bosons

stay strongly coupled in the low-energy limit. In the RPA, the boson propagator is overdamped due to Landau damping by the gapless Fermi surface. The fermion self-energy has a power-law non-Fermi-liquid frequency dependence. Further the long-wavelength density response function retains its Fermi-liquid form.

In the gauge-field problem, some of the RPA results were further substantiated³¹ through a quantum Boltzmann approach which considered the fate of various possible shape fluctuations of the Fermi surface. Smooth shape deformations of the Fermi surface (which determine long-wavelength density response and the gauge propagator) were shown to retain Fermi-liquid behavior while “rough” deformations have the potential to be non-Fermi-liquidlike. The latter determine the behavior of the single fermion Green’s function and the structure of the $2K_f$ singularities (i.e., at wave vectors connecting antipodal tangential portions of the Fermi surface) in response functions. These main results were further supported in comprehensive diagrammatic analyses³² of the model which suggested that the leading RPA answers for many quantities were in fact exact in the low-energy limit. In particular, the structure of the gauge-field propagator, the fermion self-energy, and the long-wavelength density response were argued to have the same form as the RPA result. Similar diagrammatic analyses with the same conclusions have also been reached¹⁴ for the nematic quantum critical point. In the gauge-field case, the $2K_f$ singularities in the density response function were argued to have specific non-Fermi-liquidlike power-law forms.³²

Is there a controlled limit in which the reliability of these results may be assessed? One attempted approach^{32,33} is to take the limit of N (the number of fermion species) large, and expand in powers of $1/N$. In the early work, it was argued³³ that at low energies in the large- N limit only patches of the Fermi surface with parallel normals are strongly coupled to each other. Any such patch couples strongly to a boson whose momentum is perpendicular to the normal to the Fermi surface. The low-energy physics is therefore correctly described by focusing attention on patches with parallel normals.

In some remarkable recent work, Sung-Sik Lee²⁶ reexamined the model of N fermion species coupled to a $U(1)$ gauge field in the large- N limit. He showed that even at large- N the theory remains strongly coupled, and that its solution requires nontrivial summation of an infinite number of Feynman diagrams. When only a single patch of the Fermi surface is considered, a book-keeping device was introduced to show that the $1/N$ expansion could be organized in terms of the genus of the surface in which the Feynman diagrams were drawn. Based on this the general validity of the physical picture built up by RPA and the other earlier analyses for small N has been questioned.

Even more recently in another very interesting paper Metlitski and Sachdev²⁷ studied the fate of the theory with both a Fermi-surface patch and its antipodal partner included. This is believed to be fully sufficient to correctly describe the asymptotic low-energy physics of the system. They found a number of further difficulties with the large- N expansion. Specifically higher loop corrections for the gauge propagator involved higher powers of N than the leading-

order one-loop RPA result. This unpleasant finding led them to question the existence of a well-defined large- N limit to control the theory. These authors also showed that in a perturbative loop expansion the self-energy acquires singular momentum dependence at three-loop order. However the loop expansion has no apparent control parameter.

In light of these results, it becomes important to search for alternate reliable methods to judge the validity of RPA and other diagrammatic approaches to the problem. In this paper, we introduce a controlled expansion to determine the low-energy physics of this model. We consider a family of models where the “bare” boson action is modified to

$$S_a = \int_{\vec{k}, \omega} \frac{|\vec{k}|^{z_b-1}}{e^2} |a(\vec{k}, \omega)|^2. \quad (5)$$

The number z_b (the boson “dynamical critical exponent”) equals 3 in the original model in Eq. (1). The case $z_b=2$ arises in the theory of the half-filled Landau level with long-range $1/r$ Coulomb interactions between the electrons,³ and in the theory of the bandwidth-controlled Mott transition of the half-filled Hubbard model developed in Ref. 34. We show that the large- N expansion can be controlled in the limit of small $\epsilon = z_b - 2$. Specifically we show that the limit $N \rightarrow \infty, \epsilon = z_b - 2 \rightarrow 0$ such that ϵN is finite leads to reliable answers for the low-energy behavior of the system³⁵ When can one apply this technique of introducing a control parameter by varying a dynamical exponent? We expect it to be effective when one has prior knowledge that the exponent in question is not renormalized, as is the case here. Varying the exponent so that the relevant term in the effective action is nonlocal will guarantee that the exponent will not be renormalized in perturbation theory since integrating out short-wavelength modes cannot generate the necessary IR singularity. A contrasting example to the one discussed here is the Ising fixed point: the magnetization field indeed acquires an anomalous dimension, and one would obscure the physics by artificially preventing this. We demonstrate that the RPA answers for the fermion and boson propagators are indeed exact in this limit. A systematic expansion in powers of $1/N$ is possible for small ϵ . Deviations from RPA emerge at higher orders in the $1/N$ expansion. Furthermore, differences between the gauge model and the nematic critical point also appear. At order $1/N^2$, we find a singular momentum-dependent correction to the fermion self-energy—however, in the gauge model, this singularity is subdominant to the leading-order momentum dependence so that the fermion self-energy retains its RPA form, at least to this order. Further we calculate the exponent characterizing $2K_f$ particle-hole singularities, and show in the gauge model that (depending on the value of ϵN), they may be enhanced compared to a Fermi liquid. For a quantum critical metal at a nematic transition, the fermion propagator again retains its RPA form at leading order but at $o(1/N^2)$ acquires a singular correction to the self-energy that dominates over the RPA form. This modification from the RPA is in accord with the calculation of Ref. 27 but is now performed in a controlled expansion. A further difference with the gauge-field problem is in the structure of the $2K_f$ singularities. We present calcu-

lations and physical arguments that show that the $2K_f$ singularities are weakened at the nematic quantum critical point compared with a Fermi liquid.

A crucial physical ingredient that determines the low-energy physics is the nature of the ‘‘Amperean’’ interaction between two fermions that is mediated by the boson field. The term Amperean is appropriate for the gauge-field case where the interaction is between fermion currents but we will use it to describe the nematic transition as well. In the gauge-field case, the currents of a particle in one patch of the Fermi surface are parallel to that of a hole in the antipodal patch. By Ampere’s law, the gauge-mediated interaction between such a ‘‘ $2K_f$ ’’ particle-hole pair is attractive. In contrast, the currents of two particles with one from either patch are antiparallel and the gauge-mediated interaction is repulsive in the particle-particle (Cooper) channel. The situation is reversed in the nematic model. We show how this difference between the particle-hole and Cooper channel interactions plays an important role in many aspects of the low-energy physics.

Previously Nayak and Wilczek³⁶ studied the low-energy physics of the gauge-field model for small ϵ , and finite N using perturbative renormalization-group (RG) methods. The one-loop beta function for the coupling constant e takes the form

$$\beta(e^2) = \frac{\epsilon}{2}e^2 - \frac{c}{N}e^4, \quad (6)$$

where c is a positive constant. This leads to a perturbatively accessible non-Fermi-liquid fixed point for $N\epsilon$ small and positive. Some properties of this fixed point were calculated in Ref. 36 and shown to be consistent with the RPA analysis. Calculations with this one-loop beta function are able to provide answers for the scaling exponents to order ϵ for any N . In contrast, our approach of directly solving the theory at large- N enables us to extract exponents that are high order or even nonperturbative in ϵ . Thus, for instance, the deviations from RPA (subdominant in the gauge-field model) discussed above are expected to appear in the ϵ expansion only at order ϵ^3 , and so do not show up in the $o(\epsilon)$ calculations. In the regime where they overlap (i.e., to order ϵ at fixed large N), we will show that the exponent values for many properties calculated within our approach agree with those obtained from the perturbative RG. As it does not seem to be available in the literature, we calculate the $2K_f$ exponent at this small- ϵ , finite- N fixed point. We show that the exponent has an interesting nonanalytic dependence on ϵ for small ϵ which leads in the gauge model to an enhancement of the $2K_f$ singularities compared with a Fermi liquid in this limit. In the nematic model, we find a suppression of the $2K_f$ correlations as expected on general physical grounds that we also discuss.

An important feature of the low-energy physics is that the Fermi surface is preserved and is sharp even though the Landau quasiparticle is destroyed. Further at low energies and for momenta close to the Fermi surface the fermionic spectrum is scale invariant. This was already implied by the RPA results, and survives in our treatment. A similar picture was also argued³⁷ to describe continuous phase transitions where an entire Fermi surface disappears (such as a continuous

Mott transition). Following Ref. 37 we will refer to this as a critical Fermi surface. Some (though not all) aspects of critical Fermi surfaces associated with Mott-type transitions may be expected to be shared with the Pomeranchuk transitions discussed in this paper. It is therefore useful to consider these results in terms of a general scaling form expected for fermions with a ‘‘critical Fermi surface.’’ We write for the fermion Green’s function,

$$G(\vec{K}, \omega) \sim \frac{c_0}{|\omega|^{\alpha/z}} g_0 \left(\frac{c_1 \omega}{k_{\parallel}^z} \right). \quad (7)$$

Here k_{\parallel} is the deviation of the momentum from the Fermi surface. Note that the z that enters this scaling equation is the ‘‘fermionic’’ dynamical critical exponent. For the problems studied in this paper, the RPA gives $z = \frac{z_b}{2}$ and $\alpha = 1$. The latter is a result of the absence of any singular momentum dependence in the self-energy in RPA. Our results may be viewed as a calculation of α and z within a systematic expansion. In the gauge model to $o(1/N^2)$, these exponents do not change as far as the leading singular structure is concerned. For the nematic critical point, we find that $\alpha = 1 - \eta_f$ with η_f positive.

Following the general discussion in Ref. 37, we show that the difference from the RPA result has direct and measurable consequences for the electron single-particle tunneling density of states $N(E)$ (where E is measured from the chemical potential) at the nematic quantum critical point. Within the RPA, the tunneling density of states is a constant at the Fermi level. However, beyond RPA there is a power-law suppression of $N(\omega)$,

$$N(\omega) \sim |\omega|^{\eta_f/z}. \quad (8)$$

The exponent η_f is calculated in Sec. IV. Extrapolation of the leading-order results to $z_b=3, N=2$ gives the estimate $\eta_f \approx 0.3$.

What about the fate of the large- N limit when z_b is not close to two? We suggest that recent calculations of Ref. 27 should be interpreted as an instability toward translational (and possibly other) symmetry breaking in this limit. In the nematic context, this means that there is no direct second-order quantum phase transition associated with nematic ordering in two dimensions if N is sufficiently large. Instead the transition is preempted by the appearance of density wave and possibly other orders. So if for the physical case $N=2$, there is a direct nematic transition, then it is not usefully accessed by the large- N expansion. Our approach of combining the large N with a small z_b-2 or a direct small z_b-2 perturbative RG then become the only available methods to theoretically access such a quantum critical point directly in two dimensions. Similar phenomena also happen in the gauge-field model—in the large- N limit, we propose that the uniform state is unstable to translation symmetry breaking.

It is instructive to consider the behavior of the model in a two-dimensional plane spanned by z_b and $1/N$. We show our proposed ‘‘phase diagram’’ in Figs. 1 and 2. It is clear that the approach developed in this paper is ideally suited to describing the gauge model or the nematic quantum critical point for $N=2$ if it is not part of the unstable region, i.e., if Fig. 1

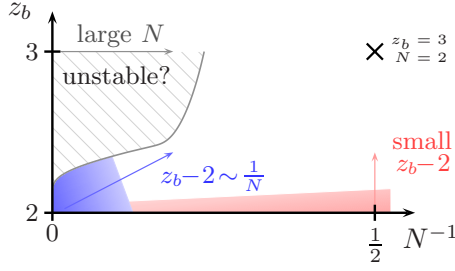


FIG. 1. (Color online) Our suggested phase diagram. Above the indicated curve, the putative critical theory is likely preempted by some other broken-symmetry state. The behavior of the proposed phase boundary at small N is one possible extrapolation.

is realized. If on the other hand $z_b=3, N=2$ belongs to the unstable region as depicted in Fig. 2, we may still hope that the approach in this paper is useful in describing the physics at temperatures above the onset of the instability.

The rest of the paper is organized as follows. In Sec. II, we begin with some preliminaries and briefly discuss the patch construction for the Fermi surface that is used in the rest of the paper. Some subtle but important aspects of the patch construction are relegated to Appendix A. Then in Sec. III, we warm up by studying the theory of just one patch and ignoring any coupling with the other antipodal patch, and show how our expansion provides a controlled answer in this simplified problem. We then study the full two-patch theory in Sec. IV and determine the singular structure of the boson and fermion propagators. In Sec. V, we present a calculation of the exponent characterizing $2K_f$ singularities within our approach. Next in Sec. VI, we explore the connections with the perturbative RG calculations of Ref. 36 and extend their results to $2K_f$ singularities. In Sec. VII, we discuss simple physical interpretation of the results of the calculations and their consequences. Section VIII describes our suggestions on a possible phase diagram. We conclude in Sec. IX with a general discussion on how our results fit in with various other related problems and theoretical descriptions of non-Fermi-liquid metals. Various appendices contain details of calculations.

II. PRELIMINARIES

As mentioned above, the low-energy physics is correctly described by focusing attention on Fermi-surface patches

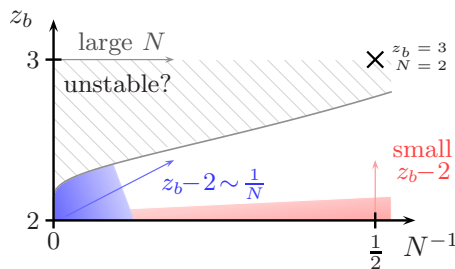


FIG. 2. (Color online) An alternate possible phase diagram. Here the interesting point $z_b=3, N=2$ is in the unstable regime; in this case, it would be best accessed starting from the correct mean-field theory for the new broken-symmetry state.

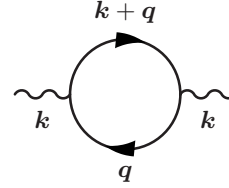


FIG. 3. One-loop boson self-energy.

with parallel normals.^{8,27,32,33} This is because the interactions mediated by the boson field are predominantly small-angle scattering processes. Furthermore, short-range four fermion interactions that couple different patches become unimportant at low energies^{8,32,33} as can be checked a posteriori after the two-patch theory is solved. Thus the universal low-energy physics of the system is correctly captured by a theory that focuses on two opposite patches of the Fermi surface. We focus henceforth on two opposite patches of Fermi surface; there are a number of subtle and important points related to the patch construction that we elaborate on in Appendix A. The patch construction also has a number of immediate consequences for the behavior of many physical properties. These will be discussed in Sec. VII.

Consider patches of the Fermi surface with normals along $\pm x$. We will denote the corresponding fermion fields $f_{R/L}$, where R denotes the right patch and L the left one. It is useful to begin by considering the boson and fermion propagators in perturbation theory keeping just the leading one-loop diagrams (Figs. 3 and 4). The imaginary frequency boson propagator $D(\vec{k}, \omega)$ becomes

$$D(\vec{k}, \omega) = \frac{1}{\gamma \frac{|\omega|}{|k_y|} + \frac{|k|^{z_b-1}}{e^2}} \tag{9}$$

with³⁸ $\gamma = \frac{1}{4\pi}$. Unless otherwise mentioned, we will henceforth set $e=1$. The fermion propagator is determined by its self-energy which at one-loop level takes the form

$$\Sigma = -i \frac{1}{\lambda N} \text{sgn}(\omega) |\omega|^{2/z_b}. \tag{10}$$

The constant λ is given by

$$\lambda = 4\pi \sin \frac{2\pi}{z_b} \gamma^{(z_b-2)/z_b} \tag{11}$$

and thus vanishes linearly as $z_b \rightarrow 2$. In terms of the scaling form in Eq. (7), this implies the fermionic dynamical critical exponent $z = \frac{z_b}{2}$ and $\alpha=1$ as promised.

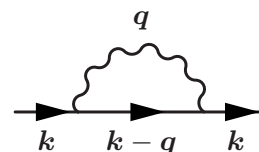


FIG. 4. One-loop fermion self-energy.

The arguments of Ref. 26 show that a minimal Euclidean action that enables correct description of the low-energy physics is given by

$$S = S_f + S_{int} + S_a, \quad (12)$$

$$S_f = \int d^2x d\tau \sum_{sa} \bar{f}_{s\alpha} (\eta \partial_\tau - is \partial_x - \partial_y^2) f_{s\alpha}, \quad (13)$$

$$S_{int} = \int d^2x d\tau \frac{S}{\sqrt{N}} a \bar{f}_{s\alpha} f_{s\alpha}, \quad (14)$$

$$S_a = \int_{\vec{k}, \omega} |k_y|^{z_b-1} |\vec{a}(\vec{k}, \omega)|^2. \quad (15)$$

Here $s=+1$ for the patch R and -1 for the patch L . The parameter η is taken to be small and positive. The field a represents just the x component of the vector field a_i . Indeed, it is just this component that couples strongly to the patches with normals along $\pm x$. Note, in particular, that the boson field couples with *opposite* sign to the two antipodal patches. If on the other hand, we were interested in the critical theory for a Pomeranchuk transition (such as a transition to an electronic d -wave nematic state in a two-dimensional metal which microscopically has square lattice symmetry), the minimal action will have a very similar form except that the boson will couple with the *same* sign to antipodal patches. While this difference is unimportant for some properties, it plays a crucial role in others. For instance, the structure of the $2K_f$ singularities is completely altered between the gauge-field and nematic models.

III. ONE-PATCH THEORY

We begin by focusing attention only on one patch, say the right one, and completely ignoring the other one. Indeed, Ref. 26 showed that the standard large- N expansion leads to an apparently strongly coupled theory already in this simplified model. The main point is that a high loop diagram may formally look like it is high order in the $1/N$ expansion. However, for many such diagrams the corresponding loop integral diverges in the $\eta \rightarrow 0$ limit. This divergence may be regularized by using the one-loop self-energy in the fermion propagator. As this is of order $1/N$, the singular η dependence is traded for an enhanced power of N in the numerator. Consequently the naive $1/N$ counting is modified and an infinite number of diagrams survive in each order of $1/N$. A systematic way to keep track of the true power of $1/N$ is obtained by using a “double-line” representation for the boson field that was previously used in the treatment of the electron-phonon interaction in metals.^{39,40} It was shown that the $1/N$ expansion could be organized as a genus expansion with all “planar” diagrams surviving to leading order. Reference 26 further established that in the large- N limit the boson propagator is unrenormalized beyond one-loop—in other words all higher loop diagrams that survive in the large- N limit give vanishing contributions. Each individual term contributing to the fermion self-energy is (if one calculates using

the one-loop fermion propagator) finite, and has the same functional form as the one-loop self-energy: formally (at $z_b = 3$),

$$\Sigma = -i \frac{1}{\lambda} \text{sgn}(\omega) |\omega|^{2/3} \sum_n \frac{a_n N^{n-1}}{N^n}. \quad (16)$$

The n th term in the sum comes from diagrams that are formally of order $1/N^n$ in the large- N expansion. However, for all planar diagrams there is a compensating enhancement factor N^{n-1} in the numerator so that each term is of order $1/N$. The worry is whether the sum over the infinite contributing diagrams leads to something singular or not.

It is straightforward to see that these results carry over to general z_b . Indeed, the kinematics leading to the divergences in the small- η limit depend only on the existence of the gapless Fermi surface and not on the detailed form of the boson propagator. When the divergence is regularized with the one-loop fermion self-energy, every $\frac{1}{\eta}$ is traded for a factor λN . Equation (16) is accordingly modified to

$$\Sigma = -i \frac{1}{\lambda} \text{sgn}(\omega) |\omega|^{2/z_b} \sum_n \frac{b_n (\lambda N)^{n-1}}{N^n}. \quad (17)$$

Here the coefficients b_n are all independent of N but in general depend on z_b . The utility of the small- z_b-2 limit where $\lambda \propto z_b-2$ is now apparent. So long as z_b-2 is of order $1/N$, the enhancement factor $(\lambda N)^{n-1}$ in the numerator of each term above is finite. If further the b_n 's are sufficiently nonsingular when $z_b \rightarrow 2$ then in the large- N limit only the $n=1$ term survives. This is just the one-loop answer which is thus exact in this limit. This claim can be illustrated explicitly by calculating a particular instance of a dangerous diagram which contributes to the series above at $n=2$, such as the one shown in Fig. 13. We do this in Appendix B and show that though it is of order $1/N$ for general z_b , in the limit $z_b-2 \propto 1/N$ it becomes higher order in $1/N$. In particular, the corresponding coefficient b_2 has a finite limit as $z_b \rightarrow 2$ so that its contribution is of order $1/N^2$. This is in fact expected to be true for all the b_n which have limits when $z_b \rightarrow 2$ such that to leading-order high- n terms in the series give subdominant powers of N to the leading-order result. Indeed, the absence of $\omega \log^2 \omega$ terms^{3,32,41} in the self-energy exactly at $z_b=2$ implies that the self-energy for small λ at fixed N has at most one inverse power of λ . Furthermore in this limit, the self-energy can be explicitly calculated using a perturbative RG technique (see Sec. VI). The answer agrees exactly with the leading-order term in the series above. This means that the small- z_b-2 behavior of b_n is such as to keep the high- n terms subdominant to the leading-order one in the large- N limit.

The boson propagator is also given exactly by the one-loop answer. For arbitrary z_b this follows from the arguments of Ref. 26 due to the vanishing of higher-order planar diagrams. However, in the small- z_b-2 limit it also follows for the same reason as above—the enhancement factors that render higher-order diagrams to be of the same nominal order in $1/N$ all become innocuous in the small- z_b-2 limit.

We emphasize that even though we have used the formal device of small z_b-2 to control the large- N expansion, the frequency dependence of the fermion self-energy $\Sigma \propto$

$-i \operatorname{sgn}(\omega)|\omega|^{2/z_b}$ is exact to all orders in z_b-2 . The smallness of z_b-2 merely assures us that the proportionality constant has a sensible $1/N$ expansion.

Similar results also apply to the fermion-boson vertex which to leading order in $1/N$ is unrenormalized. It is useful to understand the scaling structure of the low-energy physics described above. We have shown that in the one-patch theory, the large- N , small- z_b-2 limit, the low-energy physics is described by a fixed-point invariant under the scaling transformation,

$$\omega' = \omega b^{z_b/2}, \quad (18)$$

$$p'_x = p_x b, \quad (19)$$

$$p'_y = p_y b^{1/2}, \quad (20)$$

$$f'_{as}(p'_x, p'_y, \omega') = b^{-(z_b+5)/4} f_{as}(p_x, p_y, \omega), \quad (21)$$

$$a'(p'_x, p'_y, \omega') = b^{-(1+z_b)/2} a(p_x, p_y, \omega). \quad (22)$$

This is the same scaling structure that is obtained in a naive one-loop approximation. Though Ref. 26 has raised concerns over whether this scaling is internally consistent in the large- N limit, we see from the preceding analysis that it indeed is if the limit of small z_b-2 is also simultaneously taken.

The considerations of Ref. 32 can now be used to argue that this scaling structure is exact to all orders in the $1/N$ expansion (so long as z_b-2 is small) within this one-patch theory. More specifically the boson propagator and fermion self-energy will have the same functional forms [Eqs. (9) and (10)] as the one-loop answers. Thus we expect that in the N , z_b plane there is a region of finite extent where the scaling structure above is preserved. This conclusion is further bolstered by the perturbative RG analysis for finite N , small z_b-2 (see Sec. VI).

IV. TWO PATCHES

The considerations above are readily generalized to the two-patch theory. Indeed as before at each order of $1/N$, there are a number of diagrams that are divergent in the small- η limit. This is traded for an enhancement factor of λN raised to some power when the one-loop self-energy is used instead. For z_b-2 of order $1/N$, these enhancement factors become finite and a controlled $1/N$ expansion emerges. The structure of the leading-order contribution in $1/N$ to the fermion self-energy, the boson propagator, or the interaction vertex is then not modified from the one patch theory, and thus retains its RPA form. In particular, the scaling structure in Eqs. (18)–(22) is preserved. Below we examine higher-order corrections in the $1/N$ expansion, and show that new singularities appear at $o(1/N^2)$. These could potentially modify the scaling structure from that in Eqs. (18)–(22) above. However, we show that in the gauge-field model these contribute only to subdominant corrections to the one-loop fermion Green's function. On the other hand, for the nematic transition there is indeed a modification of the fermion scal-

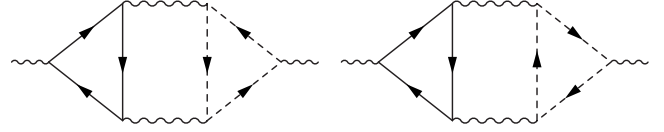


FIG. 5. Three-loop boson self-energy diagrams.

ing “dimension” at $o(1/N^2)$. Our calculations rely on the very recent impressive results of Metlitski and Sachdev²⁷ who studied the boson propagator and fermion self-energy in a direct (albeit uncontrolled) perturbative loop expansion up to three-loop order. Here we will show that calculations along the lines of those in Ref. 27 leads to controlled results for various physical quantities within our modified $1/N$ expansion.

First consider the boson propagator. At $z_b=3$, the analysis of Ref. 27 established that to three-loop order there is no shift of the true dynamical critical exponent. This was done by showing that the boson propagator at zero external frequency stays proportional to q^2 up to three loops. However, the three-loop contribution is of order \sqrt{N} bigger than the one-loop contribution thereby casting doubts on the existence of a sensible large- N limit. We will suggest an interpretation of this result in Sec. VIII. But for now we discuss the results of an identical analysis in our limit of small $z_b-2 \sim o(1/N)$. First we note that the leading-order term in the inverse gauge propagator is of order 1 with our conventions. Next higher loop diagrams clearly give subdominant powers of $1/N$ as there are no enhancement factors in the limit of small z_b-2 . At zero external frequency, the leading $1/N$ correction comes from the two diagrams shown in Fig. 5. These same diagrams were calculated in Ref. 27 at $z_b=3$. Repeating for general z_b , we find the $1/N$ correction

$$\pm c \frac{f_1(\lambda N)}{N} |q_y|^{z_b-1} \quad (23)$$

with the $-$ sign for the nematic critical point and the $+$ sign for the gauge model. The function f_1 is evaluated in Appendix C, and is readily seen to have a finite limit when $z_b \rightarrow 2$. For large λN , we have $f_1(\lambda N) \sim (\lambda N)^{z_b/2}$ in agreement with the result of Ref. 27 when $z_b=3$. We see explicitly that when z_b-2 is $o(1/N)$ these three-loop contributions are down by a factor $1/N$ compared to the one-loop term. Thus the large- N expansion is indeed well defined in this limit.

Next we consider the fermion propagator. The one-loop self-energy is inversely proportional to λN and hence is of order 1. It is easy to see by explicit calculation that the two-loop diagram shown in Fig. 6 is momentum independent, and merely provides an $o(1/N^2)$ modification of the coefficient of the frequency-dependent part of the self-energy. The most

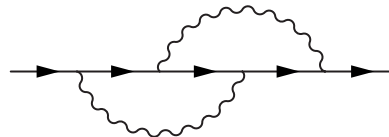


FIG. 6. Two-loop fermion self-energy diagrams merely renormalize the coefficient of $|\omega|^{2/z_b}$.

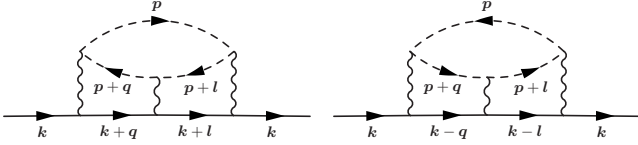


FIG. 7. Three-loop fermion self-energy diagrams involving fermions on both patches. Fermions on the right patch are denoted by solid lines and fermions on the left patch by dashed lines.

important effect at this order comes from the two graphs shown in Fig. 7. At zero external frequency, these graphs lead to singular momentum dependence.²⁷ Details are given in Appendix C. We find (for the right-moving fermion)

$$\delta\Sigma(\vec{p}, \omega=0) = \pm \frac{4}{3N^2} J(\lambda N) (p_x + p_y^2) \ln\left(\frac{\Lambda}{(p_x + p_y^2)^{z_b/2}}\right), \quad (24)$$

where the function $J(\lambda N)$ is defined in Appendix C and is positive definite. The + sign applies to the nematic critical point and the - sign to the gauge model. In a renormalization-group framework, this can be interpreted as the leading terms of a singular contribution to the self-energy of the form

$$(p_x + p_y^2)^{1 \mp 4J(\lambda N)/3N^2}. \quad (25)$$

A similar contribution exists in the frequency-dependent part as well, consistent with the dynamical scaling. In the gauge-field case, the plus sign applies and this singular correction is subdominant to the terms that already exist at leading order. Indeed we expect that any perturbation of the fixed point by irrelevant perturbations will generate an analytic contribution to the momentum dependence of the self-energy that will then dominate over the singular corrections found at $o(1/N^2)$. Though the leading-order frequency dependence is not analytic, we expect that if we use the large- N fermion propagators to calculate the effects of an irrelevant operator in perturbation theory, we will simply again generate a $|\omega|^{2/z_b}$ term. This will dominate over the singular order $1/N^2$ correction. Thus we conclude that in the gauge model the leading singularities are correctly given by the RPA forms at least to order $1/N^2$. We note that our interpretation is different from that in Ref. 27.

In the nematic case, the minus sign applies in the exponent of Eq. (25). This is more singular than the bare momentum dependence of the inverse Green's function, and consequently will dominate the low-energy physics near the Fermi surface. This singular correction can be interpreted as a shift of the scaling of the fermion fields from that in Eqs. (18)–(22). Thus to order $1/N^2$, we have

$$f'_{\alpha s}(p'_x, p'_y, \omega') = b^{-(z_b+5-\eta_f)/4} f_{\alpha s}(p_x, p_y, \omega), \quad (26)$$

where

$$\eta_f = \frac{4}{3N^2} J(\lambda N). \quad (27)$$

All the other scaling equations remain unmodified. This implies that the fermion Green's function satisfies the scaling form with $\alpha=1-\eta_f$.

What is the physical origin of the signs and the differences between the gauge and nematic models? We explain this in Sec. VII. To set the stage, we first calculate singularities in some other quantities within the general two-patch theory.

V. $2K_f$ AND OTHER SINGULARITIES

The calculations of Ref. 32 on the gauge model showed that the response to an external field that couples to the fermion density at momentum $2K_f$ is modified from that of a Fermi liquid due to the gauge interaction. The physical origin of this effect is clear. First there is a suppression of the $2K_f$ response coming from the smearing of the Landau quasiparticle due to the boson interaction. Second there is (in the gauge model) an enhancement coming from the Amperean attraction between a particle at K_f and a hole at $-K_f$. The gauge currents of such a particle and hole are parallel to each other so that the gauge interaction is attractive for such a particle-hole pair. The net modification of the $2K_f$ singularity is determined by the interplay between these two effects. We emphasize that the Amperean enhancement is specific to the gauge-field problem. For the closely analogous problem of a quantum critical point associated with a Pomeranchuk transition, the particle-hole interaction mediated by order-parameter fluctuations is repulsive. This goes in the same direction as the effect due to the smearing of the Landau quasiparticle. So there is no competition and we expect that the $2K_f$ singularities are simply suppressed when compared with the Fermi liquid at the Pomeranchuk transition.

The structure of the fixed point in the large- N , small- z_b -2 limit enables a controlled calculation of these effects. Consider an external field that couples to the $2K_f$ fermion density through the following term in the action:

$$u \int d^2x d\tau \bar{f}_L f_R + \text{H.c.} \quad (28)$$

By power counting, it is easy to see that the scaling in Eqs. (18)–(22) implies that the coupling u scales as

$$u' = ub. \quad (29)$$

In general, this will be modified at order $1/N$ as we demonstrate below. But first let us understand how the scaling of the u determines the $2K_f$ singularities. Assume in general that u scales as

$$u' = ub^{\phi_u}. \quad (30)$$

This implies that the operator $\rho_{2K_f}(\vec{x}, \tau) = \bar{f}_L(\vec{x}, \tau) f_R(\vec{x}, \tau)$ scales as

$$\rho'_{2K_f}(x', y', \tau') = b^\Delta \rho_{2K_f}(x, y, \tau) \quad (31)$$

with

$$\Delta = \frac{z_b + 3}{2} - \phi_u. \quad (32)$$

This determines the behavior of the singular part of the $2K_f$ density correlation function $C_{2K_f}(x, y, \tau) = \langle \rho_{2K_f}^*(x, y, \tau) \rho_{2K_f}(0, 0, 0) \rangle$. Its Fourier transform satisfies

$$C_{2K_f}(p_x, p_y, \omega) = b^{(3+z_b)/2-2\Delta} C'_{2K_f}(p'_x, p'_y, \omega'). \quad (33)$$

Note that the momenta p_x, p_y describe the *deviation* of the full momentum from $2K_f \hat{x}$ in this correlation function. We then immediately have the scaling form

$$C_{2K_f}(p_x, p_y, \omega) = \frac{1}{\omega^{1+(3-4\Delta)/z_b}} C\left(\frac{\omega}{|p_y|^{z_b}}, \frac{p_x}{p_y^2}\right). \quad (34)$$

Note that in the usual Fermi-liquid case, we have $z_b=2$, $\phi_u=1$ which reproduces the well-known square-root frequency dependence of the singular part of the $2K_f$ correlations.

The leading-order correction to ϕ_u in the $1/N$ expansion comes from the one-loop vertex correction diagram of Fig. 15. To calculate it, we combine the large- N expansion with an RG transformation in which internal loop integrals are performed over a thin shell in (\vec{p}, ω) space. It will be convenient to define the RG so that we integrate over arbitrary p_x, ω but over a shell in p_y where $\Lambda > |p_y| > \Lambda/b^{1/2}$. The vertex correction becomes

$$\delta u = -\frac{1}{N} \int_{\vec{p}, \omega} D(p_y, \omega) \mathcal{G}_R(p_x, p_y, \omega) \mathcal{G}_L(p_x, p_y, \omega). \quad (35)$$

Here $\mathcal{G}_{R/L}$ are the propagators of the right- and left-moving fermions, respectively. These propagators include the singular frequency-dependent self-energy discussed in previous sections. Note the minus sign in front which comes from the different signs with which the gauge field couples to the left and right fermions. In contrast at a Pomeranchuk transition, the minus sign will be absent. Consequently what is an enhanced vertex in the gauge-field problem will become a suppressed vertex at the Pomeranchuk transition.

We evaluate the integral in Appendix D. The result takes the form

$$\delta u = \frac{1}{4\pi^2 N} g(\lambda N, z_b) \ln(b), \quad (36)$$

where the function g is given by

$$g(v, z_b) = \int_0^\infty dt \left(\frac{1}{\gamma t + 1} \right) \frac{v t^{2/z_b}}{t^{4/z_b} + v^2}, \quad (37)$$

where $\gamma = \frac{1}{4\pi}$. For fixed v in the limit that $z_b=2$ this function has a sensible limit which we demote $g_2(v)$. We find

$$g_2(v) = \frac{\pi \gamma v^2}{2(\gamma^2 v^2 + 1)} \left[1 - \frac{2}{\pi \gamma v} \ln(\gamma v) \right]. \quad (38)$$

The vertex correction in Eq. (36) implies a modified scaling exponent for u ,

$$\phi_u = 1 + \frac{1}{4\pi^2 N} g(\lambda N, z_b). \quad (39)$$

As g has a finite limit when $z_b \rightarrow 2$, $N \rightarrow \infty$ but λN is finite, this modification is of order $1/N$. We note that it is sufficient to evaluate g at $z_b=2$ in this limit. To leading order in $\epsilon = z_b - 2, \frac{1}{N}$, we therefore have (using $\lambda = 2\pi^2 \epsilon$ appropriate for small ϵ)

$$\phi_u = 1 + \frac{1}{4\pi^2 N} g_2(2\pi^2 \epsilon N). \quad (40)$$

Inserting this into Eqs. (32) and (34), we see that for small $\epsilon, 1/N$ the power of ω in the $2K_f$ singularity is

$$\frac{1}{2} + \frac{\epsilon}{4} \left[1 - \frac{2g_2(2\pi^2 \epsilon N)}{\pi^2 \epsilon N} \right]. \quad (41)$$

Since $g_2(v)/v$ runs from ∞ to 0 as v runs from 0 to ∞ , the $2K_f$ singularity is suppressed over that of the Fermi liquid for $\epsilon N \rightarrow \infty$ while it is enhanced when $\epsilon N \rightarrow 0$. The competition in Eq. (41) between the first and second term in parentheses is precisely the competition between quasiparticle smearing and Amperean attraction described at the beginning of this section.

The structure of the singularities in the Cooper channel will also be modified from that of the Fermi liquid. But here the Amperean interaction is repulsive in the gauge-field case and thus the Cooper singularities will be weaker than in the Fermi liquid. On the other hand at the nematic quantum critical point, the boson-mediated interaction is attractive in the Cooper channel. Consequently there will be an enhancement of the pairing vertex due to boson exchange. This leads to an enhancement of the Cooper singularities compared with the Fermi liquid. A complete discussion of this effect is more complicated than the $2K_f$ singularities and will be presented elsewhere.⁴² Within the large- N expansion for small z_b-2 , in contrast to the $2K_f$ singularity, there is an $o(1)$ correction to the tree level scaling exponent of the Cooper vertex which receives contributions from many diagrams.⁴² Quantitative calculation of this correction as well as a study of the interesting interplay between superconductivity and critical nematic fluctuations will appear elsewhere.⁴²

VI. PERTURBATIVE FIXED POINT FOR FINITE- N , SMALL- ϵ , AND $2K_f$ SINGULARITIES

A different controlled limit was previously discussed in Ref. 36. This is obtained by considering finite N , and small ϵ where there is a perturbatively accessible renormalization-group Fermi-liquid fixed point. In this section, we study the $2K_f$ singularities at this fixed point. It will be useful for our purposes to define a RG scheme that is slightly different from Ref. 36. So let us first reproduce their main result. Consider the two-patch action

$$S = S_f + S_{int} + S_a, \quad (42)$$

$$S_f = \int d^2 x d\tau \sum_{s\alpha} \bar{f}_{s\alpha} (\partial_\tau - is\partial_x - \partial_y^2) f_{s\alpha}, \quad (43)$$

$$S_{int} = \int d^2x d\tau \frac{se}{\sqrt{N}} a \bar{f}_{s\alpha} f_{s\alpha}, \quad (44)$$

$$S_a = \int_{\vec{k}, \omega} |k_y|^{z_b-1} |\bar{a}(\vec{k}, \omega)|^2 + \dots \quad (45)$$

Compared with the “minimal” action of Sec. II above, we have set $\eta=1$ and have reinstated the boson coupling e . The ellipses in the last equation refer to other operators irrelevant at the RG fixed point to be described below (for instance, an ω^2 term in the quadratic gauge action). When $e=0$, the fermions and boson field are described by two decoupled Gaussian theories. This is at a fixed point under the scaling,

$$\tau' = \frac{\tau}{b}, \quad (46)$$

$$x' = \frac{x}{b}, \quad (47)$$

$$y' = \frac{y}{\sqrt{b}}, \quad (48)$$

$$f'(x', y', \tau') = b^{3/4} f(x, y, \tau), \quad (49)$$

$$a'(x', y', \tau') = b^{3/2-z_b/4} a(x, y, \tau). \quad (50)$$

Now turn on a small $e \neq 0$. By power counting, we find

$$e' = e b^{(z_b-2)/4}. \quad (51)$$

In differential form (if we let $b=1+dl$), we get

$$\frac{de^2}{dl} = \frac{\epsilon}{2} e^2. \quad (52)$$

Thus e^2 is relevant for $\epsilon > 0$ and irrelevant for $\epsilon < 0$. To determine the fate of the theory for $\epsilon > 0$, let us study the one-loop beta function for e^2 . We will define the RG by integrating out modes with $\Lambda > |q_y| > \frac{\Lambda}{\sqrt{b}}$. To order e^2 , the fermion self-energy is given by the integral,

$$\Sigma = \frac{e^2}{4\pi^3 N} \int_{q_x, \omega'} \int_{\Lambda/\sqrt{b}}^{\Lambda} dq_y D(q_y, \omega') \mathcal{G}(\vec{p} - \vec{q}, \omega - \omega'), \quad (53)$$

where D is the gauge propagator in the bare action. After doing the q_x integral, the remaining ω' integral only gets contributions from small ω' so we can replace the gauge propagator by its low-frequency form $\frac{1}{|q_y|^{z_b-1}}$. For small external frequency ω , we get

$$\Sigma = -\frac{ie^2 \omega}{2\pi^2 N} \int_{\Lambda/\sqrt{b}}^{\Lambda} \frac{dq_y}{q_y^{z_b-1}}. \quad (54)$$

Anticipating that there is a new fixed point when $e^2/N \sim o(\epsilon)$, we replace the integrand by its value at $z_b=2$ to get

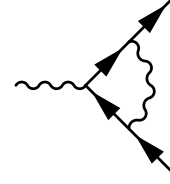


FIG. 8. One-loop correction to the fermion-boson vertex.

$$\Sigma = -\frac{ie^2 \omega}{4\pi^2 N} \ln b. \quad (55)$$

Thus the inverse fermion propagator takes the form

$$i\omega \left(1 + \frac{e^2}{4\pi^2 N} \ln b \right) - p_x - p_y^2. \quad (56)$$

The correction to the fermion-boson vertex (see Fig. 8) at order e^2 vanishes as the q_x integral has poles only on one side of the complex plane. Finally the change to the boson propagator is also zero if only modes with high $|q_y|$ are integrated out. Thus the only change is in ω dependence of the fermion propagator. This may be incorporated into a modified scaling,

$$\omega' = \omega \left(1 + \frac{e^2}{4\pi^2 N} \ln b \right) \quad (57)$$

$$= \omega b^{1+e^2/4\pi^2 N} \quad (58)$$

This implies

$$\tau' = \frac{\tau}{b^{1+e^2/4\pi^2 N}}, \quad (59)$$

$$x' = \frac{x}{b}, \quad (60)$$

$$y' = \frac{y}{\sqrt{b}}, \quad (61)$$

$$f'(x', y', \tau') = b^{3/4+e^2/8\pi^2 N} f(x, y, \tau), \quad (62)$$

$$a'(x', y', \tau') = b^{3/2-z_b/4+e^2/8\pi^2 N} a(x, y, \tau). \quad (63)$$

The modification to the flow of the coupling e is now readily obtained to be

$$e' = e b^{(z_b-2)/4 - e^2/8\pi^2 N}. \quad (64)$$

In differential form, this implies the flow equation

$$\frac{de^2}{dl} = \frac{\epsilon e^2}{2} - \frac{e^4}{4\pi^2 N}. \quad (65)$$

Thus we indeed find a fixed point when

$$e_*^2 = 2\pi^2 N \epsilon. \quad (66)$$

Right at the fixed point, the scaling equations above are identical (to within order ϵ) to those found earlier in Sec. III and

indeed to that expected based on RPA. The differences from RPA discussed in earlier sections in the fermion propagator come from three-loop calculations, and hence are not expected to show up till order ϵ^3 .

The singularities of many physical quantities can be usefully calculated within this ϵ expansion and provides an alternate controlled limit to the one we have discussed. As an illustration let us calculate the boson propagator and the fermion self-energy. Let the bare value of the electric charge at the cutoff scale be e_0 . The boson propagator is given by the usual one-loop diagram and takes the form

$$D(q_y, \omega) = \frac{1}{\frac{e_0^2}{4\pi} \frac{|\omega|}{|q_y|} + |q_y|^{1+\epsilon}}. \quad (67)$$

The Landau-damping term does not acquire any corrections from this perturbative answer at least up to the order to which the RG has been performed. To obtain the frequency dependence of the fermion propagator, we examine the flow of the coefficient of the $i\omega$ term calculated above. Let us denote this coefficient $\eta(l)$ at an RG scale l . The calculation above gives the flow equation

$$\frac{d\eta}{dl} = \frac{\eta e^2}{4\pi^2 N}. \quad (68)$$

Combining with the equation for e^2 , we obtain

$$\frac{d(\eta e^2)}{dl} = \frac{\epsilon}{2} \eta e^2. \quad (69)$$

Thus we find

$$\eta(l) e^2(l) = e_0^2 e^{\epsilon l/2}, \quad (70)$$

where we set $\eta(l=0)=1$, and e_0^2 is the bare coupling at the cutoff scale. The frequency dependence of the fermion self-energy is then obtained by setting $l = \ln \frac{\Lambda^2}{\omega}$. In the limit $\omega \rightarrow 0$, we may set $e^2(l) = e_*^2$ so that

$$\eta(\omega) = \frac{e_0^2}{2\pi^2 N \epsilon} \left(\frac{\Lambda^2}{\omega} \right)^{\epsilon/2}. \quad (71)$$

Then at order ϵ , the self-energy becomes

$$\Sigma(\omega) = -i \frac{e_0^2}{2\pi^2 N \epsilon} |\omega|^{1-\epsilon/2} \text{sgn}(\omega). \quad (72)$$

Here we have ignored a term Λ^ϵ in the numerator to this order in ϵ .

To compare with the results of previous sections, we need to set the bare coupling $e_0^2=1$. Note, in particular, that the prefactor to the frequency dependence is exactly $\frac{1}{\lambda N}$ consistent with the earlier analysis. On the other hand, to calculate the scaling dimensions of any operator directly within this epsilon expansion we need to sit right at the fixed point and perturb the theory with that operator. The fixed-point theory corresponds to setting the bare coupling $e_0=e_*$. Thus the boson propagator and fermion self-energy right at the fixed point take the forms

$$D_*(q_y, \omega) = \frac{1}{\frac{e_*^2}{4\pi} \frac{|\omega|}{|q_y|} + |q_y|^{1+\epsilon}}, \quad (73)$$

$$\Sigma_*(\omega) = -i |\omega|^{1-\epsilon/2} \text{sgn}(\omega). \quad (74)$$

Note that as $\epsilon \rightarrow 0$ these fixed-point propagators go over smoothly into those of the ‘‘decoupled’’ Gaussian fixed point, as indeed they must.

As it does not seem to be available in the literature, let us now calculate the scaling exponent for the $2K_f$ singularity at this $o(\epsilon)$ fixed point. As before, we add the term in Eq. (28) to the action. By power counting, we again find

$$u' = ub. \quad (75)$$

This is modified at leading order of ϵ by the same vertex correction diagram as before. Evaluating the integral as before, we find

$$\delta u = \frac{e_*^2 u}{4\pi^2 N} \int_{\frac{\Lambda}{\nu b}}^{\Lambda} dq_y \int_0^\infty d\omega D_*(q_y, \omega) \frac{\tilde{\Sigma}_*(\omega)}{[\tilde{\Sigma}_*(\omega)]^2 + q_y^4}. \quad (76)$$

Here we have written $\Sigma_*(\omega) = -i\tilde{\Sigma}_*(\omega)$. Naively as the vertex correction is already order $e^2 \sim \epsilon$, we should replace the integrand by its value at $\epsilon=0$, i.e., by the fermion and boson propagators at the Gaussian fixed point. However at the Gaussian fixed point $\tilde{\Sigma} = \omega$ and the frequency integral is logarithmically divergent at large ω . This signals that the leading-order ϵ correction to ϕ_u is not analytic in ϵ . To extract it we keep the correct boson propagator and fermion self-energy at the $o(\epsilon)$ fixed point calculated in Eqs. (73) and (74).

Inserting these into the integral for the vertex correction, we see that the high- ω divergence of the ω integral is cutoff by the presence of the Landau-damping term in the boson propagator. The integral is readily evaluated for small ϵ , and we find

$$\delta u = \frac{u\epsilon}{2} \ln \left(\frac{2}{\pi\epsilon N} \right) \ln b. \quad (77)$$

Thus we get the modified scaling equation

$$u' = ub^{1+\epsilon \ln(2/\pi\epsilon N)/2} \quad (78)$$

so that

$$\phi_u = 1 + \frac{\epsilon}{2} \ln \left(\frac{2}{\pi\epsilon N} \right). \quad (79)$$

Inserting into Eqs. (32) and (34), we find the power of ω in the $2K_f$ singularity

$$\frac{1}{2} + \frac{\epsilon}{4} \left(1 - 4 \ln \frac{2}{\pi\epsilon N} \right), \quad (80)$$

i.e., the $2K_f$ singularity is enhanced compared to the Fermi liquid at this fixed point. Note that this answer for ϕ_u agrees exactly with the result of Sec. V when the limit of $\epsilon N \rightarrow 0$ is taken.

VII. PHYSICAL PICTURE

In this section, we discuss some qualitative aspects of the physical picture of the low-energy physics, and show how we may understand the results of some of the detailed calculations presented in previous sections. First we notice that at low energies, the Fermi surface is sharp even though the Landau quasiparticle has been destroyed. This is qualitatively the same as in the RPA but at least in the nematic case, the detailed singular structure is modified. Despite this, as argued in many previous papers^{3,31,32} this is a compressible state. This follows immediately from the general argument in Appendix A that the compressibility does not receive any singular contributions from the low-energy scale invariant fluctuations. Hence this non-Fermi-liquid state has a finite and nonzero compressibility. Actually in the nematic case changing the chemical potential will also in general drive the system away from the critical point. This leads to a singular contribution to the ground-state energy as a function of chemical potential which could lead to a singular contribution to the compressibility.²⁷ The safe statement then is that the differential change in density in response to a change in location within the phase boundary is finite.

Actually the patch construction implies an even stronger result. Consider the susceptibility to a deformation of the Fermi surface in any angular momentum channel (i.e., the response to an external field that couples to the corresponding shape distortion of the Fermi surface). The universal singularities in this quantity are obtained by examining this coupling within the patch construction. But the patch theory does not know anything about the angular dependence of the probe field. So within each patch this external field couples in the same way as an external chemical potential. Consequently (just as for the contribution to the compressibility from the low-energy density fluctuations), there is no singular contribution to the susceptibility in any angular momentum channel which all stay finite and nonzero. For the nematic critical point, the only exception is the order-parameter channel itself ($l=2$ for the d -wave nematic). In that case, this argument implies that the critical behavior of the order-parameter susceptibility is correctly given by the mean-field Hertz answer and receives no singular corrections from the fermions. These arguments provide a simple explanation of some recent results for the nematic critical point obtained through detailed calculations.¹⁶ Consider the approach to the quantum critical point from the symmetric side where there is no nematic order. At low energies, the corresponding metal is described by Fermi-liquid theory characterized by Landau quasiparticles with an effective mass m^* and various Landau parameters. On approaching the quantum critical-point standard scaling arguments show that the effective mass diverges with an exponent as

$$m^* \sim \delta^{-(z_b-2)/(z_b-1)}. \quad (81)$$

In the Fermi-liquid phase, the compressibility is expressed in terms of m^* and the Landau parameter F_s^0 as

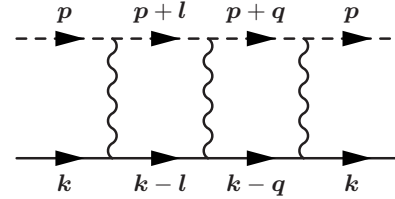


FIG. 9. The cooper-channel scattering amplitude.

$$\kappa \propto \frac{m^*}{1 + F_s^0}. \quad (82)$$

The constancy of κ as the critical point is approached implies that the Landau parameter F_s^0 diverges in exactly the same way as the effective mass. Applying this reasoning to other angular momentum channels, we see that the Landau parameters in all angular momentum channels (except the order-parameter one itself) must diverge in the same way as m^* so as to give a constant susceptibility at the critical point. This is exactly the conclusion of Ref. 16.

In either the gauge model or the nematic critical point, the patch construction implies that the only singular modification from RPA in the charge-density response happens at the $2K_f$ wave vectors (modulo the caveat discussed above for the $q=0$ response in the nematic case). As explained in detail in Sec. V apart from the $2K_f$ particle-hole correlations, the main modifications from the Fermi liquid in the two-particle response are in the structure of the pairing correlations. Whether the $2K_f$ and pairing correlations are enhanced or not compared to the Fermi liquid is largely determined by the Amperean rules. In the gauge-field case, the pair correlations are suppressed and the $2K_f$ potentially enhanced while the opposite is true for the nematic critical point.

Consider now the $1/N^2$ calculation of the fermion self-energy described in Sec. IV. The singular contribution to the self-energy comes the two diagrams shown in Fig. 7. We note that both diagrams may be expressed in terms of appropriate two-particle scattering amplitudes. Figure 9 is a scattering amplitude in the (particle-particle) Cooper channel while Fig. 10 is a scattering amplitude in the particle-hole $2K_f$ channel. Based on the physical picture dictated by the Amperean rules, we expect that in the nematic case the Cooper channel diagram by itself leads to a self-energy that is more singular than the bare terms in the action while the $2K_f$ diagram by itself leads to a singularity that is less singular than the bare term. The situation is clearly reversed in the gauge-field model. This physical picture thus enables us to understand the signs of the contributions of the two diagrams in the calculation.

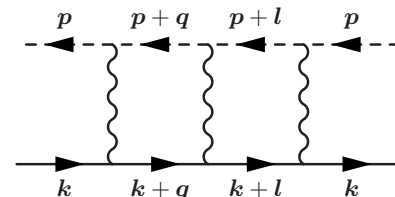


FIG. 10. The particle-hole channel scattering amplitude.

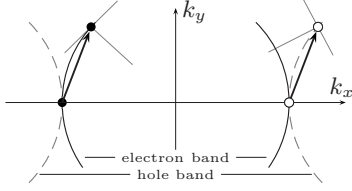


FIG. 11. After a scattering event in the $2K_f$ channel, two fermions are no longer perfectly nested.

Whether the net effect is to produce a singular correction to the self-energy that dominates over the bare one at low momenta and frequencies is determined by the competition between the Cooper and $2K_f$ contributions, i.e., by the relative magnitude of the contribution of the two diagrams. We now argue that the Cooper channel always dominates (i.e., gives the bigger contribution) consistent with the results of the actual calculation. To see this consider both two-particle scattering amplitudes when the external lines are right at the Fermi surface, and initially satisfy the “nesting” condition. In the Cooper channel, this means that the total momentum of the two incoming particles is zero. In the $2K_f$ channel, this means that the incoming particle-hole pair has momentum exactly $2K_f\hat{x}$. In the $2K_f$ channel, exchange of a boson with momentum q_y leads to a new particle-hole pair state which no longer satisfies the nesting condition (see Fig. 11). Thus after one such scattering event, the particle-hole pair is less sensitive to the Amperean attraction/repulsion mediated by subsequent boson exchange. In contrast in the Cooper channel, exchange of a boson with momentum q_y preserves the nesting condition for the resulting particle-particle pair (see Fig. 12). Thus they are able to continue to reap the benefits of the Amperean interaction in subsequent scattering events. This explains why the Cooper channel always dominates over the $2K_f$ channel. This difference between the kinematics of the Cooper and $2K_f$ scattering channels is clearly present only if the Fermi surface is curved. Thus we expect that in the artificial limit where we ignore the curvature term in the fermion Green’s function, the two diagrams will have the same magnitude and hence will cancel. Examining the relevant integrals shows that this is exactly what happens.

Let us now discuss one important physical consequence of these results. At the nematic critical point, the singular structure of the fermion Green’s function is modified from RPA at order $1/N^2$. Specifically the fermion Green’s function satisfies the scaling form of Eq. (7) with the fermionic dynamical exponent $z = \frac{z_b}{2}$ and the exponent $\alpha = 1 - \eta_f$ with η_f given in Eq. (27). Note that η_f is positive. This represents a modification of RPA which has $\eta_f = 0$. As pointed out in Ref. 37, this shows up very directly in the tunneling density of states $N(\omega)$ defined through

$$N(\omega) = \int \frac{d^2\vec{K}}{(2\pi)^2} A(\vec{K}, \omega), \quad (83)$$

where the single-particle spectra function $A(\vec{K}, \omega)$ is defined in the usual manner,

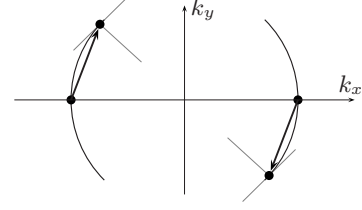


FIG. 12. After any scattering event in the Cooper channel, two fermions remain perfectly nested.

$$A(\vec{K}, \omega) = -\frac{1}{\pi} \text{Im} G(\vec{K}, i\omega \rightarrow \omega + i0^+). \quad (84)$$

Here \vec{K} is the full momentum (not linearized near the Fermi surface). Apart from being a potentially direct measure of the deviation from RPA, study of $N(\omega)$ also provides insight into the sign of η_f and some rationalization for why it is nonzero in the first place. Singular contributions to $N(\omega)$ come from momenta in the vicinity of the Fermi surface. The two-dimensional momentum integral may then be separated into an angular integral over the Fermi surface and a radial integral over just the component of the momentum normal to the Fermi surface. The former just contributes an overall constant prefactor. The latter may be directly evaluated to obtain the result advertised in Sec. I,

$$N(\omega) \sim |\omega|^{\eta_f/z}. \quad (85)$$

Thus at the nematic critical point, there is a power-law suppression of the local single-particle density of states. This suppression is of course rather natural if we remember that superconducting fluctuations are enhanced at the nematic critical point. Thus the sign of η_f may be qualitatively understood. Further the enhanced superconducting fluctuations make it plausible that there be some effect on the density of states unlike what happens in the RPA.

VIII. TOWARD A PHASE DIAGRAM

We now turn to the question of what happens for general z_b, N . The preceding sections show that so long as $z_b - 2$ is small, the theory can be controlled for any N . What happens if $z_b - 2$ is not small? As discussed Sec. I for $z_b = 3$ the possibility of using large N as a control parameter has been studied in detail recently and several difficulties have been pointed out. Here we suggest an interpretation of these difficulties. Consider first the nematic critical point. The action for the order-parameter fluctuations within the two-patch theory was calculated to three-loop order in Ref. 27. For fluctuations at zero frequency, they found

$$\int_k N(1 - c\sqrt{N})|k_y|^2 |a(\vec{k}, \omega = 0)|^2. \quad (86)$$

The second term comes from the three-loop polarizability of the fermions. The appearance of the extra factor of \sqrt{N} in the loop calculation raises concerns over the existence of a sensible large- N limit. It is currently not known what the structure of higher loop terms are, for instance, whether same or

even higher powers of N are generated by higher loop contributions. If we take the three-loop answer at face value then as $c > 0$, for large enough N , the coefficient of k_y^2 becomes negative. This signals an instability toward ordering at non-zero momentum. In particular, this means that the original assumption of a direct second-order nematic transition is not correct, and the nematic transition will be preempted by the appearance of density wave order.

Can this conclusion be changed by higher loop diagrams? One possibility is that higher loop diagrams change the sign of the coefficient c . But then a different instability will likely set in. For instance, consider diagrams with the same structure as those in Fig. 5 but with arbitrary number of boson lines connecting the right- and left-moving fermions. Whatever the sign of the sum of diagrams of this sort, so long as it has a higher power of N than the one-loop one, there will be an instability. If the sign is negative (as in the three-loop calculation), then there is an instability where the boson likes to order at nonzero wave vector. If the sign is positive then we consider the response to an external probe field that couples with *opposite* sign to the two patches. A concrete example at the nematic critical point is just an external electromagnetic gauge field, i.e., we consider the “diamagnetic” response to a static external magnetic field. This changes the sign of the external vertices in diagrams such as Fig. 5 without changing the sign of the internal vertices. The diamagnetic response to a static magnetic field then has the opposite sign from that of an ordinary metal. This signals an instability toward spontaneous flux formation, i.e., the system will likely develop a state associated with spontaneous circulating currents.

The only remaining possibility is that higher loop diagrams exactly cancel the offending \sqrt{N} term found in the three-loop calculation. While we cannot rule this out we can provide a suggestive argument against this possibility by examining the limit of small $z_b - 2$. In this limit, the static boson polarizability may formally be written as a series,

$$\Pi(k_y, \omega = 0) = |k_y|^{z_b-1} \left[1 + \sum_n \frac{f_n(\lambda N; z_b)}{N^n} \right]. \quad (87)$$

The leading $n=1$ term was calculated in Appendix C. Let us assume that the functions $f_n(x; z_b)$ all have finite limits when $z_b \rightarrow 2$,

$$\lim_{z_b \rightarrow 2} f_n(x, z_b) = F_n(x). \quad (88)$$

This is explicitly seen to be true for $n=1$, and we assume it holds for arbitrary n . Then successive terms in the series above are down by powers of $1/N$ for finite nonzero λN . So for large N , there is no instability. The instability potentially happens when λN becomes large enough that the $n=1$ term is comparable to 1, i.e., when $\lambda N \sim N^{2/z_b}$. If any higher-order term, say the n th one, is to have the same power of N in the large- λN limit (while still keeping ϵ small), then $F_n(x) \sim x^{n+z_b/2}$ for large x . But then its coefficient has a high power of $\lambda \sim \epsilon$. Therefore, any cancellation of the dangerous three-loop term by higher loop diagrams cannot in general happen for arbitrary z_b . This makes it rather likely that there is an instability for arbitrary z_b not too small.

For z_b approaching 2, comparison of the three-loop term with the leading one-loop term suggests that the instability happens when $\epsilon^{2/\epsilon} \sim \frac{1}{N}$. This leads to a “phase boundary” between the unstable and stable regions that comes in with infinite slope in the $\epsilon, \frac{1}{N}$ plane. Thus we propose the phase diagram shown in Fig. 1 or Fig. 2. Through out the unstable region there is no direct nematic transition, and it is always preempted by instability toward a different broken symmetry. It is not clear whether the unstable region encompasses the all important point $N=2, z_b=3$, i.e., whether Fig. 1 or Fig. 2 is realized. If however this point belongs to the stable region, i.e., Fig. 1 applies and there is a direct second-order nematic transition, then we have no choice but to access it from the small- ϵ region (either by combining with large N or by the perturbative RG for small N). If Fig. 2 is realized on the other hand we may still hope that our expansion captures the physics at temperatures above the instability.

Similar considerations apply to the gauge-field model. There at $z_b=3$, the gauge polarizability acquires only a positive $o(\sqrt{N})$ correction at three-loop level. However at the same order if we consider the response to an external probe that couples with the same sign to both right and left movers, then the sign of the three-loop response is reversed. In arguing for an instability, it is important that the bare zero-momentum susceptibility vanish since the result Eq. (86) applies in the scaling regime, which requires $|k_y| \ll \frac{1}{N^{3/2}\eta}$; if the bare susceptibility were nonzero and N independent, the putative instability would occur at k_y outside this regime.⁴³ A suitable choice may be a perturbation of both the volume of the system and the chemical potential, preserving the particle density. If the three-loop calculation were the full story this would again signal an instability toward a state which spontaneously orders at nonzero momentum and hence breaks translation symmetry. More generally, when higher loop terms are included the situation is similar to the discussion above for the nematic model. Consequently we suggest that the gauge-field model is also unstable toward a state with some broken symmetry at sufficiently large N when z_b is sufficiently different from 2. Thus once again this proposal would imply that if the gauge model at $N=2, z_b=3$ is stable then we have no choice but to access it as we have done from the small- ϵ region.

IX. DISCUSSION

In this concluding section we consider the implications of our results for some specific systems, and for the general theory of non-Fermi-liquid metals. An important and topical realization of the gauge model is to the theory of gapless quantum spin liquids where a gapless Fermi surface of charge neutral spin-1/2 fermionic spinons is coupled to a gapless fluctuating $U(1)$ gauge field. Note that in this example, the spin liquid is a non-Fermi-liquid metal for spin transport but is an insulator for electrical transport. Such a state has been proposed^{1,2} to describe the intermediate temperature scale physics of the layered organic Mott insulators κ -(ET)₂Cu₂(CN)₃ and EtMe₃Sb[Pd(dmit)₂]₂. Our controlled calculations merely confirm the correctness of several key results from RPA that are directly relevant to experiments—

for instance, the scaling structure of the low-energy theory implies that the specific heat follows the familiar RPA result $C_v \sim T^{2/3}$ at low temperature T . The more important contribution of the present paper to the theory of such a spin liquid is the controlled calculation of the structure of the $2K_f$ spin correlations. Detecting these in experiments would be an interesting way to “measure” the spinon Fermi surface (see Ref. 44 for a proposal). Our results also set the stage for an analysis of phase transitions from the spinon Fermi-surface state to various proximate phases with spinon pairing or other ‘order’ that may be relevant to describing the very low temperature physics of the organics.

The model of a Fermi surface coupled to a gauge field also describes algebraic charge liquid metals^{4,6,7} and the related d -wave Bose metals.⁸ An essential difference with the particular gauge model studied in this paper is that there are two species of fermions that couple with opposite gauge charges to the same fluctuating $U(1)$ gauge field. The Amperean rules are therefore different and this will lead to some differences in the results. These can be straightforwardly handled within our expansion. Similarly our methods are readily generalized to provide controlled expansions for various Pomeranchuk transitions other than the nematic example considered in detail in this paper.

It is interesting to contrast the quantum critical metal at these Pomeranchuk transitions with other examples of non-Fermi-liquid metals. One other set of examples is provided by Mott-type quantum phase transitions where an entire Fermi surface disappears continuously. Apart from continuous Mott metal-insulator transitions, these are thought to describe non-Fermi-liquid physics in heavy fermion metals near the onset of magnetic long-range order. Reference 37 argued that such Mott-type transitions will be characterized by the presence of a sharp critical Fermi surface but without a sharp Landau quasiparticle. A critical Fermi surface is also a feature of a Pomeranchuk transition if it is second order. However, we might expect that the destruction of the Landau quasiparticle is more severe at the Mott-type transitions. Indeed the explicit calculation in Ref. 34 for a continuous Mott transition found the exponent value $\alpha = -\eta$ (where η is the anomalous exponent of the boson field at the three-dimensional XY fixed point, and is known to be small and positive). This corresponds to a large fermion anomalous dimension $\eta_f = 1 - \alpha = 1 + \eta$. In contrast at the nematic critical point, the fermion anomalous dimension is small. Within RPA, it is simply 0 while the three-loop calculation of Ref. 27 as well as the controlled estimate presented in this paper give nonzero but small values. The largeness of η_f is a partial measure of the extent to which the quasiparticle is smeared (as is exemplified by the suppression of the tunneling density of states).

It is instructive to compare the non-Fermi liquids studied in this paper with those discovered recently using holographic duality,^{45–49} which may be understood heuristically as arising from a Fermi surface coupled with some bath of critical fluctuations with an infinite dynamical exponent.^{48–50} It would be interesting to find explicit field theoretical models with this feature. In the models studied here, large $z_b > 3$ is unstable because a $|q|^2$ term in the boson inverse propagator will always be generated by short-distance fluctuations and will eventually dominate.

In summary in this paper, we have developed a controlled and systematic approach to calculating the universal properties of a non-Fermi-liquid metal that arises when a gapless Fermi surface is coupled to a fluctuating gapless boson field. We illustrated our approach by studying spinon Fermi-surface spin liquids, and quantum critical metals near an electronic nematic transition in some detail. Our approach readily lends itself to the study of various closely related problems. We leave the exploration of these to the future.

ACKNOWLEDGMENTS

We thank Yong Baek Kim, Sung-Sik Lee, Max Metlitski, and Subir Sachdev for useful discussions. T.S. was supported by NSF under Grant No. DMR-0705255. J.M. and H.L. were supported by funds provided by the U.S. Department of Energy (DOE) under cooperative research Agreement No. DE-FG0205ER41360 and the OJI program (H.L.) and the Alfred P. Sloan Foundation (J.M.).

APPENDIX A: UNIVERSALITY AND THE PATCH CONSTRUCTION

In this paper, we have discussed a scaling theory which focuses on the interactions of fermions near the Fermi surface via bosons of small frequency and small momentum. The kinematics of these bosons allows us to restrict attention to one patch of the Fermi surface and its antipode. In particular, the universal singularities in the low-energy physics is correctly captured by breaking up the full Fermi surface into patches (and their antipodes), and studying the theory patch by patch. In this appendix, we briefly discuss some subtle points associated with the patch construction and the subsequent treatment of the theory for any given pair of antipodal patches. As a bonus we will show that the compressibility has no singular contributions coming from the low-energy scale invariant fluctuations.

Consider the following microscopic Lagrangian for the gauge-field problem:

$$\mathcal{L}_{UV} = f^\dagger \left[\partial_\tau - \frac{1}{2m} \left(-i\vec{\nabla} + \frac{1}{\sqrt{N}}\vec{a} \right)^2 + \mu \right] f. \quad (\text{A1})$$

This Lagrangian is gauge invariant and this forbids a mass term for the gauge field \vec{a} in the effective action.

Next consider the low-energy description which focuses on the modes near two antipodal patches of Fermi surface (with normal \hat{x} , without loss of generality), $f \approx f_R e^{ik_F x} + f_L e^{-ik_F x}$,

$$\begin{aligned} \mathcal{L}_{\text{patch}} = & f_R^\dagger \left(\partial_\tau - i v_F \partial_x - \frac{\partial_y^2}{2m} + \frac{v_F}{\sqrt{N}} a \right) f_R \\ & + f_L^\dagger \left(\partial_\tau + i v_F \partial_x - \frac{\partial_y^2}{2m} - \frac{v_F}{\sqrt{N}} a \right) f_L \end{aligned} \quad (\text{A2})$$

and $v_F = k_F/m$. Here a is the x component of the gauge field which couples strongly to this pair of patches. For other patches with normal $\hat{n}(\theta)$ at an angle θ to the x axis, it is the component $\vec{a} \cdot \hat{n}(\theta)$ that will couple strongly. The full theory

is obtained by summing over all patches and adding together the diamagnetic term

$$\mathcal{L}_{\text{dia}} \equiv \frac{a^2}{2mN} (f^\dagger f). \quad (\text{A3})$$

By itself the patch action appears to respect an ‘‘emergent’’ gauge symmetry which acts by

$$a \rightarrow a + \partial\lambda, \quad f_R \rightarrow e^{i\lambda/\sqrt{N}} f_R, \quad f_L \rightarrow e^{i\lambda/\sqrt{N}} f_L \quad (\text{A4})$$

with $\lambda = \lambda(x)$. However, we observe that the gauge field a couples in this two-patch theory to (v_F times) the *axial* current $f_R^\dagger f_R - f_L^\dagger f_L$, which is ‘‘anomalous,’’ as we now review. To diagnose this anomaly, consider the coefficient of a^2 in the effective action resulting from integrating out the fermions. In the patch approximation, the numbers of left- and right-moving fermions are separately conserved. Turning on the gauge field a violates this conservation since the gauge field couples like a chemical potential with opposite sign on the two sides. The change in the ‘‘chiral density’’ of fermions is then the density of states at the Fermi surface times the effective chemical-potential change, $\frac{v_F a}{\sqrt{N}}$.

However in the patch theory, the density of states at the Fermi surface is ill defined. It is apparent from its high-energy origin as a theory with a finite Fermi surface that this description is only valid up to some maximum deviation of the momentum from the middle of the patch. Let the patch size (which equals the cut off for q_y) be denoted as $\Lambda_y = K_F \Delta\theta$, where $\Delta\theta$ is the angular extent of the patch. The density of states in each patch is then $\frac{NK_F \Delta\theta}{4\pi^2 v_F} = \frac{Nm \Delta\theta}{4\pi^2}$. This apparently implies that the one-loop gauge-field polarizability,

$$\Pi(\vec{k}, \omega = 0) = \langle O(\vec{k}, 0) O(-\vec{k}, 0) \rangle_{\text{one-loop}} \quad (\text{A5})$$

depicted in Fig. 3 takes a *nonzero value* as $\vec{k} \rightarrow 0$,

$$\Pi(\vec{k} \rightarrow 0, \omega = 0) = -\frac{m \Delta\theta}{2\pi^2}. \quad (\text{A6})$$

Note that this is the contribution from both patches to the fermion polarizability to the coefficient of $-\frac{v_F^2}{2} a^2$ in the one-loop euclidean effective action for the gauge field. Naively this seems to be a problem as it violates the fake gauge invariance [Eq. (A4)]; more problematically, it also violates the real microscopic gauge invariance. The resolution is that the microscopic gauge invariance is obtained only when the diamagnetic term is also included in the effective action. Indeed if we sum over all patches, the contribution from the polarizability is

$$-\int_0^\pi d\theta \frac{mv_F^2}{4\pi} (\vec{a} \cdot \hat{n})^2 = -\frac{K_F^2}{4\pi m} a^2. \quad (\text{A7})$$

As the density satisfies $\rho_0 = \frac{NK_F^2}{4\pi}$, we see that the nonzero fermion polarizability at $\vec{k} \rightarrow 0, \omega = 0$ exactly cancels the diamagnetic term as required by gauge invariance.

The answer above for the fermion polarizability may be reproduced formally by considering the integral,

$$\Pi(\vec{k}, \omega = 0) \propto \int_p \mathcal{G}(k+p) \mathcal{G}(p). \quad (\text{A8})$$

Actually this integral is ill defined at short distances and depends on the order of integration. The origins of the patch theory from the original full Fermi surface means that we need to impose a cutoff on the x and y momenta but not necessarily on the frequency. Doing the frequency integral by contour integration, we reproduce (in the limit $\vec{k} \rightarrow 0$) the nonzero constant obtained above through a physical argument. Note that imposing a hard cutoff on p_x violates the fake gauge invariance of Eq. (A4). Thus a careful formulation of the patch theory that is faithful to its microscopic origins requires imposing a cutoff on the fermion momenta, and taking the frequency cut off to infinity first before sending the momenta cutoffs to infinity to define the scaling limit.

However as in any scaling theory, we expect that universal low-energy singularities are actually insensitive to how the theory is regularized at short distances (i.e., independent of the ‘‘short-distance completion’’ of the theory). Thus for the purpose of calculating the universal singularities, we are free to choose any regularization of the patch theory. A convenient choice (and the one used in this paper) is to define the scaling theory by sending the momenta cutoffs to infinity first, and then the frequency cutoffs. Then the polarization integral may be done by first integrating over p_x . The result then vanishes as both poles of p_x lie on the same side in the complex plane. Thus in this regularization of the patch theory there is no need to include a diamagnetic term to maintain gauge invariance. Indeed with this regularization, the fake gauge invariance is no longer fake and is a real property of the universal scaling theory.

Though this choice for defining the scaling theory will reproduce all the universal singularities there is no guarantee that it will correctly reproduce nonuniversal ones as it is not faithful to the original microscopic action with the full Fermi surface. A good example to illustrate this is the fermion compressibility, i.e., the response in the density of the system to a change of chemical potential. Within the patch construction, the chemical potential couples to $f_R^\dagger f_R + f_L^\dagger f_L$. This is identical to the coupling of the nematic order parameter. In the single-patch theory, the chemical-potential term is thus identical to the gauge coupling term. The contribution to the compressibility is thus given by the same $\Pi(\vec{k} \rightarrow 0, \omega = 0)$ discussed above. With the regularization actually employed in this paper (which dictates us to do the p_x integral first), this is zero as is indeed required by the gauge invariance of Eq. (A4). In the other regularization (dictated by the original microscopic situation), we do the frequency integral first and get a nonzero answer.

What is the interpretation of the zero answer that doing the p_x integral first produces? Clearly it means that there is no *singular* contribution to the compressibility in the low-energy theory. The nonzero answer found in the other regularization must be viewed as a smooth nonsingular background that is not correctly described by the scale invariant patch theory. The validity of Eq. (A4) in the scaling limit employed in the paper thus ensures that singular contribu-

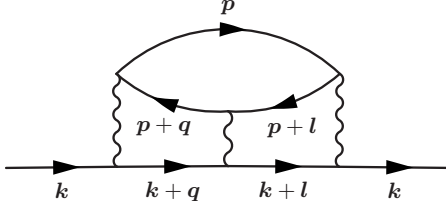


FIG. 13. Three-loop fermion self-energy containing fermions from a single patch.

tions to the compressibility vanish to all orders of perturbation theory in the one-patch theory. Actually the same result is also true in the two-patch theory. To see this note that with the short-distance completion, we have chosen the two-patch action in the presence of an external potential that couples to $f_R^\dagger f_R + f_L^\dagger f_L$ enjoys a low-energy gauge invariance similar to Eq. (A4) but where we rotate the phases of right and left movers with opposite phases $\pm \frac{\lambda}{\sqrt{N}}$. This gauge invariance ensures that the contribution to the fermion compressibility vanishes even in the two-patch theory. For the original microscopic model, this then implies that there is no singular contribution to the background nonzero compressibility.

In Appendix E, we explicitly check that short-distance sensitivity (similar to what happened in the polarizability integral) does not arise in some of the universal quantities of interest, and the answers are insensitive to the order of integration.

APPENDIX B: EXAMPLE DIAGRAM—SINGLE PATCH

Consider the three-loop self-energy contribution shown in Fig. 13 with the external Fermion on the Fermi surface, i.e., $\epsilon_k^\pm \equiv \pm k_x + k_y^2 = 0$ (“+” and “-” for the right and left patch, respectively) for $z_b \rightarrow 2$,

$$\begin{aligned} \Sigma^{(3)} &= \frac{1}{N^2} \int d\vec{l} d\vec{q} d\vec{p} D(\mathbf{q}) D(\mathbf{q} - \mathbf{l}) D(\mathbf{l}) \mathcal{G}_{R/L}(\mathbf{p}) \\ &\quad \times \mathcal{G}_{R/L}(\mathbf{k} + \mathbf{q}) \mathcal{G}_{R/L}(\mathbf{k} + \mathbf{l}) \mathcal{G}_{R/L}(\mathbf{p} + \mathbf{q}) \mathcal{G}_{R/L}(\mathbf{p} + \mathbf{l}), \end{aligned} \quad (\text{B1})$$

where

$$\mathcal{G}_{R/L}^{-1}(\mathbf{k}) = -\frac{i}{\lambda N} \text{sgn}(\omega_k) |\omega_k|^{2/z_b} + \epsilon_k^\pm \quad (\text{B2})$$

and $d\vec{l} \equiv \frac{d\omega_l dl_x dl_y}{(2\pi)^3}$.

Begin by integrating out $\mathbf{p} = (\omega_p, \vec{p})$,

$$\begin{aligned} &\int d\vec{p} \mathcal{G}_{R/L}(\mathbf{p} + \mathbf{q}) \mathcal{G}_{R/L}(\mathbf{p}) \mathcal{G}_{R/L}(\mathbf{p} + \mathbf{l}) \\ &= \frac{\omega_l f(l_y, q_y, \omega_l, \omega_q) - \omega_q f(q_y, l_y, \omega_q, \omega_l)}{q_y (\tilde{\Sigma}(\omega_l) - \epsilon_l^\pm) - l_y (\tilde{\Sigma}(\omega_q) - \epsilon_q^\pm)}, \end{aligned} \quad (\text{B3})$$

where f is a nonsingular function independent of z_b and N . The l_x and q_x integrations are straightforward and we find

$$\begin{aligned} \Sigma^{(3)} &= \frac{i \text{sgn}(\omega_k) |\omega_k|^{2/z_b} \lambda N}{N^2} \int dq_y q_l l_y d\omega_q d\omega_l \text{sgn}(q_y - l_y) \\ &\quad \times \frac{\omega_l \tilde{f}_{q,l} - \omega_q \tilde{f}_{l,q}}{\gamma |\omega_q - \omega_l| + |q_y - l_y|^2} D(\mathbf{q}) D(\mathbf{l}). \end{aligned} \quad (\text{B4})$$

One still needs to verify that the remaining integral is finite. In Ref. 26, it has been shown that all planar diagrams in the single-patch theory are UV finite so it remains to verify IR finiteness. For l finite and $q \rightarrow 0$ or $q \rightarrow l$ this is clear. For $q, l \rightarrow 0$, the integrand diverges as q_y^{-2} which is canceled by the remaining integrations. Thus the three-loop self-energy is indeed of order N^{-2} as long as λN is of order unity.

APPENDIX C: SELF-ENERGIES AT THREE LOOPS

Reference 27 identified the important diagrams that contribute to the boson and fermion self-energies at three-loop level, and evaluated them in appropriate limits when $z_b = 3$. Here we briefly sketch the modifications for general z_b close to 2. Consider the diagrams in Fig. 5 that determine the boson self-energy at three loops. For $z_b \approx 2$, we find

$$\begin{aligned} \Pi &= |k_y|^{z_b-1} \frac{2\lambda N \gamma}{N\pi} \int_0^1 ds \int_0^\infty dt \frac{(1-s)^3 t}{t+1} \frac{s}{t(1-s)^2 + s^2} \\ &\quad \times \frac{(s\gamma\lambda N)^2}{(1-s)^2 t^2 + (s\gamma\lambda N)^2}. \end{aligned} \quad (\text{C1})$$

In the physical limit $\gamma\lambda = \frac{\pi}{2}$, $N=2$ we evaluate the integral numerically to find

$$\Pi = 0.106 |k_y|^{z_b-1}. \quad (\text{C2})$$

Now consider the diagrams for the three-loop fermion self-energy contributions depicted in Fig. 7.

They are given by

$$\begin{aligned} \Sigma_{1,2}^{(3)} &= -\frac{1}{N^2} \int_{\bar{\omega}_l, \vec{l}} \int_{\bar{\omega}_q, \vec{q}} \int_{\bar{\omega}_p, \vec{p}} D(\mathbf{q}) D(\mathbf{p}) D(\mathbf{q} - \mathbf{p}) \\ &\quad \times \mathcal{G}_R(\mathbf{k} \pm \mathbf{p}) \mathcal{G}_R(\mathbf{k} \pm \mathbf{q}) \mathcal{G}_L(\mathbf{l}) \mathcal{G}_L(\mathbf{l} + \mathbf{q}) \mathcal{G}_L(\mathbf{l} + \mathbf{p}). \end{aligned} \quad (\text{C3})$$

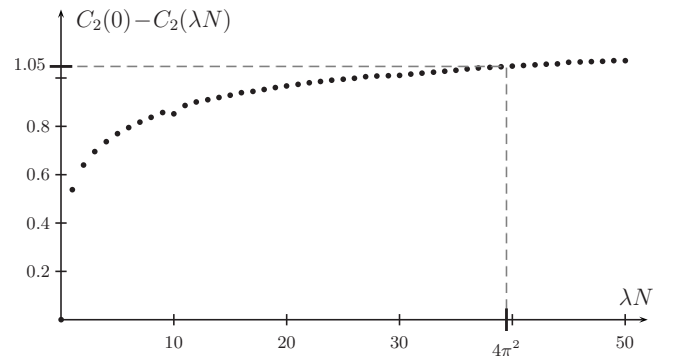


FIG. 14. Numerical integral determining the fermion anomalous dimension η_f .

The integrals may be evaluated following Ref. 27, the results are, to leading order in ϵ_k

$$\Sigma_1^{(3)}(\epsilon_k) = \frac{\lambda^2}{12\pi^4} C_{z_b}(0) \epsilon_k \ln \frac{\Lambda}{\epsilon_k}, \quad (\text{C4})$$

where

$$\Sigma_2^{(3)}(\epsilon_k) = \Sigma_2^{(3)}(0) - \frac{\lambda^2}{12\pi^4} C_{z_b}(\lambda N) \epsilon_k \ln \frac{\Lambda}{\epsilon_k}, \quad (\text{C5})$$

$$\begin{aligned} C_{z_b}(\lambda N) &= \frac{96}{(4\pi)^{4/z_b}} \int_1^\infty dx \int_0^\infty dy \int_0^\infty ds \int_s^\infty dt \frac{ts(s-t)^2}{(x+s^{z_b})(y+t^{z_b})[x+y+(t-s)^{z_b}]} \\ &\times \frac{\{t[(x-1)^{2/z_b} + x^{2/z_b} + 1] + s[(y+1)^{2/z_b} + y^{2/z_b} - 1]\}^2 - N^2 \lambda^2 (4\pi)^{-2/z_b} s^2 t^2 (s-t)^2}{(\{t[(x-1)^{2/z_b} + x^{2/z_b} + 1] + s[(y+1)^{2/z_b} + y^{2/z_b} - 1]\}^2 + N^2 \lambda^2 (4\pi)^{-2/z_b} s^2 t^2 (s-t)^2)} \\ &+ \frac{48}{(4\pi)^{4/z_b}} \int_1^\infty dx \int_1^\infty dy \int_0^\infty ds \int_0^\infty dt \frac{ts(s+t)^2}{(x+s^{z_b})(y+t^{z_b})[|x-y| + (t+s)^{z_b}]} \\ &\times \frac{\{t[(x-1)^{2/z_b} + x^{2/z_b} + 1] + s[(y-1)^{2/z_b} + y^{2/z_b} + 1]\}^2 - N^2 \lambda^2 (4\pi)^{-2/z_b} s^2 t^2 (s+t)^2}{(\{t[(x-1)^{2/z_b} + x^{2/z_b} + 1] + s[(y-1)^{2/z_b} + y^{2/z_b} + 1]\}^2 + N^2 \lambda^2 (4\pi)^{-2/z_b} s^2 t^2 (s+t)^2)}. \end{aligned} \quad (\text{C6})$$

The function $J(\lambda N)$ which appears in Eq. (24) is the sum of the singular contributions at $z_b=2$, i.e.,

$$J(\lambda N) = \left(\frac{\lambda N}{4\pi^2} \right)^2 (C_2(0) - C_2(\lambda N)) \quad (\text{C7})$$

(note that the prefactor is unity at $z_b=3$, $N=2$). A numerical estimate of this function is shown in Fig. 14.

APPENDIX D: VERTEX INTEGRAL

We want to evaluate the integrals shown in Fig. 15, where all external momenta and frequencies are put to zero. The computation is straightforward,

$$\begin{aligned} \delta u &= -\frac{1}{N} \int_{\vec{k}, \omega} \frac{D(k_y, \omega)}{\text{isgn}(\omega) \frac{|\omega|^{2/z_b}}{\lambda N} - \epsilon_k^+ \text{isgn}(\omega) \frac{|\omega|^{2/z_b}}{\lambda N} - \epsilon_k^-} = \\ &= -\frac{1}{2N} \int_{k_y, \omega} \frac{\frac{1}{\lambda N} |\omega|^{2/z_b}}{\frac{1}{\lambda^2 N^2} |\omega|^{4/z_b} + k_y^4} \frac{|k_y|}{\gamma |\omega| + |k_y|^{z_b}} = \\ &= -\frac{4}{2N(2\pi)^2} \int_{\Lambda/\sqrt{b}}^\Lambda dk_y \frac{1}{k_y} \int_0^\infty dt \frac{\frac{1}{\lambda N} t^{2/z_b}}{\frac{1}{\lambda^2 N^2} t^{4/z_b} + 1} \frac{1}{\gamma t + 1} \end{aligned}$$

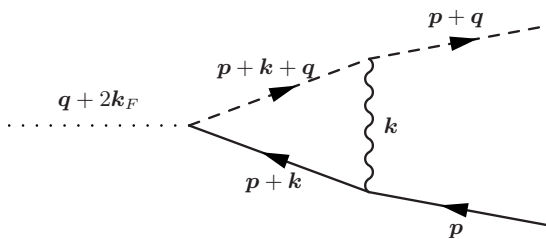


FIG. 15. One-loop correction to the $2k_F$ vertex. The dotted line denotes a $2k_F$ density fluctuation.

$$= \frac{1}{4\pi^2 N} \ln b \int_0^\infty dt \frac{\lambda N t^{2/z_b}}{t^{4/z_b} + \lambda^2 N^2} \frac{1}{\gamma t + 1}. \quad (\text{D1})$$

The corresponding expression at the perturbative (small ϵ) fixed point is given by

$$\delta u = \frac{1}{4\pi^2 N} \ln b \int_0^\infty dt \frac{t^{1-\epsilon/2}}{t^{2-\epsilon} + 1} \frac{1}{t + \frac{2}{\pi \epsilon N}}. \quad (\text{D2})$$

The integral remains finite if we take the limit $\epsilon \rightarrow 0$ in the first term in the integrand so we can evaluate it analytically. We obtain to leading order in ϵ ,

$$\delta u = \frac{\epsilon}{8\pi} \ln \frac{2}{\pi \epsilon N} \ln b. \quad (\text{D3})$$

APPENDIX E: ORDER OF INTEGRATION

Consider as a concrete example the three-loop self-energy depicted in Fig. 7. We want to keep a cutoff on they y components of momenta and verify that we can interchange the order in which we perform the integration over frequencies and momentum x components. As the only possible divergencies arise from the UV, we can set all y components as well as external momenta and frequencies to zero (the y components in the numerators of the gauge propagators are set to unity). Recall that any two integrals are independent of the order of integration, if the double integral is absolutely convergent. So consider

$$\int_{l_x, p_x, q_x, \omega_l, \omega_p, \omega_q} |D_{\omega_q} D_{\omega_q - \omega_l} D_{\omega_l} \mathcal{G}_p \mathcal{G}_q \mathcal{G}_l \mathcal{G}_{p+q} \mathcal{G}_{p+l}|, \quad (\text{E1})$$

$$\begin{aligned} & \sim \int d\omega_l \frac{|\omega_l|^{6/z_b}}{|\omega_l|^{10/z_b}} \int d\omega_p \frac{|\omega_p|^{1+6/z_b}}{|\omega_p|^{2+10/z_b}} \int d\omega_q \frac{|\omega_q|^{2+6/z_b}}{|\omega_q|^{3+10/z_b}} \\ & \times \int dp_x \frac{1}{|p_x|^2} \int dq_x \frac{1}{|q_x|^2} \int dl_x \frac{|l_x|^2}{|l_x|^5}. \end{aligned} \quad (\text{E2})$$

Since each of these integrals is individually convergent, the original integrations can be performed in any order. This

analysis applies to most of the diagrams shown here and is a result of the Landau-damping form of the gauge propagator, thus it does not extend to diagrams calculated at the small- ϵ perturbative fixed point. In our proposed expansion $N \rightarrow \infty$, $\lambda N = o(1)$, of the diagrams shown in this paper only the one-loop boson self-energy (Fig. 3) is sensitive to the order of integration.

-
- ¹O. I. Motrunich, *Phys. Rev. B* **72**, 045105 (2005).
²S.-S. Lee and P. A. Lee, *Phys. Rev. Lett.* **95**, 036403 (2005).
³B. I. Halperin, P. A. Lee, and N. Read, *Phys. Rev. B* **47**, 7312 (1993).
⁴P. A. Lee, *Phys. Rev. Lett.* **63**, 680 (1989).
⁵V. M. Galitski, G. Refael, M. P. A. Fisher, and T. Senthil, *Phys. Rev. Lett.* **95**, 077002 (2005).
⁶R. K. Kaul, A. Kolezhuk, M. Levin, S. Sachdev, and T. Senthil, *Phys. Rev. B* **75**, 235122 (2007).
⁷R. K. Kaul, Y. B. Kim, S. Sachdev, and T. Senthil, *Nat. Phys.* **4**, 28 (2008).
⁸O. I. Motrunich and M. P. A. Fisher, *Phys. Rev. B* **75**, 235116 (2007).
⁹I. J. Pomeranchuk, *Sov. Phys. JETP* **8**, 361 (1958).
¹⁰V. Oganesyan, S. A. Kivelson, and E. Fradkin, *Phys. Rev. B* **64**, 195109 (2001).
¹¹W. Metzner, D. Rohe, and S. Andergassen, *Phys. Rev. Lett.* **91**, 066402 (2003); Luca Dell'Anna and Walter Metzner, *Phys. Rev. B* **73**, 045127 (2006).
¹²H.-Y. Kee, E. H. Kim, and C.-H. Chung, *Phys. Rev. B* **68**, 245109 (2003).
¹³M. J. Lawler, D. G. Barci, V. Fernández, E. Fradkin, and L. Oxman, *Phys. Rev. B* **73**, 085101 (2006); M. J. Lawler and E. Fradkin, *ibid.* **75**, 033304 (2007).
¹⁴J. Rech, C. Pèpin, and A. V. Chubukov, *Phys. Rev. B* **74**, 195126 (2006).
¹⁵P. Wölfle and A. Rosch, *J. Low Temp. Phys.* **147**, 165 (2007).
¹⁶D. L. Maslov and A. V. Chubukov, *Phys. Rev. B* **81**, 045110 (2010).
¹⁷J. Quintanilla and A. J. Schofield, *Phys. Rev. B* **74**, 115126 (2006).
¹⁸C. A. Lamas, D. C. Cabra, and N. Grandi, *Phys. Rev. B* **78**, 115104 (2008).
¹⁹H. Yamase and W. Metzner, *Phys. Rev. B* **75**, 155117 (2007).
²⁰E. Fradkin, S. Kivelson, M. Lawler, J. Eisenstein, and A. Mackenzie, [arXiv:0910.4166](https://arxiv.org/abs/0910.4166), *Ann. Rev. of Cond. Matter Phys.* (to be published).
²¹R. Borzi, S. A. Grigera, J. Farrell, R. S. Perry, S. J. S. Lister, S. L. Lee, D. A. Tennant, Y. Maeno, and A. P. Mackenzie, *Science* **315**, 214 (2007).
²²Y. Ando, K. Segawa, S. Komiyama, and A. N. Lavrov, *Phys. Rev. Lett.* **88**, 137005 (2002).
²³V. Hinkov, D. Haug, B. Fauquè, P. Bourges, Y. Sidis, A. Ivanov, C. Bernhard, C. T. Lin, and B. Keimer, *Science* **319**, 597 (2008).
²⁴Y. Kohsaka, C. Taylor, K. Fujita, A. Schmidt, C. Lupien, T. Hanaguri, M. Azuma, M. Takano, H. Eisaki, H. Takagi, S. Uchida, and J. C. Davis, *Science* **315**, 1380 (2007).
²⁵R. Daou, J. Chang, D. LeBoeuf, O. Cyr-Choinière, F. Laliberté, N. Doiron-Leyraud, B. Ramshaw, R. Liang, D. A. Bonn, W. N. Hardy and L. Taillefer, *Nature (London)* **463**, 519 (2010).
²⁶S.-S. Lee, *Phys. Rev. B* **80**, 165102 (2009).
²⁷M. Metlitski and S. Sachdev, [arXiv:1001.1153](https://arxiv.org/abs/1001.1153) (unpublished).
²⁸T. Holstein, R. E. Norton, and P. Pincus, *Phys. Rev. B* **8**, 2649 (1973).
²⁹M. Y. Reizer, *Phys. Rev. B* **40**, 11571 (1989).
³⁰P. A. Lee and N. Nagaosa, *Phys. Rev. B* **46**, 5621 (1992).
³¹Y. B. Kim, P. A. Lee, and X.-G. Wen, *Phys. Rev. B* **52**, 17275 (1995); Y. B. Kim, A. Furusaki, X. G. Wen, and P. A. Lee, *ibid.* **50**, 17917 (1994).
³²B. L. Altshuler, L. B. Ioffe, and A. J. Millis, *Phys. Rev. B* **50**, 14048 (1994).
³³J. Polchinski, *Nucl. Phys. B* **422**, 617 (1994).
³⁴T. Senthil, *Phys. Rev. B* **78**, 045109 (2008).
³⁵When can one apply this technique of introducing a control parameter by varying a dynamical exponent? We expect it to be effective when one has prior knowledge that the exponent in question is not renormalized, as is the case here. Varying the exponent so that the relevant term in the effective action is non-local will guarantee that the exponent will not be renormalized in perturbation theory since integrating out short-wavelength modes cannot generate the necessary IR singularity. A contrasting example to the one discussed here is the Ising fixed point: the magnetization field indeed acquires an anomalous dimension, and one would obscure the physics by artificially preventing this.
³⁶C. Nayak and F. Wilczek, *Nucl. Phys. B* **417**, 359 (1994); **430**, 534 (1994).
³⁷T. Senthil, *Phys. Rev. B* **78**, 035103 (2008).
³⁸For the case of the nematic, we absorb the dependence on the angle between the nematic ordering vector and the patch in question into the coupling e . This coupling e specifies an energy scale below which our low-energy theory is applicable. Since this energy scale vanishes at the “cold spots,” we must restrict attention to a patch of the Fermi surface away from this direction.
³⁹R. Shankar, *Rev. Mod. Phys.* **66**, 129 (1994).
⁴⁰S.-W. Tsai, A. H. Castro Neto, R. Shankar, and D. K. Campbell, *Phys. Rev. B* **72**, 054531 (2005).
⁴¹A. Stern and B. I. Halperin, *Phys. Rev. B* **52**, 5890 (1995).
⁴²D. F. Mross and T. Senthil (unpublished).
⁴³We thank Max Metlitski for helpful correspondence on this point.
⁴⁴M. R. Norman and T. Micklitz, *Phys. Rev. Lett.* **102**, 067204

- (2009).
- ⁴⁵S. Lee, *Phys. Rev. D* **79**, 086006 (2009).
- ⁴⁶H. Liu, J. McGreevy, and D. Vegh, [arXiv:0903.2477](https://arxiv.org/abs/0903.2477) (unpublished).
- ⁴⁷M. Cubrovic, J. Zaanen, and K. Schalm, *Science* **325**, 439 (2009).
- ⁴⁸T. Faulkner, H. Liu, J. McGreevy, and D. Vegh, [arXiv:0907.2694](https://arxiv.org/abs/0907.2694) (unpublished).
- ⁴⁹T. Faulkner, N. Iqbal, H. Liu, J. McGreevy, and D. Vegh (unpublished).
- ⁵⁰T. Faulkner and J. Polchinski, [arXiv:1001.5049](https://arxiv.org/abs/1001.5049) (unpublished).