Two dimensional non-Fermi liquid metals: a solvable large N limit

Jeremias Aguilera Damia,[1] Shamit Kachru,[2] Srinivas Raghu,[2, 3] and Gonzalo Torroba[4]

[1]*Centro Atómico Bariloche, CNEA and CONICET, Bariloche, R8402AGP, Argentina*
[2]*Stanford Institute for Theoretical Physics, Stanford University, Stanford, CA 94305, USA*
[3]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[4]*Centro Atómico Bariloche and CONICET, R8402AGP, Argentina*
(Dated: May 22, 2019)

Significant effort has been devoted to the study of "non-Fermi liquid" (NFL) metals: gapless conducting systems that lack a quasiparticle description. One class of NFL metals involves a finite density of fermions interacting with soft order parameter fluctuations near a quantum critical point. The problem has been extensively studied in a large N limit (N corresponding to the number of fermion flavors) where universal behavior can be obtained by solving a set of coupled saddle-point equations. However a remarkable study by S.-S. Lee revealed the breakdown of such approximations in two spatial dimensions. We show that an alternate approach, in which the fermions belong to the fundamental representation of a global $SU(N)$ flavor symmetry, while the order parameter fields transform under the adjoint representation (a "matrix large N" theory), yields a tractable large N limit. At low energies, the system consists of an overdamped boson with dynamical exponent $z = 3$ coupled to a non-Fermi liquid with self energy $\Sigma(\omega) \sim \omega^{2/3}$, consistent with previous studies.

I. INTRODUCTION

In many strongly correlated quantum materials, continuous phase transitions into a broken symmetry phase occur at zero temperature as a function of pressure, doping and other non-thermal tuning parameters. At such a quantum critical point [1], the metallic fermions scatter off of nearly critical fluctuations of the order parameter, and new universal behavior, inconsistent with Landau's Fermi liquid paradigm, can occur. Understanding such non-Fermi liquid (NFL) behavior [2, 3] and its relation to high-temperature superconductivity is one of the central challenges of theoretical physics. We study a class of quantum critical points that preserve the underlying lattice translational symmetry and are not associated with a conserved order parameter –an example is the Ising nematic transition, which has been observed in several iron-based superconductors [4, 5], and may play a role in other materials as well [6, 7].

Near the quantum critical point, only the slowest modes are important; the problem can thus be recast into a quantum field theory involving fermions near the Fermi level coupled to a critical boson (order parameter) by the lowest order interaction allowed by symmetry. The leading interaction is a Yukawa-type coupling, which is relevant in the renormalization group sense below 3 space dimensions. As a consequence the theory is strongly coupled in 2 space dimensions, the limit applicable to many quasi-two dimensional quantum materials. While in recent years numerical methods have revealed a variety of strong coupling effects in two dimensions – for instance via sign problem-free quantum Monte Carlo simulations [8–10]– an analytic solution based on a controlled expansion remains elusive.

Given the absence of a perturbative coupling, it is natural to look for a large N expansion to restrict the class of quantum effects that contribute. One possibility is to extend the number of fermion spins from 2 to N, and have them interact with a singlet scalar mode; this "vector large N limit" has been intensely studied in the literature [11–18]. However, it was shown in [15] that the theory remains strongly coupled due to quantum enhancements at two and higher loops. As a result, the $1/N$ expansion is not enough to make the dynamics tractable. There exist extensions of this limit that end up being controlled, but this is achieved at the price of adding some new perturbatively small parameter by hand [19–23].

In this work we will instead focus on the "matrix large N limit," where N fermion flavors interact with an $N \times N$ matrix-valued boson. This $1/N$ expansion was originally introduced in the context of relativistic quantum field theory, in order to study Yang-Mills theory [24]. It was first applied to NFLs in [25, 26], and a controlled quantum critical point was shown to arise in an ϵ expansion around $d = 3$ spatial dimensions [27, 28]. We will study this $1/N$ expansion directly in two spatial dimensions and at zero temperature, finding an exactly solvable critical point with non-Fermi liquid behavior. The exact solution consists of an overdamped order parameter field with dynamical exponent $z_b = 3$, coupled to a non-Fermi liquid metal with fermion dynamical exponent $z_f = 3/2$. Similar solutions have been obtained both in direct perturbation theory [14], and in the vector large N limit.[1] Here, however, they correspond to a controlled and asymptotically exact solution of an infrared fixed point. Our results thus provide a controlled framework for understanding non-Fermi liquid behavior.

The paper is organized as follows. In Sec. II we present the model and discuss the one-loop QCP. In Sec. III we extend the validity of the QCP to all orders in the $1/N$ expansion. We do this by determining a low energy limit

---

[1] See also [29, 30] for other methods that give similar self-energy effects.

where the standard large N counting of planar and non-planar diagrams applies. In Sec. IV we compare with the vector large N expansion, which remains strongly coupled; we track the difference to the qualitatively different behavior of the 't Hooft coupling. We also compare our framework to the holographic approach to non-Fermi liquids, and propose future directions of research. In the Appendices we present an alternative and equivalent renormalization-group analysis, as well as a scaling analysis of a more general model that includes the vector and matrix large N expansions.

II. THE ONE LOOP CRITICAL POINT

Our euclidean action involves a two-dimensional system consisting of fermions $(\psi, \bar{\psi})$ at finite density interacting with a critical boson ϕ:

$$
\begin{aligned}
S \; = \; & \int d\tau d^2x \left\{ \frac{1}{2} \text{Tr} \left[\frac{1}{c^2} \left(\partial_\tau \phi \right)^2 + \left(\nabla \phi \right)^2 \right] \right. \\
& \left. + \; \psi_i^\dagger \left(\partial_\tau + \varepsilon(i\nabla) - \mu_F \right) \psi^i + \frac{g}{\sqrt{N}} \phi_j^i \psi_i^\dagger \psi^j \right\}.
\end{aligned}
\tag{1}
$$

To facilitate an asymptotically exact solution, we impose a global $SU(N)$ flavor symmetry, with $\psi_i, i = 1, \cdots N$ transforming in the fundamental, and $\phi_j^i, i, j = 1 \cdots N$ in the adjoint representation. Here, we have tuned to criticality by switching off the boson mass, c is the boson speed, the fermion has a dispersion relation $\varepsilon(\vec{k})$ and chemical potential μ_F, and the two fields are coupled via a cubic Yukawa interaction. This is the most relevant interaction consistent with the symmetries, and we will show that other interactions, such as the boson ϕ^4 and the BCS coupling, are irrelevant at the fixed point.

We will first analyze the critical point that arises at one loop, and in Sec. III we will show that all the other corrections vanish in $1/N$. So the fixed point will turn out to be one-loop exact in $1/N$.

The kinematics of the Fermi surface and its coupling to the boson will play an important role in the long distance dynamics. So let us first review the decomposition of fermionic and bosonic momenta. A given point on the one-dimensional Fermi surface is parametrized by the Fermi surface radius k_F and a unit vector $\hat{n}$. The fermionic momentum is then written as a radial fluctuation [31], $\vec{p} = \hat{n}(k_F + p_\perp)$. The Yukawa interaction implies that the boson momentum $\vec{q}$ behaves as a difference of fermion momenta. Near the point $\hat{n}$ on the Fermi surface, we will decompose $\vec{q} = q_\perp \hat{n} + \vec{q}_\parallel$, and will often denote the relative angle by $\cos\theta = \vec{q} \cdot \hat{n}/q$.

One loop quantum effects induce boson and fermion self-energy corrections; see Fig. 1. A standard calculation gives the boson self-energy (Landau damping)

$$
\Pi(q_0, q) = \frac{k_F}{N} \frac{g^2}{2\pi v} \frac{|q_0|}{\sqrt{q_0^2 + (vq)^2}}.
\tag{2}
$$

While this is a $1/N$ effect, we will include it because it dominates at low energies. Including the effects of
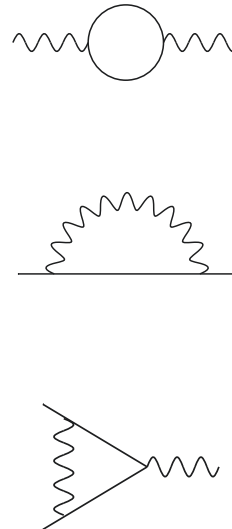

FIG. 1: One-loop quantum effects: boson self-energy (top), fermion self-energy (middle), and vertex renormalization (bottom). Boson and fermion propagators are denoted by wavy lines and straight lines, respectively.

$\Pi(q_0, q)$, the boson spectral weight dominantly arises from the kinematic regime $|q_0| < vq$, where the characteristic boson speed is slow compared to that of the fermion, and where the boson mixes with the continuum of particle-hole excitations of the Fermi surface. As a result, the boson gets overdamped, and combining (2) with (1) gives, at low energies, a boson with $z = 3$ scaling, $q^3 \sim M_D^2 |q_0|$. Here we have introduced the Landau damping scale

$$
M_D^2 \equiv \frac{k_F}{N} \frac{g^2}{2\pi v^2}.
\tag{3}
$$

We will then work with the resummed bosonic propagator [1, 27]

$$
D^{-1}(q_0, q) \approx q^2 + M_D^2 \frac{|q_0|}{q},
\tag{4}
$$

which will be shown to be self-consistent.

The computation of the fermion self-energy using this resummed overdamped boson propagator is standard and results in the following expression:

$$
\Sigma(p_0) = \frac{g^2}{2\pi\sqrt{3}v} \frac{1}{M_D^{2/3}} \text{sgn}(p_0) \, |p_0|^{2/3}.
\tag{5}
$$

The self-energy is a regular function of momentum, which we have not included here, since it becomes irrelevant due to the $z = 3$ scaling of the boson internal line. The remaining one loop effect, the vertex correction, is suppressed by $1/N$, analogous to a "Migdal" approximation in the electron-phonon problem, and can be neglected.

Eqs. (4) and (5) describe a nontrivial QCP, where the radial fermionic momentum scales differently from the bosonic momentum [27, 28, 32]. It is not hard to check that the scale transformations

$$\omega \to \lambda\omega \,, \ q_\perp \to \lambda^{2/3}q_\perp \,, \ q_\parallel \to \lambda^{1/3}q_\parallel \,, \qquad (6)$$

and

$$\phi(q_0, q) \to \lambda^{-4/3}\phi(q_0, q) \,, \ \psi(q_0, q) \to \lambda^{-7/6}\psi(q_0, q) \quad (7)$$

leave the IR effective action (which includes the above self-energy corrections) invariant. As a result, we obtain a quantum critical point where the boson has scaling dimension and dynamical exponent $(\Delta_\phi = -\frac{4}{3}, \ z_b = 3)$, and for the fermion, $(\Delta_\psi = -\frac{7}{6}, \ z_f = \frac{3}{2})$.[2] The only relevant coupling (besides the chemical potential) is the boson mass, which we tune to criticality. The Yukawa interaction becomes marginal at the fixed point, while 4-boson and 4-Fermi interactions are irrelevant. (This is why we neglected them from the beginning). We note that this fixed point agrees with the $\epsilon = 1$ limit of the NFL studied in [27, 28], in $d = 3 - \epsilon$ dimensions.

Finally, let us determine the energy scale below which we flow to the one-loop QCP. This is the crossover at which the quantum self-energies begin to dominate over the tree-level kinetic terms. This happens when the $z = 3$ regime is reached, which requires $q_0^2/c^2 \lesssim \Pi(q_0, q)$ and $q_0^2 \lesssim (vq)^2$. Assuming we are near the mass-shell condition $q^3 \sim M_D^2 q_0$, this gives an energy scale

$$E \lesssim c\frac{\sqrt{k_F g^2}}{N^{1/2}} \ \min(\frac{c^{1/2}}{v}, \frac{v^{1/2}}{c})\,. \qquad (8)$$

III. QUANTUM CRITICALITY AT ALL ORDERS IN $1/N$

Including the self-energy effects described in the previous section, we obtain a one-loop QCP with effective Lagrangian

$$L_{eff} = L_f + L_b + L_Y \qquad (9)$$

where

$$L_f = \int dp_\perp (k_F d\hat{n}) \psi_{\hat{n}}^\dagger \left(i\beta N^{1/3}\mathrm{sgn}(p_0)|p_0|^{2/3} - vp_\perp\right) \psi_{\hat{n}}$$

$$L_b = \int dq_\perp dq_\parallel \phi \left(q^2 + \frac{\gamma}{N}\frac{|q_0|}{q}\right) \phi \qquad (10)$$

$$L_Y = \frac{g}{\sqrt{N}} \int dq_\perp dq_\parallel dp_\perp \, k_F d\hat{n} \, \phi(q)\psi_{\hat{n}}^\dagger(p+q)\psi_{\hat{n}}(p)\,.$$

Here we have introduced the combinations

$$\beta = \frac{1}{(2\pi)^{2/3}3^{1/2}} \left(\frac{g^4}{vk_F}\right)^{1/3} \,, \ \gamma = \frac{k_F g^2}{2\pi v^2}\,. \qquad (11)$$

Using the $1/N$ expansion, we now want to extend this to all loop orders. This, however, encounters some problems due to the fact that the explicit N dependence in the propagators precludes the standard large N counting of planar and non-planar diagrams. In particular, some terms that are irrelevant by the power-counting of (6), are actually enhanced by N. A simple example occurs in the bosonic propagator. Here $q_\perp^2$ is irrelevant compared to $q_\parallel^2$, but the N-scaling, dictated by the on-shell conditions, is $q_\perp \sim N^{1/3}q_0^{2/3}$, $q_\parallel \sim 1/N^{1/3}q_0^{1/3}$. So $q_\perp^2 \gg q_\parallel^2$ at fixed energy. In other words, the low energy limit does not commute with the large N limit.

We will now argue that the low energy and large N limits can be taken simultaneously, if the external frequencies and momenta scale in a specific way with N. To see this, we note that the previous problem –the large N limit ruining the $z = 3$ scaling– is resolved if the low energy limit is taken as $q_0 \sim 1/N^2$. Indeed, this makes $q_\perp$ and $q_\parallel$ above scale with the same power of N. Therefore, we will consider the redefinition

$$p_0 = \frac{1}{N^2}\tilde{p}_0 \,, \ p_\perp = \frac{\beta}{N}\tilde{p}_\perp \,, \ p_\parallel = \frac{\gamma^{1/3}}{N}\tilde{p}_\parallel\,. \qquad (12)$$

We will show that correlation functions with fixed $(\tilde{p}_0, \tilde{p}_i)$ are described by a QCP that is one-loop exact in the $1/N$ expansion. Before proceeding, we also note that we have introduced factors of β, γ in (12), so that the engineering dimensions of the new variables, $[\tilde{p}_0] = 1$, $[\tilde{p}_\perp] = 2/3$, $[\tilde{p}_\parallel] = 1/3$, match the scaling dimensions (6) of the one-loop fixed point.

The redefinition (12) produces overall powers of N and (β, γ) in the two-point functions. However, these factors cause no problem, as they can be absorbed into the redefinition of fields. The canonically normalized fields, where these factors are absorbed, become

$$\chi_{\hat{n}} = \frac{\beta k_F^{1/2}}{N^2}\psi_{\hat{n}} \,, \ \varphi = \frac{(\beta\gamma)^{1/2}}{N^3}\phi\,. \qquad (13)$$

Given the engineering dimensions (in Fourier space) $[\psi] = -2$, $[\phi] = -5/2$, the dimensions of the canonical fields become $[\chi] = -7/6$, $[\varphi] = -4/3$. As expected, these agree with the scaling dimensions (7). The last step replaces these redefinitions in the Yukawa coupling; the resulting effective action $S_{eff} = S_f + S_b + S_Y$ reads

$$S_f = \int dp_0 dp_\perp \, d\hat{n} \, \chi_{\hat{n}}^\dagger \left(i\mathrm{sgn}(p_0)|p_0|^{2/3} - vp_\perp\right) \chi_{\hat{n}}$$

$$S_b = \int dq_0 dq_\perp dq_\parallel \, \varphi \left(q^2 + \frac{|q_0|}{q}\right) \varphi \qquad (14)$$

$$S_Y = \frac{g_*}{\sqrt{N}} \int dq_0 dp_0 dq_\perp dq_\parallel dp_\perp \, d\hat{n} \, \varphi(q)\chi_{\hat{n}}^\dagger(pq)\chi_{\hat{n}}(p)\,,$$

and we have dropped all the tildes from the frequencies and momenta. The coupling evaluates to

$$\frac{g_*^2}{v} = 2\pi\sqrt{3}\,. \qquad (15)$$

---

[2] These are the dimensions in momentum space representation, as in (7).

This plays the role of the 't Hooft coupling at the fixed point. In Appendix A, we show that the above fixed point action (14) can equally well be captured by a renormalization group treatment – see, for instance Eq. (A6). Indeed, the scalings and redefinitions that we just performed are automatically included in the RG approach in terms of the running parameters.

$$\Pi(q_0, q) = \frac{g^2}{N} \int \frac{dk_0}{2\pi} \frac{dk_\perp}{2\pi} \frac{d\theta}{2\pi} \frac{1}{ik_0 + i\Sigma(k_0) - vk_\perp} \frac{1}{i(k_0 + q_0) + i\Sigma(k_0 + q_0) - v(k_\perp + q\cos\theta)}$$
$$i\Sigma(p_0) = -g^2 \int \frac{dq_0}{2\pi} \frac{qdq}{2\pi} \frac{d\theta}{2\pi} \frac{1}{q^2 + \Pi(q_0, q)} \frac{1}{iq_0 + i\Sigma(q_0) - vq\cos\theta}. \tag{16}$$
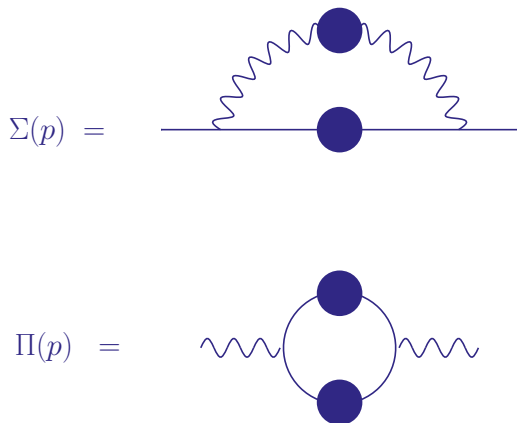
$$\Sigma(p) =$$

$$\Pi(p) =$$

FIG. 2: Schwinger-Dyson equations for the fermion and boson self-energies, neglecting vertex corrections in the large N theory.

Diagrammatically, replacing the one-loop contributions in the right hand side of (16) gives rise to the two-loop rainbow diagrams for Π and Σ, and this continues by induction to higher-loop planar diagrams. This is illustrated diagrammatically in Fig. 2. By explicit calculation, (4) and (5) provide a solution to (16), so long as the low energy limit is taken as in (12).[3] In summary, the one loop result is a self-consistent solution to the Scwhinger-Dyson equations, and all planar contributions beyond one loop vanish in the low energy limit (12).

On the other hand, all nonplanar corrections to the QCP are explicitly suppressed by powers of $1/N$. This can be seen directly from (14): the usual large N counting of diagrams applies, because N only appears in the cubic

---

[3] Above this window, the $z_b = 3$ and $z_f = 3/2$ scalings are not preserved.

Since the fixed point theory has an order one 't Hooft coupling, we expect that we have to resum all planar diagrams that contribute to (14). Fortunately, they all vanish beyond one loop. This can be seen by noting that planar corrections to the self-energies are resummed in terms of the Schwinger-Dyson equations

interaction and not inside the two-point functions. This is a consequence of the way in which the low energy and large N limits are taken in (12).

Let us also mention that the tree-level irrelevant contributions to the kinetic terms –the boson and fermion frequency terms, and the higher order term $p_\perp^2/k_F$ in the fermion dispersion relation– can also be seen to be suppressed by powers of $1/N$ compared to the critical terms. The same occurs with terms that have four or more fields in the action. As a result, none of the irrelevant corrections to the QCP are enhanced by powers of N.

We conclude that the one-loop QCP (14) is exact to all orders in $1/N$. It arises in the simultaneous large N and low energy limit dictated by (12), with $(\tilde{p}_0, \tilde{p}_i)$ fixed. This QCP thus provides an example of a solvable non-Fermi liquid in two spatial dimensions.

IV. DISCUSSION AND CONCLUSIONS

We have shown above that the matrix large N limit provides a controlled set of solutions describing the two dimensional quantum critical metal. This was achieved by taking a simultaneous large N and low energy limit (12). The solvability of the $1/N$ expansion may appear surprising, both from previous results on the vector large N limit [15], and because in general it is very hard to resum the planar expansion in relativistic quantum field theory [24, 33, 34]. In order to address this, let us now briefly discuss the problem from the viewpoint of the renormalization group (RG).

The self-consistency of the quantum effective action (14) implies an IR stable RG fixed point. In App. A, we show how this result can equally well be captured by a renormalization group treatment; see for instance Eq. (A6). We summarize here the essential features. In the vicinity of the fixed point, the one-loop beta function for the combination $\alpha \sim g^2/v$ is

$$-\mu \frac{d\alpha}{d\mu} = c_1\alpha - c_2\alpha^2, \tag{17}$$

where, μ is the sliding energy scale (the RG flow parameter), and c_1, c_2 are positive order one constants. The first term above describes the tree-level scaling behavior of α at low energies, while the second term contains the effects of quantum self-energy corrections (recall that vertex corrections can be neglected in the large N limit). As a consequence, there is an IR stable fixed point with an order unity fixed point value $\alpha_* \sim \mathcal{O}(1)$. This fixed point precisely corresponds to the action (14), where a $z = 3$ boson is coupled to a non-Fermi liquid with an order unity 't Hooft coupling (15).

By contrast, in the vector large N limit, the fermion self-energy is a $1/N$ correction. As a consequence, the analogous RG flows are described by an equation of the form

$$-\mu\frac{d\alpha}{d\mu} = c_1'\alpha - \frac{c_2'}{N}\alpha^2, \tag{18}$$

and the resulting fixed point value corresponds to $\alpha_* \sim N$. This theory remains strongly coupled at the purported fixed point and we lose theoretical control. This is the essence of the problem noted in [15]. We explore this further in App. B, where we construct a scaling theory of the vector large N limit by rescaling momenta, frequency and redefining fields. This rederives an action analogous to Eq. (14) with a 't Hooft coupling of order N, showing that the theory flows to strong coupling even at leading order in the large N expansion.

Let us also compare our results with the planar limit of non-abelian gauge theories, and more generally with large N conformal field theories (CFTs). In this case, there is an infinite number of planar diagrams, whose resummation can often be described by a classical gravitational theory in one more dimension [33] (and see e.g. [34] for a review). In contrast, here we have found a finite number of planar diagrams that are ultimately responsible for the QCP. The main difference is that in relativistic theories it is necessary to resum the effect of relevant single-trace interactions of the matrix fields, such as $\mathrm{tr}(\phi^4)$. This gives rise to an infinite class of planar graphs that contribute. On the other hand, in the nonrelativistic setup of this work, the analog single-trace interactions are irrelevant. This leads to a finite class of diagrams whose effects can be taken into account exactly in the $1/N$ expansion.

In recent years, gauge/gravity duality has provided another framework for obtaining NFLs. See [35–38] for some of the original works, and [39–41] for reviews with additional references. These NFLs can be minimally described by coupling a strongly interacting large N CFT to a Fermi surface [42]. The CFT dresses the Fermi surface into a NFL with self-energy $\sim \omega^{2\Delta-1}$, where Δ is the dimension of the CFT operator that couples to the fermions. On the other hand, the backreaction of the fermions on the CFT is a negligible $1/N$ effect. Here we find some similarities with our framework, where the $N \times N$ order parameter ϕ gives rise to a NFL behavior $\sim \omega^{2/3}$. One important difference, however, is that the dynamics of the ϕ field itself is produced by its coupling to the Fermi surface, and does not need to be put in by hand. In any case, the flexibility of these semi-holographic Fermi liquids suggests generalizations of the theory studied in this work, where an overdamped $N \times N$ boson with dynamical exponent z_b is coupled to a Fermi surface with N fermion flavors. We hope to consider this in future work.

To conclude, we have identified a solvable matrix large N limit in which a two dimensional non-Fermi liquid arises at a quantum critical point. The theory has identical universal power laws to those conjectured in the vector large N theories. In the future, we wish to study the interplay between non-Fermi liquid behavior and superconductivity in such systems, as well as to study finite temperature thermodynamic and transport properties using the large N expansion. Lastly, we comment here that while the solvable large N limit provides insights into the nature of quantum materials, it remains unknown whether the precise power laws are the same in realistic systems with $N \sim 1$.

Appendix A: Renormalization group analysis

In this Appendix we present a renormalization-group treatment of the QCP discussed in the main text.

It is convenient to introduce the combination

$$\alpha \equiv \frac{1}{6\pi\sqrt{3}}\frac{g^2}{v}, \tag{A1}$$

in terms of which the fermion wavefunction renormalization reads

$$Z(p_0) = 1 + \frac{\Sigma(p_0)}{p_0} = 1 + \frac{3\alpha}{M_D^{2/3}|p_0|^{1/3}}. \tag{A2}$$

The RG approach focuses on running (or renormalized) couplings and fields. The renormalized fermion field is related to the original one by $\psi_r = Z^{1/2}\psi$, which makes

the kinetic term canonical. This redefinition gives rise to a running velocity and coupling

$$v_r(\mu) = Z^{-1}(\mu)v \ , \ g_r(\mu) = Z^{-1}(\mu)g \, . \qquad \text{(A3)}$$

Then $g_r(\mu)^2/v_r(\mu) \sim Z^{-1}(\mu)$. Since this combination scales as $\mu^{1/3}$ in the $z = 3$ regime, we will introduce the dimensionless running coupling $\alpha_r(\mu)$ defined as

$$\alpha = M_D^{2/3}\mu^{1/3} Z(\mu)\, \alpha_r(\mu) \, , \qquad \text{(A4)}$$

with the power of M_D chosen to match engineering dimensions. Using (A2), we find

$$\alpha_r(\mu) = \frac{1}{M_D^{2/3}\mu^{1/3}} \frac{\alpha}{1 + \frac{3\alpha}{M_D^{2/3}\mu^{1/3}}} \, . \qquad \text{(A5)}$$

In the IR, this flows to the fixed-point value

$$\alpha_r(0) = \frac{1}{3} \, . \qquad \text{(A6)}$$

These results can be equivalently obtained from the one-loop beta functions

$$\mu\frac{d\alpha_r}{d\mu} = -\frac{1}{3}\alpha_r + 2\gamma\alpha_r \, ,$$
$$2\gamma = -\mu\frac{d\log Z}{d\mu} \, . \qquad \text{(A7)}$$

The first term in the coupling beta function is simply the classical scaling dimension at short distance, while the second term is due to the fermion anomalous dimension. The fixed point value is (A6), with anomalous dimension $\gamma(0) = 1/6$. This gives rise to a quantum kinetic term $\sim \omega^{1-2\gamma} = \omega^{2/3}$.

1. Comparison with the vector large N limit

Let us compare these results with the vector large N expansion, where N fermion flavors interact with a singlet bosonic mode. We choose a Yukawa coupling $g/\sqrt{N}$, with g fixed at large N.

The first difference with the matrix limit is that Landau damping is enhanced by the fermion flavors running

in the loop, and this gives rise to a damping scale M_D independent of N. On the other hand, the fermion self-energy is now suppressed by $1/N$, obtaining

$$\Sigma(p_0) = \frac{g^2}{2\pi\sqrt{3}v\, N} \frac{1}{M_D^{2/3}} \text{sgn}(p_0)\, |p_0|^{2/3} \, . \qquad \text{(A8)}$$

This NFL contribution starts to dominate over the tree-level kinetic term for energies

$$E \lesssim \frac{(g^2/v)^3}{M_D^2} \frac{1}{N^3} \, . \qquad \text{(A9)}$$

Following the RG approach described above gives a beta function for the running coupling

$$\mu\frac{d\alpha_r}{d\mu} = -\frac{1}{3}\alpha_r + \frac{1}{N}\alpha_r^2 \, , \qquad \text{(A10)}$$

and we recall that $\alpha \sim g^2/v$, see (A1). We then find that the IR fixed point has a 't Hooft coupling $g^2 \sim N$. Since this grows with N, we do not expect a well-defined $1/N$ expansion. This is where the crucial difference with the matrix large N limit lies, since the latter has an order one 't Hooft coupling. In Appendix B, we explore a general redefinition along the lines of (12) for the vector large N limit, finding agreement with the present RG result.

The growth $g^2 \sim N$ implies a proliferation of higher loop diagrams, and a strongly coupled large N limit. And indeed, this is the main result of [15]. While this reference introduced a different genus expansion to take into account this effect, it is currently not known how to resum the leading contributions in the vector large N limit.

Appendix B: Large N field redefinitions

In this Appendix we analyze the field redefinitions required at large N, that were used in the main text. This approach will serve to exhibit the qualitatively different behavior of the matrix and vector large N limits.

For generality, we consider a model that can capture both the matrix and vector large N limits:

$$S = \int dp_0 dp_\perp\, (k_F d\hat{n})\, \psi_{\hat{n}}^\dagger \left(iN^x\text{sgn}(p_0)|p_0|^{2/3} - vp_\perp\right)\psi_{\hat{n}} + \int dq_0 dq_\perp dq_\parallel\, \phi\left(q^2 + N^y\frac{|q_0|}{q}\right)\phi$$

$$+ \ \frac{g}{\sqrt{N}}\int dq_0 dp_0 dq_\perp dq_\parallel dp_\perp\, k_F d\hat{n}\, \phi(q)\psi_{\hat{n}}^\dagger(p+q)\psi_{\hat{n}}(p,p) \, . \qquad \text{(B1)}$$

The matrix large N model is recovered for $(x = 1/3, y = -1)$, whereas the vector large N corresponds to $(x =$

$-1, y = 0)$. We have also set to one dimensionful combinations β, γ such as those in (10).

As in the main text, we seek a redefinition that eliminates the N-dependence of the propagators and which preserves the on-shell scalings

$$v p_\perp \sim N^x |p_0|^{2/3} \;,\; q^3 \sim N^y |q_0| \,. \qquad (B2)$$

For the boson, we do not need in general to redefine $q_\perp$ and $q_\parallel$ by the same factor, but the parallel component should dominate in the kinetic term so that the $z = 3$ scaling is maintained. Furthermore, since the bosonic momentum $q_\perp$ adds to the fermion momentum in the Yukawa coupling, $q_\perp$ and $p_\perp$ should be redefined by the same factor. On the other hand, we do not scale $\hat{n}$, which is the position on the Fermi surface.

Let us then look for a redefinition

$$(q_0, p_0) = N^a (\tilde{q}_0, \tilde{p}_0) \,,\; (q_\perp, p_\perp) = N^b (\tilde{q}_\perp, \tilde{p}_\perp) \,,\; q_\parallel = N^c \tilde{q}_\parallel$$
$$\phi = N^\rho \varphi \;,\; \psi = N^\eta \chi \,. \qquad (B3)$$

Preservation of the $z = 3$ scaling at large N requires

$$c \le b \;\Rightarrow\; \tilde{q}^2 \approx \tilde{q}_\parallel^2 \,. \qquad (B4)$$

Homogeneous N-scaling of both bosonic and fermionic kinetic terms implies

$$a = \frac{3}{2}(b - x) \;,\; c = \frac{1}{3}(a + y) \,, \qquad (B5)$$

for which condition (B4) reads

$$b \le -x + \frac{2}{3}y = \begin{cases} -1 \;,\; x = 1/3 \;,\; y = -1 \\ 1 \;,\;\; x = -1 \;,\; y = 0 \end{cases} \qquad (B6)$$

For the large N matrix model in the main text we adopted $b = -1$, which gives the largest energy window where the fixed point is valid.

Given this and imposing N-independence of both bosonic and fermionic kinetic terms determines the field redefinitions

$$\rho = -\frac{1}{2}(4b - 3x + y) \;,\; \eta = -\frac{1}{4}(7b - 3x) \,. \qquad (B7)$$

The resulting action has canonically-normalized two-point functions, and a cubic interaction proportional to $(g/\sqrt{N})N^{-\frac{x}{2} - \frac{y}{6}}$. Therefore,

$$\lambda = \frac{g^2}{N^{x + \frac{y}{3}}} \,, \qquad (B8)$$

plays the role of the 't Hooft coupling, which measures the strength of the coupling in the large N limit. Note the above expression is independent of our choice of b that determines how the external frequencies and momenta scale to zero.

In the matrix large N, with $(x = 1/3, y = -1)$, the 't Hooft coupling becomes $\lambda = g^2$, independent of N. So there is a well-defined $1/N$ expansion. In contrast, in the vector large N theory we have $(x = -1, y = 0)$, and the strength of the coupling is

$$\lambda = g^2 N \,. \qquad (B9)$$

So we do not expect a well-defined large N expansion, in agreement with [15]. In this way, we find a qualitative difference in the matrix and vector large N expansions.

[1] J. Hertz, "Quantum critical phenomena," *Phys. Rev. B* **14** (Aug, 1976) 1165–1184. http://link.aps.org/doi/10.1103/PhysRevB.14.1165.

[2] A. Schofield, "Non-fermi liquids," *Contemporary Physics* **40** no. 2, (1999) 95–115.

[3] C. Varma, Z. Nussinov, and W. van Saarloos, "Singular or non-fermi liquids," *Physics Reports* **361** no. 5, (2002) 267–417.

[4] T. Shibauchi, A. Carrington, and Y. Matsuda, "Quantum critical point lying beneath the superconducting dome in iron-pnictides," *arXiv preprint arXiv:1304.6387* (2013) .

[5] H. Kuo, J. Chu, J. Palmstrom, S. Kivelson, and I. Fisher, "Ubiquitous signatures of nematic quantum criticality in optimally doped fe-based superconductors," *Science* **352** no. 6288, (2016) 958–962.

[6] G. Stewart, "Heavy-fermion systems," *Reviews of Modern Physics* **56** no. 4, (1984) 755.

[7] H. Löhneysen, A. Rosch, M. Vojta, and P. Wölfle,

"Fermi-liquid instabilities at magnetic quantum phase transitions," *Rev. Mod. Phys.* **79** (Aug, 2007) 1015–1075. http://link.aps.org/doi/10.1103/RevModPhys.79.1015.

[8] Y. Schattner, S. Lederer, S. Kivelson, and E. Berg, "Ising nematic quantum critical point in a metal: A monte carlo study," *Phys. Rev. X* **6** (Aug, 2016) 031028. http://link.aps.org/doi/10.1103/PhysRevX.6.031028.

[9] E. Berg, S. Lederer, Y. Schattner, and S. Trebst, "Monte carlo studies of quantum critical metals," *Annual Review of Condensed Matter Physics* **10** no. 1, (2019) 63–84. https://doi.org/10.1146/annurev-conmatphys-031218-013339.

[10] X. Y. Xu, Z. H. Liu, G. Pan, Y. Qi, K. Sun, and Z. Y. Meng, "Revealing Fermionic Quantum Criticality from New Monte Carlo Techniques," *arXiv e-prints* (Apr, 2019) arXiv:1904.07355, arXiv:1904.07355 [cond-mat.str-el].

[11] J. Polchinski, "Low-energy dynamics of the

spinon-gauge system," *Nuclear Physics B* **422** no. 3, (1994) 617–633.

[12] Y. B. Kim, A. Furusaki, X.-G. Wen, and P. A. Lee, "Gauge-invariant response functions of fermions coupled to a gauge field," *Phys. Rev. B* **50** (Dec, 1994) 17917–17932. https://link.aps.org/doi/10.1103/PhysRevB.50.17917.

[13] B. L. Altshuler, L. B. Ioffe, and A. J. Millis, "Low-energy properties of fermions with singular interactions," *Phys. Rev. B* **50** (Nov, 1994) 14048–14064. http://link.aps.org/doi/10.1103/PhysRevB.50.14048.

[14] J. Rech, C. Pépin, and A. Chubukov, "Quantum critical behavior in itinerant electron systems: Eliashberg theory and instability of a ferromagnetic quantum critical point," *Phys. Rev. B* **74** (Nov, 2006) 195126. http://link.aps.org/doi/10.1103/PhysRevB.74.195126.

[15] S.-S. Lee, "Low-energy effective theory of fermi surface coupled with u(1) gauge field in $2+1$ dimensions," *Phys. Rev. B* **80** (Oct, 2009) 165102. http://link.aps.org/doi/10.1103/PhysRevB.80.165102.

[16] M. Metlitski and S. Sachdev, "Quantum phase transitions of metals in two spatial dimensions. i. ising-nematic order," *Phys. Rev. B* **82** (Aug, 2010) 075127. http://link.aps.org/doi/10.1103/PhysRevB.82.075127.

[17] S. Sur and S.-S. Lee, "Chiral non-fermi liquids," *Phys. Rev. B* **90** (Jul, 2014) 045121. http://link.aps.org/doi/10.1103/PhysRevB.90.045121.

[18] S.-S. Lee, "Recent developments in non-fermi liquid theory," *arXiv preprint arXiv:1703.08172* (2017) .

[19] C. Nayak and F. Wilczek, "Non-fermi liquid fixed point in 2+ 1 dimensions," *Nuclear Physics B* **417** no. 3, (1994) 359–373.

[20] C. Nayak and F. Wilczek, "Renormalization group approach to low temperature properties of a non-fermi liquid metal," *Nuclear Physics B* **430** no. 3, (1994) 534–562.

[21] T. Senthil and R. Shankar, "Fermi surfaces in general codimension and a new controlled nontrivial fixed point," *Phys. Rev. Lett.* **102** (Jan, 2009) 046406. http://link.aps.org/doi/10.1103/PhysRevLett.102.046406.

[22] D. Mross, J. McGreevy, H. Liu, and T. Senthil, "Controlled expansion for certain non-fermi-liquid metals," *Phys. Rev. B* **82** (Jul, 2010) 045121. http://link.aps.org/doi/10.1103/PhysRevB.82.045121.

[23] D. Dalidovich and S.-S. Lee, "Perturbative non-fermi liquids from dimensional regularization," *Phys. Rev. B* **88** (Dec, 2013) 245106. http://link.aps.org/doi/10.1103/PhysRevB.88.245106.

[24] G. 't Hooft, "A Planar Diagram Theory for Strong Interactions," *Nucl. Phys.* **B72** (1974) 461. [,337(1973)].

[25] R. Mahajan, D. M. Ramirez, S. Kachru, and S. Raghu, "Quantum critical metals in $d = 3 + 1$ dimensions," *Phys. Rev. B* **88** (Sep, 2013) 115116. http://link.aps.org/doi/10.1103/PhysRevB.88.115116.

[26] A. L. Fitzpatrick, S. Kachru, J. Kaplan, and S. Raghu, "Non-Fermi-liquid behavior of large-N_B quantum critical metals," *Phys. Rev. B* **89** no. 16, (Apr., 2014) 165114, arXiv:1312.3321 [cond-mat.str-el].

[27] G. Torroba and H. Wang, "Quantum critical metals in $4 - \epsilon$ dimensions," *Phys.Rev.* **B90** no. 16, (2014) 165144, arXiv:1406.3029 [cond-mat.str-el].

[28] S. Raghu, G. Torroba, and H. Wang, "Metallic quantum critical points with finite BCS couplings," *Phys. Rev.* **B92** no. 20, (2015) 205104, arXiv:1507.06652 [cond-mat.str-el].

[29] S. Chakravarty, R. Norton, and O. Syljuasen, "Transverse gauge interactions and the vanquished fermi liquid," *Phys. Rev. Lett.* **74** (February, 1995) 1423.

[30] V. Oganesyan, S. Kivelson, and E. Fradkin, "Quantum theory of a nematic fermi fluid," *Phys. Rev. B* **64** (Oct, 2001) 195109. http://link.aps.org/doi/10.1103/PhysRevB.64.195109.

[31] R. Shankar, "Renormalization group approach to interacting fermions," *Rev.Mod.Phys.* **66** (1994) 129–192.

[32] A. Fitzpatrick, G. Torroba, and H. Wang, "Aspects of Renormalization in Finite Density Field Theory," *Phys.Rev.* **B91** (2015) 195135, arXiv:1410.6811 [cond-mat.str-el].

[33] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," *Int. J. Theor. Phys.* **38** (1999) 1113–1133, arXiv:hep-th/9711200 [hep-th]. [Adv. Theor. Math. Phys.2,231(1998)].

[34] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, "Large N field theories, string theory and gravity," *Phys. Rept.* **323** (2000) 183–386, arXiv:hep-th/9905111 [hep-th].

[35] S.-S. Lee, "A Non-Fermi Liquid from a Charged Black Hole: A Critical Fermi Ball," *Phys. Rev.* **D79** (2009) 086006, arXiv:0809.3402 [hep-th].

[36] H. Liu, J. McGreevy, and D. Vegh, "Non-Fermi liquids from holography," *Phys. Rev.* **D83** (2011) 065029, arXiv:0903.2477 [hep-th].

[37] M. Cubrovic, J. Zaanen, and K. Schalm, "String Theory, Quantum Phase Transitions and the Emergent Fermi-Liquid," *Science* **325** (2009) 439–444, arXiv:0904.1993 [hep-th].

[38] T. Faulkner, H. Liu, J. McGreevy, and D. Vegh, "Emergent quantum criticality, Fermi surfaces, and AdS(2)," *Phys. Rev.* **D83** (2011) 125002, arXiv:0907.2694 [hep-th].

[39] S. A. Hartnoll, "Lectures on holographic methods for condensed matter physics," *Class. Quant. Grav.* **26** (2009) 224002, arXiv:0903.3246 [hep-th].

[40] C. P. Herzog, "Lectures on Holographic Superfluidity and Superconductivity," *J. Phys.* **A42** (2009) 343001, arXiv:0904.1975 [hep-th].

[41] J. McGreevy, "Holographic duality with a view toward many-body physics," *Adv. High Energy Phys.* **2010** (2010) 723105, arXiv:0909.0518 [hep-th].

[42] T. Faulkner and J. Polchinski, "Semi-Holographic Fermi Liquids," *JHEP* **06** (2011) 012, arXiv:1001.5049 [hep-th].