Upper bounds on the superfluid stiffness and superconducting $T_c$: Applications to twisted-bilayer graphene and ultra-cold Fermi gases

Tamagnha Hazra, Nishchhal Verma, and Mohit Randeria

Department of Physics, The Ohio State University, Columbus, Ohio 43210

Understanding the material parameters that control the superconducting (SC) transition temperature $T_c$ is a problem of fundamental importance. In many novel superconductors, phase fluctuations of the SC order parameter determine $T_c$, rather than the mean field collapse of the amplitude due to pair breaking. We derive rigorous upper bounds on the superfluid phase stiffness $D_s$ valid in any dimension. This in turn leads to an upper bound on $T_c$ in 2D, which holds irrespective of mechanism, strength of pairing interaction, or order-parameter symmetry. These bounds lead to stringent constraints for the strongly correlated regime of low-density and narrow-band systems. We show that $k_B T_c \leq E_F/8$ across the 2D BCS-BEC crossover in ultra-cold Fermi gases. For magic-angle twisted bilayer graphene (MA-TBG), the band structure constrains the maximum possible $T_c$ to be close to the experimentally observed value, demonstrating that MA-TBG is in a phase-fluctuation dominated regime. Finally, we discuss the question of deriving rigorous upper bounds on $T_c$ in 3D.

Our work is motivated by the fundamental question: what limits the superconducting (SC) transition temperature $T_c$? Within BCS mean-field theory, and its extensions like Eliashberg theory, the amplitude of the SC order parameter is destroyed by the breaking of pairs, and $T_c$ scales with the pairing gap $\Delta$. The material parameters that control the mean-field $T_c$ are the electronic density of states (DOS) at the chemical potential $N(0)$ and the effective interaction, determined by the spectrum of fluctuations that mediate pairing.

Beginning with the pioneering experiments of Uemura [1] and theoretical ideas of Emery and Kivelson [2] on underdoped cuprates, it became clear that the mean field picture of $T_c$ scaling with the pairing gap is simply not valid in many novel superconductors. The loss of SC order is then governed by fluctuations of the phase of the order parameter, rather than the suppression of its amplitude, and $T_c$ is related to the superfluid stiffness $D_s$. The material parameters that determine $D_s$ are rather different from those that determine the pairing gap $\Delta$.

The question of mean field amplitude collapse versus phase fluctuation dominated SC transition is brought into sharp focus by a variety of recent experiments in narrow band and low density systems. One of the most exciting recent developments is the observation of very narrow bands in magic-angle twisted bilayer graphene (MA-TBG) leading to correlation-induced ”Mott” insulating states [3] and superconductivity [4] in their vicinity. Flat bands are also also expected to arise in various topological states of matter [5–8]. BCS theory-based intuition suggests that narrow bands have a large DOS $N(0)$ and lead to high temperature superconductivity. Is this true or do phase fluctuations limit the $T_c$?

There has also been a burst of recent interest in superconductivity or paired superfluidity in low density systems, both in quantum materials (e.g., doped SrTiO$_3$, the LaAlO$_3$/SrTiO$_3$ interface) and quantum gases. Ultra-cold Fermi gases in the strongly interacting regime of the BCS-BEC crossover [9, 10] exhibit $k_B T_c/E_F$ values [11] larger than those observed in the solid state. Is there an intrinsic limit on the $k_B T_c/E_F$ ratio?

In this paper, we obtain sharp answers to these questions, especially in 2D. We show that the superfluid stiffness $D_s$ constrains $T_c$ in the strongly correlated regime of narrow band and low density systems. First, we derive an upper bound on the superfluid stiffness $D_s(T) \leq \tilde{D}(T)$, where $\tilde{D}$ is proportional to the optical sum rule spectral weight. This inequality is valid in all dimensions for multi-band systems with arbitrary interactions. Next, we show that $\tilde{D}$ is necessarily ”small” in low density and in narrow band systems. In many cases of interest, we show that $\tilde{D}$ is essentially determined by the non-interacting band structure. Finally, we use $\tilde{D}$ to obtain the upper bound $k_B T_c \leq \epsilon_D(0)/2$, a step that is rigorously controlled in 2D superconductors exhibiting a Berezinskii-Kosterlitz-Thouless (BKT) transition.

We illustrate the usefulness of our bounds for a variety of systems. For MA-TBG, we use available band structure results [12–16] to constrain $T_c$ without any knowledge of the pairing mechanism or order-parameter symmetry. We obtain a rigorous (but weak) bound of $\sim 20$ K. Using physically motivated approximations, we estimate a bound on $T_c$ as low as 6.7 K.

For the 2D BCS-BEC crossover in ultra cold atoms we show that the maximum $k_B T_c \leq E_F/8$, an exact result that poses a stringent constraint on $T_c$ in the strongly interacting regime. We also describe bounds on $T_c$ for the attractive Hubbard model, relevant for current optical lattice experiments [17], that demonstrate the tension between pair breaking and phase fluctuations, and highlight the connection with a pairing pseudogap [18, 19].

Finally, we discuss the question of deriving similar bounds in 3D. We show that the presence of non-universal pre-factors in the relation between $T_c$ and $D_s$, as well their scaling near a SC quantum critical point, pose difficulties in deriving a rigorous bound in 3D.

Results: We first outline our main results and then give a detailed derivation and specific applications. We consider a Fermi system described by the Hamiltonian

$$\mathcal{H} = \mathcal{H}_K + \mathcal{H}_{\text{int}}; \quad \mathcal{H}_K = \sum_{\mathbf{k},m,\sigma} \epsilon_m(\mathbf{k}) c^\dagger_{\mathbf{k}m\sigma} c^\phantom{\dagger}_{\mathbf{k}m\sigma}$$

(1)
where \( \mathbf{k} \) is crystal momentum, \( m \) is a band label, and \( \sigma \) the spin. \( \mathcal{H}_K \) is the kinetic energy and \( \mathcal{H}_{\text{int}} \) describes interactions (electron-phonon, electron-electron, etc.), including those that give rise to superconductivity. For now, we ignore disorder and return to it at the end.

The macroscopic superfluid stiffness \( D_s \) determines the free energy cost of distorting the phase of the SC order parameter as given by \( |\Delta\rangle e^{i\theta} \) via the Boltzmann factor \[ -D_s \int d^d r \sqrt{|\theta|^2/2k_B T}. \] It is related to the London penetration depth \( \lambda_L^2 = (4\mu_0 e^2/\hbar^2)D_s \) in 3D \((16\pi^2 e^2/h^2)D_s \) in Gaussian. Microscopically, \( D_s \) can be calculated as the static, long wavelength limit of the transverse current response to a vector potential \( \mathbf{A} \), which enters \( \mathcal{H} \) through a Peierls' substitution in the tight-binding representation of \( \mathcal{H}_K \). (Our results are equally valid for neutral superfluids with rotation playing the role of the magnetic field.) We obtain a rigorous upper bound valid in any dimension

\[
D_s(T) \leq \tilde{D}(T) = \frac{\hbar^2}{4\Omega} \sum_{\mathbf{k},m,m',\sigma} M_{m,m'}^{-1}(\mathbf{k}) |\langle \mathbf{c}_{\mathbf{k}m\sigma}^\dagger \mathbf{c}_{\mathbf{k}m'\sigma} \rangle|^2 \tag{2}
\]

where \( \Omega \) is the volume of the system and \( M_{m,m'}^{-1}(\mathbf{k}) \) is a mass tensor that depends only on the electronic structure of \( \mathcal{H}_K \); see eq. (5) below. The temperature and interactions impact \( \tilde{D} \) only through \( |\langle \mathbf{c}_{\mathbf{k}m\sigma}^\dagger \mathbf{c}_{\mathbf{k}m'\sigma} \rangle|^2 \).

We next use \( \tilde{D} \) to provide an upper bound on the SC transition temperature in 2D. Using the Nelson-Kosterlitz [20] universal relation, we obtain

\[
k_B T_{c,2D} \leq \pi \tilde{D}(0)/2 \tag{3}
\]

For weak coupling, \( T_c \) is well described by mean field theory and our result, though valid as an upper bound, may not be very useful. However, as we show below, for strongly interacting systems our bound gives insight into both the value of \( T_c \) and on parameter dependence.

**Bound on superfluid stiffness:** The intuitive idea behind \( D_s \leq \tilde{D} \) is as follows. \( (2\pi^2 e^2/\hbar^2)D_s \) is the optical conductivity spectral weight integrated over the bands in (1), and \( (4\pi^2 e^2/\hbar^2)D_s \) is the coefficient of the \( \delta(\omega) \) piece in \( \text{Re} \sigma(\omega) \) in the SC state; (note: \( \int_0^\infty d\omega \delta(\omega) = 1/2 \)). The inequality (2) says that the weight in the SC delta-function must be less than or equal to the total spectral weight.

To derive (2), we use the Kubo formula for \( D_s \) as a linear response [21] to an external vector potential

\[
D_s \sim \tilde{D} - \left( \frac{\hbar^2}{4e^2} \right) \chi_{Ja,Ja}^{\perp}(q \to 0, \omega = 0). \tag{4}
\]

Here \( \tilde{D} \sim \delta^2 \mathcal{H} / \delta \mathcal{A}_{\mathbf{k}z}^2 \) is the diamagnetic response and \( \chi^{\perp} \) is the transverse \((\perp)\) paramagnetic current correlator. Note that \( D_{s,a} \) and \( D_{\perp} \) depend on \( a = x, y, z \) in a spatially anisotropic system, but we omit the subscript \( a \) unless essential. From its Lehmann representation we see that \( \chi_{Ja,Ja}^{\perp}(\mathbf{q},0) \geq 0 \) at all \( T \); see Appendix B. We thus obtain \( D_{s,a} \leq D_{\perp} \).

As shown in Appendix A, \( D_s \) is given by eq. (2) with

\[
M_{m,m'}^{-1}(\mathbf{k}) = \sum_{\alpha,\beta} U_{m,\alpha}(\mathbf{k}) \frac{\partial^2 t_{\alpha\beta}(\mathbf{k})}{\partial (\mathbf{k} h)^2} U_{\beta,m'}(\mathbf{k}). \tag{5}
\]

Here \( \alpha, \beta \) label orbitals/sites within a unit cell of a Bravais lattice, \( t_{\alpha\beta}(\mathbf{r}) \) is the Fourier transform of the hopping matrix element \( t_{\alpha\beta}(\mathbf{r}_i - \mathbf{r}_j) \), and \( U_{m,\alpha}(\mathbf{k}) \) is the unitary transformation that diagonalizes \( t_{\alpha\beta}(\mathbf{k}) \) to \( \epsilon_m(\mathbf{k}) \delta_{m,m'} \).

For a single band, eqs. (2) and (5) simplify to \( \tilde{D} = (4\pi^2 e^2/\hbar^2) \sum_{\mathbf{k},\sigma} (\partial^2 \epsilon(\mathbf{k}) / \partial k^2) n_{\sigma}(\mathbf{k}) \) where the momentum distribution \( n_{\sigma}(\mathbf{k}) = \langle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma} \rangle \). For parabolic dispersion \( \epsilon(\mathbf{k}) = \hbar^2 k^2/2m \) this yields \( \tilde{D} = \hbar^2 n/4m \) independent of \( T \) and of interactions. Here \( D_s(T) = \hbar^2 n_s(T) / 4m \), and our bound simply says that the superfluid density \( n_s(T) \leq n \) the total density, with \( n_s(0) = n \) guaranteed by continuous translation invariance.

For materials with non-parabolic dispersion and/or multiple bands, \( D \) depends both on \( T \) and on interactions, since the thermal average in (2) is calculated using the full \( \mathcal{H} \). It is illuminating to derive an upper bound for \( D \) which is independent of both \( T \) and of interactions. We describe the single band result here; the multi-band generalization is in Appendix D. We write \( \mathcal{H}_K = -\sum_{\mathbf{R}\delta\sigma} [t(\delta) \mathbf{c}_{\mathbf{R}\delta\sigma}^\dagger \mathbf{c}_{\mathbf{R}\delta\sigma} + \text{h.c.}] \) with translationally invariant hopping amplitudes \( t(\delta) \) that depend only the vector \( \delta \) connecting lattice \( \mathbf{R} \) and \( \mathbf{R}' \) on an arbitrary lattice. We couple the system to a vector potential and compute \( \tilde{D} \), which involves terms like \( \sum_{i,j} \delta_{ij}^2 t(\delta)(\mathbf{c}_{i}^\dagger \mathbf{c}_j) \) with \( \delta = i - j \) (schematically). We note that \( \tilde{D} \geq 0 \), since it is the sum rule for a dissipative response function \( \text{Res}(\omega) \geq 0 \). We then use the triangle inequality and Cauchy-Schwarz \( |\langle \mathbf{c}_i^\dagger \mathbf{c}_j \rangle| \leq \sqrt{\langle n_i \rangle \langle n_j \rangle} = n \) to obtain \( D_s \leq \tilde{D} = n \sum_i \delta_{ij}^2 |t(\delta)|/2 \). This shows that for small hopping and/or low density, one necessarily has a small \( D_s \).

**\( T_c \) bound in 2D:** For a BKT transition in 2D, we have a universal ratio [20] of \( k_B T_c / D_s(T_c^\perp) = \pi/2 \). We can combine this with our \( \tilde{D} \) bound in two different ways. Either we can use \( D_s(T_c^\perp) \leq \tilde{D}(T_c) \). Alternatively, we can use the fact that \( D_0(T) \) is a decreasing function of temperature (which we have not attempted to prove rigorously, but find no reason to doubt) to get \( D_s(T_c^\perp) \leq D_s(0) \leq \tilde{D}(0) \). This immediately leads to the bound (3), whose consequences we now explore for various systems of interest. In anisotropic systems \( D = \max \{D_s, D_{\perp}\} \) suffices, but we can argue for the much stronger result \( D = \tilde{D}^2 \) in 2D; see Appendix H.

**2D BCS-BEC crossover:** Let us first consider the simplest case of a single band with parabolic dispersion \( \epsilon(\mathbf{k}) = \hbar^2 \mathbf{k}^2/2m \), which leads to a new exact result for the 2D BCS-BEC crossover in ultra-cold Fermi gases. The dimensionless interaction in a dilute gas in 2D is \( \log(E_p/E_F) \), where \( E_F = \pi\hbar^2 n/m \) is the Fermi energy and \( E_b \) the binding energy of the two-body bound state in vacuum [22]. Using \( \tilde{D} = \hbar^2 n/4m \), we predict \( k_B T_c \leq E_F/8 \) for all \( E_p/E_F \). Note that there are very few exact results in the strongly interacting regime \( E_b \sim E_F \).
tain magic angles $\theta$ with a bandwidth less than 10 meV by choosing cer-
in the weak-coupling BCS limit $\omega \ll 1$, the mean
field $k_B T_c$ is $\sim \sqrt{E_F E_b}$ [22] with a prefactor, includ-
ing Gorkov-Melik-Barkhudarov corrections [23], that is
known. This $T_c$ has an essential singularity in the attraction and is much smaller than our bound. In the BEC limit our bound is also larger than the result [24] $k_B T_c = E_F/[21 \log \log(2n_{\xi}^2)]$, valid in the asymptotic regime $\log \log \gg 1$. Here the bosons have mass $2m$, density $n/2$ and an inter-boson scattering length $a_b$, which has been computed [23] in terms of $E_b$. Our bound
cautions against a naive extrapolation of the BEC limit result into the strong interaction regime. The existing experiment [25] seems to be inconsistent with our bound. We note, however, that the determination of the BKT $\theta$ is a delicate matter in a cold atom experiment on a quasi-
2D system in a harmonic trap, while our bound is for a strictly 2D system in the thermodynamic limit.

**Magic angle twisted bilayer graphene:** The existence of very narrow bands in MA-TBG was predicted by continuum electronic structure calculations [12, 13] that pointed out the crucial role of $\alpha = w/\hbar v_F^0 K \theta$, where $\theta$ is the twist angle between the two layers, $w$ is the interlayer tunneling, and $v_F^0$ the bare Fermi velocity and $K$ the Dirac-node location in monolayer graphene. It was predicted that $v_F$ in TBG can be tuned to zero [12], with a bandwidth less than 10 meV by choosing certain magic angles $\theta$, the largest of which $\approx 1.1^\circ$ has now been achieved in experiments [3, 4]. Recently, pressure-
tuning of $w$ has also resulted in very narrow bands [26].

Little is known at this time about the nature of the SC state or the pairing mechanism. Proximity to a “Mott” insulator and narrow bandwidth suggest the importance of electron correlations, while the extreme sensitivity of the dispersion to structure suggests that electron-
phonon interactions could also be important. We ar-
gue here that simply using the best available electronic structure information [14–16] for MA-TBG, and without any prejudice about the interactions responsible for SC, we can put strong constraints on its superconducting $T_c$.

There are two bands for each of the two valleys, one above and the other below the charge neutrality point (CNP). Each band has a two-fold spin degeneracy, with bands for one valley related to those of the other by time-
reversal. We include these bands in the $\sum_{m,m'}$ in eq. (2), while the $\Sigma_k$ is over the moiré Brillouin zone, a hexagon with side $2K \sin(\theta/2) \approx K \theta$. We use the tight-binding model of ref. [14], a multi-parameter fit to the continuum model dispersion [12], to calculate $M^{-1}(\mathbf{k})$ of (5).

To bound $T_c$ for the SC state near half-filling on the hole-doped side of the CNP, we take $n(\mathbf{k}) = 0$ for the empty bands above the CNP. This implies that only $m = m'$ terms survive in (2), since $\langle c_{m\sigma}^\dagger c_{m\sigma}' \rangle = 0$ for $m \neq m'$, when either of the bands is empty or is completely filled. The empty band result follows from the Cauchy-Schwartz inequality $|\langle c_{m\sigma}^\dagger c_{m'\sigma}' \rangle|^2 \leq n_{m\sigma}(\mathbf{k}) n_{m'\sigma}(\mathbf{k})$. A similar argument works for the filled case after a particle-hole transformation; see Appendix E.

Since $\tilde{D} \geq 0$, we use the triangle inequality to obtain $\tilde{D} \leq (\hbar^2/4\Omega) \sum_{\mathbf{k},m,\sigma} |M^{-1}_{m,m'}(\mathbf{k})| n_{m\sigma}(\mathbf{k})$. Using $n(\mathbf{k}) \leq 1$ for the bands below CNP, we obtain the bound $T_c \leq 19.5K$ near half-filling for hole doping using the tight-binding model of ref. [14]. A similar calculation leads to $T_c \leq 15.7K$ near half-filling for electron doping. We note that using $|M^{-1}|$ and general constraints on $n(\mathbf{k})$ leads to rigorous bounds, but weakens the result.

Next we make a physically motivated estimate of $\tilde{D}$, which allows us to improve the result, but making it an “approximate bound”. We use the band theory result $\langle c_{m\sigma}^\dagger c_{m'\sigma}' \rangle = \delta_{m,m'} \Theta(\mu - m(\mathbf{k}))$ at $T = 0$, with the chemical potential $\mu$ determined by the density $\Omega^{-1} \sum_{\mathbf{k},m,\sigma} n_{m\sigma}(\mathbf{k})$. This, together with $M^{-1}_{m,m'}(\mathbf{k})$ calculated from the tight binding model of ref. [14], leads to the density-dependent estimate of $\tilde{D}$ plotted in Fig. 1.

The integrated optical spectral weight is given by $(2\pi^2 \hbar^2) \tilde{D}$ as noted above. Note that it vanishes at the band insulators when all bands are either filled or empty. Clearly our band-structure based estimate does not know about the “Mott” insulating states at half-
filling away from CNP, where in reality the (low-energy) optical spectral weight should go to zero.

$(\pi/2)$ times the $\tilde{D}$ plotted in Fig. 1 is an estimated upper bound on the SC $T_c$. The system is not SC over most of the doping range, but our bound is the maximum attainable $T_c$ if the system were to exhibit superconduc-
tivity. We find the maximum $T_c$ to be about 6.7K, close to the 3K found in experiment [26]. The fact that the measured $T_c$ is close to our stiffness bound implies that MA-TBG is a strongly correlated SC in a phase fluctuation dominated regime, which suggests that it should exhibit spectral anomalies such as the pseudogap.

**2D attractive Hubbard model and optical lattices:** We next obtain important insights on the value of $T_c$ and its interaction-dependence for the 2D attrac-
tive Hubbard model, where we can compare our bound with Quantum Monte Carlo (QMC) simulations [27] free of the fermion sign problem. This system has also been investigated in recent optical lattice experiments [17].

Consider nearest-neighbor (NN) hopping on a square lattice with \( \mathcal{H} = -t \sum_{\langle i,j \rangle} c_{i \sigma}^\dagger c_{j \sigma} + \text{h.c.} - |U| \sum_i (n_i - 1/2)(n_i - 1/2) \). For \( n \neq 1 \) this model has a SC ground state, exhibiting a crossover from a weak coupling BCS state \((|U|/t \ll 1)\) to a BEC of hard-core on-site bosons \((|U|/t \gg 1)\). The QMC estimate [27] of \( T_c \), obtained from the BKT jump in the \( D_s \), is a non-monotonic function of \(|U|/t\) at a fixed density \( n \); see Fig. 2. The BCS mean field \( T_c^{\text{MFT}} \) correctly describes the weak coupling \( T_c \), but for \(|U|/t \gtrsim 2\) lies well above the actual \( T_c \) and is only a pair breaking scale. In the limit \(|U|/t \gg 1\) we see \( T_c \sim t^2/|U| \), the effective boson hopping.

Our bound permits us to understand \( T_c(|U|/t) \) in the intermediate coupling regime where there are no other reliable analytical estimates. For NN hopping on a square lattice, \( \delta^2 \epsilon(k)/\partial k_z^2 \sim \epsilon(k) \), and \( \delta \) is proportional to the kinetic energy. To estimate \( \delta \), we need to make an approximation for \( n(k) \). If we choose a step-function (as we did for the MA-TBG) we get an estimated bound of \( \approx 0.3t \) for \( n = 0.7 \), independent of \(|U|/t| \).

To obtain a better estimate, we note that, as \(|U|/t| \) increases, the pair-size shrinks and \( n(k) \) broadens. In the extreme \(|U|/t| \)-limit of on-site bosons, \( n(k) \) is flat (\( k \)-independent), leading to \( \delta \to 0^+ \), since \( \delta^2 \epsilon(k)/\partial k_z^2 \) is a zero mean, periodic function whose \( k \)-sum vanishes. To model this broadening of \( n(k) \), we use the results of the \( T \) = 0 BCS-Leggett crossover theory (Appendix D). This gives us the (approximate) bound plotted in Fig. 2, which has the correct \( t^2/|U| \) asymptotic behavior at large \(|U|\).

In general we find \( T_c \leq \min \{ T_c^{\text{MFT}}, \pi \delta B(0)/2kB \} \). For temperatures between the pairing scale \( T_c^{\text{MFT}} \) and the \( T_c \) at which phase coherence sets in, the “normal state” exhibits a pseudogap due to pre-formed pairs [18, 19].

**Three dimensional systems:** Experiments suggest that there may also be a 3D version of the relationship between \( T_c \) and the superfluid stiffness; see, e.g., the extensive compilation of data in a Uemura-like plot in Fig. 6 of ref. [4]. We should note, however, that this figure is a log-log plot of \( T_c \) versus “Fermi energy \( E_F \),” and there is considerable leeway in defining a single “\( E_F \)” in multi-band materials.

We have not succeeded in deriving a rigorous bound on the 3D \( T_c \), unlike in 2D. There are two problems that one faces in trying to derive a bound in 3D: one is a question of the magnitude of \( T_c \) and the other a question about the functional form of its dependence on \( D_s \).

Following Emery and Kivelson (EK) [2], we focus on the 3D phase ordering temperature \( k_BT \) = \( AD_s(0)\bar{\sigma} \), which could provide a bound on \( T_c \). Here \( A \) is a (dimensionless) constant and \( \bar{\sigma} \) is the length-scale up to which one has to coarse-grain to derive an effective XY model. EK use \( \pi\xi^2 = \pi\zeta^2 \), where \( \zeta \) is the coherence length, and suggest, based on Monte Carlo results for classical XY models, that \( A \approx 4.4 \) gave a reasonable account of experiments on underdoped cuprates and other materials.

However, the coefficient \( A \) is non-universal and can vary from one system to another. Consider the 3D problem of the BCS-BEC crossover in ultra-cold Fermi gases [10] with \( h^2k^2/2m \) dispersion and interaction, characterized by the s-wave scattering length \( a_s \), tuned using a Feshbach resonance. It is known from experiments [11] at unitarity \((a_s \to \infty)\) that \( k_BT_c = 0.17E_F \). QMC estimates [28, 29] range from \( k_BT_c \approx 0.15E_F \to 0.17E_F \) at unitarity, and also show the expected non-monotonic behavior of \( k_BT_c/E_F \) as a function of \( k_BT \), with a maximum \( k_BT_c/E_F \approx 0.22 \) at a small positive \( 1/k_BT \). We choose \( \zeta = k_F^{-1} \) near unitarity [30] and try to use \( k_BT = A(h^2n/4m\sqrt{\zeta}) \) as a bound on \( T_c \). Consistency with the observed \( k_BT_c/E_F \approx 0.22 \) then requires \( A \approx 7.4 \), quite different from the 4.4 quoted above. We do not know if there is a definite value of \( A \) that would give a “phase-ordering” upper bound on \( T_c \) in 3D.

The following argument suggests that there may, in fact, be no general bound on \( T_c \) that is linear in \( D_s(0) \) in 3D. From a practical point of view, one is interested in learning about the highest \( T_c \) in a class of materials. But, if a general bound were to exist, it should be equally valid in situations where both \( T_c \) and \( D_s(0) \) are driven to zero by tuning a (dimensionless) parameter \( \delta \to 0^+ \) toward a quantum critical point (QCP). From the action \( S = -\frac{1}{2}\int d\tau d^d\eta \left\{ \Delta^2 \frac{\partial \phi}{\partial \tau}^2 - \frac{\partial \phi}{\partial \tau} \right\} + \delta \sum_{\langle i,j \rangle} \phi_i \phi_j \) describing SC phase fluctuations, we get the quantum Josephson scaling relation [31] \( D_s(0) \sim \delta^{(2+2d-2)} \). One also obtains, as usual, \( T_c \sim \delta^{z\nu} \), where \( z \) and \( \nu \) are the dynamical and correlation length exponents in \( d \) spatial dimensions.

Thus we find \( T_c \sim |D_s(0)|^{(2+2d-2)/2} \) near the QCP. In 2D, this implies \( T_c \sim D_s(0) \) consistent with the bound we derived above. However, in 3D we get \( T_c \sim D_s(0)^{2/3} \) which will necessarily violate an upper bound on \( T_c \).
that scales linearly with $D_d(0)$ sufficiently close to the QCP. This is not just an academic issue, as experiments see precisely such a deviation from linear scaling with $T_c \sim \sqrt{D_d(0)}$, consistent with $z = 1$, both in highly underdoped [32, 33] and in highly overdoped [34, 35] cuprates. It would be very interesting to see if there is a suitable formulation of the result that precludes the QCP regime and allows us to focus on the “large” stiffness case to derive a phase dominated bound for “large” $T_c$’s in 3D.

**Concluding remarks:** We have thus far ignored disorder, which always reduces $D_d$ [36]. Thus our upper bounds continue to be valid, although they can they can be improved. Although we focused on narrow band and low density systems, our bounds have also important implications for systems close to insulating states, either correlation-driven or disorder-driven. In either case, if there is a continuous superconductor to insulator transition, the superfluid stiffness will eventually become smaller than the energy gap and control the SC $T_c$.

As a design principle, it is interesting to ask if one can have multi-band systems where a narrow band has a large energy gap and large “mean field” $T_c$ interacting with a broad band that makes a large contribution to the superfluid stiffness, thus getting the best of both worlds. This is left for future investigations.

**Acknowledgements:** We are grateful to P. Törnä, S. Peotta, and L. Liang for pointing out an error in an earlier version of our paper that led us to the correct multi-band result presented here. We thank S. Kivelson, T. Senthil and C. Jayaprakash for discussions, and acknowledge support from NSF DMR-1410364.

[1] Y. J. Uemura, G. M. Luke, B. J. Sternlieb, J. H. Brewer, J. F. Carolan, W. N. Hardy, R. Kadono, J. R. Kempton, R. F. Kiefl, S. R. Kreitzman, P. Mulhern, T. M. Riseman, D. L. Williams, B. X. Yang, S. Uchida, H. Takagi, J. Gopalakrishnan, A. W. Sleight, M. A. Subramanian, C. L. Chien, M. Z. Cieplak, G. Xiao, V. Y. Lee, B. W. Statt, C. E. Stronach, W. J. Kossler, and X. H. Yu, Physical Review Letters 62, 2317 (1989).
[2] V. J. Emery and S. A. Kivelson, Nature 374, 434 (1995).
[3] Y. Cao, V. Fatemi, A. Demir, S. Fang, S. L. Tomarken, J. Y. Luo, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, E. Kaxiras, R. C. Ashoori, and P. Jarillo-Herrero, Nature 556, 80 (2018).
[4] Y. Cao, V. Fatemi, S. Fang, K. Watanabe, T. Taniguchi, E. Kaxiras, and P. Jarillo-Herrero, Nature 556, 43 (2018).
[5] N. B. Kopnin, T. T. Heikkilä, and G. E. Volovik, Phys. Rev. B 83, 220503 (2011).
[6] E. Tang and L. Fu, Nature Physics 10, 964 (2014).
[7] S. Peotta and P. Törnä, Nature Communications 6, 8944 (2015).
[8] L. Liang, T. I. Vanhala, S. Peotta, T. Siro, A. Harju, and P. Törnä, Phys. Rev. B 95, 024515 (2017).
[9] W. Ketterle and M. W. Zwierlein, in Ultracold Fermi Gases, Proc. of the Int. Sch. of Phys. “Enrico Fermi”, Course CLXIV, Varenna, 2006, edited by M. Inguscio, W. Ketterle, and C. Salomon (IOS Press, Amsterdam, 2008).
[10] M. Randeria and E. Taylor, Annual Review of Condensed Matter Physics 5, 209 (2014).
[11] M. J. H. Ku, A. T. Sommer, L. W. Cheuk, and M. W. Zwierlein, Science 335, 563 (2012).
[12] R. Bistritzer and A. H. MacDonald, Proceedings of the National Academy of Sciences 108, 12233 (2011).
[13] J. M. B. Lopes dos Santos, N. M. R. Peres, and A. H. Castro Neto, Physical Review Letters 99, 256802 (2007).
[14] M. Koshino, N. F. Q. Yuan, T. Koretsune, M. Ochi, K. Kuroki, and L. Fu, Physical Review X 8, 031087 (2018).
[15] J. Kang and O. Vafek, Physical Review X 8, 031088 (2018).
[16] H. C. Po, L. Zou, A. Vishwanath, and T. Senthil, Physical Review X 8, 031089 (2018).
[17] D. Mitra, P. Brown, E. Guardado-Sanchez, S. S. Kondov, T. Devakul, D. A. Huse, P. Schauss, and W. Bakr, Nature Physics 14, 173 (2018).
[18] M. Randeria, N. Trivedi, A. Moreo, and R. T. Scalettar, Physical Review Letters 69, 2001 (1992).
[19] N. Trivedi and M. Randeria, Physical Review Letters 75, 312 (1995).
[20] D. R. Nelson and J. M. Kosterlitz, Physical Review Letters 39, 1201 (1977).
[21] D. J. Scalapino, S. R. White, and S. Zhang, Physical Review B 47, 7995 (1993).
[22] M. Randeria, J.-M. Duan, and L.-Y. Shieh, Physical Review Letters 62, 981 (1989).
[23] D. S. Petrov, M. A. Baranov, and G. V. Shlyapnikov, Physical Review A 67, 031601 (2003).
[24] D. S. Fisher and P. C. Hohenberg, Physical Review B 37, 4936 (1988).
[25] M. G. Ries, A. N. Wenz, G. Zürn, L. Bayha, I. Boettcher, D. Kedar, P. A. Murthy, M. Neidig, T. Lonpe, and S. Jochim, Physical Review Letters 114, 230401 (2015).
[26] M. Yankowitz, S. Chen, H. Polshyn, K. Watanabe, T. Taniguchi, D. Graf, A. F. Young, and C. R. Dean, arXiv:1808.07865 [cond-mat] (2018), arXiv:1808.07865 [cond-mat].
[27] T. Paiva, R. Scalettar, M. Randeria, and N. Trivedi, Physical Review Letters 104, 066406 (2010).
[28] E. Burovski, E. Kozik, N. Proko’ev, B. Svistunov, and M. Troyer, Physical Review Letters 101, 090402 (2008).
[29] O. Goulko and M. Wingate, Physical Review A 82, 053621 (2010).
[30] J. R. Engelbrecht, M. Randeria, and C. A. R. Sáde Melo, Physical Review B 55, 15153 (1997).
[31] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Physical Review B 40, 546 (1989).
[32] I. Hetel, T. R. Lemberger, and M. Randeria, Nature Physics 3, 700 (2007).
[33] D. M. Broun, W. A. Huttema, P. J. Turner, S. Özcan, B. Morgan, R. Liang, W. N. Hardy, and D. A. Bonn, Physical Review Letters 99, 237003 (2007).
[34] T. R. Lemberger, I. Hetel, A. Tsukada, M. Naito, and M. Randeria, Physical Review B 83, 140507 (2011).
[35] I. Božović, X. He, J. Wu, and A. T. Bollinger, Nature 536, 309 (2016).
[36] A. Paramekanti, N. Trivedi, and M. Randeria, Physical Review B 57, 11639 (1998).
**Appendix A: Linear response, $D_s$ and $\bar{D}$**

Let us consider the general Hamiltonian

$$\mathcal{H} = \mathcal{H}_K + \mathcal{H}_{\text{int}}$$  \hspace{1cm} (A1)

where $\mathcal{H}_{\text{int}}$ represents arbitrary interactions, including those that give rise to superconductivity, and $\mathcal{H}_K$ is the most general single particle Hamiltonian for a multi-band/multi-orbital lattice model

$$\mathcal{H}_K = \sum_{i\beta j\alpha} t_{i\beta j\alpha} (\mathbf{r}_{i\alpha} - \mathbf{r}_{j\beta}) c_{i\alpha}^\dagger c_{j\beta}.$$  \hspace{1cm} (A2)

Here $t_{i\beta j\alpha} (\mathbf{r}_{i\alpha} - \mathbf{r}_{j\beta})$ represents the hopping matrix element from orbital $\beta$ in unit cell $j$ to orbital $\alpha$ in unit cell $i$ with $i,j$ spanning all unit cells, including $i = j$. We omit the spin label $\sigma$ only to simplify notation but we are not ignoring spin, as emphasized by the spin sum. In the presence of an external vector potential $\mathbf{A}$, the hopping picks up the Peierls phase

$$\mathcal{H}_K \rightarrow \mathcal{H}_K = \sum_{\mathbf{R},i\beta j\alpha} t_{i\beta j\alpha} (\mathbf{r}_{i\alpha} + \mathbf{R}) e^{-i e \mathbf{A}(\mathbf{R})/\hbar} c_{i\alpha}^\dagger c_{j\beta}$$  \hspace{1cm} (A3)

where we use the notation $\mathbf{R} = (\mathbf{r}_{i\alpha} + \mathbf{r}_{j\beta})/2$ and $\mathbf{r} = \mathbf{r}_{i\alpha} - \mathbf{r}_{j\beta}$ for simplicity. Since we are eventually interested in the long wavelength limit $\mathbf{q} \rightarrow 0$, we choose a very slowly varying vector potential and write $\mathbf{r}_{i\alpha} \cdot \mathbf{A} \cdot d\mathbf{l} = \mathbf{A}(\mathbf{R}) \cdot \mathbf{r}$.

Within linear response theory we can Taylor expand the exponential retaining terms which are linear (paramagnetic) and quadratic (diamagnetic) in $\mathbf{A}$. We transform to Fourier space using $t_{i\beta j\alpha} (\mathbf{k}) = \sum_{\mathbf{r}} t_{i\beta j\alpha} (\mathbf{r}) e^{-i \mathbf{k} \cdot \mathbf{r}}$ and $c_{i\alpha} = \Omega^{-1/2} \sum_{\mathbf{k}} e^{i \mathbf{k} \cdot \mathbf{r}_{i\alpha}} d_{\mathbf{k} \alpha}$. We can then write the current operator $j_x = \delta \mathcal{H}_K/\delta A_x$ as the sum of the paramagnetic ($P$) and diamagnetic ($D$) current operators given by

$$j_x^P (\mathbf{q}) = \frac{e}{\hbar} \sum_{a\beta, k \alpha} \frac{\partial t_{a\beta j\alpha}}{\partial k_x} d_{k \alpha}^\dagger q_{2, \alpha} d_{k \beta} - q_{2, \beta}$$  \hspace{1cm} (A4)

$$j_x^D (\mathbf{q}) = \frac{e^2}{\hbar^2} \sum_{a\beta, k \alpha} \frac{\partial^2 t_{a\beta j\alpha}}{\partial k_x^2} d_{k \alpha}^\dagger d_{k \beta} A_x (\mathbf{q}),$$  \hspace{1cm} (A5)

where we only show the $x$-component for simplicity. Note that the paramagnetic current operator, when transformed to the band basis, will in general have interband matrix elements [7, 8]. The only property of $j_x^P (\mathbf{q})$ that we will need to use below, however, is that it is a Hermitian operator; see eq. (B1).

The superfluid stiffness $D_s$ is defined as the static long-wavelength limit of the transverse response of the current density $j$ to a vector potential $\mathbf{A}$

$$\langle j_x (\mathbf{q}, \omega) \rangle = -\frac{4 e^2}{\hbar^2} D_s A_x (\mathbf{q}, \omega), \quad q_x = 0, q_\perp \rightarrow 0, \omega = 0$$  \hspace{1cm} (A6)

here $\perp$ represents the orthogonal directions to $x$. Standard linear response theory leads to the Kubo formula

$$D_s = \bar{D} - \frac{\hbar^2}{4 e^2} \chi_{j_x j_x} (\mathbf{q} \rightarrow 0, \omega = 0)$$  \hspace{1cm} (A7)

where the first term is the diamagnetic term, which is of central interest in this work, and the second is the transverse paramagnetic current-current correlation function. We will focus on the latter in Appendix B, where we show that $\chi_{j_x j_x} \geq 0$ at all temperatures.

Here we focus on the first term that can be read off from the form of the diamagnetic current operator. We find it convenient to write it in the band basis as

$$\bar{D} = \frac{\hbar^2}{4 \Omega} \sum_{mm', kk} M_{mm'}^{-1} (\mathbf{k}) \langle c_{km}^\dagger c_{km'} \rangle$$  \hspace{1cm} (A8)

with the inverse mass tensor given by

$$M_{mm'}^{-1} (\mathbf{k}) = \sum_{a\beta} U^*_{m,ak} (\mathbf{k}) \frac{\partial^2 t_{a\beta j\alpha}}{\partial (hk_x)^2} U_{\beta,ak'} (\mathbf{k}),$$  \hspace{1cm} (A9)
The unitary transformation $U$ that transforms from the orbital to the band basis is defined by
\[\sum_{\alpha\beta} U^*_{m,\alpha}(\mathbf{k}) t_{\alpha\beta}(\mathbf{k}) U_{\alpha,m}(\mathbf{k}) = \epsilon_m(\mathbf{k}) \delta_{m,m'} .\] (A10)

This allows us to write the final result in the band basis using
\[d_{k\alpha} = \sum_m U_{\alpha,m}(k)c_{km}.\] (A11)

We note several important points about the inverse mass tensor $M^{-1}_{\alpha\beta}(\mathbf{k})$. (i) It depends only on the bare band structure, and is independent of temperature and interactions, (ii) it has both diagonal and off-diagonal terms in the band indices. and (iii) it is not simply related to the curvature of the bands $\partial^2 \epsilon_m(\mathbf{k})/\partial k_x^2$, in contrast to the single-band case in eq. (A12).

The standard reference on the formalism for calculating the superfluid stiffness in lattice systems is Scalapino, White and Zhang (SWZ) [21]. Our normalization conventions differ from them and, more importantly, they focus on the special case of a single band model with nearest-neighbor (NN) hopping on a square (or cubic) lattice. Thus it may be useful for us to provide a “dictionary” relating our results to theirs.

In the single-band case our expression for $\bar{D}$ reduces to
\[\bar{D} = \frac{1}{4\Omega} \sum_{k\omega} \frac{\partial^2 \epsilon(k)}{\partial k_x^2} n(k)\] (A12)
where the momentum distribution
\[n(k) = \langle \mathbf{c}_k^+ \mathbf{c}_k \rangle .\] (A13)
This result is valid for arbitrary one-band dispersion. For the special case of nearest-neighbor (NN) hopping on a square (or cubic) lattice, it is easy to see that the right hand side of eq. (A12) is proportional to the kinetic energy in the $x$-direction, $\langle -K_x \rangle$ in the notation of SWZ. Our result thus reduces to
\[\bar{D} \rightarrow \langle -K_x \rangle/4 .\] (A14)

Finally, we note that our superfluid stiffness $D_s$ is related to that of SWZ by
\[D_s = (\hbar^2/4\pi e^2) \bar{D}^{SWZ} .\] (A15)

**Appendix B: Derivation of Bound $D_s \leq \bar{D}$**

We show that $\chi_{j,j}(\mathbf{q},\omega = 0) \geq 0$ at any temperature. This follows directly from the Lehmann representation
\[\chi_{j,j}(\mathbf{q},\omega = 0) = \sum \frac{e^{-\beta E_i} - e^{-\beta E_j}}{E_j - E_i} \langle i | j^P_x(\mathbf{q}) | j \rangle^2 \geq 0\] (B1)
where $|i\rangle$ and $|j\rangle$ are exact eigenstates of the full Hamiltonian $\mathcal{H}$ in Eq. (A1) with eigenvalues $E_i, E_j$ and $Z = \text{Tr}[e^{-\beta \mathcal{H}}]$. The last inequality follows from $(e^{-x} - e^{-y})(y - x) \geq 0$. At zero temperature, this expression reduces to
\[\chi_{j,j}(\mathbf{q},\omega = 0) = 2\sum_i (E_i - E_0)^{-1} | \langle i | j^P_x(\mathbf{q}) | 0 \rangle |^2 \geq 0\] (B2)
where $|0\rangle$ is the ground state. From Eq. (A7), we thus conclude that
\[D_s \leq \bar{D} .\] (B3)

**Appendix C: Relation between $\bar{D}$ and optical spectral weight**

To see that $\bar{D}$ is proportional to the optical sum rule spectral weight, we identify the dynamical conductivity $\sigma(\omega)$ as the current response to an electric field $\mathbf{E} = -\partial_t \mathbf{A}$
\[i\omega \sigma(\omega) = \chi_{j,j}(\mathbf{q} = 0, \omega) - \frac{4e^2}{\hbar^2} \bar{D}\] (C1)
Using the Kramers-Krönig relation

$$\omega \text{ Im } \sigma(\omega) = -\frac{2}{\pi} P \int_0^\infty d\omega' \text{ Re } \sigma(\omega') \frac{\omega'^2}{\omega'^2 - \omega^2}$$  \hspace{1cm} (C2)

and $\text{Re} \chi_{jx,jx}(\omega \to \infty) \to 0$, we obtain the sum rule for the optical conductivity as

$$\int_0^\infty d\omega \text{ Re } \sigma(\omega) = \frac{2\pi e^2}{\hbar^2} \tilde{D}$$  \hspace{1cm} (C3)

**Appendix D: Real space bound on $$\tilde{D}$$**

Except in the case of a single parabolic band, $\tilde{D}$ depends in general on both the $T$ and the interactions, since the thermal average in $\langle c_{km}^\dagger c_{km'} \rangle$ is calculated using the full $H$. It is thus illuminating to derive an upper bound for $\tilde{D}$ which shows that $\tilde{D}$ must become small when the densities are low or if all the hopping parameters are small. Such a bound for the single-band case with arbitrary dispersion was sketched in the paper. Here we turn to the multi band case.

It is convenient to start with the real space representation

$$\tilde{D} = \frac{1}{4\Omega} \sum_{\mathbf{R},a,\alpha} r_x^2 t_{a\beta}(\mathbf{r}) \langle c_{ia}^\dagger c_{j\beta} \rangle.$$  \hspace{1cm} (D1)

Here both forward and backward hopping are accounted for in $\sum_{\mathbf{R}}$ with $t_{a\beta}(\mathbf{r}) = t_{\beta a}^*(\mathbf{r})$. Since $\tilde{D} \geq 0$ we can use the triangle inequality. Further using the Cauchy-Schwarz inequality we get

$$\tilde{D} \leq \frac{1}{4\Omega} \sum_{\mathbf{R},a,\alpha} r_x^2 \left| t_{a\beta}(\mathbf{r}) \langle c_{ia}^\dagger c_{j\beta} \rangle \right| \leq \frac{1}{4\Omega} \sum_{\mathbf{R},a,\alpha} r_x^2 \left| t_{a\beta}(\mathbf{r}) \right| \sqrt{n_{ia} n_{j\beta}}$$  \hspace{1cm} (D2)

where $n_{ia} = \langle c_{ia}^\dagger c_{ia} \rangle$.

Here and below we define an inner product for operators $A,B$ in terms of the thermal expectation value $\langle A B \rangle$, which allows us to use the Cauchy-Schwarz inequality $|\langle A B \rangle|^2 \leq \langle A^2 \rangle \langle B^2 \rangle$.

**Appendix E: Interband contributions to $$\tilde{D}$$**

We discuss here the conditions under which we can ignore the inter-band contributions to $\tilde{D}$ given by

$$\tilde{D} = \frac{\hbar^2}{4\Omega} \sum_{mm',k}\left(M_{mm'}^{-1}(k) \langle c_{km}^\dagger c_{km'} \rangle \right)$$  \hspace{1cm} (E1)

This requires us to understand when $\langle c_{km}^\dagger c_{km'} \rangle = 0$ for $m \neq m'$. We show here that this is the case, independent of interactions, when (a) either one of the two bands in empty, and (b) when either one of the two bands is fully filled.

We use the Cauchy-Schwarz inequality (see end of Appendix D) to obtain

$$\left| \langle c_{km}^\dagger c_{km'} \rangle \right| \leq \sqrt{n_m(k) n_{m'}(k)}$$  \hspace{1cm} (E2)

where $n_m(k) = \langle c_{mk}^\dagger c_{mk} \rangle$ is the momentum distribution function, and equality holds for $m = m'$. For $m \neq m'$, if either band is completely empty, $n_m(k) = 0$ for all $k$ and the inter-band contribution to $\tilde{D}$ in Eq. (E1) vanishes.

A similar argument for completely filled bands follows from a particle-hole transformation $c_{mk} \rightarrow h_{mk}^\dagger$. Since

$$\langle c_{km}^\dagger c_{km'} \rangle = -\langle h_{km}^\dagger h_{km'} \rangle,$$

$$\left| \langle c_{km}^\dagger c_{km'} \rangle \right| = \left| \langle h_{km}^\dagger h_{km'} \rangle \right| \leq \sqrt{n_m^h(k) n_{m'}^h(k)} = \sqrt{(1 - n_m(k))(1 - n_{m'}(k))}.$$  \hspace{1cm} (E3)
Thus we conclude that for filled and empty bands, the inter-band terms do not contribute to the sum in Eq. (E1), even in the presence of arbitrary interactions.

Finally, we note the simple fact that within band theory there are no inter-band contributions to \( \tilde{D} \). In the absence of interactions (denoted by subscript 0) we obtain

\[
\langle c_{k}^{\dagger} c_{k'} \rangle_0 = f(\epsilon_{m}(k)) \delta_{m,m'}
\]  

(E4)

where \( f \) is the Fermi function.

**Appendix F: Magic Angle Twisted Bilayer Graphene (MA-TBG)**

Magic angles in twisted bilayer graphene were first predicted by the continuum model [12]. Following up on the experimental discovery of correlation-induced superconductivity and superconductivity in MA-TBG, there has been considerable progress in understanding just the electronic structure. The continuum model dispersion is accurately reproduced by the multi-parameter tight binding fit of Koshino et al. [14], which takes into account hopping over distances up to \( 9|L_M| \) where \( L_M \) is the moiré lattice vector. We use the hopping integrals presented in the Supplementary Information file eff_hopping_ver2.dat of ref. [14] to construct the non-interacting Hamiltonian \( H_{K} \) of Eq. (A2). We then identify the unitary matrix \( U(k) \) that diagonalizes \( t_{\alpha\beta}(k) \) (see eq. (A10)) and use it together with \( t_{\alpha\beta}(k) \) to compute the inverse mass tensor

\[
M^{-1}_{mm',\alpha}(k) = \sum_{\alpha'} U_{m,\alpha}(k) \frac{\partial^2 t_{\alpha'\beta}(k)}{\partial (\hbar k_a)^2} U_{\beta,m',\alpha}(k).
\]  

(F1)

Note that we have made explicit here the direction \( a = x, y \) as an additional subscript on \( M^{-1} \).

The inverse mass tensor, obtained from the band structure information as described above, is used to compute \( \tilde{D}_x \) and \( \tilde{D}_y \) and bound \( T_c \) as described in the paper. The additional input needed to determine \( \tilde{D} \) using eq. (A8) is \( \langle c_{k}^{\dagger} c_{k'} \rangle \), and we took two different approaches to compute this.

In the first approach, we looked at SC near half-filling on the hole-doped side of the CNP, and argued that the chemical potential was sufficiently far from the CNP that we can take the band above the CNP to be empty. Then using the result of Appendix E we can ignore all inter-band terms with \( m \neq m' \). For the occupied band we only used the general constraint that \( n(k) \leq 1 \). Using the triangle inequality, we then obtain

\[
\tilde{D}_a \leq \frac{\hbar^2}{4\Omega} \sum_{k_{m,a}} |M^{-1}_{m,m',a}(k)|.
\]  

(F2)

where the empty bands above the CNP are excluded from the sum.

A similar reasoning also works for SC in the vicinity of half-filling on the electron-doped side of the CNP, where we need to use the fact that the bands below CNP are filled to eliminate inter-band terms following Appendix E. We
use a particle hole transformation $c_{mk}^\dagger \to h_{mk}^\dagger$, under which $t_{\alpha\beta}(\mathbf{k}) \to -t_{\alpha\beta}(\mathbf{k})$ and thus $M^{-1} \to -M^{-1}$. We write $\tilde{D}$ in terms of the hole momentum distribution functions $n_{m,\mathbf{k}}^h = \langle h_{mk}^\dagger h_{mk} \rangle$ to get

$$\tilde{D} = \frac{\hbar^2}{4\Omega} \sum_{m,\mathbf{k},\sigma} M^{-1}_{mm,\sigma}(\mathbf{k}) \left[n_{m,\mathbf{k}}^h(\mathbf{k}) - 1\right].$$

We then show that the second term on the right hand side vanishes as follows:

$$\sum_{m,\mathbf{k}} M^{-1}_{mm,\sigma}(\mathbf{k}) = \sum_{\mathbf{k},\alpha\beta} \frac{\partial^2 t_{\alpha\beta}(\mathbf{k})}{\partial (\hbar k_a)^2} \sum_m U_{m,\alpha,\sigma}(\mathbf{k}) U_{m,\beta,\sigma}(\mathbf{k}) = \sum_{\mathbf{k},\alpha} \frac{\partial^2 t_{\alpha\alpha}(\mathbf{k})}{\partial (\hbar k_a)^2} = 0.$$  

We have just used $\sum_m U_{\beta,m}(\mathbf{k}) U_{m,\alpha}(\mathbf{k}) = \delta_{\beta,\alpha}$, which follows from the unitarity of $U$, and then the fact that $\partial^2 t_{\alpha\alpha}(\mathbf{k})/\partial (\hbar k)^2$ is a periodic function with zero mean, whose $\sum_{\mathbf{k}}$ vanishes. Using the triangle inequality and the general constraint $n_{m,\mathbf{k}}^h(\mathbf{k}) \leq 1$, we obtain an expression for electron doping which is similar to the hole-doped case:

$$\tilde{D}_a \leq \frac{\hbar^2}{4\Omega} \sum_{\mathbf{k},m,\sigma} |M_{mm,\sigma}(\mathbf{k})|$$

where now the filled bands below the CNP are excluded from the sum. These bounds, though rigorous, are weak because they involve $|M^{-1}|$ and only very general constraints on $n(\mathbf{k})$.

The second (approximate) approach was to simply use a $T = 0$ (non-interacting) band-theory estimate. We thus use eq. (E4) to obtain

$$\tilde{D}_a \approx \frac{\hbar^2}{4\Omega} \sum_{\mathbf{k},m,\sigma} M_{mm,\sigma}(\mathbf{k}) \Theta(\mu - \epsilon_m(\mathbf{k}))$$

with the chemical potential $\mu$ determined by the density. Using $M_{mm,\sigma}(\mathbf{k})$ (with $a = x,y$) calculated from the tight binding model of ref. [14], this leads to the density-dependent estimates of $\tilde{D}_x$ and $\tilde{D}_y$ shown in Fig. 1 in our paper.

**Appendix G: Attractive Hubbard Model**

It is interesting to ask how our bound on SC $T_c$ in 2D depends on interactions. We use the attractive Hubbard model on a square lattice as a concrete example to understand these trends, and to compare our bound with estimates of $T_c$ from sign-problem free quantum Monte Carlo simulations.

Our bound is $k_B T_c \leq \pi/(8\Omega) \sum_{\mathbf{k},\sigma} \left[ \frac{\partial^2}{\partial k_a^2} \epsilon(\mathbf{k}) \right] n_{\sigma}(\mathbf{k})$. This result can be written in terms of the kinetic energy $\langle -K_\sigma \rangle$ as discussed at the end of Appendix A. The interaction-dependence is contained in the momentum distribution function $n_{\sigma}(\mathbf{k})$ which, as we argued in the paper, must become increasingly broader and flatter as $|U|/t$ increases. In the weak coupling BCS limit (small $|U|/t$) $n_{\sigma}(\mathbf{k})$ is almost like the Fermi function at $T = 0$, very slightly broadened by the superconductivity. On the other hand in the extreme BEC limit (large $|U|/t$) of nearly on-site bosons, the $n_{\sigma}(\mathbf{k})$ of the constituent fermions is essentially flat.

We model this $|U|/t$ trend in the momentum distribution using the BCS-Leggett crossover theory expression

$$n_{\sigma}(\mathbf{k}) = \frac{1}{2} \left[ 1 - \frac{\epsilon(\mathbf{k}) - \mu}{E(\mathbf{k})} \right]$$

where $E(\mathbf{k}) = \sqrt{(\epsilon(\mathbf{k}) - \mu)^2 + \Delta^2}$ is the Bogoliubov quasiparticle energy. The chemical potential $\mu$ and the pair potential $\Delta$ are determined self-consistently for a given density $n$ and attraction $|U|$ by solving the $T = 0$ gap and number equations

$$\frac{1}{|U|} = \frac{1}{\Omega} \sum_{\mathbf{k},\sigma} \frac{1}{2E(\mathbf{k})}$$

$$n = \frac{1}{\Omega} \sum_{\mathbf{k},\sigma} n_{\sigma}(\mathbf{k})$$

We see from Fig. 2 that the $T_c$ obtained from QMC data [27] is always lower than $T_c^{\text{bound}}$. Fig. 2 also shows that the bound is most useful in the intermediate to strong coupling regime, and less useful in the weak coupling regime where $T_c$ is, in fact, well described by $T_c^{\text{MFT}}$, the pair breaking energy scale.
Appendix H: $T_c$ Bounds in spatially anisotropic systems

We collect here some results on the role of spatial anisotropy focusing mainly on 2D. We note that various quantities that we have considered are different in different directions labeled by $a = x, y$. We have shown that

$$D_{s,a}(T) \leq \tilde{D}_{a}(T). \quad (H1)$$

The most conservative bound on $T_c$ in 2D is then

$$k_B T_c \leq \frac{\pi}{2} \max \{\tilde{D}_x, \tilde{D}_y\}. \quad (H2)$$

Clearly this bound is not optimal because we expect $T_c$ to go to zero if either $D_{s,x}$ or $D_{s,y}$ goes to zero. Using BKT theory we can show that

$$k_B T_c = \frac{\pi}{2} \left( D_{s,x}(T_c) D_{s,y}(T_c) \right)^{1/2} \quad (H3)$$

which leads to the improved bound

$$k_B T_c \leq \frac{\pi}{2} \left( \tilde{D}_x(0) \tilde{D}_y(0) \right)^{1/2} \quad (H4)$$

To derive eq. (H3) we start with the Free energy for phase fluctuations

$$\mathcal{F} = \frac{1}{2} \int dx \, dy \left[ D_{s,x}(\partial_x \theta)^2 + D_{s,y}(\partial_y \theta)^2 \right]. \quad (H5)$$

We then rescale lengths using $x' = (D_0/D_{s,x})^{1/2} x$ and $y' = (D_0/D_{s,y})^{1/2} y$, where $D_0$ is any convenient energy scale for normalization, to obtain

$$\mathcal{F} = \frac{1}{2} \left( D_{s,x} D_{s,y} \right)^{1/2} \int dx' \, dy' \left[ (\partial_x \theta)^2 + (\partial_y \theta)^2 \right]. \quad (H6)$$

This immediately leads to the generalization of the Nelson-Kosterlitz result in eq. (H3). We emphasize that the reason this seemingly naive argument works is that the line of fixed points below $T_c$ are actually described by a Gaussian theory and the BKT $T_c$ is precisely when vortex-antivortex unbinding becomes relevant at a Gaussian fixed point. We thank Steve Kivelson and C. Jayaprakash for very useful conversations related to this argument.