Mean-field theory for superconductivity in twisted bilayer graphene

Teemu J. Peltonen,1 Risto Ojajärvi,1 and Tero T. Heikkilä1

1Department of Physics and Nanoscience Center, University of Jyväskylä, P.O. Box 35 (YFL), FI-40014 University of Jyväskylä, Finland

(Dated: May 14, 2018)

Recent experiments show how a bilayer graphene twisted around a certain magic angle becomes superconducting as it is doped into a region with approximate flat bands. We investigate the mean-field superconducting state in such a system and show how the state evolves as the twist angle is tuned, and as a function of the doping level. We argue that part of the experimental findings could well be understood to result from an attractive electron–electron interaction mediated by electron–phonon coupling, but the flat-band nature of the excitation spectrum makes also superconductivity quite unusual. For example, as the flat-band states are highly localized around certain spots in the structure, also the superconducting order parameter becomes strongly inhomogeneous.

Experiments on strongly doped graphene [1–4] have shown that with proper preparations, graphene can be driven to the superconducting state. Such experiments indicate that the lack of superconductivity in undoped graphene is not necessarily due to a lack of an (effective) attractive electron–electron interaction with strength λ that would drive graphene superconducting, but rather the small density of states (DOS) close to the Dirac point. Technically, in contrast to the Cooper instability for metals taking place with arbitrarily small λ, superconductivity in an electron system with a massless Dirac dispersion $c^2 = v_F^2 p^2$ and an energy cutoff $\epsilon_c$ has a quantum critical point $\lambda_c = \pi \hbar^2 v_F^2/(2epsilon_c)$ [5] such that for $\lambda < \lambda_c$, mean-field superconductivity does not show up at any temperature. From this perspective, doping to a potential $\mu$ leads to an increased DOS, and thereby to a non-vanishing critical temperature $T_c \approx |\mu| \exp\left(-\lambda_c/\lambda - 1\right)/\left|\mu\right| - 1$. An alternative approach would be to change the spectrum and increase the density of states close to the Dirac point. The extreme limit would be an approximatively flat band of size $\Omega_{FB}$, where the group velocity tends to zero. In such systems the critical temperature is a linear function of the coupling strength $T_c = \lambda \Omega_{FB} / \pi^2$ [6] and quite high $T_c$ can be expected even without extra doping [7–13].

Recent observations of superconductivity in twisted bilayer graphene [TBG, see Fig. 1(a)] [14] take place in systems where theoretical studies have predicted the occurrence of asymptotically flat bands [15–25]. Here we study the mean field theory of superconductivity in such systems, starting from the hypothesis [26] that the observations could be explained from the flat-band perspective. In particular, we use the model of Refs. [15, 20] for the spectrum of the twisted bilayer, add an on-site attractive interaction of strength $\lambda$, and evaluate the mean-field order parameter profile.

We find that the order parameter, and along with it the mean-field critical temperature, have a similar non-monotonous behavior with respect to the twist angle as in the experiments. We also predict the behavior of the density of states in the superconducting state, resulting from the peculiarities of the flat-band eigenstates and from the position dependence of the superconducting gap [Fig. 1(b)]. Even if our pairing interaction is quite simple, the resulting energy dependent density of states is quite unusual. In addition, we show how doping away from the flat band eventually destroys superconductivity.

Normal state. We describe the normal state of TBG with the model of Refs. [15, 20]. With this model, we can describe the twist angles $\theta$ at which the lattices $L$ and $L^\theta$ of the two graphene layers are commensurate, so that the system as a whole is periodic in the moiré superlattice $SL$. For small $\theta$, the set of all twist angles is dense in commensurate angles, in which case we can approximate an incommensurate structure by a commensurate one [20]. Hence for simplicity we study only the structures for which the exact superlattice matches the approximate superlattice. These structures are characterized by a rotation parameter $m \in \mathbb{N}$, for which the
rotation angle is given by
\[ \cos(\theta) = \frac{3m^2 + 3m + 1/2}{3m^2 + 3m + 1}. \] (1)

The primitive vectors of the superlattice SL are given by
\[ t_1 = ma_1 + (m+1)a_2, \quad t_2 = -(m+1)a_1 + (2m+1)a_2 \]
and the primitive vectors of the reciprocal superlattice SL* are
\[ G_1 = \frac{4\pi}{|t_1|}((3m+1)a_1 + a_2), \quad G_2 = \frac{4\pi}{|t_1|}(-(3m+2)a_1 + (3m+1)a_2), \]
where the lattice constant of the superlattice is \(|t_1| = \sqrt{3m^2 + 3m + 1}a\) and the graphene lattice primitive vectors are \(a_1 = (1, \sqrt{3}a)/2\) and \(a_2 = (-1, \sqrt{3})a/2\) with \(a\) the lattice constant \([15]\). The first Brillouin zone of the TBG is denoted by \(\mathbb{R}^2/SL^*\). In the following, we assume that \(G \in SL^*\) and \(k \in \mathbb{R}^2/SL^*\), and also that the sums and integrals are restricted to these sets.

In the normal state, TBG is described by a low energy effective Hamiltonian \([15]\)

\[ \mathcal{H}_{pk}(G, G') = \begin{cases} \hbar v_F \sigma^\rho \cdot (k + G + \rho \Delta K/2) - \mu |\delta_{G,G'}| & \text{for } 1, \\ \hbar v_F \sigma^\rho_0 \cdot (k + G - \rho \Delta K/2) - \mu |\delta_{G,G'}| & \text{for } 2, \end{cases} \] (2)

where the matrix element corresponds to the layer structure and \(\rho \in \{+, -\}\) is the valley index with + corresponding to \(K\) and - to \(K' = -K\). Furthermore, each entry is a \(2 \times 2\) matrix due to the sublattice structure in graphene. The diagonal terms in Eq. (2) describe the Dirac dispersion in the two layers and are diagonal also in \(G\). Here, \(\sigma^\rho = (\rho \sigma_x, \rho \sigma_y)\). For the second layer we include the rotation \(\theta\) so that \(\sigma^\rho_0 = e^{i\theta |\sigma_0|/2} \sigma^\rho e^{-i\theta |\sigma_0|/2}\).

\(\Delta K = K' - K\) is the relative shift of the Dirac cones between the layers. The coordinates correspond to those of layer 1 as measured from the \(K\)-point, but shifted with a vector \(+\Delta K/2\) for layer 1 and \(-\Delta K/2\) for layer 2. With this choice, the relative momentum \(k\) on both layers corresponds to the same absolute momentum. Furthermore \(\mu\) is the chemical potential describing the effect of doping, here taken to be equal in both layers.

The off-diagonal terms in the Hamiltonian describe the coupling between the two layers. The matrix element at valley \(\rho\) between a state in sublattice \(\alpha\) in layer 1 and a state in sublattice \(\beta\) in layer 2 is

\[ \delta^\rho_{\alpha \beta}(G) = \frac{1}{N} \sum_r e^{-iG \cdot (r + \delta_{\alpha \beta})} e^{i\rho \delta_{\alpha \beta}(r)} \mathcal{K}^{\alpha \beta}(r) t_1(\delta_{\alpha \beta}(r)), \] (3)

where \(\delta_{\alpha \beta}(r)\) is the horizontal displacement vector between the site at \(r\), sublattice \(\alpha\) in layer 1 and the nearest neighbor at sublattice \(\beta\) in layer 2. \(\delta_1\) denotes one of the nearest neighbor vectors connecting the graphene A and B sublattices. The sum is over the graphene lattice sites in the superlattice unit cell, and \(N\) denotes the number of these sites. For the interlayer hopping energy \(\mathcal{H}_1(\delta)\) we use the same Slater–Koster parameterization as in Ref. \([15]\). Furthermore we approximate the interlayer coupling by only including the matrix elements with \(G \in \{0, -G_1, -G_1 - G_2\}\) (valley \(K\)) or \(G \in \{0, G_1, G_1 + G_2\}\) (valley \(K'\)), since they are an order of magnitude larger than the rest.

For \(\theta \approx 1^\circ\), the electronic dispersion becomes almost flat \([19]\) and the group velocity \(dE/d\rho\) tends towards zero. In Fig. 2 we plot the resulting normal-state dispersion (a–c) and the (local and total) density of states (d–i) close to this “magic” angle. The exact value of this magic angle depends on the details of the hopping model. In our case it is around 0.96°, i.e., somewhat lower than what was found in \([19]\). However, the qualitative behavior of the local density of states is rather similar to the previous models. In particular, there are two closely spaced DOS peaks signifying the flattening of the bands. The local density of states is plotted along the line shown in Fig. 1, including three high-symmetry points with AB, AA, and BA stacking. These correspond to \(r = -1/3, 0,\) and \(1/3\), respectively.

**Superconducting state.** We assume that there is a local attractive interaction with strength \(\lambda\) (see Appendix for...
further details). We do not consider the specific nature of the pairing interaction and for the purposes of this paper it can be mediated by phonons or other bosonic modes. This model disregards the retardation effects due to such a mechanism, but is a valid approximation to the more general Eliashberg approach for weak coupling [27, 28]. However, we also disregard direct Coulomb interaction. As shown in [28], within mean field theory and for weak coupling, it leads to an effective attractive interaction strength λ_{eff}, smaller than the bare λ. However, as long as λ_{eff} > 0, there is a possibility for a superconducting state.

Within a mean-field theory in the Cooper channel we find a self-consistency equation for a local superconducting order parameter (see Appendix for further details). Assuming furthermore that this order parameter shares the periodicity of the moiré superlattice, we find the self-consistency equation for the superconducting order parameter,

\[
\Delta_{\alpha i}(G) = \lambda \sum_{\rho, b} \sum_{G'} \int \frac{dk}{(2\pi)^2} \tanh \left( \frac{E_{\rho bk}}{2k_B T} \right) \left( \frac{\hbar}{\Delta(G') - \Delta(G-G')} \right) u_{\rho bk, \alpha i}(G') v_{\rho bk, \alpha i}(G-G'),
\]

where the band sum b is calculated over the positive energy bands, α ∈ {A, B} is the sublattice index and i ∈ {1, 2} is the layer index. u_{\rho bk} and v_{\rho bk} are the eigenvectors of the Bogoliubov–de Gennes equation,

\[
\sum_{G'} \left( \frac{\hbar}{\Delta(G'-G)} - \Delta^*(G'-G) \right) \left( \begin{array}{c} u_{\rho bk}(G') \\ v_{\rho bk}(G') \end{array} \right) = E_{\rho bk} \left( \begin{array}{c} u_{\rho bk}(G) \\ v_{\rho bk}(G) \end{array} \right).
\]

We solve this self-consistent order parameter with a few values of the interaction constant λ and for a few different twist angles close to the magic angle. We include in the sum the energy levels closest to zero energy. We have checked that the results do not change much if the higher-energy levels are also included, up to energy cutoffs of the order of 100 meV. For comparison between different angles, we choose the chemical potential corresponding to the charge neutrality, marked in Figs. 2(g–i) with a dashed line. The resulting total density of states is plotted in Fig. 2(j–k), to allow for a comparison to the normal state. The corresponding local density of states (not shown) has the same localized structure as in the normal state, but the energy-dependence is modified similarly as the total DOS. The maximum of the position-dependent Δ, which according to numerics is equal in both layers and sublattices, is plotted in Fig. 3(a) for different values of the twist angle and for four different coupling strengths. The precise angle for the maximum depends a bit on the chosen coupling strength. Moreover max(Δ) is almost a linear function of λ [see Fig. 3(b)], as appropriate for a flat-band superconductor. Far from the magic angle, the Fermi speed v_F(θ) increases so that the chosen λ is below the critical value λ_c. This is why Δ vanishes for angles θ ≥ 1.1°.

We can analyze the resulting magnitude of Δ based on a flat-band result (assuming a constant Δ and E_{ρ bk} ≈ Δ for an extreme flat band) according to which (see Appendix for further details) Δ = Ω_{FB}/π^2, where Ω_{FB} = Ω_{min} = 8π^2/(\sqrt{3}|t_1|^2). This yields Δ = 1.3×10^{-3}λ/a^2 for m = 34 corresponding to the magic angle. For comparison a linear fit to the linear region in Fig. 3(b) gives max(Δ) = -0.07 meV + 0.96×10^{-3}λ/a^2. The magnitude hence agrees very well with this simple model.

In Fig. 4 we show the temperature dependence of Δ for m = 34, from which we may infer the approximate lineare relation k_BT_c ≈ 0.25 max(Δ(T = 0)) for the critical temperature. The prefactor is somewhat lower than for an extreme flat band with a constant Δ, for which k_BT_c = Δ/2 (see Appendix for further de-
The difference is most likely explained by the nonvanishing bandwidth and the position-dependent $\Delta$ of our model. The maximum critical temperatures for the models calculated in Fig. 3(a) range from 3 K to about 20 K. The lower end of these values, calculated with $\lambda = 1{\text{ eV}^2/a^2}$, is thus quite close to that found in [14]. We stress that this is the mean-field critical temperature; the observed resistance transition is most likely rather a Berezinskii–Kosterlitz–Thouless transition and hence somewhat lower than the mean-field one. Note that despite the flatness of the bands, the supercurrent can be non-vanishing in the case when the eigenstate Wannier functions overlap [29] as is the case for TBG.

Dependence on electrostatic doping. Besides $\theta$-dependence, we can check how doping away from the center of the two DOS peaks affects the superconducting state. Figure 5(a) shows the charge density vs. chemical potential in the normal state. In Fig. 5(b) we plot the order parameter $\max(\Delta(\delta\mu))$ for different values of the doping $\delta\mu$ as measured from the charge neutrality point. Close to the magic angle, the energy scale of superconductivity exceeds that of the normal-state dispersion, and hence the only effect of the doping is to move away from the flat-band regime, suppressing superconductivity [30]. For the lowest value of $\lambda$ in the plot, $\max(\Delta)$ is smaller than the bandwidth, and hence doping to the DOS peaks from their center can enhance superconductivity, resulting in a non-monotonous doping dependence. Note that as the effect of direct Coulomb interaction neglected here depends on the amount of screening, it also most likely depends on the charge density. As a result, this may lead to a more complicated $\Delta(\delta\mu)$. However, as this depends a lot on the electromagnetic environment of the system, we will study this effect in a further work.

Concluding, we find that a BCS-type mean field model with relatively weak attractive interaction constant possibly even due to electron–phonon coupling (see Appendix for further details) can explain the occurrence of superconductivity in twisted bilayer graphene. We also make a number of predictions concerning the mean-field superconducting state, in particular the density of states and doping dependence. Our results form hence a check-point for further studies that use a simplified picture of the TBG flat band states, or consider mechanisms beyond the one in this paper. Our results could also have relevance in explaining the observations of superconductivity in twisted interfaces of graphite [31–33]. Clearly our mean-field theory fails to explain the insulator state [34] found experimentally in TBG at $n \approx \pm 2e/A_{\text{moiré}}$ as well as the lack of superconductivity for hole doping [14, 35]. There have been many suggestions of an unconventional (or more unconventional than here) superconducting state both for regular graphene [36, 37] and for TBG [24, 38–46], typically directly related with the Coulomb interaction, and in some cases related with non-local interactions. We point out that our simple BCS model disregards the strain effects in moiré bands, as well as the possible dependence of the interaction constant on the twist angle. Whereas such mechanisms may play a role in TBG, we believe that the simplest BCS-type mean field superconductivity should also be considered as a viable effective model of the observations. Nevertheless, even in this case superconductivity would be highly exceptional, for example because it can be so strongly controlled by electrostatic doping.

This project was supported by Academy of Finland Key Funding project (Project No. 305256), and the Center of Excellence program (Project No. 284594).
Derivation of the self-consistency equation

The Hamiltonian for the attractive interaction is

$$H_{\text{int}} = \frac{\lambda}{2} \sum_{\sigma, \rho, \rho'} \int d\mathbf{r} \psi_{\sigma \rho \alpha i}^\dagger(\mathbf{r}) \psi_{\sigma \rho' \alpha i}^\dagger(\mathbf{r}) \psi_{\sigma \rho' \alpha i}(\mathbf{r}) \psi_{\sigma \rho \alpha i}(\mathbf{r}),$$

(A1)

where $\alpha \in \{A, B\}$ is the sublattice index and $i \in \{1, 2\}$ is the layer index. Doing the mean field approximation in the Cooper channel for Eq. (A1), we find a selfconsistency equation for a local superconducting order parameter

$$\Delta_{\alpha i}(\mathbf{r}) = \lambda \sum_{\rho, b} \int \frac{d\mathbf{k}}{(2\pi)^2} \tanh \left(\frac{E_{\rho B k}}{2k_B T}\right) \times u_{\rho \rho', \alpha i}(\mathbf{r}) v_{\rho' \rho, \alpha i}(\mathbf{r}),$$

(A2)

where $u_{\rho \rho', \alpha i}$ is the $(\alpha, i)$-component of the spinor $u_{\rho \rho'}$. The spinors $u_{\rho \rho'}$ and $v_{\rho \rho'}$ are determined by solving the Bogoliubov–de Gennes equation

$$\begin{pmatrix} \mathcal{H}_{\rho}(\mathbf{r}) - \Delta(\mathbf{r}) & u_{\rho \rho'}(\mathbf{r}) \\ \Delta^*(\mathbf{r}) - \mathcal{H}_{\rho'}(\mathbf{r}) & v_{\rho \rho'}(\mathbf{r}) \end{pmatrix} = E_{\rho \rho'} \begin{pmatrix} u_{\rho \rho'}(\mathbf{r}) \\ v_{\rho \rho'}(\mathbf{r}) \end{pmatrix}.$$

(A3)

Substituting the Bloch wave expansion

$$\begin{pmatrix} u_{\rho \rho'}(\mathbf{r}) \\ v_{\rho \rho'}(\mathbf{r}) \end{pmatrix} = e^{i \mathbf{k} \cdot \mathbf{r}} \sum_{\mathbf{G}'} e^{i \mathbf{G}' \cdot \mathbf{r}} \begin{pmatrix} u_{\rho \rho'}(\mathbf{G}') \\ v_{\rho \rho'}(\mathbf{G}') \end{pmatrix},$$

(A4)

for the eigenstates into Eq. (A3) and assuming $\Delta(\mathbf{r})$ to be periodic in the superlattice, we find the Fourier space Bogoliubov–de Gennes equation [Eq. (5) in the main text] and the Fourier space version of the self-consistency equation [Eq. (4) in the main text].

Charge density

The non-coupled system of twisted bilayer graphene is charge neutral at the chemical potential $\mu = 0$. The charge density due to the electrons at that potential is

$$n_0 = 2e V \sum_{b \in B, \mathbf{k}} f_0(\epsilon_{\alpha b \mathbf{k}}) = 2e V \sum_{b \in \Omega, \mathbf{k}} f_0(\epsilon_{\alpha b \mathbf{k}}) + n_{\text{high}}$$

(A5)

with

$$n_{\text{high}} = 2e V \sum_{b \in B' \Omega, \mathbf{k}} f_0(\epsilon_{\alpha b \mathbf{k}}),$$
dependent, the Bogoliubov–de Gennes equation can be
notation of the main paper, along with some estimates.
low especially the treatments in [5, 30] and adopt to the
scribe the observations of superconductivity in twisted
claiming that a quite simple BCS-style model could de-
known. Here we reconcile this notion with our results
of bands in one valley.

The charge neutrality point \( \mu^* \) is determined from the
equation \( n(\mu^*) = 0 \). It is shown for different twist angles
in Figs. 2(g–i) of the main text, and is always located in
the middle between the two DOS peaks.

### Simplified model of the superconducting state

The notion of weak or absent electron–phonon me-
diated superconductivity in pristine graphene is widely
known. Here we reconcile this notion with our results
claiming that a quite simple BCS-style model could de-
scribe the observations of superconductivity in twisted
bilayer graphene. These results are not new, but we fol-
low especially the treatments in [5, 30] and adopt to the
notation of the main paper, along with some estimates.

We start from the generic self-consistency equation for
the mean-field order parameter \( \Delta \). If \( \Delta \) is position
independent, the Bogoliubov–de Gennes equation can be
solved to yield

\[
\Delta = 4\lambda \int_{k}^{k_c} \frac{d\mathbf{k}}{(2\pi)^2} \frac{\Delta}{E_k} \frac{1}{\tanh \left( \frac{E_k}{2kB_T} \right)},
\]

where the prefactor 4 comes from summation over the
valley and band indices, where in the band sum we in-
clude only the doubly degenerate lowest positive energy
band. The cutoff \( k_c \) is specified more below. We there-
more assume that \( E_k = \sqrt{\epsilon_k^2 + \Delta^2} \). Here and below,
without loss of generality we assume \( \Delta = |\Delta| \geq 0 \). Our idea
is to solve the self-consistency equation in three cases: (i)
at the Dirac point for a Dirac dispersion \( \epsilon_k^2 = \hbar^2v_F^2k^2 \),
(ii) for a Dirac dispersion at non-zero doping \( \mu \), i.e.,
\( \epsilon_k^2 = (\hbar v_F k - \mu)^2 \), and (iii) for a flat band with and
without doping, \( \epsilon_k \approx \mu \) for \( k \in \Omega_{FB} \). In each case we
have the normal-state solution \( \Delta = 0 \), which we exclude
by dividing both sides in Eq. (A8) by \( \Delta \).

Note that Eq. (A8) does not represent the full self-
consistency equation solved in the main text. Rather, we
use it here simply to provide estimates of the behavior of
\( \Delta \) in various limits.

### Linear dispersion, no doping

Far away from the magic angle, the twisted bilayer beha-
ves as if the two graphene layers would be almost un-
coupled. This means that the low-energy dispersion
exhibits two separate copies of the graphene Dirac dis-
persion. Inserting an ultraviolet energy cutoff \( \epsilon_c = \hbar v_F k_c \)
and performing the integral for the \( T = 0 \) gap function,
the self-consistency equation goes to the form

\[
\frac{\pi \hbar^2 v_F^2}{2\lambda} = -\Delta + \sqrt{\Delta^2 + \epsilon_c^2},
\]

or

\[
\Delta = \frac{\pi \hbar^2 v_F^2}{4} \frac{\lambda^2 - \lambda_c^2}{\lambda \lambda_c^2},
\]

where \( \lambda_c = \pi \hbar v_F^2/(2\epsilon_c) \). Since \( \Delta \geq 0 \), this solution
makes sense only if \( \lambda > \lambda_c \), and otherwise the only pos-
sible solution is the normal state \( \Delta = 0 \).

In pristine graphene, the critical interaction strength
be can written also in terms of the nearest-neighbour
hopping term \( \gamma_0 \approx 3 \text{eV} \) [47]. Namely, within a
nearest-neighbour tight-binding model the Fermi speed
of graphene is \( v_F = \sqrt{3\gamma_0 a/(2\hbar)} \), where \( a \) is the graphene
lattice constant. We hence get

\[
\lambda_c = \frac{3\pi \gamma_0}{8} \frac{\gamma_0 a^2}{\epsilon_c}.
\]

If the attractive interaction results from electron–phonon
coupling, a typical cutoff energy could be of the or-
der of the Debye energy 200 meV [48]. In this case
\( \lambda_c \approx 50 \text{eV}a^2 \), 5 to 50 times larger than the values of
λ used in our work. 100 to 200 meV is also the range of the maximum cutoff energy that we have used in our numerical results when including the contribution from higher bands. Even if the cutoff \( \epsilon_c \) would be of the order of \( \gamma_0 \), the resulting \( \lambda_c \) would be one order of magnitude larger than the smallest \( \lambda \) used in our results.

**Linear dispersion, with doping**

Let us try to reconcile the observations of superconductivity in Li or Ca doped graphene with the above idea. These cases are more accurately described by [49] within the Eliashberg theory. Here we just show in which sense doping fits into the above picture. Assuming \( \epsilon_k = \pm \hbar v_F k - \mu \) and \( \Delta < \epsilon_c \), and cutting the integral at \( \epsilon_k = \epsilon_c \) the self-consistency equation at \( T = 0 \) becomes [5]

\[
\frac{\pi \hbar^2 v_F^2}{2\lambda} = \sqrt{\Delta^2 + \epsilon_c^2} - \sqrt{\Delta^2 + \mu^2 + |\mu| \ln \frac{|\mu| + \sqrt{\Delta^2 + \mu^2}}{\Delta}}.
\]

Let us assume that \( \Delta \ll |\mu|, \epsilon_c \) so that we can expand the right hand side in \( \Delta \). In this case we find an analytic solution for \( \Delta \),

\[
\Delta = 2|\mu| \exp \left[ -\frac{\epsilon_c}{|\mu|} \left( \frac{\lambda_c}{\lambda} - 1 \right) - 1 \right].
\]

Let us assume a cutoff energy \( \epsilon_c = 200 \text{ meV} \) and a coupling strength \( \lambda = \lambda_c/22 \) (corresponding to about 5 eV/a^2 with the above estimates). In this case, with \( \mu = 0.7 \text{ eV} \) we would get \( \Delta = 1.3 \text{ meV} \). This corresponds to a critical temperature of 9 K, in the same range as the one that was measured in Li or Ca doped graphene [1–4].

**Flat band estimate**

Let us now make similar estimates for the flat-band case of the moiré superlattice. In this case, we assume that \( \Delta \) is larger than the bandwidth of the lowest-energy band. Within that band, we can hence approximate \( E_k \approx \Delta \) in Eq. (A8) and at \( T = 0 \) the integral is over a constant function. As a result, we get

\[
\Delta_{\text{FB}} = \frac{\lambda}{\pi^2} \Omega_{\text{FB}} = \frac{8\lambda}{\sqrt{3}(3m^2 + 3m + 1)a^2}, \quad (A14)
\]

where \( \Omega_{\text{FB}} = \frac{8\pi^2}{\sqrt{3}(3m^2 + 3m + 1)a^2} \) is the area of the first Brillouin zone of the moiré superlattice. Within the model adapted in the main text, the magic angle is around \( m \approx 34 \), in which case we would get \( \Delta_{\text{FB}} = 1.3 \times 10^{-3} \lambda/a^2 \). In Fig. 3(b) of the main text, the solid line has a slope of \( 1 \times 10^{-3} \lambda/a^2 \), i.e., very close to this simple estimate.

The temperature dependent \( \Delta \) in the flat-band case is obtained by solving

\[
\Delta = \Delta_{\text{FB}} \tanh \left( \frac{\Delta}{2k_B T} \right). \quad (A15)
\]

At the critical temperature, \( \Delta \to 0 \), and we can hence expand the right hand side to the linear order in \( \Delta/(2k_B T_c) \). This directly yields \( T_c = \Delta_{\text{FB}}/(2k_B T) \).

In the case of a non-zero potential \( \mu \), we can use \( E_k \approx \sqrt{\mu^2 + \Delta^2} \) in the self-consistency equation. It then becomes (for \( \Delta > 0 \))

\[
\Delta = \Delta_{\text{FB}} \frac{\Delta}{\sqrt{\mu^2 + \Delta^2}}, \quad \text{or} \quad \Delta = \sqrt{\Delta_{\text{FB}}^2 - \mu^2}. \quad (A16)
\]

In this case superconductivity is hence suppressed when the absolute value of the chemical potential is larger than \( \Delta_{\text{FB}} \).