Cooper pairs [1] of mutually attracting electrons form the building blocks of superconductivity [2]. Thirty years after the discovery of high-temperature superconductivity in cuprates [3], many details of the pairs remain unknown, including their size and shape. Here we apply brand new ARPES-based methods [4] that allow us to reconstruct the shape and size of the pairs in Bi$_2$Sr$_2$CaCu$_2$O$_{8+δ}$. The pairs are seen to form a characteristic “starfish” shape that is very long (>50Å) in the near-nodal direction but extremely short (~4.5Å) in the antinodal (Cu-Cu) direction. We find that this ultrashort antinodal length scale, which is of order a lattice constant, is approximately constant over a wide range of doping levels even as many other parameters including the pairing strength change. This suggests that this new length scale, along with the pair shape, is one of the most fundamental characteristics of the pairs. Further, the shape and ultrashort length scale should make the pairs create or intertwine with variations in charge and pair density, center on various types of lattice positions, and potentially explain aspects of the nematic order in these materials.
The mechanism of high-temperature superconductivity in cuprates remains one of the defining problems in condensed matter physics. By exploiting recent developments that greatly improve the quantitative accuracy of angle-resolved photoemission spectroscopy (ARPES), especially in its ability to extract the self-energies [4], we are able to directly measure all key electronic parameters and pairing interactions as a function of angle. This gives us the ability, for the first time, to reconstruct the angular-dependent pairing length scales in a superconductor (Fig. 1). In contrast to most previous concepts that hold that the pairs in the cuprates are round as in conventional superconductors [5], we find that the pair length scales have the shape of a 4-legged starfish. The starfish-shaped pairing lengths have very long (>50Å) arms near the node (45° to the Cu-Cu bonds) and a very short body of about 4.5Å along the Cu-Cu bonds. Previous results only indicated the overall average superconducting coherence length of the pairs to be around 20Å – a properly weighted averaging of our result returns a similar value (see supplementary information S1 and Fig. S1). As a function of doping, our data shows the surprising finding that the antinodal pair size remains approximately constant and is pinned near the lattice parameter of 3.8Å – a result that further indicates that this short length scale is not accidental but rather is likely a defining or driving characteristic of the pairing interaction.

For conventional superconductors [2], where the pair wave functions are well-known, the superconducting pair size can be defined as the expectation value of the root-mean-square radius, in either real-space or k-space, as [6,7]

\[ \xi_{\text{pair}} = \frac{\int dr |\psi(r)|^2}{\int dr |\psi(r)|^2} = \frac{\int dk \left| \partial_k \psi(k) \right|^2}{\int dk \left| \psi(k) \right|^2} \]

where \( \psi(r) \) is the Cooper-pair wave function in real space, \( \partial_k \) is the gradient operator in momentum space (k space), and \( \psi(k) \) is the Fourier transform of \( \psi(r) \) to k space [8, 9]. The wavefunction is \( \psi(k) = \Delta(k)/E_k \) with \( E_k = \sqrt{\varepsilon_k^2 + \Delta(k)^2} \) [8, 9], where \( \varepsilon_k \) is the non-gapped dispersion.
The expression of the pair size includes multiple terms (see all terms including their derivation in supplementary information S2), but the dominant term of the pair size is written as:

$$
\xi_{\text{pair}}(\theta) \approx \sqrt{\int dk_\perp \frac{\Delta_\theta^2}{E_k^6} \left( \partial_\perp \epsilon_k \right)^2}
$$

where \( \theta \) is the Fermi surface angle, \( E_k = \sqrt{\epsilon_k^2 + \Delta_\theta^2} \), and the k-integrals and the gradients are taken in the direction perpendicular to the Fermi surface (see the justification for the dominant term in supplementary information S2 and Fig. S2). The important point is that all parameters necessary to compute Eqn. 2 are now directly accessible from ARPES experiments, especially after the recent introduction of a reliable method for the 2-dimensional analysis of ARPES data [4], which delivers consistent and accurate results for the band energy \( \epsilon_k \), the pairing gaps \( \Delta_\theta \), and especially the single-particle self-energies \( \Sigma'(k, \omega, T) \) and \( \Sigma''(k, \omega, T) \) and the related renormalization parameter \( Z(k, \omega, T) \) enter into the evaluation of \( \epsilon_k \) and \( E_k \). Utilizing previously published low-temperature ARPES data from a \( T_c=85K \) underdoped sample we extract the pair size as a function of angle, as plotted in Fig. 1a and 1b. It is seen that the pair length scales have a shape reminiscent of a four-armed starfish, with the arms extending to very large distances near the nodes, with a compact body of 4.5 Å in the antinodal direction. As the gap goes to zero at the node, the pair size concept is more complicated - thus, in Fig. 1a and 1b, we omitted the exact nodal direction indicated by the blank in the figure.

The above result is a natural consequence of the d-wave symmetry of the pairing interaction strengths and can be compared to the more commonly discussed Pippard coherence length [10] that is the length
scale over which the superconducting state recovers when it is destroyed locally. For a BCS superconductor, the coherence length is given by

$$\xi_{cl}(\theta) = \frac{\hbar v_F}{\pi \Delta_\theta}$$

(3)

where $v_F$ is the Fermi velocity and $\Delta$ the superconducting gap (pairing strength) [11]. The large gap in the antinodal direction therefore naturally corresponds to the smaller length scale of the starfish body, while the very small gap near the nodes corresponds to the very long arms. Supplementary Fig. S1 shows our ARPES-based result of the coherence length. While previous measurements only indicated the overall average superconducting coherence length of the pairs to be around 20Å [12, 13], our result, when properly weighted by an average around the zone, returns a similar value (see detailed discussion in supplementary information S1).

In addition to the UD85K sample of Fig. 1a and 1b, we present the antinodal pair size from 5 other superstructure-free (Bi,Pb)$_2$Sr$_2$CaCu$_2$O$_{8+\delta}$ samples and one Bi$_2$Sr$_2$CaCu$_2$O$_{8+\delta}$ sample with optimal doping (Fig. 1c), with the 7 samples having dopings that span the underdoped and overdoped regimes. The antinodal pair sizes are extremely short and doping independent. These antinodal pair sizes are extracted from the spectral data (Figs. 2a-2g) from all samples with temperature well below $T_C$ for the antinodal region of the Brillouin zone where the gap is largest (red line in Fig. 2o). The photon energy has been chosen so as to emphasize the antibonding band of the bilayer-split bands [14-15] – a comment about the bonding band will come later. 2D fits to these data following the procedure of ref [4] are shown in Fig. 2h-2n. Fig. 2p and 2q also show the $k_F$ EDCs from each set of experimental data and fits [16].

The individual parameters extracted from the fits of Fig. 2 are plotted as a function of doping in Fig. 3. With these individual parameters, we are then able to write the antinodal pair wavefunction and thus
extract the low temperature antinodal pair size for each doping level using Eqn. 2. The results of these calculations for a wide range of doping levels are plotted in Fig. 1c, showing that the antinodal pair size is essentially unaffected by doping level, even though the various constituent parameters of Fig. 3a-d vary significantly with doping. For a few doping levels we have separately extracted similar data on the bonding band of the bilayer split bands. These bands are farther from the Fermi energy ($E_{BB}$ and $k_F$ are larger than the ones for the antibonding bands), and though they have the same gap sizes as their antibonding counterparts we find that the antinodal pair sizes are significantly larger for the bonding bands (see supplementary information S4 and Fig. S3, Fig. S4).

The most surprising aspect of these results is the extremely short and universal antinodal pair size of order 4.5Å. This is the case even though the individual parameters determining this quantity (Eqn. 2) all vary with doping (Fig 3), in some cases strongly. We therefore argue that this constant ultrashort length scale, which is close to the lattice parameter of 3.8 Å, can not be a coincidence but rather must be a natural and defining aspect of the pairing interactions in the cuprates. Such a short pairing length scale is not consistent with a conventional pairing mechanism that exchanges long-wavelength (low-q) phonons, or low-q spin fluctuations, but would likely admit a purely electronic mechanism [4] or perhaps one that exchanges zone-boundary bosons.

The extremely short length scale of the pairs is extremely different from that of conventional superconducting pairs that extend over many lattice sites, and therefore allows a new type of question to be asked that would be meaningless for a system with large pairs: what is the location of the pair center relative to a specific lattice site, i.e. are the pairs centered on individual Cu atoms (Fig 4a), bonds (Fig 4b), or plaquette centers (Fig 4c)? This is important since we should intuitively expect a different type of pairing interaction to be applicable depending upon which of the centering options of Fig. 4 is realized, so this is addressing the most fundamental local nature of the pairing interaction. This
question also gets at the very heart of the nature of the carriers in the cuprates – for example, even though spectroscopy indicates that the doped carriers (holes) predominantly go onto oxygen sites \(^{17,18}\), strong theoretical arguments by Zhang and Rice indicate that the oxygen degrees of freedom can be projected out so that the effective carriers can be considered as Cu-centered \(^{19}\). If the carriers are indeed Cu-centered Zhang-Rice singlets then we expect the pairs to be bond-centered (Fig 4b), meaning they could go onto either horizontal or vertical bonds and among other things should lead to various nematic ordering tendencies similar to what has been observed in STM experiments \(^{20}\). On the other hand, if the carriers stay on the oxygen sites during the pairing \(^{21}\) then one of Figs 4a or 4c would be expected. Future experiments, for example scanning Josephson tunneling microscopy experiments \(^{22}\), might be able to resolve which of the centering options of Fig 4 the pairs choose.

Further, the packing of starfish-shaped pairs onto the lattice would naturally lead to weak variations in charge density and pair density, potentially giving rise to charge density waves \(^{22}\) and pair density waves \(^{22}\) in the cuprates - a clear example of an intertwined order \(^{23}\) between the superconducting and charge channels.

Because this pairing short length is so constant across the range of doping levels, it is plausible to consider it one of the defining characteristics of the pairing in the cuprates, perhaps more so than the strength of the pairing (superconducting gap) itself. Indeed, the strongly increasing pairing strength with underdoping, known for almost two decades \(^{24}\), should perhaps no longer be considered to be a signal of an intrinsically stronger interaction in the underdoped regime, but rather as a means of fixing the antinodal pair size at this lattice-based value. A positive feedback on the pairing interactions, proposed recently \(^{4}\), is one such mechanism that would allow or maybe force the pairing strength to grow until it sets the pair size at the optimal value, i.e. it enforces a type of speed-limit on the pairing enforced by the lattice parameters. Regardless of the mechanism of the pairing, our reconstruction of the shape and size of the pairs, including especially the universal antinodal length scale, will bring us
much closer to understanding the various cuprate phenomenology as well as the microscopic interactions that drive the pairing in these materials.

**Methods**

Pb-doping renders BSCCO free of the superstructure effects which contaminate most ARPES data on BSCCO cuprates. Thus we achieve data with very high resolution and low background, which is essential for self-energy analysis but is otherwise fully consistent with the large body of data taken by many groups over the years. ARPES experiments were performed at SSRL and HiSOR. In our measurements we selected incident photon energies to separate the response of the bilayer-split antibonding band from that of the bonding band; results from former are shown here and those from the latter (which are qualitatively similar) are shown in supplementary information Fig. S3 and S4, and discussed in the supplemental information S4.

We have performed high resolution ARPES on a series of superstructure-free bilayer (Bi,Pb)$_2$Sr$_2$CaCu$_2$O$_8$ (Bi2212) samples in both the overdoped and underdoped regimes and with $T_C$’s ranging from 69 to 91K, and with temperatures from 10K to 250K, though for the present work we focus on the low temperature superconducting-state spectra. A key enabler of the present work was the recent development of a new two-dimensional fitting procedure for the ARPES data [4] that no longer must focus on the one-dimensional MDCs (Momentum Distribution Curve) or EDCs (Energy Distribution Curves), but instead treats these simultaneously and on an equal footing. Our analysis assumes (successfully so far) that spectral broadening and band renormalizations (the effects of electronic correlations) can be treated within the conventional many-body language of Green’s functions and electronic self energies, though we do not restrict the system to have self-energies that are small compared to the energy of the states, i.e. there is no requirement that we reside in the quasiparticle or Fermi-liquid frameworks. With this we are able (successfully so far) to describe all
modifications to the spectra due to the onset of superconductivity within the framework of the conventional Nambu-Gor'kov formalism. Another new aspect of this procedure is that the real and imaginary parts of the extracted electronic self-energies $\Sigma'(k,\omega,T)$ and $\Sigma''(k,\omega,T)$ are intrinsically constrained to be causal (obey the Kramers-Kronig relations), as opposed to previous efforts in which the Kramers-Kronig relations were taken in one direction (e.g. taken from $\Sigma''$ to obtain information about $\Sigma'$, or vice-versa). This treatment allows us to extract, naturally and simultaneously, the superconducting gaps, the coherence factors for particle-hole mixing (which determine the Bogoliubov quasiparticles), and both real and imaginary parts, $\Sigma'$ and $\Sigma''$, of the complex self-energy. All of these quantities were obtained at many momentum points throughout the Brillouin zone and for many temperatures across the pairing and superconducting regimes.

Acknowledgements

We thank Bruce Normand and Rahul Nandkishore for very useful conversations, Mark Golden for putting us in touch with Y. Huang for sample growth, and D. H. Lu and M. Hashimoto for assistance at SSRL. Funding for this research was provided by DOE Grant No. DE-FG02-03ER46066 and by the Center for Experiments on Quantum Materials. SSRL is operated by the Office of Basic Energy Sciences of the US Department of Energy and HiSOR is supported by the Japanese Ministry for Science. Work at Brookhaven was supported by the Office of Basic Energy Sciences (BES), Division of Materials Science and Engineering, U.S. Department of Energy (DOE), through Contract No. DE-sc0012704. R. D. Z. and J. S. were supported by the Center for Emergent Superconductivity, an Energy Frontier Research Center funded by BES.

Author contributions

The project was conceived and directed by DD. The crystals were grown by HB, YH, RDZ, JS, and
GDG. ARPES measurements were performed by HL and XZ with assistance from SP. The theoretical framework was provided by HL, GA, and DD. Data refinement and 2D fitting were performed by HL.

The text was written by HL and DD with comments from all authors.

**Competing interests:** The authors declare no competing interests.

Correspondence and requests for materials should be addressed to D.S.D. (dessau@colorado.edu)
Figure 1 | The angular dependence of the Cooper pair length scales as extracted from ARPES data. **a.** Pair size (e-e separation expectation value using center-of-mass (CM) coordinates) as a function of direction measured from a Tc=85K under doped Bi$_2$Sr$_2$CaCu$_2$O$_8$ sample (UD85K) at 15K. The pairs extend to >50Å near the nodal direction, with drastically decreasing size when moving towards the antinode. **b.** A zoom in to the central region of a pair. An ultrashort length scale of 4.5Å is seen for the antinodal pairs – a length on the order of the lattice parameter. **c.** The ultrashort antinodal length scale is maintained for all doping levels studied indicating this is likely a universal property of the Cooper pairs. Raw ARPES data and relevant parameters throughout the Brillouin zone are obtained from ref [4]. See text and supplemental information for details of how the pair size is obtained as well as the impact of the bilayer splitting (the present data is only from the antibonding or AB band). The error bars of the pair size are based on the uncertainty of the parameters of the calculation (shown in Fig. 3).
Figure 2 | Antinodal spectra and fits over a wide doping range at temperatures well below $T_c$. a-g. Experimental spectra, with photon energy chosen so as to emphasize the antibonding band. h-n. The corresponding 2D fitting results plotted together with the extracted bare bands (black dashed curves). Intensity scales between doping levels are constrained by the theoretical fits and so are the measured spectra. o. Fermi surface plot including the location of the cuts of panels a-g as shown by the red line. p-q. $k_F$ EDCs and fits taken from the top panels. UD79 means under-doped sample with $T_c$ of 79K, OP91 means optimal doped sample with $T_c$ of 91K, OD85 means over-doped sample with $T_c$ of 85K, etc.
Figure 3 | The key parameters vs doping extracted from the 2D fitting to the antinodal spectra. a-b. Two parameters that we used to construct the bare band for the antinodal cut— the bare band bottom (panel a), and the momentum at the Fermi level $k_F$ (panel b). Both of these parameters display an expected decreasing trend as moving towards the over doped side. c. the superconducting gap size. It follows a linear trend with the doping, which is consistent with the previous findings in the literature [25]. d. Renormalization factor $Z$, extracted from the real part of the self-energy at $E_F$ (see supplementary S5 and Fig. S5 for detailed discussion.). The error bars in panel a and panel c include the uncertainty of the $E_F$ ($\pm 0.5$ meV), and the $3\sigma$ return from fitting. The error bars in panel b contain the momentum resolution ($\pm 0.006$ 1/Å), and the $3\sigma$ return from fitting. The error bars in panel d reflect the standard deviation of the $Z$ value extracted from different energy range near the Fermi level (within the superconducting gap energy scale).
**Figure 4** | Various centering options for the pairs as drawn using CM coordinates. Because the pair length scales are so short it is reasonable to expect them to be centered at specific locations within the Cu–O lattice, with the three main options shown here: a. A Cu-centered pair. b. A horizontal bond-centered pair. Vertical-bond centered pairs should be energetically degenerate with this configuration. c. A plaquette-centered pair. Cu atoms are shown as brown $d_{x^2-y^2}$ orbitals, oxygen as beige p orbitals.

**References**

1. Cooper, L. N. Bound Electron Pairs in a Degenerate Fermi Gas. *Phys. Rev.* **104**, 1189-1190 (1956).
2. Bardeen, J., Cooper, L. N., and Schrieffer, J. R. Theory of Superconductivity. *Phys. Rev.* **108**, 1175-1204 (1957).
3. Bednorz, J. G., and Müller, K. A. Possible high $T_c$ superconductivity in the Ba–La–Cu–O system. *Z. Phys. B.* **64**, 189-193 (1986).
4. Li, H. *et al.*, Coherent organization of electronic correlations as a mechanism to enhance and stabilize high-$T_c$ cuprate superconductivity, *Nat. Commun.* **9**, 26 (2018).
5. Y.J. Uemura Superfluid density of high-Tc cuprate systems: implication on condensation mechanisms, heterogeneity and phase diagram. Solid State Communications 126, 23 (2003).
6. Benfatto, L., Toschi, A., Caprara, S., and Castellani, C., Coherence length in superconductors from weak to strong coupling. *Phys. Rev. B* **66**, 054515 (2002).
7. Casas, M. *et al.*, BCS-Bose theory of the superconducting coherence length. *Sol. State. Commun.* **94**, 419-423 (1995).
8. Schrieffer, J. R., *Theory of Superconductivity*, Addison-Wesley, New York, (1964).
9. Kleinert, H., Collective quantum fields, *Fortschritte der Physik* **26**, 565-671 (1972).
10. Pippard, A. B., An experimental and theoretical study of the relation between magnetic field and current in a superconductor, *Proc. R. Soc. London, Ser. A* **216**, 547 (1953).
11. Tinkham, M., *Introduction to Superconductivity: Second Edition*, Dover Books on Physics (2004)
12. Wang, Y. et al., Dependence of upper critical field and pairing strength on doping in cuprates. *Science* **299**, 86-89 (2003).

13. J. P. Hinton et al. The rate of quasiparticle recombination probes the onset of coherence in cuprate superconductors. *Sci. rep.* **6**, 23610 (2016).

14. Chuang, Y.-D. et al., Doubling of the bands in overdoped Bi$_2$Sr$_2$CaCu$_2$O$_{8+δ}$ - evidence for c-axis bilayer coupling. *Phys. Rev. Lett.* **87**, 117002 (2001).

15. Feng, D.L. et al., Bilayer Splitting in the Electronic Structure of Heavily Overdoped Bi$_2$Sr$_2$CaCu$_2$O$_8$+d. *Phys. Rev. Lett.* **86**, 5550 (2001).

16. Feng, D. L. et al., Signature of Superfluid Density in the Single-Particle Excitation Spectrum of Bi$_2$Sr$_2$CaCu$_2$O$_8$+δ. *Science* **289**, 277-281 (2000).

17. Nücker, N., Fink, J., Fuggle, J. C., Durham, P. J., and Temmerman W. M., Evidence for holes on oxygen sites in the high-Tc superconductors La$_{2-x}$Sr$_x$CuO$_4$ and YBa$_2$Cu$_3$O$_7$, *Phys. Rev. B* **37**, 5158 (1988).

18. Chen, C. T. et al., Electronic states in La$_{2-x}$Sr$_x$CuO$_4$ probed by soft-x-ray absorption, *Phys. Rev. Lett.* **66**, 104 (1991).

19. Zhang, F.C. & Rice, T.M. Effective Hamiltonian for the superconducting Cu oxides, *Phys. Rev. B* **37**, 3759 (1988).

20. Lawler, M. J., et al., Intra-unit-cell electronic nematicity of the high-Tc copper-oxide pseudogap states, *Nature* **466**, 347-351 (2010).

21. Hirsch, J. E. and Marsiglio, F. Superconducting state in an oxygen hole metal. *Phys. Rev. B* **39**, 11515 (1989).

22. Hamidian, M. H. et al., Detection of a Cooper-pair density wave in Bi$_2$Sr$_2$CaCu$_2$O$_{8+δ}$. *Nature* **532**, 343 (2016).

23. Fradkin, E., Kivelson, S. A., and Tranquada, J. M., Colloquium: Theory of intertwined orders in high temperature superconductors. *Rev. Mod. Phys.* **87**, 457 (2015).

24. Ding, H. et al., ARPES study of the superconducting gap and pseudogap in Bi$_2$Sr$_2$CaCu$_2$O$_{8+δ}$ *J. Phys. Chem Solids* **59**, 1888–1891, (1998).

25. Hashimoto, M., et al. Energy gaps in high-transition-temperature cuprate superconductors. *Nat. Phys.* **10**, 483–495 (2014).
Supplementary Information

S1. Weighted average of Pippard coherence length

For a BCS superconductor, the coherence length is given by

\[ \xi_{\text{CL}}(\theta) = \frac{hv_F}{\pi \Delta} \]  \hspace{1cm} (S1)

In the presence of strong correlations, the Fermi velocity \( v_F \) is given by \( v_F = \frac{v_{BB}}{Z} \) with \( v_{BB} \) the bare band velocity and \( Z \) the renormalization factor. Thus, \( v_F \) is an effective quantity whose strong suppression is contained in the renormalization parameter \( Z \) which may also be considered as an enhancement of the single-particle mass. Fig. S1 shows the coherence length (panel a) of sample UD85K calculated according to Eqn. S1, using the bare bands \( \varepsilon_{0k}, \Delta, \Sigma'' \), and \( Z \) that are extracted from the 2D fitting. To get a weighting for the coherence length at different Fermi surface angles, we consider the condensate density of Cooper pairs \( N_c = \Delta(k)^2/E_k^2 \) \([1,2]\), where \( \Delta \) is the superconducting gap, \( E_k = \sqrt{\varepsilon_k^2 + \Delta(k)^2} \) with \( \varepsilon_k \) is the non-gapped dispersion (with renormalization). We can calculate the average coherence length weighted by the condensate density to be 13Å for the UD85 sample and 17Å for the OD74 sample. We plot these values in Fig. S1b (red points) together with values from the literature extracted from Hc2 measurements from reference \([3]\) (blue dots). Converting the Hc2 data to coherence length using the formula \( \xi_{\text{CL}} = \frac{\Phi_0}{2\pi H_{c2}} \) gives the dashed red line of Fig. S1b, which within the error bar (up to a factor of two in the Hc2 measurements) compares reasonably well to the ARPES-extracted coherence lengths.
S2. The functional form of pair size

The general expression of the characteristic size of the Cooper pairs is:

$$\xi_{\text{pair}} = \sqrt{\int dr |\psi(r)|^2} = \sqrt{\int dk |\partial_k \psi(k)|^2}$$

(S2)

where $\psi(r)$ is the Cooper-pair wave function in real space, $\partial_k$ is the gradient operator in momentum space ($k$ space), and $\psi(k)$ is the Fourier transform of $\psi(r)$ to $k$ space [4]. We utilize the standard result $\psi(k) = \Delta(k)/E_k$, with $E_k = \sqrt{\varepsilon_k^2 + \Delta(k)^2}$ [4], where $\varepsilon_k$ is the non-gapped dispersion that includes the renormalization effect due to the many-body interactions. $\varepsilon_{0k}$ is the bare band dispersion. For one ARPES cut (an antinodal cut as shown in Fig. 2), we can approximate the bare band as a parabola, i.e., $\varepsilon_{0k} = -\frac{E_{BB}}{k_F^2} \omega^2 + E_{BB}$ with parameters of the Fermi momentum $k_F$ and bare band bottom $E_{BB}$, where the $k_F$ and $E_{BB}$ of the antinodal cut of various doping are displayed in Fig. 3. We will evaluate Eqn. S2 for every angle $\theta$, discussing the denominator and numerator separately. The denominator of the pair size as a function of angle is:

$$D_\theta = \sqrt{\int dk \frac{\Delta^2}{(\varepsilon_{0k}/Z_\theta)^2 + \Delta^2}}$$

(S3)

where the $k$ integral is perpendicular to the Fermi surface at an angle $\theta$, the integral range is within the first Brillouin zone, and $\varepsilon_{0k}$, $\Delta_\theta$, and $Z_\theta$ are the corresponding bare band, superconducting gap and renormalization factor for the ARPES cut at certain Fermi surface angles (the antinodal cut is shown in Fig. 2), and $E_k = \sqrt{\varepsilon_k^2 + \Delta^2} = \sqrt{(\varepsilon_{0k}/Z_\theta)^2 + \Delta^2}$. The numerator of the pair size functional form is:
\[ N_\theta = \frac{1}{Z^2} \sqrt{ \int d \mathbf{k}_\perp \frac{\Delta_k^2}{E_k^6} (\partial_k \varepsilon_{0k})^2 + \frac{\Delta_k^2}{E_k^6} (\partial_k \Delta_k)^2 - 2 \frac{\Delta_k^2}{E_k^6} (\partial_k \varepsilon_{0k})(\partial_k \Delta_k) } \]  

with the same integral condition with the numerator described above. Here \( \partial_k = \partial_\perp + \partial_\parallel \) where \( \partial_\perp \) is the derivative normal to the Fermi surface and \( \partial_\parallel \) is the derivative tangential to the Fermi surface. However, as the variation of \( \Delta(k) \) (within 40 meV) over the whole Brillouin zone is much smaller than the variation of \( \varepsilon_{0k} \) (within 0.5 eV), thus \( \partial_k \Delta \ll \partial_k \varepsilon_{0k} \) and the first term in the numerator dominates. On the other hand, \( \partial_\parallel \varepsilon_{0k} \ll \partial_\perp \varepsilon_{0k} \) so the dominant term (subscript D for dominant added) in the numerator can be approximated as

\[ N_{\theta} \approx N_{\theta D} = \frac{1}{Z^2} \sqrt{ \int d \mathbf{k}_\perp \frac{\Delta_\theta^2}{E_k^6} (\partial_\perp \varepsilon_{0k})^2 } \]  

To further justify the approximation of \( N_{\theta} \approx N_{\theta D} \), Fig. S2 shows a simulation result of the percentage variation of \( N_D \) to \( N \) as \( \frac{|N_{\theta} - N_{\theta D}|}{N_{\theta}} \times 100\% \) throughout the Brillouin zone. Even at the largest variation at the antinode, there’s only a 3% variation, which is much smaller than our systematic error for the pair size.

**S3. Cooper pair size and coherence length**

In the BCS limit (band bottom energy much larger than the superconducting gap size \( E_{BB} \ll \Delta \)), the Cooper pair size \( \xi_{\text{pair}} \) is quite similar to the coherence length \( \xi_{\text{CL}} \approx 1.11 \xi_{\text{CL}} \) as can be obtained by a direct calculation using Eqn. S5. However, in cuprate superconductors, when moving towards the antinode, \( \Delta \) increases, and the band bottom decreases, the ratio of pair size to coherence length deviates from the BCS limit. And in fact, at the antinodal region, samples of all dopings that we measured show a strong deviation from the BCS limit due to the shallow band bottom that is comparable or even smaller (OD69) than the superconducting gap.
S4. Bilayer splitting and Bonding and Antibonding bands.

Due to the fact that there are two CuO$_2$ planes per unit cell (a bilayer) with some coupling between them, the bands split into a lower or bonding band (B) and an upper or antibonding band (AB) \cite{5,6}. By utilizing a photon energy of 24 eV the AB band has a dominant matrix element (experimental intensity) and there is almost no contamination from the B band. This is the photon energy that was utilized for all data shown in the main text. Utilizing the photon energy of 20 eV both the B and AB bands are visible, as shown in Fig. S3c for the OD69 sample. The very clean separation of the two bands for this sample and this photon energy allows us to do a proper 2D fitting and extract all relevant parameters, as summarized by the green data points in Fig. S4. While a similar effect is observed for the other dopings, the distinction between the two bands is less clear and so an accurate fitting is presently more difficult. As expected the bare band bottom and $k_F$ are very different for the B band, while the extracted gap values are essentially identical. The renormalization parameter $Z$ is perhaps slightly smaller than for the AB band, though the difference is minimal. The B band pair size is however dramatically larger than the AB band pair size.

S5. Extraction of the parameter $Z$ from the slope of $\Sigma'$ vs $\omega$.

It is perhaps surprising that the renormalization parameter $Z$ has only a weak doping dependence (Fig 3d). The way this is extracted from our data is shown in Fig. S5. First $\Sigma''(\omega)$ and $\Sigma'(\omega)$ were extracted from the low temperature antinodal data from each sample utilizing the recently-developed 2-dimensional fitting technology \cite{7}. The value $Z$ is defined $Z = 1 + |\partial \Sigma' / \partial \omega|_{\omega \rightarrow 0}$, i.e. it is related to the slope of $\Sigma'$ vs omega at $E_F$. This slope is highlighted by the black dashed
line in Fig. S5a. It is observed that the low energy slope is approximately independent of doping for this low T antinodal data, explaining the approximately constant value of $Z$ shown in Fig 3d.

S6. Pair size, pair radius, and the Center of Momentum frame.

Equations S2 and S5 show that the pair size is calculated as the rms expectation value of the radius $r$. Utilizing the experimentally determined pair wavefunction, a direct calculation returns values that range from $r=4.5$ Å at the antinode. This value of $r$ is the inter-electron separation within a pair, i.e. it gives the separation of one of the electrons from the other electron, just as it would give the separation of an electron from the nucleus if a one-electron hydrogenic wavefunction was utilized in Eqn. S2. In contrast to a H atom with a heavy nucleus that has the center of momentum (CM) at the nucleus, the two electrons in the Cooper pair have equal mass so a reduced mass scheme with the CM midway between the two electrons is more intuitive for visualization. This is the scheme utilized for Fig 4, i.e. the expectation values of the pair are observed to go between $\pm r/2$. 
**Figure S1 | Coherence length of the sample UD85K at 15K.** a. Coherence length vs Fermi surface angle, calculated using Eqn. S1 with all the parameters extracted from 2D fitting method (purple square). The corresponding ARPES data and the fit results are presented in ref. [5]. The black curve is an interpolation of the data points. The error bars are based on the uncertainty of the parameters extracted from the 2D fitting. Using the interpolated curves shown in the figure, we can then calculate the properly averaged coherence lengths, which we compare to Hc2 measurements [3] in panel b.

**Figure S2 | Simulation result of the percentage variation of the approximate numerator terms to full term in the pair size functional form.** The maximum variation is only ~3%.
**Figure S3 | Bilayer split bands.** Antibonding=AB, Bonding=B. Antinodal ARPES from the OD69 sample taken with different photon energies so as to emphasize the AB band only (panel b) or both the AB and B bands (panel c). Panel a shows the experimental Fermi surfaces for both the AB and B bands. $h\nu=24$ eV was utilized to access the AB band while 20 eV was used to emphasize both of them.
Figure S4. Comparison of bonding band parameters (green triangle) vs. antibonding (black square). The bonding band pair sizes are much larger than the antibonding pairs.
Figure S5 | Self energy of the antinodal cut with different dopings at temperature well below $T_c$. The real part (panel a) and the imaginary part (panel b) of the self-energy extracted from 2D fitting results shown in Fig. 2. The slope of $\Sigma'$ near $E_F$ (black dashed line) is approximately independent of doping, and is the reason that $Z$ is roughly constant with doping.
References

1. Yang, C. N., Concept of off-diagonal long-range order and the quantum phases of liquid He and of superconductors. *Rev. Mod. Phys.* 34, 694 (1962).

2. Fukushima, N., Ohashi, Y., Taylor, E., and Griffin, A., Superfluid density and condensate fraction in the BCS-BEC crossover regime at finite temperatures. *Phys. Rev. A* 75, 033609 (2007).

3. Wang, Y. *et al.*, Dependence of upper critical field and pairing strength on doping in cuprates. *Science* 299, 86-89 (2003).

4. Schrieffer, J. R., *Theory of Superconductivity*, Addison-Wesley, New York, (1964).

5. Chuang, Y.-D. *et al.*, Doubling of the bands in overdoped Bi2Sr2CaCu2O8+δ - evidence for c-axis bilayer coupling *Phys. Rev. Lett.* 87, 117002 (2001).

6. Feng D.L. *et al.*, Bilayer Splitting in the Electronic Structure of Heavily Overdoped Bi2Sr2CaCu2O8+d. *Phys. Rev. Lett.* 86, 5550 (2001).

7. Li, H. *et al.*, Coherent organization of electronic correlations as a mechanism to enhance and stabilize high-$T_C$ cuprate superconductivity, *Nat. Commun.* 9, 26 (2018).