How does the high temperature superconductivity (HTSC) emerge with decreasing temperature? This highly debated question is crucial for understanding the mechanism of HTSC. The superconducting transition in HTSC appears to be unusual, with a seeming lack of changes in the quasiparticle (QP) density of states at $T_c$ and persistence of the normal state pseudogap above $T_c$ [1, 2]. However, there is no consensus about the $T$-evolution of the energy gap. Results obtained by different experimental techniques range from complete $T$-independence of the superconducting gap, $\Delta_{SG}$, in the whole $T$-range [3], to strong $T$-dependence at $T = T_c$ [4, 5, 6, 7, 8, 9]. In some experiments, coexistence of the two energy scales below $T_c$ was reported [2, 8, 9]. The existing controversy is one of the major obstacles for understanding HTSC and requires further $T$-dependent studies.

Intrinsic tunneling spectroscopy utilizes the weak interlayer ($c$-axis) coupling in layered HTSC. Unlike surface probe techniques, it is perfectly suited for $T$-dependent studies and provides information about bulk electronic properties of HTSC. In recent years different ways of improving this technique were employed [7, 8, 11, 12, 13, 14]. Yet, interpretation of intrinsic tunneling characteristics may not be free from misconceptions, associated both with the problem of self-heating [15] and the lack of complete understanding of the interlayer transport mechanism.

Here we study temperature, bias and doping dependencies of intrinsic tunneling in small $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ (Bi-2212) mesa structures. We observe an abrupt crossover in the shape of current-voltage characteristics (IVC’s), from thermal activation (TA) like at $T > T_c$, to $T$-independent quantum tunneling like at $T < T_c$. Our data indicates that the interlayer transport mechanism in Bi-2212 changes in underdoped Bi-2212, together with development of the pseudogap. The observed simple TA behavior in the whole normal state region puts strong constraints on the nature of the $c$-axis pseudogap. In the supplementary information [15] we discuss single QP characteristics, self-heating and clarify determination of parameters.

We study two batches of crystals: the Y-doped Bi(Y)-2212 with the maximum $T_c \approx 94.5$K and the pure Bi-2212 with the maximum $T_c \approx 86$K. Details of mesa fabrication and characterization are described elsewhere [7, 8, 12].

Figs. 1 a) and b) show the $T$-variation of $I - V$ and $dI/dV(V)$ curves for a Bi-2212 mesa, containing $N = 7$ intrinsic Josephson junctions, see Fig. 1 c). The following characteristic features are seen in $dI/dV$ curves: The sharp "coherence" peak appears at $T < T_c$ and is, in analogy with conventional low-$T_c$ junctions, attributed to the sum-gap voltage $V_g = 2N\Delta_{SG}/e$ [7, 8, 11, 14]. A broader hump is seen in the whole $T$-range and is attributed to the $c$-axis pseudogap [7, 8, 11]. At elevated $T$ the $dI/dV$ curves cross in nearly one point [3, 14]. The $T$-dependence of these features is plotted in Fig. 1 d).

Figure 2 a) shows $dI/dV(V)$ curves (in a semi-log scale) for the most underdoped Bi(Y)-2212 mesa with $N = 34$ junctions. The thin and thick lines in Fig. 2 a) represent $dI/dV(V)$ curves below and above $T_c$, respectively. It is seen that the shape of interlayer characteristics exhibits a remarkable qualitative change at $T_c$:

(i) Above $T_c$ the slope of $\ln[dI/dV(V)]$ curves decreases with increasing $T$, while the voltage scale remains almost the same. The curves lean towards the horizontal line $dI/dV = \sigma_N$, corresponding to the normal state conductance, and cross almost in one point.

(ii) Below $T_c$ the $\ln[dI/dV(V)]$ curves have the same $T$-independent slope. With increasing $T$ the curves remain parallel and move towards the vertical axis $V = 0$ (i.e., the voltage scale decreases).

Experimental curves from Fig. 2 a) have a characteristic V-shape with voltage-independent slope. The overall shape of the curves at $T > T_c$ closely resembles thermal activation characteristics [16], which can be described by a simple expression [15]:

$$
\frac{dI}{dV}(T, V) \propto \frac{n(T)}{T} \exp \left[ \frac{U_{TA}}{k_B T} \right] \cosh \left[ \frac{eV}{2k_B T} \right],
$$

where $n(T)$ is the density of states at $T$ and $U_{TA}$ is the thermal activation energy.
where \( n \) is the concentration of mobile charge carriers, and \( U_{TA} \) is the TA barrier. Indeed, the \( \cos \) term reproduces the rounded V-shape of \( \ln[|dI/dV|](V) \) curves with the slope that monotonously increases as \( 1/T \).

In Fig. S10 [15] we show that the data from Fig. 2 a) in the whole normal state, \( T > T_c \), is very well described by Eq. (1) with constant \( U_{TA} \simeq 24meV \) and \( T \)–linear \( n(T) \). It also reproduces the crossing point [10], which according to Eq. (1) occurs at \( eV \simeq 2U_{TA} \). Apparently, the TA barrier should be associated with the \( c \)–axis pseudogap. Thus, the appearance of the crossing point is simply the consequence of a fairly \( T \)–independent \( c \)–axis pseudogap [1, 3, 7, 8], while the crossing voltage represents the pseudogap energy. Indeed, a correlation between the pseudogap energy \([1, 3, 7, 8]\) and the crossing voltage is seen from Fig. 1 d).

According to Eq. (1) the slope of the curves in Fig. 2 a) should be proportional to the reciprocal temperature: \( d/dV(\ln[|dI/dV|]) = (e/2k_BT) \tanh(eV/2k_BT) \simeq e/2k_BT \). Fig. 2 b) shows the effective temperature, \( T_{eff} \), obtained from the slopes at \( V/N = 30mV \) [15]. It is seen that in the normal state \( T_{eff} \simeq T \), confirming the TA nature of interlayer transport at \( T > T_c \). However, at \( T < T_c \) an abrupt saturation of \( T_{eff}(T) \) occurs, typical for quantum tunneling transport [11, 12, 13]. This is the central observation of this work. Remarkably, the crossover is clearly distinguishable at \( T \rightarrow T_c \), although no other spectroscopic features can be resolved in \( dI/dV(V) \) curves.

So far we discussed \( dI/dV(V) \) at finite bias. Another straightforward way to investigate the TA behavior is to consider \( T \)–dependence of zero-bias resistance, \( R_0 \), (the \( \exp \) term in Eq. (1)). Experiments on small mesas provide a possibility to measure the small-bias QP resistance, \( R^{QP}_{QP} \), in the superconducting state \([11, 12, 13]\), which is otherwise shunted by the supercurrent [20]. This is facilitated by the large specific capacitance (high quality factor) of intrinsic junctions [18], which causes the pronounced hysteresis in the IVC’s and allows junctions to remain in the resistive state even below the critical current, see Fig. 1 c). Therefore, bias yield an additional parameter for our studies, which may render crucial for correct interpretation of the data.

Fig. 3 a) shows the dc– resistance \( R_{dc} = V/I \) at different \( I \) for the same Bi-2212 mesa as in Fig. 1. Measurements were done by first applying a large current, sufficient for switching to the last QP branch, see Fig. 1 c), and then ramping it back to the desired value. The \( R_{dc} \) at the smallest current \( I = 5\mu A \) coincides with the conventional ac– resistance, \( R_{ac} \). It drops at \( T < T_c \) since the finite quality factor of the junctions causes retrapping to the superconducting branch at a finite current [18]. From Fig. 3a) it is seen how non-linearity develops with decreasing both bias and \( T \). At small \( I \) the curves approach an asymptotic, allowing a confident estimate of the zero-bias quasiparticle resistance, \( R^{QP}_0 \), without a need for extrapolation.
The $R_{0}^{QP}$ grows with decreasing $T$ at $T < T_c$, and is usually described either in terms of power-law dependence [20], inherent for single QP tunneling in the presence of a $d$-wave gap (see Fig. S7 [15]), or in terms of TA [21]. The double-logarithmic plot, Fig. 3 b), demonstrates that $R_0(T)$ is not described by the power law in any extended $T$-range for our mesas. Neither is it perfectly described by the Arrhenius law $exp(U_{TA}/k_B T)$, as demonstrated in Fig. 3 c). On the other hand, Fig. 3 d) demonstrates that the ratio $R_0/T$ follows very accurately the Arrhenius law (dashed line) in the whole normal region, consistent with the TA expression Eq. 11 and with the finite bias behavior, Fig. 2 b). However, at $T < T_c$, $R_{0}^{QP}/T$ deviates downwards from the Arrhenius law. This is consistent with saturation of the effective $T_{eff}(T < T_c)$, as shown in Fig. 2 b). Therefore, it is a consequence of the same crossover.

In Fig. 4 we analyze doping dependence of $R_{0}(T)/T$. In the normal state, $R_0(T)$ varies in the TA manner for all doping levels. In the superconducting state, $R_0^{QP}$ also continues to grow with decreasing $T$, but the rate of the growth with respect to the TA behavior, $R_{TA}$, (dashed lines) strongly depends on doping:

For the overdoped (OD) mesa, $R_{0}^{QP}(T < T_c)$ grows much faster than in the normal state; i.e., there is a sharp onset of the excess resistance at $T = T_c$. However, the excess resistance at $T < T_c$ decreases rapidly with underdoping. It becomes small in the near optimally doped (OP) mesa and turns negative already in moderately underdoped (UD) mesas.

Inset in Fig. 4 shows the magnitude of the excess QP resistance with respect to the TA fit, $R_{TA}$, in the normal state (dashed lines). The quality of the fit could be judged from the flatness of the normal state region. A progressive decrease of the excess QP resistance at $T < T_c$ with decreasing doping is obvious.

We start discussion of the observed phenomena by recollecting the expected $T$ variation for conventional single $QP$ tunneling characteristics. As discussed in sec. I of the supplementary information, provided the physical requirement that $\Delta_{SC}$ and the quasiparticle lifetime increase with decreasing $T$, all single QP tunneling characteristics (including $d$-wave) exhibit the following universal $T$-dependence:

A) Opening of the superconducting gap at $T < T_c$ leads to appearance of the large excess QP resistance due to rapid freezing out of QP’s upon their condensation into Cooper pairs (see inset in Fig. 3 B).

B) $R_{0}^{QP}$ continues to grow with decreasing $T$ both due to freezing-out of thermal QP’s and due to growth of the $QP$ lifetime, which decreases the number of available subgap $QP$ states. All this leads to a progressive growth of the slope of $dI/dV(V)$ characteristics with decreasing $T$.

The sudden appearance of excess resistance at $T < T_c$ in OD mesas indicates that the superconducting transition here is conventional, in a sense that it is accompanied...
by opening of the superconducting gap, and is also consistent with single QP tunneling mechanism of interlayer transport.

However, the superconducting transition in under-doped mesas is clearly abnormal. We emphasize again that appearance of the negative excess resistance and the crossover to $T$-independent slope reflect the same phenomenon at zero and finite bias, respectively. Therefore, any explanation of the observed unusual $T$-dependence must account for both zero and finite bias behavior.

At the first glance, the lack of excess resistance at $T < T_c$ in UD mesas could be consistent with the precurser superconductivity scenario of the pseudogap [3], according to which the gap does not open and QP's do not start to pair at $T_c$, but at much higher $T^*$. Then one could argue that the lack of excess resistance reflects simply the lack of dramatic changes in the QP spectrum upon establishing of the global phase coherence at $T_c$. However, the same argument would make it difficult to explain the abrupt crossover to $T$-independent slope. Furthermore, as we noted above, the slope of $dI/dV(V)$ should continue to increase with decreasing $T$ for single QP tunneling. Therefore, the observed crossover to constant slope can hardly be explained in terms of single QP tunneling [15].

This conclusion brings us to one possible interpretation of the observed phenomena. If the $c$-axis transport in UD mesas is not solely due to single QP's, it must also involve pairs. The corresponding multi-particle tunneling [22, 23] and the multiple Andreev reflection [24, 25] processes are well studied for conventional low-$T_c$ junctions. Both processes are similar [22] and are almost $T$-independent, because they do not rely on the presence of thermally excited QP's, but directly involve Cooper pairs from the Fermi level. The multi-particle current occurs via elastic conversion (dissociation or recombination) of Cooper pairs into QP's; while the multiple Andreev reflection process is due to transmutation of a quasielectron into a quasihole with creation of a Cooper pair and can be universally described in terms of inelastic tunneling in the presence of time-dependent phase difference due to the ac-Josephson effect [25].

Importantly, the multi-particle processes result in almost $T$-independent and $V$-exponential subgap current [22, 23], which essentially coincides with the description of the observed crossover, see Fig. 2. Furthermore, the abruptness of the crossover at $T_c$, points towards the coherent multiple Andreev reflection mechanism of the interlayer current, because it requires the time-periodic ac-Josephson effect [27] and, consequently, abruptly disappears simultaneously with phase coherence at $T_c$.

We may further speculate why the multiparticle processes become progressively more important with underdoping. The multiparticle current decrease much faster than the single QP current with decreasing the interface transparency [22, 25]. Therefore, the single QP dominates over multiparticle current, unless there are micro-shots in the tunnel barrier [25]. The required micro-shots in underdoped Bi-2212 could arise from doping-dependent nano-scale inhomogeneity, observed at the surface of Bi-2212 [26, 27], which in this case should persists also in the bulk of Bi-2212 crystals.

Such interpretation implies that not only the electronic structure, but also the $c$-axis transport mechanism changes with underdoping: from coherent and directional [15] single QP tunneling in overdoped, to progressively increasing pair contribution in underdoped Bi-2212. The latter is apparent only in the phase coherent state at $T < T_c$, is almost $T$-independent, and is consistent with multiple Andreev reflection mechanism of the interlayer transport.

Finally we note that the reported amazingly trivial thermal activation behavior in the whole normal state region, puts strong constrains on the nature of the $c$-axis pseudogap. Even though, the similar TA-behavior can be ascribed to several processes (e.g., inelastic tunneling via the impurity [27, 28], or elastic tunneling via a resonant state in the tunnel barrier, or Coulomb blocking of tunneling [3]), it does not involve any gap, nor angular dependence in the QP spectrum, but assumes instead that there is some constant blocking barrier for interlayer hoping [10]. This may indicate that the “large” $c$-axis pseudogap is not a pairing gap, which would naturally explain its indifference to such classical depairing factors as temperature and magnetic field.

We are grateful to E.Silva for stimulating discussions and to A.Yurgens, T.Benseman and G.Balakrishnan for providing Bi(Y)-2212 and Bi-2212 single crystals. Financial support from the K.&A. Wallenberg foundation and the Swedish Research Council is gratefully acknowledged.

I. SUPPLEMENTARY INFORMATION

A. Nomenclature

ARPES: angular resolved photoemission spectroscopy, Bi-2212: Bi$_2$Sr$_2$CaCu$_2$O$_{8+δ}$, Bi(Y)-2212: Y-doped Bi-2212, DoS: density of states, $\Delta$: energy gap, $\Delta T$: self-heating, $\Gamma$: depairing factor, reciprocal quasiparticle lifetime; HTSC: high temperature superconductor, $I$: current, IJJ: intrinsic Josephson junction, IVC: Current-Voltage characteristic, $N$: number of IJJ's in the mesa, $n$: concentration of mobile charge carriers, OD, OP, UD: overdoped, optimally doped, underdoped, respectively; QP: quasiparticle, $R_0$: zero bias resistance (in the normal state), $R_0^{QP}$: zero bias quasiparticle resistance in the superconducting state, is different from $R_0$ because of the pronounced hysteresis in the IVC's at $T < T_c$. 

\[ R_n: \text{tunnel (normal) resistance,} \]
\[ R_{th}: \text{thermal resistance of a mesa,} \]
\[ \text{SIS: Superconductor-Insulator-Superconductor,} \]
\[ \text{SIN: Superconductor-Insulator-Normal metal,} \]
\[ \text{STM: scanning tunneling microscope,} \]
\[ \text{TA: thermal activation,} \]
\[ T_c: \text{superconducting critical temperature,} \]
\[ U_{TA}: \text{thermal activation barrier,} \]
\[ V: \text{voltage,} \]
\[ V_g: \text{the sum-gap voltage } V_g = 2N\Delta/e \]

The capital M/S in front of the figure or equation number indicates that the reference is made to the Manuscript or the Supplementary information, respectively.

B. Influence of the quasiparticle life time on \( dI/dV \) characteristics

We have argued in the manuscript that the observed crossover to \( T \)-independent slope and the correlated appearance of the negative excess resistance are inconsistent with the single QP tunneling mechanism of the \( c \)-axis transport. However, the interlayer tunneling in Bi-2212 may depend on a number of parameters \( 31, 33 \):

(i) The QP life time, \( 1/\Gamma \), which, according to ARPES data, has a substantial \( T \)-dependence close to \( T_c \), where it increases roughly linearly with \( T \): \( 9, 33 \).

(ii) Momentum conservation upon tunneling (coherent vs. incoherent tunneling);

(iii) The directionality of \( c \)-axis tunneling, i.e., the angular dependence of the tunneling matrix element \( t(\varphi_1, \varphi_2) \), caused by a non-spherical Fermi surface \( 31, 32 \).

(iv) Temperature and (v) angular dependence of the gap in the DoS and the points (i-iii) above. Deviations from pure \( d \)-wave symmetry were reported in recent ARPES experiments \( 9 \).

It may not be obvious that experimental data can not be described by the single QP tunneling with a fortunate combination of all those parameters. Therefore, below we want to dwell upon this statement.

We start with considering the consequences of points (ii, iii) and (v), i.e., coherence and directionality of tunneling and the symmetry of the order parameter. The single QP tunneling current is given by:

\[
I(V) = \int_0^{2\pi} \int_0^{2\pi} \int_{-\infty}^{+\infty} d\varphi_1 d\varphi_2 dE \cdot t(\varphi_1, \varphi_2)\rho(E, \varphi_1)\rho(E + eV, \varphi_2)f(E)\left[1 - f(E + eV)\right],
\]

where \( E \) is the energy of the QP with respect to the Fermi surface, \( \varphi_{1,2} \) are the angles in the momentum space of the initial and the final state of the QP, and \( \rho(E, \varphi_1) \) \( \rho(E + eV, \varphi_2) \) are the corresponding QP DoS:

\[
\rho(E, \varphi) = \Re[(E-i\Gamma)/\sqrt{(E-i\Gamma)^2 - \Delta(\varphi)^2}],
\]

In Fig. S5 a) and b) we show numerically simulated SIS characteristics for different single QP tunneling scenarios. We assumed, \( t(\varphi_1, \varphi_2) = \text{const for incoherent-nondirectional} \), \( t(\varphi_1, \varphi_2) \propto \delta(\varphi_1 - \varphi_2) \) for coherent-nondirectional and \( t(\varphi_1, \varphi_2) \propto [\cos(k_x) - \cos(k_y)]^2 \delta(\varphi_1 - \varphi_2) \) for coherent-directional tunneling \( 31 \).

The simulations were made for \( \Delta(\varphi = 0) = 35\text{meV} \), at low \( T \) and small depairing \( \Gamma \ll \Delta(0) \). Detailed discussion of \( dI/dV \) characteristics can be found in Ref. \( 30 \).

Irrespective of the scenario, the zero-bias resistance diverges (the conductance tends to zero) at \( T \to 0 \) and \( \Gamma = 0 \). The divergence is removed by the finite \( \Gamma \), see the inset in Fig. S6 b). Therefore, \( dI/dV(V = 0, T \to 0) \) provides the information about the value of \( \Gamma \). We also note that \( dI/dV(\Gamma) \) characteristics exhibit a sharp peak at the sum-gap voltage \( V_g = 2\Delta/e \) (except for the curve A).

It is expected that interlayer QP tunneling in Bi-2212 single crystals should be predominantly coherent (provided that the single crystal is pure enough so that there is no momentum scattering upon tunneling) and strongly directional, with dominating antinodal tunneling \( 31, 32 \). Indeed, the curve C in Fig. S5 resembles most closely the experimental characteristics (cf. with the curves at \( T = 13.4K \) from Fig. M1a and b), see also a discussion in Refs. \( 31, 34 \).

In what follows, we will analyze the incoherent-nondirectional single QP tunneling characteristics, representing an "extreme" \( d \)-wave case with the weakest sum-gap peak. The conclusions will, however, be universal for any single QP tunneling scenario.

Now we consider the effect of \( T \) (point iv) and \( \Gamma \) (point i) on \( dI/dV \) characteristics. It is clear that both parameters will smoothen \( dI/dV(V) \) characteristics and fill-in the dip in conductance at \( V = 0 \). However, they do this in a slightly different manner. This is illustrated in Figs. S6 and S7.

The increase of \( \Gamma \) leads to smearing of the QP gap singularity and simultaneous filling-in of the sub-gap states in DoS, Eq.(S2). This naturally leads to smearing of the sum-gap peak and to filling-in of the zero-bias dip in \( dI/dV \), as shown in the inset in Fig. S6 b).

The increase of \( T \) leads to the decrease of \( \Delta \). Both factors result in the increased number of excited QP’s above the gap, which leads to rapid filling-in of the zero bias dip in conductance. However, the increase of \( T \) (unlike \( \Gamma \)) keeps the gap singularity in DoS unchanged. From Fig. S7 it is seen that this leads not only to filling-in of the dip, but also to development of the maximum at \( V = 0 \), representing the so called zero-bias logarithmic singularity \( 35 \).

The origin of the singularity is very simple: at elevated \( T \) there is a substantial amount of thermally excited QP’s just above the gap. At \( V = 0 \) the partly filled gap singularities in the two electrodes are co-aligned, causing a large current flow from one electrode to another, which is exactly compensated by the counterflow from the second electrode. However, exact cancelation is lifted at an arbi-
FIG. 5: (Color online). Simulated single QP characteristics for different tunneling scenarios. A: incoherent, non-directional, $d$–wave; B: coherent, non-directional, $d$–wave; C: coherent, directional, $d$–wave; D: $s$–wave. Note that the sharp sum-gap peak and the half-gap singularity in $dI/dV$ are inherent to most scenarios.

An arbitrary small voltage across the junction, leading to a sharp maximum in $dI/dV$. From Fig. S7 it is seen that the zero-bias maximum in $d$–wave junctions is pronounced even in the extreme case of weakest sum-gap peak. We have checked that the zero-bias logarithmic singularity is inherent for single QP characteristics, irrespective of the tunneling scenario.

Interestingly, the experimental characteristics do not exhibit the zero-bias logarithmic singularity, see Fig. M2a). Within the single QP tunneling scenario this is only possible if the depairing factor increases with $T$ at such a rate that it smears out the gap singularity in DoS and thus suppresses development of the singularity. We estimate that the corresponding $\Gamma$ at $T = T_c$ should be in the range 2-5 meV, somewhat smaller than deduced from ARPES [9, 33].

Now we can substantiate our statement. In Fig. S8 we show the attempt of maintaining the same slope of $\ln dI/dV(V)$ curves for progressively decreasing $\Delta$. Since increase of both $\Gamma$ and $T$ leads to filling-in of the zero bias dip, i.e., to decrease of the slope, the slope can be maintained only if $\Gamma$ and $T$ move in the opposite directions.

It is impossible to maintain the slope under the physical requirements that $\Gamma$ increases and $\Delta$ decrease simultaneously with increasing $T$ because all of the three parameters work in the same direction - decrease the slope. The inset in Fig. S7 shows the calculated zero bias resistance vs. $1/T$ for incoherent, non-directional, $d$–wave single QP tunneling and different $\Gamma$. It is seen that a prominent excess resistance appears even in this extreme $d$–wave case and even for very large $\Gamma$. The dashed line shows that the excess resistance grows approximately as $T^{-2}$ in this case. The abrupt appearance of the excess QP resistance is a mere reflection of the abrupt opening of the superconducting gap at $T_c$. Therefore, the observed $T$–independent slope and the absence of excess resistance, observed for UD mesas, are inconsistent with the single QP tunneling and points towards the doping-dependent change in the interlayer transport mechanism.

FIG. 6: (color online). Effect of the depairing factor on single QP characteristics. Shown are simulated characteristics at constant $T$ and $\Delta$ for varying $\Gamma$. Increase of $\Gamma$ tends to reduce the slope of $dI/dV$ characteristics. Inset in panel b) indicates that filling-in of the zero bias dip with $\Gamma$ occurs in a power law manner.
C. Analysis of self-heating and non-equilibrium phenomena

Self-heating can distort IVC’s of Josephson junctions at large bias. The temperature rise due to self-heating is given by a simple expression

$$\Delta T = PR_{th}(T),$$

where $P = IV$ is the dissipated power and $R_{th}$ is the effective thermal resistance of the junctions, which is $T$-dependent and, therefore, bias dependent $[36]$. The influence of self-heating on IVC’s of our mesa structures was thoroughly studied in Refs. $[34, 37, 38]$ (see also a discussion of scaling of IVC’s in the inset of Fig. 1 and Fig. 3 from Ref. $[7]$).

Despite relative simplicity of the phenomenon, discussion of self-heating in intrinsic tunneling spectroscopy has caused a considerable confusion, a large part of which has been caused by a series of publications by V. Zavaritsky $[39]$, in which he “explained” the non-linearity of intrinsic tunneling characteristics by assuming that there is no intrinsic tunneling. The irrelevance of this model to our subject was discussed in Ref. $[38]$.

A certain confusion might be also caused by a large spread in $R_{th}$, reported by different groups $[14, 36, 40, 41, 42]$. For the sake of clarity it should be emphasized that those measurements were made on samples of different geometries. It is clear that $R_{th}$ depends strongly on the geometry $[36, 37]$, e.g., $R_{th}$ can be much larger in suspended junctions with poor thermal link to the substrate $[42]$ than in the case when both top and bottom surfaces of the junctions are well thermally anchored to the heat bath $[14]$. For mesa structures similar to those used in this study (a few $\mu m$ in-plane size, containing $N \simeq 10$ IJJ’s), there is a consensus that $R_{th}(4.2 K) \sim 30 – 70 K/mW$ (depending on bias) $[36, 41]$ and $R_{th}(90 K) \sim 5 – 10 K/mW$ $[36]$. Larger values $R_{th} > 100 K/mW$ claimed by some authors $[10]$ are unrealistic for our mesas because they can withstand dissipated powers in excess of $10 mW$ without being melted.

Yet, we note that talking about a typical value of $R_{th}$ is equally senseless as talking about a typical value of a contact (Maxwell) electrical resistance: both depend on the geometry. Therefore, reduction of mesa sizes provides a simple way for reduction of self-heating $[37]$. Consequently, variation of $dI/dV$ characteristics with the junction size and geometry provides an unambiguous way of discriminating artifacts of self-heating from the spectroscopic features $[36]$.

How self-heating can distort the IVC’s of Josephson junctions is obvious: since self-heating rises the effective $T$ it may affect the IVC only via $T$-dependent parameters. There are three such parameters: the quasiparticle resistance, the superconducting switching, and the superconducting gap. They will affect the IVC’s of Bi-2212 mesas, containing several stacked Josephson junctions, in the following manner:

The consecutive increase of $T$ upon sequential switching of IJJ’s from the superconducting to the resistive state will distort the periodicity of quasiparticle branches. Each consecutive QP branch will have a smaller QP resistance (smaller $V$ at given $I$) and smaller switching current, see the discussion in Refs. $[7, 38]$. This type of distortion of the QP branches becomes clearly visible (at base $T = 4.2 K$) when $\Delta T \gtrsim 20 K$ $[37, 38]$.

The $T$-dependence of $\Delta$ may lead to appearance of
back-bending of the IVC at the sum-gap knee. Fig. S9 reproduces the results of numerical simulation of such the distortion, made specifically for the case of Bi-2212 mesa with the corresponding \( T \)-dependent parameters (see Ref. [34] for details). Fig. S9 a) shows a set of simulated undistorted IVC’s at different \( T \) for coherent, directional, \( d \)-wave tunneling. b) distorted IVC’s at the same base temperatures; c) The mesa temperature as a function of bias. d) Temperature dependence of the genuine superconducting gap (solid line) and the "measured" gap obtained from distorted IVC’s (dashed curve). Note that even strong self-heating (\( T \) reaches \( T_c/2 \) at \( V_g \) at 4.2K) does not cause considerable distortion of the measured gap. Data from Ref. [34]

FIG. 9: (Color online). Simulated distortion of SIS tunneling characteristics by strong self-heating. a) Undistorted IVC’s at different \( T \). Simulations were made for typical parameters of our mesas, \( T \)-dependent thermal conductivity of Bi-2212, and for coherent, directional, \( d \)-wave tunneling. b) distorted IVC’s at the same base temperatures; c) The mesa temperature as a function of bias. d) Temperature dependence of the genuine superconducting gap (solid line) and the "measured" gap obtained from distorted IVC’s (dashed curve). Note that even strong self-heating (\( T \) reaches \( T_c/2 \) at \( V_g \) at 4.2K) does not cause considerable distortion of the measured gap. Data from Ref. [34]

of self-heating the "measured" gap vanishes at \( T = T_c \), i.e., simultaneously with the true gap. Thus trivial self-heating simply can not "hide" the qualitative \( \Delta(T) \) dependence. For example, there is no way in which one can get the vanishing "measured" (self-heating affected) gap if the true gap is \( T \)-independent. Therefore, the discrepancy in the measured strong \( T \)-dependence of the gap at \( T_c \) in intrinsic tunneling (and recent ARPES [2]) or complete \( T \)-independence in STM measurements [3] can not be attributed to self-heating.

To quantitatively analyze the significance of self-heating in our mesas, we provide values of the dissipation power \( P = IV \) for the studied mesas: For the near-OP Bi(Y)-2212 mesa at \( T = 4.9K \) dissipation at the sum-gap peak is \( 0.82 mW \). From the previous analysis of the size dependence of \( V_g \) for similar mesas [36] it was observed that \( V_g \) becomes size-independent, and therefore not affected by self-heating, for small mesas with \( P(V_g) < 1 mW \). All the mesas studied in this manuscript fall into this category. All of them exhibit perfect periodicity of quasiparticle branches [7] and none of them exhibit back-bending at any \( T \), which according to Fig. S9, implies that self-heating at \( V_g \) is less than \( T_c/2 \) at \( T = 4.2K \). This is consistent with previous in-situ measurements [33], \( R_{th}(T = 4.2K, V = V_g) \approx 30-40 K/mW \).

The rest of the data presented in the manuscript is not affected by self-heating for the following reasons:

The effective \( T_{eff} \) in Fig. M2b) was obtained at \( V/N = 30 meV \) at which the dissipation power was only \( P \approx 3.8 \mu W \) at \( T = 5.5K \) (for comparison, \( P(V_g) = 0.21 mW \) for the same curve). Therefore, self-heating here is negligible (sub-Kelvin). Besides, as argued in the manuscript, the slope of \( \ln dI/dV(V) \) in Fig. M2 is apparently \( V \) and \( P \) independent, therefore the same data could be obtained at much smaller or larger \( P \). The data in Figs. M3 b-d) and M4 is free from self-heating because it was obtained at zero bias.

Finally, it is important to emphasize that the concept of heat diffusion is applicable for small Bi-2212 mesas containing only few atomic layers. The phonon transport in this case is ballistic [34, 45] and the energy flow from the mesa is determined not by collisions between the tunneled non-equilibrium QP with thermal phonons, but by spontaneous emission of a phonon upon relaxation of the non-equilibrium QP [13]. This process is not hindered at \( T = 0 \). Therefore, the effective \( R_{th} \) (and self-heating) can be much smaller because it is not limited by poor thermal conductivity at \( T = 0 \), but is determined by the fast, almost \( T \)-independent, non-equilibrium QP relaxation time. The concept of self-heating becomes adequate only in the bulk of the Bi-2212 crystal, where the dissipation power density and the temperature rise are much smaller due to the much larger area of the crystal. For more details see the discussion in Ref. [36]. The non-equilibrium energy transfer channel is specific for atomic scale intrinsic Josephson junctions made of perfect single crystals. It can explain a remarkably low self-heating at very high bias [13].
that for strongly UD mesas it is $R_0$ (circles in Fig. S10 b), rather than $R_0/T$, that is more accurately described by the Arrhenius law (dashed line).

From comparison of Figs. M2a) and S10a) it is seen that Eq.(M1) does provide a good fit to experimental data at $T > T_c$, including the crossing point. Simultaneously a dramatic discrepancy is seen below $T_c$. This indicates that the slopes of $dI/dV(V)$ curves in the superconducting state are not determined by the real $T$, but by some constant effective temperature, as shown in Fig. M2 b).

Fig. S10 b) shows the slope of experimental curves from Fig. M2a) at $V/N = 30 mV$ (red squares) and the logarithm of zero bias resistance (green circles) as a function of $1/T$. Clear linear dependence of both at $T > T_c$ indicates that the IVC’s in the normal state have thermal activation nature both at zero and finite bias, consistent with Eq.(M1). According to Eq.(M1) the effective temperature can be explicitly obtained from the slope $dI/dV[\ln(dI/dV)] = (e/2k_B T) \tanh(eV/2k_B T) \simeq e/2k_B T$. Here there are two slight obstacles: First, the $\tanh$ term deviated from unity at high $T$ leading to eventual saturation of the slope at $1/T \to 0$. The exact $T$–dependence according to Eq.(M1) is shown by the solid (magenta) line in Fig. S10 b). We emphasize that there are no fitting parameters in this curve. It is seen that Eq.(M1) reproduces quite well $T$–dependence of the slopes, except for a small offset. This additional offset is caused by the fact that the slope of experimental curves becomes negative at large $T$ [3]. The simple TA expression, Eq.(M1), does not reproduce this negative slope. However, more advanced TA simulations do reproduce the negative slope and, in fact, all the details of experimental $dI/dV$ characteristics at $T > T_c$ [10].

It should be noted that TA-like behavior is quite universal for many process. Except for pure thermal activation over the finite barrier without the gap in the electronic DoS [10], it may appear in pure tunneling characteristics at the presence of the gap in electronic DoS, due to TA-like behavior of the Fermi factor (even though this would require a very specific correlation between $T$–dependent factors mentioned in sec.I); as a result of inelastic tunneling via the impurity [28], or elastic tunneling via a resonant state [29] in the tunnel barrier. Abrikosov has shown that the later scenario can quantitatively reproduce the interlayer characteristics of HTSC in the normal state [24]. The blue dash-dotted line in Fig. S10 b) represents the slope $dI/dV[\ln(dI/dV)]$ in the case of resonant tunneling with the appropriate energy of the resonant state. It follows the simple linear TA behavior at $k_B T$ lower than the resonant energy and saturates at higher $T$. Comparison with the resonant tunneling calculation shows that the saturation of the slope at high $T$ results in appearance of a negative offset in the linear $dI/dV[\ln(dI/dV)]$ vs $1/T$ dependence at $1/T \to 0$. Below the saturation temperature, the actual temperature can be easily extracted from the $dI/dV[\ln(dI/dV)]$ slopes, compensated by the offset at $1/T \to 0$, as indicated in
Fig. S10 b) :

\[ T_{\text{eff}} = \frac{e}{2k_B \left[ d/dV \ln(dI/dV) \right] - \text{Offset}(T^{-1} \rightarrow 0)}. \]

This expression was used for obtaining the effective \( T \) shown in Fig. M2b).

From Fig. S10 b) it is seen that at \( T < T_c \) both \( \ln R_0 \) and the slope of experimental curves simultaneously deviate downwards with respect to the linear \( TA \) behavior. Therefore, both are consequences of the same phenomenon (which we attributed to doping-induced change in the \( c \)-axis transport) only at zero and finite bias, respectively.

[1] J.L. Tallon and J.W. Loram, Physica C 349, 53 (2001)
[2] D.N. Basov and T. Timusk, Rev.Mod.Phys. 77, 721 (2005).
[3] Ch. Renner, et al., Phys.Rev.Lett. 80, 149 (1998)
[4] S.I. Vedeneev, et al., Phys.Rev.B 49, 9823 (1994);
[5] Ya.G. Ponomarev, et al., Sol.St.Commun. 111, 513 (1999)
[6] V.M. Svistunov, et al., JETP Lett. 71, 289 (2000)
[7] V.M. Krasnov, et al., Phys.Rev.Lett. 84, 5860 (2000);
[8] V.M. Krasnov, Phys.Rev.B 65, 140504(R) (2002).
[9] W.S. Lee, et al., Nature 450, 81 (2007); T. Kondo, et al., Phys.Rev.Lett. 98, 267004 (2007).
[10] M. LeTacon, et al., Nature Phys. 2, 537 (2006)
[11] Y. Yamada, et al., Phys.Rev.B 68, 054533 (2003)
[12] Yu.I. Latyshev, et al., Physica C 362, 156 (2001)
[13] V.M. Krasnov, Phys.Rev.Lett. 97, 257003 (2006)
[14] M.H. Bae, et al., J. Korean Phys. Soc. 48, 1017 (2006).
[15] For details see the on-line Supplementary Information.
[16] M. Giura, et al., Supercond. Sci. Technol. 20, S54 (2007)
[17] J.M. Martinis, et al., Phys. Rev. B 35, 4682 (1987)
[18] V.M. Krasnov, et al., Phys.Rev.Lett. 95, 157002 (2005)
[19] A. Yurgens, et al., Phys.Rev.Lett. 79, 5122 (1997)
[20] N. Morozov, et al., Phys.Rev.Lett. 84, 1784 (2000)
[21] T. Watanabe, et al., Phys.Rev.Lett. 79, 2113 (1997).
[22] B.N. Taylor and E. Burstein, Phys.Rev.Lett. 10, 14 (1963)
[23] J.M. Rowell and W.L. Feldmann, Phys. Rev. 172, 393 (1968).
[24] A.W. Kleinsasser, et al., Phys.Rev. 72, 1738 (1994)
[25] E.N. Bratus’, et al., Phys.Rev.Lett. 74, 2110 (1995)
[26] T. Cren, et al., Europhys. Lett. 54, 84 (2001);
[27] J. Lee, et al., Nature 442, 546 (2006)
[28] J. Lambe, and R.C. Jaklevic, Phys.Rev. 165, 821 (1965);
[29] A.A. Abrikosov, Physica C 317-318, 154 (1999).
[30] Y.Yamada and M.Suzuki, Phys.Rev. B 66, 132507 (2002)
[31] L.B.Ioffe and A.J.Mills, Phys.Rev.B 58, 11631 (1998).
[32] W.Kim and J.P.Carbotte, Phys.Rev.B. 63, 054526 (2001)
[33] M.R.Norman, M.Randeria, H.Ding, and J.C.Campuzano, Phys.Rev.B. 57, R11093 (1998)
[34] V.M.Krasnov, Physica C 372-376, 103 (2002)
[35] I.I.Larkin and Yu.N.Ovchinnikov, Sov.Phys.JETP 24, 1035 (1967)
[36] V.M.Krasnov, M.Sandberg, and I.Zogaj, Phys.Rev.Lett. 94, 077003 (2005)
[37] V.M.Krasnov, A.Yurgens, D.Winkler and P.Delsing, J.Appl.Phys. 89, 5578 (2001); ibid. 93, 1329 (2003).
[38] V.M.Krasnov, Phys.Rev.B 75, 146501 (2007)
[39] V.N. Zavartisky, Phys. Rev. Lett. 92, 259701 (2004); Physica C 341-348, 440 (2004); Phys. Rev. B 72, 094503 (2005).
[40] P.J.Thomas, J.C.Fenton, G.Yang, and C.E.Gough, Physica C 341-348, 1547 (2000); J.C.Fenton, P.J. Thomas, G. Yang, and C.E. Gough, Appl.Phys.Lett 80, 2535 (2002)
[41] A.Yurgens, D.Winkler, T.Claeson, S.Ono and Y.Ando, Phys.Rev.Lett. 92, 259702 (2004); cond-mat/0309131
[42] H.B.Wang, T.Hatano, T.Yamashita, P.H.Wu, and P.Mullet, Appl.Phys.Lett. 86, 023504 (2005);
[43] V.M. Krasnov, T.Bauch, and P.Delsing, Phys.Rev.B 72, 012512 (2005)
[44] V.M.Krasnov, A.E.Kovalev, A.Yurgens, and D.Winkler, Phys.Rev.Lett. 86, 2657 (2001)
[45] M.F.Crommie and A.Zettl, Phys.Rev.B 43 (1991) 408.
[46] F.F.Balakirev, et al., Nature 424, 912 (2003); S.Ono, S.Komiya and Y.Ando, Phys.Rev.B 75, 024515 (2007)