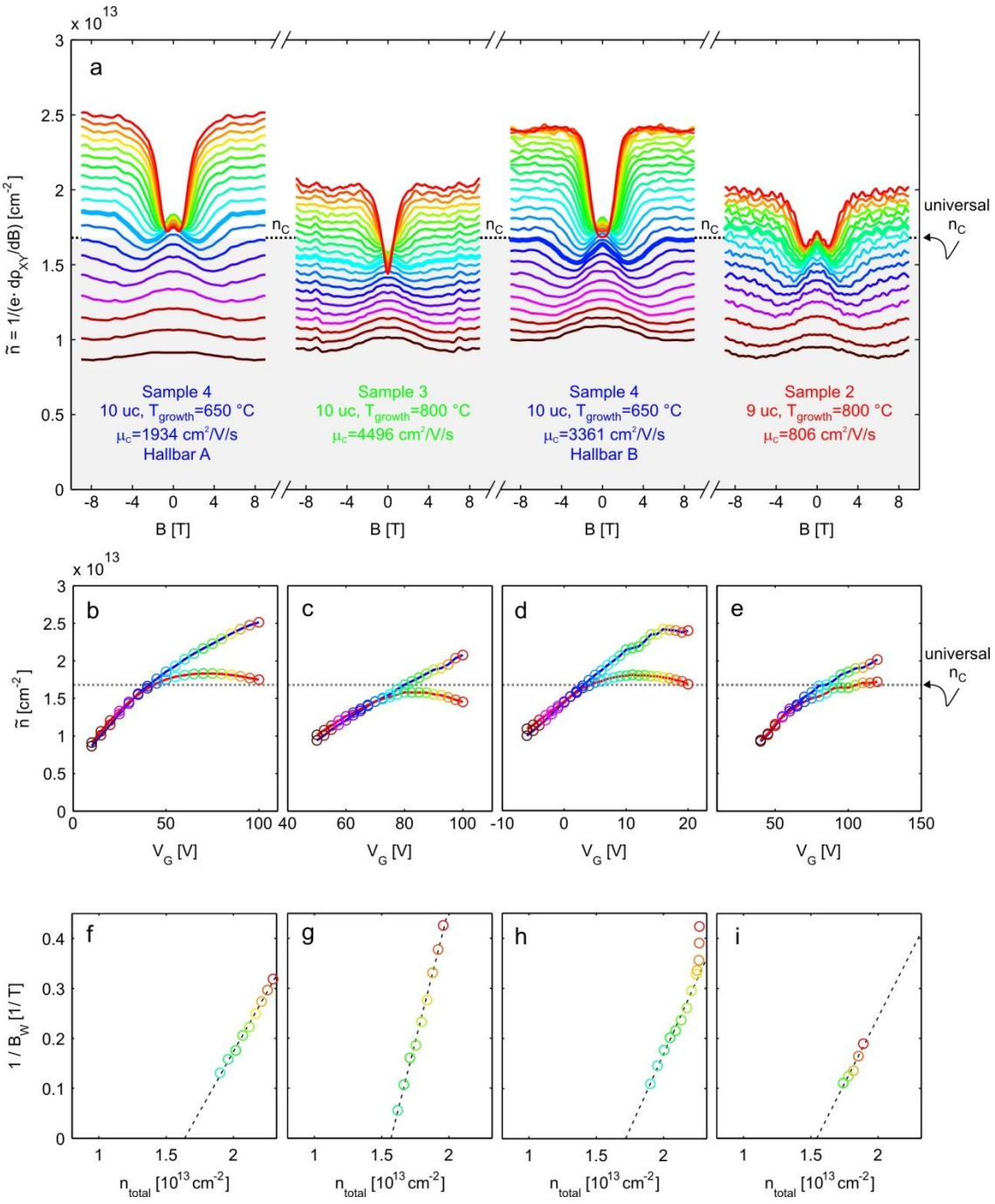
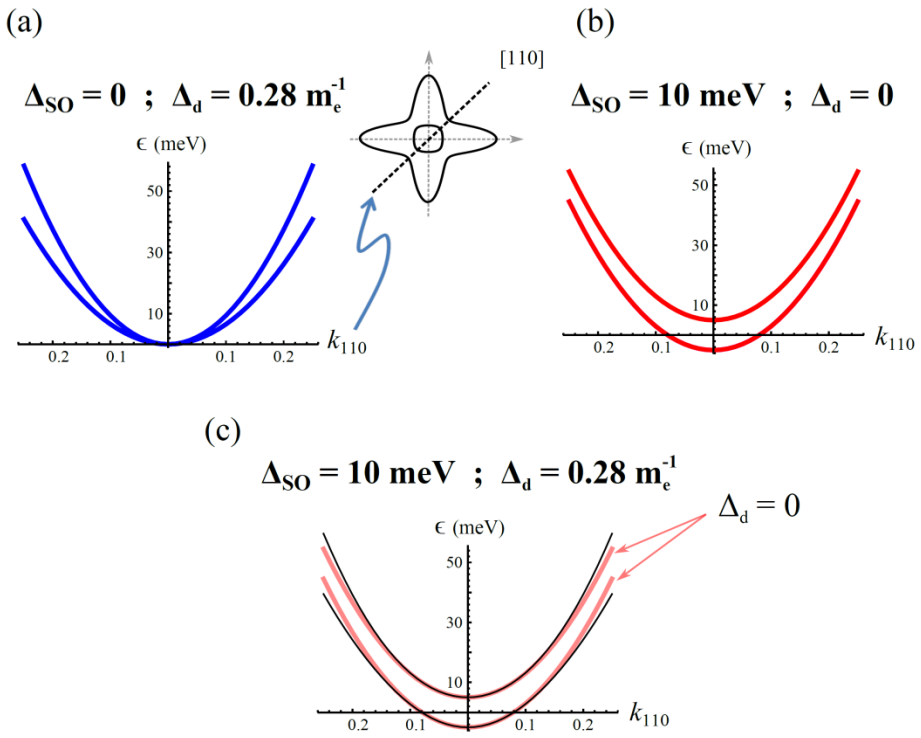
SUPPLEMENTARY FIGURES

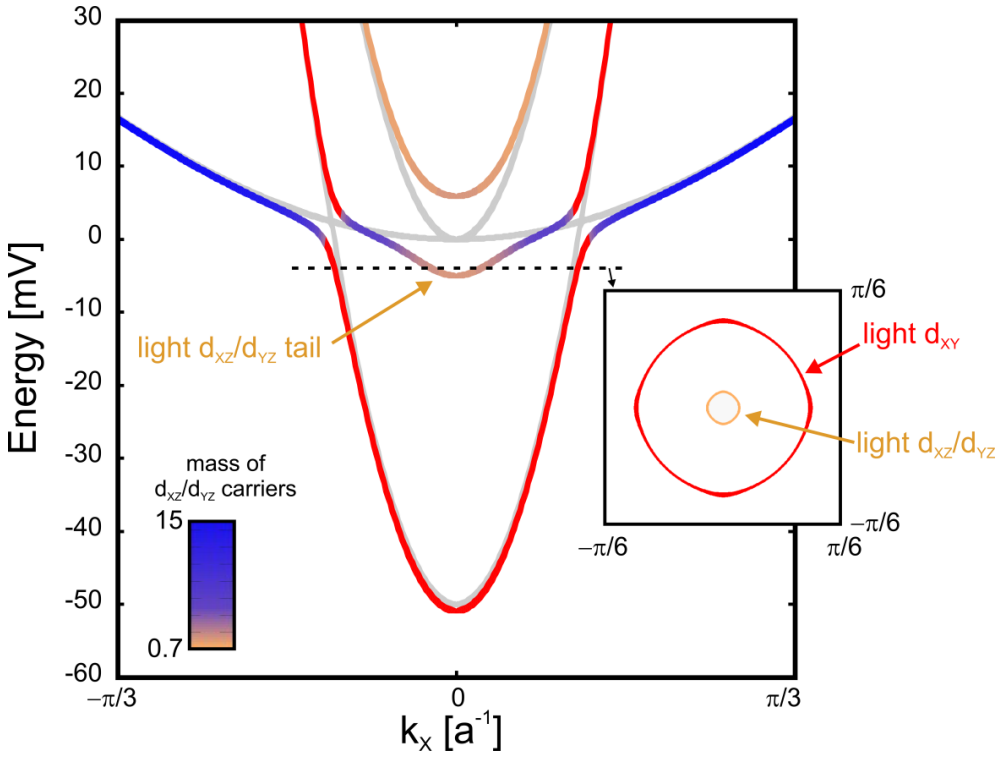


Supplementary Figure S1: The universality of the critical density and of Hall traces across multiple samples. The figure shows the data for samples 2 – 4 (similar data from sample 1 is shown in the main text). a) The effective density, $\tilde{n} = 1/(e \cdot d\rho_{XY}/dB)$, calculated from the derivative of the measured Hall resistance with respect to magnetic field, $d\rho_{XY}/dB$, is plotted as a function of B, for various gate

voltages, V_G (brown to red go from the lowest to the highest V_G). For each sample we indicate the LAO layer thickness, growth temperature and measured mobility at the critical transition (see main text). The dashed black line marks the universal critical density, $n_C = 1.68 \cdot 10^{13}\ cm^{-2}$, derived in the main paper. In all samples, independent of mobility and LAO thickness, we observe a clear change of the characteristic traces around the universal density (see text). b) - e) Corresponding plots of the effective density at zero field, $\tilde{n}(B = 0T)$ [red line], and high field, $\tilde{n}(B = 9T)$ [blue line], as a function of V_G for the samples in panel a. For all samples these two density traces depart near the universal critical density, marked by a grey line in the figures. Above this density, $\tilde{n}(B = 9T)$ continues to rise whereas $\tilde{n}(B = 0T)$ remains roughly constant. f) - i) Inverse of the scaling width, B_W, defined in the main paper, as a function of total density obtained from the Hall coefficient at $B = 9T$. Once again, in all the samples, independent of mobility and LAO thickness, $1/B_W$ extrapolates to zero indicating a divergence of B_W at the universal critical density, $n_C = 1.68 \pm 0.18 \cdot 10^{13}\ cm^{-2}$. Note that the vertical scale of f), h) and i) ranges between 0 and 0.45, whereas it ranges between 0 and 0.9 for g).



Supplementary Figure S2: Comparing the relative contributions of ASO and diagonal hopping terms to the energy dispersion of the heavy bands. a) Calculated bands with diagonal hopping but without ASO along the [110] direction in *k*-space (inset), along which the band mixing is important, using the parameters determined by fitting the DFT calculation in Fig.6 of Ref. 13 to the tight binding model described in the Supplementary Methods. The gap opened by the diagonal hopping goes as k^2 and therefore at the Γ-point the bands are degenerate. b) Similar calculation, but only with ASO. Here the gap is independent of *k* c) A calculation with both ASO and diagonal hopping terms included (black lines). In this case the gap goes as $\sqrt{4\Delta_d^2\, k^4 + \Delta_{SO}^2}$. For comparison we also plot in red lines the bands calculated only with ASO (as in panel b), clearly showing that the effect of diagonal hopping is negligible several tens of meV above the Lifshitz point.



Supplementary Figure S3: The band structure at the LAO/STO interface with atomic spin-orbit (ASO) interactions. The main panel shows the energy bands without ASO (gray) and with ASO (colored) calculated with the ARPES measured[28] values for the light and heavy masses, $m_l = 0.7m_e$ and $m_h = 15m_e$ (m_e is the electronic mass), an energy splitting of $\Delta_E = 50meV$, and taking the strength of ASO to be $\Delta_{ASO} = 10meV$. In the calculation with ASO the states are colored according to their orbital content and effective mass: States whose d_{XY} content is larger than 20% are colored red. The remaining d_{XZ}/d_{YZ} states are colored according to their effective mass, ranging from brown (light mass) to blue (heavy mass) (see colorbar). The mass plotted is that along the x direction, $m^* = \hbar^2 k_x/(dE/dk_x)$, but qualitatively similar behavior is obtained when we plot the mass averaged over the entire Fermi surface. ASO modifies the bands mostly around their degeneracy points. One consequence of that is that the heavy (blue) d_{XZ}/d_{YZ} carriers become light (brown) at the bottom of these bands. Inset: The Fermi surfaces 1meV above the bottom of the d_{XZ}/d_{YZ} bands, showing large circular Fermi surface of light d_{XY} carriers and a small circular Fermi surface of light d_{XZ}/d_{YZ} carriers.

SUPPLEMENTARY METHODS

Data from additional samples.

Figure 3f of the main text collected data from four LAO/STO samples that differed in the number of LAO unit cells and in their mobilities. All these samples showed a critical transition at the same universal density. In Fig. 1 we showed the detailed Hall traces of sample 1 with 6 unit cells of LAO. In this section we present the detailed Hall traces of samples 2 - 4 and analyze them in light of the ASO model described in the last section of the Supplementary Methods.

In the main text we demonstrated that both the Hall resistance, ρ_{XY}, and the Hall coefficient, $R_H = \rho_{XY}/B$, show clear signatures of the critical density. To make the identification of this density even more intuitive we plot in Supplementary Figure S1a a related quantity, the effective density, $\tilde{n} = 1/(e \cdot d\rho_{XY}/dB)$, determined from the derivative of ρ_{XY} with respect to B. Compared to R_H, this derivative is a local quantity and thus accentuates the magnetic field-dependent features in the curves. For each sample, we plot in Supplementary Figure S1a, $\tilde{n}$ as a function of B for various gate voltages, V_G. Notably, in all samples we can identify a clear critical transition density when we go from low V_G (brown traces) to high V_G (red traces), as is explained below.

There are two sharp criteria to identify the critical density, n_c. The first is that below n_c the Hall slopes (and correspondingly $\tilde{n}$) are identical at zero and high field, whereas above n_c they are different. The second is that once n_c is crossed, the low field density, $\tilde{n}(B = 0T)$, remains stuck at n_c. Clearly, all the samples in Supplementary Figure S1a exhibit these two changes around the universal value of n_c (dashed line in the figure). Below n_c, in all the curves $\tilde{n}(B = 9T)$ is roughly equal to $\tilde{n}(B = 0T)$ whereas above n_c the former becomes larger than the latter, and their difference grows with increasing V_G. Furthermore, once $\tilde{n}(B = 0T)$ reaches n_c it remains stuck at this value although the total density continues to rise, as is evident from the crowding of curves near n_c at $B = 0T$. These two changes are seen more directly when we plot $\tilde{n}(B = 0T)$ and $\tilde{n}(B = 9T)$ as a function of V_G for all samples (red and blue traces respectively, in Supplementary Figures S1b-S1e).

Further evidence for a clear signature of the critical density in these additional samples comes from extracting the characteristic width, B_W, of their R_H traces. This width was obtained by fitting the R_H trace at each V_G to a Lorentzian, as was shown in Fig. 2c. Supplementary Figures S1f - S1i show the result of the fit as a function of the total density, $n_{total} = 1/(eR_H(B = 9T))$. The scaling width reflects the mobility of the system (see Eq. S4 below) and therefore not only changes from sample to sample but also at each gate voltage V_G since the mobility changes as V_G is varied. However, in all samples the inverse of the scaling width, $1/B_W$, appears to diverge as the total density is lowered to the critical density, $n_C = 1.68 \pm 0.18\ cm^{-2}$, thus highlighting the universal nature of the Lifshitz transition.

Careful examination of the curves reveals that at intermediate magnetic fields there are fine features, which are not captured by the above analysis. These are apparent as small dips in $\tilde{n}$, which are equivalent to the peaks of R_H at $B = B_P$ shown in Fig. 1c for sample 1. The physics of these dips in $\tilde{n}$ (peaks in R_H) will be described in a future publication. Here, however, we want to comment on their relation to the multiple-band physics. Above n_c these dips show a strong correlation with the signatures of multiple bands. As a function of magnetic field these dips always appear right before the sharp rise in $\tilde{n}$, associated with the multiple bands. The magnetic field position of this rise varies substantially with gate voltage, however, as can be seen from Supplementary Figure S1a the dips in $\tilde{n}$ always perfectly track the position of this rise. The exact same behavior was observed for sample 1 in the main text, where we showed that peaks in R_H and the R_H falloff associated with multiband physics scale together in magnetic field (Fig. 2c).

Interestingly, we can see in Supplementary Figure S1a that the dips in $\tilde{n}$ appear also below n_c. Moreover, if we track the magnetic field position of these dips we see that they evolve continuously across the transition. These observations hint that even below n_c there is a mild multiple-band behavior which evolves continuously into the strong multiple-band behavior commencing at n_c. One might expect weak multiple-band signatures from higher d_{XY} subbands, however, since these subbands are independent of the d_{XZ}/d_{YZ} bands that are responsible for the strong multiple-band behavior above n_c, generally the former would not continuously evolve to the latter. A much more natural

explanation for the above observation is given by the ASO model described in the last section of the Supplementary Methods. Within this model the mass of the d_{XZ}/d_{YZ} carriers changes rapidly but continuously from light to heavy as a function of the filling of these bands. It is thus straightforward to assign the weak multi-band signatures below n_c to the light d_{XZ}/d_{YZ} carriers and the strong multi-band signatures above n_c to the heavy d_{XZ}/d_{YZ} carriers at higher filling. Within this model the evolution would indeed be sharp yet continuous as we see in our measurements.

The two-band approximation to the transport above the critical density.

The simplest model for understanding the behavior of the transport above the critical density is the two-band approximation[24]. This model assumes that two types of carriers are conducting in parallel, each having its own density and mobility. It also assumes that the parameters of these bands are independent of B. This model is often used in the LAO/STO literature to extract the mobilities and densities of two carriers when Hall resistance traces are S-shaped. In this section we show that this model should be used with caution. While it can capture some basic features of the data, it can also lead to large mistakes in some of the extracted parameters. In the main text we were therefore careful to use only the robust predictions of this model that are insensitive to its assumptions which might be violated.

Assuming two parallel conducting bands whose densities and mobilities are n_1, μ_1, n_2, μ_2 the model predicts that the Hall coefficient of the combined system, R_H, and its longitudinal resistance, ρ_{XX}, should have Lorentzian shapes as a function of B:

$$R_H = R_\infty + \frac{R_0 - R_\infty}{1+\left(B/B_W\right)^2}, \qquad \rho_{xx} = \rho_\infty + \frac{\rho_0 - \rho_\infty}{1+\left(B/B_W\right)^2} \qquad \text{(Eq. S1)}$$

where the asymptotic values of these two quantities at zero and infinite fields can be written in terms of the parameters of the bands as:

$$R_0 = \frac{n_1\mu_1^{\,2} + n_2\mu_2^{\,2}}{\left(n_1\mu_1 + n_2\mu_2\right)^2}, \qquad R_\infty = \frac{1}{\left(n_1 + n_2\right)} \qquad \text{(Eq. S2)}$$

and:

$$\rho_0 = \frac{1}{n_1\mu_1 + n_2\mu_2}, \qquad \rho_\infty = \frac{\left(n_1\mu_2 + n_2\mu_1\right)}{\left(n_1 + n_2\right)^2 \mu_1\mu_2} \qquad \text{(Eq. S3)}$$

The model further predicts that the Lorentzians of both R_H and ρ_{XX} should have the same characteristic width in magnetic field:

$$B_W = \frac{n_1\mu_1 + n_2\mu_2}{\left(n_1 + n_2\right)\mu_1\mu_2} \qquad \text{(Eq. S4)}$$

In Figs. 2c and 2d in the main text we compare the measured R_H and ρ_{XX} traces at gate voltages above the critical density with a Lorentzian fit and indeed find a reasonable agreement with this functional form. Moreover, as is predicted by the model, we see that R_H and ρ_{XX} share the same characteristic field scale, B_W. We demonstrated this by showing that if for every gate voltage we use the characteristic field with which we scaled R_H to scale the ρ_{XX} at the same gate voltage, we get that all the ρ_{XX} traces perfectly collapse onto a universal curve. Thus, overall it seems like this simple model can capture the basic features in our data.

However, if we carefully examine the fits of the model we can clearly identify systematic deviations. For example, we see that at low field both R_H and ρ_{XX} deviate from Lorentzians – the former peaks at a finite field $\left(B_P\right)$ and not at zero, and the latter is extremely flat up to B_P, both deviating from the simple Lorentzian dependence. Clear deviations are also observed at high fields, especially in ρ_{XX}, which shows a linear increase as a function of *B* up to the highest field in our measurements, inconsistent with the Lorentzian shape. One might consider improving the model by implementing more realistic bands: three bands instead of two, elliptical Fermi surfaces instead of spherical, and even including the effect of ASO. We found, however, that even when we include these details, the above deviations could still not be explained.

In the literature the above details are often ignored, and a rough fit to kinked Hall curves is used to extract two densities and two mobilities for the carriers. Such fits can give apparently reasonable values of mobilities and densities, especially at high total densities. We find, though, that using this model brute-force to systematically study the gate dependence of the extracted parameters leads to unphysical results. For example, we get that the total density in the bands decreases with increasing gate voltage, which is clearly wrong. The most significant error in the model is in its prediction of the crossover field. As can be seen from Eq. S4, when the density of the second band, n_2, is low (as is the case near the critical point) the model predicts a crossover field at $B_W \approx 1/\mu_2$, the inverse of the lower mobility of the two bands, whereas in our data we consistently find that the crossover appear at $B_W \approx 1/\mu_1$, the inverse of the higher mobility of the two bands.

In a future publication we will describe a more elaborate model which we believe can account for the above observations. Here, however we do want to note that the main problem of the simple two-band model is its assumption that the parameters of the bands are independent of magnetic field. As a result, one has to be cautious when using this model, and extract from it only robust parameters that are insensitive to this assumption. The parameters that are most sensitive to this assumption are the mobilities of the two bands, which are directly affected by the inconsistencies between the model and the data in the value of B_W, described above. Thus, imposing the two band model on the data would often result in large errors in the extracted mobilities. In the main text we were therefore careful not to extract the mobilities, but extracted only the densities which are less sensitive to the errors of the model. Specifically, to extract the densities we have used in the main text two identities: The first relates the Hall coefficient at large fields with the total density, $R_\infty = 1/(en_{total})$. As can be seen from Eq. S2 above, this identity is always valid independent of the mobilities of the two bands. The second relates the Hall coefficient at B=0 to the density of the high mobility band, $R_0 = 1/(en_{high})$. This identity is mathematically correct in two independent cases: when the mobilities of the two bands are substantially different, or when the density of the second band is low enough

$n_2 << n_1\mu_1 / \mu_2$, as is the case near the critical density. If our assumption of substantially different mobilities in the two bands is not accurate, this could modify our extracted densities much above the critical density, however, near the critical density the picture that we presented would still be completely accurate. Specifically, the determination of the universal density, which is a main point in this paper is completely robust and would not depend on any assumptions about the mobilities.

Details of the calculation of the energy bands with atomic spin-orbit interactions

In the absence of spin-orbit coupling, the effective Hamiltonian is decoupled from the spin degree of freedom. Near the bottom of the three t_{2g} bands it is given by

$$H_0 = \begin{bmatrix} \frac{k_x^2}{2m_h} + \frac{k_y^2}{2m_l} & \Delta_d k_x k_y & 0 \\ \Delta_d k_x k_y & \frac{k_x^2}{2m_l} + \frac{k_y^2}{2m_h} & 0 \\ 0 & 0 & \frac{k_x^2}{2m_l} + \frac{k_y^2}{2m_l} - \Delta_E \end{bmatrix}$$

where the three dimensional Hilbert space represents the d_{YZ}, d_{XZ}, and d_{XY} bands respectively, m_l is the light mass, m_h the heavy mass, Δ_E the splitting of the d_{XY} from the other two bands due to the transverse confinement, and Δ_d is a band mixing term resulting from diagonal next-nearest-neighbor hopping between d_{XZ} and d_{YZ} orbitals. The last term is the dominant orbital mixing term between the heavy bands and is known to be significant in the band structure of transition metals. To quantitatively determine the actual size of this term in the LAO/STO system we fitted the DFT calculation in Ref. 13 to the tight binding model above, yielding $\Delta_d = 0.28 m_e^{-1}$. The fit describes accurately the heavy bands over the entire Brillouin zone including their splitting along the diagonal direction in *k*-space which directly reflects the diagonal hopping term.

The same Hamiltonian in the full six dimensional Hilbert space, including spin, can be written as $\mathcal{H}_0 = H_0 \otimes \sigma^0$, where σ^0 is the 2 by 2 identity matrix. The atomic spin-orbit term is written in the same space as:

$$\mathcal{H}_{SO} = \Delta_{SO} L \cdot \sigma = \Delta_{SO} \sum_{\alpha=1}^{3} L^{\alpha} \otimes \sigma^{\alpha} \ ,$$

where σ^{α} are the Pauli matrices acting on the spin and L^{α} are angular momentum-2 matrices, projected to the space of the three t_{2g} orbitals. Because the spin-orbit coupling is completely local the Kramer's degeneracy at the atomic level is carried over to all wave-vectors.

Another type of coupling between the t_{2g} orbitals can be induced because the transverse confinement is not symmetric to reflection $(z \to -z)$. This allows coupling between the d_{XY} and d_{XZ} through anti-symmetric hopping along y and similarly between the d_{XY} and d_{YZ} in hopping along x. Such couplings are encapsulated in the term:

$$\mathcal{H}_{z} = \Delta_{z} \begin{bmatrix} 0 & 0 & -i\,k_x \\ 0 & 0 & i\,k_y \\ i\,k_x & -i\,k_y & 0 \end{bmatrix} \otimes \sigma^{0}$$

Since this term is not fully local (i.e. it is momentum dependent), it splits the band degeneracy. In fact, taken together with the ASO this last term generates an effective Rashba coupling near the bottom of the d_{XY} band. We note however that under realistic conditions $\Delta_z k_F$ is much smaller than Δ_{SO} and therefore the Rashba splitting of the bands will be almost unnoticeable on the scale of Supplementary Figure S3.

To understand the relative importance of the different band-mixing terms we plot in Supplementary Figure S2 the energy bands calculated with realistic parameters for three different cases: Including diagonal hopping but without ASO (Supplementary Figure S2a), including ASO but without diagonal hopping (Supplementary Figure S2b) and including both ASO and diagonal hopping (Supplementary Figure S2c). The energy bands in the figure are plotted along the [110] direction in *k*-space, along which the heavy bands are degenerate and their mixing is important.

Supplementary Figure S2c shows that when both ASO and diagonal hopping are present the effect of the latter near the bottom of the heavy bands is negligible. The reasons for that are simple: First, unlike the ASO-induced mixing, which is *k*-independent and therefore sizable for all Fermi energies, the diagonal hopping term goes as k^2 and thus

disappears near the bottom of the bands. Furthermore, the contributions of these two terms to the energy are not additive, but rather go as $\sqrt{4\Delta_d^2\, k^4 + \Delta_{SO}^2}$, reducing the effect of diagonal hopping even further whenever it is smaller than the ASO. For typical carrier densities in LAO/STO the Fermi energy lies only few meV above the bottom of the heavy bands, and in this range diagonal hopping has a negligible effect on the bands calculated with ASO which are presented in the main text.

The band structure with atomic spin-orbit interactions and its relation to the observed transport properties.

In the first part of the main text we used the simplest band model that captured the essential features of the Lifshitz transition. We then showed that the inclusion of atomic spin-orbit (ASO) interactions can further explain the strong correlation between the transition point (the bottom of the d_{XZ} and d_{YZ} bands) and strong spin-orbit interactions, as observed experimentally. In this section we demonstrate that the refined band structure with ASO has additional features that may explain other surprising observations in the LAO/STO system. For example, we argue that such a model can naturally explain why carrier densities extracted from Shubnikov-de-Haas oscillations are so much lower than those determined by Hall measurements[18,19,36-38].

Supplementary Figure S3 shows the energy bands calculated with and without ASO (colored and grey lines correspondingly). The details of the calculations were explained in the previous section. As can be seen in the figure, the effect of ASO is strongest near band degeneracy points. At these points ASO hybridizes the electronic states and opens energy gaps in the spectrum. Such degeneracies exist between the d_{XZ} and d_{YZ} bands at the Γ point of the Brillouin zone, and between the d_{XY} and the other two bands at their crossing points. Since the heavy band is very shallow, these three types of crossings appear close in energy, around the transition point.

Note that the full band structure, calculated by DFT[13], has also other non-SO terms that mix the d_{XZ} and d_{YZ} bands and open a hybridization gap between them. These are fully captured by a second-nearest-neighbor diagonal hopping term, Δ_d, that mixes the

d_{XZ} and d_{YZ} orbitals. When we add this term to our Hamiltonian (previous section) we can accurately reproduce the DFT bands at low Fermi energies (<100meV) relevant to our experiments. In fact, since the SO term is independent of the momentum whereas the diagonal hopping term is quadratic in it, the effect of the latter is negligible in all relevant Fermi energy and thus the band structure in Supplementary Figure S3 is a reliable representation of the full band structure.

ASO forms the basis for all spin-orbit interactions in the system. Specifically Rashba SO that appears when inversion symmetry is broken, is just a second order perturbation process in ASO. Thus the dependence of ASO on the Fermi Energy determines directly the dependence of any SO interactions on the Fermi Energy, including that of the Rashba and other equivalent momentum-dependent SO terms.

ASO has two important consequences that significantly affect the transport in this system:

1. Spin is strongly coupled to the orbital momentum near the Lifshitz critical density. Around this point the d_{XY}, d_{XZ} and d_{YZ} bands cross, allowing the spin-orbit coupling to strongly hybridize them and form superpositions with well-defined atomic orbital momenta. For example, the degenerate d_{XZ}/d_{YZ} states at the Γ point hybridize to form the $d_{\mp} = d_{XZ} \pm i d_{YZ}$ states, with a well-defined projection of the orbital momentum along the z direction, $m_z = \pm 1$. At the same time, the spin becomes strongly coupled to this orbital momentum. Since the orbital momentum has a preferred axis of orientation (e.g. the z-axis for the example above), the spin is also preferably aligned along this axis, and an energy of the order of Δ_{ASO} is required to polarize it away from this axis.

When the Fermi energy is increased above the degeneracy points, the band splittings due to the different dispersion of the d_{XY}, d_{XZ} and d_{YZ} orbitals become increasingly dominant over the ASO coupling and the effect of the latter decays. In Fig. 3k of the paper the enhancement and subsequent decay of the characteristic spin-orbit scale is seen from the variation with the Fermi energy of the expectation value $\langle L \cdot S \rangle$, integrated around the Fermi surface.

Further refinement to this picture is given by the transverse confining potential, which breaks the inversion symmetry perpendicular to the plane. This allows for a small tunneling matrix element between d-orbitals of different symmetries on neighboring atoms. Combined with the strong ASO coupling, this additional tunneling leads to a small, Rashba-like splitting of the energy bands away from the Γ point. This effective Rashba coupling is a direct consequence of the ASO coupling and is thus also peaked near the critical density. However, since it relies on small, field-induced hopping elements, its effect would be much weaker than that of the ASO coupling.

2. The heavy d_{XZ}/d_{YZ} carriers become light at the bottom of their bands. The d_{XZ}/d_{YZ} carriers play a central role in the observations reported in this paper. These carriers have a much larger in-plane mass than the light d_{XY} carriers. Therefore their population is associated with a large jump in the DOS, which is consistent with the transition that we measure in transport. In a simple band model without ASO this jump is abrupt, occurring when the Fermi energy crosses the bottom of the d_{XZ}/d_{YZ} bands. ASO, however, modifies this picture by drastically changing the mass of the d_{XZ}/d_{YZ} carriers near the bottom of their bands.

To demonstrate this effect we have colored the eigenstates in Supplementary Figure S3 according to their orbital content and effective masses. The d_{XY} carriers are colored red, and the d_{XZ}/d_{YZ} are colored by a continuum of colors that represent their effective mass, ranging from brown (light mass) to blue (heavy mass) (see caption for details). At high energies the d_{XZ}/d_{YZ} states have indeed a heavy mass (blue). However, close to the bottom of the band we can clearly identify a low-energy tail in which the d_{XZ}/d_{YZ} carriers are light. This effect is a robust consequence of the hybridization gap created by ASO at the degeneracy point between the d_{XZ}/d_{YZ} orbitals. While the amount of light carriers contained in this low-energy tail is proportional to the strength of ASO, the light mass obtained at its bottom is always the same. In fact, it is straight forward to show that this mass is approximately twice the light mass, $m^* = 2({m_h}^{-1} + {m_l}^{-1})^{-1} \approx 2m_l$ (where m_l and m_h are the light and heavy masses in the absence of ASO). In contrast to the case without ASO, here the mass of the d_{XZ}/d_{YZ} carriers depend on the Fermi energy,

evolving continuously from a light mass at the bottom of the band to a heavy mass at higher Fermi energies.

To further demonstrate the light nature of the d_{XZ}/d_{YZ} carriers near the bottom of the band, we plot in the inset to Supplementary Figure S3 the Fermi surfaces at low band filling. In addition to the circular d_{XY} Fermi surface we see a small, circular, Fermi pocket that corresponds to the light d_{XZ}/d_{YZ} carriers. This is in sharp contrast to the picture without ASO where the d_{XZ}/d_{YZ} Fermi surfaces are elliptical and heavy even at the bottom of their band. While the actual size of ASO in this system is yet to be determined experimentally, from transport experiments[10,12] and ab-intio calculations[33] we expect it to be in the range $\Delta_{\mathrm{ASO}} \approx 10 - 25\mathrm{meV}$. For such ASO strengths the density of light carriers that the d_{XZ}/d_{YZ} band can accommodate before it becomes heavy is about $1.5 - 5 \cdot 10^{12}\mathrm{cm}^{-2}$.

The existence of light d_{XZ}/d_{YZ} carriers has important implications for the transport in this system. On Hall measurements we expect the effect of these carriers to be rather small. The Hall resistance measures the parallel addition of these carriers and the d_{XY} carriers. Since the effective masses of these two types of carriers are comparable and thus also their mobilities, even when these two types of carriers coexist, their combined Hall resistance would behave as the Hall resistance of a single carrier-type with the total density of both. A dramatically different result will occur when one measures the Shubnikov-de-Haas oscillations. These oscillations measure the individual bands independently, giving a different oscillation frequency for each carrier type, which is directly proportional to their individual densities. Since the density of the light d_{XZ}/d_{YZ} carriers is significantly smaller than that of the d_{XY} carriers, it would result in substantially different oscillation period as a function of 1/B. In fact, all Shubnikov-de-Haas measurements in LAO/STO interfaces[18,19] as well as those in δ-doped STO[36-38] have robustly observed a pocket of light mass carriers with densities in the range $1.3 - 4.8 \cdot 10^{12}\mathrm{cm}^{-2}$. These measurements extracted a mass of $m^* \approx 1.2 - 1.4m_e$, which is twice that measured by ARPES[28,29]. The factor two discrepancy between these two measurement techniques is in fact consistent with the ASO model; ARPES experiments, done on a bare STO surface with much higher electron density, determines the masses from high energies and thus yields the bare light mass, whereas transport is sensitive to

the low energies dispersion which is renormalized by ASO. So far the low density pockets measured by Shubnikov-de-Haas were explained by higher subbands of d_{XY} carriers. However, as we see here, the effect of ASO provides an alternative, robust, and intrinsic mechanism for the existence of such light carriers, of the d_{XZ}/d_{YZ} type.

There are two important features of the light d_{XZ}/d_{YZ} carriers that set them apart from light d_{XY} carriers at higher subbands. First, in the ASO model the appearance of the light pocket is intimately tied to the bottom of the d_{XZ}/d_{YZ} bands, namely to the critical density. This is in contrast to higher d_{XY} subbands whose energies do not have to coincide with the d_{XZ}/d_{YZ} band bottom. Secondly, a scenario that includes several d_{XY} subbands would involve several discrete transitions as a function of the Fermi energy, e.g. a transition from a single light carrier type at the lowest d_{XY} subband to two carrier types in the two lowest d_{XY} subbands, and finally to three carriers types that include also the heavy d_{XZ}/d_{YZ} carriers. In contrast, in the ASO model when the d_{XZ}/d_{YZ} start populating there are always only two carrier types. At low filling these are the d_{XY} and light d_{XZ}/d_{YZ} carriers and at high filling they are the d_{XY} and heavy d_{XZ}/d_{YZ} carriers. Although the transition between these two cases is rather sharp as a function of increasing Fermi energy, it is a *continuous* crossover. The strong two-band signatures that we observed in the paper should appear only when heavy carriers are populated. However, if this model is correct, we might expect to see weak signatures of the light mass tail even below the critical density and these signatures should evolve *continuously* to the strong two-band features above the critical density. In the first section of the Supplementary Methods, we presented detailed Hall resistance data that complemented the one shown in the main text and analyzed it in light of these predictions.

SUPPLEMENTARY DISCUSSION

Universal shape of the confining potential at the LAO/STO interface

The two-dimensional electrons near the LAO/STO interface are confined by a potential well whose shape can depend on several parameters including the thickness of the LAO, the electric field from the back gate, and the nonlinear dielectric constant of STO. This shape affects the energy splitting between the d_{XY} and d_{XZ}/d_{YZ} bands, and the splitting in turn determines the carrier density at the Lifshitz transition. Thus, if the well shape can change between samples of different LAO thicknesses or even within a single sample as a function of the back gate voltage, then why is the observed Lifshitz transition density universal?

Naively, one would expect the Lifshitz density to be universal only if the confinement well shape is fixed. But this simple picture is not true: the Lifshitz transition density can be universal even if the confinement well shape is changing, as long as *the shape of the confinement well depends only on the confined carrier density*. Namely, for different carrier densities the shape of the potential might be different, but for any given density this shape is always the same. If this is the case, then *at the critical density*, the well shape and energy band splitting are always the same and therefore (self consistently) the critical density is uniquely defined, although the shape of the well changes with carrier density. But why would the shape of the well be uniquely determined by the carrier density?

In general one might expect the well shape to depend on two independent parameters: The electric field created from the LAO side and the field from the back gate side. If this is the case then one could get the same carrier density with different potential wells thereby leading to a non-universal Lifshitz density. Experimentally, however, transport measurements[43] demonstrate that for an LAO thickness above 4 unit cell the densities and mobilities of the conducting 2D electrons do not change as a function of LAO thickness, showing that the thickness of LAO has small effect over the well shape thereby supporting our universal density scenario.

One possible way to understand this effect is to consider the nonlinearity of the dielectric constant of STO. When the LAO layer is thicker than 4 unit cell electrons are transferred from the LAO surface to the interface to reduce the built-in dipole inside the LAO, such that both the valence band of LAO at the surface and the conduction band of STO at the interface remain at the Fermi energy. The amount of transferred charge (conducting and localized) is correspondingly[44]:

$$n_{transfer} = \frac{1}{2a^2}\left(1 - \frac{d_C}{d_{LAO}}\right), \qquad \text{(Eq. S5)}$$

where a is the lattice constant of LAO, d_{LAO} is the thickness of the LAO layer and $d_C \approx 4uc$ is the critical thickness for the polarization "catastrophe". In the limit of thick LAO this formula gives half an electron per unit cell. The field, E, created at the STO side of the interface by this transferred charge is:

$$E = \frac{en_{transfer}}{\varepsilon_{STO}} = \frac{e}{\varepsilon_{STO}}\frac{1}{2a^2}\left(1 - \frac{d_C}{d_{LAO}}\right). \qquad \text{(Eq. S6)}$$

This large field has a dramatic effect on the dielectric constant of STO, ε_{STO}, near the interface. This constant is known to strongly depend on the electric field[45]. At low fields ε_{STO} is very large and has weak field dependence. Around a critical field, E_C, it drops sharply to $\varepsilon_{STO} \approx 1$ and it remains at this value for all fields above that. To understand the possible consequences of this strong nonlinearity we approximate this dependence by a very crude model that assumes that ε_{STO} drops from a large value (~10,000) to unity as a step function at the field E_C. Under this simplifying assumption the potential well will have two distinctive regions: Near the interface the field is always above E_C resulting in $\varepsilon_{STO} \approx 1$. Since the field is not screened in this region, the potential well is steep. This sharp part of the potential well is believed to confine most of the localized electrons to the TiO_2 plane nearest the interface, which is the majority of the transferred carriers. Once the field crosses E_C it becomes strongly screened and the well becomes much

shallower. This part of the well is the one that confines the mobile electrons. Notably, within this model, the boundary condition for the confining potential for the mobile electrons is always given by E_C, independent of the amount of charge that is transferred from the LAO side. Although this model is over-simplified, especially since the non-linearity of ε_{STO} should play an important role even in the shallow part of the well, it still captures its leading effect on the confinement potential shape and demonstrates why in the presence of non-linear ε_{STO} this shape will depend weakly on the LAO thickness, thereby suggesting why the Lifshitz transition density could be universal, as we observe in our experiments.

SUPPLEMENTARY REFERENCES

43. Cancellieri, C. *et al.* Electrostriction at the $LaAlO_3/SrTiO_3$ Interface. *Phys. Rev. Lett.* **107**, 056102 (2011).

44. Son, W., Cho, E., Lee, B., Lee, J. & Han, S. Density and spatial distribution of charge carriers in the intrinsic n-type $LaAlO_3$-$SrTiO_3$ interface. *Phys. Rev. B* **79**, 245511 (2009).

45. Copie, O. *et al.* Towards Two-Dimensional Metallic Behavior at $LaAlO_3/SrTiO_3$ Interfaces. *Phys. Rev. Lett.* **102**, 216804 (2009).