Two-dimensional excitonic quasiparticles in a three-dimensional crystal: The case of anatase TiO_2

Edoardo Baldini[1, 2], Letizia Chiodo[3,4], Adriel Dominguez[5], Maurizia Palummo[6], Simon Moser[7], Meghdad Yazdi[8], Gerald Auböck[1], Benjamin P. P. Mallett[8], Helmuth Berger[9], Arnaud Magrez[9], Christian Bernhard[8], Marco Grioni[7], Angel Rubio[5, 10, 11], and Majed Chergui[1]

[1] *Laboratory of Ultrafast Spectroscopy, ISIC and LACUS, EPFL, CH-1015 Lausanne, Switzerland*

[2] *Laboratory for Ultrafast Microscopy and Electron Scattering, ICMP and LACUS, EPFL, CH-1015 Lausanne, Switzerland*

[3] *Unit of Nonlinear Physics and Mathematical Modeling, Department of Engineering, Università Campus Bio-Medico di Roma, Via Álvaro del Portillo 21, I-00128, Rome, Italy*

[4] *Center for Life Nano Science, Sapienza, IIT, Viale Regina Elena 291, I-00161, Rome, Italy*

[5] *Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, Germany*

[6] *Dipartimento di Fisica, Universitá "Tor Vergata", Via della Ricerca Scientifica 1, 00133 Roma, Italy*

[7] *Laboratory of Electron Spectroscopy, ICMP and LACUS, EPFL, CH-1015 Lausanne, Switzerland*

[8] *Department of Physics, University of Fribourg, Chemin du Musée 3, CH-1700 Fribourg, Switzerland*

[9] *Laboratory of Physics of Complex Matter Physics, ICMP, EPFL, Lausanne, Switzerland*

[10] *Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195, Berlin, Germany*

[11] *Departamento Física de Materiales, Universidad del País Vasco, Av. Tolosa 72, E-20018, San Sebastián, Spain*

Bound electronic excitations play a major role in the electrodynamics of insulators and are typically described by the concept of Wannier-Mott and Frenkel excitons [1]. The former represent hydrogenic electron-hole pairs delocalized over several unit cells of a crystal and they occur in materials with high dielectric constant; the latter correspond to a correlated electron-hole pair localized on a single lattice site and they mostly prevail in molecular solids. Between these two extremes, an intermediate type of excitons exists, typically referred to as charge-transfer excitons [2, 3, 4]. A prototypical system in which these rare quasiparticles have been theoretically predicted [5] but never experimentally confirmed is the anatase polymorph of TiO_2, which is one of the most promising material for light-energy conversion applications. Here, we combine angle-resolved photoemission and optical spectroscopies, along with *ab initio* state-of-the-art theoretical calculations, to demonstrate that the direct optical gap of anatase TiO_2 is dominated by a charge-transfer exciton band rising over the continuum of indirect interband transitions. In particular, we find that the lowest exciton possesses a two-dimensional nature and is characterized by a giant binding energy of $\sim$ 300 meV. The universality of these findings is proven in highly defective samples used in light-energy conversion applications, by interrogating these systems out-of-equilibrium via ultrafast two-dimensional UV spectroscopy.

The field of excitonics has gained increased attention in the last years, due to the unique properties that excitons manifest in the conversion and transport of energy. Key to these developments is the ability to exploit exciton physics in materials that are easily fabricated and widely available. Anatase TiO_2 belongs to the class of solids with superior functionalities for the conversion of light into other forms of energy [6, 7, 8], but the nature of its fundamental optical excitations is still unclear and heavily debated. Despite the absence of strong electron-electron correlations, the main source of uncertainty remains the electronic structure and its relationship to the optical spectra. It is therefore pivotal to clarify the single- and two-particle excitation spectra of anatase TiO_2 and to establish the nature of the charge excitations produced upon photon absorption.

Anatase TiO_2 crystallizes in a tetragonal unit cell, built on a network of corner- or edge-sharing TiO_6 octahedra (Fig. 1A). The substantial difference between the lattice constants a = 3.78 Å and c = 9.51 Å results in a strong anisotropy of the electronic and optical properties. First significant steps towards the understanding of the electronic structure were achieved by experimental probes such as angle-resolved photoemission (ARPES) [9, 10] and optical spectroscopies [11, 12, 13]. Recent ARPES studies revealed that the fundamental bandgap of anatase TiO_2 has an indirect character, since the valence band (VB) maximum lies close to the X point of the Brillouin Zone (BZ) and the conduction band (CB) minimum is at the Γ point [9, 10]; consequently, the lowest optical absorption edge can be described in terms of an Urbach tail caused by the phonon-induced localization of excitons [11]. This Urbach tail displays two independent responses for light polarized perpendicular or parallel to the c-axis, as expected from the intrinsic structural anisotropy of the system. Less attention, however, has been paid to the detailed characterization of the optical response above the absorption threshold, where anisotropy effects become more pronounced [12, 13]. In particular, the role played by many-body correlations in the optical properties has remained elusive to experimental probes, leading to a lack of knowledge about the character of the elementary direct charge excitations in this material.

On the theory side, Density Functional Theory (DFT) with many-body perturbation-theory corrections provided a proper description of the material's dielectric function [5, 14, 15, 16]. The

diagonalization of the Bethe-Salpeter Hamiltonian predicted several direct optical transitions at energies well below the direct electronic gap computed at the GW level. The existence of bound excitons in anatase TiO_2 is, however, still awaiting experimental verification, due to the difficulty of measuring the exciton binding energy (E_B). Indeed, conventional experimental methods like optical absorption [17], photoluminescence [18, 19] and magneto-optics [20] are not suitable for an indirect gap material, as the onset of direct band-to-band transitions cannot be identified; moreover, to derive E_B, these methods often rely on approximate models valid only for hydrogenic Wannier-Mott excitons [21].

Here, we address the importance of many-body correlations in anatase TiO_2 by a combination of ARPES and spectroscopic ellipsometry (SE), which allows us to clarify the values of the direct gap for both charged and neutral excitations. Their difference quantifies the interaction energy between the electron and hole generated by photon absorption. Importantly, we demonstrate the existence of strongly bound charge-transfer excitons in anatase TiO_2 and offer an accurate and rigorous estimate of E_B for a direct exciton rising over the continuum of indirect interband transitions, free from assumptions on the nature of the excitonic species under study. The results are corroborated by many-body perturbation-theory, revealing the two-dimensional (2D) spatial distribution of the excitonic states on the (001) plane and the localized nature of the excitonic states for light polarized along the [001] direction. Finally, the stability of the localized bound excitonic states in samples of different character, ranging from single-crystals with varying degrees of defects to colloidal nanoparticles (NPs), is demonstrated by ultrafast 2D spectroscopy in the deep UV spectral range.

The first ingredient to determine the exciton E_B is an accurate evaluation of the direct electronic band gap of the material. To estimate the direct electronic bandgap at the Γ point via ARPES measurements, we inject excess electrons into the CB by inducing oxygen vacancies in an (001)-oriented single-crystal (see Ref. [10]). The new Fermi level at the CB minimum represents a solid reference energy to evaluate the gap with respect to the VB. Indeed, because of the weak electronic correlations, the electron doping provided via oxygen vacancies does not significantly renormalize the electronic structure of anatase TiO_2 [10]. Fig. 1B shows the second derivative ARPES data of the valence states at 20 K between the Γ and X points of

the three-dimensional (3D) BZ. The dashed blue line, highlighting the top of the VB, has been drawn as a guide to the eye. As expected, the onset of the VB occurs close to the X point, and precedes the rise at Γ by $\sim$ 500 meV. The direct inspection of the band structure yields a first highly dispersing band close to the VB upper edges, whose maxima in the vicinity of the X and Γ points sit at -3.55 $\pm$ 0.05 eV and -4.05 $\pm$ 0.05 eV, respectively. Our primary interest here is the value of the direct gap at the Γ point, which can be estimated around 4.05 eV.

Further insight into the nature of the elementary excitations is obtained by SE, which is a very reliable technique for measuring the dielectric function $\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega)$ of a material, allowing the determination of the steady-state absorption and reflectance spectra. Fig. 2A-B show the imaginary part of the dielectric function, $\epsilon_2(\omega)$, measured at 20 K in a (010)-oriented single crystal, with light polarized perpendicular ($\mathbf{E} \perp$ c) and parallel ($\mathbf{E} \parallel$ c) to the c-axis, respectively. We also performed analogous measurements on the single-crystals with various concentrations of oxygen vacancies used in the ARPES experiments and observed no apparent shift in the energy position of the peaks (Fig. S4C-D). In Fig. 2A, the direct absorption is characterized by the presence of a sharp peak at 3.79 eV (I), preceded by a long Urbach tail at lower energies [11]. A second, broader charge excitation (II) lies at 4.61 eV and extends up to 5 eV. The c-axis response (Fig. 2B) is instead characterized by a feature peaking at 4.13 eV (III) with a shoulder at 5 eV. Remarkably, all these excitations are still clearly distinguished at room temperature (RT) (Fig. S5A-B).

To rationalize the experimental data, we calculated the anatase TiO_2 band structure at the GW level of theory (Fig. 1C and Fig. S9). Grey diamonds in Fig. 1C denote the values obtained within GW for VB and CB at the Γ and X points. These are also reported in Fig. 1B for a direct comparison with experimental values from ARPES, showing excellent agreement not only in the overall band shape across the explored region of the BZ, but also for the values of the band gaps, the direct one at the Γ point (E_{dir} = 4.07 eV), and the indirect one from $\sim$ X to Γ (E_{ind} = 3.61 eV). These fully converged GW values are consistently reproduced within different computational approaches (see Fig. S10). The full calculations of the electronic band structure in Fig. 1C also yield important information about the Γ-Z direction, which is crucial for understanding the optical transitions of the material. The CB and VB dispersions between

Γ and Z are approximately parallel, providing a large joint density of states for the occurrence of optical transitions. Here, we show that this peculiar dispersion leads to intense excitonic transitions in the optical absorption spectrum for both considered light polarizations.

To unravel the nature of the optical excitations, we calculated $\epsilon_2(\omega)$ with and without many-body electron-hole correlations. The optical spectra in the uncorrelated-particle picture are obtained within the random-phase approximation (RPA) on top of the GW quasiparticle description, while the absorption spectra are calculated by solving the Bethe-Salpeter Equation (BSE) as implemented in the BerkeleyGW code [22] (see SI for details of the calculations). The simulated spectra with (violet lines) and without (red lines) electron-hole correlations are shown in Fig. 2A-B. By including many-body effects [5, 14, 15, 16], we obtained an excellent agreement with the high-resolution SE spectra, and we performed a detailed analysis of optical transitions in real and reciprocal space. In particular, for $\mathbf{E} \perp$ c, the calculations clearly indicate an absorption maximum (exciton I) at 3.76 eV, well below the direct valence-to-conduction optical transition evaluated at the independent particles level. Such a feature is associated with an exciton band, with the contributing exciton transitions retaining a strongly bound character. The points of the BZ involved in these transitions are the ones along the Γ-Z direction, *i.e.* the region where the parallelism of the CB and VB dispersions largely contributes to the enhancement of the oscillator strength (inset of Fig. 1C, violet arrow). Due to the almost parallel band structure for CB and VB along the Γ-Z direction, the electronic gap at the Γ point (coincident with the continuum absorption rise) can be used as a reference energy in the evaluation of E_B, leading to $E_B \sim 300$ meV. The combination of ARPES and SE here allows us to confirm the bound exciton scenario for the excitation observed at 3.76 eV, and to provide the first experimental estimate of E_B for anatase TiO_2. The energy, shape and reciprocal space contributions for exciton II highlight its bulk-resonance character, most evident as its offset coincides with the RPA @ GW absorption rise. For $\mathbf{E} \parallel$ c, the strong absorption maximum in the computational spectrum at 4.28 eV is also in very good agreement with the SE results. The energy position and peak shape of exciton III indicate a mixing of localized and bulk-resonant contributions, as the continuum onset in RPA @ GW undergoes an intensity enhancement. The k-points contributing to this charge excitation are both localized along the Γ-Z line and

distributed in the whole BZ. The mixed nature of exciton III makes it less straightforward to estimate its E_B. Assuming the RPA @ GW onset at 4.4 eV for $\mathbf{E} \parallel$ c as the reference energy for evaluating E_B (marked as E_{dir} in Fig. 2B), we can estimate $E_B \sim 150$ meV, which is well above the standard exciton binding energies observed in bulk semiconductors.

The large E_B of the excitons in anatase TiO_2 leads us to a deeper investigation of the nature of such anomalously bound quasiparticles. A certain degree of spatial localization has to be expected for these neutral excitations. The character of excitons I and III can be inferred by the spatial distribution of the excitonic wavefunctions in a frozen lattice. In Fig. 3 the relevant excitonic wavefunctions are shown on an isosurface, representing the electronic configuration when the hole of the considered excitonic pair is localized on a specific site. The first bright exciton observed for $\mathbf{E} \perp$ c (first BSE eigenvalue at 3.76 eV) extends two-dimensionally on the (001) plane with a radius of 15 Å, being almost confined to a single atomic plane along the c-axis (Fig. 3A and 3D). Fig. 3B depicts the spatial distribution of the charge excitations associated with peak II of Fig. 2A, showing significant contributions from extended bulk states. Hence, this peak is ascribed to a resonant excitation which does not evolve into a bound state. Its electron wavefunction appears completely delocalized over many lattice constants around the hole. Finally, in Fig. 3C, the linear combination of excitonic wavefunctions (with eigenvalues in the range 4.15-4.35 eV for $\mathbf{E} \parallel$ c) contributing to peak III leads to a highly localized exciton in all three directions, with an average radius of 20 Å.

Typically, the character of an exciton (Wannier-Mott, Frenkel, or an intermediate one) is driven by the combined effects of the crystal geometry and the screening properties. Our analysis reveals that excitons in anatase TiO_2 can be regarded as peculiar quasiparticles with an intermediate character between the Frenkel and Wannier-Mott regimes, since the spatial extent is not restricted to a single unit cell nor is it fully delocalized in space. These excitons result from the interaction between an electron in Ti $3d_{xy}$ orbitals and a hole in O $2p$ orbitals and therefore are of the charge-transfer type. Since our *ab initio* calculations are performed within the frozen lattice approximation, we are neglecting possible phonon effects that play a crucial role in the experimental optical spectra. These arise both in the Urbach tail of the spectrum, dominated by phonon-assisted interband transitions, and in the lineshapes of the

excitonic features, which retain a broad character even at 20 K. Indeed, it is extremely unlikely that such a large linewidth can be explained only via the residual exciton-defect scattering in a high-quality single-crystal. Therefore, we conclude that most of the linewidth is intrinsic and arises from a series of sidebands due to coupling to several modes of the phonon field [23]. As a consequence, such strong exciton-phonon coupling may lead to the formation of polaronic excitons. The presence of an intermediate-to-strong electron-phonon coupling in anatase TiO_2 has already been invoked to explain experimental results naturally pointing to the polaron concept [24, 25]. Recent transport [26], ARPES [10] and Scanning Tunneling Microscopy [27] data detected signatures of large polaronic electrons in the CB of this material, while earlier threshold absorption [11] and photoluminescence [28] data provided clear evidence for Toyozawa's theory of the Urbach tail and radiative recombination of self-trapped excitons [23].

The above description dealt with bulk single-crystals of anatase TiO_2, but in most applications [6, 8, 7], highly-defective samples are used at RT and ambient pressure, *e.g.* NPs or mesoporous films. The possible extension of our findings in such samples would have a strong impact on the common understanding of charge interaction and energy transport in applications involving anatase TiO_2. However, one expects that the carriers released by defects and the local electric fields generated by charged impurities screen the Coulomb interaction in the exciton, leading to the cancellation of the binding forces. Moreover, strong exciton-defect and exciton-impurity scattering can also cause an extreme broadening of the exciton linewidth, hiding the characteristic exciton feature into the continuum of indirect interband excitations. These ideas are reinforced upon direct inspection of the equilibrium absorption spectrum of colloidal anatase TiO_2 NPs, which does not show any obvious signature of excitonic transitions (Fig. S2 and black curve in Fig. 4A) [29, 30]. In addition, this spectrum is strongly affected by the scattering of the incident light. One may therefore question the validity of the above conclusions to the actual systems used in applications.

Here, instead, we demonstrate that the bound excitons existing in the bulk emerge with the same phenomenology in anatase TiO_2 NP at RT. A powerful approach to address the existence of excitonic transitions in colloidal anatase TiO_2 is to interrogate the system out-of-equilibrium via ultrafast 2D UV transient absorption spectroscopy [32], since this technique subtracts the

scattered light and provides a better contrast for resolving hidden features. Typically, the exciton lineshapes can be identified through the pump-induced transparency of the excitonic peak, which is referred to as "exciton bleaching". This nonlinear optical process is intrinsically related to a many-body phenomenon: its manifestation depends not only on the final-state interactions of electron and hole involved in the excitonic state but also on the interaction with all other particles in the material, that can contribute to screening or blocking the excitonic transition [31]. The 2D UV experiment covers the spectral region of the excitonic resonances. Fig. 4A and 4B respectively show the transient absorption (normalized ΔA) of anatase TiO_2 NPs in aqueous solution and the transient reflectivity (ΔR/R) of anatase TiO_2 single-crystals (with varying degrees of oxygen vacancies), as a function of the probe photon energy, at a time delay of 1 ps and for different pump photon energies. All the transients exhibit a negative (bleach) signal, and display two prominent features at 3.88 eV and 4.35 eV. While the former excitation is present at all pump photon energies, the latter becomes more prominent for pump photon energies beyond 4.1 eV, *i.e.* in correspondence with the threshold for accessing the c-axis optical response. Both features have a relatively narrow width of 250 meV.

To assign them, we computed ΔA from the ΔR/R measured in anatase TiO_2 single-crystals (see details in the SI). Fig. 4B shows the derived ΔA along the a- and c-axis, for a 4.4 eV pump photon energy and a time delay of 1 ps. In these separate polarization channels, two negative features appear at 3.88 eV and 4.32 eV, similarly to the ΔA spectra of NPs. It appears that these energies match those of the first excitonic peaks in the equilibrium absorption spectrum along the a- and c-axis (note that the values of the absorption peaks slightly differ from the peaks in the dielectric function). Hence, this analysis leads us to disentangle the a-axis and c-axis contributions in the ΔA response of the NPs. Both excitons are found to appear in the latter due to the lack of anisotropy in the randomly oriented NPs in solution. Therefore, the excitonic features are also present in the equilibrium absorption spectrum of anatase TiO_2 NPs, but they are obscured by the strong light scattering.

Establishing unambiguously the existence of excitons in bulk single crystals of anatase TiO_2 and more importantly in NPs, provides novel perspectives on the optical response of this widely used material. In particular, the 2D wavefunction of the lowest charge-transfer exciton represents

a fascinating aspect, because it does not occur in a quasi-2D material. In the last years, several 2D excitons were discovered in materials such as hexagonal boron nitride [33] and transition metal dichalcogenides [34, 35, 36], which are layered 2D systems held together by Van der Waals forces to form a 3D lattice. The lack of strong interaction between the layers yields to a physical behaviour which reflects the predominantly 2D structure. The situation of anatase TiO_2 is radically different in that a genuine 3D material exhibits a 2D exciton wavefunction on a specific lattice plane.

Besides the fundamental interest in this exotic species, the correct description of the equilibrium optical spectrum can improve the comprehension of the enhanced photocatalytic activity of anatase TiO_2 (001) surface compared to other facets: to this purpose, it is crucial to establish how molecules adsorbed to this surface react to specific charge excitations generated by tunable incident light. From the point of view of devices, the newly discovered bound excitons may provide a significant source of optical nonlinearity, paving the route for the development of electro-optical or all-optical switches in the UV. Also engineered nanostructures exposing a large percentage of (001) facets can be useful in guiding the energy at the nanoscale in a selective way [37, 38, 39]. Finally, due to the important contribution that phonons play in the exciton linewidth and lineshape, we expect that the optical properties of anatase TiO_2 can be effectively altered by tuning the exciton-phonon coupling, *e.g.* through the applications of mechanical strain or via dielectric screening. In this regard, new insights from many-body theory will be crucial for evaluating the transport of these excitonic species, their coupling to the vibrational degrees of freedom and their reaction to various external stimuli.

Acknowledgments: We thank Prof. Fabrizio Carbone for the useful discussions and Dr. Luca Moreschini for the technical support. We acknowledge financial support from the Swiss NSF via the NCCR:MUST and the contracts No. 206021-157773 and 200020-153660, the European Research Council Advanced Grant DYNamo ERC-2010-AdG-267374, Spanish Grant FIS2013-46159-C3-1-P, Grupos Consolidados del Gobierno Vasco (IT578-13), COST Actions CM1204 (XLIC), MP1306 (EUSpec), and the Austrian Science Fund (FWF P25739-N27).

References

[1] M. Ueta, *Excitonic Processes in Solids* (Springer-Verlag, Berlin, 1986).

[2] S. Knox, *Theory of Excitons* (Academic Press, New York, 1963).

[3] E. Collart *et al.*, Localized and delocalized excitons: Resonant inelastic x-ray scattering in $La_{2-x}Sr_xNiO_4$ and $La_{2-x}Sr_xCuO_4$. *Phys. Rev. Lett.* **96**, 157004 (2006).

[4] D. S. Ellis *et al.*, Charge-transfer exciton in La_2CuO_4 probed with resonant inelastic x-ray scattering. *Phys. Rev. B* **77**, 060501 (2008).

[5] L. Chiodo *et al.*, Self-energy and excitonic effects in the electronic and optical properties of TiO_2 crystalline phases. *Phys. Rev. B* **82**, 045207 (2010).

[6] A. Fujishima, K. Honda, Electrochemical photolysis of water at a semiconductor electrode. *Nature* **238**, 37-38 (1972).

[7] B. O'Regan, M. Grätzel, A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO_2 films. *Nature* **353**, 737-740 (1991).

[8] E. Pelizzetti, C. Minero, Mechanism of the photo-oxidative degradation of organic pollutants over TiO_2 particles. *Electroch. Acta* **38**, 47-55 (1993).

[9] M. Emori *et al.*, Electronic structure of epitaxial anatase TiO_2 films: Angle-resolved photoelectron spectroscopy study. *Phys. Rev. B* **85**, 035129 (2012).

[10] S. Moser *et al.*, Tunable polaronic conduction in anatase TiO_2. *Phys. Rev. Lett.* **110**, 196403 (2013).

[11] H. Tang, F. Lévy, H. Berger, P. E. Schmid, Urbach tail of anatase TiO_2. *Phys. Rev. B* **52**, 11, 7771-7774 (1995).

[12] N. Hosaka, T. Sekiya, C. Satoko, S. Kurita, Optical properties of single-crystal anatase TiO_2. *J. Phys. Soc. Jap.* **66**, 3, 877-880 (1997).

[13] T. Sekiya, M. Igarashi, S. Kurita, S. Takekawa, M. Fujisawa, Structure dependence of reflection spectra of TiO_2 single crystals. *J. of Elec. Spectr. and Rel. Phen.* **92** 247-250 (1998).

[14] H. M. Lawler *et al.*, Optical to UV spectra and birefringence of SiO_2 and TiO_2: First-principles calculations with excitonic effects. *Phys. Rev. B* **78**, 205108 (2008).

[15] W. Kang, M. S. Hybertsen, Quasiparticle and optical properties of rutile and anatase TiO_2. *Phys. Rev. B* **82**, 085203 (2010).

[16] M. Landmann, E. Rauls, W. G. Schmidt, The electronic structure and optical response of rutile, anatase and brookite TiO_2. *J. Phys. Condens. Matter* **24**, 195503 (2012).

[17] J. F. Muth *et al.*, Absorption coefficient, energy gap, exciton binding energy, and recombination lifetime of GaN obtained from transmission measurements. *Appl. Phys. Lett.* **71**, 18, 2572 (1997).

[18] D. D. Sell, Resolved free-exciton transitions in the optical-absorption spectrum of GaAs. *Phys. Rev. B* **6**, 10, 3750 (1972).

[19] R. Rinaldi *et al.*, Exciton binding energy in GaAs V-shaped quantum wires. *Phys. Rev. Lett.* **73**, 21, 2899 (1994).

[20] J. C. Maan, G. Belle, A. Fasolino, M. Altarelli, K. Ploog, Magneto-optical determination of exciton binding energy in GaAs-$Ga_{1-x}Al_xAs$ quantum wells. *Phys. Rev. B* **30**, 4, 2253 (1984).

[21] R. J. Elliott, Intensity of optical absorption by excitons. *Phys. Rev.* **108**, 6, 1384 (1957).

[22] J. Deslippe *et al.*, BerkeleyGW: A Massively Parallel Computer Package for the Calculation of the Quasiparticle and Optical Properties of Materials and Nanostructures. *Comput. Phys. Commun.* **183**, 1269 (2012).

[23] Y. Toyozawa, *Optical Processes in Solids* (Cambridge Univ. Press, New York, 2003), chap. 10.

[24] N. A. Deskins, N. Dupuis, Electron transport via polaron hopping in bulk TiO_2: A density functional theory characterization. *Phys. Rev. B* **75**, 195212 (2007).

[25] C. Di Valentin, A. Selloni, Bulk and surface polarons in photoexcited anatase TiO_2. *J. Phys. Chem. Lett.* **2**, 2223-2228 (2011).

[26] J. Jaćimović *et al.*, Pressure dependence of the large-polaron transport in anatase TiO_2 single crystals. *Europhys. Lett.* **99**, 57005 (2012).

[27] M. Setvin *et al.*, Direct view at excess electrons in TiO_2 rutile and anatase. *Phys. Rev. Lett.* **113**, 086402 (2014).

[28] H. Tang, H. Berger, P. E. Schmid, F. Lévy, G. Burri, Photoluminescence in TiO_2 anatase single crystals. *Sol. Stat. Comm.* **87**, 9, 847 (1993).

[29] N. Serpone, D. Lawless, R. Khairutdinov, Size effects on the photophysical properties of colloidal anatase TiO_2 particles: Size quantization or direct transitions in this indirect semiconductor? *J. Phys. Chem.* **99**, 16646 (1995).

[30] S. Monticone, R. Tufeu, A. V. Kanaev, E. Scolan, C. Sanchez, Quantum size effect in TiO_2 nanoparticles: Does it exist? *Appl. Surf. Sci.* **162**, 565 (2000).

[31] H. Haug, S. Schmitt-Rink, Basic mechanisms of the optical nonlinearities of semiconductors near the band edge. *J. Opt. Soc. Am. B* **2**, 7, 1135 (1985).

[32] G. Auböck, C. Consani, F. van Mourik, M. Chergui, Ultrabroadband femtosecond two-dimensional ultraviolet transient absorption. *Opt. Lett.* **12**, 2337 (2012).

[33] S. Galambosi *et al.*, Anisotropic excitonic effects in the energy loss function of hexagonal boron nitride. *Phys. Rev. B* **83**, 081413(R) (2011).

[34] D. Y. Qiu, F. H. da Jornada, S. G. Louie, Optical spectrum of MoS_2: Many-body effects and diversity of exciton states. *Phys. Rev. Lett.* **111**, 216805 (2013).

[35] K. He *et al.*, Tightly bound excitons in monolayer WSe_2. *Phys. Rev. Lett.* **113**, 026803 (2014).

[36] M. M. Ugeda *et al.*, Giant bandgap renormalization and excitonic effects in a monolayer transition metal dichalcogenide semiconductor. *Nat. Mat.* **13**, 1091-1095 (2014).

[37] H. G. Yang *et al.*, Anatase TiO_2 single crystals with a large percentage of reactive facets. *Nature* **453**, 638-642 (2008).

[38] G. Giorgi, M. Palummo, L. Chiodo, K. Yamashita, Excitons at the (001) surface of anatase: Spatial behavior and optical signatures. *Phys. Rev. B* **84**, 073404 (2011).

[39] M. Palummo, G. Giorgi, L. Chiodo, A. Rubio, K. Yamashita, The nature of radiative transitions in TiO_2-based nanosheets. *J. Phys. Chem. C* **116**, 18495 (2012).

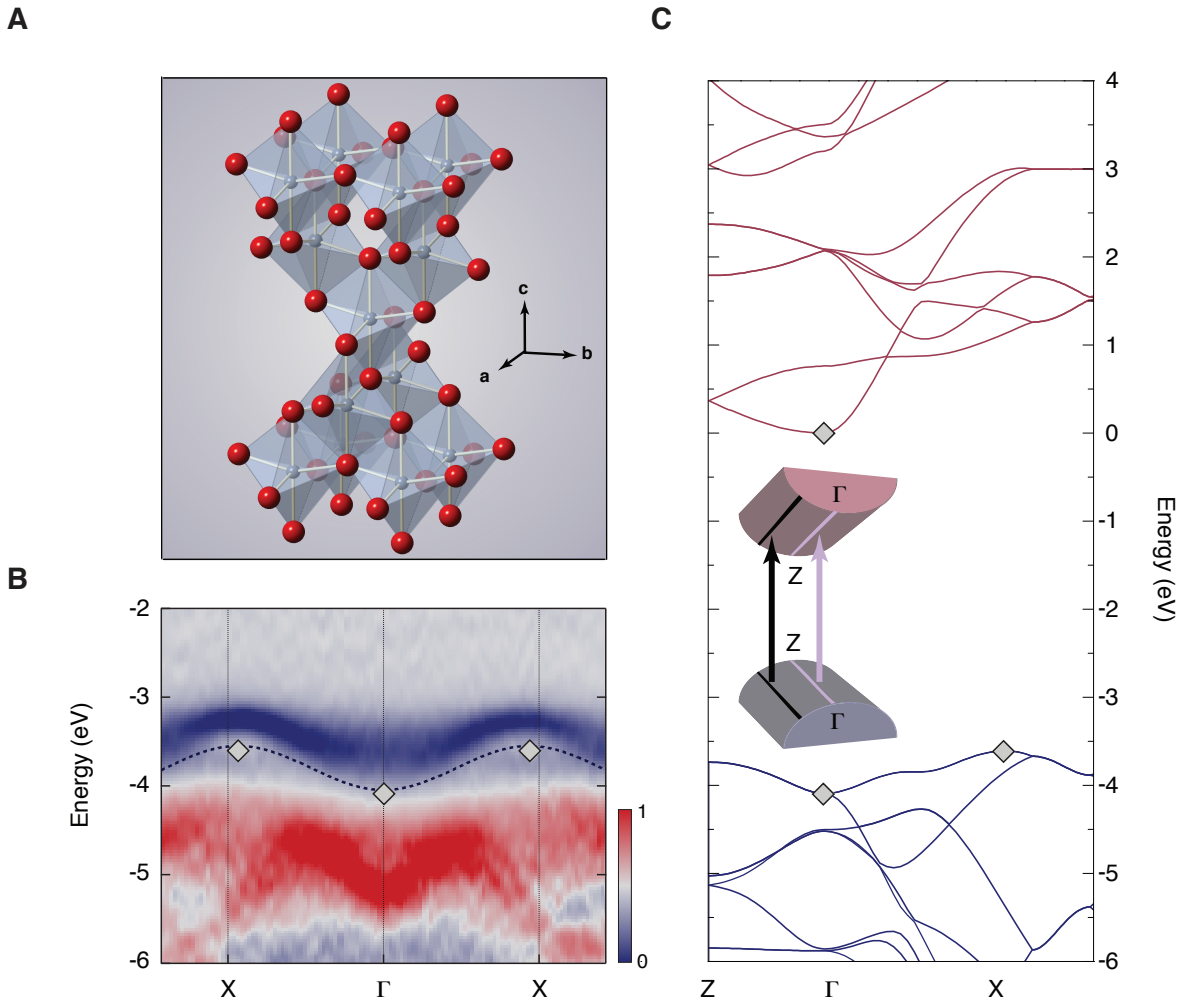

Figure 1: (**A**) Crystallographic structure of anatase TiO_2 with highlighted TiO_6 polyhedra. Blue atoms represent titanium, red atoms represent oxygen. (**B**) Second derivative ARPES data of the electronic structure at the top of the VB between Γ and X. Dashed blue lines are reported as a guide to the eye. Grey diamonds indicate the values obtained with the GW corrections. The spectrum is referenced to the Fermi energy, which lies at the bottom of the CB at Γ. (**C**) Electronic band structure of bulk anatase TiO_2 along the Z - Γ - X directions of the first BZ. The lines indicate the DFT calculations corrected by the GW values. Grey diamond dots indicate the values obtained with the GW corrections at the high symmetry points, while the values along the high symmetry directions are obtained by correcting with linearly interpolated values. The inset shows a schematic representation of the three-dimensional electronic structure between the Γ and Z points of the BZ, where the CB and VB dispersions are approximately parallel. The violet and black arrow highlight the region involved in the optical transitions which contribute to exciton I and III respectively. The gap at Γ is a good reference to highlight the bound nature of exciton I in Fig. 2.

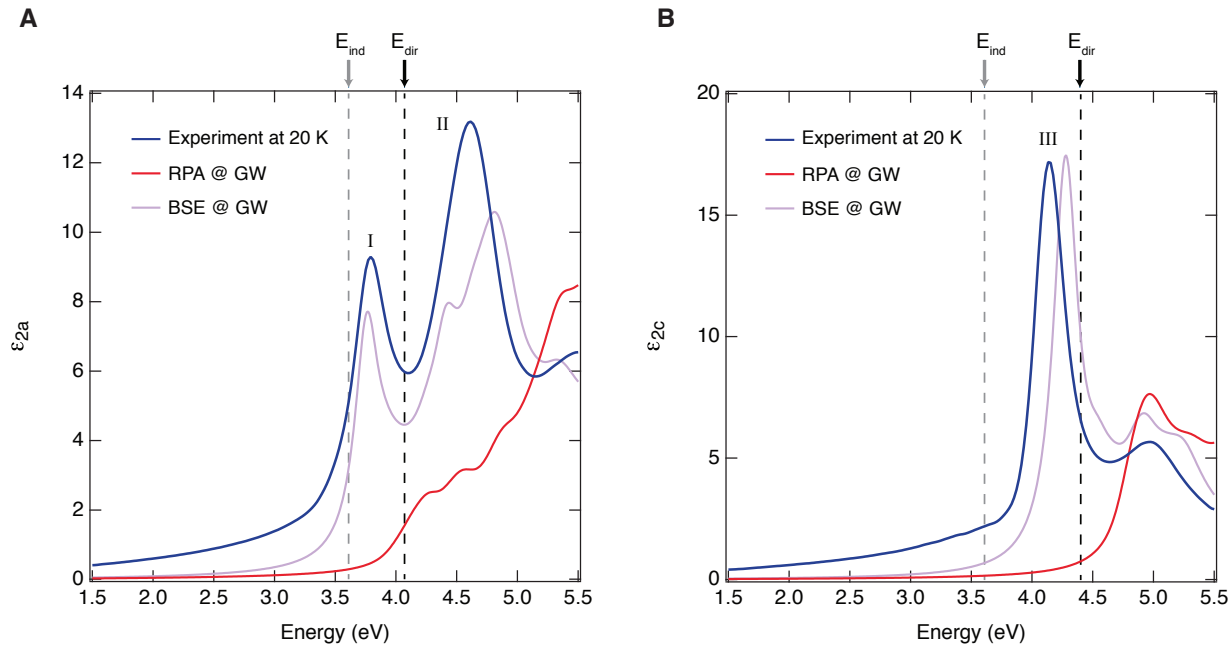

Figure 2: Imaginary part of the dielectric function at 20 K with the electric field polarized along (**A**) the a-axis and (**B**) the c-axis. The experimental data measured by spectroscopic ellipsometry are reported in blue, the calculated spectra in the RPA @ GW scheme in red and the calculated spectra in the BSE @ GW scheme in violet. The quasiparticle indirect gap E_{ind} = 3.61 eV is indicated by a dashed grey vertical line; the direct gaps E_{dir} = 4.07 eV (at the Γ point) for **E** ⊥ c and E_{dir} = 4.4 eV (coincident with the onset of the RPA @ GW) for **E** ∥ c are indicated by dashed black vertical lines in Fig. (**A**) and (**B**) respectively.

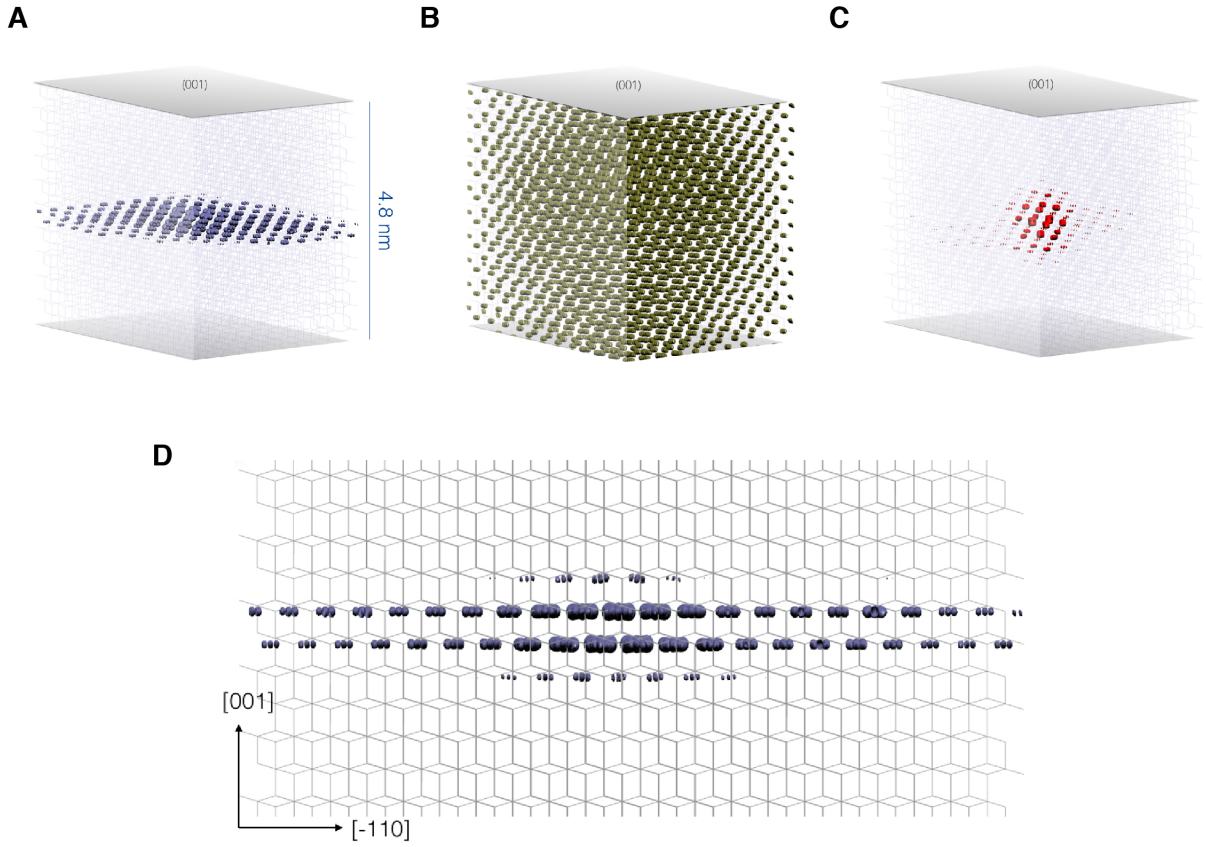

Figure 3: Isosurface representation for the wavefunction of the fundamental charge excitations in anatase TiO_2. (**A**) Charge-transfer exciton I at 3.76 eV. (**B**) Resonance II at 4.37 eV. (**C**) Charge-transfer exciton III at 4.25 eV. (**D**) Side-view of the charge-transfer exciton I at 3.76 eV.

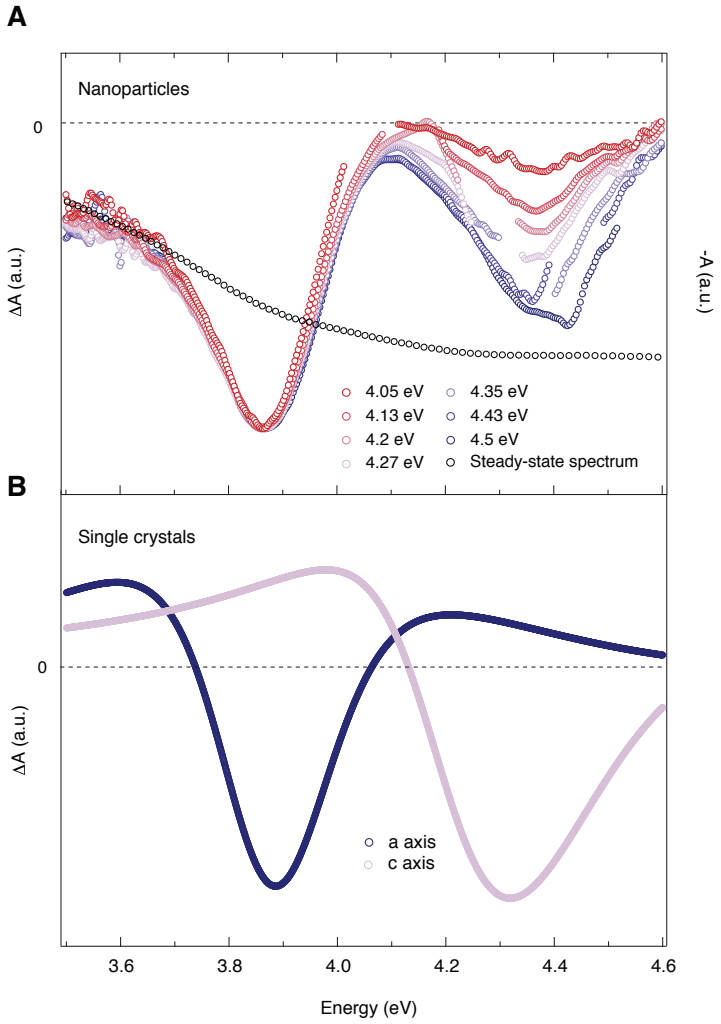

Figure 4: (**A**) Normalized transient absorption spectra of room-temperature colloidal solution of anatase TiO_2 NPs at a fixed time-delay of 1 ps and for different pump photon energies (indicated in the figure). Each trace is normalized with respect to the minimum of the main feature at 3.88 eV. For comparison, the black trace shows the inverted steady-state absorption spectrum. (**B**) Transient absorption spectrum of room-temperature anatase TiO_2 single-crystals along the a- and c-axis at a fixed time delay of 1 ps. For this experiment, the pump photon energy is 4.4 eV.

S1 Experimental methods

S1.1 Single crystal growth and characterization

High-quality single-crystals of anatase TiO_2 were produced by a chemical transport method from anatase powder and NH_4Cl as transport agent, similar to the procedure described in Ref. [40]. In detail, 0.5 g of high-purity anatase powder were sealed in a 3 mm thick, 2 cm large and 20 cm long quartz ampoule together with 150 mg of NH_4Cl, previously dried at 60 °C under dynamic vacuum for one night, and 400 mbar of electronic grade HCl. The ampoules were placed in a horizontal tubular two-zone furnace and heated very slowly to 740 °C at the source, and 610 °C at the deposition zone. After two weeks, millimeter-sized crystals with a bi-pyramidal shape were collected and cut into rectangular bars (typically $0.8 \times 0.6 \times 0.15$ mm^3). We characterized the raw anatase single-crystals through transport and Reflection High-Energy Electron Diffraction (RHEED) measurements, as in Ref. [10]. Copper-doped anatase TiO_2 single-crystals were obtained by annealing raw anatase single-crystals in O_2 at 700 °C for 6 days in the presence of Cu vapors. The reduced form of anatase TiO_2 was instead obtained by annealing the raw anatase TiO_2 crystals at 700 °C for 10 days under 950 mbar of CO.

The reduced anatase TiO_2 single-crystal was also polished along a (010)-oriented surface, in order to be used in the spectroscopic ellipsometry (SE) measurement. The surface roughness of this crystal was estimated via Atomic Force Microscopy (AFM). Fig. S1A-B shows two images taken under the AFM. The average surface roughness of the polished surface was found around 0.9 nm.

S1.2 Nanoparticles synthesis and characterization

TiO_2 nanoparticles (NPs) were prepared by the sol-gel method [41]. The synthesis was carried out in a glove box under argon atmosphere. We used titanium isopropoxide (Sigma Aldrich, 99.999% purity) as precursor, which was mixed with 10 ml of 2-propanol. This mixture was added dropwise under vigorous stirring to cold acidic water (2 °C, 250 ml H_2O, 18 MΩ, mixed with 80 ml glacial acetic acid, final pH 2). At the beginning the mixture looked turbid, but after stirring it in an ice bath for 12 hours, it became transparent as the amorphous NPs

were formed. Half of the mixture was left stirring for days to stabilize the NPs. The other half was peptized at 80 °C for about 2 hours until the liquid turned into a transparent gel. The gel was autoclaved at 230 °C for 12 hours. During this process the previous amorphous sample became denser and underwent a phase transition, resulting in anatase TiO_2 NPs. After the autoclave the NPs have precipitated to the bottom of the container. They were separated from the supernatant and added to 100 ml acidic water (pH 2) to obtain a white colloidal solution with a final concentration of ca. 337 mM. In Ref. [42], we reported the details of the sample characterization by means of X-ray diffraction and transmission electron microscopy. These techniques enabled us to demonstrate the good quality of the anatase phase and the spherical shape (with an average diameter of approximately 25 nm) of the NPs.

The steady-state absorption spectrum of the colloidal solution of anatase TiO_2 was recorded at room temperature using a commercial UV-VIS-NIR spectrometer (Shimadzu, UV-3600). Before measuring the absorption spectrum of the sample, a reference spectrum of the pure solvent (acidic water, pH 2) was taken to check its transparency in the investigated spectral range. Fig. S2 shows the absorption spectrum of the colloidal solution in a range between 2.6 eV and 4.8 eV. This spectrum is comparable to the one reported in previous works [29, 30].

S1.3 Angle-resolved photoemission spectroscopy

The Angle-Resolved Photoemission Spectroscopy (ARPES) measurements were performed at the Electronic Structure Factory endstation on beamline 7.0.1 at the Advanced Light Source, Berkeley, USA. A raw anatase TiO_2 single-crystal was polished and cleaned in a buffered 5% fluoridic acid solution before introducing it into the UHV system ($<10^{-10}$ mbar). The crystal was annealed in 35 mbar of oxygen at 400 °C for 30 minutes before the ARPES experiments.

S1.4 Steady-state spectroscopic ellipsometry

Using SE, we measured the complex dielectric function of the sample, covering the spectral range from 1.5 eV to 5.5 eV. We used a Woollam VASE ellipsometer at the University of Fribourg, Switzerland. The sample was mounted in a helium flow cryostat, allowing measurements from room temperature down to 10 K. When at cryogenic temperatures, these latter

measurements were performed at $<10^{-8}$ mbar to prevent measurable ice-condensation onto the sample. Anisotropy corrections were performed using standard numerical procedures [44] and diffraction effects at low frequency were accounted for using the procedure developed by Humliček *et al.* [45]. The data have been further corrected to account for the surface roughness of the single-crystal, which was estimated around 0.9 nm by means of AFM.

S1.5 Ultrafast two-dimensional UV spectroscopy

The ultrafast optical experiments have been performed using a novel set-up of tunable UV pump and broadband UV probe, described in detail in [32]. A 20 kHz Ti:sapphire regenerative amplifier (KMLabs, Halcyon + Wyvern500), providing pulses at 1.55 eV, with typically 0.6 mJ energy and around 50 fs duration, pumps a noncollinear optical parametric amplifier (NOPA) (TOPAS white - Light Conversion) to generate sub-90 fs visible pulses (1.77 - 2.3 eV range). The typical output energy per pulse is 13 μJ. Around 60% of the output of the NOPA is used to generate the narrowband pump pulses. The visible beam, after passing through a chopper, operating at 10 kHz and phase-locked to the laser system, is focused onto a 2 mm thick BBO crystal for nonlinear frequency doubling. The pump photon energy is controlled by the rotation of the crystal around the ordinary axis and can be tuned in a spectral range up to $\sim$ 0.9 eV ($\sim$ 60 nm) wide. The typical pump bandwidth is 0.02 eV (1.5 nm) and the maximum excitation energy is about 120 nJ. The pump power is recorded on a shot-to-shot basis by a calibrated photodiode for each pump photon energy, allowing for the normalization of the data for the pump power. The remaining NOPA output is used to generate the broadband UV probe pulses with $\sim$ 1.3 eV ($\sim$ 100 nm) bandwidth through an achromatic doubling scheme. Pump and probe pulses, which have the same polarization, are focused onto the sample, where they are spatially and temporally overlapped. The typical spot size of the pump and the probe are 100 μm and 40 μm FWHM respectively, resulting in a homogeneous illumination of the probed region.

This setup can be used either in transmission or in reflection configuration. The anatase TiO_2 single-crystals were studied by detecting their transient reflectivity (ΔR/R) upon photoexcitation, while the NPs were investigated by recording their transient absorption (ΔA). In the case of the measurements on the anatase TiO_2 single-crystals, the specimens were mounted on a rotating

sample holder, in order to explore the $\Delta R/R$ response along the desired crystalline axis. The measurements were performed on the three different classes of samples (raw anatase, copper doped anatase and reduced anatase). The crystals were polished in order to obtain optical-grade surfaces. The portion of the probe beam reflected by the surface of the crystal has been detected and the time evolution of the difference in the UV probe reflection with and without the pump pulse reconstructed. All the experiments have been performed at room temperature. Concerning the measurements on the anatase NPs, the sample consisted of anatase TiO_2 NPs dispersed in an aqueous solution (20% acetic acid and 80% water) to avoid interparticle charge-transfer. The colloidal solution circulated into a 0.2 mm thick quartz flow-cell to prevent photo-damage and its concentration has been adjusted to provide an optical density of approximately 0.4. Alternatively, for limiting the issues of cross-phase modulation in the sample and achieving a better time resolution, the NPs solution could be made flowing in the form of 0.2 mm thick jet. The probe was measured after its transmission through the sample and its detection synchronized with the laser repetition rate. The difference of the probe absorption with and without the pump pulse has been measured at different time delays between the pump and the probe, thanks to a motorized delay line in the probe path.

After the sample, the transmitted/reflected broadband probe beam is focused in a multi-mode optical fiber (100 μm), which is coupled to the entrance slit of a 0.25 m imaging spectrograph (Chromex 250is). The beam is dispersed by a 150 gr/mm holographic grating and imaged onto a multichannel detector consisting of a 512 pixel CMOS linear sensor (Hamamatsu S11105, 12.5 $\times$ 250 μm pixel size) with up to 50 MHz pixel readout, so the maximum read-out rate per spectrum (almost 100 kHz) allows us to perform shot to-shot detection easily. The described experimental setup typically offers a time resolution of 150 fs, but can be improved to 80 fs with the adoption of a prism compressor on the pump pulses, at the expenses of a reduction of the probe bandwidth in the achromatic frequency doubling scheme.

S2 Additional experimental results

S2.1 Angle Resolved Photoemission Spectroscopy

The ARPES experiment has been performed on an (001)-oriented single-crystal at T = 20 K. The presence of oxygen vacancies leads the injection of electrons in the conduction band (CB), causing a shift of the Fermi level (E_F) to approximately 80 meV above the bottom of the CB at Γ. Fig. S3A is a constant energy map measured at the E_F with a photon energy $h\nu$ = 85 eV. The partially resolved circular contours at the center of the BZ correspond to a small electron pocket at the bottom of the CB. Further details on the characterization of this electron pocket are reported in Ref. [10]. Fig. S3B shows the ARPES data of the valence states at 20 K between the Γ and X points of the three-dimensional BZ. The data have been measured with a photon energy $h\nu$ = 128 eV. Fig. 1B in the main text represents the second-derivative of these data along the energy direction. This method has been proven effective for resolving dispersive features in ARPES maps [43].

S2.2 Steady-state spectroscopic ellipsometry

The SE measurements have been performed on two classes of (010)-oriented single-crystals, namely reduced anatase (transparent crystal) and raw anatase (black crystal).
The reflectivity response derived from SE on the transparent (violet lines) and black (blue lines) single crystals is plotted in Fig. S4A-B together with the data of Hosaka *et al.* (red lines) [12]. The reflectivities for light polarized along the a- and c-axis of the crystals are shown in Fig. S4A and S4B, respectively. In order to establish a direct comparison with the normal-incidence reflectivity data of Hosaka *et al.* [12], all reported traces have been measured at 100 K. We observe the overall agreement of our results with the ones reported in literature, with slight shifts in the reflectivity peak energies. However, given the sharpness of the reflectivity lineshapes, we can also deduce that the quality of our transparent crystal is much higher than the one employed by Hosaka *et al.* [12]. Indeed, the reflectivity spectrum of the latter is closer to the response measured on our black crystal, which has a large percentage of oxygen vacancies.

Fig. S4C-D display the normalized absorption spectra of our anatase TiO_2 single-crystals

at 300 K. The a- and c-axis responses are respectively shown in Fig. S4C and S4D. Here, our interest is to reveal the robustness of the excitonic peaks even at room temperature and to highlight the role played by the presence of oxygen vacancies in the black sample. Hence, for this purpose, the data have been normalized with respect to the lowest direct exciton peaks. We observe that in both the a- and c-axis absorption spectra the exciton peak energies are not renormalized by the presence of oxygen vacancies, while the linewidth becomes clearly broader in the black crystal. We associate this effect with the higher exciton-defect scattering, which reduces the lifetime of the excitonic species.

Fig. S5A and S5B display $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$ calculated by applying the KK analysis on the reflectivity spectrum at room temperature. This technique represents a valuable test to evaluate the consistency with our original $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$ values. We observe that the effect of the KK transformation on both the in-plane and out-of-plane $\epsilon_2(\omega)$ is to modify the lineshape and the intensity of the peaks and to change the spectral-weight in the Urbach tail. This behavior proves that the KK analysis, even performed on the broadest possible spectral range, is not a precise approach to treat the normal-incidence reflectivity data.

S2.3 Ultrafast transient-reflectivity of single crystals

The femtosecond transient-reflectivity experiments have been performed on three different types of (001)-oriented single-crystals, namely raw anatase (black crystal), copper-doped anatase (red crystal) and reduced anatase (transparent crystal), and on a (010)-oriented raw anatase single-crystal. In the first experiment we monitored the ΔR/R signal in a broadband UV range (3.75 - 4.35 eV) for the three classes of (001)-oriented samples along the ab planes. Both pump and probe polarizations were set in a parallel configuration. An isotropic optical response is found when the (001)-oriented crystals are rotated about the c axis. The pump energy was set at 4.1 eV, in order to selectively perturb the spectral region above the first excitonic peak.

Fig. S6A compares the transient spectrum obtained from the three specimens at a fixed time-delay of 6 ps. The ΔR/R spectrum is bipolar, showing an inversion point around 3.96 eV. The inflection that is present in the signal obtained from the Cu-doped sample around 4.1 eV is an artefact produced by the scattering of the pump beam. A similar response can be found for

all the temporal delays between pump and probe up to 1 ns. The main difference is displayed by the intensity of the signal at long time delays, since it depends on the rate of the carrier recombination process. Fig S6B shows three temporal traces up to 100 ps probed around 3.82 eV, in which this effect is most visible. The decay of the nonequilibrium signal in the black crystal is faster than the other responses, since the increased density of in-gap states facilitates charge carrier recombination across the bandgap. The detailed analysis of the kinetics will be subject of a future separate publication.

A second set of experiments was subsequently performed on (001)- and (010)-oriented black single-crystals of anatase TiO_2, in order to access the anisotropic dynamics along a- and c-axes. Fig. S7A and S7C display the transient-reflectivity maps of the in-plane and out-of-plane response of the (010)-oriented single-crystal, as a function of the probe energy and of the time delay between pump and probe. Although the kinetics have been registered up to 1 ns, the two maps are displayed up to 10 ps. The a-axis response was measured upon photoexcitation at 4.4 eV, with the broadband probe covering the range 3.7 eV - 4.65 eV. These results have also been reproduced by polarizing the beams along the [100] axis of the (001)-oriented samples. The higher-energy region of the spectrum does not evolve in time, remaining unaffected by the photoexcitation process. The c-axis response (Fig. S7C) is obtained through the measurement of the (010)-oriented single-crystal with a pump and probe polarizations set along the c-axis. Also in this case the pump energy is at 4.4 eV, but the probed range is shifted to 4.05 eV - 4.8 eV. The transient spectrum of Fig. S7D strongly differs from the in-plane one, consisting of a negative contribution set around 4.28 eV and a tail extending to 4.6 eV in the high-energy range. A conclusive $\Delta R/R$ map with both pump (at 4.4 eV) and probe (between 3.7 and 4.6 eV) polarized along the c-axis of the (010)-oriented single-crystal is reported in Fig. S8. This demonstrates the absence of c-axis spectral features at low energies and confirms the finding of the pump-probe experiment along the c-axis reported in Fig. S7C. The difference in terms of intensity with respect to Fig. S7C has to be attributed to the reduced pump intensity (of $\sim$ 1/3) due to constraints in the generation of the broadband probe beam.

S3 Theory and data analysis

S3.1 Extracting ΔA from ΔR/R in single-crystals

In our ultrafast measurements, we probed ΔR/R of anatase TiO_2 single-crystals and ΔA of a colloidal solution of NPs. Thus, it is useful to analyse the time-resolved dynamics of the single-crystals in terms of their transient-absorption, in order to establish a link with the data on NPs. However, while ΔA is directly proportional to $\Delta\epsilon_2$, ΔR/R has a non-trivial relationship with both $\Delta\epsilon_1$ and $\Delta\epsilon_2$ in the probed spectral range. Indeed, in the UV, the real and imaginary part of the dielectric constant have rather similar absolute values. For this reason, the optical reflectivity is equally sensitive to the reactive and the absorptive components of the dielectric function.

Hence, in order to calculate the pump-induced evolution of the ΔA from the ΔR/R data, we proceed as follows. We model the static ellipsometry data using a set of Lorentz oscillators, we calculate the equilibrium reflectivity (R_0), and we fit the measured transient reflectance $R_{exp}(t)/R_{exp}$ with a differential model (R(t) - R_0)/R_0, where R(t) is a model for the perturbed reflectivity obtained by variation of the parameters used to fit the equilibrium data as a function of the pump-probe delay t. We adopted this approach to treat our ΔR/R data in order to avoid possible systematic errors that can be produced by the typical analysis through Kramers-Kronig transformations. The SE spectra were fitted using the RefFIT program [46]. The dielectric function is given by:

$$\epsilon(\omega) = \epsilon_\infty + \sum_i \frac{\omega_{pi}^2}{\omega_{0i}^2 - \omega^2 - i\Gamma_i\omega} \tag{1}$$

where ϵ_∞ is the high-frequency dielectric constant, which takes into account the contribution of all oscillators at very high frequencies and ω_{pi}^2, ω_{0i}^2, Γ_i are, respectively, the plasma frequency, the transverse frequency and the linewidth (scattering rate) of the i-th Lorentz oscillator. The absorbance is then given by: $A(\omega) = \omega\ \mathrm{Im}\sqrt{\epsilon}$.

For the fitting of the transient data, the Lorentz oscillators in our experimental range are allowed to change in order to reproduce the dynamical reflectivity. This procedure enables to extract the transient dielectric function $\Delta\epsilon(\omega,\tau) = \Delta\epsilon_1(\omega,\tau) + i\Delta\epsilon_2(\omega,\tau)$ and finally leads to the evaluation of the ΔA for the single-crystals.

S3.2 *Ab initio* calculations: Computational details

Many-body perturbation theory was employed to compute the band structure and the dielectric response of bulk anatase TiO_2. The GW and Bethe Salpeter Equation (BSE) calculations [47, 48, 49], used for evaluating the electronic and optical properties respectively, were performed on-top of eigenvalues and eigenfunctions obtained from density functional theory (DFT). We used the planewave pseudopotential implementation of DFT as provided by the package Quantum Espresso. GW and BSE calculations were performed with two different implementations, the one in the BerkeleyGW package [22] and the one in Yambo [50].

DFT calculations rely on the generalized gradient approximation as in the Perdew-Burke-Ernzerhof (PBE) scheme for the exchange-correlation functional. The Ti norm-conserving pseudopotential includes semicore states $3s$ and $3p$. While standard structural and electronic quantities are already converged at the DFT level of description with an energy cutoff of 90 Ry, the energy cutoff has here been raised up to 160 Ry to properly include the high number of bands necessary to reach convergence for the many-body evaluated properties. Bulk anatase TiO_2 was modelled on a body-centred tetragonal lattice containing 2 Ti atoms and 4 O atoms (primitive cell) with lattice parameters (optimised at the PBE-DFT level) a = b = 3.79 Å and c = 9.66 Å. These optimized lattice constants are very close to the experimental values a = b = 3.78 Å and c = 9.51 Å.

The ground state electronic density is properly described with a coarse 4×4×4 k-point grid for sampling of the BZ. Finer grids are really needed to get a proper description of screening effects and many-body corrections at the GW and BSE levels of theory. GW calculations were performed by evaluating the quasiparticle corrections to DFT eigenvalues at the G_0W_0 one-shot level of theory [47]. The coarse 4×4×4 k-point grid sampling is sufficient for the description of the quasiparticle corrected band structure. The full converged calculations shown in the main text were obtained with BerkeleyGW. Below we also show that the results are code-independent by comparing the BerkeleyGW results with the ones obtained with the Yambo code from a non-converged case (as full convergence was not possible to achieve with the present implementation in Yambo).

The GW results were then interpolated into a finer grid of 16×16×16. For the solution of

the BSE, we performed DFT calculations on a finer and randomly shifted grid, up to 16×16×16 k-points (4096 irreducible k-points), within BerkeleyGW.

In BerkeleyGW, for the computation of the polarizability and inverse dielectric matrices, we employed a total of 2474 conduction bands and G-vectors with kinetic energies up to 46 Ry, whereas the self-energy operator was computed using 2472 unoccupied bands and a G-vector cutoff energy of 46 Ry and 160 Ry for the screened and bare Coulomb matrices, respectively. For the solution of the BSE, the six lowest conduction bands and six topmost valence bands were included. All employed energy cutoffs, number of empty bands and number of k-points were carefully determined after exhaustive convergence tests, and are needed to have a fully converged absorption spectra. The excitonic Hamiltonian has been evaluated by including the resonant part, and also by adding the antiresonant term. Moreover, calculations of spin-polarized optical properties, done to highlight possible dark excitons due to triplet excitations, have been performed. No measurable differences among the three kinds of calculations have been obtained, therefore in this Article results including only the resonant part of the Hamiltonian are reported and discussed. Details on convergence for specific quantities of interest are reported in the following section.

In all calculations shown in the main text and in the SI, a phenomenological Lorentzian broadening with an energy width of 0.12 eV has been applied to mimic the experimental broadening of the absorption spectra.

S3.3 *Ab initio* calculations: Results

The calculated GW direct bandgap at the Γ point is 4.07 eV, at Z is 4.13 eV, and the indirect bandgap (between Γ and a k-point close to X) is 3.61 eV (see Fig. S9 below). These values have been converged up to 0.05 eV, and the two coincide, besides the use of a different plasmon pole models. The present fully converged mimimum GW quasiparticle correction amounts to 1.4 eV, that is slightly smaller than the value of 1.69 eV from Ref. [5] (the difference comes from the smaller number of bands and k-points used in that calculation due to computational limitations inherent to the Yambo code at the time). The careful convergence evaluation done in the present work, the use of two different implementations of the theory (see Fig. S10),

and the very good agreement of the value of the electronic gap at Γ between theoretical and experimental data, allow to assess the reliability of the above reported values for describing the electronic properties of anatase TiO_2.

The symmetry-line along Γ-Z features nearly parallel dispersion curves for the conduction and valence band edges. Therefore, using the direct gap at Γ of 4.07 eV as reference to measure the bound character of the excitons is a very good approximation (the smallest gap along that direction is ∼4 eV). The peculiar shape of electronic states along Γ-Z is closely related to the optical properties of the material. Due to the band structure shape along the Γ-Z high-symmetry direction, we notice that the k-point sampling is especially critical for the optical spectra quality, since the main excitons are built up from optical transitions with contributions from a small region of the BZ. This region displays a wormlike shape and is located exactly along the Γ-Z direction.

A number of ∼ 500 k-points (obtained from the 12×12×12 grid in Yambo or from the randomly shifted 8×8×8 grid in BerkeleyGW) provides a spectrum that is still not in perfect agreement with the experimental data (despite the improvements with respect to the one published in Ref. [5], which was calculated with a less refined grid of 8×8×8 unshifted points). The absorption edge is converged to few meV already with the above cited grid, while the peak maximum of exciton I is converged within 10 meV for a grid of 12×12×12. Most importantly, the main effect of the stringent convergence obtained here with respect to k-points and number of bands is given by the shape of exciton I (see Fig. S11). This charge excitation, split in two small peaks at low convergence [5] (or a main peak with a shoulder), becomes a unique, uniform peak, similar to the experimental shape (dark red curve in Fig. S11). The k-sampling needs such a fine grid, since the main optical transitions contributing to exciton I comes from the Γ-Z line, with bands almost parallel and flat. The two-dimensional charge transfer exciton I for $\mathbf{E} \perp c$ (at 3.76 eV) has indeed a major contribution from the transition from the top of VB to the bottom of CB at the middle point in the Γ-Z line. To a lesser extent, significant contributions come from the k-points lying along the Γ-Z line and close to it in every direction. The contribution increases gradually when approaching the aforementioned Γ-Z middle point. Even when the GW electronic band gap of 3.96 eV at the middle point along the Γ-Z line is

used as a reference energy to calculate the exciton binding energy (E_B), the bound nature of exciton I is confirmed (with E_B = 200 meV). A phenomenological Lorentzian broadening of 0.12 eV was then applied to reproduce the experimental spectra (see Fig. S12 for comparison with BSE eigenvalues, showing that mainly one eigenvalue is contributing to the exciton peak I). The energy, shape and reciprocal space contributions for exciton II highlights its bulk-resonance character, most evident as its offset coincides with the RPA @ GW absorption rise. On the other hand, we found that the excitonic peak (exciton III) for **E** ∥ c (at 4.28 eV) is of a more complex nature, with a mixed contribution of bound excitons and delocalized resonant transitions. The former have the dominant character and the contributing transitions are found throughout the BZ, while the main contribution for the latter is concentrated in few points close to the Γ-Z line in the region around Z. The energy and shape of exciton III confirm the analysis of a mixed bulk resonance and localized character, as the continuum onset in RPA @ GW appears to undergo an intensity enhancement. The mixed bulk-localized nature makes it less straightforward to estimate its E_B. Assuming the RPA @ GW onset at 4.4 eV for **E** ∥ c as the reference energy for evaluating E_B, we estimate $E_B \sim 150$ meV.

The slight shift of exciton III peak maximum with respect to the experimental value (0.1 eV) has also been investigated in detail. Increasing the convergence on k-points and bands did not allow to obtain an as perfect matching as for exciton I. We could also exclude possible effects of anisotropic screening by increasing the local field effects parameters and by separating the screening along the in-plane and out-of-plane components, since this does not lead to significant changes in the spectrum. The peak maximum seems instead to be related, in a nonlinear manner, to the lattice constants. The a-axis lattice constant from *ab initio* optimization is in excellent agreement with experimental data (3.79 Å *vs* 3.78 Å), while the c-axis lattice constant is slightly (1%) overestimated. The energy of the peak maximum shifts of 0.2 eV when the experimental or PBE optimized lattices are used; however, the PBE lattice constants give the best agreement with the experimental data for the exciton III position.

Finally, the presence of dark excitons in anatase TiO_2 at energies below the bright exciton I at 3.76 eV has been ruled out by our spin-resolved optical BSE calculations (the lowest exciton in anatase TiO_2 is a singlet and it is optically active).

Additional References

[40] H. Berger, H. Tang, Lévy, Growth and Raman spectroscopic characterization of TiO_2 anatase single crystals. *J. Cryst. Growth* **130**, 108-112 (1993).

[41] S. Mahshid, M. Askari, M. Sasani Ghamsari, Synthesis of TiO_2 nanoparticles by hydrolysis and peptization of titanium isopropoxide solution. *J. Mater. Process. Tech.* **189**, 296-300 (2007).

[42] M. H. Rittmann-Frank *et al.*, Mapping of the photoinduced electron traps in TiO_2 by picosecond x-ray absorption spectroscopy. *Angew. Chem.* **53**, 23, 5858-5862 (2014).

[43] P. Zhang *et al.*, A precise method for visualizing dispersive features in image plots. *Rev. Sci. Instrum.* **82**, 043712 (2011).

[44] D. E. Aspnes, Approximate solution of ellipsometric equations for optically biaxial crystals. *J. Opt. Soc. Am.* **70**, 1275-1277 (1980).

[45] J. Humlíček, C. Bernhard, Diffraction effects in infrared ellipsometry of conducting samples. *Thin Solid Films* **455**, 177-182 (2004).

[46] A. B. Kuzmenko, Guide to Reffit: Software to fit optical spectra (2004) <http://optics.unige. ch/alexey/reffit.html>.

[47] G. Onida, L. Reining, A. Rubio, Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **74**, 601 (2002).

[48] L. Hedin, New Method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **139**, A796 (1965).

[49] L. Hedin, S. Lundqvist, *Solid State Physics*, (Academic Press, New York, 1969), Vol. 23, 1.

[50] A. Marini, C. Hogan, M. Grüning, D. Varsano, Yambo: An *ab initio* tool for excited state calculations. *Comp. Phys. Comm.* **180**, 1392 (2009).

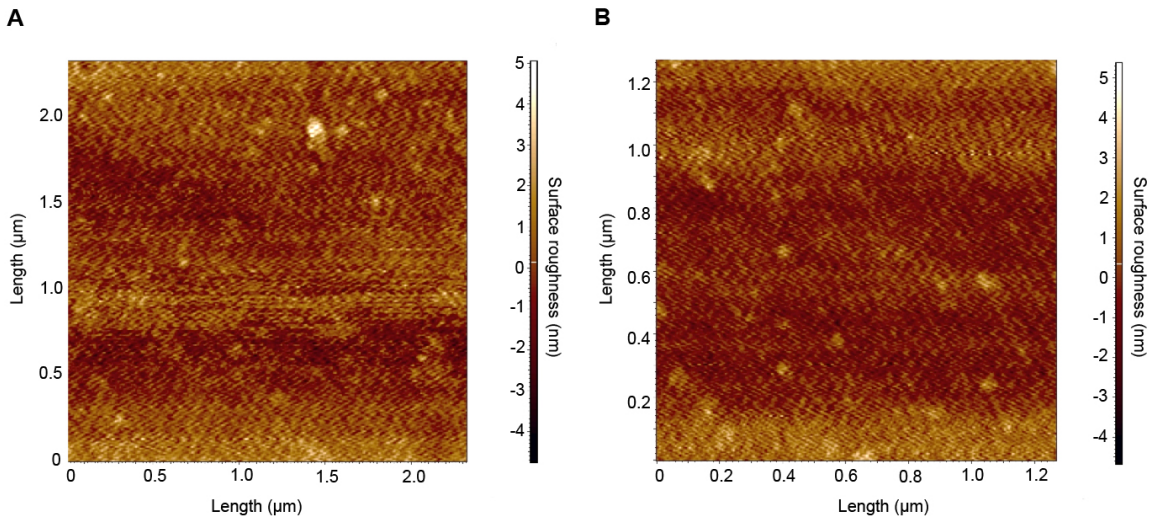

Figure S1: (A - B) Surface roughness characterization of the (010)-oriented polished surface of the reduced anatase TiO_2 single-crystal used for the spectroscopic ellipsometry measurement. The images are taken using Atomic Force Microscopy and the average surface roughness is estimated around 0.9 nm.

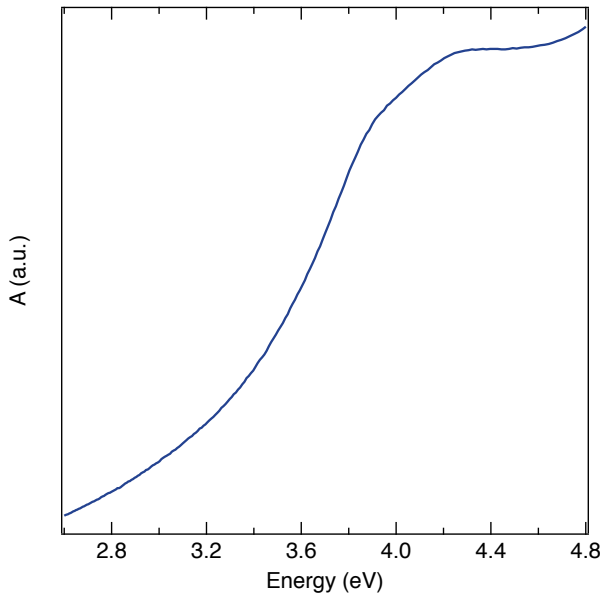

Figure S2: Room temperature steady-state absorption spectrum of anatase TiO_2 NPs dispersed in aqueous solution.

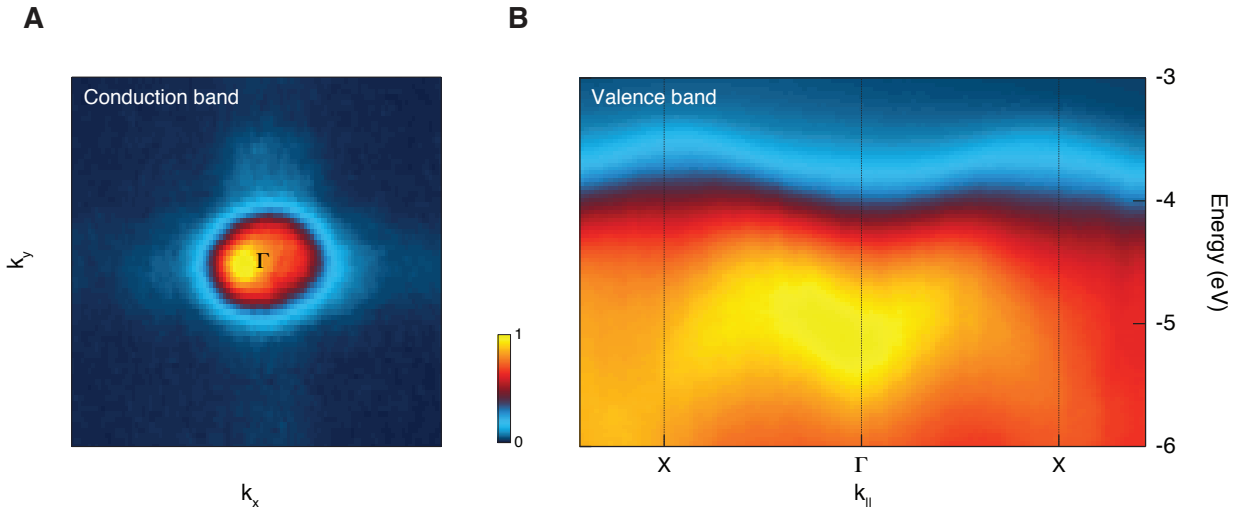

Figure S3: (A) Fermi surface in the $k_x k_y$ plane (T = 20 K, $h\nu$ = 85 eV) of an (001)-oriented single crystal of anatase TiO_2 with oxygen vacancies. The total carrier density is estimated to be 2×10^{20} cm^{-3}. An electron pocket is found at the Γ point of the BZ. Further details are given in Ref. [10]. (**B**) ARPES data of the electronic structure at the top of the VB between Γ and X. The spectrum is referenced to the Fermi energy, which lies at the bottom of the CB at Γ.

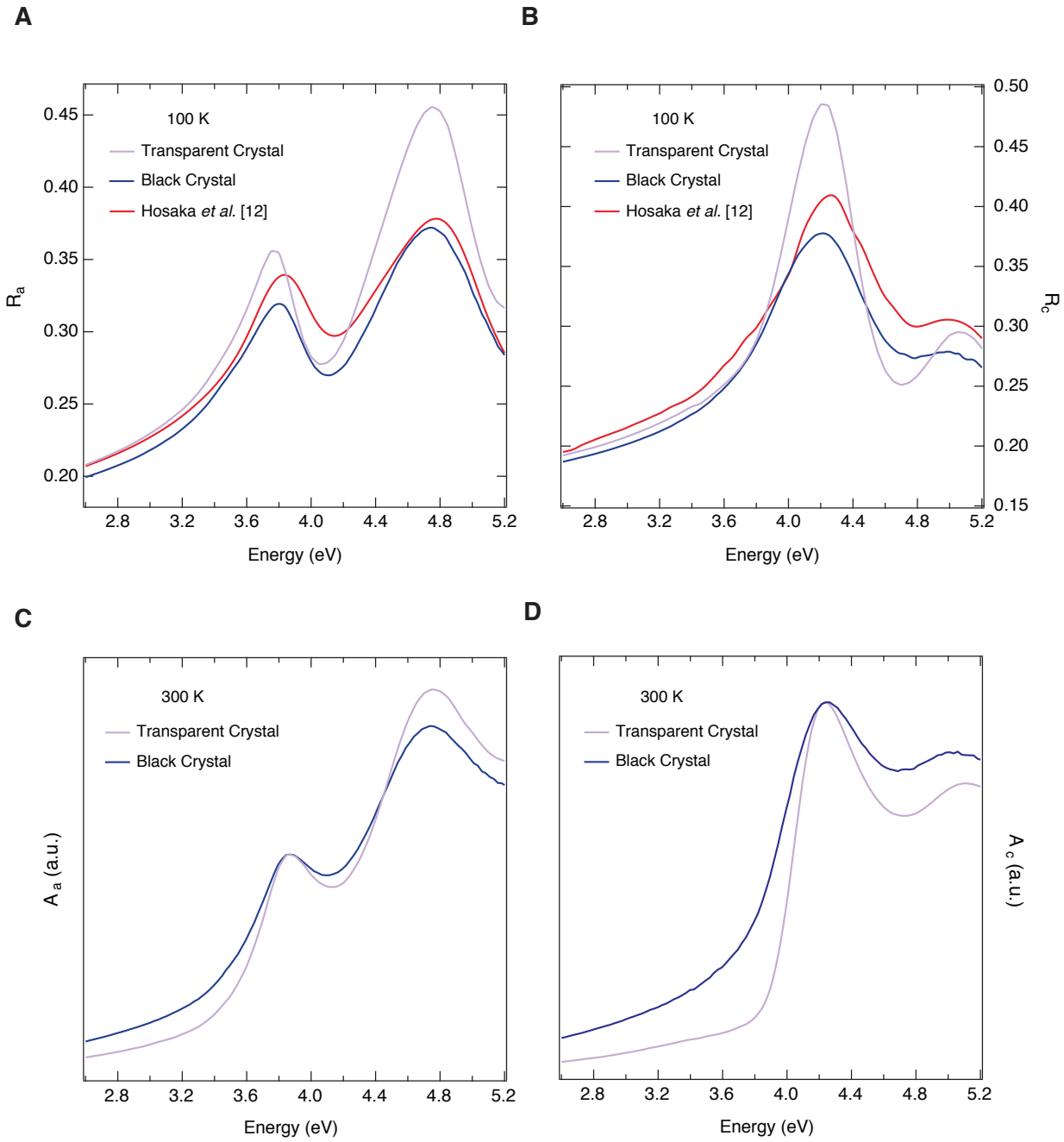

Figure S4: (**A-B**) Static reflectivity spectra of (010)-oriented anatase TiO_2 single-crystals at 100 K. The electric field is polarized along (**A**) the a-axis and (**B**) the c-axis of the crystals. The data derived from our SE measurements are depicted in violet for the transparent crystal and in blue for the black crystal, while the reflectivity measured in Ref. [12] is reported in red lines. (**C-D**) Normalized static absorption spectra of (010)-oriented anatase TiO_2 single-crystals at 300 K. The electric field is polarized along (**C**) the a-axis and (**D**) the c-axis. The data derived from our SE measurements are depicted in violet for the transparent crystal and in blue for the black crystal.

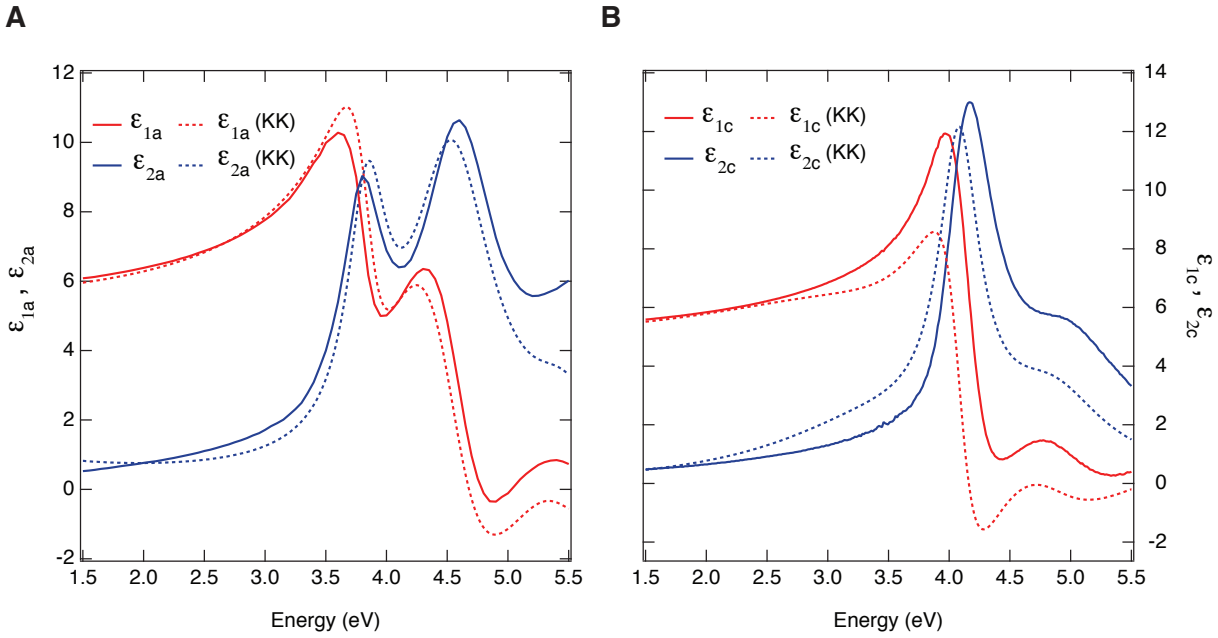

Figure S5: Complex dielectric function of the (010)-oriented anatase TiO_2 single-crystal at room temperature. The electric field is polarized along (**A**) the a-axis and (**B**) the c-axis. The real part, $\epsilon_1(\omega)$, is plotted in red, while the imaginary part, $\epsilon_2(\omega)$, in blue. The solid-line curves depict the data directly extracted from spectroscopic ellipsometry, while the dashed lines are calculated by KK analysis of reflectivity.

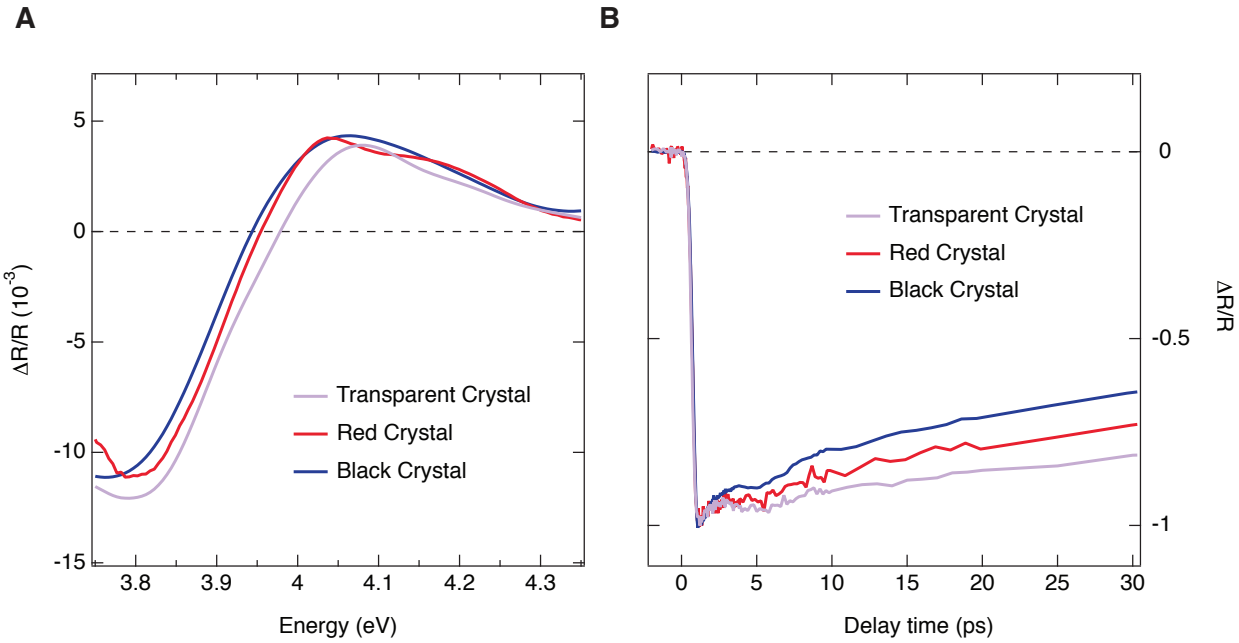

Figure S6: UV TR measurements on different classes of (001)-oriented anatase TiO_2 single crystals at RT: the signals measured from raw crystals are depicted in black, from Cu-doped crystals in red and from reduced crystals in violet. Both pump and probe polarizations lie along the a-axis and the pump energy is set at 4.1 eV. (**A**) Transient spectrum at the fixed time delay of 6 ps. (**B**) Normalized temporal traces at a fixed probe energy of 3.82 eV.

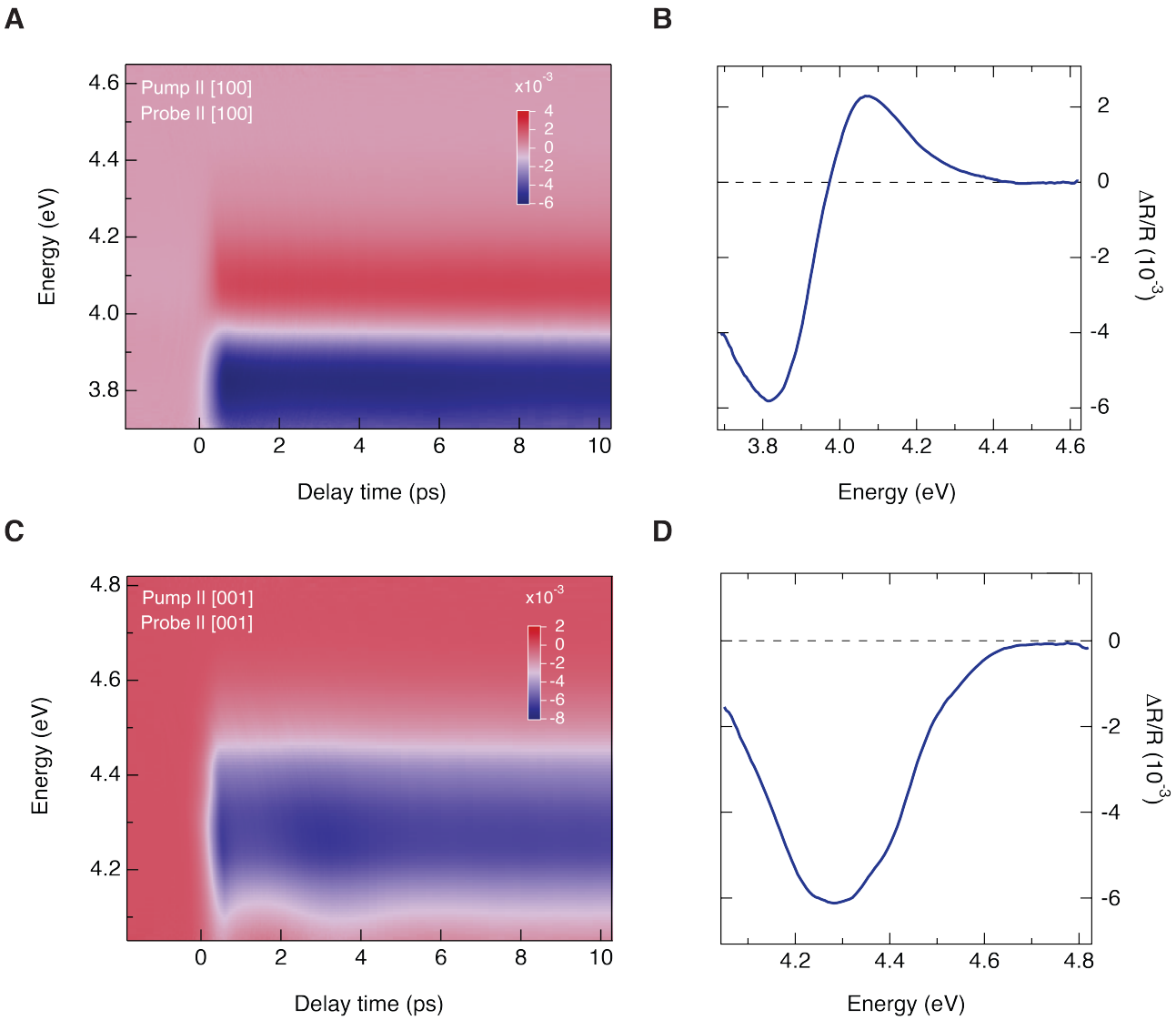

Figure S7: UV TR measurements on (010)-oriented anatase TiO_2 single-crystals at RT. (**A-B**) Color-coded maps of $\Delta R/R$ measured upon photoexcitation at 4.4 eV. (**C-D**) Transient spectra, obtained from a cut at 1 ps in the experimental conditions reported for (**A**) and (**B**) respectively.

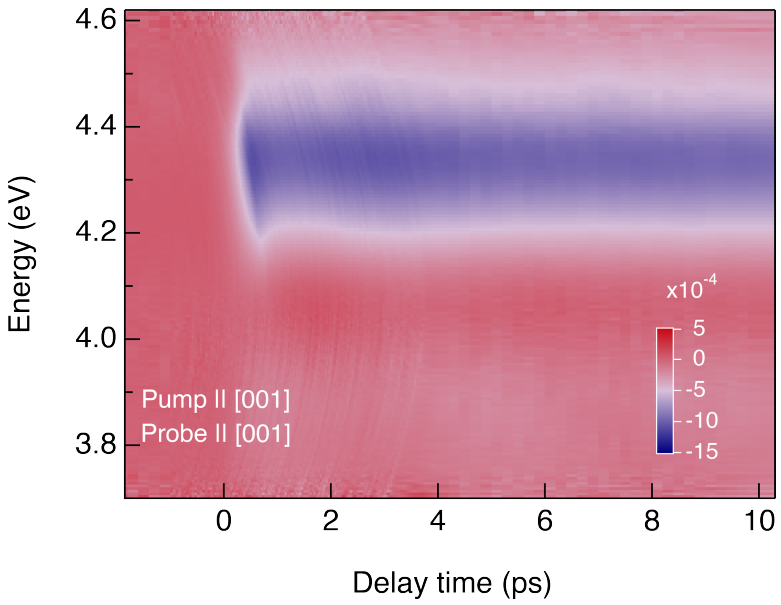

Figure S8: Color-coded map of $\Delta R/R$ from the (010)-oriented single-crystal measured at RT upon photoexcitation at 4.4 eV and with pump and probe beams polarized along the c-axis. The probe photon energy covers the spectral range 3.7 - 4.6 eV, which demonstrates the absence of emerging features at low energies.

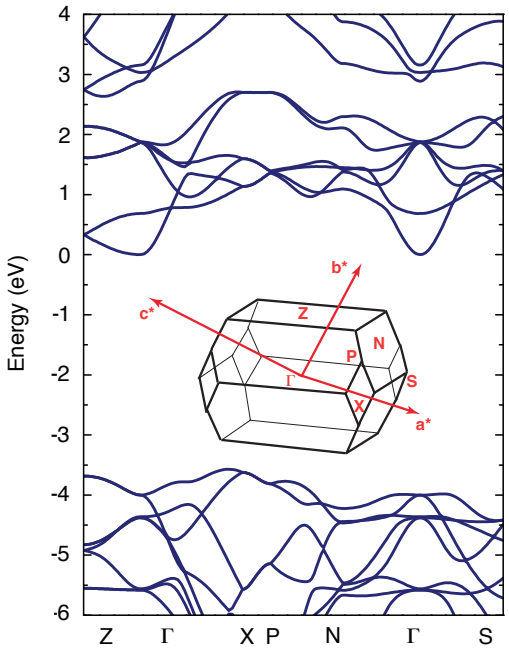

Figure S9: Complete GW electronic band structure of bulk anatase TiO_2. The zero of the energy is set at the minimum of the CB at the Γ point. The direct gap at Γ sets a reference for defining the bound character of the excitons (see the discussion in the main text). The inset shows the 3D BZ of anatase TiO_2.

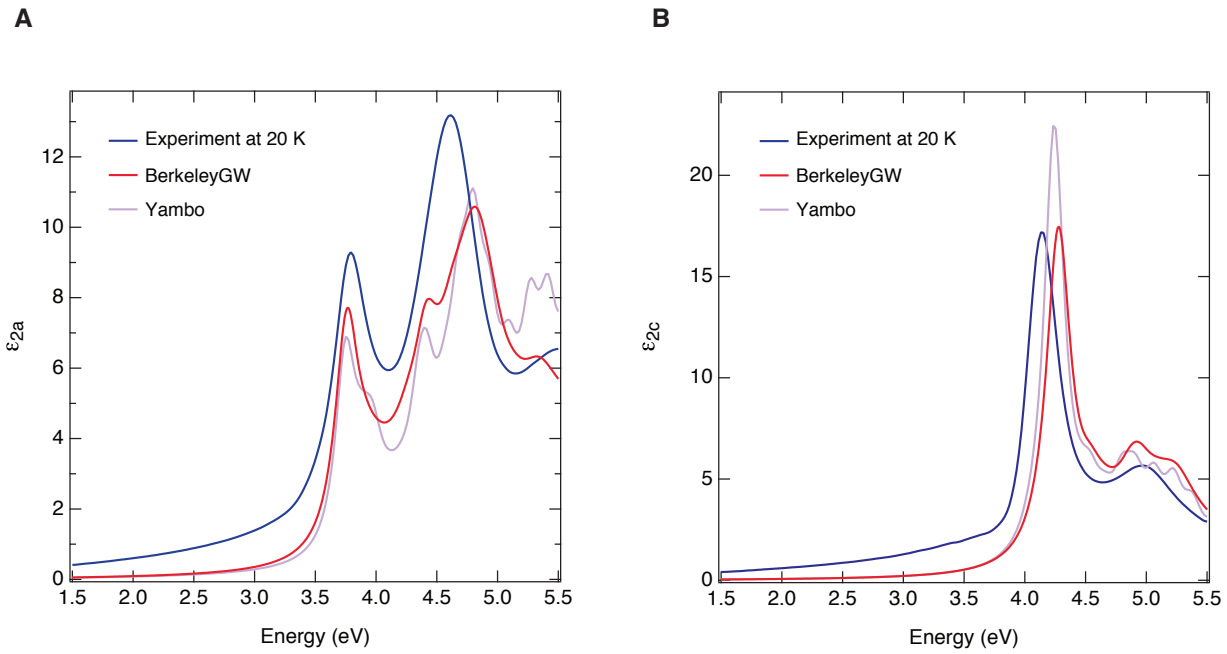

Figure S10: Comparison between the experimental SE data (blue curve) and BSE calculations, for light polarized along (**A**) the a-axis (**B**) the c-axis. Both BerkeleyGW (red curve) and Yambo (violet curve) data are evaluated using the highest convergence parameter values described in the text. For BerkeleyGW they correspond to the converged spectra (both for peaks shape and position); for Yambo these are the maxima computational feasible values, allowing to obtain a reasonable spectra convergence but showing a spurious shoulder above the main exciton peak in (**A**) (as in previously published works). The fully-converged spectra (red surves) show a single peak in agreement with the experimental data. For light polarized along the a-axis the agreement between the two calculations is excellent.

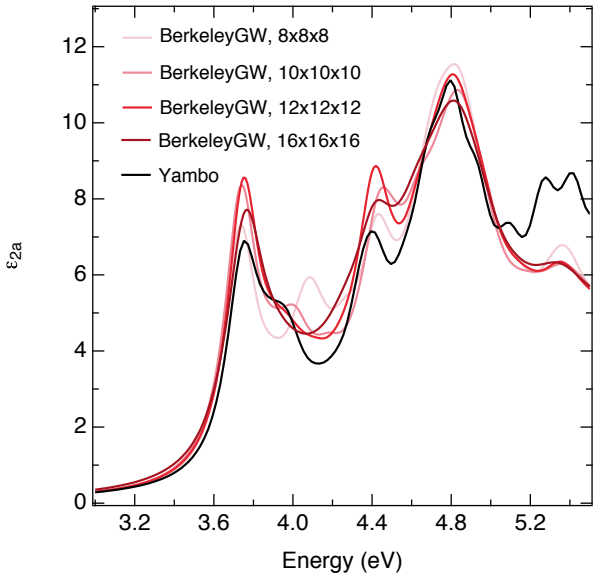

Figure S11: Comparison between the calculated ϵ_{2a} using BerkeleyGW and Yambo. The convergence of the spectrum for increasing k-points grid dimension is reported in a red-color scale for the BerkeleyGW data, while the best converged spectrum from Yambo is shown in black.

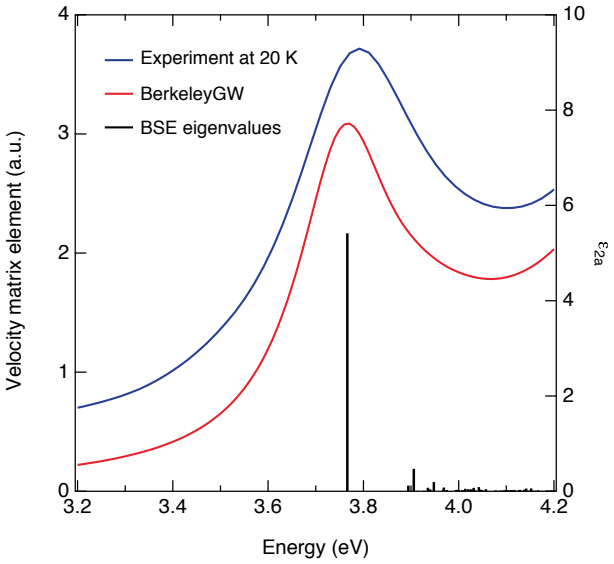

Figure S12: Black bars represent the square of the transition matrix elements of the velocity operator along the a-axis ($|\mathbf{E}_a \cdot \langle 0|\ \mathbf{v}\ |S\rangle\,|^2$), corresponding to exciton states S contributing to the peak I. This quantity is related to the oscillator strength f_S by $f_S = (2 \mid \mathbf{E}_a \cdot \langle 0|\ \mathbf{v}\ |S\rangle\,|^2)/E_S$, where E_S is the excitation energy corresponding to exciton state S [22]. Comparison with the experimental ϵ_{2a} (blue curve) and the full-converged BerkeleyGW calculations with a phenomenological Lorentzian broadening of 0.12 eV (red curve).