Ultrafast X-Ray hyperspectral imaging of a photo-induced phase transition with nanometer space and femtosecond time resolution

Allan S. Johnson1,2, Daniel Pérez-Salinas1, Khalid M. Siddiqui2, Sungwon Kim3, Sungwook Choi3, Klara Volckaert2, Paulina E. Majchrzak2, Søren Ulstrup2, Kent Hallman4, Richard F. Heglund Jr.4, Christian M. Günther5, Bastian Pfau6, Stefan Eisebitt6,7, Dirk Backes6, Francesco Maccherozzi8, Ann Fitzpatrick8, Sarnjeet Dhesi9, Pierluigi Gargiani9, Manuel Valvidares9, Hyeongi Choi10, Dogeun Jang10, Abhishek Katoch10, Soonnam Kwon10, Sang Han Park10, Hyunjung Kim3, Simon E. Wall1,2,*

1 ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860, Castelldefels, Barcelona, Spain.
2 Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Aarhus C, Denmark.
3 Department of Physics, Sogang University, Seoul 04107, Korea.
4 Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235-1807, USA.
5 Zentraleinrichtung Elektronenmikroskopie (ZELMI), Technische Universität Berlin, 10623 Berlin, Germany.
6 Max-Born-Institut, 12489 Berlin, Germany.
7 Institut für Optik und Atomare Physik, Technische Universität Berlin, 10623 Berlin, Germany.
8 Diamond Light Source, Chilton, Didcot, Oxfordshire, OX11 0DE, United Kingdom.
9 ALBA Synchrotron Light Source, Barcelona, Spain.
10 Pohang Accelerator Laboratory, Pohang, Gyeongbuk 37673, Korea.

*Corresponding authors: allan.johnson@icfo.eu and simon.wall@phys.au.dk

Using light to control transient phases in quantum materials is an emerging route to engineer new properties and functionality1,2, with both thermal and non-thermal phases observed out-of-equilibrium3–6. Transient phases are expected to be heterogeneous, either through photo-generated domain growth7–10 or through the generation of topological defects11,12, which impacts the dynamics of the system. Yet to date this nanoscale heterogeneity has not been directly observed. Here, we use time- and spectrally-resolved coherent X-ray imaging to track the prototypical light-induced insulator-to-metal phase transition in vanadium dioxide (VO2) on the nanoscale with femtosecond time resolution. Remarkably, the early-time dynamics are independent of the initial spatial heterogeneity. We observe a sub-140-fs switch to the metallic phase which launches strain waves that propagate on the few picosecond timescale into the sample, with a heterogeneous response emerging only after hundreds of picoseconds. This direct method of imaging nanoscale dynamics reveals an interpretation of the transition in VO2 in strong contrast to those based on spatially averaged probes5,6,13 and suggest that photo-excitation generates a highly-orthorhombically-strained rutile metallic phase. Our results demonstrate the critical importance of spatially and spectrally resolved measurements for understanding and interpreting the transient phases of quantum materials.

Direct observation of the nucleation and growth of a new phase is the solid-state equivalent of the molecular movie in chemistry14. However, this is challenging to realize in solids due to the range of length-scales involved, with dynamics occurring from the atomic scale up to the macroscale. Remarkable progress has been made in understanding atomic scale dynamics on the ultrafast timescale by using diffraction-based probes, which can reveal atomic dynamics beyond the mean response15 or dynamics of the average structure over nanoscale regions16. However, electronic probes are needed to understand the functionality of these states. This is particularly
important in quantum materials, where transient states can be created by light that have electronic properties not found in equilibrium. In many cases, transient states are believed to be heterogeneous at the nanoscale, both because of the inhomogeneous excitation profiles generated by the pump beam in the depth of the material, but also due to the intrinsic heterogeneity of many quantum materials. A key challenge to understanding these phases is therefore to isolate the photo-induced state and to directly probe its properties at the nanoscale. Here, we use time- and energy-resolved coherent resonant soft X-ray imaging to observe the ultrafast insulator-metal phase transition in the prototypical quantum material vanadium dioxide (VO₂), over a macroscopic area with sub-50-nm spatial resolution and 140 fs time resolution, returning full spectroscopic information on transient states at the nanoscale.

The light-induced phase transition in VO₂ has been particularly influential in our understanding of optically driven quantum materials. At room temperature, VO₂ is in a monoclinic-insulating (M1) phase characterized by dimerized pairs of vanadium ions. Light can be used to break these dimers and drive the transition on the ultrafast timescale to the high temperature rutile-metallic (R) phase. The study of this transition has driven the development of multiple new techniques: it was the first solid-solid phase transition to be tracked by time-resolved X-ray diffraction, while ultrafast X-ray absorption techniques were first used to understand its electronic nature. The transition to the rutile phase has been shown to proceed by a disordering of the vanadium pairs on a sub-100 fs timescale, with the band gap collapsing on a similar timescale. However, heterogeneity in the transient state is believed to play a key role in the dynamics. Although the lattice and electronic properties can change on the ultrafast timescale, analysis of the THz conductivity suggests that the rutile metallic phase locally nucleates and grows on a timescale of tens of picoseconds. In addition, electron diffraction results have suggested that a non-equilibrium monoclinic metallic phase can be generated locally at the nanoscale, which is not found in equilibrium. However, the existence of non-equilibrium phases remains debated because to date such processes have not been observed directly.

To address the role of nanoscale heterogeneity and phase separation, we use time- and spectrally-resolved resonant soft X-ray coherent imaging at the PAL-XFEL to image the light-induced phase transition on the ultrafast timescale with nanometer spatial resolution. The power of this technique lies in the fact that it is a wide-field imaging technique that can exploit resonant X-ray spectroscopy both to provide a contrast mechanism between phases and to enable the extraction of quantitative spectral information to aid phase identification on the nanoscale. We report time-resolved imaging using two modes of operation, Fourier transform holography (FTH) and coherent diffractive imaging (CDI). In FTH, scattering patterns are inverted directly through the use of a fast Fourier transform and require a single exposure. This enables rapid data collection, but comes at the expense of losing the absolute values of the complex transmission. CDI images make use of multiple exposures, to increase the dynamic range of detected scattering pattern, so that images can be obtained via iterative phase retrieval algorithms in order to obtain the quantitative absolute transmission of the sample.

Figure 1a shows a multi-energy FTH-image of VO₂ recorded at the vanadium L- (517 eV, red channel) and oxygen K- (529.5 and 531.25 eV, blue and green channels) edges at 325 K, a temperature at which the insulating and metallic phase coexists. These energies are chosen because the absorption coefficient shows large changes depending on the phase of the material, with the 529.5 eV signal showing a decrease in transmission and the 531.25 eV signal an increase when the system changes from M1 to R. These changes can be used to provide contrast in imaging. The RGB image of Fig 1a shows the full topography of our sample, which consists of a range of crystallite sizes and grain boundaries (white). These boundaries are known to pin the position of the metallic domains and in larger crystallites clear domain coexistence of M1 (purple) and rutile metallic (green) phases can be seen.
In Fig 1b we further verify the initial domain assignment by performing a temperature dependent measurement. By subtracting the green and blue channels from the RGB image, we can remove the temperature independent topography. However, as these channels show opposite changes in crossing $T_c$, the domain structure is still preserved. Here we can clearly see the R phase nucleating and forming a stripe state with M1, which is common in nanocrystals\textsuperscript{28}.

We now examine the dynamics of the heterogeneous state at 325 K, in order to observe the disappearance of the M1 phase, growth of the R phase and, or, the nucleation of new transient phases. We excite the system with 24mJ/cm\textsuperscript{2}, 800-nm pulses and perform time-resolved FTH imaging at 529.5 eV photon energy. At this fluence we are in the saturation regime of the phase transition, where increasing fluence no longer results in substantial changes, but the initial domain structure still recovers between photo-excitation events (See Supplementary note 1 and extended data figure 1). Large changes are found all over the sample, and in Figure 1c we focus on two regions of interest (ROIs) which are cuts across approximately 50 nm wide rutile metallic domains surrounded by the M1 phase, which appears as a dip in the image intensity. In both cases, the contrast between the metal and insulating states is strongly, but not completely, lost within the first 150 fs, after which subsequent dynamics are only observed on the hundreds of picosecond timescale.

**Figure 1**: Time dependent X-ray holographic imaging of VO\textsubscript{2}. a False colour composite FTH image of VO\textsubscript{2} from images recorded on the VO\textsubscript{2} soft X-ray resonance (below) at 517 (red), 529.5 (blue) and 531.25 eV (green). The metallic R-phase appears green and insulating M1-phase appears purple. b Temperature dependent domain growth highlighted through the subtraction of the blue and green channels, $\Delta s$, which removes the sample morphology. The ROI used is indicated by the white dotted region in panel a. c Transmission dynamics of two line-outs spanning R regions surrounded by the M1 phase. Their positions are indicated in panel a and colour-coded. The domain structure, initially ~50 nm, is promptly lost. Background is shaded according to state of the material as a guide to the eye.
We next examine the spatial dependence of the dynamics more closely. Figure 2a shows the pump-induced changes in the domain structure across the full-field of view measured at 529.5 eV, relative to the state of the sample recorded at -8.5 ps. Regions of increased transmission are shown in red, with decreased transmission in blue. Changes are observed across the entire sample with nanoscale texturing, but notably the spatial dependence of the pattern is roughly independent of pump-probe delay after the initial changes. A comparison to the static domain pattern (Fig 2b) shows that the regions where the transmission decreased correspond to regions that started in the green-R (purple-M1) phase. These early time dynamics can be modelled as a double 1/t decay, with a prompt, resolution limited change (140 fs) followed by a slower 4.75 ps evolution.

Instead, the changes observed at the R phase are mainly artificial, resulting from the loss of the DC component in the FTH images, which can cause correlated dynamics across the whole image (see supplementary note 2 and extended data figure 2). Therefore, in order to identify how many unique processes are occurring spatially, we perform a principal components (PC) analysis of the dynamic images. This process breaks down the transmission dynamics $T$, into a series of “eigen” spatial and temporal functions of the form $T(x,y,t) = \Sigma_i A_i(x,y)f_i(t)$.

For times up to 20 ps, we find that only a single PC is needed to describe the dynamical evolution of the images, i.e. $T(x,y,t) = A(x,y)f(t)$. Only when data beyond ~100 ps are included in the analysis are additional terms needed, see extended data figures 3-5 and supplementary notes 3 and 4. The spatial and temporal response of the initial dynamics are plotted in Fig 2b. The spatial pattern shows a clear correlation with the initial domain structure, demonstrating that the observed dynamics are indeed the result of the M1 regions of the sample switching to the metallic R phase, and that no other local dynamics occur. A fit to the time trace for this process reveals two
time-constants, with the first a resolution-limited 132 fs fall time, consistent with the ultrafast nature of the structural\textsuperscript{15} and electronic\textsuperscript{19,20} changes during the M1 to R phase transition. In addition, a second 4.75 ps time constant is also observed. Importantly, as only one principal component is needed to describe the dynamics these two timescales occur in identical regions of the sample. This secondary picosecond time-constant has been seen in multiple experiments, which have shown it to be fluence dependent\textsuperscript{5,7,13,32,33}. However, the interpretation of this time-constant has been debated. In some cases, it has been taken as evidence for a non-equilibrium, monoclinic metallic phase\textsuperscript{5,13}, while others have interpreted it as nucleation and growth of the metallic phase\textsuperscript{7}. The monoclinic metallic phase is proposed to occur in regions that do not switch to the rutile metallic phase, and thus the two timescales should occur at spatially distinct regions. Similarly, in the nucleation and growth picture, the fast timescale should only be observed at the initial nucleation site, from which the metallic phase grows. Both these scenarios are at odds with the data presented here and would require more than the single principal component to describe.

To further confirm this assignment and to better understand the nature of the transient state formed after the picosecond evolution, we use spectrally resolved CDI imaging to recover the full spectrum of the newly-switched regions\textsuperscript{29}. We acquire a hyperspectral image by scanning the X-ray probe wavelength, with images taken at 31 photon energies across the oxygen K-edge at a delay of 20 ps after photoexcitation (Methods). The resulting spectrally-integrated image is shown in Fig 3a, which shows that the sample is remarkably homogenous after excitation. However, as already noted in Fig 1, markers of the initial phase are still observed at key energies. To elucidate the origin of this effect, we then extract the transient spectra from all regions of the sample that were initially insulating and compare them to those that were initial metallic (Fig 3b), enabling us to understand how the transient metallic state differs from the insulating state. The resulting spectra, and the differential, are shown in Figs 3c and 3d.

Both regions are remarkably similar and the resulting difference, less than 0.5% at 530.5 eV, is significantly smaller than the changes found from the insulator-metal transition, which are of order 10\%\textsuperscript{29}, showing the regions that were initially M1 has switched to the R phase. One explanation for this small residual signal could be temperature, as regions that are initially metallic will absorb a different amount of laser energy than those that start in the insulating phase. However, the spectral signatures observed are not consistent with thermal differences in the metallic phase\textsuperscript{23,34}. Instead, we suggest that these differences result from strain generated during the phase transition. The volume of VO\textsubscript{2} increases by ≈1\% when switched from M1 to the R phase in equilibrium\textsuperscript{35}. Here, however, although the dimerization can be lost on the ultrafast time scale, this volume expansion cannot occur. As a result, the metallic phase that is photo-generated is more strained than the regions that were initially metallic. Because of the anisotropy of the initial M1 phase, the light-induced rutile state will then be anisotropically stressed, resulting in an orthorhombic unit cell. The orthorhombic cell will then slowly relax to the tetragonal rutile structure, which may be an alternative explanation for long-lived diffraction features that have been interpreted as a monoclinic metal\textsuperscript{5,6,13}. This interpretation is further supported by the fact that the observed spectral shifts of the $\pi^*$ state at 530 eV (primarily in-plane here)\textsuperscript{28,29} are consistent with XAS signals from statically strained samples\textsuperscript{36}. In comparison, the $\sigma^*$ state at 533 eV (primarily out-of-plane) shows no significant differences.
Figure 3: Nanoscale X-ray spectroscopy of transient phases. a Frequency integrated hyperspectral CDI image at 20 ps pump-probe delay, showing clear topography but no signs of domains. b Masks dividing the sample into regions that started in the R phase (blue), and thus have not switched phase, and in M1 phase (red). c Average complex transmission spectrum (transmission-continuous lines and phase-dashed lines) of blue and red regions at 20 ps delay. Shaded regions correspond to one standard deviation across all sample points. d Difference between switched and un-switched regions. Shaded regions correspond to one standard error. Differences are consistent with a small red-shift of the π* state, while the σ* resonance is largely unaffected.

In addition, propagation of this strain into the depth of the film can explain the slower picosecond time-constant observed. The large volume change will generate discontinuities at the surfaces of the thin film, leading to shock waves. The time for a shockwave to cross our sample when entirely switched to the R-phase is around 7 ps, consistent with our observed timescale (see Supplementary note 5 and extended data figure 6). Furthermore, due to the finite penetration depth of the pump, at lower fluences the film will partially switch from the front surface. This leads to large, fast moving, short-lived strain waves launched from the vacuum-R-phase and R-to-M1 phase interfaces into the out-of-plane direction. As the excitation fluence is increased, the R-phase will be switched deeper into the material. As a result, the acoustic dynamics will speed up due to the higher speed of sound in the R phase, consistent with the fluence dependence of this feature as seen in other measurements. However, although the out-of-plane strain propagation is fast, the in-plane strain remains because the larger length scales involved in the in-plane direction result in significantly longer dynamics.

In conclusion, we have used time-resolved X-ray spectroscopy to image the light-induced phase transition in VO₂ on the nanoscale for the first time. Despite substantial heterogeneity in our initial state, the dynamics of the transition are largely homogenous, with all M1 insulating regions governed by the same dynamical process within the first 20 ps. The lack of spatial variation enables us to exclude the local generation of non-thermal phases or a nucleation-and-growth mechanism on this timescale. Instead, we show that the photoexcited phase generated is rutile-like, but highly strained, with the out-of-plane strain relaxing rapidly on the picosecond timescale, but the in-plane strain remaining for hundreds of picoseconds, consistent with a transient orthorhombic, rather than true rutile, structure. By performing time and spectrally resolved CDI we can now track and identify new phases with nanoscale resolution. In addition, we find no evidence for the ultrafast generation of topological defects, which
have been indirectly observed through other methods in different materials\textsuperscript{11,12}. This may be because of the high fluence excitation used in this experiment, resulting in a homogenous nucleation of the R phase, but our results demonstrate that these emergent topological defects could now be imaged in the time domain\textsuperscript{39}. Furthermore, our approach can easily be extended to characterize filament growth after applying electric fields, which will enable spectroscopy of field-induced states driven by quasi-DC fields or THz pulses\textsuperscript{5,40}. As our technique can be straightforwardly applied to other quantum materials, these results demonstrate a new approach to explore heterogeneous transient states in quantum materials.

**Methods**

**Sample Preparation and Measurements**

Samples consisted of 75-nm-thick layers of VO\textsubscript{2} prepared on Si\textsubscript{3}N\textsubscript{4} membranes by pulsed laser deposition and subsequent annealing. A [Cr(5 nm)/Au(55 nm)]\textsubscript{20} multilayer (\textapprox 1.1 µm thickness) was deposited on the opposite side and a focused ion beam was used to mill the mask structure. A 2 µm-diameter aperture was milled to define the field of view, along with five 50- to 90-nm-diameter reference apertures 5 µm away from the central aperture.

Pump-probe experiments were performed at the FTH end-station of the SXS beamline at the Pohang Accelerator Laboratory X-ray Free Electron laser operating at 60 Hz repetition rate. Time-resolved images were taken by alternating between positive and negative time delays to ensure that the initial domain structure did not change. Each image averages 9000 XFEL shots. The XFEL beam was focused to a 50 µm x 50 µm spot size at the sample position using a Kirkpatrick-Baez mirror pair. To prevent the sample from being damaged the XFEL was attenuated to have \textapprox 1.2x10\textsuperscript{8} photons/pulse. Optical pump only backgrounds were subtracted from each image before either iterative reconstruction or holographic inversion. Laser pulses with a central wavelength of 800-nm were focused to a spot size of 200 µm full-width at half maximum at the sample with a small (\textapprox 1°) crossing angle relative to the normal incidence X-rays. The overall temporal resolution was around 150 fs, limited by the relative timing jitter of the optical and X-ray beams. Both beams impinged on the sample VO\textsubscript{2}-side first to reduce the potential impact of plasmonic effects in the Cr/Au multilayer introducing additional inhomogeneity\textsuperscript{16}.

**FTH and CDI imaging**

In FTH a beam block can be added to the scattering field to block the intense low-momentum scatter, allowing single exposures to capture the high-momentum scattering needed to observe the nanoscale domains. However, the presence of beam block in the central part of the Fourier plane removes the DC component and acts as a high-pass filter. As a result, the absolute values of the complex transmission of the sample are lost. In CDI, multiple exposures are combined to improve dynamic-range, enabling better sampling of both low- and high-q scattering. Although, in principle the beam block can be removed to capture the DC component, it is not necessary for CDI to reconstruct the DC level if the X-ray beam has sufficient coherence, which is the case at the FEL. Instead, multiple exposures with the beam block in place were sufficient to produce good CDI results via iterative phase retrieval algorithms\textsuperscript{29}.

**Principal Components Analysis and Data Fitting**

The dynamics of the real space images were analysed using principal components (PCs) analysis. First all negative-time-delay (probe before pump images) from each time trace were grouped and decomposed into PCs; as these images should be constant over time, the amplitude of the 2\textsuperscript{nd} PC was used to define a noise threshold for the pump-probe measurements. Then each of the two time-traces (0-1 ps and 0-20 ps) were decomposed into PCs and any PC with an amplitude below the noise threshold was discarded. Only two PCs were found to be significant for each time trace; one with a spatial pattern resembling the negative time delay structure with a weak structure in time, and one approximately constant at negative time delays and exhibiting a sharp drop at temporal overlap. We identified the first as the static background, with dynamics resulting only from fluctuations in the overall X-FEL brightness, and the second as the dynamics of the phase transition. We replaced the first component with its
time-averaged mean value and kept the second. We note that including the other, weaker PCs barely affects the time traces; therefore, we ascribed all time dynamics to the second principal component (extended data figure 4). The remaining time dynamics can be described as a purely real function. Including images taken at longer time delays (100, 250, or 500 ps) breaks this simple description, and additional PCs are required to describe the dynamics, corresponding to the onset of spatial dynamics (extended data figure 3).

While the short and long time traces cannot be directly combined because of a change in the initial domain structure, which we attribute to a random fluctuation in the XFEL intensity introducing irreversible changes, examination of PC corresponding to the dynamics shows they are the same in both cases (extended data figure 5). As such, we fit the time dependence of both traces simultaneously with a double 1/t function and a freely varying amplitude.

Reconstruction of the transient phase

Full amplitude and phase images of transient phase were recovered using partially-coherent iterative phase reconstruction algorithms as described in Johnson et al. Long and short exposures were combined to provide high dynamic range images at 31 photon energies across the oxygen K-edge, which were each used to reconstruct real space images independently. The object constraint was determined from the known mask geometry and the FTH reconstructions; a data constraint mask was used to allow the reconstruction to freely vary in regions where the beamblock masked the detector or where additional background light or detector damage was present, with the additional constraint that the blocked low-\(q\) response obeyed circular symmetry. The properties of the transient state were determined by comparing the average spectral response of all sample regions which began in the R-phase and those which were switched from the M-phase. These regions were determined by combining FTH images at positive and negative time delays at three different photon energies; only regions where all three photon energies agreed were preserved for further analysis. These regions also agree clearly with a PC analysis of the spectrogram, though the degree of sample inhomogeneity prevented a conventional clustering analysis.

Extended Data Figure 1: Fluence dependence of the transient dynamics. a Differential signal at 20 ps time delay and 529.5 eV photon energy as a function of fluence. Changes rise rapidly around 20 mJ/cm\(^2\) and saturate. Changes above 24 mJ/cm\(^2\) are associated with changes in the initial state rather than dynamics, as shown in panel b. b Pre-time-zero images at 36 mJ/cm\(^2\) and 39.6 mJ/cm\(^2\). When the fluence is increased the initial domain pattern undergoes an irreversible change due to the increased heat load.
Extended data figure 2: Demonstration of the effects of the beam block on FTH images. a Hypothetical transmission of a sample in the case of a homogenous metal (green) and a mixed-phase (orange) of metal and insulating regions (insulating regions have a lower transmission). b Change in transmission when going from the mixed-phase to the metallic phase; only the insulating phase shows changes. c FTH image corresponding to the transmission function in panel a. d Change in transmission for the FTH image, in this case both insulating (indicated by the blue arrows) and metallic (red arrows) can show an apparent change.

Extended data figure 3: Amplitude of the principal components. a Relative amplitude of the principal components considering only data in the first picosecond (truncated) and including data at 250 and 500 ps (with long times). Also shown is the noise floor. b,d PC1 (top) and PC2 (bottom) using truncated time data. c,e PC1 and PC2 including data out to 500 ps.
Extended data figure 4: residual signal after subtracting the two main principal components. Each image uses its own floating colour scale to best highlight the differences.

Extended data figure 5: Comparison of short and long time data. Left, the long-time scan data used for fitting the slow 4.5 ps time constant. Right, reproduction of the short time data shown in Figure 2 of the main manuscript. Insert shows the spatial dependence of the principal component.
Extended data figure 6: Simulated thermal and acoustic propagation in VO2 thin films. a Temperature profile in the sample as a function of time. b Corresponding metallic fraction vs time. c Initial stress field from a homogeneously switched film. d Resulting stress relaxation dynamics in the out-of-plane direction.

Acknowledgements
We acknowledge the support of Josh Turner and Alex Reid with an early iteration of this experiment. This work was funded through the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (Grant Agreement No. 758461) and PGC2018-097027-B-I00 project funded by MCIN/ AEI /10.13039/501100011033/ FEDER “A way to make Europe” and CEX2019-000910-S [MCIN/ AEI/10.13039/501100011033], Fundació Cellex, Fundació Mir-Puig, and Generalitat de Catalunya through CERCA. By the Independent Research Fund Denmark under the Sapere Aude program (Grant No. 9064-00057B) and VILLUM FONDEN under the Young Investigator Program (Grant No. 15375). S.K., S.C., and H. K. acknowledge support from the National Research Foundation of Korea (NRF-2020R1A2C1007416, NRF-2021R1A3B1077076). Part of this work was carried out with the support of Diamond Light Source, instrument I06 (proposal MM22048). This work was also funded by a basic science research program funded by the Ministry of Education of Korea (NRF-2020R1A2C1007416). KAH and RFH were supported by the United States National Science Foundation (EECS-1509740). ASJ acknowledges support of a fellowship from “la Caixa” Foundation (ID 100010434), fellowship code LCF/BQ/PR21/11840013, and support from the Marie Skłodowska-Curie grant agreement no. 754510 (PROBIST).

Author Contributions
ASJ, DPS, KMS, SK, SC SHP performed experiments at PAL FEL and were remotely assisted by KV, PEM and SEW. The beamline and scattering chamber were designed and built by AK, SK and SHP. The laser was run by HC and DJ. Data analysis and software development was led by ASJ with help from DPS. Samples were prepared and characterized by KAH and RFH and processed for holography by CMG. Further characterization of the samples was performed at the ALBA synchrotron with collaboration by ASJ and DP with PG and MV and remote support from KV and PEM. Domain stability measurements were performed at the Diamond light source by ASJ and DP with support from DB, FM, AF and SD. The project was conceived by SEW and planned and designed by ASJ, SU, BP, SE, HK and SEW. ASJ, DP and SEW wrote the manuscript with input from all authors.

The authors declare no competing interests

Supplementary information is available for this paper
Correspondence and requests for materials should be addressed to Allan S. Johnson or Simon E. Wall.

Bibliography
1. Basov, D. N., Averitt, R. D. & Hsieh, D. Towards properties on demand in quantum materials. Nat. Mater. 16, 1077–1088 (2017).
2. de la Torre, A. et al. Colloquium: Nonthermal pathways to ultrafast control in quantum materials. Rev. Mod. Phys. 93, 041002 (2021).
3. Ichikawa, H. et al. Transient photoinduced ‘hidden’ phase in a manganite. Nat. Mater. 10, 101–105 (2011).
4. Fausti, D. et al. Light-induced superconductivity in a stripe-ordered cuprate. Science 331, 189–91 (2011).
5. Morrison, V. R. et al. A photoinduced metal-like phase of monoclinic VO₂ revealed by ultrafast electron diffraction. Science 346, 445–448 (2014).
6. Sood, A. et al. Universal phase dynamics in VO₂ switches revealed by ultrafast operando diffraction. Science 373, 352–355 (2021).
7. Abreu, E. et al. Ultrafast electron-lattice coupling dynamics in VO₂ and V₂O₃ thin films. Phys. Rev. B 96, 094309 (2017).
8. Cocker, T. L. et al. Phase diagram of the ultrafast photoinduced insulator-metal transition in vanadium dioxide. Phys. Rev. B - Condens. Matter Mater. Phys. 85, 155120 (2012).
9. Haupt, K. et al. Ultrafast Metamorphosis of a Complex Charge-Density Wave. Phys. Rev. Lett. 116, 016402 (2016).
10. Laulhé, C. et al. Ultrafast Formation of a Charge Density Wave State in 1T − TaS₂ : Observation at Nanometer Scales Using Time-Resolved X-Ray Diffraction. Phys. Rev. Lett. 118, 247401 (2017).
11. Lee, W. S. et al. Phase fluctuations and the absence of topological defects in a photo-excited charge-ordered nickelate. Nat. Commun. 3, 838 (2012).
12. Zong, A. et al. Evidence for topological defects in a photoinduced phase transition. Nat. Phys. 15 27-31 (2019).
13. Otto, M. R. et al. How optical excitation controls the structure and properties of vanadium dioxide. Proc. Natl. Acad. Sci. U. S. A. 116, 450–455 (2019).
14. Dwyer, J. R. et al. Femtosecond electron diffraction: ‘making the molecular movie’. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. 364, 741–778 (2006).
15. Wall, S. et al. Ultrafast disordering of vanadium dimers in photoexcited VO₂. *Science* **362**, 572–576 (2018).

16. Danz, T., Domröse, T. & Ropers, C. Ultrafast nanoimaging of the order parameter in a structural phase transition. *Science* **371**, 371–374 (2021).

17. Cavalleri, A. et al. Femtosecond Structural Dynamics in VO₂ during an Ultrafast Solid-Solid Phase Transition. *Phys. Rev. Lett.* **87**, 237401 (2001).

18. Cavalleri, A. et al. Picosecond soft x-ray absorption measurement of the photoinduced insulator-to-metal transition in VO₂. *Phys. Rev. B* **69**, 153106 (2004).

19. Cavalleri, A. et al. Band-Selective Measurements of Electron Dynamics in VO₂ Using Femtosecond Near-Edge X-Ray Absorption. *Phys. Rev. Lett.* **95**, 067405 (2005).

20. Wegkamp, D. et al. Instantaneous band gap collapse in photoexcited monoclinic VO₂ due to photocarrier doping. *Phys. Rev. Lett.* **113**, 216401 (2014).

21. Abreu, E. et al. Dynamic conductivity scaling in photoexcited V₂O₃ thin films. *Phys. Rev. B* **92**, 085130 (2015).

22. Hilton, D. J. et al. Enhanced photosusceptibility near Tc for the light-induced insulator-to-metal phase transition in vanadium dioxide. *Phys. Rev. Lett.* **99**, 226401 (2007).

23. Vidas, L. et al. Does VO₂ Host a Transient Monoclinic Metallic Phase? *Phys. Rev. X* **10**, 031047 (2020).

24. Kang, H.-S. et al. Hard X-ray free-electron laser with femtosecond-scale timing jitter. *Nat. Photonics* **11**, 708–713 (2017).

25. Park, S. H. et al. PAL-XFEL soft X-ray scientific instruments and X-ray optics: First commissioning results. *Rev. Sci. Instrum.* **89**, 055105 (2018).

26. Jang, H. et al. Time-resolved resonant elastic soft x-ray scattering at Pohang Accelerator Laboratory X-ray Free Electron Laser. *Rev. Sci. Instrum.* **91**, 083904 (2020).

27. Park, S. H. et al. Scientific instruments for soft X-ray photon-in/photon-out spectroscopy on the PAL-XFEL. *J. Synchrotron Radiat.* **26**, 1031–1036 (2019).

28. Vidas, L. et al. Imaging Nanometer Phase Coexistence at Defects During the Insulator–Metal Phase Transformation in VO₂ Thin Films by Resonant Soft X-ray Holography. *Nano Lett.* **18**, 3449–3453 (2018).
29. Johnson, A. S. et al. Quantitative hyperspectral coherent diffractive imaging spectroscopy of a solid-state phase transition in vanadium dioxide. *Sci. Adv.* 7, eabf1386 (2021).

30. Koethe, T. C. et al. Transfer of spectral weight and symmetry across the metal-insulator transition in VO$_2$. *Phys. Rev. Lett.* 97, 116402 (2006).

31. Haverkort, M. W. et al. Orbital-assisted metal-insulator transition in VO$_2$. *Phys. Rev. Lett.* 95, 196404 (2005).

32. Wall, S. et al. Tracking the evolution of electronic and structural properties of VO$_2$ during the ultrafast photoinduced insulator-metal transition. *Phys. Rev. B* 87, 115126 (2013).

33. Baum, P., Yang, D.-S. & Zewail, A. H. 4D Visualization of Transitional Structures in Phase Transformations by Electron Diffraction. *Science* 318, 788–792 (2007).

34. Tobase, T., Yoshiasa, A., Hiratoko, T. & Nakatsuka, A. Rutile- and anatase-type temperature-dependent pre-edge peak intensities in K-edge XANES spectra for AO (A = Mn), A$_2$O$_3$ (A = Sc, Cr and Mn) and AO$_2$ (A = Ti and V). *J. Synchrotron Radiat.* 25, 1129–1134 (2018).

35. Kucharczyk, D. & Niklewski, Z. Accurate X-ray Determination of the Lattice Parameters and the Thermal Expansion Coefficients of VO$_2$ near the Transition Temperature. *J. Appl. Crystallogr.* 12, 370–373 (1979).

36. Laverock, J. et al. Strain dependence of bonding and hybridization across the metal-insulator transition of VO$_2$. *Phys. Rev. B* 85, 081104 (2012).

37. Thomsen, C., Grahn, H. T., Maris, H. J. & Tauc, J. Surface generation and detection of phonons by picosecond light pulses. *Phys. Rev. B* 34, 4129–4138 (1986).

38. Schick, D. et al. Ultrafast lattice response of photoexcited thin films studied by X-ray diffraction. *Struct. Dyn.* 1, (2014).

39. Büttner, F. et al. Observation of fluctuation-mediated picosecond nucleation of a topological phase. *Nat. Mater.* 20, 30-37 (2021).

40. Del Valle, J. et al. Spatiotemporal characterization of the field-induced insulator-to-metal transition. *Science* 373, 907–911 (2021).

41. Wang, T. et al. Femtosecond Single-Shot Imaging of Nanoscale Ferromagnetic Order in Co / Pd Multilayers Using Resonant X-Ray Holography. *Phys. Rev. Lett.* 108, 267403 (2012).
Supplementary Note 1: Fluence dependence of the phase transition and initial domain stability
For pump-probe experiments, it is vital that the initial state is stable to repeated exposures. Experiments were performed at the I06 beamline at the Diamond light source in which the static domain structure was compared before and after laser excitation. It was found that while a single pulse could change the domain pattern, such changes only occurred in the first laser pulse and the new domain structure was then stable to subsequent laser pulses. However, increasing the laser fluence could induce new changes in the structure. This result was confirmed at PAL-XFEL as shown in extended data figure 1, where we constantly monitored the domain structure at negative time delays to ensure permanent changes were not being detected as time-resolved changes. Occasionally permanent changes were observed. These occurred normally when the laser fluence was changed significantly, but could also change during a run, which were attributed to a higher than average fluence X-ray shot\textsuperscript{11}.

Supplementary Note 2: Effects of global changes in FTH images
Extended data figure 2 shows the effect of the beam block on images when recorded with Holography. Extended data figure 2a shows the actual transmission of the sample. In one case (orange, mixed), the sample consists of two states (insulator and metal). The insulating state has a lower transmission than the metallic state. The second case, (green, metal) is that of a homogenous state. If the mixed state is converted to a homogenous metallic phase, then the only positive changes should be seen in regions that were initially in the insulating phase. Regions that started off metallic remain metallic and thus no changes in transmission occur. As a result, the DC transmission of the sample has changed. When holographic imaging is performed with a beam block, the central region of the scattering pattern is obscured, which acts as a high pass filter and enhances edge contrast. The result of this filter is shown in Extended data figure 2c, which shows the real part of the FTH image for the data shown in Extended data figure 2a. The high pass filter removes the DC component of the image, making homogenous samples look flat and enhancing the contrast at the edges. Importantly the domain structure can still be seen in the heterogeneous sample. When comparing the difference between the homogenous metal and the mixed phase, regions that were initial metallic now show a negative change, even though no real change occurred in order to compensate for a change in the DC level.

Supplementary Note 3: PCA analysis
Principal component analysis decomposes a signal into a series of linear functions which are linearly uncorrelated and when summed describe the total dynamics. Individual PCs are not necessarily directly physically meaningful; however in our case we find a single component that describes all dynamics. We first analyze the noise in our data by examining the PCs of ten repeated measurements at \(-10\) ps over a period of 6 hours. All PCs showed only random fluctuations with laboratory time, while all meaningful structural information was contained in the 0\textsuperscript{th} PC component (PC0). Thus we define our noise floor as the amplitude of the second largest PC (PC1). In analyzing the pump-probe time traces, we discard all PCs below this noise threshold, as shown in Extended data figure 3a. For the short time trace this leaves only two significant PCs. The first (PC0) shows only random fluctuations with pump-probe delay and is again attributed to the instability of the free electron laser intensity, while the second (PC1) exhibits the dynamics discussed in the main text. When including longer time delays a third PC (PC2), is comparable to the noise floor indicating that more spatial components are needed for longer dynamics. This is further supported by explicit examination of the spatial modes (Extended data figure 3b-e), where for short times PC2 already resembles noise while including long times results in PC2 showing strong structure.
We verify that all short-time dynamics are captured in PC0 and PC1 by plotting the difference between the full pump-probe trace and one constructed from only the two most significant PCs. The result is plotted in Extended data figure 4, and shows no significant spatial structure at any time delay, justifying the approach.

**Supplementary Note 4: Analysis of the time dependent data**

The time dependent data were obtained from analysing two experimental runs. One with high time-resolution data over the first picosecond and a second with picosecond data over 20 ps. Between measurements, the initial pre-time-zero domain structure changed. As a result, the two runs could not be directly combined and the PCA amplitude as a function of time needs to be scaled in order to compare the data sets. To this end, a global fit was used to simultaneously fit the dynamics of both data sets. The long-time data, measured out to 20 ps, is shown in Extended data figure 5.

**Supplementary Note 5: In and out of plane strain propagation**

To understand the slower timescale in our data (4.75 ps) we have simulated the effect of both thermal and acoustic propagation in our sample. For thermal effects we solve the 1D heat diffusion equation using the values tabulated in reference 42. We neglect emissivity effects, which should be negligible on the timescales considered here, and focus on heat propagation into the bulk. We consider the situation where the deposited fluence is enough to partially switch the sample, and heat propagation into the insulating region can initiate further changes, leading to a time-dependent metallic fraction. Extended data figure 6a shows the temperature profile as a function of time. As heat propagates from the front surface to the back more of the sample heats to above $T_c$. The volume fraction above $T_c$ is plotted in Extended data figure 6b and shows a slow increase over tens of picoseconds, in disagreement with our observed timescales. We assumed no thermal transmission into the weakly absorbing silicon nitride substrate. Including this factor further reduces the heating rate.

The agreement with the acoustic propagation is far better. We solve the coupled acoustic wave equation in 1D according to the method described in references 37,38, but modified to include a discontinuous volume change at the phase transition in the source term of the stress tensor and damping of the propagating waves. The speed and damping rate of the acoustic waves in both phases are taken from Abreu et al. 7, while the volume discontinuity is taken from diffraction measurements 35. The initial stress field for a homogenously switched sample is shown in Extended data figure 6c, and shows the overall stress field is dominated by the phase transition related volume expansion rather than the thermal effect. The stress immediately begins to relax by the volume expanding into the vacuum and silicon nitride substrate; the mean stress in the VO$_2$ is plotted in Extended data figure 6d and shows a rapid decrease over the first few-picoseconds, undergoing only a single notable reflection before damping away. We note here a perfect boundary is assumed; imperfect acoustic surfaces and sample inhomogeneity will both blur the perfect acoustic dynamics plotted here in real samples. If the sample is only partially switched more complex dynamics are observed due to the mixture of fast and slow propagation in the metallic and insulating volume fractions, respectively.

Note that both these simulations consider propagation in the bulk direction. In-plane, these dynamics are markedly slower due to the disparate length scales. For a pump spot size of 200 µm FWHM, the strain wave would take several nanoseconds to propagate from the edges of the photo-transformed region to the centre; these dynamics are furthermore much more sensitive to experimental geometry than the absorption length or film thickness limited out-of-plane dynamics. Defects could significantly reduce the transverse timescale by providing additional routes to strain relaxation, and may be important for spatially dependent dynamics observed at the 100s of picoseconds timescale. However, transverse thermal propagation will still be slow even in this case.