Light from van der Waals quantum tunneling devices

Markus Parzefall [1], Áron Szabó[2], Takashi Taniguchi[3], Kenji Watanabe [3], Mathieu Luisier[2] & Lukas Novotny[1]

The understanding of and control over light emission from quantum tunneling has challenged researchers for more than four decades due to the intricate interplay of electrical and optical properties in atomic scale volumes. Here we introduce a device architecture that allows for the disentanglement of electronic and photonic pathways—van der Waals quantum tunneling devices. The electronic properties are defined by a stack of two-dimensional atomic crystals whereas the optical properties are controlled via an external photonic architecture. In van der Waals heterostructures made of gold, hexagonal boron nitride and graphene we find that inelastic tunneling results in the emission of photons and surface plasmon polaritons. By coupling these heterostructures to optical nanocube antennas we achieve resonant enhancement of the photon emission rate in narrow frequency bands by four orders of magnitude. Our results lead the way towards a new generation of nanophotonic devices that are driven by quantum tunneling.

[1] Photonics Laboratory, ETH Zürich, 8093 Zürich, Switzerland. [2] Integrated Systems Laboratory, ETH Zürich, 8092 Zürich, Switzerland. [3] National Institute for Material Science, 1-1 Namiki, Tsukuba 305-0044, Japan. Correspondence and requests for materials should be addressed to M.P. (email: mparzefall@ethz.ch) or to L.N. (email: lnovotny@ethz.ch)

The desire to develop an optical implementation of the classical antenna in its role as a transducer between an electronic source and electromagnetic radiation has revived interest in light emission from inelastic electron tunneling (IET). This process, originally discovered in the 1970s[1], allows for the conversion of electrical energy into optical excitations on length scales considerably smaller than what is achievable with other optoelectronic devices. Recently, IET-driven optical antennas were demonstrated[2,3] and continued development has lead to the achievement of greatly improved device efficiencies[4] and directionality[5], the development of scalable fabrication processes[6,7] as well as the exploration of new applications in sensing[8] and catalysis[9].

The vast majority of IET studies, both on integrated devices[1,10–13] as well as in scanning tunneling microscopy (STM)[14–19], are centered around metal-insulator-metal (MIM) tunnel junctions. In these devices the tunnel junction defines both the electrical circuit and the optical mode confinement, which severely limits the degrees of freedom for device design and optimization. Furthermore, as light emission from MIM devices is the result of a two-step process in which the tunneling electrons first excite a propagating or localized surface plasmon polariton (SPP) mode that subsequently decays via radiation, it is difficult to disentangle the influence of optical and electrical device properties on the light emission process.

A device structure that allows the optical mode to be defined independently of the tunnel junction would provide additional degrees of freedom for device design, adaptability and optimization of performance. Here we introduce and present a first realization of van der Waals quantum tunneling (vdWQT) devices, hybrid van der Waals heterostructures[20,21] comprised of a vertical stack of gold (plasmonic metal), hexagonal boron nitride (h-BN, two-dimensional insulator) and graphene (Gr, two-dimensional semimetal), as illustrated in Fig. 1a. We show that the optical mode confinement in these vdWQT devices can be defined independently of the electrical tunnel junction—setting a new paradigm in interfacing optics and electronics at the nanoscale. As graphene is almost transparent to light; only 2.3% of photons propagating through it are being absorbed[22], we observe—in addition to the two-step process—the direct emission of photons from tunnel junctions that are not coupled to MIM cavities. A comparison to a theoretical model allows us to determine IET to be the source of the observed emission.

We further demonstrate that photon emission from IET can be locally enhanced by coupling to an optical antenna, cf. Fig. 1a. vdWQT device-coupled nanocube antennas emit light from a nanoscopic volume, resulting in the diffraction-limited emission pattern shown in Fig. 1b. Nanocube antennas provide a high local density of optical states (LDOS) and give rise to a narrow emission spectrum, cf. Fig. 1c. On resonance, we achieve an increase of the spectral photon emission rate by more than four orders of magnitude.

To fully understand these results and their implications for future research we first analyze the photon emission from vdWQT devices in the absence of optical antennas before studying the effect of antenna-coupling.

Results

Light from uncoupled vdWQT devices. Figure 2a shows an optical microscope image of a set of devices. They are composed of a vertical stack of a gold electrode, a few-layer h-BN crystal and monolayer graphene (cf. Fig. 1a), assembled on top of a glass (SiO$_2$) substrate. Details on the fabrication process are described in the Methods section. The tunneling devices are formed in the regions where all of the three constituting materials overlap, i.e., at the end of the three left electrodes in Fig. 2a. The electrodes on the right-hand side serve as electrical contacts to the graphene sheet.
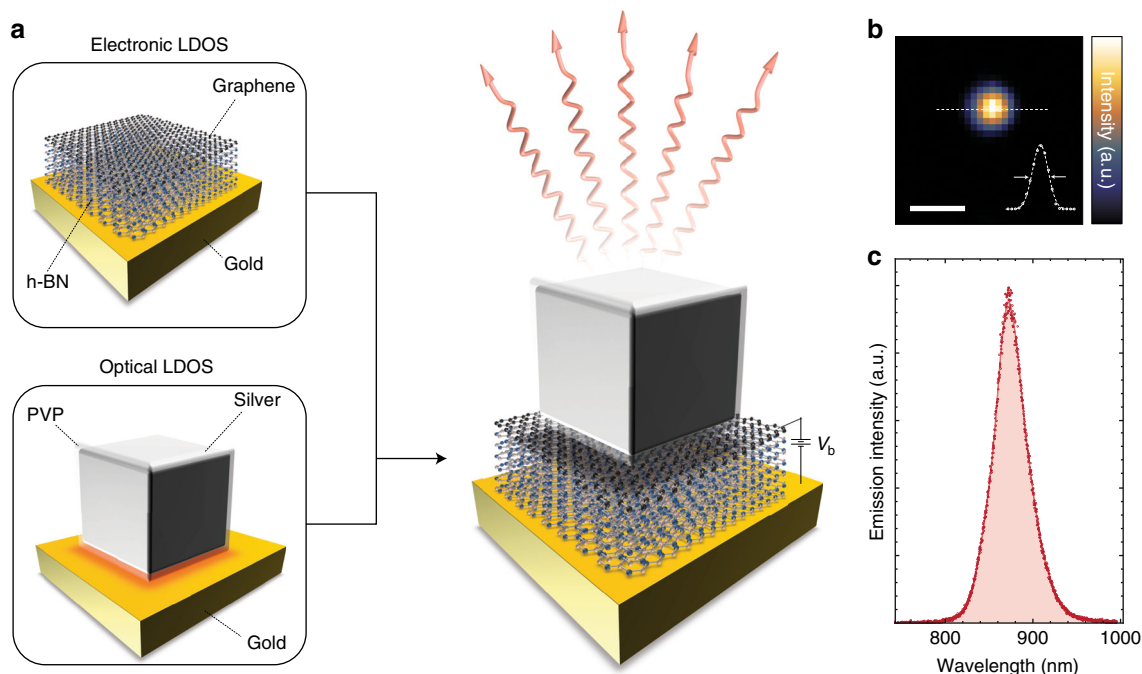

Fig. 1 Visualization of the vdWQT device concept. **a** Illustration (not to scale) of a gold–few-layer h-BN–graphene vdWQT device, integrated with a (silver, PVP-coated) nanocube antenna. In this device configuration, the electronic LDOS is controlled by the hybrid vdW heterostructure whereas the optical LDOS is governed by the nanocube antenna. Applying a voltage V_b across the insulating few-layer h-BN crystal results in antenna-mediated photon emission (wavy arrows) due to quantum tunneling. **b**, **c** Measured spatial (**b**) and spectral (**c**) photon distribution from a nanocube antenna coupled to a vdWQT device, demonstrating a diffraction-limited spot and a narrow emission spectrum. The inset in **b** shows a line-cut, featuring a line-width (FWHM) of ~460 nm, close to the expected value of $\lambda/(2NA)$ ~ 480 nm. Scale bar: 1 μm
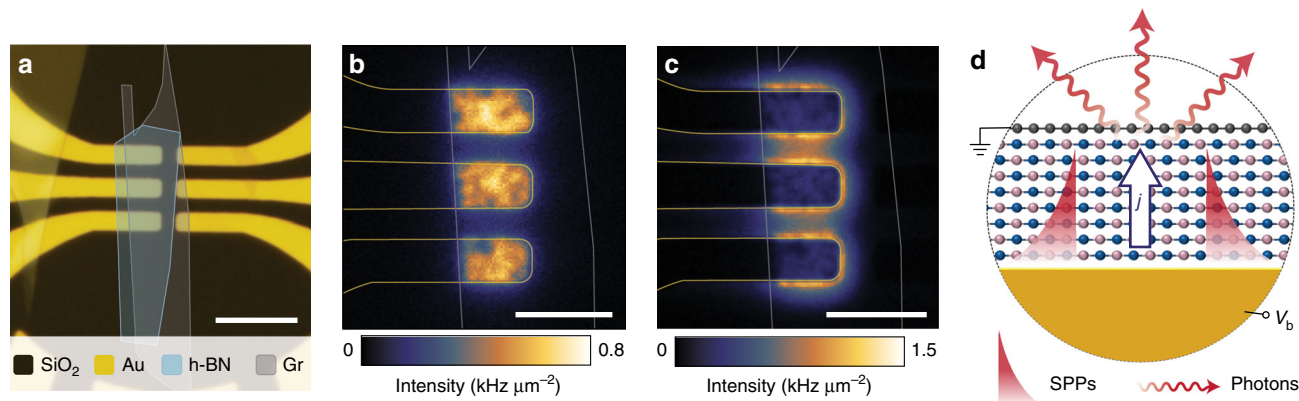
Fig. 2 Light emission from vdWQT devices. **a** Optical microscope image of a set of devices. The three areas on the left, where gold electrodes (50 nm), h-BN crystal ($N = 7$) and graphene overlap, form the tunneling devices. Electrodes on the right serve as electrical contacts to the graphene sheet. Scale bar: 10 μm. **b** Spatial distribution of photons emitted into the air/upper half-space above the graphene sheet from the devices at an applied voltage of 2.0 V. Light is emitted from the entire area of the devices. Scale bar: 5 μm. **c** Corresponding distribution of photons emitted into to substrate/lower half-space. Light is observed primarily at the edges of the devices. Scale bar: 5 μm. Straight lines in **b**, **c** mark the outlines of gold electrodes (gold) and graphene (gray). Intensity units are as measured and not corrected for system efficiency. **d** Sketch of the vertical cross-section through the device, illustrating the mechanism of light generation. The inelastic component of the tunneling current j couples to the optical modes of the vdWQT device, i.e., the free-space continuum of optical states as well as surface-bound SPP modes (not to scale). Photons emitted into the air half-space give rise to the image in **b**, SPP scattering at the edges of the gold electrodes generate the image in **c**

When a voltage between the graphene and gold electrodes is applied, we observe photon emission into both half-spaces, i.e., the air above the device and the glass substrate. Our analysis regarding the origin of this emission is segmented into two parts. First, we analyze our results in terms of the optical modes that are excited. Second, we will determine which physical mechanism causes the mode excitation.

Direct and indirect photon emission. Photons that are emitted into the air half-space give rise to the image shown in Fig. 2b. Evidently, photons are emitted from the entire area of the devices. On the contrary, photons emitted into the glass half-space give rise to the image shown in Fig. 2c, where the emission is localized near the edges of the device. The origin of these images is found by identifying the optical modes of the planar geometry. As we show in Supplementary Note 4, these modes are the radiation continuum of free space and SPP modes bound to the gold surfaces, as illustrated in Fig. 2d (additionally, our geometry supports graphene plasmons which are not considered here, cf. Supplementary Note 4).

On the one hand, tunneling electrons may couple to the radiation continuum, emitting photons without an intermediate excitation. This direct emission process, enabled by the transparency of the top graphene electrode, explains the emission into the top half-space from the device area seen in Fig. 2b. It is important to note that this observation cannot be explained by the random scattering of SPPs by surface roughness as the emission area is restricted to the area of the tunnel junction whereas SPPs are free to propagate across the left edge of the junction area due to the negligible mode mismatch (cf. Supplementary Fig. 3f). On the other hand, photons may also be emitted indirectly as the result of a two-step process. First, a SPP is excited by a tunneling event. Second, since the SPP is bound to the surface it may only emit a photon when it is scattered, that is, at the edges of the gold electrodes. Since photons are preferentially scattered into the medium with the higher refractive index, i.e., glass, we observe photon emission from the edges of the device in Fig. 2c. The relative strength of the coupling to the radiation continuum and the SPP modes is governed by the LDOS associated with each mode. Due to the low
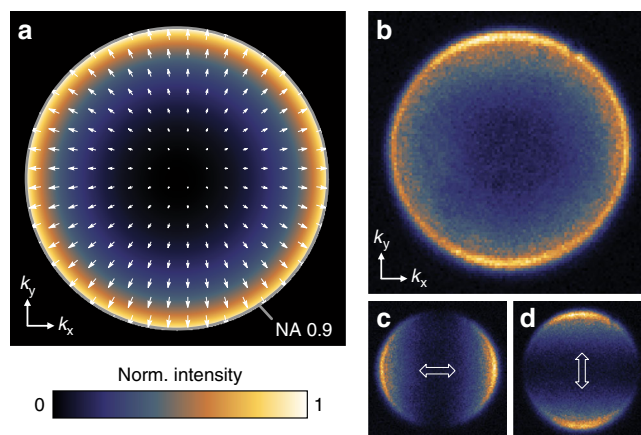

Fig. 3 Radiation pattern of vdWQT device emission. **a** Calculated emission pattern of an electric point dipole located at a distance of 1 nm above a gold half-space, radiating a free space wavelength of 800 nm. Arrows indicate the position-dependent orientation of the electric field vector. The maximum value of the in-plane wave vector component $k_{||} = (k_x^2 + k_y^2)^{1/2}$ is given by $k_{||,max} = \mathrm{NA} \times k_0$ where $\mathrm{NA} = 0.9$ is the numerical aperture of the objective and $k_0 = 2\pi/\lambda$. **b** Measured angular distribution of photons emitted into the upper half-space. **c**, **d** As **b** when measured through a polarizer whose transmitting axis is oriented horizontally/vertically, respectively

mode confinement of the SPP mode the associated LDOS is of the same order of magnitude as the LDOS associated with direct photon emission (cf. Supplementary Note 4), resulting in similar intensities of the two emission pathways, cf. Fig. 2b, c.

To provide further evidence for the direct photon emission process we analyze the angular distribution of photons emitted into the air half-space, shown in Fig. 3. If photons are indeed directly emitted by tunneling electrons their radiation pattern should resemble the radiation pattern of a vertical dipole above a metal surface (cf. Fig. 2d), as calculated in Fig. 3a[23]. The calculation shows that photons are primarily emitted into high angles. Furthermore, as indicated by the white arrows, the
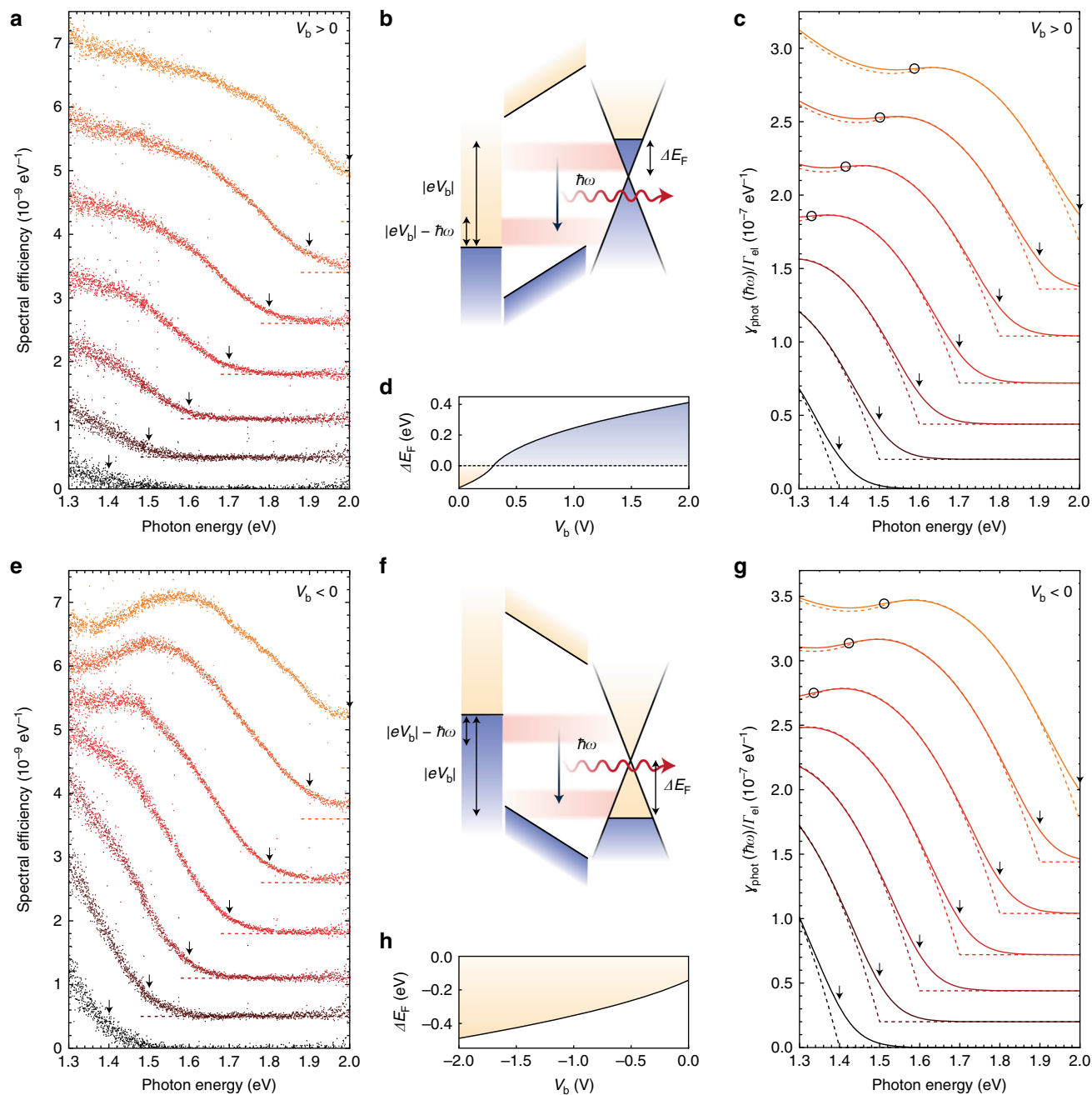
Fig. 4 Spectral distribution of directly emitted photons from vdWQT devices. **a**, **e** Measured efficiency spectra (spectral distribution of emitted photons per tunneling electron, cf. Supplementary Note 1) for positive/negative voltages V_b from a Au–7L h-BN–Gr device. Subsequent spectra are acquired with increasing voltage from $|V_b| = 1.4$ V (black) to $|V_b| = 2.0$ V (orange) in steps of 0.1 V. Spectra are offset with respect to each other, dashed lines indicate the respective zero line. Arrows indicate the spectral position where $\hbar\omega = eV_b$. **b**, **f** Schematic of inelastic electron tunneling at positive/negative voltages. Optical mode excitations (here: photons) are generated by spontaneous transitions from occupied states in one electrode to unoccupied states in the other electrode. The energy window available for spontaneous excitation of a mode of energy $\hbar\omega$ is given by $|eV_b| - \hbar\omega$. **c**, **g** Theoretical photon emission rate spectra ($\rho_{rad} = 0.04\rho_O$) for positive/negative voltages V_b, normalized to the elastic tunneling rate Γ_{el}. Spectra are calculated for voltages from $|V_b| = 1.4$ V to $|V_b| = 2.0$ V in steps of 0.1 V. Straight and dashed lines show calculations for $T = 300$ K and $T = 0$ K, respectively. Black circles mark the positions on each curve for which the condition $|eV_b| - \hbar\omega = |\Delta E_F|$ is fulfilled. Arrows indicate the spectral position where $\hbar\omega = eV_b$. **d**, **h** Electrostatically induced graphene Fermi level shift ΔE_F (cf. Methods) as a function of (positive/negative) applied voltage

emission is radially polarized. Experimentally, this distribution is obtained by imaging the back focal plane of the microscope objective that collects the photons. The measured radiation pattern, as well as the decomposition into two orthogonal polarization directions, is shown in Fig. 3b–d. The experimental results stand in excellent agreement with the direct emission mechanism.

Mode excitation by quantum tunneling. The remaining question relates to the nature of the emission: What is the physical process that leads to the observed light emission? While it is generally accepted that inelastic electron tunneling—where one[1,15] or multiple[24,25] electrons generate an optical excitation while traversing the barrier—is the dominating mechanism in light emission from MIM tunnel junctions[13,15], it is not a given

that the same is true for the devices presented here due to the two-dimensional nature of the graphene electrode and its unique band structure[26]. The different mechanisms that are discussed in the context of light emission from tunnel junctions are

1. Inelastic electron tunneling.
2. Hot electron photoemission, where the electron first traverses the barrier and subsequently transitions to an energetically lower-lying unoccupied state through the generation of an optical excitation[27], and
3. Thermal emission, where electron tunneling gives rise to a 'hot' electron distribution, which in turn generates blackbody-like radiation[28,29].

Irrespective of the mechanism, the spectrum of the light that is emitted is ultimately not only determined by the mechanism itself but also by the frequency-dependence of the LDOS (ρ_{opt}) and radiation efficiency (η_{rad}) of the associated modes in a given sample geometry. In particular, for the case of light emission from MIM tunnel junctions, this interplay of a series of frequency-dependent factors has made it historically difficult to distinguish between the different physical mechanisms that might be at work[13,27]. The direct photon emission from our devices is governed by an optical mode density that is nearly constant in frequency (cf. Supplementary Note 4), allowing us to get an unobstructed view on the underlying mechanisms.

To identify the dominating emission mechanism, we now turn to the spectral distribution of these directly emitted photons, shown for applied voltages in the range from 1.4 to 2.0 V in Fig. 4a, e. First, we find that the spectral efficiency of light emission is very similar for both bias polarities. In MIM devices this finding is in agreement with both inelastic electron tunneling and hot electron photoemission because of the symmetry of the device. The efficiency of hot electron photoemission, that relies on the spontaneous generation of an optical excitation after reaching the counter electrode, depends on the efficacy of competing decay channels. With hot carrier lifetimes in graphene and gold being of the order of 100 fs[30] and 10 fs[31], respectively, one would expect a pronounced bias asymmetry in the emission efficiency, which is not observed. Second, we observe a clear correlation between the spectral cutoff $\hbar\omega_{max}$ and the applied voltage V_b given by $\hbar\omega_{max} \sim |eV_b|$. This indicates that multiple-electron IET processes do not play a dominant role[24,25]. Furthermore, it is not compliant with blackbody-like thermal emission, where the emitted spectrum is determined by the effective temperature of the electron gas rather than the applied voltage[28,29]. Hence, we identify single-electron IET to be the most likely mechanism underlying the observed emission.

Theoretical device modeling. To validate this hypothesis we compare our experimental results with theoretical calculations, describing the IET process—as depicted in Fig. 4b, f—as a spontaneous transition between occupied electronic states of one electrode to unoccupied states of the other electrode[13,32–35]. Based on Fermi's golden rule and within the validity range of the dipole approximation, we may express the spectral rate of inelastic electron tunneling as[35]

$$\gamma_{inel}(\hbar\omega) = \frac{\pi e^2}{3\hbar\omega m^2 \varepsilon_0} \rho_{opt}(\hbar\omega)$$
$$\times \int_{\hbar\omega}^{eV_b} |\mathcal{P}(E, \hbar\omega)|^2 \rho_{Au}(E - \hbar\omega)\rho_{Gr}(E)\mathrm{d}E, \qquad (1)$$

where ρ_{opt} is the partial LDOS along the direction of electron tunneling, $\mathcal{P}(E, \hbar\omega)$ is the momentum matrix element (cf. Supplementary Note 2), ω is the angular frequency, $\hbar\omega$ the energy of the optical excitation and $\rho_{Au/Gr}$ is the electronic density of states of gold/graphene. The equation as stated applies to the positive

voltage range. For negative voltages, the roles of the two electronic state densities are interchanged. The process is schematically depicted in Fig. 4b, f for positive/negative bias. At positive voltages electrons excite optical modes by transitioning from occupied states in the graphene electrode to unoccupied states in the gold electrode. Correspondingly, at negative voltages, optical modes are excited by transitions from occupied states in the gold electrode to unoccupied states in the graphene electrode. At any given mode energy $\hbar\omega$ the energetic range over which inelastic electron tunneling may occur is given by the applied voltage as $|eV_b| - \hbar\omega$. While $\mathcal{P}$ and ρ_{Au} are approximately constant (cf. Supplementary Note 3), ρ_{Gr} depends linearly on ΔE_F, cf. Eq. (4). It is important to realize that the gold layer does not only act as one of the electrodes of the tunnel junction but also as a gate electrode[36,37]. This causes the charge carrier density of the graphene sheet and hence the shift of the Fermi level with respect to the Dirac point (ΔE_F) to depend on the tunnel voltage, as depicted in Fig. 4d, h, cf. Methods. Differences in the band alignment and dipole layers at the gold/h-BN and graphene/h-BN interfaces result in a finite value of ΔE_F even at $V_b = 0$, cf. Supplementary Note 1.

From Eq. (1) we calculate the spectral photon emission rate as $\gamma_{phot} = \eta_{rad} \times \gamma_{inel} = \rho_{rad} \times \gamma_{inel}/\rho_{opt}$, where η_{rad} is the radiation efficiency and $\rho_{rad} = \rho_{opt} \times \eta_{rad}$ is the radiative LDOS. For the planar heterostructure we find that $\rho_{rad} = 0.04\rho_0$, where $\rho_0 = \omega^2 \pi^{-2} c^{-3}$ is the vacuum LDOS and c is the speed of light in vacuum (cf. Supplementary Note 4). The resulting γ_{phot} spectra, normalized by the calculated elastic tunneling rate Γ_{el} (cf. Supplementary Note 1), are displayed in Fig. 4c, where dashed and straight lines correspond to calculations for $T = 0$ K and $T = 300$ K, respectively. At $T = 0$ K, the quantum cut-off condition $\hbar\omega_{max} = |eV_b|$ is perfectly fulfilled such that the emission rate is zero for energies larger than the applied voltage. At finite temperatures this condition is relaxed by the thermal broadening due to the Fermi-Dirac distribution function, in agreement with experimental results. Most importantly, we are able to reproduce the experimentally observed spectral shape, confirming IET to be the underlying physical mechanism. It is determined by the electronic density of states of graphene, in particular by the voltage-dependent energetic position of the Dirac point (cf. Fig. 4d, h and Methods). Black circles on each curve mark the points for which the condition $|eV_b| - \hbar\omega = |\Delta E_F|$ is fulfilled. This condition determines the spectral location of the observed inflection point. The effect of the density of states minimum is more strongly pronounced for negative voltages. In our modeling we assume a symmetric and energy-independent electronic wave function decay (cf. Supplementary Note 2). Deviations from this assumption, as well as potential effects related to the gate-dependence of the barrier height at the graphene–h-BN interface[38], are possible causes for the observed asymmetry. We emphasize that no fitting parameters are employed in the theoretical model with the exception of the extraction of the residual Fermi level shift at zero applied bias (cf. Supplementary Note 1).

Comparing the absolute values of the experimental and theoretical spectra of Fig. 4a, c, respectively, we find that the experimental efficiencies are lower than the theoretically calculated ones by a factor of ~40. This deviation is caused by the presence of an additional phonon-enhanced tunneling channel in metal–graphene tunnel junctions[39–42], see Supplementary Note 1 for details.

vdWQT devices coupled to nanocube antennas. Until now we have analyzed the coupling of tunneling electrons to photons and SPPs of a planar, layered heterostructure—the optical properties
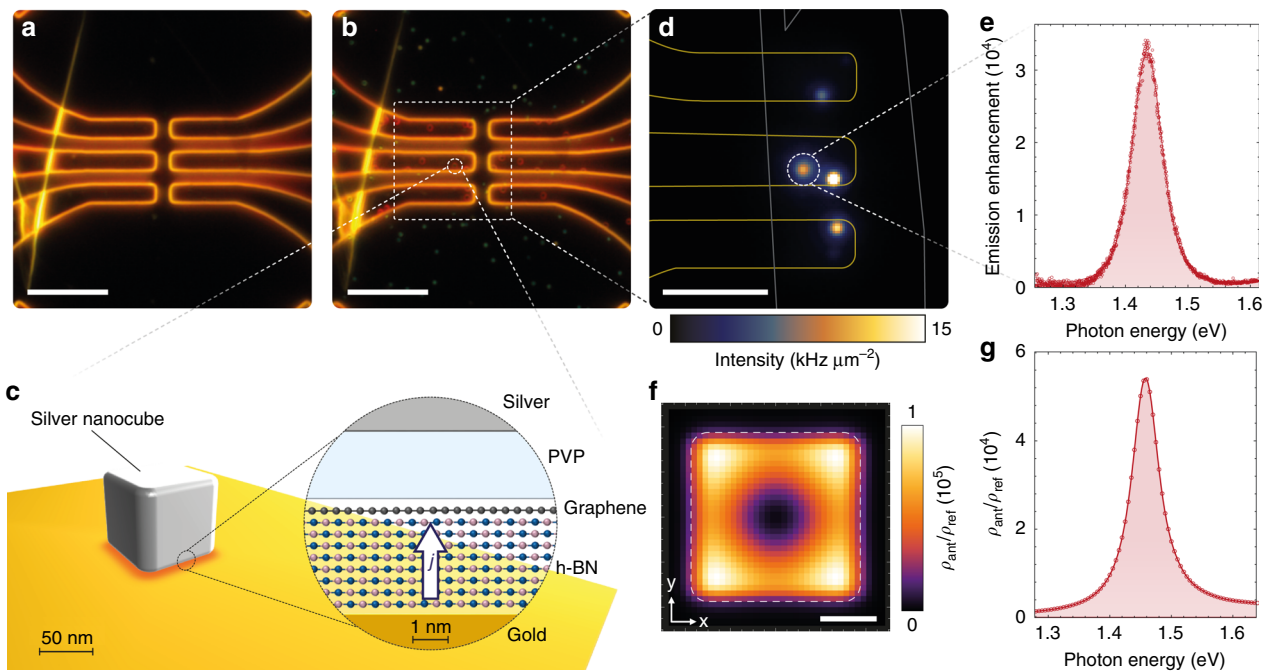
Fig. 5 Nanocube antennas driven by a vdWQT device. **a** Darkfield microscope image of the same devices as shown in Fig. 2a. Scale bar: 10 µm. **b** Darkfield image after the deposition of silver nanocubes onto the sample. The red ring-like patterns on top of the electrodes indicate the presence of nanocube antennas. Scale bar: 10 µm. **c** Schematic illustration of the nanocube antenna configuration. The inset shows a vertical cross-section of the modified heterostructure configuration which determines the optical modes "seen" by tunneling electrons. **d** Spatial intensity distribution of light emission from nanocube antennas. Outlines indicate the spatial location of the edges of graphene and gold electrodes. Scale bar: 5 µm. **e** Emission spectrum of the antenna marked by a dashed circle in **d** at $V_b = 2$ V, normalized by the corresponding emission spectrum of the planar heterostructure (Fig. 4a). **f** Spatial (on resonance) and **g** spectral dependence of the simulated, radiative fraction of the optical mode density of the nanocube antenna ρ_{ant} for a vertically oriented dipole located in the center of the h-BN crystal, normalized by the radiative LDOS of the planar geometry ρ_{ref}, cf. Supplementary Note 4. The white dashed line in **f** marks the outline of the silver nanocube (75 nm with a 7.5 nm edge radius)

of which are characteristic of a single gold film with a nearly energy-independent LDOS. Our findings demonstrate that we can view tunneling electrons as quantum emitters with emission properties that depend on the electronic density of states as well as the optical LDOS, cf. Eq. (1), and can thus be controlled by coupling to optical antennas[43,44] or cavities[45] to increase emission rates and shape emission spectra and emission pattern. To illustrate this concept we couple the vdWQT devices to nanocube antennas[46,47].

Silver nanocubes are deposited on top of the vdWQT devices from solution (cf. Methods) and can be directly observed with a darkfield microscope. Before nanocube deposition, as seen in Fig. 5a, scattering is dominated by the edges of the gold film. After nanocube deposition, as seen in Fig. 5b, we additionally observe randomly distributed spots, which are associated with optical resonances formed in the junctions between nanocubes and gold electrodes, as schematically illustrated in Fig. 5c. It is interesting to note that nanocubes sitting on top of the gold electrodes and the heterostructure produce a red ring-shaped pattern whereas (single) nanocubes sitting on the glass surface appear green. The ring-like pattern is caused by the symmetry of a higher-order MIM resonator mode that dominates the scattering response in the visible spectral range, cf. ref. [48]. The green color of nanocubes on glass originates from the nanocubes dipolar resonance. The distance between the two metallic domains is determined by the number of h-BN layers and the thickness of the PVP coating (~3 nm) that stabilizes the nanocubes in solution.

While the formation of nanocube antennas leaves the electronic properties, i.e., the elastic and phonon-enhanced tunneling channels (c.f. Supplementary Note 1), largely

unaffected—it strongly alters the light emission properties of the device. Without nanocubes we observe emission from the entire device area, cf. Fig. 2b. After antenna-coupling, as shown in Fig. 5d, the spatial intensity distribution is dominated by point-like sources, the size of which is given by the point-spread function of our imaging system. In the following we will focus on and selectively analyze (cf. Methods) the emission spot marked by the dashed circle in Fig. 5d. Further cube antennas are discussed in Supplementary Note 5. Figure 5e shows the photon emission spectrum of this particular nanocube antenna, normalized by the corresponding emission spectrum of the planar heterostructure (Fig. 4a) from an area equivalent to the nanocube footprint (75 × 75 nm², cf. Supplementary Note 5). Previously, before the deposition of nanocubes, the emission spectrum was broad (cf. Fig. 4a) and primarily determined by the electronic density of states of the electrodes, whereas now photons are emitted into a narrow spectral window centered at ~1.43 eV with a full width at half maximum (FWHM) of ~57 meV. Compared to the emission spectrum of the bare vdWQT device, we find that—on resonance —the probability for an electron tunneling event, occurring within the active area of the nanocube antenna, to result in the generation of a photon is enhanced by more than four orders of magnitude. This enhancement is independent of applied voltage and polarity (cf. Supplementary Note 5).

The origin of these alterations is found by analyzing the modified optical properties of the system. According to Eq. (1), the spectral rate of photon emission by inelastic electron tunneling is determined by the optical mode density $\rho_{opt}(\hbar\omega)$. The formation of nanocube antennas modifies $\rho_{opt}(\hbar\omega)$ in two ways. Firstly, SPPs are confined to the nanoscale gap between the silver nanocube and the gold surface. The optical mode density of

these MIM-SPPs is strongly enhanced as a result of the spatial field confinement and the increased propagation constant. Additionally, the termination of the MIM gap by the facets of the cube act as reflectors—forming a resonator—which leads to an additional increase in the mode density by the Purcell effect[49–51]. While ρ_{opt} determines the rate of inelastic electron tunneling γ_{inel}, the fraction associated with photon emission γ_{phot} is determined by the radiative LDOS ρ_{rad}.

Classically, the LDOS enhancement ρ_{opt}/ρ_0 corresponds to an increase of the power dissipated by a point dipole P/P_0 due to its interaction with the environment[23]. This correspondence allows us to determine the LDOS by means of classical electrodynamics (cf. Methods and Supplementary Note 4). The spectral variation of the radiative LDOS of the nanocube antenna ρ_{ant}, normalized by the radiative LDOS of the planar geometry ρ_{ref} (cf. Supplementary Note 4), is shown in Fig. 5g. The radiative LDOS is resonantly enhanced at a photon energy of $\hbar\omega_{res}$ ~1.46 eV and over a spectral range given by the FWHM of ~57 meV. We further analyze the nature of the enhancement by varying the position of the dipole, oscillating at ω_{res}, as shown in Fig. 5f. The lateral variation of ρ_{rad} maps out the fundamental mode of the MIM resonator. The average LDOS enhancement is found to be ~5×10^4, in excellent agreement with the experimentally observed photon emission rate enhancement of ~3×10^4. These findings validate the dependence of the strength of the inelastic electron tunneling channel on the LDOS experienced by tunneling electrons, as suggested by Eq. (1).

Discussion

In conclusion, we have introduced and experimentally demonstrated a device platform that allows for the excitation of optical modes at the atomic scale by inelastic electron tunneling—van der Waals quantum tunneling devices. In particular, we have demonstrated the independent control over the electronic and optical properties of the system, both of which determine the spectral rate of inelastic electron tunneling in the form of the respective densities of states, cf. Eq. (1).

The electronic properties and the optical source spectrum are determined by a hybrid van der Waals heterostructure, here: gold–hexagonal boron nitride–graphene. The optical modes of this structure are the photon modes of free space and SPP modes. Based on a careful analysis of the spatial, angular and spectral distribution of the emitted photons, we unambiguously determine the direct emission of light from tunneling electrons to be the source of the observed emission. The emission is well described by a spontaneous emission model, confirming inelastic electron tunneling as its source. The spectral flatness of the optical mode density of the planar heterostructure allows us to observe the influence of the electronic density of states on the emission spectrum.

The optical properties are further modified by an external photonic architecture, here: nanocube antennas. The integration with these optical antennas leads to a spectral shaping and enhancement of the photon emission rate. Nanocube antennas constitute nanoscopic resonators for MIM-SPPs, accompanied by a local enhancement of the LDOS. Additionally, the geometry allows for the efficient coupling of MIM-SPPs to photons, resulting in the observed enhancement of the photon emission rate by more than four orders of magnitude.

This concept is easily translated to more complex systems. To optimize the source efficiency or introduce other device functionalities one may choose materials constituting the van der Waals heterostructure[20,21] from an entire library of two-dimensional atomic crystals[52]. For example, utilizing

semiconducting atomic crystals enables the shaping of the energy-dependence of the electronic density of states and may allow for the suppression of the elastic tunneling channel in favor of the inelastic process. In the devices presented here the source efficiency is reduced compared to traditional MIM tunneling devices due to the dominance of the phonon-assisted tunneling channel (cf. Supplementary Note 1). However, this may be overcome in future device configurations since tunneling between identical atomic crystals may allow for significantly higher efficiencies due to favorable momentum selection rules that have not yet been harnessed in any other material system[53–55].

The individual two-dimensional crystals constituting vdWQT devices very weakly interact with optical modes that are polarized perpendicular to the sample plane as shown here for the case of the long-range SPPs of the planar geometry and the MIM-SPP mode inside the nanocube antenna. The same modes however couple most efficiently to tunneling electrons. This feature of vdWQT devices may be further explored by their integration with optical systems that support corresponding modes such as certain types of photonic cavities[45], hybrid plasmonic waveguide resonators[56,57], or SPP resonators[58–60].

Methods

Device fabrication. Devices are fabricated on commercially available glass coverslips ($22 \times 22 \times 0.13$ mm^3). Electrodes are defined by photolithography and electron beam evaporation of 1 nm titanium (Ti, 0.05 nm/s) and 50 nm gold (Au, 0.1 nm/s) at a pressure of $<10^{-7}$ mbar. Lift-off is performed sequentially in acetone, isopropyl alcohol and deionized water. The RMS roughness of the gold electrodes is 0.4–0.5 nm.

Graphene and few-layer h-BN are mechanically exfoliated from bulk crystals onto oxidized silicon wafers (100 nm SiO$_2$). h-BN crystals were grown as described in ref. [61]. Suitable crystals are identified optically. Their thickness, i.e., the number of layers, is determined by a combination of optical contrast and atomic force microscopy. h-BN crystals with thicknesses of 6–7 layers (corresponding to ~2–2.3 nm) are chosen as a compromise between the voltage that may be applied without reaching the electrostatic breakdown field of ~1 V/nm and the tunneling current which decreases exponentially with crystal thickness[62,63]. To assemble the tunneling devices, graphene and few-layer h-BN are picked up successively with a polydimethylsiloxane (PDMS)/polycarbonate (PC) stack. The vdW stack is subsequently transferred on top of the gold electrodes. For pickup and transfer we follow the procedure developed by Zomer et al.[64]. The PC film employed as a carrier is dissolved in chloroform after successful transfer. Finally, the samples are annealed at 200 °C in a tube furnace while flowing 380 sccm argon and 20 sccm hydrogen to reduce residues on top of the heterostructure.

Nanocube antennas are assembled by drop-casting 50 μL of silver nanocubes (size: 78 ± 5 nm) in ethanol (concentration approximately 0.1 mg/mL, purchased from Nanocomposix) on top of the gold–h-BN–graphene assembly. After a short incubation time of several seconds the sample is spun dry at ~3000 rpm.

Following this approach we have assembled ten gold–h-BN–graphene heterostructures with varying numbers of individual devices and nanocube antennas, the properties of which are in agreement with the representative set of experimental data and their evaluation shown in this work.

Optical and electrical device characterization. The devices are characterized on a customized, inverted Nikon TE300 microscope under ambient conditions. Electrical measurements are carried out using a Keithley 2602B source meter unit. Light emitted from the devices into the air half-space is collected by an air objective (Nikon, 100×, NA 0.9) whereas light emitted into the glass half-space is collected by an oil-immersion objective (Nikon, 100×, NA 1.4). For real space and Fourier space (backfocal plane) imaging, the respective plane is imaged onto a electron-multiplying charge-coupled device (EMCCD, Andor iXon Ultra). The emission from individual nanocube antennas is selectively analyzed by spatial filtering in an intermediate image plane using a pinhole with a diameter of 100 μm, corresponding to a diameter of 1 μm in the sample plane. Spectra are acquired using a Princeton Instruments Acton SpectraPro 300i spectrometer. Spatial intensity distribution maps as shown in Figs. 2b, c and 5d are not corrected for system transmission, detection efficiency and the point spread function of the imaging system that determines the size of the emission spots in Fig. 5d. Spectral efficiency and photon flux spectra as shown in Figs. 4a, e and 5e and Supplementary Figs. 4a, 5a are corrected for the transmission function of the optical system and detection efficiency. They are further normalized to the area from which the photons are emitted, i.e., the entire device area in Fig. 4a and the footprint of the nanocube antenna in Fig. 5e as well as Supplementary Figs. 4a, 5a.

Device electrostatics. The dependence of the charge carrier density of graphene (n_{Gr}) on the applied voltage can be described as

$$e(V_b - V_0) = \frac{e^2 n_{Gr}}{c_g} + \Delta E_F, \qquad (2)$$

where $c_g = \varepsilon_0 \varepsilon_{hBN}/d_{hBN}$ is the geometric capacitance per unit area, ε_0 is the vacuum permittivity and ε_{hBN} and $d_{hBN} = N \times d_{il}$ are the DC permittivity and thickness of the h-BN crystal, respectively, with N being the number of h-BN layers and $d_{il} = 0.33$ nm being the h-BN interlayer distance[65]. The second term, ΔE_F, accounts for the quantum capacitance of the two-dimensional graphene sheet[66,67]. The charge carrier density is related to the electronic density of states (ρ_{Gr}) through

$$n_{Gr} = \int_0^{\Delta E_F} \rho_{Gr}(E) dE. \qquad (3)$$

Near the Dirac point energy E_D, ρ_{Gr} depends linearly on energy as[26]

$$\rho_{Gr}(E) = \frac{2|E - E_D|}{\pi \hbar^2 v_F^2}, \qquad (4)$$

where v_F is the Fermi velocity of graphene.

Combining Eqs. 2–4 results in a quadratic equation in ΔE_F that allows us to explicitly calculate ΔE_F as a function of voltage. For calculations, we assume $v_F = 1.1 \times 10^6$ m/s[68] and $\varepsilon_{hBN} = 3.5$[69].

Local density of optical states. The inelastic electron tunneling rate, described in Eq. (1), depends on the LDOS ρ_{opt}. For uncoupled vdWQT devices we calculate the LDOS semi-analytically (Supplementary Note 4). The LDOS of antenna-coupled vdWQT devices is calculated numerically (Supplementary Note 4). In both approaches we make use of the fact that the normalized LDOS ρ_{opt}/ρ_0 is equivalent to the normalized power P/P_0 dissipated by a classical point dipole with moment $\mathbf{p}$[23].

Data availability
The data that support the findings of this study are available from the corresponding author upon reasonable request.

References
1. Lambe, J. & McCarthy, S. L. Light emission from inelastic electron tunneling. *Phys. Rev. Lett.* **37**, 923–925 (1976).
2. Parzefall, M. et al. Antenna-coupled photon emission from hexagonal boron nitride tunnel junctions. *Nat. Nanotechnol.* **10**, 1058–1063 (2015).
3. Kern, J. et al. Electrically driven optical antennas. *Nat. Photon.* **9**, 582–586 (2015).
4. Qian, H. et al. Efficient light generation from enhanced inelastic electron tunnelling. *Nat. Photon.* **12**, 485–488 (2018).
5. Gurunarayanan, S. P. et al. Electrically driven unidirectional optical nanoantennas. *Nano Lett.* **17**, 7433–7439 (2017).
6. Du, W. et al. On-chip molecular electronic plasmon sources based on self-assembled monolayer tunnel junctions. *Nat. Photon.* **10**, 274–280 (2016).
7. Namgung, S. et al. Ultrasmall plasmonic single nanoparticle light source driven by a graphene tunnel junction. *ACS Nano* **12**, 2780–2788 (2018).
8. Dathe, A., Ziegler, M., Hübner, U., Fritzsche, W. & Stranik, O. Electrically excited plasmonic nanoruler for biomolecule detection. *Nano Lett.* **16**, 5728–5736 (2016).
9. Wang, P., Krasavin, A. V., Nasir, M. E., Dickson, W. & Zayats, A. V. Reactive tunnel junctions in electrically driven plasmonic nanorod metamaterials. *Nat. Nanotechnol.* **13**, 159 (2018).
10. Davis, L. C. Theory of surface-plasmon excitation in metal-insulator-metal tunnel junctions. *Phys. Rev. B* **16**, 2482–2490 (1977).
11. Hansma, P. K. & Broida, H. P. Light emission from gold particles excited by electron tunneling. *Appl. Phys. Lett.* **32**, 545–547 (1978).
12. Hone, D., Mühlschlegel, B. & Scalapino, D. J. Theory of light emission from small particle tunnel junctions. *Appl. Phys. Lett.* **33**, 203–204 (1978).
13. Persson, B. N. J. & Baratoff, A. Theory of photon emission in electron tunneling to metallic particles. *Phys. Rev. Lett.* **68**, 3224–3227 (1992).
14. Johansson, P., Monreal, R. & Apell, P. Theory for light emission from a scanning tunneling microscope. *Phys. Rev. B* **42**, 9210–9213 (1990).
15. Berndt, R., Gimzewski, J. K. & Johansson, P. Inelastic tunneling excitation of tip-induced plasmon modes on noble-metal surfaces. *Phys. Rev. Lett.* **67**, 3796–3799 (1991).
16. Berndt, R. & Gimzewski, J. K. Photon emission in scanning tunneling microscopy: Interpretation of photon maps of metallic systems. *Phys. Rev. B* **48**, 4746–4754 (1993).
17. Sakurai, M., Thirstrup, C. & Aono, M. New aspects of light emission from STM. *Appl. Phys. A* **80**, 1153–1160 (2005).
18. Rossel, F., Pivetta, M. & Schneider, W.-D. Luminescence experiments on supported molecules with the scanning tunneling microscope. *Surf. Sci. Rep.* **65**, 129–144 (2010).
19. Bharadwaj, P., Bouhelier, A. & Novotny, L. Electrical excitation of surface plasmons. *Phys. Rev. Lett.* **106**, 226802 (2011).
20. Geim, A. & Grigorieva, I. Van der Waals heterostructures. *Nature* **499**, 419–425 (2013).
21. Novoselov, K. S., Mishchenko, A., Carvalho, A. & Castro Neto, A. H. 2D materials and van der Waals heterostructures. *Science* **353**, aac9439 (2016).
22. Nair, R. R. et al. Fine structure constant defines visual transparency of graphene. *Science* **320**, 1308–1308 (2008).
23. Novotny, L. & Hecht, B. *Principles of Nano-Optics*. 2nd edn, (Cambridge University Press, Cambridge, 2012).
24. Schull, G., Néel, N., Johansson, P. & Berndt, R. Electron-plasmon and electron-electron interactions at a single atom contact. *Phys. Rev. Lett.* **102**, 057401 (2009).
25. Peters, P.-J. et al. Quantum coherent multielectron processes in an atomic scale contact. *Phys. Rev. Lett.* **119**, 066803 (2017).
26. Castro Neto, A. H., Guinea, F., Peres, N. M. R., Novoselov, K. S. & Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109–162 (2009).
27. Kirtley, J., Theis, T., Tsang, J. & DiMaria, D. Hot-electron picture of light emission from tunnel junctions. *Phys. Rev. B* **27**, 4601–4611 (1983).
28. Downes, A., Dumas, P. & Welland, M. E. Measurement of high electron temperatures in single atom metal point contacts by light emission. *Appl. Phys. Lett.* **81**, 1252–1254 (2002).
29. Buret, M. et al. Spontaneous hot-electron light emission from electron-fed optical antennas. *Nano Lett.* **15**, 5811–5818 (2015).
30. Breusing, M. et al. Ultrafast nonequilibrium carrier dynamics in a single graphene layer. *Phys. Rev. B* **83**, 153410 (2011).
31. Brown, A. M., Sundararaman, R., Narang, P., Goddard, W. A. III & Atwater, H. A. Nonradiative plasmon decay and hot carrier dynamics: effects of phonons, surfaces, and geometry. *ACS Nano* **10**, 957–966 (2015).
32. Schimizu, T., Kobayashi, K. & Tsukada, M. Theoretical approach to the microscopic mechanism of light emission from a scanning tunneling microscope. *Appl. Surf. Sci.* **60-61**, 454–459 (1992).
33. Downes, A., Taylor, M. E. & Welland, M. E. Two-sphere model of photon emission from the scanning tunneling microscope. *Phys. Rev. B* **57**, 6706–6714 (1998).
34. Uskov, A. V., Khurgin, J. B., Protsenko, I. E., Smetanin, I. V. & Bouhelier, A. Excitation of plasmonic nanoantennas with nonresonant and resonant electron tunnelling. *Nanoscale* **8**, 14573 (2016).
35. Parzefall, M., Bharadwaj, P. & Novotny, L. in *Quantum Plasmonics* (eds Bozhevolnyi, S. I., Martin-Moreno, L. & Garcia-Vidal, F.) 211–236 (Springer, Cham, 2017).
36. Bokdam, M., Khomyakov, P. A., Brocks, G., Zhong, Z. & Kelly, P. J. Electrostatic doping of graphene through ultrathin hexagonal boron nitride films. *Nano Lett.* **11**, 4631–4635 (2011).
37. Britnell, L. et al. Field-effect tunneling transistor based on vertical graphene heterostructures. *Science* **335**, 947–950 (2012).
38. Ponomarenko, L. et al. Field-effect control of tunneling barrier height by exploiting graphene's low density of states. *J. Appl. Phys.* **113**, 136502 (2013).
39. Brar, V. W. et al. Scanning tunneling spectroscopy of inhomogeneous electronic structure in monolayer and bilayer graphene on sic. *Appl. Phys. Lett.* **91**, 122102 (2007).
40. Zhang, Y. et al. Giant phonon-induced conductance and scanning tunneling spectroscopy of gate-tunable graphene. *Nat. Phys.* **4**, 627–630 (2008).
41. Wehling, T. O., Grigorenko, I., Lichtenstein, A. I. & Balatsky, A. V. Phonon-mediated tunneling into graphene. *Phys. Rev. Lett.* **101**, 216803 (2008).
42. Decker, R. et al. Local electronic properties of graphene on a BN substrate via scanning tunneling microscopy. *Nano Lett.* **11**, 2291–2295 (2011).
43. Bharadwaj, P., Deutsch, B. & Novotny, L. Optical antennas. *Adv. Opt. Phot.* **1**, 438–483 (2009).
44. Biagioni, P., Huang, J.-S. & Hecht, B. Nanoantennas for visible and infrared radiation. *Rep. Prog. Phys.* **75**, 024402 (2012).
45. Vahala, K. J. Optical microcavities. *Nature* **424**, 839 (2003).
46. Moreau, A. et al. Controlled-reflectance surfaces with film-coupled colloidal nanoantennas. *Nature* **492**, 86 (2012).
47. Akselrod, G. M. et al. Probing the mechanisms of large purcell enhancement in plasmonic nanoantennas. *Nat. Photon.* **8**, 835–840 (2014).

48. Chikkaraddy, R. et al. How ultranarrow gap symmetries control plasmonic nanocavity modes: from cubes to spheres in the nanoparticle-on-mirror. *ACS Photonics* **4**, 469–475 (2017).

49. Bozhevolnyi, S. I. & Søndergaard, T. General properties of slow-plasmon resonant nanostructures: nano-antennas and resonators. *Opt. Expr.* **15**, 10869–10877 (2007).

50. Kuttge, M., Garcia de Abajo, F. J. & Polman, A. Ultrasmall mode volume plasmonic nanodisk resonators. *Nano Lett.* **10**, 1537–1541 (2009).

51. Faggiani, R., Yang, J. & Lalanne, P. Quenching, plasmonic, and radiative decays in nanogap emitting devices. *ACS Photonics* **2**, 1739–1744 (2015).

52. Mounet, N. et al. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nat. Nanotechnol.* **13**, 246 (2018).

53. Svintsov, D., Vyurkov, V., Ryzhii, V. & Otsuji, T. Voltage-controlled surface plasmon-polaritons in double graphene layer structures. *J. Appl. Phys.* **113**, 053701 (2013).

54. Svintsov, D., Devizorova, Z., Otsuji, T. & Ryzhii, V. Plasmons in tunnel-coupled graphene layers: Backward waves with quantum cascade gain. *Phys. Rev. B* **94**, 115301 (2016).

55. de Vega, S. & Garcia de Abajo, F. J. Plasmon generation through electron tunneling in graphene. *ACS Photonics* **4**, 2367–2375 (2017).

56. Oulton, R. F. et al. Plasmon lasers at deep subwavelength scale. *Nature* **461**, 629–632 (2009).

57. Lu, Y.-J. et al. Plasmonic nanolaser using epitaxially grown silver film. *Science* **337**, 450–453 (2012).

58. Kress, S. J. et al. Wedge waveguides and resonators for quantum plasmonics. *Nano Lett.* **15**, 6267–6275 (2015).

59. Kress, S. J. et al. A customizable class of colloidal-quantum-dot spasers and plasmonic amplifiers. *Sci. Adv.* **3**, e1700688 (2017).

60. Zhu, W. et al. Surface plasmon polariton laser based on a metallic trench fabry-perot resonator. *Sci. Adv.* **3**, e1700909 (2017).

61. Taniguchi, T. & Watanabe, K. Synthesis of high-purity boron nitride single crystals under high pressure by using Ba–BN solvent. *J. Cryst. Growth* **303**, 525–529 (2007).

62. Lee, G.-H. et al. Electron tunneling through atomically flat and ultrathin hexagonal boron nitride. *Appl. Phys. Lett.* **99**, 243114 (2011).

63. Britnell, L. et al. Electron tunneling through ultrathin boron nitride crystalline barriers. *Nano Lett.* **12**, 1707–1710 (2012).

64. Zomer, P., Guimarães, M., Brant, J., Tombros, N. & Van Wees, B. Fast pick up technique for high quality heterostructures of bilayer graphene and hexagonal boron nitride. *Appl. Phys. Lett.* **105**, 013101 (2014).

65. Levinshtein, M. E., Rumyantsev, S. L. & Shur, M. S. *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe* (Wiley, Hoboken, 2001).

66. Luryi, S. Quantum capacitance devices. *Appl. Phys. Lett.* **52**, 501–503 (1988).

67. Dröscher, S. et al. Quantum capacitance and density of states of graphene. *Appl. Phys. Lett.* **96**, 152104 (2010).

68. Ponomarenko, L. A. et al. Density of states and zero landau level probed through capacitance of graphene. *Phys. Rev. Lett.* **105**, 136801 (2010).

69. Jang, S. K., Youn, J., Song, Y. J. & Lee, S. Synthesis and characterization of hexagonal boron nitride as a gate dielectric. *Sci. Rep.* **6**, 30449 (2016).

Author contributions

M.P. and L.N. conceived the device concept. M.P. fabricated and characterized the devices. M.P. performed all analytical calculations and numerical simulations. Á.Sz. and M.L. performed ab initio device simulations. K.W. and T.T. synthesized the h-BN crystals. M.P. and L.N. discussed results and co-wrote the paper with contributions from Á.Sz. and M.L.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-08266-8.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: *Nature Communications* thanks Reuven Gordon, Jean-Jacques Greffet and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.