Wedge Waveguides and Resonators for Quantum Plasmonics

Stephan J. P. Kress,† Felipe V. Antolinez,† Patrizia Richner,‡ Sriharsha V. Jayanti,†,# David K. Kim,† Ferry Prins,† Andreas Riedinger,† Maximilian P. C. Fischer,† Stefan Meyer,† Kevin M. McPeak,† Dimos Poulikakos,‡ and David J. Norris*,†

†Optical Materials Engineering Laboratory, ETH Zurich, 8092 Zurich, Switzerland
‡Laboratory of Thermodynamics in Emerging Technologies, ETH Zurich, 8092 Zurich, Switzerland
#Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455, United States

*Corresponding author. E-mail: dnorris@ethz.ch

Supporting Information

1. Materials

Materials for nanocrystal synthesis. Cadmium oxide (CdO, 99.999%, 48-0800) was purchased from Strem Chemicals. n-dodecylphosphonic acid (DDPA, 98%, 08118) was purchased from Epsilon Chimie. 1-butanol (ACS Grade, 101990) was purchased from Merck KGaA. Diphenylphosphine (DPP, 98%, 252964), ethanol (96%, 02850), hexadecylamine (HDA, 90%, H7408), hexane (95%, 208752), methanol (99.9%, 34860), selenium shot (99.999%, 204307), sulfur (99.5%, 84683), trioctylphosphine (TOP, 97%, 718165), trioctylphosphine (TOP, 90%, technical, 117854), trioctylphosphine oxide (TOPO, 90%, 346187), 1-octadecene (ODE, 90%, O806), 1-octanethiol (98.5%, 471836), oleylamine (OAm, 70%, O7805), oleic acid (OLA, 90%, O7805).
Materials for waveguide and resonator fabrication. Sulfuric acid (95-98%, 258105) and tetradecane (>99%, 87140) were purchased from Sigma. Isopropyl alcohol (IPA) was purchased from Thommen-Furler AG. Hydrogen peroxide (H₂O₂, 30%, AnalaR Normapur) was purchased from VWR. Potassium hydroxide solution (KOH, 30%) was purchased from Technic France. Hydrofluoric acid (HF, 49%) was purchased from Fluka. Gold (Au, 99.999%) and silver (Ag, 99.99%) pellets were purchased from Kurt Lesker. Tungsten dimple boats (49 x 12 x 0.4 mm) were purchased from Umicore. Two-inch, single-side-polished, single-crystalline Si(100) wafers with <0.4 nm root-mean-square (RMS) roughness were purchased from Silicon Valley Microelectronics. 18.2 MΩ deionized water (DI) was obtained from a MilliQ Advantage A10 System. All chemicals and substrates were used as received without further purification, unless noted below.

2. Deposition of high-quality metallic films and wedge-tip radius determination.

Ag-film deposition. A Kurt J. Lesker Nano36 thermal evaporator equipped with a custom-built Meissner trap was used for the Ag-film deposition. Evaporation onto a native-oxide-covered Si(100) wafer was performed with 99.99% Ag pellets (Kurt J. Lesker) in a tungsten dimple boat at high-rates (>25 Å/s) and low residual gas pressures (3x10⁻⁸ Torr obtained via the liquid-N₂-cooled Meissner trap). This provided films with optical properties close to single-crystalline values. Further details regarding the setup and the deposited films may be found in McPeak et al.¹¹

Au-film deposition. The same apparatus was used to deposit Au films. The chamber pressure was ~5x10⁻⁷ Torr and the deposition rate was >1 Å/s. Marginally better optical properties can be obtained under the conditions discussed in McPeak et al.¹¹ However, Au was not the focus of the present work.
**Film characterization.** For optical characterization, we used a J. A. Woollam V-VASE variable-angle spectroscopic ellipsometer with rotating analyzer to extract the dielectric functions. Figure S1 shows the measured Ag dielectric functions.

**Wedge tip measurement.** For the tip radius determination, 50 nm of Ag was evaporated as described above (>25 Å/s, 3x10⁻⁸ Torr) onto the wedge template followed by evaporation of 500 nm Ge onto the wafer (>30 Å/s, 3x10⁻⁸ Torr). The Ag/Ge film was then stripped from the wedge template using ultraviolet-light-curable epoxy and a backing glass slide. Finally, cross sections were obtained by breaking the glass slide with the help of a tungsten carbide pen and then imaged using a scanning electron microscope (Figure S3). The thin, ductile Ag-layer and thick, brittle Ge-layer ensured that the tip did not deform in a ductile fashion while breaking and thus provided a reliable measure of the tip radius. Tip radii were highly reproducible at 20±1 nm for our process parameters.

### 3. Electromagnetic simulations of waveguides and resonators

**Simulations in two (2D) and three dimensions (3D).** Numerical finite-element simulations were performed with COMSOL Multiphysics (V4.4, RF module) to solve the inhomogeneous wave equations in the frequency domain. The waveguide was modeled as a lossy, Ag wedge surrounded by air with a wedge angle, \( \theta \), and translational invariance along \( x \). Solutions therefore take the form,

\[
E(x, y, z, t) = \bar{E}(y, z)e^{i\beta x}e^{-i\omega t},
\]

with \( \bar{E}(y, z) \) the modal electric field distribution at \( x, t = 0 \) and \( \beta \) the complex-valued propagation constant. In all simulations, the radius of curvature at the wedge apex, \( r \), was 20 nm and the metal assumed to be non-magnetic (\( \mu_r = 1 \)) with a frequency-dependent, complex-valued dielectric function, \( \varepsilon(\omega) \), experimentally determined by ellipsometry (Figure S1).

To identify the electric- and magnetic-field distributions and determine the propagation constant of the waveguided wedge-plasmon-polaritons (WPPs), we used the 2D-eigenmode solver. For these 2D simulations, we surrounded the simulation domain (width \( b_{\text{sim}} = 10 \lambda_{\text{vac}} \), where \( \lambda_{\text{vac}} \)
is the vacuum wavelength) by a perfectly matched layer (PML, width $b_{\text{PML}} = \lambda_{\text{vac}}$) (see Figure S2a). Next, we extracted the propagation length, $L_{\text{WPP}}$, and wavelength, $\lambda_{\text{WPP}}$, of the WPP mode from the propagation constant as $L_{\text{WPP}} = 1/(2\text{Im}\{\beta\})$ and $\lambda_{\text{WPP}} = 2\pi/\text{Re}\{\beta\}$.

In the 3D simulations (Figure S2b), the WPP mode was excited at one boundary ($x = 0$) and terminated with a PML at the other boundary ($x = l$). The 3D simulations used a smaller width of the simulation domain (width $b_{\text{sim}} = 4 \lambda_{\text{vac}}$). The structures were generally meshed with a maximum element size of a quarter of $\lambda_{\text{vac}}$ and a finer mesh (of ~5 nm element size) in the near-tip region.

**Figure-of-merit calculations for wedge waveguides.** The figure-of-merit (FOM) was defined as $\text{FOM} = \frac{L_{\text{WPP}}^2}{A_{\text{WPP}}}$, i.e. the WPP propagation length squared, $L_{\text{WPP}}^2$, divided by its effective modal area, $A_{\text{WPP}}$. The propagation length, $L_{\text{WPP}} = \frac{1}{2\text{Im}\{\beta\}}$, is calculated from the complex propagation constant of the WPP, as discussed above. The effective modal area, $A_{\text{WPP}} = \frac{\int |\varepsilon(r)|^2 dA}{\max [\varepsilon|E(r)|^2]}$, is inversely proportional to the maximum spontaneous emission enhancement for a dipolar emitter coupled to the waveguide. The propagation lengths and effective modal areas for three different vacuum wavelengths (550, 600, and 650 nm) are shown in Figure S4.

**Dispersion relation for wedge waveguides.** Figure S5 shows the dispersion relation for photons (in air), SPPs, and WPPs (for a wedge angle of 70.5°). The dispersion for SPPs, $k_x = \frac{\omega}{c} \sqrt{\varepsilon_m \varepsilon_d / \varepsilon_m + \varepsilon_d}$, was determined from ellipsometric data (Figure S1) assuming $\varepsilon_d = 1$ for air. WPPs were simulated for vacuum wavelengths from 400 (top right in Figure S5) to 800 nm (bottom left). For each wavelength, the wave vector was extracted from the real part of the simulated propagation constant $\beta$.

**Relative dipole-waveguide coupling.** A quantum emitter (modeled as a time-harmonic line current of length 2 nm) was placed at a height of 10 nm above the surface of the wedge waveguide (Figure
Using a simulation at $\lambda_{\text{vac}} = 600$ nm, the Poynting vector (along the waveguide) was integrated over a circular disk ($r = \lambda_{\text{vac}}/4$) centered on the apex at a distance of 15 $\mu$m from the dipolar emitter (dashed red circle in Figure S6a and the red circles in the insets of Figure S6b). When the dipole was laterally displaced away from the apex along the wedge face, the integrated power flux decreased significantly. The red and green curves in Figure S6b show the trend for a dipole oriented either vertically or in the surface-normal direction (see sketch in Figure S6a). For both orientations the dipole-waveguide coupling decreases rapidly as the dipole is moved away from the apex.

**Simulation of reflectivities of block reflectors on wedge waveguides.** Reflectivities at $\lambda_{\text{vac}} = 630$ nm were calculated from 3D simulations in which the mode (excited at $x = 0$ $\mu$m) propagates towards a rectangular block (see Figure S10a). Because the transverse magnetic field along the apex of the waveguide can also be modeled as a superposition of two counter-propagating waves, $H_y(x) = H_0 \left[ e^{i\beta x} + r_F e^{i\beta(2d-x)} \right]$, where $r_F$ is the reflection coefficient and $d$ is the effective reflector position (i.e. the actual position of the reflector plus a small correction for field penetration into the block). In this formula, we could exploit the propagation constant $\beta$ obtained from the 2D-mode analysis. This analytical expression could then be applied via $H_y(x)H^*_y(x)$ to fit (MATLAB, R2014a) the 3D-simulated transverse magnetic-field intensity (at 2 nm above the apex) to obtain $r_F$ (see Figure S10b). The reflection coefficient for a reflector height of 600 nm was determined to be 97% which translates into a reflectivity, $R = r_F^2$, of 94%.

**Estimation of the Purcell spontaneous emission enhancement for wedge waveguides and resonators in the frequency domain, i.e. for an infinitely narrow emitter.** Purcell emission-enhancement factors for the wedge waveguides and wedge resonators were determined from full-3D electromagnetic simulations of the emitted power in a dipole/vacuum, dipole/waveguide, and dipole/resonator structure. For the dipole/vacuum system, the dipole source (approximated as a 2-nm-long, single-frequency, time-harmonic, line current) was placed in vacuum and the total emitted...
power was integrated over the simulation domain boundaries. For the dipole/waveguide system, the dipole source was placed 10 nm above the apex of the wedge and the total emitted power was computed as the power flux across the simulation domain boundaries plus the power emitted into the waveguide. The same was performed for a wedge resonator with a cavity length of 7.5 µm and reflectors with a height of 600 nm and length of 1 µm. The 2-nm line current was placed at the center of the resonator between the reflectors where an electric field antinode is expected for every second longitudinal resonator mode. Simulations were performed on the cavity resonance, which in this particular case occurred at $\lambda_{\text{vac}} = 640$ nm, and the total emitted power was determined from the entire simulation domain. To estimate the Purcell factors, we normalized the emitted power from dipole/waveguide and the dipole/resonator to the power emitted by the dipole in vacuum. The values obtained were 16 and 106.

We stress that since these computations are done in the frequency domain and on resonance, the reported Purcell-enhancement factors are only valid for emitters that have a much narrower linewidth than the cavity resonance. This condition would be satisfied for our structures only for narrow atomic emitters or for solid-state emitters (such as quantum dots) at cryogenic temperatures. For the quantum dots used here, which are at room temperature, another treatment is necessary, as discussed in detail in Section 13 below.

**Calculation of properties of wedge waveguides and resonators at visible (630 nm, 800 nm) and telecom wavelengths (1550 nm) for Ag, Au, and Cu.** 2D-eigenmode analyses were performed for wedge waveguides with a wedge angle of 70.5°, tip radius of 20 nm, and dielectric properties obtained by ellipsometry. This was performed for Ag, Au, and Cu, each at three wavelengths (630, 800, and 1550 nm). The dielectric data for Ag at visible wavelengths are shown in Figure S1. All other dielectric data were taken from McPeak et al., where values for films deposited in our laboratory are summarized. From the eigenmode analyses, $\lambda_{\text{WPP}}, L_{\text{WPP}},$ and $A_{\text{WPP}}$ were extracted, as described above. Based on these values, $\text{FOM} = \frac{L_{\text{WPP}}^2}{A_{\text{WPP}}}$ was calculated. These along with $\lambda_{\text{vac}}$
and the dielectric functions (\(\varepsilon_1, \varepsilon_2\)) are listed in Table S1, which characterizes our wedge waveguides. Particularly at 1550 nm, we find that all waveguides have \(L_{WPP}\) above 300 \(\mu\text{m}\). For Ag, this value is almost 0.5 mm. Such long propagation lengths are expected while the mode remains strongly confined to a modal area more than an order of magnitude smaller than \(\lambda_{\text{vac}}^2\).

In order to calculate the performance limits of the wedge resonators, we started with the values for \(\lambda_{WPP}, L_{WPP},\) and \(A_{WPP}\). The maximum attainable quality factor, \(Q_{\text{max}}\), for a plasmonic wedge resonator is then given by,

\[
Q_{\text{max}} = \frac{2\pi L_{WPP}}{\lambda_{WPP}}
\]

Values for \(Q_{\text{max}}\) are listed in Table S2. For 630 and 800 nm, Ag provides \(Q_{\text{max}}\) of 311 and 588, respectively. At telecom wavelengths (1550 nm), values above 1000 are obtained for all metals; Ag is even beyond 2000.

Another important parameter for the wedge resonator is the smallest modal volume, \(V_{\text{min}}\). This can be calculated by integrating over the standing wave pattern of a resonator of size \(\frac{\lambda_{WPP}}{2}\). Alternatively, this integral may be simplified since the electric-field intensity profile follows that of a sine wave squared. Thus, we can multiply the modal area, \(A_{WPP}\), by \(\frac{\lambda_{WPP}}{2}\) and account for the fact that the integration of a sine wave squared gives us an additional factor of \(\frac{1}{2}\), leading to

\[
V_{\text{min}} = A_{WPP} \frac{\lambda_{WPP}}{4}.
\]

Dimensionless values for both \(A_{WPP}/\lambda_{\text{vac}}^2\) and \(V_{\text{min}}/\lambda_{\text{vac}}^3\) are listed in Table S2.

However, due to scattering at the imperfect reflectors, resonators cannot be expected in practice to achieve \(Q_{\text{max}}\) and \(V_{\text{min}}\) simultaneously. To estimate a more realistic value for the maximum attainable \(Q/V\) we determined \(Q_{\text{pract}}\), the actual quality factor that should be possible for a wedge resonator with \(V_{\text{min}}\). For \(Q_{\text{pract}}\), we assumed finite reflectivities of 95\%, as obtained in simulations for high block reflectors (heights >1 \(\mu\text{m}\)). We then calculated the classical Purcell expression for...
the achievable emission enhancement for an emitter that has a linewidth significantly narrower than the cavity resonance as:\cite{S4,S5}

$$\text{Purcell}_{\text{max}} = \frac{3}{4\pi^2 V_{\text{min}}} Q_{\text{pract}} \lambda_{\text{vac}}^3.$$  

We note that the linewidths of our fabricated resonators are sharp compared to many quantum emitters, particularly solid-state emitters such as nitrogen-vacancy centers or quantum dots at room temperature. The reported Purcell factors in Table S2 are only valid for emitters that are narrower than the cavity resonance, \textit{i.e.} for atomic-like emitters or for solid-state emitters (such as quantum dots) at cryogenic temperatures where the emitter linewidth is significantly reduced. For a more detailed discussion the reader is referred to Section 13 below.

Dimensionless values for $Q_{\text{pract}}/[V_{\text{min}}/\lambda_{\text{vac}}^3]$ and $\text{Purcell}_{\text{max}}$ are also listed in Table S2. $Q_{\text{pract}}/[V_{\text{min}}/\lambda_{\text{vac}}^3]$ reaches values above 10000 for all metals at 630 and 800 nm. The resonators can therefore lead to Purcell emission enhancements approaching 1000 at these wavelengths. This value might even be surpassed if novel reflector geometries (which can provide the Fresnel reflection coefficient of 99\%) can be developed. For reflectivities of 99\%, values of 3000 and above are expected for $\text{Purcell}_{\text{max}}$. In either case (\textit{i.e.} for reflectivities of 95\% or 99\%), high quality factors are predicted while the mode volume, $V_{\text{min}}$, remains deep sub-diffraction ($V_{\text{min}}/\lambda_{\text{vac}}^3 < 0.007$). This is a clear advantage over photonic microcavities, which are restricted by diffraction to much higher modal volumes.

4. Fabrication of wedge waveguides and resonators

\textbf{Fabrication of waveguide template.} Two-inch Si(001) wafers with a native-oxide layer were overcoated in a plasma-enhanced chemical vapor deposition (PECVD, Oxford Instruments, PECVD 80+) process with 100 nm SiN$_x$ that later serves as the mask for the anisotropic wet etch. After a pre-bake (hotplate, 120 °C, 60 s), the wafer was spin-coated (4000 r.p.m., 30 s) with an approximately 500-nm thick layer of AZ 1505 (AZ Electronic Materials). This was followed by a softbake (hotplate, contact, 100 °C, 60 s) and exposure with a MA6 mask aligner (Karl Susse,
vacuum contact, constant-power mode, 3 s) to generate line patterns (20 \( \mu \text{m} \) width, 300 \( \mu \text{m} \) length, and 40 \( \mu \text{m} \) periodicity) using a chrome-on-lime-glass photolithography mask (University of Minnesota, Nano Center). During exposure, alignment of the Si crystalline planes with the waveguide is essential so that the anisotropic etching yields crystallographic planes without step edges. This was achieved by optically aligning the primary standard wafer-flat Si(110) with the waveguide axis. Next the photoresist was developed in undiluted AZ 726 MiF (AZ Electronic Materials) for 60 s, rinsed in DI (60 s, puddle), and dried using a \( \text{N}_2 \) gun. Subsequently, the exposed Si\(_{\text{Nx}}\) lines were etched in a reactive-ion-etching process (RIE, Oxford Instrument, RIE 80 Plus) to reveal the underlying Si. After the RIE dry etch, the photoresist was removed in acetone using an ultrasonic bath, and the exposed lines of Si were etched in a KOH-IPA-DI (109 g KOH, 50 g IPA, and 341 g DI) solution under vigorous stirring (1000 r.p.m.) on a hotplate (60 °C) for 60 min. When alignment was good, this formed defect-free waveguide templates (i.e. with triangular trenches etched in the Si). In a last step, the remaining Si\(_{\text{Nx}}\) mask was removed [HF(49%):DI, 1:5, 60 s] and the wafer cleaned for 10 min in piranha solution (50 mL of 97% \( \text{H}_2\text{SO}_4 \) and 50 mL of 30% \( \text{H}_2\text{O}_2 \)). (Caution: this solution reacts violently with solvents and other organic material.) In a final step the wafer is rinsed for 30 min in DI, sonicated for 10 min in DI, and finally dried with a \( \text{N}_2 \) gun.

**Incorporation of bump lines and block reflectors into the Si waveguide templates.** For the bump lines and block reflectors, we further patterned the Si templates. The wet-etched Si trenches were milled using a dual-beam focused-ion beam (FIB, FEI Helios NanoLab 450S) at an acceleration voltage of 30 kV and an ion-beam current of 80 pA. For the block reflectors, rectangular depressions 1 \( \mu \text{m} \) in width and 5 \( \mu \text{m} \) in length (centered in the trench bottom) were milled into the Si. The depth was either 600, 900, or 1200 nm. These values represented the height of the block after template stripping.

**Ag deposition and template stripping.** After template patterning, an optically thick (~350 nm) Ag film was evaporated onto the Si template, as discussed in Section 2. The deposited Ag film was
subsequently template-stripped by bonding the Ag film to a soda-lime-glass microscope slide using an ultraviolet-light-curable epoxy (EpoTek), cured for 60 min under an ultraviolet lamp.\textsuperscript{56} The Ag-epoxy-glass stack was then stripped from the substrate to reveal waveguides and resonators with incorporated bump lines and reflectors. Au structures were made similarly.

**Fabrication of a wedge-waveguide control sample.** To demonstrate how critical the deposition of the Ag is for the wedge-waveguide performance, we repeated the above fabrication procedure with Ag deposited at a more commonly used rate (0.5 Å/s) and chamber pressure (>5×10\textsuperscript{−7} Torr). A comparison of the optical performance of non-optimized and optimized waveguides (both made via template stripping) is shown in Figure S8.

5. **Synthesis and characterization of colloidal semiconductor quantum dots (QDs)**

**Synthesis of core CdSe QDs.** CdSe cores with their lowest-energy absorption peak at 589 nm (approximately 4.1 nm diameter\textsuperscript{57}) were synthesized by modifying a published procedure.\textsuperscript{58} Briefly, 822 mg of CdO, 16.2 g of TOPO, 37 g of HDA, and 3.215 g of DDPA were combined in a 250-mL four-neck round-bottom flask and heated to 90 °C under N\textsubscript{2}. At 90 °C, the mixture was degassed three times to below 0.1 Torr (0.133 mbar) with stirring at 1000 r.p.m. The flask was returned to N\textsubscript{2}, heated to 320 °C until the solution turned clear and colorless, and held for 15 minutes. The temperature was reduced to 260 °C, the stir rate reduced to 300 r.p.m., and 8 mL of 1M Se in 97% TOP (1M TOP:Se) with 85 µL of DPP was swiftly injected. After injection, the stir rate was increased to 1000 r.p.m. and the temperature was maintained at 260 °C for 155 min to reach a size of 4.1 nm.\textsuperscript{57} The reaction was quenched by removing the heating mantle, submerging the flask in a water bath when the reaction mixture reached 200 °C, and adding 40 mL of 1-butanol at 130 °C to prevent solidification.

The mixture was evenly split between six 50-mL centrifuge tubes, and methanol was then added to each tube to yield a total volume of 50 mL and centrifuged at 4000 r.p.m. for 10 min. The clear supernatant was discarded. The precipitate in each tube was redispersed in 20 mL of hexane.
(120 mL total) and left undisturbed overnight. The following day, the mixture was centrifuged once more at 4000 r.p.m. for 20 min to precipitate unreacted material. The supernatant was saved, transferred to six fresh 50-mL centrifuge tubes, precipitated with ethanol, and centrifuged at 4000 r.p.m. for 10 min once more. The supernatant was discarded. The precipitate in each tube was redispersed in 4 mL of hexane (24 mL total) and precipitated again with ethanol and centrifugation at 4000 r.p.m. for 10 min. The supernatant was discarded. To yield a concentrated dispersion, the nanocrystals were redispersed in 20 mL of hexane and stored in the dark until further use.

**Cadmium-oleate stock preparation for core/shell synthesis.** Cadmium-oleate stock was prepared by combining 0.256 g of CdO, 2.6 mL of OLA, and 20 mL ODE in a three-neck 100-mL round-bottom flask and degassed three times at room temperature to below 0.1 Torr. The mixture was returned to N₂, raised to 270 °C until a clear, colorless solution was formed, and held for thirty minutes. The temperature was reduced to 150 °C and 1.3 mL of degassed OAm was added to prevent solidification at room temperature. The temperature was then reduced to 100 °C, degassed for 30 min, and then transferred into a N₂ glovebox for future use. Due to the addition of OAm, this stock does not set into a white solid, which is typical with cadmium oleate, but rather it remains an easily usable liquid.

**Zinc-oleate stock preparation for core/shell synthesis.** Zinc-oleate stock was prepared by combining 1.1 g of Zn(ac)₂, 3.8 mL of OLA, and 21 mL of ODE in a three-neck 100-mL round-bottom flask and degassed three times at room temperature to below 0.1 Torr while stirring at 800 r.p.m. The mixture was returned to N₂, raised to 200 °C until a clear, colorless solution was formed, and held for an additional 30 min. The temperature was reduced to 150 °C and 3.95 mL of degassed OAm was added to prevent solidification at room temperature. The temperature was then reduced to 100 °C, degassed for 30 min, and then transferred into a N₂ glovebox for future use. Due to the addition of OAm, this stock does not set into a white solid, which is typical with zinc oleate, but rather it remains an easily usable liquid.
Growth of CdS/ZnS shell/shell on CdSe cores to obtain red-emitting QDs. The CdS/ZnS shell/shell was grown on CdSe cores following a published procedure. Briefly, 100 nmol of CdSe cores in hexane, 3 mL of ODE, and 3 mL of OAm were added to a three-neck 100-mL round-bottom flask while stirring at 800 r.p.m. The mixture was degassed at room temperature for an hour, then at 120 °C for 20 min while stirring at 800 r.p.m. The mixture was switched to N₂, then heated to 305 °C at a rate of 16 °C/min. At 210 °C, two separate syringes of cadmium oleate (0.22 mmol) and octanethiol (a 1.2-fold excess, 0.264 mmol) precursors were each diluted in ODE to give a total volume of 3 mL and injected with a syringe pump at a rate of 1.5 mL/h for two hours. After precursor injection, the temperature was lowered to 200 °C, 1 mL of degassed OLA was added dropwise, and the mixture annealed for an hour at 200 °C.

After annealing, the temperature was lowered to 120 °C and the reaction mixture was degassed for 30 min under vacuum. After degassing, the flask was switched to N₂ and the temperature was raised to 280 °C at a rate of 16 °C/min. At 210 °C, two separate syringes of zinc oleate (0.24 mmol) and octanethiol (a 2-fold excess) precursors were each diluted in ODE to give a total volume of 3 mL and injected with a syringe pump at a rate of 2.5 mL/hour for 72 min. After precursor injection, the temperature was reduced to 100 °C, degassed for 10 min under vacuum, and then cooled to room temperature under N₂.

To precipitate and clean the QDs, an equivalent amount of ethanol to reaction mixture was added and the combination was centrifuged at 4000 r.p.m. for 10 min. The clear supernatant was discarded, the bright precipitate redispersed in 4 mL of hexane, and precipitated again with 10 mL of ethanol and 4000 r.p.m. for 10 min. Again, the colorless supernatant was discarded, the precipitate redispersed in 2 mL of hexane, and precipitated with 5 mL of ethanol and 4000 r.p.m. for 10 min. The clear supernatant was discarded and the particles dispersed in hexane and stored in the dark until future use. The recipe resulted in core/shell/shell QDs with their lowest-energy absorption peak at 614 nm and an emission maximum at 627 nm in hexane (see Figure S14). The emission peak shifted to 630 nm on the Ag waveguides.
Growth of CdSe/ZnS core/shell QDs with a composition gradient to obtain green-emitting QDs. The core/shell QDs with a composition gradient were prepared by modifying a published procedure. Briefly, 51.4 mg of CdO, 734 mg of Zn(acet)$_2$, and 2 mL of OLA were combined in a three-neck 100-mL round-bottom flask. The mixture was degassed three times at room temperature at pressures below 0.1 Torr, then at 150 °C for 30 min while stirring at 1000 r.p.m. The flask was switched to N$_2$ and 15 mL of degassed ODE was added. The temperature was raised to 310 °C to form a clear, colorless solution of cadmium oleate and zinc oleate. At 310 °C, 3 mL of TOP (90%) containing 0.4 mmol selenium and 3 mmol sulfur was swiftly injected into the reaction mixture. The temperature was maintained at 300 °C for 10 min, then lowered to 200 °C. At 200 °C, 3 mL of OLA was added dropwise and the mixture was annealed for an hour at 200 °C. After annealing, the mixture was cooled to room temperature with a water bath.

To precipitate and clean the particles, the reaction mixture was transferred to a 50-mL centrifuge tube, filled with ethanol to yield a total volume of 50 mL, and centrifuged at 4000 r.p.m. for 10 min. The clear supernatant was discarded. The bright precipitate was redispersed in 5 mL of hexane and 3 mL of ethanol was added. The mixture was centrifuged at 4000 r.p.m. for 10 min. The precipitate was discarded, not the supernatant. The colored supernatant was transferred to a new centrifuge tube, 2 mL of ethanol was added, and the mixture was centrifuged at 4000 r.p.m. for 10 min. The clear supernatant was discarded. The precipitate was dispersed in 4 mL of hexane and precipitated again with 4 mL of ethanol and 4000 r.p.m. for 10 min. The clear supernatant was discarded. The particles were dispersed in 4 mL of hexane and stored in the dark until future use. The recipe resulted in core/shell QDs with their lowest-energy absorption peak at 550 nm and an emission maximum at 564 nm (see Figure S14).

Absorption and fluorescence spectra of the QDs. Absorbance spectra of the QDs were obtained using a UV-visible spectrophotometer (UV-VIS, Varian Cary 50). Photoluminescence spectra were collected with a Spex Fluorolog-2 spectrofluorometer equipped with two double monochromators.
(0.22 m, SPEX 1680) and a 450 W xenon lamp as the excitation source. See Figure S14 for absorbance and photoluminescence spectra for the two types of QDs used in this study.

6. Placement of semiconductor QDs

Electrohydrodynamic printing setup. A description of the printing setup can be found in Galliker et al.\textsuperscript{S11} and Kress et al.\textsuperscript{S12}

QD-ink preparation. We transferred the QDs from hexane to tetradecane by selective evaporation prior to printing and adjusted the optical density (OD) at the longest-wavelength absorption peak to OD 0.5 for a 1-mm path-length cuvette.\textsuperscript{S12} This ensured that evaporative time scales were reduced before the QD dispersion was placed in the printing nozzle.

Printing onto wedge waveguides and resonators. We transferred the QD ink to a gold-coated glass capillary (1 µm opening diameter) that serves as the printing nozzle. Using a piezo-stage, the waveguides were brought within a few microns (5 µm) of the printing capillary, and a voltage pulse (200 V) was applied to eject attoliter, QD-containing droplets (~100-nm diameter). QD patterns were then generated by translating the substrate using the piezo stage. For the waveguide propagation-length and scattering measurements (Figure 3 in the main text), a QD-ink diluted 1:1 with tetradecane was used to print lines that are 2 or 5 µm in length (2x overprinting). The sparse QD deposits (~100 QDs) for the resonators (Figures 1f,g and 5 of the main text) were obtained with a QD ink diluted 1:10 with tetradecane and voltage pulses of only 20 ms duration. Additional details may be found in a preceding publication.\textsuperscript{S12}

Printing of individual quantum dots on wedge waveguides. For the placement of individual QDs on the wedge waveguides, a QD ink diluted 1:30 with tetradecane was prepared and the printing nozzle continuously translated across the wedge waveguides. Electrostatic autofocusing then guided the droplets containing individual quantum dots into the desired near-apex region.
7. Plasmonic resonator model

Photonic and plasmonic Fabry-Pérot resonator model. Both the dispersive and non-dispersive model for our plasmonic resonators are based on the classical Fabry-Pérot resonator,

\[ I(\nu) = \frac{I_0}{(1-r)^2 + 4r \sin \left(\frac{\pi \nu}{\nu_{FSR}}\right)}, \]

which describes the frequency-dependent mode intensity distribution, \( I(\nu) \), of a Fabry-Pérot resonator. \( I_0 \) is the intensity of the initial wave present at one boundary of the cavity.\(^{S13} \) The free spectral range, \( \nu_{\text{FSR}} = \frac{c}{2L} \), quantifies the mode spacing in frequency, \( \nu \), and is inversely proportional to the cavity size, \( L \), and proportional to the phase velocity, \( c \). The round-trip-loss factor, \( r = \sqrt{RR} \), is the factor by which the field magnitude decreases for a round trip in the resonator. For a typical photonic Fabry-Pérot resonator, it is only dependent on the reflectivity, \( R \).\(^{S13} \) However, for plasmonic resonators, additional losses (propagation and scattering losses) as well as dispersion [particularly for the more confined wedge-plasmon polaritons (WPPs)] must be considered.\(^{S14,S15,S16} \) Hence, we developed additional models for our wedge resonators that include plasmon-related effects. Specifically, this included both a non-dispersive and a dispersive model, as described below.

Non-dispersive plasmonic-resonator model. In the non-dispersive model, it is assumed that both the propagation length, \( L_{\text{WPP}} \), and phase velocity, \( c_{\text{WPP}} \), of the WPPs are free of dispersion (wavelength and frequency independent). This allows us to define a constant free spectral range, \( \nu_{\text{WPP}} = \frac{c_{\text{WPP}}}{2L} \), and enables us to lump the losses occurring during propagation (damping in the metal and scattering into photons) and reflection into a single loss factor for a plasmonic resonator of length, \( L \), with propagation length, \( L_{\text{WPP}} \), and reflectors of reflectivity, \( R \),

\[ r_{\text{WPP}} = R \cdot e^{-\frac{L}{L_{\text{WPP}}}}. \]

Based on this loss factor, \( r_{\text{WPP}} \), the treatment is very similar to that of the classical photonic Fabry-Pérot resonator,
\[ I_{\text{non-disp}}(\nu) = \frac{I_0}{(1-r_{\text{WPP}})^2 + 4r_{\text{WPP}} \sin \left( \frac{\nu}{\nu_{\text{WPP}}} \right)} \]

which allows us to easily calculate both the finesse,

\[ F = \frac{\pi \sqrt{r_{\text{WPP}}}}{1 - r_{\text{WPP}}} \]

and quality factor,

\[ Q = 2\pi \nu \tau_p, \]

for various cavity lengths, \( L \), and reflectivities, \( R \), (as done in Figure 4e,f of the main text). In the above formula, the photon lifetime,

\[ \tau_p = -\frac{L}{c_{\text{WPP}} \ln r_{\text{WPP}}}, \]

can be determined from the WPP phase velocity, \( c_{\text{WPP}} \). For the simulations of \( Q \) and \( F \), we used the experimentally determined propagation length (19 \( \mu \)m) and reflectivity (93%) for \( \lambda_{\text{vac}} = 630 \) nm (i.e. for the red-emitting QDs as the plasmonic source), obtained from scattering measurements of the waveguide and spectral measurements of the resonators, respectively.

**Dispersive plasmonic-resonator model.** For the dispersive plasmonic resonator, it was assumed that \( L_{\text{WPP}}(\nu) \) and \( c_{\text{WPP}}(\nu) \), are dispersive. Hence, the loss factor,

\[ r_{\text{WPP}}(\nu) = R \cdot e^{-\frac{L}{L_{\text{WPP}}(\nu)}}, \]

and free spectral range,

\[ \nu_{\text{WPP}}(\nu) = \frac{c_{\text{WPP}}(\nu)}{2L}, \]

are also dispersive. In order to determine how dispersive the WPPs are for our geometry, we performed a 2D-modal analysis (as described above in Section 3) and extracted \( L_{\text{WPP}}(\nu) \) and \( c_{\text{WPP}}(\nu) \). To compare with our experiments, the calculated \( L_{\text{WPP}}(\nu) \) were reduced by 36%. In Figure S15 we find that both the loss factor and phase velocity have a linear wavelength dependence over the emission spectrum of the red QDs (630 nm, 25 nm FWHM).
Once the dispersive character was quantified, we could build a dispersive plasmonic-resonator model,

\[ I_{\text{disp}}(\nu) = \frac{I_0}{(1 - r_{\text{WPP}}(\nu))^2 + 4r_{\text{WPP}} \sin\left(\frac{\pi \nu}{v_{\text{WPP}}(\nu)}\right)} \]

Because we found (Figure S15) that the round-trip-loss factor and phase velocity (and thus the free spectral range) of wedge waveguides are approximately linear in wavelength (and frequency) we assumed that the round-trip-loss factor and free spectral range are first order polynomials of the frequency. With this, we then computed (as previously for the non-dispersive model) the modes of a dispersive plasmonic wedge resonator using the experimentally-determined reflectivity (see Figure 4 in main text).

**Quantum-dot plasmonic resonators.** The expression for the intensity of the non-dispersive or the dispersive plasmonic resonator quantifies the mode density, \( I_{\text{res}} \). When a QD with an approximately Gaussian and normalized emission,

\[ E_{\text{QD}} = e^{-\frac{(\nu - \nu_{\text{peak}})^2}{2\sigma^2}} \]

is coupled to this cavity, the emission of the QD into the cavity should be proportional to \( I_{\text{res}} \). Hence, we expect a spectrum given by,

\[ I_{\text{QD-res}}(\nu) = I_{\text{res}} E_{\text{QD}} = I_{\text{res}} e^{-\frac{(\nu - \nu_{\text{peak}})^2}{2\sigma^2}} \]

where \( I_{\text{res}} \) is the expression either for the non-dispersive-, \( I_{\text{non-disp}} \), or dispersive-, \( I_{\text{disp}} \), resonator model. The expression for the QD-dispersive-plasmonic resonator with the experimentally determined WPP propagation length (19 \( \mu \)m), block reflectivity (93%), and emission of the QDs (630 nm peak, 25 nm FWHM) was used to compute the spectra in Figure 4 of the main text.

**Accounting for photons in the plasmonic wedge resonators.** Photons that do not interact with the cavity but are simply emitted by the QD and then scatter from the point of observation (i.e. at the block reflector) must be included for a quantitative fit of the data in Figure 5 of the main text.
Therefore, a contribution to the detected intensity due to these photons, \( I_{\text{phot}} \), was also included in our model,

\[
I_{\text{phot}} = \kappa E_{\text{QD}},
\]

where \( \kappa \) is a fitting parameter. The expression for the combined collected signal of non-modulated photons, \( I_{\text{phot}} \), and modulated plasmons, \( I_{\text{res}} \), was then,

\[
I_{\text{QD-res-phot}}(\nu) = (I_{\text{res}} + I_{\text{phot}}) = (I_{\text{res}} + \kappa)e^{-\frac{(\nu - \nu_{\text{peak}})^2}{2\sigma^2}}.
\]

**Experimental determination of block reflectivities from resonator spectra.** In order to experimentally determine the block reflectivities, we used the non-dispersive plasmonic resonator model because of the excellent fit to the data and fewer free parameters compared to the dispersive model (which makes extraction of values more unique and reliable). In order to account for photons not modulated but scattered by the cavity we added an additional term to the QD plasmonic-resonator model (see previous paragraph). For the fit, close bounds were set for all parameters that may either be determined experimentally (cavity length) or have physical limits (reflectivities). This allowed us to extract reliable values for all parameters in the expression for the resonator. For the calculation of the reflectivities, we started with the loss factor, \( r_{\text{WPP}} = R \cdot e^{-\frac{L}{L_{\text{WPP}}}} \), determined by the fit and used the nominal cavity length, \( L \), and experimental propagation length, \( L_{\text{WPP}} = 19 \mu\text{m} \), to extract block reflectivities of 93%.

8. **Intensity measurements**

**Setup.** For the intensity measurements (Figure 3 in the main text), we used an inverted microscope (Nikon, Ti-U) in epi-fluorescence mode equipped with a high-numerical-aperture (NA) air objective (Nikon, TU Plan Fluor 100x, 0.9 NA, no coverslip) and an air-cooled, linear, scientific complementary-metal-oxide-semiconductor camera (sCMOS, Flash 4.0, Hamamatsu, pixel size 6.5 \( \mu\text{m} \), quantum yield >70% for the spectral region studied) that was attached to the left optical port of the microscope. The sample was illuminated through the objective using a collimated light-emitting
diode (LED) emitting at 470 nm (470 nm CoolLED, AHF Analysentechnik), which was spectrally narrowed using a band-pass-filter (Brightline HC 435/40). The collected light was passed through both a 510-nm long-pass beam splitter (HC BS 510) and a 500-nm long-pass emission filter (Brightline HC 500 LP). Quantitative analysis of the images was possible because of the high quantum yield and linearity of the camera.

9. Scattering signal analysis

Optical analysis of scattering centers. Fluorescence counts across the scattering centers were extracted using ImageJ software over a length of 100 pixels (equivalent to 6.5 µm in the object plane) with a width of a single pixel in both the x- and y-direction. These spatial intensity cross sections strongly resemble (within 10%) that of a dipole emitting at 630 nm (approximated as spatial Gaussian with FWHM of 355 nm).

Propagation length determination. To determine the WPP propagation length, we used ImageJ and extracted the intensity of the wedge-plasmon-polariton signal along the waveguide. Since a diffraction-limited image should fall within approximately 6 pixels for our camera (6 x 65 nm = 390 nm), we used a line with width of 6 pixels to extract the intensity. Fitting an exponential decay to the values that are between 20 and 50 µm from the dipolar source (for these values an exponential decay is expected), we determined the propagation lengths for WPPs generated by the green- (564 nm, 15.4 µm) and red- (630 nm, 19 µm) emitting QDs.

10. QD-WPP coupling efficiency measurements

Sample. To determine the coupling efficiency of QDs to WPPs, two scattering centers of approximately 100 nm in height and 200 nm in width were placed at a separation of 20 µm on top of the wedge by FIB milling and subsequent template-stripping (see Section 4). Such scattering centers were chosen since they scatter a large proportion of the wedge plasmons into freely propagating photons. A separation of 20 µm was chosen such that higher-order effects could be neglected in the analysis. Subsequently, ~100 QDs were printed on the apex of the wedge in
between the two scattering centers. The false-color fluorescence micrograph of the structure generated is shown in Figure S9.

**Setup.** For the coupling measurements, fluorescence-intensity micrographs of the apex of the wedge were recorded using a Nikon Ti-U in epi-fluorescence mode with an electron-multiplying charge-coupled device camera (EM-CCD, Andor, iXon 888 Ultra, 16bit, 13 µm pixel size, quantum yield >93%) air-cooled to −75 °C. Images were recorded with an electron-multiplying gain of 25 and read out at 10 MHz.

**Analysis.** In the intensity maps (false-color image in Figure S9) one finds three spots corresponding to the two scattering centers (labeled Scatterer 1 and Scatterer 2) and the direct photon emission from the QDs. For the determination of the coupling efficiency, the intensities were extracted using ImageJ software. As all of the emission spots are nearly diffraction-limited and recorded with a 100x 0.9 NA air objective, a large fraction of their intensity resides within an area corresponding to approximately 6 x 6 pixels (6 x 0.13 µm = 0.78 µm) on the camera. Hence, all the intensities were extracted from this size region to obtain the highest signal to noise.

These values were then corrected for the background taken from a similarly sized region in the image. Assuming similar collection efficiencies for both the QD emission and the scattering centers due to the high-NA air objective and neglecting any higher-order effects (such as reflection from Scatterer 1, propagation to Scatterer 2, and subsequent scattering from Scatterer 2), we derived a relatively simple expression for the intensity of WPPs launched by the QDs,

\[ I_{wpp,0} = \frac{1}{\eta_{scatt}} I_{scatt,1} e^{\frac{l_1}{L_{wpp}}} + \frac{1}{\eta_{scatt}} I_{scatt,2} e^{\frac{l_2}{L_{wpp}}}. \]

This expression corrects for the propagation losses (WPP propagation length, \( L_{wpp} = 19 \) µm at a wavelength of 630 nm) and the finite scattering efficiency, \( \eta_{scatt} \), of the bumps. \( I_{scatt,1} \) and \( I_{scatt,2} \) are the intensities measured from the two scatterers, which are at distances \( l_1 \) and \( l_2 \) from the point of emission, i.e. the point where the QDs were printed. The scattering efficiency \( \eta_{scatt} \) is
conservatively estimated to be 0.50, but may vary between 0.25 and 0.75.\textsuperscript{17} When the obtained intensity for the WPPs launched, $I_{\text{wpp},0}$, is compared to the photons emitted directly from the QDs, $I_{\text{phot},0}$, we find that 2.5 times more WPPs are launched than photons into free-space for the conservative assumption of a 50% scattering efficiency, $\eta_{\text{scatt}}$. When we assume a lower scattering efficiency of 0.25 this ratio increases to a factor of 5, \textit{i.e.} five times as many wedge plasmons are launched than photons. In either case, these values are impressive in that they represent average values for a large number of randomly oriented QDs, all of which might not couple well to the wedge. Individual well-oriented QDs should couple more efficiently than the values stated above. When we assume that non-radiative losses can be reduced to an insignificant fraction by introduction of an additional spacer layer and proper alignment of the dipoles, we can obtain beta factors that reach values from 63\% to 83\% (depending on the assumed scattering efficiency of 0.75 or 0.25). These results agree with earlier predictions.\textsuperscript{18} In summary, we find that emission of QDs into the single-mode waveguide can be highly efficient. However, we stress that our values are only estimates, as a precise determination of the beta factor requires knowledge of changes in the difficult-to-measure non-radiative component.

11. **Spectral measurements**

**Setup.** For the spectral measurements (Figure 5 in the main text), we used a Nikon TE 200 inverted microscope. To excite our QDs, we used the 436-nm line from a 100 W Nikon high-pressure mercury-arc lamp mounted on the epi-illumination port. The emission from the printed QDs and leakage from the cavities was collected using a high-NA air objective (Nikon, TU, Plan Fluor, 100x, 0.9 NA) and sent to an imaging spectrometer (Horiba Jobin Yvon, Triax 320) equipped with a liquid-N\textsubscript{2}-cooled CCD camera (Princeton Instruments, Spec-10). A 600 lines/mm grating blazed at 500 nm was used. Integration across 5-20 rows of pixels yielded the spectrum.
12. Lifetime measurements

Setup. For lifetime measurements, we used an inverted microscope (Nikon, TE200) together with a 50x 0.8 NA air objective (TU Plan Fluor, 50x, 0.8 NA, no coverslip). A 405 nm picosecond pulsed laser (Picoquant, LDH-D-C 405) was used to illuminate the sample. Before entering the microscope, the laser was defocused by a convex lens to illuminate the sample with a large spot of about 50 µm diameter. The fluence was <1 nJ/cm², far below the onset of multi-exciton generation in these QDs. The emission of the sample was then directed to the active area (50 x 50 µm) of a single-photon-counting avalanche photodiode (Picoquant, MPD-SPAD) placed at the image plane outside the microscope. This setup allowed us to spatially resolve the spontaneous emission lifetime by moving the sample position using a piezo nanopositioning system (Physik Instrumente, P-615.3CD NanoCube). Measurements were taken at 5 MHz and 10 MHz repetition rates with 32 ps binning intervals using a time-correlated single-photon-counting (TCSPC) module (Picoquant, Picoharp 300). The width of the overall instrument response function was ~290 ps at FWHM. Both the TCSPC module and the piezo stage were controlled by a LabView program.

Analysis. Lifetime values were extracted by fitting a single-exponential decay model convoluted with the measured instrument response function (IRF) of our system to the measured data (MATLAB, R2014a). Only the data up to the decay by one order of magnitude from the initial value was considered by the fitting routine.

13. Purcell enhancement of QDs in plasmonic-wedge-waveguide cavities

Purcell enhancement for narrow emitters. The general Purcell formula,

$$\text{Purcell} = \frac{3}{4\pi^2} \frac{\lambda^3}{V_{cav}} Q_{cav},$$

assumes that the emitter is in resonance with a cavity mode of quality factor, $Q_{cav}$, and modal volume, $V_{cav}$. For the above equation to be valid, two conditions are required regarding the emitter. First, its emission wavelength, $\lambda_{em}$, must overlap with the cavity resonance, $\lambda_{em} = \lambda_{cav} = \lambda$. Second, its emission linewidth must be narrower than the cavity resonance, $\delta\lambda_{em} \ll \delta\lambda_{cav}$, so that
the entire emission spectrum is equally enhanced. In other words, the quality factor of the emitter, \( Q_{\text{em}} = \frac{\lambda_{\text{em}}}{\delta \lambda_{\text{em}}} \), must be much higher than the quality factor of the cavity, \( Q_{\text{cav}} = \frac{\lambda_{\text{cav}}}{\delta \lambda_{\text{cav}}} \), or

\[
\frac{\lambda_{\text{em}}}{\delta \lambda_{\text{em}}} = Q_{\text{em}} \gg Q_{\text{cav}} = \frac{\lambda_{\text{cav}}}{\delta \lambda_{\text{cav}}}.
\]

The above condition can be fulfilled if the emitter linewidth is extremely narrow (e.g. for atomic electronic transitions) or if the cavity has a low \( Q \). The latter is commonly assumed for plasmonic cavities since the resonances typically have quality factors of \( \sim 10 \). However, this condition is not satisfied in our wedge waveguides, which have resonances that are much sharper than our QD emitters: \( Q_{\text{em}} \approx 25 \) and \( Q_{\text{cav}} \approx 200 \).

**Purcell enhancement for broad emitters.** When the condition \( Q_{\text{em}} \gg Q_{\text{cav}} \) is not fulfilled, the assumptions of the classical Purcell formula are violated, implying that modifications are necessary. This case has been treated in the literature,\(^{19,20}\) where it is shown that when \( Q_{\text{em}} \ll Q_{\text{cav}} \), the Purcell enhancement is limited by \( Q_{\text{em}} \). Hence, one replaces \( Q_{\text{cav}} \) with \( Q_{\text{em}} \) and obtains the modified Purcell expression,

\[
P_{\text{broad emitter}} = \frac{3}{4\pi^2} \frac{\lambda^3}{V_{\text{cav}}} Q_{\text{em}}.
\]

**Purcell enhancement for QDs on our 6.5 \( \mu \)m wedge cavity.** In order to determine the spectrally-averaged Purcell enhancement for an emitter broader than the cavity resonance we can use (as described above),

\[
P_{\text{broad emitter}} = \frac{3}{4\pi^2} \frac{\lambda^3}{V_{\text{cav}}} Q_{\text{em}}.
\]

The quality factor of our QDs can be easily computed from the emission peak position and its linewidth \( Q_{\text{em}} = \frac{\lambda_{\text{em}}}{\delta \lambda_{\text{em}}} \approx 630 \text{ nm} / 25 \text{ nm} \approx 25 \) (see Figure S14 for spectra). Further we can determine the cavity volume for a 6.5 \( \mu \)m cavity from \( V_{\text{cav}} = \frac{1}{2} A_{\text{eff}} L \), where the prefactor of one half arises due to the integration of the sine-squared intensity distribution across the resonator (see
Section 3 above), and the modal area, $A_{\text{eff}}$, is given in Table S2. When we evaluate the above expression, we find that spectrally-averaged Purcell enhancement factor for our red-emitting QDs in our $L = 6.5 \, \mu\text{m}$ wedge-waveguide cavity is,

$$\text{Purcell}_{\text{QD}} = \frac{3}{4\pi^2} \frac{\lambda_{\text{em}}^2}{V_{\text{cav}}} Q_{\text{em}} = 22.8.$$ 

**Deduction of the Purcell enhancement factor from spontaneous emission lifetime measurements.** Experimental lifetime-reduction measurements can only be compared with computationally obtained Purcell factors if non-radiative rates are negligible (i.e. when the quantum yield approaches unity). Figure S11 shows the measured photoluminescence lifetime traces for quantum dots in different environments. The QDs dispersed in tetradecane show a lifetime of 16.7±0.2 ns, which is consistent with literature values for these types of QDs with near-unity quantum yield.\textsuperscript{S9} When the quantum dots are printed on glass, the lifetime reduces to 9.15±0.09 ns due to additional non-radiative pathways (e.g. exciton diffusion) for QDs in an ensemble.\textsuperscript{S21} From these numbers we conclude that the added non-radiative rate reduces the quantum yield from near unity in dispersion to 55% in the ensemble.

Placing the QDs in the resonator yields a lifetime of 736±38 ps when measured at the reflector. Thus, this lifetime arises from QDs that are coupled to the cavity. This lifetime reduction can be predominantly attributed to an increased radiative rate in the presence of a very high density of states in the near-field of the wedge resonator, while the ensemble non-radiative decay rates remain unchanged. This accelerated radiative rate brings the quantum yield of the coupled QDs back to a value close to unity. In the limit of near-unity quantum yield in dispersion and strongly enhanced radiative rates in the resonator, the experimental total lifetime reduction of 22.6 (dispersion over resonator) can be compared to the Purcell factor of 22.8 obtained from computations. In this case, we are neglecting any changes due to quenching of the QDs due to the proximity of the metal. This is unavoidable because we cannot measure this effect directly. Thus, our treatment can be
considered as a zeroth-order approach for comparing experiments with expectations from the Purcell effect.

14. Single-QD experiments

Sample preparation. Placement of individual QDs on the plasmonic wedges was achieved by further diluting the QD ink to a ratio of 1:30 in tetradecane. In the experiments we used narrower (2.5-5.0 µm) and shorter wedges (25 µm) to exploit the wedge ends as scattering centers for the WPPs launched by the individual QDs. For placement of the QDs, the nozzle was translated perpendicular to the apex of the wedge. Autofocusing then guided the droplets containing individual QDs to the desired locations on the apex.

Optical setup. For the observation of single QDs, the sample was illuminated with an intense LED light source (Lumencor, Sola SE II light engine) through a 100x high-NA air objective (Nikon, TU Plan Fluor, 100x, 0.9 NA) in epi-fluorescence mode. Fluorescence images were recorded with an air-cooled electron-multiplying CCD camera (EM-CCD, Andor, iXon3 888 Ultra, 16bit, 13 µm pixel size, quantum yield >93%) that was attached to an imaging spectrometer (Andor, Shamrock 303i) mounted on the left port of an inverted microscope (Nikon, Ti-U). For the blinking studies 1000 frames, each with 1 s exposure, were recorded using an electron-multiplying gain of 200 at a readout rate of 10 MHz and a sensor temperature of −75 °C.

Single-QD blinking analysis. The stacks of tiff images that were recorded were then further analysed with ImageJ software (see Figures S12 and S13 for fluorescence images and intensity time traces). In order to get the highest signal to noise, the counts were read out for an area 6 by 6 pixels corresponding to an area of 0.78 by 0.78 µm at the object plane (an area containing most of the emission from a diffraction-limited image for the objective used here). The extracted intensities for the wedge end, QD1, QD2, and QD3 were then corrected for the time-averaged background for an area of the same size (0.78 by 0.78 µm). This time series was subsequently smoothened with a moving average of 3 s (MATLAB, R2014b). In order to generate the false-color fluorescence
images with reduced noise for Figure 6 of the main text, frames were averaged over 20 individual frames (corresponding to a total integration time of 20 s) for the on and off state (Figure 6f,g of the main text) and for the entire 1000 frames recorded for the overview picture (Figure 6a of the main text). Results of this analysis with the blinking traces can be found in Figure S13.

15. Supporting references

S1. McPeak, K. M.; Jayanti, S. V.; Kress, S. J. P.; Meyer, S.; Iotti, S.; Rossinelli, A.; Norris, D. *ACS Photonics* 2015, 2, 326–333.
S2. Jackson, J. D., *Classical Electrodynamics*. 3rd ed.; Wiley: New York, 1998.
S3. Oulton, R. F.; Bartal, G.; Pile, D. F. P.; Zhang, X. *New J. Phys.* 2008, 10, 105018.
S4. Vahala, K. J. *Nature* 2003, 424, 839-846.
S5. Sauvan, C.; Hugonin, J. P.; Maksymov, I. S.; Lalanne, P. *Phys. Rev. Lett.* 2013, 110, 237401.
S6. Nagpal, P.; Lindquist, N. C.; Oh, S.-H.; Norris, D. J. *Science* 2009, 325, 594-597.
S7. Yu, W. W.; Qu, L. H.; Guo, W. Z.; Peng, X. G. *Chem. Mater.* 2003, 15, 2854-2860.
S8. Reiss, P.; Bleuse, J.; Pron, A. *Nano Lett.* 2002, 2, 781-784.
S9. Boldt, K.; Kirkwood, N.; Beane, G. A.; Mulvaney, P. *Chem. Mater.* 2013, 25, 4731-4738.
S10. Bae, W. K.; Char, K.; Hur, H.; Lee, S. *Chem. Mater.* 2008, 20, 531-539.
S11. Galliker, P.; Schneider, J.; Eghlidi, H.; Kress, S.; Sandoghdar, V.; Poulikakos, D. *Nat. Commun.* 2012, 3, 890.
S12. Kress, S. J. P.; Richner, P.; Jayanti, S. V.; Galliker, P.; Kim, D. K.; Poulikakos, D.; Norris, D. J. *Nano Lett.* 2014, 14, 5827-5833.
S13. Saleh, B. E. A.; Teich, M. C., *Fundamentals of Photonics*. 2nd ed.; Wiley: Hoboken, 2007.
S14. Dittbacher, H.; Hohenau, A.; Wagner, D.; Kreibig, U.; Rogers, M.; Hofer, F.; Aussenegg, F.; Krenn, J. *Phys. Rev. Lett.* 2005, 95, 257403.
S15. Sorger, V. J.; Oulton, R. F.; Yao, J.; Bartal, G.; Zhang, X. *Nano Lett.* 2009, 9, 3489-3493.
S16. de Leon, N. P.; Shields, B. J.; Yu, C. L.; Englund, D. E.; Akimov, A. V.; Lukin, M. D.; Park, H. *Phys. Rev. Lett.* 2012, 108, 226803.
S17. Brucoli, G.; Martin-Moreno, L. *Phys. Rev. B* 2011, 83, 075433.
S18. Martin-Cano, D.; Martin-Moreno, L.; Garcia-Vidal, F. J.; Moreno, E. *Nano Lett.* 2010, 10, 3129-3134.
S19. Coccioli, R.; Boroditsky, M.; Kim, K. W.; Rahmat-Samii, Y.; Yablonovitch, E. *IEE Proc.-Optoelectron.* 1998, 145, 391-397.
S20. Poitras, C. B.; Lipson, M.; Du, H.; Hahn, M. A.; Krauss, T. D. *Appl. Phys. Lett.* 2003, 82, 4032-4034.
S21. Kagan, C. R.; Murray, C. B.; Bawendi, M. G. *Phys. Rev. B* 1996, 54, 8633-8643.
S22. Moreno, E.; Rodrigo, S.; Bozhevolnyi, S.; Martin-Moreno, L.; Garcia-Vidal, F. *Phys. Rev. Lett.* 2008, 100, 023901.
### Table S1

Expected properties for wedge waveguides calculated for common plasmonic metals and wavelengths. Wedge-plasmon polariton wavelength, \( \lambda_{WPP} \), propagation length, \( L_{WPP} \), and modal area, \( A_{WPP} \), obtained from 2D-eigenmode simulations of wedge waveguides. The calculated figure of merit, FOM, is also shown. Simulations were performed for Ag, Au, and Cu, each at three wavelengths commonly used in plasmonics (630, 800, and 1550 nm). Dielectric data for simulations were taken from McPeak et al.\textsuperscript{S1} except for Ag at 630 and 800 nm, which are plotted in Figure S1. The table also lists the vacuum wavelength, \( \lambda_{vac} \), the real part of the dielectric constant, \( \varepsilon_1 \), and the imaginary part of dielectric constant, \( \varepsilon_2 \). Details on the computation may be found in the Section 3 of this document.

| Material | \( \lambda_{vac} \) (nm) | \( \varepsilon_1 \) | \( \varepsilon_2 \) | \( \lambda_{WPP} \) (nm) | \( L_{WPP} \) (\( \mu \)m) | \( A_{WPP} \) (\( \mu \)m\(^2\)) | FOM |
|----------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----|
| Ag       | 630             | -18.4           | 0.439           | 592             | 29.3            | 0.00644         | 133000 |
| Au       | -12.8           | 1.05            | 574             | 5.41            | 0.00461         | 6340            |
| Cu       | -12.6           | 0.781           | 573             | 6.95            | 0.00452         | 10700           |
| Ag       | 800             | -32.2           | 0.726           | 772             | 72.3            | 0.0129          | 404000           |
| Au       | -27.3           | 1.09            | 767             | 33.9            | 0.0114          | 101000           |
| Cu       | -26.4           | 1.08            | 766             | 31.9            | 0.0111          | 91800           |
| Ag       | 1550            | -134            | 3.62            | 1540            | 493             | 0.0724          | 3360000          |
| Au       | -127            | 5.37            | 1540            | 299             | 0.0699          | 1280000         |
| Cu       | -119            | 4.24            | 1540            | 332             | 0.0671          | 1650000         |
Table S2. Dimensionless parameters for wedge resonators calculated for common plasmonic metals and wavelengths. This table uses values from Table S1 to calculate the dimensionless properties of wedge resonators in terms of the vacuum wavelength, $\lambda_{\text{vac}}$. Derived from the simulated values are the maximum quality factor, $Q_{\text{max}}$, the minimum normalized modal volume, $V_{\text{min}}/\lambda_{\text{vac}}^3$, and the Purcell factor for the smallest cavity size, assuming a reflectivity of 95%. We list the dimensionless propagation length, $L_{\text{WPP}}/\lambda_{\text{WPP}}$, the dimensionless modal area, $A_{\text{WPP}}/\lambda_{\text{vac}}^2$, and an estimate for the highest quality-factor-over-volume attainable in practice, or $\frac{Q_{\text{pract}}}{V_{\text{min}}/\lambda_{\text{vac}}^3}$. This latter parameter was used in the Purcell formula to estimate the maximum obtainable Purcell enhancement, $\text{Purcell}_{\text{max}}$. More details may be found in Section 3 of this document.
17. Supporting figures

Figure S1. Measured optical data for Ag. Wavelength-dependence of (a) the real component ($\varepsilon_1$) and (b) the imaginary component ($\varepsilon_2$) of the dielectric function for Ag, obtained from ellipsometry. Low deposition-chamber pressures ($3\times10^{-8}$ Torr) and high deposition rates (>25 Å/s) enable optical properties comparable to single-crystalline Ag. The films also provide excellent geometrical fidelity that results in long plasmon propagation lengths on the waveguides (see also Table S1). These data were used for all electromagnetic simulations and calculations for Ag at visible or near-visible wavelengths.
Figure S2. Sketch of the wedge geometry for 2D and 3D electromagnetic simulations. (a) For the 2D-eigenmode analysis the waveguide was modeled as a Ag wedge in air with a wedge angle of $\theta$ and a radius of curvature of 20 nm. The size of the computational domain is ten times the vacuum wavelength, $b_{\text{sim}} = 10 \lambda_{\text{vac}}$. The domain is surrounded by a perfectly-matched-layer (PML) of size, $b_{\text{PML}} = \lambda_{\text{vac}}$. (b) The 3D geometry is the same as for the 2D-eigenmode analysis, except that the simulation domain was reduced to $b_{\text{sim}} = 4 \lambda_{\text{vac}}$, and the wedge of finite length, $l$, is excited at one end, $x = 0$, and terminated by a PML at the other, $x = l$. See Section 3 of this document for further details.
Figure S3. Wedge-tip radius determination. (a,b) Cross-sectional scanning electron micrographs of two different Ag wedges. To obtain an accurate value for the tip radius, the Ag film (50 nm) was backed with a thick Ge layer (500 nm) before template stripping. The brittleness and good adhesive properties of the Ge layer enabled us to obtain a clean break without ductile deformation. (c) Analyses of the scanning-electron-micrograph cross sections [here for the tip shown in (b)] gave highly reproducible tip radii of 20±1 nm. The scale bars are 50 nm.
Figure S4. Predicted propagation lengths and effective modal areas of wedge-plasmon polaritons. (a) Propagation length, $L_{WPP}$, and (b) effective modal area, $A_{WPP}$, for plasmons propagating along the apex of wedge waveguides of various wedge angles (30, 50, 70, 90, 110, 130, 150, and 170°) and vacuum wavelengths [550 (green), 600 (yellow), and 650 nm (red)]. As seen in earlier work, $L_{WPP}$ increases for blunter wedges due to reduced confinement. However, $A_{WPP}$ increases much more rapidly than $L_{WPP}$ for increasing wedge angles; the mode becomes unconfined above 130° (seen by the rapidly increasing modal area). The combined changes in $L_{WPP}$ and $A_{WPP}$ versus angle explain the maximum predicted for the figure-of-merit, $FOM = \frac{L_{WPP}^2}{A_{WPP}}$, in Figure 2b of the main text between 90 and 110°.
Figure S5. Dispersion relation for photons, SPPs, and WPPs. Dispersion plot (angular frequency versus wave vector) for photons (black solid line), surface-plasmon polaritons (SPPs, propagating at a flat Ag/air interface, black solid line with dots), and wedge-plasmon polaritons (WPP, propagating along the apex of a Ag wedge with a radius of curvature of 20 nm at the apex and a wedge angle of 70.5°, red line with dots). For the photon and SPP dispersions, analytical solutions were used. For the WPP dispersion a 2D-eigenmode analysis was performed. The dielectric functions in Figure S1 were assumed.
Figure S6. Simulated dipole-waveguide coupling. (a) Model for quantifying the sensitivity of emitter-waveguide coupling to the quantum-dot (QD) placement. The QD was approximated as a dipole emitting at a vacuum wavelength, $\lambda_{\text{vac}}$, of 600 nm. The dipole-to-waveguide coupling was calculated for different separations between the emitter and the apex. The QD was modeled as a 2-nm line current and placed 10 nm above the waveguide surface and translated away from the apex along the wedge face at a constant height, $h = 10$ nm. The dipole was translated from 0 to 600 nm from the apex, for two different dipole orientations: vertically (red) and perpendicular to the wedge surface (green). The red dashed circle shows the area used to quantify the power emitted into the waveguide. This disk is placed at a distance of 15 $\mu$m from the emitter. Its diameter ($\lambda_{\text{vac}}/2$) was chosen because it captures half of the power transmitted by the WPP mode. (b) For both orientations (red and green), the emitter-waveguide coupling decreases dramatically when the dipole is displaced by only a few tens of nanometers. The red disk used to determine the integrated power flux into the waveguide is shown in the inset.
Figure S7. Wedge-plasmon polaritons (WPPs) on Au. (a) Scanning-electron micrograph (SEM) of three Au wedges. The waveguides include bump lines (seen as bright lines). Red-emitting QDs were printed along the wedge apex (arrows) at different distances from the bump lines (scale bar, 10 µm). (b) Magnified SEM of one of the Au wedges in a (scale bar, 2 µm). (c–e) Fluorescence micrographs of the same structures in (a) (scale bar, 5 µm). Photoexcited QDs launch WPPs that scatter from the bumps. (f–h) Magnified images of the scattering from the bumps in (c–e) (scale bar, 500 nm). The WPPs decay with propagation distance. (i) The signals in (f–h) confirm a sub-diffraction waveguide mode that scatters only at the apex. (j) If the propagating mode also had a component along the face of the wedge, the scattered signal would broaden transverse to the apex. (k–m) Magnified images as in (f–h), but normalized (scale bar, 500 nm). (n) Schematic of the diffraction-limited mode propagating along the wedge apex (scale bar, 2 µm). (o–q) Spatial cross sections of the signals in (k–m) in the x- (blue) and y-directions (red), compared to the expected signal from an ideal point dipole emitting at 630 nm (black line), approximated as a Gaussian.
Figure S8. Fluorescence and scattering signals of wedge-plasmon polaritons (WPPs) propagating on non-optimized and optimized Ag wedges. False-color fluorescence micrographs of WPPs propagating away (horizontal streaks) from photoexcited green-emitting quantum dots (QDs, bright spot in middle) on (a) non-optimized and (b) optimized Ag wedge waveguides (see Section 4 of this document). (c,d) The intensity profile along the wedge (for distances between 20 and 50 µm from the QDs), extracted from the micrographs in (a,b), respectively. (e-h) Similar data for photoexcited red-emitting QDs. Non-optimized wedges clearly make determination of experimental parameters such as the WPP propagation length highly challenging, if not impossible. This demonstrates the importance of high-fidelity waveguides for wQED.
Figure S9. Estimating the QD-wedge-plasmon coupling efficiency. False-color fluorescence micrograph of approximately 100 QDs printed on the apex of a Ag wedge close to two extended scattering centers (bump lines that are 200 nm wide and 100 nm high). Dashed white lines denote the position of the wedge. The two scattering centers are separated by 20 µm (scale bar, 2 µm). In the fluorescence micrograph one finds three signals: two from the bump lines (labeled Scatterer 1 and Scatterer 2) that convert propagating wedge-plasmon polaritons (WPPs) into photons and one from direct photon emission from the ~100 QDs on the wedge. Considering the damping of the wedge plasmons (assuming a wedge-plasmon propagation length of 19 µm) and the similar magnitude of the scattered and direct photon signals, one can see directly from the image that the coupling of QDs to the wedge must be efficient. Indeed, a detailed analysis shows that WPPs are much more efficiently launched than photons, by a ratio up to 5:1 (see Section 10 of this document), which translates into a beta factor of 83% if non-radiative losses are neglected. We note that WPPs propagating on the apex are not visible in the image. Streaks, as in Figures 3j,k, are approximately three orders of magnitude weaker than the signals scattered from the bump lines. This highlights the fidelity of the wedges, which allows an intense beam of WPPs to propagate without significant unintended scattering. Also, the absence of any other scattering signals (beyond Scatterer 1 and Scatterer 2) demonstrates that the WPP mode is the predominant channel of plasmonic emission. No other plasmon polaritons are generated with significant efficiency.
Figure S10. Determination of reflectivities for block reflectors on wedge waveguides. (a) Contour plot of the logarithm of the transverse magnetic-field intensity at $\lambda_{\text{vac}} = 630$ nm along the apex of a wedge waveguide ($\theta = 70.5^\circ$). The WPP mode is excited at $x = 0$ $\mu$m and propagates towards the reflector of 600-nm height placed at $x = 7$ $\mu$m. (b) The transverse magnetic-field intensity 2 nm above the apex is plotted (red dots) and compared to an analytical expression (black line) described in Section 3 of this document. This yields a reflection coefficient of 97%, or a reflectivity of 94%. This is in excellent agreement with the experimentally determined reflectivity of 93% (Figure 5e of the main text).
Figure S11. Photoluminescence lifetime data for quantum dots (QDs) in different environments (in tetradecane, on glass, on silver, on the Ag wedge, and in the Ag resonator). Time-resolved photoluminescence of red-emitting QDs sitting in different environments. Colloidal QDs dispersed in tetradecane (blue) show a single-exponential-decay with a lifetime of 16.7 ± 0.2 ns. When the QDs are printed on borosilicate glass (green) and on flat template-stripped Ag (orange), the lifetime is reduced to 9.15 ± 0.09 ns and 4.45 ± 0.05 ns, respectively. This behavior is expected due to an increase in the non-radiative rate in a QD ensemble and the emergence of additional plasmonic modes (flat silver) in the near field of the emitter. The time-resolved photoluminescence measured directly from the QDs printed in the resonator (magenta) yields a lifetime of 2.72 ± 0.08 ns. The emission from the same cavity (6.5 µm in length) but measured via the WPPs scattered from the reflector (red) shows a much faster lifetime of 0.736 ± 0.038 ns. This difference suggests that a distribution of different lifetimes exists in the QD ensemble placed in the resonator. Only the emitters that are well coupled to the cavity are maximally enhanced. Such QDs contribute heavily to the scattering signal from the reflector. The instrument response function (IRF) of our system has a width of 290 ps at FWHM (black, dashed line). See Section 12 of this document for further details.
Figure S12. Photoluminescence intensity trace for a single quantum dot (QD) placed near the apex of a Ag plasmonic wedge. (a) Fluorescence-intensity time trace for a single QD in the near-apex region. Clearly resolved are the on and off periods, which are characteristic of an individual quantum emitter. (b-k) False-color fluorescence micrographs for individual frames during on and off periods at the times noted. The scale bars are 1 µm. These data demonstrate the deposition of individual emitters on the wedges.
Figure S13. Photoluminescence signals for three individual quantum dots (QDs) on a Ag wedge waveguide. False-color fluorescence micrograph (a) of three individual QDs (QD1, QD2, and QD3). In addition to direct emission of photons, the QDs launch wedge-plasmon polaritons (WPPs) along the wedge that (inefficiently) scatter out at the wedge end (scale bar, 1 µm). The time trace for the collected intensity at the wedge end (b) is correlated with the three individual QDs (c-e). The individual QDs (c-e) show on/off (blinking) behavior that is characteristic for individual quantum emitters. Even though the wedge end (b) is correlated with the time traces of all three QDs (c-e), it is particularly correlated with QD2 (d) as this emitter seems to couple well to the wedge (cross correlation >80% for QD2 with the wedge end), most likely due to beneficial orientation of the dipole of QD2. The blinking behavior and correlation may also be observed in the intensity histograms of the wedge end, QD1, QD2, and QD3 (f-i), respectively. Particularly the wedge end, QD2, and QD3 show bimodal intensity distributions that are characteristic for the on and off states of single QDs. These data indicate that individual QDs couple to WPPs and thus single-emitter quantum plasmonics is possible with this system.
Figure S14. Absorbance and fluorescence spectra for our colloidal semiconductor QDs. (a) Absorbance (darker line) and photoluminescence (brighter line, excitation at 350 nm) spectra for CdSe/ZnS core/shell QDs with a composition gradient in dispersion. The maximum emission is at 564 nm. (b) Similar spectra for CdSe/CdS/ZnS core/shell/shell QDs with a maximum emission at 627 nm. This value shifted to 630 nm when the QDs were placed on the Ag structures.
Figure S15. Plot of the dispersive properties of wedge waveguides. (a) Dispersive phase velocity and (b) round-trip loss factor simulated for a Ag wedge resonator with 10-μm cavity length, 70.5° wedge angle, and 20-nm tip radius of curvature. We find that both the phase velocity and the loss factor change approximately linearly over the entire range of emission for the red-emitting QDs (630 nm, FWHM 25 nm). Hence, we assume for the dispersive-resonator model that both the phase velocity and loss factor are linear with wavelength (which is also true for frequency over the range plotted).