# Supplementary Information for

# Lasing at the nanoscale: Coherent emission of surface plasmons by an electrically driven nanolaser

Dmitry Yu. Fedyanin,[1*] Alexey V. Krasavin,[2] Aleksey V. Arsenin[1] and Anatoly V. Zayats[2]

[1]Laboratory of Nanooptics and Plasmonics, Moscow Institute of Physics and Technology, 9 Institutsky Lane, Dolgoprudny 141700, Russian Federation

[2]Department of Physics, King's College London, Strand, London WC2R 2LS, United Kingdom

[*]e-mail: dmitry.fedyanin@phystech.edu

## 1. Amplification scheme

One of the main advantages of plasmonic nanostructures for optoelectronics is the possibility to use the SPP supporting metal-semiconductor interface for electrical injection. However, it is extremely difficult to realize this concept in practice because the best metals—gold, silver and copper, while other plasmonic materials (materials with a negative dielectric function) introduce high absorption increasing modal losses by orders of magnitude. The recently proposed so-called 'alternative' plasmonic materials [1] are good prospects for epsilon-near-zero and negative-refractive-index metamaterials, but in the infrared they are too lossy and cannot compete with gold and silver in deep-subwavelength plasmonic waveguides and cavities. For example, for titanium nitride, $\varepsilon_{TiN}$=–75 + 23i at $\lambda$ = 1.55 μm [2], while for gold, the imaginary part of the dielectric function, which responsible for losses, is at least twice lower: $\varepsilon_{Au}$=–115 + 11i [3]). In addition, the electrical conductivity of the alternative plasmonic materials is much lower than that of silver and gold, which prevents their use as efficient electrical contacts.

The use of metal-semiconductor Schottky contacts [4,5] is the most natural realization of electrical pumping in plasmonic devices. This approach provides population inversion in the immediate vicinity of the metal-semiconductor interface that is extremely important, since plasmonic modes are highly confined to the metal surface. But such a pumping scheme is not without its shortcomings. To the best of our knowledge, only one binary semiconductor material (InAs) forms Schottky contacts to gold and silver with a barrier height greater than the bandgap energy of the semiconductor [6,7], which is required for amplification [4]. This limits the operation wavelength to around 3 μm and operation temperature to below 150 K as consequences of the small bandgap energy and high Auger recombination in InAs, respectively. Inability of the Schottky barrier to block majority carriers under forward bias leads to a quite high leakage current even in a single heterostructure SPP amplifier [8] that

makes the amplification scheme power-inefficient at small intensities of the surface plasmon field.

Double heterostructures and multi-quantum wells could potentially solve this problem, but it is not easy to provide a high overlap of the plasmonic mode with the active region of the semiconductor structure [9,10]. The reason is that gold, silver and copper typically form rectifying Schottky contacts to direct-bandgap semiconductors, while low-resistance ohmic contacts are needed for creation of the population inversion. Common techniques of producing Au ohmic contacts to direct-bandgap III-V semiconductors are based on Au alloys (e.g. Au-Sn, Au-Be, Au-Zn, Au-Ge) or non-alloyed multilayer metallic structures such as Cr/Au, Ni/Au, Ti/Au and Ti/Pt/Au [11]. Despite the fact that the latter was successfully used in metal-coated nanopillar lasers [12,13], the above mentioned approaches cannot be utilized for loss compensation in nanoscale plasmonic waveguides, where the metal-semiconductor contact is used for SPP guiding and, therefore, the introduced high ohmic losses are highly undesirable. This problem is no less important in integrated on-chip photon and SPP sources, which are crucial for nanophotonic circuitry, where in-plane integrability is essential. While most of the other configurations suitable for the implementation of electrical pumping can emit light only perpendicular to the wafer [12-14] or demonstrate a dipolar behavior [15], emission parallel to the wafer is easily achievable in a waveguiding geometry [16,17].

In an active SPP waveguide, the signal propagates along the metal-semiconductor interface and a significant amount of the SPP field is concentrated in the metal, which primarily determines modal losses. At the same time, this interface should serve as a good ohmic contact for efficient electron or hole injection. Unfortunately, Au alloys, which form ohmic contacts to III-V semiconductors, can hardly compete with gold and silver in terms of optical losses in the infrared [18]. In addition, annealing after metal deposition drastically changes the contact structure via group-V elements outdiffusion, Au indiffusion and chemical reaction at the interface that seriously deteriorates the interface quality. In the second approach, apart from Ti, Ni, Cr being very absorptive materials, non-alloyed contacts require thick heavily doped cap semiconductor layers [9,10,19] to guarantee pure tunneling behaviour of the metal-semiconductor contact. These layers prevent achieving high confinement of the plasmonic mode to the active region of the semiconductor structure and introduce additional absorption losses that are especially critical, since the SPP propagation losses at the metal-semiconductor interface exceed $10^3$ $cm^{-1}$ at telecommunication wavelengths.

The proposed amplification scheme is based on an Au/InAsP/InGaAs/AlInAs double-heterostructure tunneling Schottky barrier diode. The lightly doped InGaAs layer sandwiched between $p^+$-AlInAs and $n^+$-InAsP layers with the bandgap energies exceeding that of InGaAs acts as an active region of the semiconductor structure. Under high forward bias, electrons and holes are injected into InGaAs from the InAsP and AlInAs sides, respectively, while high potential barriers for electrons at the InGaAs/AlInAs heterojunction and for holes at the InAsP/InGaAs heterojunction confine the excess carriers to the active region. Gold deposited on top of the electron injection $n^+$-InAsP layer is used to form an n-type ohmic contact. Despite the fact that the Au/n-InP Schottky barrier is as high as 0.5 eV [20], Fermi level in Au/InAs contacts is anomalously pinned to the conduction band of InAs [6,7] and a wide range of InAsP ternary alloys form contacts to gold with a quite small barrier height. In particular,

the height of the Au/n-$InAs_{0.4}P_{0.6}$ barrier does not exceed 0.2 eV [21,22] that combined with an extremely small effective mass of electrons allows these carriers to tunnel freely through the Schottky barrier even at moderate doping of $InAs_{0.4}P_{0.6}$ ($N_D^{InAsP}$=3×10$^{18}$ cm$^{-3}$) (see the next section for details). The voltage drop across the contact does not exceed 7 mV even at a current density of 50 kA/cm$^2$, which ensures nearly ideal ohmic characteristics.

## 2. Ohmic contact for double-heterostructure-based metal-semiconductor nanolasers

Plasmonic metals: silver and gold, – form rectifying Schottky contacts to most of direct-bandgap semiconductors [24], which creates a significant problem for optoelectronic devices. The only unique binary III-V semiconductor is InAs: the Fermi level at the Au/InAs interface is above the conduction band edge of InAs and n-type InAs forms an ohmic contact to Au [22,23]. However, the bandgap energy of InAs is equal to only 0.36 eV at room temperature, therefore, it cannot be considered as a cladding layer of the double-heterostructure optoelectronic device. On the other hand, the bandgap energy of InP is 1.34 eV, but n-type InP forms a Schottky contact to Au with a barrier height of about 0.5 eV. In this regard, $InAs_xP_{1-x}$ alloys are of a great interest, since it is possible to find a semiconductor material with a relatively large bandgap energy and a small barrier height, so that Au/$n^+$-$InAs_xP_{1-x}$ contact is ohmic (Figure S1) due to electron tunneling through the barrier. This gives the possibility to use the $In_yGa_{1-y}As$ alloy lattice matched to $InAs_xP_{1-x}$ as an active layer of the double heterostructure, while the $InAs_xP_{1-x}$ layer efficiently injects electrons and blocks holes because of the large valence band discontinuity between $InAs_xP_{1-x}$ and $In_yGa_{1-y}As$ (Figure S1).

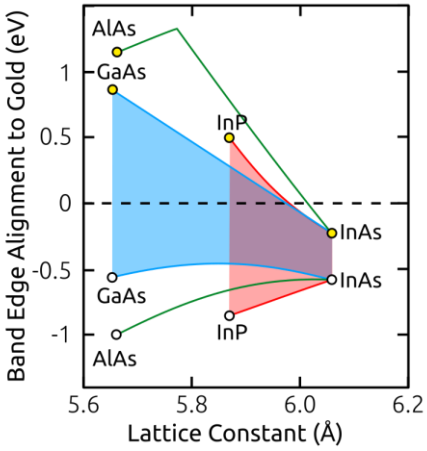



**Figure S1.** Conduction and valence band edge energies versus a lattice constant for InAsP, InGaAs and AlInAs ternary alloys [24]. Yellow and white circles denote the conduction and valence band edges of the binary semiconductors, respectively.

$InAs_{0.4}P_{0.6}$ is a good candidate for a cladding layer with an ohmic contact in double-heterostructure nanoscale optoelectronic devices. The Schottky barrier height $\varphi_B$ at the interface between gold and $InAs_{0.4}P_{0.6}$ is as low as 0.15 – 0.2 eV [24,25], therefore electrons can

tunnel through this barrier. To calculate the tunneling and over-barrier current, we evaluate the barrier transmission by solving the Schrödinger equation for a parabolic potential barrier and calculate the electron flux [26]:

$$J = q\rho_{\mathrm{mf}} \int_{-\infty}^{+\infty} dE\left[f(E-qV) - f(E)\right] \int_{0}^{\pi/2} \sin\theta \left|T_{\mathrm{m}\to\mathrm{s}}(\theta,E)\right|^2 \frac{\hbar k_{\mathrm{zs}}(\theta,E)}{m_{\mathrm{es}}} d\theta. \tag{S1}$$

In the above expression, $q$ is the electron charge, $V$ is the bias voltage, $m_{\mathrm{es}}$ is the effective electron mass in the semiconductor, $\rho_{\mathrm{mf}}$ is the density of states in the metal at the Fermi energy, $f(E)$ is the Fermi distribution function, $T_{\mathrm{m}\to\mathrm{s}}$ is the barrier transmission coefficient and

$$k_{\mathrm{zs}}(\theta,E) = \frac{1}{\hbar}\sqrt{2(E-qV)m_{\mathrm{es}} - \left(\upsilon_{\mathrm{f}} m_{\mathrm{em}} \sin\theta\right)^2}, \tag{S2}$$

where $\upsilon_{\mathrm{f}}$ and $m_{\mathrm{em}}$ are the Fermi velocity and effective electron mass in the metal, respectively. The barrier transmission coefficient can be expressed as follows [26]

$$\begin{aligned} T_{\mathrm{m}\to\mathrm{s}}(\theta,E) = \sqrt{\frac{8}{\pi}}\Bigg\{ &\left[U(-\kappa_{\mathrm{zs}}^2,\zeta_0) + \frac{i}{\kappa_{\mathrm{zm}}} U'(-\kappa_{\mathrm{zs}}^2,\zeta_0)\right] \times \left[i\kappa_{\mathrm{zs}} V(-\kappa_{\mathrm{zs}}^2,0) + V'(-\kappa_{\mathrm{zs}}^2,0)\right] \\ &- \left[V(-\kappa_{\mathrm{zs}}^2,\zeta_0) + \frac{i}{\kappa_{\mathrm{zm}}} V'(-\kappa_{\mathrm{zs}}^2,\zeta_0)\right] \times \left[i\kappa_{\mathrm{zs}} U(-\kappa_{\mathrm{zs}}^2,0) + U'(-\kappa_{\mathrm{zs}}^2,0)\right] \Bigg\}^{-1}. \end{aligned} \tag{S3}$$

Here, $U(x,y)$ and $V(x,y)$ are the parabolic cylinder functions,

$$\kappa_{\mathrm{zs}}(\theta,E) = \vartheta k_{\mathrm{zs}}(\theta,E), \tag{S4}$$

$$\kappa_{\mathrm{zm}}(\theta,E) = \frac{m_{\mathrm{es}}}{m_{\mathrm{em}}} \vartheta \frac{\upsilon_{\mathrm{f}} m_{\mathrm{em}} \sin\theta}{\hbar}, \tag{S5}$$

$$\vartheta = \sqrt[4]{\frac{\hbar^2 \varepsilon_{\mathrm{st}}}{16\pi q^2 m_{\mathrm{es}} (N_{\mathrm{D}} - N_{\mathrm{A}})}}, \tag{S6}$$

where $\varepsilon_{\mathrm{st}}$ is the static dielectric constant of the semiconductor, $N_{\mathrm{D}}$ and $N_{\mathrm{A}}$ are the donor and acceptor concentration.

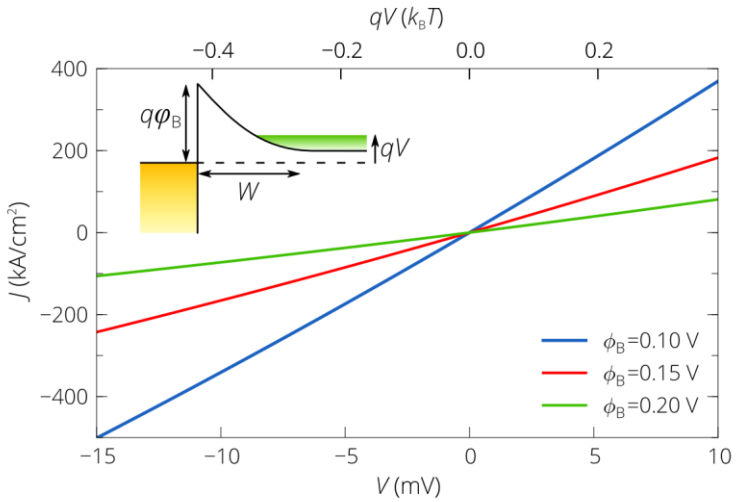



**Figure S2.** Current-voltage characteristics of the Au/$n^+$-$InAs_{0.4}P_{0.6}$ contact for different metal-semiconductor barrier heights $\varphi_B$, $T$ = 300 K, $N_D$–$N_A$ = 3×$10^{18}$ $cm^{-3}$, $\upsilon_f$ =1.4×$10^8$ cm/s, $m_{es}$ = 0.057$m_{e0}$ [27], $m_{em}$ = $m_{e0}$, where $m_{e0}$ is the free-electron mass. Insert: Schematic illustration of the Schottky barrier. At room temperature, the barrier width is equal to 13 nm.

Calculations show that the Au/$n^+$-$InAs_{0.4}P_{0.6}$ contact exhibits a perfect ohmic behavior even in the case of a significantly high barrier height ($\varphi_B$ = 0.15 – 0.2 eV) (Figure S2), the specific contact resistance does not exceed 1.3×$10^{-7}$ Ω $cm^2$ and does not change appreciably with the temperature. This provides ideal conditions for the design of coherent light and SPP sources operating at moderate and high current densities, since the voltage drop across the Au/$n^+$-$InAs_{0.4}P_{0.6}$ contact is less than 7 mV even at a current density of 50 kA/$cm^2$. It should be emphasized that such a small resistance is achieved with a moderate donor concentration of 3×$10^{18}$ $cm^{-3}$. In comparison, InP with the same donor concentration forms a rectifying contact to gold (Figure S3).

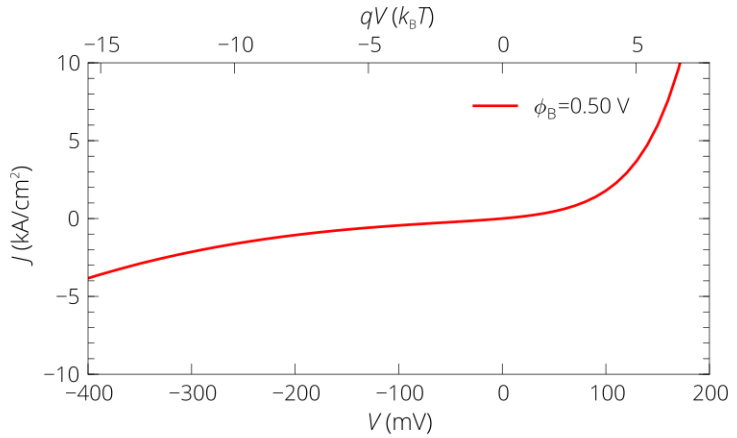



**Figure S3.** Current-voltage characteristics of the Au/$n^+$-InP at room temperature, $N_D$–$N_A$ = 3×$10^{18}$ $cm^{-3}$, $m_{es}$ = 0.079$m_{e0}$ [27].

## 3. Mode structure and spectral properties of the SPP nanolaser.

### *3.1. Calculation of the dielectric function of compound semiconductors*

Relative permittivity of the medium containing a mixture of molecules of $i^{th}$ type, can be linked to their polarizabilities $\alpha_i$ and densities $N_i$ through the Clausius-Mossotti relation [28]

$$\frac{\varepsilon-1}{\varepsilon+2}=\frac{1}{3}\sum_i N_i\alpha_i. \tag{S7}$$

Similar expressions are derived for each of the permittivities of the uniform individual components:

$$\frac{\varepsilon_i-1}{\varepsilon_i+2}=\frac{1}{3}N\alpha_i, \tag{S8}$$

where $N=\sum_i N_i$. Combining the above equations, one can express the permittivity of the compound thorough the permittivities of the components:

$$\frac{\varepsilon-1}{\varepsilon+2}=\frac{1}{3}\sum_i N_i\alpha_i=\sum_i\frac{N_i}{N}\cdot\frac{1}{3}N\alpha_i=\sum_i\eta_i\frac{\varepsilon_i-1}{\varepsilon_i+2}, \tag{S9}$$

where $\eta_i=N_i/N$ is the relative density fractions of the $i^{th}$ component. In the case of a binary compound $A_xB_{1-x}C$ equation (S9) is transformed into

$$\frac{\varepsilon-1}{\varepsilon+2}=x\frac{\varepsilon_{AC}-1}{\varepsilon_{AC}+2}+(x-1)\frac{\varepsilon_{BC}-1}{\varepsilon_{BC}+2}. \tag{S10}$$

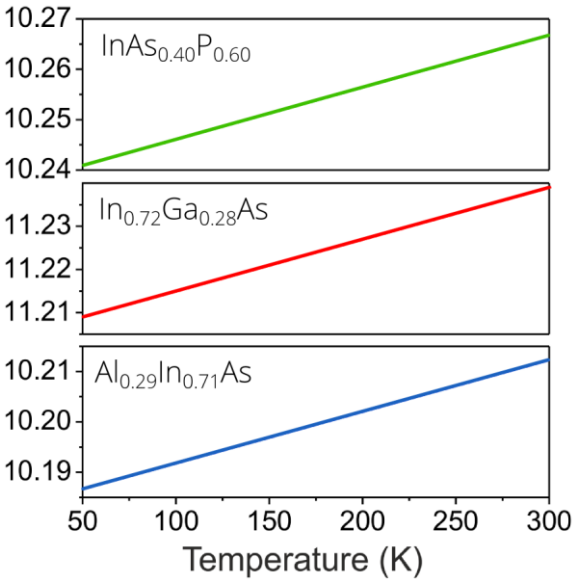
This equation is used to calculate the dielectric functions of all involved compound semiconductors, e.g. for $In_{0.72}Ga_{0.28}As$ $x = 0.72$, $\varepsilon_{AC}=\varepsilon_{InAs}$ and $\varepsilon_{BC}=\varepsilon_{GaAs}$. The dielectric constants for GaAs, InAs, InP as well as their spectral dispersion are derived from the experimental data reported in Ref. [29], while Ref. [30] is used for AlAs.

It should be noted that the dielectric function of the semiconductor changes noticeably as temperature decreases down to the cryogenic level. In order to take into account this phenomenon in the calculations, we employ an empirical model elaborated in Ref. [31], which gives the following expression for the temperature dependence of the relative permittivity:

$$\varepsilon(\omega,T)=1+\left(\frac{A}{E_g(T)+B(\omega,T)}\right)^2. \tag{S11}$$

Here, $E_g(T)$ is the bandgap energy of the semiconductor, $A$ = 13.6 eV and $B(\omega,T)= B_0 + B_1\omega + B_2T$, with $B_2$=2.5×10$^{-5}$ eV/K [31,32]. Introducing in the equation the bandgap energies of the semiconductors at the temperatures of interest (Table S1) and deriving the only unknown coefficients $B_0$ and $B_1$ by fitting the spectral dependence of the relative permittivity at room temperature (equation (S10)), the spectral dependences of the involved compound semiconductors are calculated at 77, 150 and 300 K.

**Table S1**. Values of the bandgap energies used in the calculations of $InAs_{0.40}P_{0.60}$, $In_{0.72}Ga_{0.28}As$ and $Al_{0.29}In_{0.71}As$ [27] and temperature dependences of the semiconductor permittivities.

| | Bandgap energy at different temperatures | | | Permittivity |
|---|---|---|---|---|
| | 77 K | 150 K | 300 K | |
| $InAs_{0.40}P_{0.60}$ | 0.987 eV | 0.971 eV | 0.924 eV |  |
| $In_{0.72}Ga_{0.28}As$ | 0.621 eV | 0.604 eV | 0.557 eV | |
| $Al_{0.29}In_{0.71}As$ | 1.086 eV | 1.067 eV | 1.000 eV | |



### *3.2. Calculation of the dielectric function of gold*

Dielectric function of gold is calculated in the infrared employing the Drude model $\varepsilon_{Au}(\omega) = 1 - \omega_p^2/[\omega^2 + i\Gamma(T)\omega]$, where $\omega_p = 1.37\times10^{16}$ s$^{-1}$ is the plasma frequency, which is practically temperature-independent, and $\Gamma(T) = \gamma_p(T) + \gamma_e(T,\omega) + \gamma_g(T) + \gamma_s(T)$ is the temperature-dependent damping frequency, which consists of several terms, representing electron-phonon $\gamma_p(T)$, electron-electron $\gamma_e(T,\omega)$, grain boundary $\gamma_g(T)$, and surface scattering $\gamma_s(T)$. The temperature dependence of these scattering rates was carefully

determined and the model was verified by the experimental data from Johnson and Christy [3] for the appropriate sample parameters used there in the measurements, for details see Ref. [33] and its supplementary information. This model is applied to calculate the dielectric constants of gold in the studied case (the film thickness is 100 nm, the grain size is equal to 60 nm) for 77 K, 150 K and 300 K temperatures.

### *3.3. Mode structure of the T-shaped plasmonic waveguide*

To investigate the mode structure of the T-shaped plasmonic waveguide and optimize its geometrical parameters with respect to the desired modal characteristics, 2D eigenmode finite element method (FEM) numerical simulations have been performed using COMSOL Multiphysics software. They reveal a family of modes of various orders having both TM and TE polarizations (Figure S4). The waveguide geometrical parameters defining the localization and propagation characteristics of the mode are optimized to give the fundamental $TM_{00}$ mode (the mode, which will be later used as the operational mode) an advantage over all the other modes. The height of the waveguide is adjusted to be 1 μm. This value is large enough to leave the properties of the $TM_{00}$ mode practically unaffected by the presence of the semiconductor substrate and prevent its radiation into the substrate, but small enough, so all other modes are extremely leaky with very shorter propagation lengths (Figure S4). The waveguide width *w* essentially determines the localization of modes in the horizontal direction and the mode effective index, both of them become especially important, when a ring resonator structure is implemented, since they eventually determine the out-of bend radiation losses. From this perspective, the decrease of the waveguide width down to 200 – 400 nm particularly deteriorates characteristics of the TE photonic modes, which become close to the cut-off. For the $TM_{00}$ mode, however, the width optimization requires careful balancing: the increase of the radiation losses for smaller widths is accompanied by the decrease of Ohmic ones, since a bigger portion of the mode becomes localized at the less absorptive $SiO_2$/Au interface. To take into account all the involved effects, the 2D eigenmode simulations and full 3D modelling (see Section 3.4) of the ring modes in different width waveguides have been performed, and the optimal value for the waveguide width $w$ = 300 nm was found.

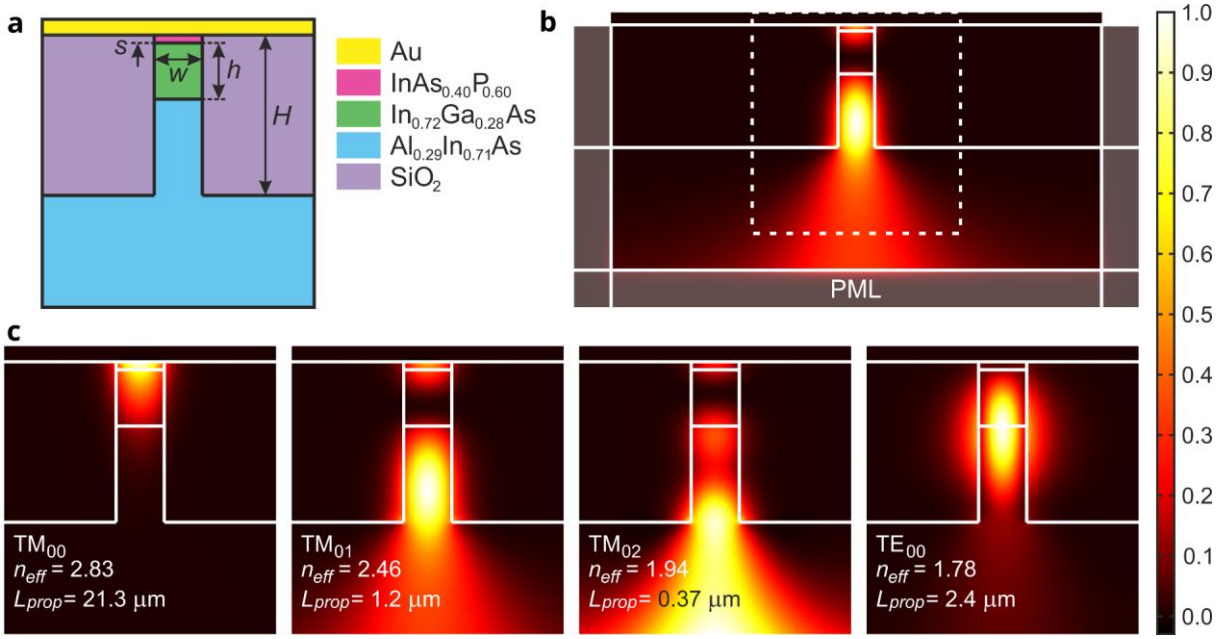



**Figure S4.** (a) Cross section of the T-shaped plasmonic waveguide with the geometrical parameters $H$ = 1 μm, $s$ = 50 nm, $h$ = 350 nm and $w$ = 300 nm. (b) FEM simulation geometry for finding eigenmodes of the waveguide: perfectly matched layers are placed around the waveguide to absorb the leakage radiation. The dashed line marks the area presented in c. (c) Distribution of the electric field intensity $|E|^2$ in a T-shaped plasmonic waveguide for $TM_{00}$, $TM_{01}$, $TM_{02}$ and $TE_{00}$ modes at a frequency of $\hbar\omega$ = 0.62 eV. $n_{\text{eff}}$ and $L_{\text{prop}}$ denote the effective mode index and propagation length, respectively.

### *3.4. 3D eigenmode simulations of a waveguide ring resonator*

The mode structure of a ring resonator based on the T-shaped plasmonic waveguide (Section 3.3) is analyzed using the three-dimensional eigenmode FEM simulations (Figure S5). The heights of the $InAs_{0.40}P_{0.60}$ spacer and $In_{0.72}Ga_{0.28}As$ active region are kept constant, while the waveguide width $w$ and the ring radius $R$ are varied to find the optimal operational optical mode and provide it with favorable resonance characteristics in comparison with other modes supported by the resonator. The ring structure is separated from the outer boundaries of the simulation domain by a distance large enough to ensure the absence of the overlap between them and the modes' near-fields. The radiation components were efficiently absorbed by the perfectly matched layer, partly shown in Figure S5(a) in green.

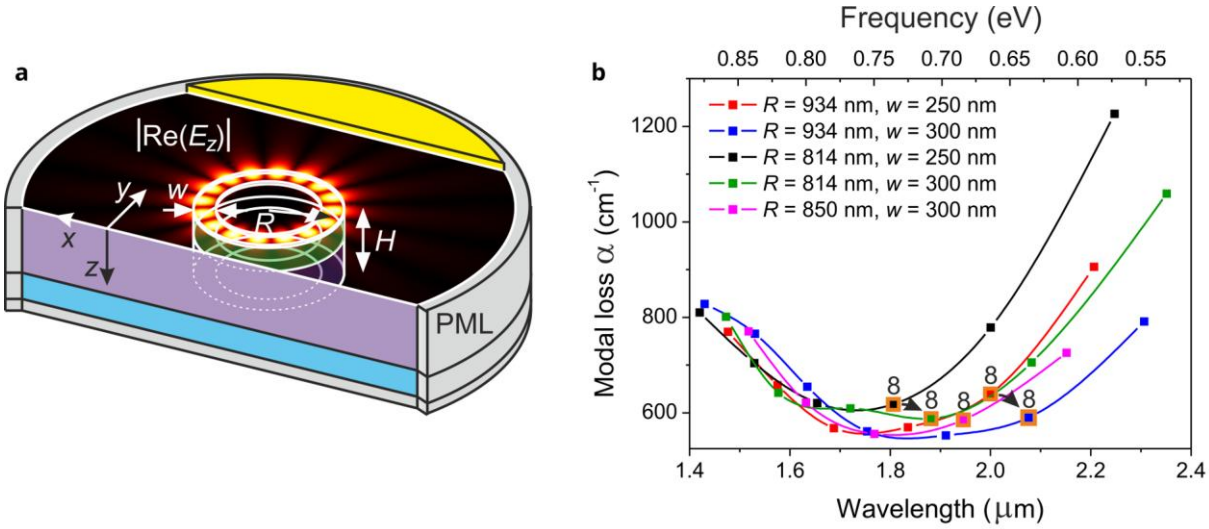



**Figure S5.** (a) Waveguide ring-resonator structure implemented in the 3D eigenmode numerical simulations. The field map shows the simulated distribution of the electric field perpendicular to the metal surface $|\mathrm{Re}(E_z)|$ at distance of 1 nm below it, for the $TM^8_{00}$ mode. The part of the PML domain (marked in grey color) is not shown to expose the structure. (b) The dependence of the modal losses of the $TM_{00}$ modes of various orders on their resonant wavelength for different ring resonators radii and waveguide widths. The points corresponding to the $TM^8_{00}$ mode are marked with orange squares.

In the ring geometry, straight-waveguide modes described in the previous section produce resonances (ring modes) of various orders, e.g. $TM^7_{00}$, $TM^8_{00}$, $TM^6_{01}$, $TE^5_{00}$, with specific frequencies (The top index in the mode notation corresponds to the number of mode periods along the ring and the bottom index indicates the order of the original mode in the T-shaped waveguide) The spectral range studied in the simulations extends from 0.5 eV ($\lambda \approx 2.4$ μm) to 0.9 eV ($\lambda \approx 1.4$ μm), covering the entire amplification bandwidth of $In_{0.72}Ga_{0.28}As$ (Section 4).

For each mode, we can find the quality factor as a ratio of the energy stored in the resonator to the energy loss per radian of oscillation

$$Q = \frac{\omega}{2\delta}, \tag{S12}$$

and mode energy loss per centimeter

$$\alpha = \frac{2\delta}{\upsilon_g}, \tag{S13}$$

which is more convenient to use in a steady-state regime. $Q$ and $\alpha$ are calculated using the complex frequency $\Omega = \omega + \delta \times i$ returned by the 3D eigenvalue solver and the group velocity determined from the mode dispersion, found in the designated 2D eigenmode simulations.

In the 3D simulations, we have observed that the ring geometry provides the same level of loss discrimination between $TM_{00}$ and $TM_{01}$, $TM_{02}$, and $TE_{00}$ modes as it was observed in 2D simulations (Section 3.3). In addition to these modes, the simulations reveal other

families of modes (e.g., $TM_{10}$ and $TE_{02}$ modes), but their quality factors are negligibly small due to high radiation losses. On the other hand, $TM_{00}$ modes of different orders show comparable quality factors (Figure 2(a) in the main article). For this reason, special attention should be paid to the choice of the operational mode for the proposed coherent SPP source on the basis of the modal loss and spectral position of the mode.

It can be seen that that the modal loss as a function of the mode order (resonance wavelength) shows a pronounced minimum (Figure S5(b)). This is the result of a trade-off between two main sources of loss. At high frequencies, the electromagnetic field of the $TM_{00}$ mode is highly confined to the Au/InAsP interface, which results in high ohmic loss. At low frequencies, the electromagnetic field is less localized, the effective mode index decreases, which eventually gives rise to radiation loss. To preserve the readability of the graph, only the points corresponding to $TM_{00}$ modes of the 8$^{th}$ order are marked, the orders of the other modes can be determined adding 1 for each step to the left and subtracting 1 for each step to the right on the same curve. The second observation is that the waveguide width of 300 nm is more advantageous than that of 250 nm due to lower modal loss (Figure S5(b)). Among the modes with the lowest modal loss in the spectral region 1.85 – 2.1 μm, the 8$^{th}$ order mode shows the best performance being the most close to the peak amplification wavelength (Section 4), while the modes of the 7$^{th}$ and 9$^{th}$ orders will be suppressed by low material gain at these wavelengths (Figure 2(b) in the main text). Therefore, the ring resonator design with a ring radius of 850 nm and a waveguide width of 300 nm is found to be optimal, in particular, because it ensures a subwavelength size of the device.

## 4. Spontaneous emission and material gain of the semiconductor

### *4.1. Interband transitions*

The material gain connected with interband transitions can be written as [34,35]

$$g_{\mathrm{mat}}^{\mathrm{c-h}}(n,p,\hbar\omega)=g_{\mathrm{mat}}^{\mathrm{c-h}}\left(n(F_{\mathrm{n}}),p(F_{\mathrm{p}}),\hbar\omega\right)=g_{\mathrm{mat}}^{\mathrm{c-h}}(F_{\mathrm{n}},F_{\mathrm{p}},\hbar\omega)=\frac{4\pi^2e^2}{c\bar{n}m_{\mathrm{e0}}^2\omega}\left|M_{\mathrm{ch}}\right|^2$$
$$\times\int_{-\infty}^{+\infty}\left[\left|M_{\mathrm{env}}^{\mathrm{c\text{-}hh}}(E,E-\hbar\omega)\right|^2\rho_{\mathrm{v}}^{\mathrm{hh}}(E_{\mathrm{v}}-E+\hbar\omega)+\left|M_{\mathrm{env}}^{\mathrm{c\text{-}lh}}(E,E-\hbar\omega)\right|^2\rho_{\mathrm{v}}^{\mathrm{lh}}(E_{\mathrm{v}}-E+\hbar\omega)\right] \quad \text{(S14)}$$
$$\times\rho_{\mathrm{c}}(E-E_{\mathrm{c}})\left[\frac{1}{1+\exp\left(\frac{E-F_{\mathrm{n}}}{k_{\mathrm{B}}T}\right)}-\frac{1}{1+\exp\left(\frac{F_{\mathrm{p}}-E+\hbar\omega}{k_{\mathrm{B}}T}\right)}\right]dE,$$

where $m_{\mathrm{e0}}$ is the free-electron mass, $F_{\mathrm{n}}$ and $F_{\mathrm{p}}$ are the quasi-Fermi levels for electrons and holes, $\bar{n}$ is the real part of the refractive index of the semiconductor, $M_{\mathrm{ch}}$ is the average matrix element connecting Bloch states near the band edges of the conduction and valence bands

($M_{ch} = m_{e0}E_p/12$, where $E_p$ is the Kane energy), $\rho_c$ is the density of states in the conduction band, $\rho_v^{hh}$ and $\rho_v^{lh}$ are the densities of states in the valence band for heavy and light holes, respectively, finally $M_{env}^{c\text{-}hh}$ and $M_{env}^{c\text{-}lh}$ are the envelope matrix elements for the transitions between the conduction and heavy-hole (HH) bands and the conduction and light-hole (LH) bands, respectively. In the case of the lightly doped or undoped semiconductor, the envelope matrix element and the material gain can be easily calculated using the *k*-selection rule and the parabolic band approximation (Figure S6).

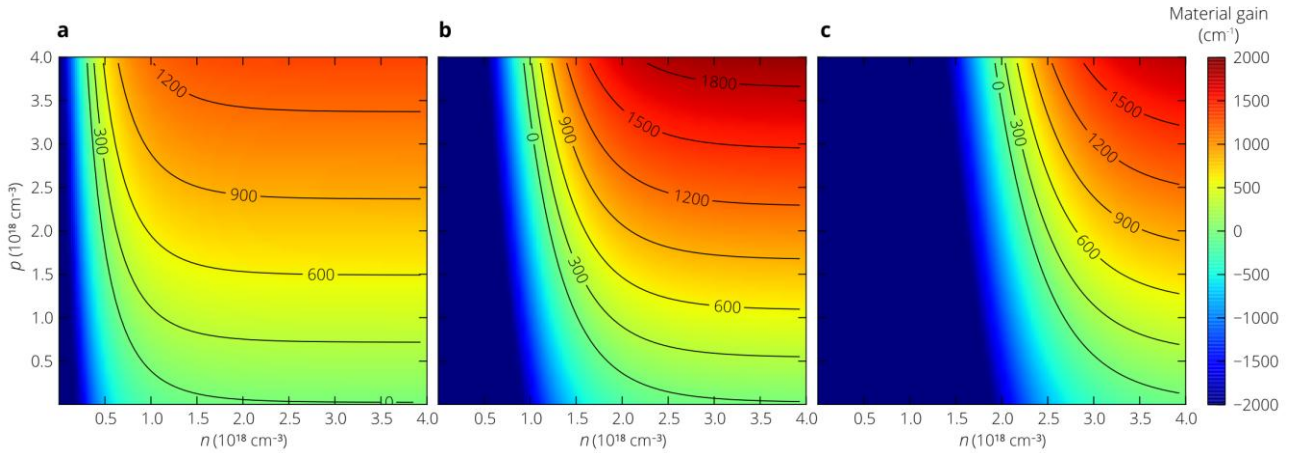



**Figure S6.** Dependence of the material gain due to interband transitions $g_{mat}^{c-h}$ on the electron and hole concentration in the undoped $In_{0.72}Ga_{0.28}As$ for different frequencies at room temperature: (a) $\hbar\omega$ = 0.576 eV, (b) $\hbar\omega$ = 0.637 eV, (c) $\hbar\omega$ = 0.701 eV. The following parameters are used in the calculations for $In_{0.72}Ga_{0.28}As$: $m_e = 0.036m_{e0}$, $m_{hh} = 0.413m_{e0}$, $m_{lh} = 0.042m_{e0}$ , $E_g$ = 0.557 eV [27].

The rate of spontaneous emission in the frequency range from $\hbar\omega$ to $\hbar\omega + d\hbar\omega$ per unit volume per unit time is given by

$$r_{spont}^{bulk}(n,p,\hbar\omega)d(\hbar\omega) = r_{spont}^{bulk}(F_n,F_p,\hbar\omega)d(\hbar\omega) = \frac{\bar{n}^2\omega^2}{\pi^2\hbar c^2}\frac{g_{mat}^{c\text{-}h}(F_n,F_p,\hbar\omega)}{1-\exp\left(\frac{\hbar\omega - F_n + F_p}{k_B T}\right)}d(\hbar\omega), \quad \text{(S15)}$$

and the total spontaneous emission recombination rate is equal to

$$U_{spont}^{bulk}(n,p) = \int_0^{+\infty} r_{spont}^{bulk}(n,p,\hbar\omega)d(\hbar\omega) = B_{bulk}(n,p)\times(np - n_{eq}p_{eq}), \quad \text{(S16)}$$

where $B_{bulk}(n,p)$ is the bimolecular recombination coefficient, $n_{eq}$ and $p_{eq}$ are the equilibrium electron and hole concentrations, respectively. In order to use expression (S16) for the recombination rate in the finite difference solver, $B_{bulk}(n,p)$ is fitted with the polynomial function of *n* and *p* (Figure S7).

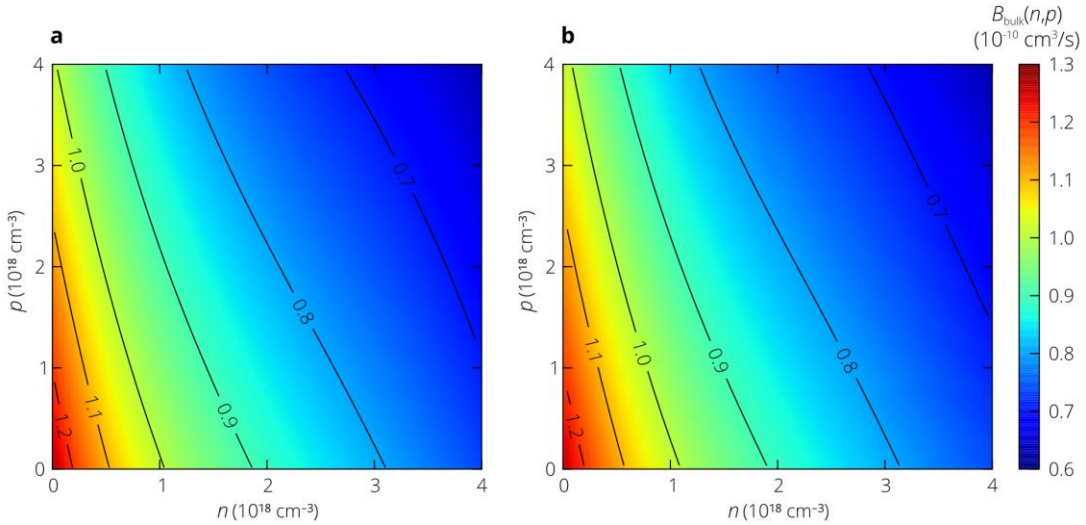



**Figure S7.** $B_{bulk}(n,p)$ dependence on the electron and hole concentration in the undoped $In_{0.72}Ga_{0.28}As$ at (a) room temperature and (b) the 4$^{th}$ order polynomial fit of this dependence.

### *4.2. Intervalence band absorption*

The split-off (SO) energy $\Delta_{so}$ in $In_{0.72}Ga_{0.28}As$ is equal to 0.38 eV and the intervalence band absorption (IVBA) can decrease material gain in the photon energy range 0.38 – 0.7 eV, where transitions between HH and SO bands play a notable role and LH-HH and LH-SO transitions are insignificant [36]. The IVBA coefficient $\alpha_{mat}^{hh-so}$ can be found similar to equation (S14). The main difference is that the average matrix element connecting Bloch states near the band edges of the valence and split-off bands is expressed as $M_{sh} = \hbar^2 k^2 E_p^2 / \left(E_g + \Delta_{so}\right)^2 72$ [37], where $\hbar k = \sqrt{2m_{hh}\left(E_v - E\right)}$ is the hole momentum. Figure S8 shows the calculated dependence of of $\alpha_{mat}^{hh-so}$ on the hole concentration.

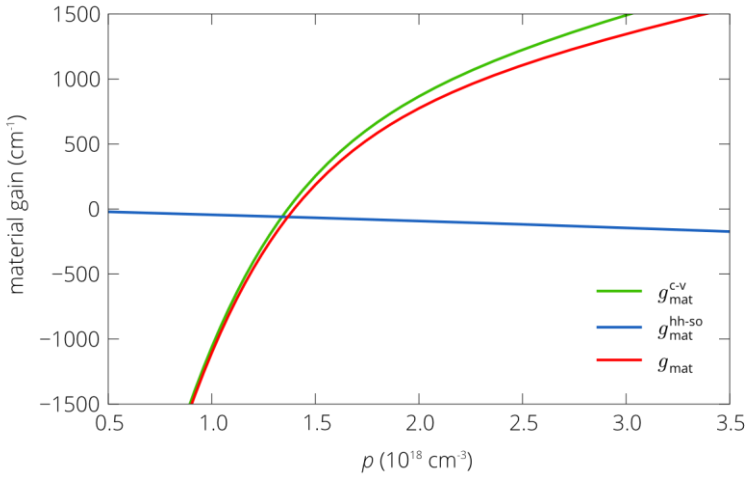



**Figure S8.** Material gain due to band-to-band and inter valence band transitions and net material gain as functions of the hole concentration (the electron concentration is equal to the hole

concentration) at a photon energy of $\hbar\omega$ = 0.637 eV. The split-off hole effective mass is equal to $0.15m_{e0}$ [27]. The photon energy is well above $\Delta_{so}$ and the IVBA does not significantly affect net material gain.

## 5. Spontaneous and stimulated emission in waveguides and resonators

### *5.1. Modal gain and spontaneous emission into waveguide modes*

Spontaneous emission rate $r_{\text{spont}}$ at the photon energy $\hbar\omega$ is connected with the optical gain $g$ at the same frequency through the Einstein coefficients *A* and *B* [34]. In a waveguide, the ratio of these two coefficients is equal to the density of modes per unit energy per unit waveguide length:

$$\frac{A(\hbar\omega)}{B(\hbar\omega)} = \rho_{\text{waveguide}}(\hbar\omega) = \frac{1}{\pi\hbar\upsilon_{\text{g}}(\hbar\omega)}, \tag{S17}$$

and the modal gain is expressed as

$$g_{\text{modal}}(\hbar\omega) = \frac{c}{8\pi\upsilon_{\text{g}}(\hbar\omega)} \frac{\iint_{\text{active}} \bar{n}(x,y,\hbar\omega) g_{\text{mat}}\left(F_{\text{n}}(x,z), F_{\text{p}}(x,z), \hbar\omega\right) \left|E(x,z,\hbar\omega)\right|^2 dxdz}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W(x,z,\hbar\omega)dxdz}. \tag{S18}$$

In the above expression, $\rho_{\text{waveguide}}(\hbar\omega)$ is the density of modes per unit energy per unit waveguide length, $E(x,z)$ and $W(x,z)$ are the distributions of the complex electric field and the energy density of the guided mode in the waveguide cross-section, respectively. The integral in the numerator of equation (S18) is taken over the active region of the waveguide. This leads us to the expression for the spontaneous emission into the guided mode

$$r_{\text{spont}}^{\text{modal}}(\hbar\omega) = \frac{c \iint_{\text{active}} \bar{n}(x,z,\hbar\omega) \dfrac{g_{\text{mat}}\left(F_{\text{n}}(x,z), F_{\text{p}}(x,z), \hbar\omega\right)}{1-\exp\left(\dfrac{\hbar\omega - F_{\text{n}}(x,z) + F_{\text{p}}(x,z)}{k_{\text{B}}T}\right)} \left|E(x,z,\hbar\omega)\right|^2 dxdz}{8\pi^2\hbar\upsilon_{\text{g}}(\hbar\omega)\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W(x,z,\hbar\omega)dxdz} \tag{S19}$$

that can be written in a more compact form:

$$r_{\text{spont}}^{\text{modal}}(\hbar\omega) = \frac{\iint_{\text{active}} r_{\text{spont}}^{\text{local}}\left(F_{\text{n}}(x,z), F_{\text{p}}(x,z), \hbar\omega\right) \left|E(x,z,\hbar\omega)\right|^2 dxdz}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W(x,z,\hbar\omega)dxdz}, \tag{S20}$$

where

$$r_{\text{spont}}^{\text{local}}\left(F_{\text{n}}(x,z),F_{\text{p}}(x,z),\hbar\omega\right)=\frac{c\bar{n}(x,z,\hbar\omega)}{8\pi^2\hbar\upsilon_{\text{g}}(\hbar\omega)}\frac{g_{\text{mat}}\left(F_{\text{n}}(x,z),F_{\text{p}}(x,z),\hbar\omega\right)}{1-\exp\left(\dfrac{\hbar\omega-F_{\text{n}}(x,z)+F_{\text{p}}(x,z)}{k_{\text{B}}T}\right)} \tag{S21}$$

can be calculated in the same way as $r_{\text{spont}}^{\text{bulk}}(n,p,\hbar\omega)$.

### *5.2. Spontaneous and stimulated emission in a resonator*

In the ring resonator, the density of modes can be written as

$$\rho_{\text{cavity}}(\hbar\omega)=2\sum_i\delta(\hbar\omega-\hbar\omega_i) \tag{S22}$$

where $\omega_i$ are the eigenfrequencies of the ring resonator and the factor of 2 comes from two directions (clockwise and anticlockwise). It gives the expression for the spontaneous emission into the cavity mode with the eigenfrequency $\omega_i$ per unit ring length per unit time

$$\begin{aligned}R_{\text{spont}}^{i}&=\int_{\hbar(\omega_i-\Delta\omega)}^{\hbar(\omega_i+\Delta\omega)}d(\hbar\omega)\delta(\hbar\omega-\hbar\omega_i)\frac{c}{4\pi L}\\&\times\frac{\iint\limits_{\text{active}}dxdz\,\bar{n}(x,z,\hbar\omega)\dfrac{g_{\text{mat}}\left(F_{\text{n}}(x,z),F_{\text{p}}(x,z),\hbar\omega\right)}{1-\exp\left(\dfrac{\hbar\omega-F_{\text{n}}(x,z)+F_{\text{p}}(x,z)}{k_{\text{B}}T}\right)}\left|E(x,z,\hbar\omega)\right|^2}{\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}dxdz\,W(x,z,\hbar\omega)}\\&=\frac{c}{4\pi L}\frac{\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\bar{n}(x,z,\hbar\omega_i)\dfrac{g_{\text{mat}}\left(F_{\text{n}}(x,z),F_{\text{p}}(x,z),\hbar\omega_i\right)}{1-\exp\left(\dfrac{\hbar\omega_i-F_{\text{n}}(x,z)+F_{\text{p}}(x,z)}{k_{\text{B}}T}\right)}\left|E(x,z,\hbar\omega_i)\right|^2dxdz}{\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}W(x,z,\hbar\omega_i)dxdz}\end{aligned} \tag{S23}$$

where $L$ is the length of the ring resonator ($L = 2\pi R$, where $R$ is the radius of the ring resonator). For convenience, we rewrite equation (S23) in the form of equation (S20)

$$R_{\text{spont}}^{i}=\frac{\iint\limits_{\text{active}}R_{\text{spont}}^{i\,\text{local}}\left(F_{\text{n}}(x,z),F_{\text{p}}(x,z),\hbar\omega_i\right)\left|E(x,z,\hbar\omega_i)\right|^2dxdz}{\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}W(x,z,\hbar\omega_i)dxdz}, \tag{S24}$$

where

$$R_{\text{spont}}^{i\,\text{local}}\left(F_{\text{n}}(x,z),F_{\text{p}}(x,z)\right)=\frac{c\bar{n}\left(x,z,\hbar\omega_i\right)}{4\pi L}\frac{g_{\text{mat}}\left(F_{\text{n}}(x,z),F_{\text{p}}(x,z),\hbar\omega_i\right)}{1-\exp\left(\dfrac{\hbar\omega_i-F_{\text{n}}(x,z)+F_{\text{p}}(x,z)}{k_{\text{B}}T}\right)}. \qquad \text{(S25)}$$

Figure S9 shows the dependence of $R_{\text{spont}}^{i\,\text{local}}$ on the electron and hole concentrations in the plasmonic ring resonator with a ring radius of 850 nm. It should be noted that the dependence of $R_{\text{spont}}^{i\,\text{local}}$ on *n* and *p* is remarkably different from that of the spontaneous emission in a bulk semiconductor (equation (S16)). This can be easily seen, if one rewrites equation (S25) in the form of equation (S16)

$$R_{\text{spont}}^{i\,\text{local}}\left(F_{\text{n}},F_{\text{p}}\right)=R_{\text{spont}}^{i\,\text{local}}[n(F_{\text{n}}),\,p(F_{\text{p}})]=B_i(n,p)\times\left(np-n_{\text{eq}}\,p_{\text{eq}}\right) \qquad \text{(S26)}$$

and plot $B_i(n,p)$ versus the electron and hole concentrations (Figure S10). The dependence of $B_i$ on the carrier concentrations gives a pronounced peak, since the photons or SPP quanta are emitted into the resonator mode at a frequency of about 0.637 eV in contrast to the case of a bulk material, where the radiation spectrum is very broad (Figure 2(b) in the main text).

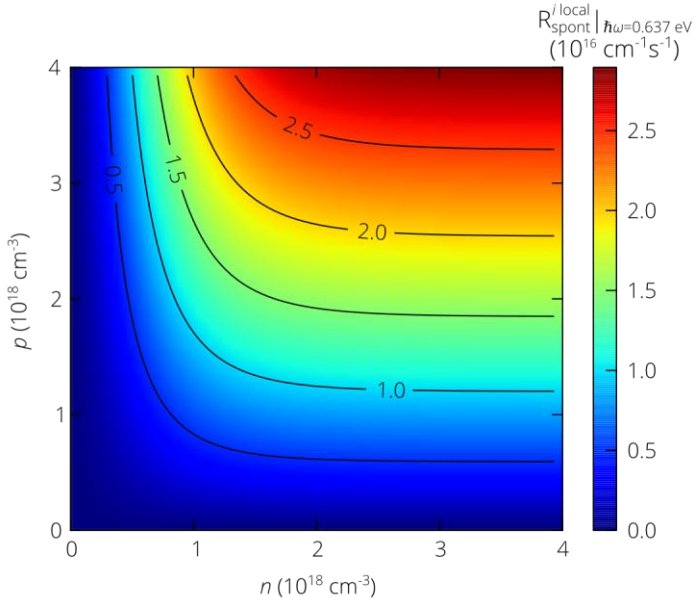



**Figure S9.** The dependence of $R_{\text{spont}}^{i\,\text{local}}$ on the electron and hole concentrations in $In_{0.72}Ga_{0.28}As$ calculated at a frequency of 0.637 eV for a ring radius of 850 nm.

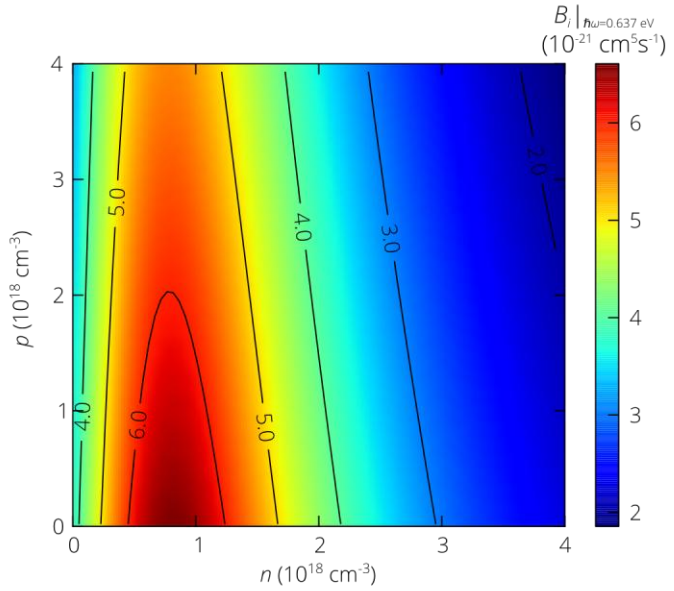



**Figure S10.** The dependence of $B_i(n,p)$ on the electron and hole concentrations in $In_{0.72}Ga_{0.28}As$ calculated at a frequency of 0.637 eV for a ring radius of 850 nm.

### *5.3. Enhanced spontaneous emission and Purcell effect*

Highly localized waveguide and resonator modes can enhance spontaneous emission affecting the carrier recombination dynamics, and therefore the overall process of amplification of plasmonic modes. At the same time, the decay rate of an excited state of a dipolar emitter is highly dependent of the electromagnetic environment around the emitter, the phenomenon which is known as the Purcell effect [38]. In our case, the environment is quite complex, including the semiconductor channel structure, composed of the layers of different refractive indexes, surrounded by low-refractive index dielectrics, with a semiconductor substrate on one side and a metallic film on the other. This produces an elaborated mode structure, creating new channels in which the emitter can radiate. Furthermore, the presence of the metal can noticeably affect the relaxation dynamics, introducing a direct quenching mechanism, the phenomenon specifically featuring the process of amplification of SPPs [39].

From equations (S15), (S20) and (S21), one can obtain the Purcell factor in a waveguide for the guided mode at a frequency $\hbar\omega$:

$$P_{j\,\mathrm{mode}}(x,z,\hbar\omega)=\frac{c^3}{8\omega^2\upsilon_{\mathrm{g}}^{j\,\mathrm{mode}}(\hbar\omega)\overline{n}(x,z,\hbar\omega)}\times\frac{\left|E_{j\,\mathrm{mode}}(x,z,\hbar\omega)\right|^2}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}W_{j\,\mathrm{mode}}(x,z,\hbar\omega)dxdz},\qquad\text{(S27)}$$

where $\upsilon_{\mathrm{g}}^{j\,\mathrm{mode}}$ is the group velocity of the $j^{\mathrm{th}}$ mode. Net Purcell factor at a frequency $\hbar\omega$ is equal to the sum of all the Purcell factors for guided and radiation modes:

$$P(x,z,\hbar\omega) = \sum_{j=1}^{M_{\text{guided}}} P_{j\,\text{mode}}(x,z,\hbar\omega) + P_{\text{radiation modes}}, \tag{S28}$$

where $M_{\text{guided}}$ is the number of guided modes at a frequency $\hbar\omega$ and $P_{\text{radiation modes}}$ is attributed to radiation modes. In a complex structure, $P_{\text{radiation modes}}$ cannot be easily estimated, and, instead of calculating $P_{\text{radiation modes}}$, we use 3D eigenmode FEM numerical simulations to evaluate $P(x,z,\hbar\omega)$.

The decay rate $\gamma(\mathbf{r})$ of a dipole $\boldsymbol{\mu}$ in an arbitrary electromagnetic environment is proportional to the power $S^{\text{str}}(\mathbf{r})$ emitted by it. The latter can be determined in a straightforward way if the electric field at the dipole position $\mathbf{E}_{\text{tot}}^{\text{str}}(\mathbf{r})$ is known, including the electric field of the dipole itself and all the electric field components resulting from interaction of the emitted field with the surrounding structure [40]:

$$S^{\text{str}}(\mathbf{r}) = \frac{\omega}{2}\,\text{Im}[\boldsymbol{\mu}\cdot\mathbf{E}_{\text{tot}}^{\text{str}}(\mathbf{r})]. \tag{S29}$$

Then, Purcell factor can be calculated normalizing $S^{\text{str}}(\mathbf{r})$ to the power emitted by the dipole placed in uniform medium $S^{\text{unif}}$ (vacuum or uniform dielectric, depending on the reference):

$$P(\mathbf{r}) = \frac{\gamma^{\text{str}}(\mathbf{r})}{\gamma^{\text{unif}}} = \frac{S^{\text{srt}}(\mathbf{r})}{S^{\text{unif}}} = \frac{\text{Im}[\boldsymbol{\mu}\cdot\mathbf{E}_{\text{tot}}^{\text{str}}(\mathbf{r})]}{\text{Im}[\boldsymbol{\mu}\cdot\mathbf{E}_{\text{tot}}^{\text{unif}}]}, \tag{S30}$$

from where we can obtain

$$\gamma^{\text{str}}(\mathbf{r}) = \frac{\text{Im}[\boldsymbol{\mu}\cdot\mathbf{E}_{\text{tot}}^{\text{str}}(\mathbf{r})]}{\text{Im}[\boldsymbol{\mu}\cdot\mathbf{E}_{\text{tot}}^{\text{unif}}]}\gamma^{\text{unif}} = P(\mathbf{r})\gamma^{\text{unif}}. \tag{S31}$$

The fields involved in equation (S31) can be directly obtained from the full 3D numerical simulations of the radiating dipole in the required electromagnetic environment. To prove the validity of this approach, it was tested on the physical system of a radiating dipole near a Si/Air interface. The obtained Purcell factor is found to be in excellent agreement with the exact analytical solution which can be derived in that case [40].

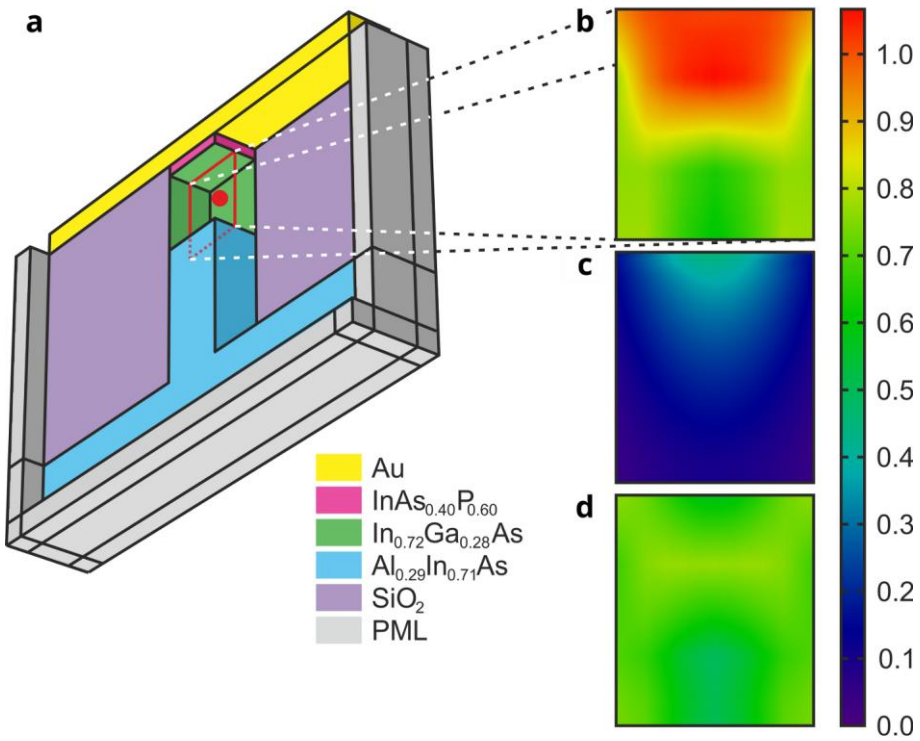



**Figure S11.** (a) Layout of the numerical simulation model of a dipole radiation inside the T-shaped plasmonic waveguide. The position of the dipole, marked by a red mark, was scanned across the active $In_{0.72}Ga_{0.28}As$ region, marked by a red rectangular. (b) The Purcell factor map in the active $In_{0.72}Ga_{0.28}As$ region of the T-shaped plasmonic waveguide derived from the numerical simulation results. (c) The map of the Purcel factor component originating from coupling to the $TM_{00}$ waveguide mode, calculated using equation (S27). (d) Difference between panel b and panel c.

As the next step, this approach is implemented to determine the Purcell factor in the T-shaped plasmonic waveguide. The 3D numerical simulations are performed for various dipole positions $(x_d, z_d)$ in the $In_{0.72}Ga_{0.28}As$ active region and monitoring the electric field $\mathbf{E}(x_d, z_d)$ (Figure S11(a)).

Then, employing equation (S30) the Purcell factor $P(\mathbf{r}) = \gamma^{str}(\mathbf{r}) / \gamma^{InGaAs}(\mathbf{r})$ map in the active region is created (Figure S11(b)). Generally, and especially at the upper part of the region, the magnitude of the Purcell factor is lower than that in the uniform $In_{0.72}Ga_{0.28}As$. This can be explained by the fact that the waveguide width is subwavelength and, therefore, the Purcell factor can be expected to be in the interval between the one in uniform $In_{0.72}Ga_{0.28}As$ $\left(P^{In_{0.72}Ga_{0.28}As} = 1\right)$ and the one in uniform $SiO_2$ ($P^{SiO_2} = n^{SiO_2} / n^{In_{0.72}Ga_{0.28}A} \times P^{In_{0.72}Ga_{0.28}As} \approx 0.4$). On the other hand, the most pronounced feature of the Purcell factor distribution is its correlation with the intensity of the main $TM_{00}$ mode, which is in agreement with what can be expected from equation (S27). From this formula, the Purcell factor component originating from coupling to $TM_{00}$ waveguide mode was calculated and its map in the active region was plotted (Figure S11(c)). After subtraction of this component from the all-inclusive numerical result (Figure S11(b)), practically uniform distribution of Purcell factor was obtained (Figure S11(d)).

## 6. Self-consistent numerical simulation of the SPP nanolaser

### *6.1. Electronic model*

To simulate the carrier behavior within the semiconductor, we use the two-component drift-diffusion model. Since the carrier flow in the structure is predominantly directed along the vertical *z* axis (Figure S12), one-dimensional electronic simulations are appropriate for the numerical simulation of the proposed amplification scheme and the proposed SPP nanolaer based on it [33,41]. In this case, the carrier transport in the steady-state regime is described by six first order non-linear differential equations

$$\begin{cases} \dfrac{d\varphi}{dz} = -E_z, \\ \dfrac{dE_z}{dz} = \dfrac{4\pi q}{\varepsilon}\left(p - n + N_{\mathrm{D}} - N_{\mathrm{A}}\right), \\ \dfrac{dn}{dz} = \dfrac{1}{qD_{\mathrm{n}}} J_{\mathrm{n}} - \dfrac{\mu_{\mathrm{n}} n}{D_{\mathrm{n}}} E_z, \\ \dfrac{dp}{dz} = -\dfrac{1}{qD_{\mathrm{p}}} J_{\mathrm{p}} + \dfrac{\mu_{\mathrm{p}} p}{D_{\mathrm{p}}} E_z, \\ \dfrac{dJ_{\mathrm{n}}}{dz} = qU = q\left(U_{\mathrm{stim}} + U_{\mathrm{spont}} + U_{\mathrm{Auger}}\right), \\ \dfrac{dJ_{\mathrm{p}}}{dz} = -qU = -q\left(U_{\mathrm{stim}} + U_{\mathrm{spont}} + U_{\mathrm{Auger}}\right), \end{cases} \tag{S32}$$

where all notations have their usual meaning [42], i.e. $\varphi$ is the electrostatic potential; $E_z$ is the static electric field; $q$ is the electron charge; $\varepsilon$ is the static dielectric constant; $p$ and $n$ are the concentrations of holes and electrons, respectively; $N_{\mathrm{D}}$ and $N_{\mathrm{A}}$ are the donor and acceptor impurity concentrations; $D_{\mathrm{p}}$ and $D_{\mathrm{n}}$, $\mu_{\mathrm{p}}$ and $\mu_{\mathrm{n}}$ are the diffusion coefficients and mobilities for holes and electrons, respectively; $J_{\mathrm{n}}$ and $J_{\mathrm{p}}$ are the electron and hole current densities; $U$ is the electron-hole recombination rate that includes the stimulated emission ($U_{\mathrm{stim}}$), spontaneous emission ($U_{\mathrm{spont}}$) and non-radiative Auger ($U_{\mathrm{Auger}}$) recombination rates. These differential equations must be completed by 18 interface boundary conditions at 4 interfaces: three – at the metal-semiconductor contact ($z = 0$), six – at the InAsP/InGaAs heterojunction ($z = s$), six – at the InGaAs/AlInAs heterojunction ($z = s+h$) and three – at the back contact ($z = H$).

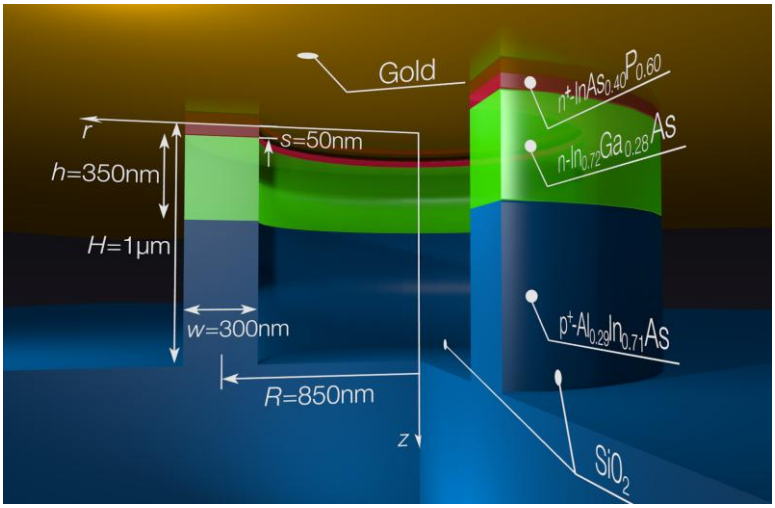



**Figure S12.** Structure of the electrically driven coherent SPP source. Subwavelength ring resonator of the electrically driven coherent SPP source is based on a T-shaped plasmonic waveguide.

The back contact is modeled as an ideal Ohmic contact:

$$\begin{cases} \varphi|_{z=H-0} = V, \\ n|_{z=H-0} = n_{\text{eq}}^{\text{AlInAs}}, \\ p|_{z=H-0} = p_{\text{eq}}^{\text{AlInAs}}, \end{cases} \tag{S33}$$

where $V$ is the bias voltage, $n_{\text{eq}}^{\text{AlInAs}}$ and $p_{\text{eq}}^{\text{AlInAs}}$ are the equilibrium electron and hole concentrations in bulk AlInAs.

Boundary conditions for the tunneling Schottky contact can be derived as follows. If the semiconductor is heavily doped, the Schottky barrier height and barrier width are small, carrier tunneling through the barrier dominates over other transport mechanisms and contact exhibits well pronounced ohmic properties as was shown in Section 2. The current density through the contact can be expressed as

$$J_{\text{contact}} = \frac{1}{\rho_{\text{contact}}} V_{\text{contact}}, \tag{S34}$$

where $V_{\text{contact}}$ is the voltage drop across the contact and $\rho_{\text{contact}}$ is the specific contact resistance, which in general is voltage dependent $\rho_{\text{contact}}(V_{\text{contact}})$. Accordingly, at a distance of the Schottky barrier thickness $W$, $\varphi|_{z=W} = \varphi|_{z=0} + V_{\text{contact}}$. This allows us to implement proper boundary conditions at the tunneling Schottky contact in the drift-diffusion model by placing the actual boundary for the simulation domain at the distance $W$ from the metal-semiconductor interface [43]. In the low resistance and heavy doping limit, the boundary conditions at the Au/n$^+$-InAsP interface can be written as

$$\begin{cases} \varphi|_{z=W} = J_{\text{n}} \times \rho_{\text{n}}^{\text{contact}}(J_{\text{n}}), \\ n|_{z=W} = n_{\text{eq}}^{\text{InAsP}}. \end{cases} \tag{S35}$$

Here $\rho_n^{contact}(J_n)$ is the specific contact resistance attributed to electron tunneling through the Schottky barrier at the Au/$n^+$-InAsP contact, which can be calculated by solving of the Schrödinger equation for a parabolic potential barrier (see Section 2) and $n_{eq}^{InAsP}$ is the equilibrium electron concentrations in bulk InAsP. It is easy to see that equation (S35) can approximated by the ideal-ohmic-contact boundary conditions

$$\begin{cases} \varphi|_{z=W} = 0, \\ n|_{z=W} = n_{eq}^{InAsP}, \end{cases} \tag{S36}$$

as the contact resistance tends to zero. For InAsP with a donor concentration of $3\times10^{18}$ cm$^{-3}$, the specific contact resistance is less than $1.5\times10^{-7}$ Ω cm$^2$ and $\varphi|_{z=W} < 0.015$ V at a current density of 100 kA/cm$^2$ (this value is well above the operating current of the device under consideration). Since the voltage drop across the barrier region does not exceed $k_BT$ at room temperature, ideal-ohmic-contact approximation can be used. At liquid nitrogen temperature, $q\varphi|_{z=W} > 4k_BT$ at $J$ = 100 kA/cm$^2$, and the boundary conditions in equation (S35) are preferred. In the case of the double-heterostructure tunneling Schottky barrier diode, when the threshold current is significantly reduced as temperature decreases, the ideal ohmic contact approximation gives the same result as the non-ideal contact model. In contrast to electrons, holes, being minority carriers in $n^+$-InAsP, do not experience tunneling and the boundary conditions for them should be given in accordance with the thermionic emission theory [42,44]:

$$J_p|_{z=+0} = -q\upsilon_{pr}(p|_{y=+0} - p_0), \tag{S37}$$

where $p_0 = N_v^{InAsP} F_{1/2}\left[\left(E_v|_{z=+0} - F_m\right)/k_BT\right]$ is the quasi-equilibrium hole concentrations at $z$ = 0 ($F_{1/2}$ is the Fermi-Dirac integral, $F_m$ is the Fermi level in gold, $E_v|_{z=0}$, is the valence band edge at $z$ = 0, $N_v^{InAsP}$ is the effective density of states in the valence band of InAsP), $\upsilon_{pr}$ is the effective recombination or collection velocity for holes at the metal-semiconductor interface. Since the boundary condition for holes must be given at $z = W$, equation (S37) is difficult to use in numerical simulations. If the concentration of holes (minority carriers in InAsP) is much less than $N_v^{InAsP}$, which is valid in the $n^+$-layer of the double heterostructure, we can write boundary condition (S37) in the equivalent form at $z = W$ [42,43]:

$$J_p|_{z=W} = -\frac{4\pi q m_p k_B^2}{(2\pi\hbar)^3} T^2 \exp\left(-\frac{E_g - \left(E_c|_{z=W} - F_n|_{z=W}\right) - V_{contact}}{k_BT}\right) \times\left\{\exp\left(\frac{F_m - F_p|_{z=W}}{k_BT}\right) - 1\right\}. \tag{S38}$$

Here, $F_n|_{z=W} = k_BT\,\mathrm{invF}_{1/2}\left(n|_{z=W}/N_c^{InAsP}\right)$ and $F_p|_{z=W} = k_BT\,\mathrm{invF}_{1/2}\left(p|_{z=W}/N_v^{InAsP}\right)$ are quasi-Fermi levels at $z = W$ (invF$_{1/2}$($x$) is the inverse Fermi-Dirac integral of order ½), $E_c|_{z=W}$ is the conduction band edge at $z = W$ and $F_m$ is the Fermi level in gold. Since in the $V_{contact}$ is smaller than $k_BT$, the boundary condition for holes can be simplified to

$$J_{\mathrm{p}}\big|_{z=W} = -\frac{4\pi q m_{\mathrm{p}} k_{\mathrm{B}}^2}{(2\pi\hbar)^3} T^2 \exp\left(-\frac{E_{\mathrm{g}} - E_{\mathrm{c}}\big|_{z=W} + F_{\mathrm{m}}}{k_{\mathrm{B}}T}\right) \times \left\{\exp\left(\frac{F_{\mathrm{m}} - F_{\mathrm{p}}\big|_{z=W}}{k_{\mathrm{B}}T}\right) - 1\right\}. \qquad \text{(S39)}$$

If the hole current density is smaller than the expression before the braces in equation (S38), the boundary condition can also be used in the form of $F_{\mathrm{p}}|_{z=W} = F_{\mathrm{m}}$,.

The heterojunctions are simulated using thermionic emission boundary conditions [45,46]. Thus, at the InAsP/InGaAs heterojunction, we obtain

$$\begin{cases} \varphi|_{z=s-0} = \varphi|_{z=s+0}, \\ \varepsilon^{\mathrm{InAsP}} E_z|_{z=s-0} = \varepsilon^{\mathrm{InGaAs}} E_z|_{z=s+0}, \\ J_{\mathrm{n}}|_{z=s-0} = J_{\mathrm{n}}|_{z=s+0}, \\ J_{\mathrm{n}}|_{z=s-0} = -q\upsilon_{\mathrm{nr}}^{\mathrm{InAsP}} n|_{z=s-0} + \dfrac{m_{\mathrm{n}}^{\mathrm{InAsP}}}{m_{\mathrm{n}}^{\mathrm{InGaAs}}} q\upsilon_{\mathrm{nr}}^{\mathrm{InGaAs}} N_{\mathrm{c}}^{\mathrm{InGaAs}} \mathrm{F}_{1/2}\left(\mathrm{invF}_{1/2}\left(\dfrac{n|_{z=s+0}}{N_{\mathrm{c}}^{\mathrm{InGaAs}}}\right) + \dfrac{\Delta E_{\mathrm{c}}^{\mathrm{InAsP-InGaAs}}}{k_{\mathrm{B}}T}\right), \\ J_{\mathrm{p}}|_{z=s-0} = J_{\mathrm{p}}|_{z=s+0}, \\ J_{\mathrm{p}}|_{z=s-0} = q\upsilon_{\mathrm{pr}}^{\mathrm{InAsP}} p|_{z=s-0} - \dfrac{m_{\mathrm{p}}^{\mathrm{InAsP}}}{m_{\mathrm{p}}^{\mathrm{InGaAs}}} q\upsilon_{\mathrm{pr}}^{\mathrm{InGaAs}} N_{\mathrm{v}}^{\mathrm{InGaAs}} \mathrm{F}_{1/2}\left(\mathrm{invF}_{1/2}\left(\dfrac{p|_{z=s+0}}{N_{\mathrm{v}}^{\mathrm{InGaAs}}}\right) + \dfrac{\Delta E_{\mathrm{v}}^{\mathrm{InAsP-InGaAs}}}{k_{\mathrm{B}}T}\right), \end{cases} \qquad \text{(S40)}$$

where $m_{\mathrm{n}}^{\mathrm{InAsP}}$ ($m_{\mathrm{n}}^{\mathrm{InGaAs}}$) and $m_{\mathrm{p}}^{\mathrm{InAsP}}$ ($m_{\mathrm{p}}^{\mathrm{InGaAs}}$) are the effective electron and hole masses in InAsP (InGaAs), $\varepsilon^{\mathrm{InAsP}}$ and $\varepsilon^{\mathrm{InGaAs}}$ are the static dielectric constants of InAsP and InGaAs, $\Delta E_{\mathrm{c}}^{\mathrm{InAsP\text{-}InGaAs}}$ and $\Delta E_{\mathrm{v}}^{\mathrm{InAsP\text{-}InGaAs}} = E_{\mathrm{g}}^{\mathrm{InGaAs}} - E_{\mathrm{g}}^{\mathrm{InAsP}} - \Delta E_{\mathrm{c}}^{\mathrm{InAsP\text{-}InGaAs}}$ are the conduction- and valence-band discontinuities, i.e. the differences in energy of the conduction and valence band edges in $InAs_{0.4}P_{0.6}$ and $In_{0.72}Ga_{0.28}As$ at the heterojunction (at room temperature $\Delta E_{\mathrm{c}}^{\mathrm{InAsP\text{-}InGaAs}} = -0.13$ eV [24,47]), $\upsilon_{\mathrm{nr}}^{\mathrm{InAsP}}$ ($\upsilon_{\mathrm{nr}}^{\mathrm{InGaAs}}$) and $\upsilon_{\mathrm{pr}}^{\mathrm{InAsP}}$ ($\upsilon_{\mathrm{pr}}^{\mathrm{InGaAs}}$) are the effective recombination or collection velocities for electrons and holes in InAsP (InGaAs), which are equal to the quarter of the corresponding average thermal velocities in InAsP (InGaAs). Boundary conditions at the second (InGaAs/AlInAs) heterojunction are written in the same way.

## *6.2. Optoelectronic model*

As it was discussed above, the continuity equations for electron and hole generation and recombination involves three processes: non-radiative Auger recombination ($U_{\mathrm{Auger}}$) and recombination for the spontaneous ($U_{\mathrm{spont}}$) and stimulated ($U_{\mathrm{stim}}$) emission. The latter two connect electrical and optical properties of the structure.

Recombination rate $U_{\mathrm{Auger}}$ attributed to Auger process is proportional to the third power of excess carrier concentration and is expressed as

$$U_{\mathrm{Auger}}(z) = [C_{\mathrm{p}}(z)p(z) + C_{\mathrm{n}}(z)n(z)] \times [n(z)p(z) - n_{\mathrm{eq}}(z)p_{\mathrm{eq}}(z)], \qquad \text{(S41)}$$

where $n$ and $p$ are the electron and hole concentrations, $n_{\mathrm{eq}}$ and $p_{\mathrm{eq}}$ are the equilibrium electron and hole concentrations, $C_{\mathrm{n}}$ and $C_{\mathrm{p}}$ are the electron and hole Auger recombination coefficients. In $In_{0.72}Ga_{0.28}As$ at room temperature, $C_{\mathrm{n}} + C_{\mathrm{p}} = C = 3.8\times10^{-28}$ cm$^6$s$^{-1}$ [48] and the

Auger recombination is mainly induced by the conduction-to-heavy-hole recombination accompanied by the heavy-to-split-off transitions, therefore, $C_n << C_p \approx C$.

Spontaneous emission includes emission into free space $\left(U_{\text{spont}}^{\text{free}}\right)$ and into the resonator modes $\left(U_{\text{spont}}^{\text{resonator}}\right)$:

$$U_{\text{spont}} = U_{\text{spont}}^{\text{free}} + U_{\text{spont}}^{\text{resonator}}. \tag{S42}$$

Recombination rate associated with emission into free space can be expressed as

$$U_{\text{spont}}^{\text{free}}(z) = P_{\text{free}}^{\text{s}}(z) B_{\text{bulk}}\left(n(z), p(z)\right) \times [n(z)\, p(z) - n_{\text{eq}}(z)\, p_{\text{eq}}(z)], \tag{S43}$$

where function $B_{\text{bulk}}(n,p)$ is calculated separately and fitted with the polynomial function of $n$ and $p$ (Section 5.1) and $P^{\text{s}}_{\text{free}}$ is the normalized Purcell factor for emission into free space obtained from the 2D eigenmode FEM simulations

$$P_{\text{free}}^{\text{s}}(z) = \frac{1}{w} \int_{-w/2}^{w/2} P_{\text{free}}(x, z)\, dx. \tag{S44}$$

For each mode of the resonator (see Figure 2 in the main article), spontaneous emission rate $U_i^{\text{spont}}$ is calculated explicitly using equations (S24, S25) and the distribution of the electromagnetic field of the mode in the waveguide cross-section of the ring resonator obtained in the 3D eigenmode simulations. The total recombination rate attributed to spontaneous emission into the resonator modes is equal to the sum over all modes of the resonator:

$$U_{\text{spont}}^{\text{resonator}}(z) = \sum_{i=0}^{M} U_i^{\text{spont}}(z), \tag{S45}$$

where $M$ is the number of modes.

Since the thickness of the InGaAs layer is only 350 nm, spontaneous emission into free space does not experience noticeable attenuation or amplification. The emission into resonator modes is directed parallel to the wafer, which provides favorable conditions for strong light matter interaction. Stimulated emission into the $i^{\text{th}}$ resonator mode can be written as

$$U_i^{\text{stim}}(z) = \frac{c\bar{n}(z)}{8\pi} g_{\text{mat}}\left(n(z), p(z), \hbar\omega_i\right) \frac{S_i}{\upsilon_i^{\text{E}} \hbar\omega_i} \times \frac{\frac{1}{w} \int_{R-w/2}^{R+w/2} \left|E_i(r,z)\right|^2 dr}{\int_{-\infty}^{+\infty} dz \int_{0}^{+\infty} dr W_i(r,z)}, \tag{S46}$$

where $R$ is the radius of the ring, $g_{\text{mat}}$ is the material gain calculated explicitly using expression (S14), $\omega_i$ and $S_i$ are the frequency and power of the $i^{\text{th}}$ resonator mode, respectively, $\upsilon_i^{\text{E}}$ is the energy velocity of the $i^{\text{th}}$ resonator mode obtained in the 2D eigenmode simulations, $E_i$ and $W_i$ are the distributions of complex electric field and the energy density of the $i^{\text{th}}$

resonator mode in the waveguide cross-section. Finally, we need to sum the stimulated emission over all resonator modes:

$$U_{\text{stim}}(z) = \sum_{i=0}^{M} U_i^{\text{stim}}(z). \tag{S47}$$

To complete the self-consistent model, six non-linear differential equations (S32) must be supplemented with $M$ rate equations for $M$ resonator modes, which can be written as follows:

$$\alpha_i^{\text{modal}} S_i = w \int_{-\infty}^{+\infty} \hbar\omega_i U_i^{\text{spont}}(z) dz. \tag{S48}$$

In the above equation, the modal loss of the $i^{\text{th}}$ resonator mode $\alpha_i^{\text{modal}}$ is the sum of $\alpha_i^{\text{S modal}} = -g_i^{\text{S modal}}$ representing loss or gain in the semiconductor, $\alpha_i^{\text{O modal}}$ attributed to the ohmic and radiation losses of the ring-resonator mode and $\alpha_i^{\text{C modal}}$ arising from the emission coupling to the bus plasmonic waveguide (Figure 3(a) in the main article):

$$\alpha_i^{\text{modal}} = -g_i^{\text{S modal}} + \alpha_i^{\text{O modal}} + \alpha_i^{\text{C modal}}. \tag{S49}$$

Here, $g_i^{\text{S modal}}$ can be expressed explicitly

$$g_i^{\text{S modal}} = \frac{c}{8\pi \upsilon_i^{\text{E}} \int_{-\infty}^{+\infty} dz \int_{0}^{+\infty} dr W_i(r,z)} \int_{-\infty}^{+\infty} \bar{n}(z) \left[ \int_{R-w/2}^{R+w/2} |E_i(r,z)|^2 dr \right] g_{\text{mat}}\left(n(z), p(z), \hbar\omega_i\right) dz, \tag{S50}$$

where $\alpha_i^{\text{O modal}}$ is taken from 3D eigenmode simulations (Section 3.4), while $\alpha_i^{\text{C modal}}$ should be evaluated from the 3D FEM simulations. In order to do the latter, a numerical simulation model has been created (Figure S13(a)), in which the ring-resonator modes are excited at their resonant frequencies (these frequencies are found in the 3D eigenmode simulations in Section 3.4) using an input bus waveguide, while the output waveguide was placed on the opposite side of the ring (shown on the top). The field map shown in Figure S13(a) corresponds to the coupling of the $TM^{8}_{00}$ mode, which is the operational mode of the proposed coherent SPP source (see the main article). The output monitoring distance $L_{\text{monit}}$ is set to be sufficiently large in order to avoid the influence of small radiation losses of the ring-resonator mode on the output power. The power flow in the ring is probed a quarter of the circle before the point of the smallest separation to monitor it well beyond the region where the coupling is important. The coupling coefficient is then calculated with a proper correction for the attenuation of the mode in the ring and in the output waveguide section:

$$C^{\text{coupl}} = \frac{I_{\text{out}} \exp\left( \dfrac{L_{\text{out}}}{L_{\text{prop}}^{\text{straight}}} \right)}{I_{\text{ring}} \exp\left( -\dfrac{\pi R}{2 L_{\text{prop}}^{\text{ring}}} \right)}, \tag{S51}$$

where $I_{\text{ring}}$ and $I_{\text{out}}$ are the power flow integrals over the $InAs_{0.40}P_{0.60}$ spacer (domain, where the influence of the scattered fields is minimal) at the corresponding probing positions (Figure S13(a)) and $L_{\text{prop}}^{\text{ring}}$ and $L_{\text{prop}}^{\text{straight}}$ are the mode propagation lengths in the ring and the straight waveguide, respectively.

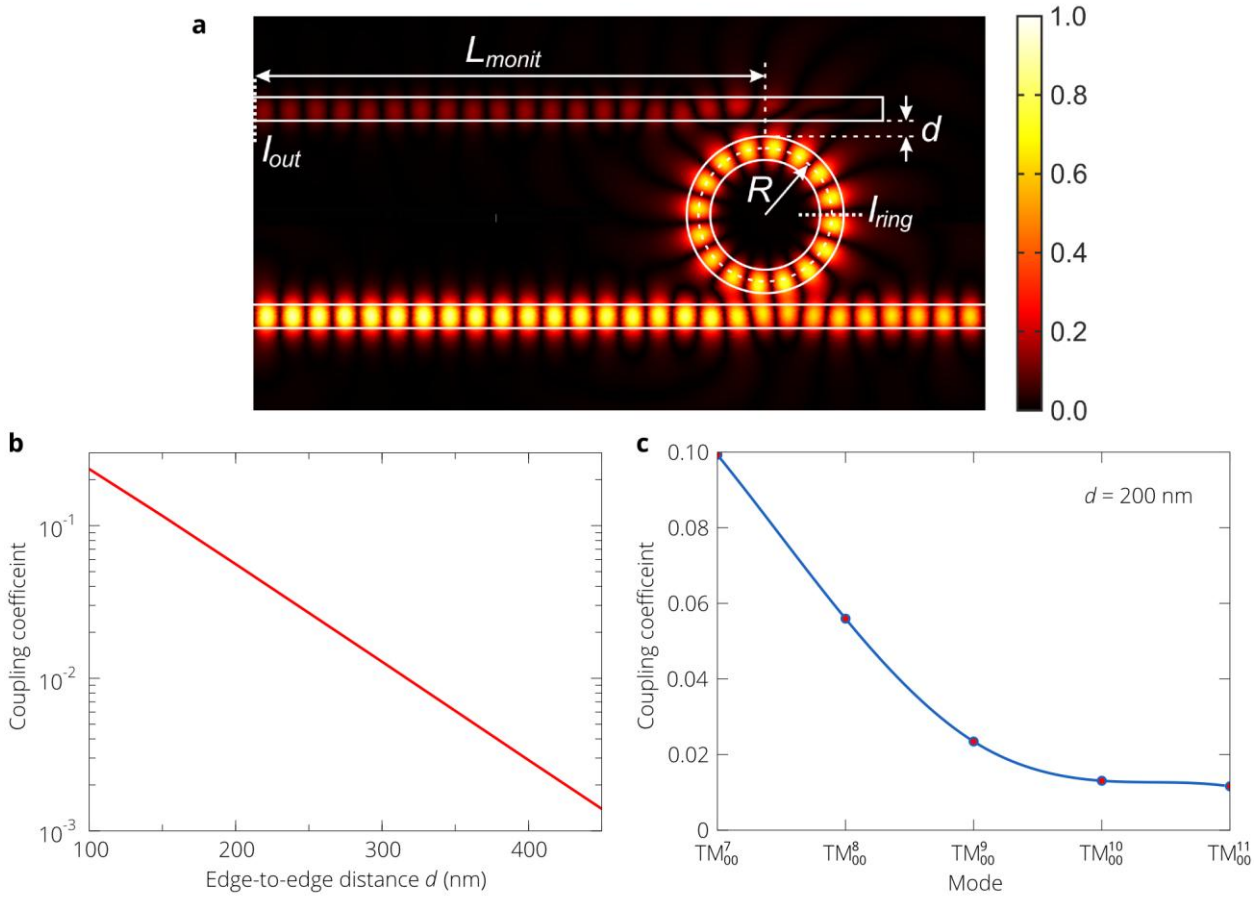



**Figure S13.** (a) Schematics of the geometry for evaluation of the coupling ratio to a straight waveguide (on top) overlaid with the simulated distribution of electric field $|\text{Re}(E_z)|$, when the $TM^8_{00}$ ring-resonator mode is excited using the input bus waveguide. The edge-to-edge distance between the input bus waveguide and the ring resonator is equal to 150 nm and the output bus waveguide is separated from the resonator by a distance of $d$ = 200 nm (b) Coupling coefficient for the $TM^8_{00}$ mode as a function of the edge-to-edge separation distance between the ring resonator and the output but waveguide. (c) Coupling coefficient between the ring resonator with a radius of 850 nm and the output bus waveguide calculated for $TM_{00}$ modes of different orders.

The coupling coefficient for the operational $TM^8_{00}$ mode is studied as a function of the edge-to-edge separation distance between the ring resonator and the output bus waveguide. As can be expected, the coupling ratio decreases exponentially as the separation distance decreases (Figure S13(b)) [49]. On the basis of the results, in the design of the coherent plasmonic source a separation distance of 200 nm is used providing the optimal trade-off between the magnitude of the output signal and the inflicted "useful" cavity losses.

As the next step, we compare the coupling coefficients $C^{coupl}$ for $TM^{7}_{00}$, $TM^{8}_{00}$ and $TM^{9}_{00}$ modes, which compete for the gain provided by the InGaAs active medium (Figure S13(c)). Since the fields are extended out of the semiconductor core further for lower orders of $TM_{00}$ modes (the modes with the lower resonant frequencies), their coupling coefficient is higher at the fixed value of the separation distance between the ring resonator and the output bus waveguide. In contrast, high frequency $TM^{9}_{00}$ and $TM^{10}_{00}$ modes have better localization of the electromagnetic field and lower coupling.

Finally, we use the coupling coefficient $C^{couple}$ to calculate the distributed coupling loss $\alpha^{C\ modal}$ for each mode of the ring resonator:

$$\alpha^{C\,\mathrm{modal}} = -\frac{\ln\left(1 - C^{\mathrm{coupl}}\right)}{2\pi R}. \tag{S52}$$

For each resonator mode, the output power of the proposed SPP source can be found as

$$S_i^{\mathrm{output}} = S_i C_i^{\mathrm{coupl}} = S_i\left|1 - \exp\left(-2\pi R\alpha_i^{C\,\mathrm{modal}}\right)\right|, \tag{S53}$$

where $C_i^{coupl}$ is coupling coefficient between the ring resonator and the output bus waveguide for the $i^{th}$ resonator mode.

## 6. Surface recombination

Non-radiative surface recombination can make a significant contribution to the total current and dramatically decrease the efficiency of nanoscale electronic and optoelectronic devices. The influence of electron and hole trapping by defects at the interfaces on the characteristics of the proposed coherent SPP source is evaluated by introducing a surface recombination in the model in a self-consistent way.

We treat the surface recombination rate using the Shockley-Read-Hall model [50]:

$$U_{\mathrm{surf}}(n,p) = \frac{np - n_{\mathrm{eq}}p_{\mathrm{eq}}}{\frac{1}{S_{\mathrm{p}}}(n+n_1) + \frac{1}{S_{\mathrm{n}}}(p+p_1)}, \tag{S54}$$

where $n$ and $p$ are the electron and hole concentrations nearby the interface, $n_{eq}$ and $p_{eq}$ are the equilibrium carrier concentration, and $n_1$ and $p_1$ are the equilibrium concentration of electron and holes, when the Fermi level at the interface coincides with the energy level of the trap center, $S_n$ and $S_p$ are the SRVs for electrons and holes, respectively, which can be also expressed as

$$S_{\mathrm{n}} = N_{\mathrm{trap}}\sigma_{\mathrm{n}}\upsilon_{\mathrm{Tn}}, \tag{S55}$$

$$S_{\mathrm{p}} = N_{\mathrm{trap}}\sigma_{\mathrm{p}}\upsilon_{\mathrm{Tp}}. \tag{S56}$$

Here, $N_{trap}$ is the surface density of trapping centers at the interface, $\sigma_n$ and $\sigma_p$ are the capture cross-sections for electron and holes, $\upsilon_{Tn}$ and $\upsilon_{Tp}$ are the thermal velocities of electrons and holes, respectively. Accordingly, the boundary conditions at the interfaces are changed appreciably. For the InAsP/InGaAs heterojunction we obtain

$$\begin{cases} \varphi\big|_{z=s-0} = \varphi\big|_{z=s+0}, \\ \varepsilon^{\mathrm{InAsP}} E_z\big|_{z=s-0} = \varepsilon^{\mathrm{InGaAs}} E_z\big|_{z=s+0}, \\ J_n\big|_{z=s+0} = J_n\big|_{z=s-0} + qU_{surf}\big|_{z=s-0} + qU_{surf}\big|_{z=s+0}, \\ J_p\big|_{z=s+0} = J_p\big|_{z=s-0} - qU_{surf}\big|_{z=s-0} - qU_{surf}\big|_{z=s+0}, \\ J_n\big|_{z=s-0} = \dfrac{m_n^{\mathrm{InAsP}}}{m_n^{\mathrm{InGaAs}}} q\upsilon_{nr}^{\mathrm{InGaAs}} N_c^{\mathrm{InGaAs}} \mathrm{F}_{1/2}\left(\mathrm{invF}_{1/2}\left(\dfrac{n\big|_{z=s+0}}{N_c^{\mathrm{InGaAs}}}\right) + \dfrac{\Delta E_c^{\mathrm{InAsP-InGaAs}}}{k_B T}\right) \\ \quad - q\upsilon_{nr}^{\mathrm{InAsP}} n\big|_{z=s-0} - qU_{surf}\big|_{z=s-0}, \\ J_p\big|_{z=s-0} = -\dfrac{m_p^{\mathrm{InAsP}}}{m_p^{\mathrm{InGaAs}}} q\upsilon_{pr}^{\mathrm{InGaAs}} N_v^{\mathrm{InGaAs}} \mathrm{F}_{1/2}\left(\mathrm{invF}_{1/2}\left(\dfrac{p\big|_{z=s+0}}{N_v^{\mathrm{InGaAs}}}\right) + \dfrac{\Delta E_v^{\mathrm{InAsP-InGaAs}}}{k_B T}\right) \\ \quad + q\upsilon_{pr}^{\mathrm{InAsP}} p\big|_{z=s-0} + qU_{surf}\big|_{z=s-0}, \end{cases} \tag{S57}$$

where

$$U_{surf}\big|_{z=s-0} = \frac{n\big|_{z=s-0} p\big|_{z=s-0} - n_{eq}\big|_{z=s-0} p_{eq}\big|_{z=s-0}}{\dfrac{1}{S_p^{\mathrm{InAsP\text{-}InGaAs}}}\left(n\big|_{z=s-0} + n_1\big|_{z=s-0}\right) + \dfrac{1}{S_n^{\mathrm{InAsP\text{-}InGaAs}}}\left(p\big|_{z=s-0} + p_1\big|_{z=s-0}\right)}, \tag{S58}$$

and

$$U_{surf}\big|_{z=s+0} = \frac{n\big|_{z=s+0} p\big|_{z=s+0} - n_{eq}\big|_{z=s+0} p_{eq}\big|_{z=s+0}}{\dfrac{1}{S_p^{\mathrm{InGaAs\text{-}InAsP}}}\left(n\big|_{z=s+0} + n_1\big|_{z=s+0}\right) + \dfrac{1}{S_n^{\mathrm{InGaAs\text{-}InAsP}}}\left(p\big|_{z=s+0} + p_1\big|_{z=s+0}\right)}. \tag{S59}$$

The boundary conditions at the InGaAs/AlInAs heterojunction are written in the same way. In the above expressions, $S_n^{\mathrm{InGaAs\text{-}InAsP}}$ and $S_p^{\mathrm{InGaAs\text{-}InAsP}}$ are the surface recombination velocities (SRVs) for electrons and holes in InGaAs at the InAsP/InGaAs heterojunction. Similarly, $S_n^{\mathrm{InAsP\text{-}InGaAs}}$ and $S_p^{\mathrm{InAsP\text{-}InGaAs}}$ are the SRVs for electrons and holes in InAsP at the InAsP/InGaAs interface. For electrons and holes in $In_{0.72}Ga_{0.28}As$, the SRVs at the $InAs_{0.40}P_{0.60}/In_{0.72}Ga_{0.28}As$ and $In_{0.72}Ga_{0.28}As/Al_{0.29}In_{0.71}As$ heterojunctios do not exceed $2\times10^3$ cm/s [51]. In our simulations, we also use this value ($2\times10^3$ cm/s) for recombination in $InAs_{0.40}P_{0.60}$ and $Al_{0.29}In_{0.71}As$, since surface recombination plays a significant role only in the active InGaAs region, where there are very high concentrations of excess electrons and holes. This guarantees that the first term in the numerator of equation (S54) is much larger than the second one. In contrast, the concentrations electrons and holes in InAsP and AlInAs are close to equilibrium values, therefore, $U_{surf}\big|_{z=s-0}$ and $U_{surf}\big|_{z=h+0}$ are much smaller than $U_{surf}\big|_{z=s+0}$ and

$U_{surf}|_{z=h\text{-}0}$ and do not have an impact on the injection current and other characteristics of the proposed coherent SPP source.

Figure S15 presents the simulated distribution of electrons and holes near the InAsP/InGaAs and InGaAs/AlInAs interfaces at a current density of 30 kA/cm$^2$. It is clearly seen that the minority carrier concentrations in InAsP and AlInAs are much smaller than the concentrations of majority carrier even in the regions near the heterojunction. In the active InGaAs layer, the electron and hole concentrations are almost equal to each other (Figures S15), except very thin regions near the heterojunction. It should be noted that concentrations in the bulk of the active layer (e.g., in the center of this layer) are typically used to estimate the influence of surface recombination on the device efficiency. However, Figure S15(c,d) demonstrates that this approach can hardly be used for quantitative calculations. In our structure, at current density of 30 kA/cm$^2$, the surface recombination rate at the InAsP/InGaAs interface is $1\times10^{21}$ cm$^{-2}$s$^{-1}$, it is twice as small as that calculated using the concentrations in the bulk of the InGaAs layer Figure S15(c). At the InGaAs/AlInAs interface, this effect is not well pronounced under the same injection current (Figure S15(d)). The net surface recombination rate is about $3.3\times10^{21}$ cm$^{-2}$s$^{-1}$ corresponding to a current density of 530 A/cm$^2$, which is negligible compared with the total current $J$ = 30 kA/cm$^2$.

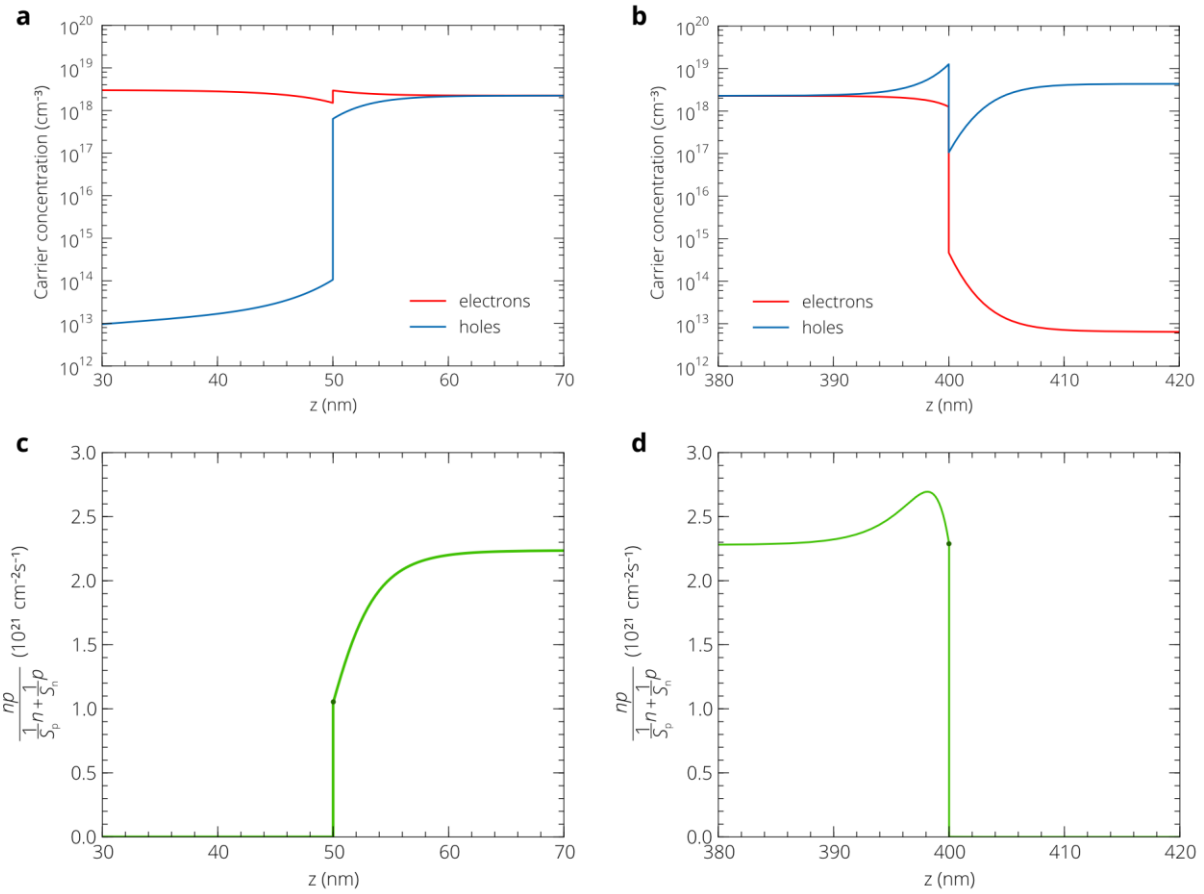



**Figure S15.** (a,b) Simulated distribution of carrier concentration and (c,d) profiles of the Shockely-Read-Hall rate equation $np/\left(n/S_{\mathrm{p}}+p/S_{\mathrm{n}}\right)$ in the vicinity of (a,c) InAsP/InGaAs and (b,d) InGaAs/AlInAs heterojunction at a current density of 30 kA/cm$^2$.

Figure S16 shows that surface recombination does not have a substantial influence on the output characteristics of the proposed SPP source. At a current density of 30 kA/cm$^2$, the contribution of surface recombination to the total current is only 1.8%. We should emphasize that poor-quality interfaces can give rise to the surface density of trapping centers and increase the SRV up to 1×10$^4$ cm/s. Nevertheless, even in this case, the contribution of surface recombination will not exceed 3 kA/cm$^2$, which is only about 10% of the total current.

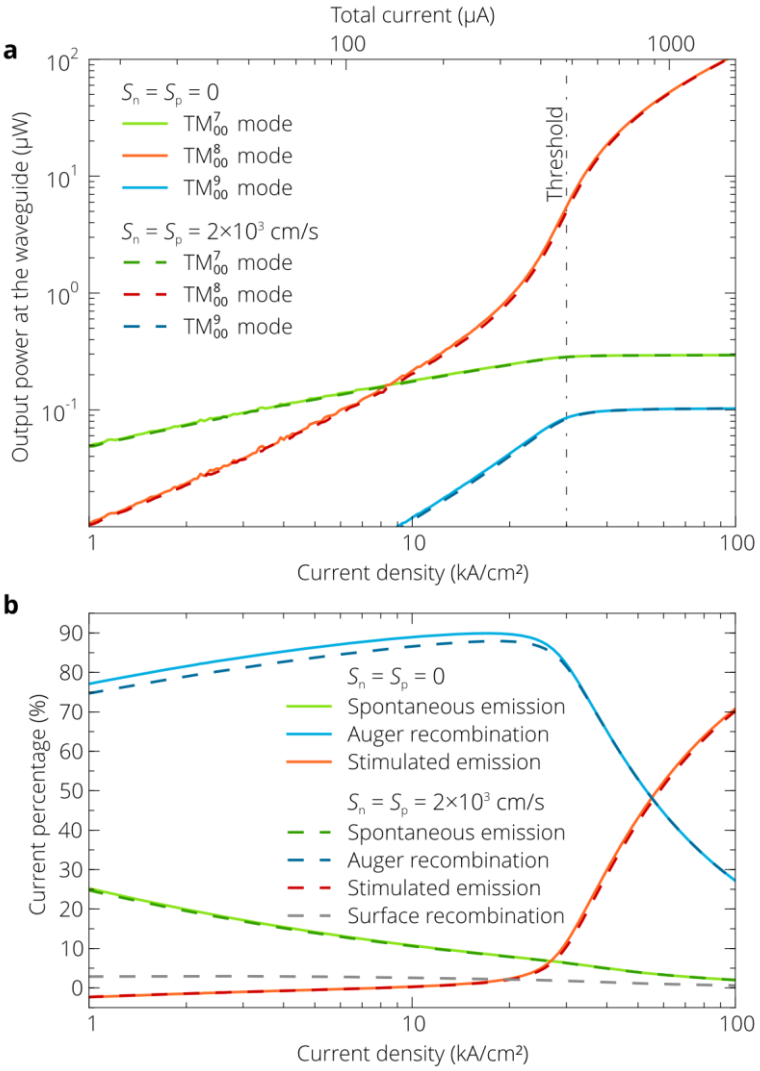



**Figure S16.** The dependence of (a) output power in the waveguide and (b) percentage contribution of different recombination processes (Auger recombination, spontaneous emission, stimulated emission and surface recombination) in the total current on the injection current. The current percentage for stimulated emission is negative below $J$ = 8 kA/cm$^2$ where absorption prevails over stimulated emission.

## 7. Coupling to photonic waveguide

Coupling of the nanolaser output to photonic waveguides instead of plasmonic ones has a lot of advantages, such as compatibility with a standard photonic circuitry and low propagation loss. To compare the coupling efficiency from the nanolaser to the native plasmonic and Si photonic waveguides, eigenmode numerical simulations of the resonator coupled to an output bus waveguide were performed. For a fare comparison, the imaginary part of the eigenfrequency (related to the combined output power into the bus waveguide, ohmic, scattering and radiation losses) was equalised and set to the one in the lasing regime considered above. The parameters and position of the photonic waveguide were optimized to obtain the best coupling efficiency. The Si photonic waveguide has a cross-section of 400 nm×400 nm and supports one TE and one TM mode. It was placed at an edge-to-edge distance of 350 nm from the nanolaser and 75 nm above the metal surface supporting the SPP mode (Figure S17). In the absence of a continuous plasmonic contact layer in this case, the nanolaser was placed on a 100 nm thick metallic island of a circular shape ($R_{\mathrm{met}} = 1.65\,\mu\mathrm{m}$), cut below the facing edge of the photonic waveguide. The optical power coupled to the output Si waveguide was calculated to be about 5 times lower than in the case of a native plasmonic waveguide, considered in the main text. The underlying reason is the mismatch between the modes in the ring-resonator and the Si waveguide leading to an additional scattering loss, while in the case of a plasmonic waveguide they are perfectly matched, both in terms of the modal index and spatial distribution. Therefore, albeit of a lower efficiency, the coupling from the proposed nanolaser to photonic waveguides is possible and due to the advantages for the latter may be useful for certain applications.

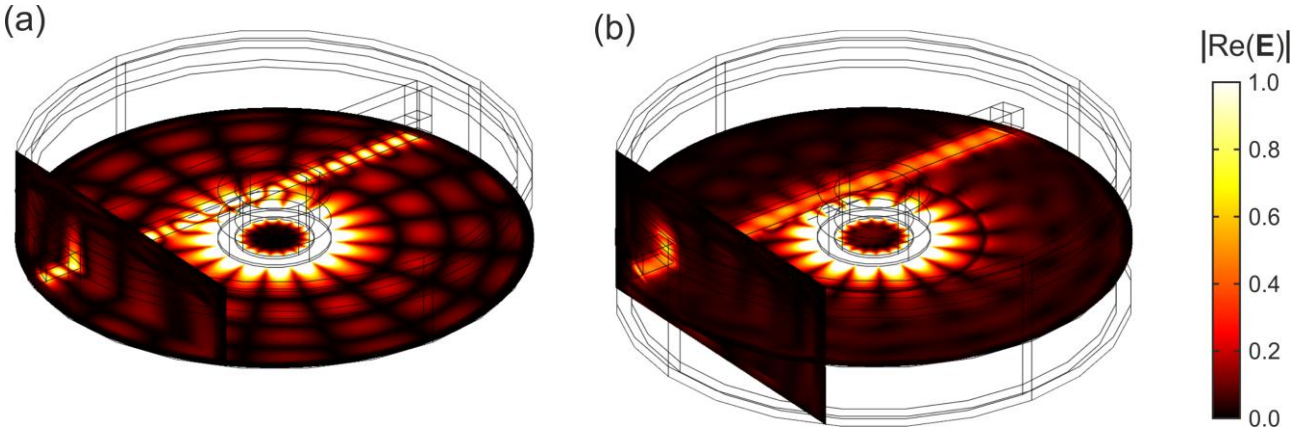



**Figure S17.** $\left|\mathrm{Re}(\mathbf{E})\right|$ fieldmaps (a view from the bottom) plotted at a 25 nm height above the metal surface for the nanolaser coupled to (a) a native plasmonic (with sizes given in Fig. 1) and (b) a silicon (400 nm×400 nm) waveguides obtained using the eigenmode numerical simulations. The scale is saturated to visualize the coupling region and the output mode profiles.

## References


[1] Naik, G. V., Shalaev, V. M., Boltasseva, A. Alternative Plasmonic Materials: Beyond Gold and Silver. *Adv. Mater.* **25**, 3264−3294 (2013).

[2] Kinsey, N., Ferrera, M., Naik, G. V., Babicheva, V. E., Shalaev, V. M., Boltasseva, A. Experimental demonstration of titanium nitride plasmonic interconnects. *Opt. Express* **22**, 12238−12247 (2014).

[3] Johnson, P. B. & Christy, R. W. Optical constants of the noble metals. *Phys. Rev. B* **6**, 4370-4379 (1972).

[4] Fedyanin, D. Y. & Arsenin A. V. Surface plasmon polariton amplification in metal-semiconductor structures. *Opt. Express* **19**, 12524-12531 (2011).

[5] Fedyanin, D. Y. Toward an electrically pumped spaser. *Opt. Lett.* **37**, 404-406 (2012).

[6] Mead, C. A. & Spitzer, W. G. Fermi level position at semiconductor surfaces. *Phys. Rev. Lett.* **10**, 471-472 (1963).

[7] Walpole, J. N. & Nill, K. W. Capacitance-voltage characteristics of metal barriers on p PbTe and p InAs: effects of the inversion layer. *J. Appl. Phys.* **42**, 5609-5617 (1971).

[8] Fedyanin, D. Y., Krasavin, A. V., Arsenin, A. V., & Zayats, A. V. Surface plasmon polariton amplification upon electrical injection in highly integrated plasmonic circuits, *Nano Lett.* **12**, 2459-2463 (2012).

[9] Costantini, D. Generation and Amplification of Surface Plasmon Polaritons at Telecom Wavelength with Compact Semiconductor-Based Devices (Université Paris Sud - Paris XI, 2013).

[10] Costantini, D., Bousseksou, A., Fevrier, M., Dagens, B. & Colombelli, R. Loss and gain measurements of tensile-strained quantum well diode lasers for plasmonic devices at telecom wavelengths. *IEEE J. Quantum Electron.* **48**, 73–78 (2012).

[11] Baca, A. G., Ren, F., Zolper, J. C., Briggs, R. D. & Pearton, S. J. A survey of ohmic contacts to III-V compound semiconductors. *Thin Solid Films* **308-309**, 599–606 (1997).

[12] Hill, M. T. *et al.* Lasing in metallic-coated nanocavities. *Nat. Photonics* **1**, 589-594 (2007).

[13] Ding, K. *et al.* Room-temperature continuous wave lasing in deep-subwavelength metallic cavities under electrical injection. *Phys. Rev. B* **85**, 041301(R) (2012).

[14] Suh, J. Y. *et al.* Plasmonic bowtie nanolaser arrays. *Nano Lett.* **12**, 5769-5774 (2012).

[15] Khurgin, J. B. & Sun, G. How small can “Nano” be in a “Nanolaser”? *Nanophotonics* **1**, 3-8 (2012).

[16] Takeda, K. *et al.* Few-fJ/bit data transmissions using directly modulated lambda-scale embedded active region photonic-crystal lasers. *Nat. Photonics* **7**, 569-575 (2013).

[17] Matsuo, S. *et al.* Photonic crystal lasers using wavelength-scale embedded active region. J. Phys. D: Appl. Phys. **47**, 023001 (2014).

[18] Blaber, M. G., Arnold, M. D. & Ford, M. J. A review of the optical properties of alloys and intermetallics for plasmonics. *J. Phys. Condens. Matter* **22**, 143201 (2010).

[19] Stareev, G., Kunzel, H. & Dortmann, G. A controllable mechanism of forming extremely low-resistance nonalloyed ohmic contacts to group III-V compound semiconductors. *J. Appl. Phys.* **74**, 7344-7356 (1993).

[20] Cetin, H. & Ayyildiz, E. Temperature dependence of electrical parameters of the Au/n-InP Schottky barrier diodes. *Semicond. Sci. Technol.* **20**, 625-631 (2005).

[21] Kajiyama, K., Mizushima, Y. & Sakata, S. Schottky barrier height of n-InxGa1-xAs diodes. *Appl. Phys. Lett.* **23**, 458-459 (1973).

[22] Bhargava, S., Blank, H.-R., Narayanamurti, V. & Kroeme, H. Fermi-level pinning position at the Au-InAs interface determined using ballistic electron emission microscopy. *Appl. Phys. Lett.* **70**, 759-761 (1997).

[23] Ohler, C., Daniels, C., Forster, A. & Luth, H. Barrier height at clean Au/InAs(100) interfaces. *J. Vac. Sci. Technol. B* **15**, 702-706 (1997).

[24] Tiwari, S. & Frank, D. J. Empirical fit to band discontinuities and barrier heights in III–V alloy systems. *Appl. Phys. Lett.* **60**, 630-632 (1992).

[25] Kajiyama, K., Mizushima, Y. & Sakata, S. Schottky barrier height of n-InxGa1-xAs diodes. *Appl. Phys. Lett.* **23**, 458-459 (1973).

[26] Conley, J. W., Duke, C. B., Mahan, G. D., & Tiemann, J. J. Electron tunneling, in metal-semiconductor barriers. *Phys. Rev.* **150**, 466-469 (1966).

[27] Kasap, S. & Capper, P. (Eds.), *Handbook of Electronic and Photonic Materials* (Springer, New York, 2007).

[28] Adachi, S. Physical Properties of III-V Semiconductor Compounds: InP, InAs, GaAs, GaP, InGaAs, and InGaAsP (Wiley, New York, 1992).

[29] Adachi, S. Optical dispersion relations for GaP, GaAs, GaSb, InP, InAs, InSb, AlxGa1−xAs, and In1−xGaxAsyP1−y. *J. Appl. Phys.* **66**, 6030-6040 (1989).

[30] Palik, E. D. *Handbook of Optical Constants of Solids II* (Elsevier Science, 1991).

[31] Herve, P. J. L. & Vandamme, L. K. J. Empirical temperature dependence of the refractive index of semiconductors. *J. Appl. Phys.* **77**, 5476-5477 (1995).

[32] Herve, P. J. L. & Vandamme, L. K. J. in *Proceedings of OPTO 94*, 371-374 (Paris, France, 1994).

[33] Fedyanin, D. Y., Krasavin, A. V., Arsenin, A. V., & Zayats, A. V. Surface plasmon polariton amplification upon electrical injection in highly integrated plasmonic circuits, *Nano Lett.* **12**, 2459-2463 (2012).

[34] Casey, H. C. & Panish, M. B. *Heterostructure Lasers Part A* (Academic, New York, 1978).

[35] Casey, H. C., Stern, F. Concentration-dependent absorption and spontaneous emission on heavily doped GaAs. *J. Appl. Phys.* **47**, 631-643 (1976).

[36] Krishnamurthy, S., Yu, Z. G., Gonzalez, L. P. & Guha, S. Temperature- and wavelength-dependent two-photon and free-carrier absorption in GaAs, InP, GaInAs, and InAsP. *J. Appl. Phys.* **109**, 033102 (2011).

[37] Haug, A. Free-carrier absorption in semiconductor lasers. *Semicond. Sci. Technol.* **7**, 373-378 (1992).

[38] Purcell, E. M. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.* **69**, 681 (1946).

[39] De Leon, I. & Berrini, P. Theory of surface plasmon-polariton amplification in planar structures incorporating dipolar gain media. *Phys. Rev. B* **78**, 161401 (2008).

[40] Ford, G. W. & Weber, W. H. Electromagnetic interactions of molecules with metal surfaces. *Phys. Rep.* **113**, 195-287 (1984).

[41] Snowden, C. M. *Introduction to Semiconductor Device Modelling* (World Scientific, Singapore, 1998).

[42] Sze, S. M. *Physics of Semiconductor Devices* (Wiley, New York, 1981).

[43] Schroeder, D. Modelling of Interface Carrier Transport for Semiconductor Device Simulation (Springer-Verlag, Vienna, 1994).

[44] Rhoderic, E. H. *Metal Semiconductor Contacts* (Clarendon Press, Oxford, 1978).

[45] Horio, K. & Yanai, H. Numerical modeling of heterojunctions including the thermionic emission mechanism at the heterojunction interface. *IEEE Trans. Electron Devices* **37**, 1093-1098 (1990).

[46] Lopez-Gonzalez, J. M. & Prat, L. Numerical modelling of abrupt InP/InGaAs HBTs. *Solid-St. Electron.* **39**, 523-527 (1996).

[47] Lim, A. C. H. *et al.* Determination of InAsP/InP and InGaAs/InP band offsets using blue shifting type II asymmetric multiple quantum wells. *AIP Conf. Proc.* **772**, 977 (2005).

[48] Gfroerer, T. H., Priestley, L. P., Fairley, M. F. & Wanlass, M. W. Temperature dependence of nonradiative recombination in low-band gap InxGa1-xAs/InAsyP-y double heterostructures grown on InP substrates. *J. Appl. Phys.* **94**, 1738-1743 (2003).

[49] Krasavin, A. V. & Zayats, A. V. Passive photonic elements based on dielectric-loaded surface plasmon polariton waveguides. *Appl. Phys. Lett.* **90**, 211101 (2007).

[50] Selberherr, S. *Analysis and Simulation of Semiconductor Devices* (Springer-Verlag, Wien, 1984).

[51] Ning, S., Fay, P., Sinharoy, S., Forbes, D. & Scheiman, D. Characterization and modeling of InGaAs/InAsP thermophotovoltaic converters under high illumination intensities. *J. Appl. Phys.* **101**, 064511 (2007).