Supporting Information for
Ultra-Subwavelength Two-Dimensional Plasmon Electronics

William F. Andress, Hosang Yoon, Kitty Y. M. Yeung, Ling Qin, Ken West, Loren Pfeiffer, Donhee Ham*

* E-mail: donhee@seas.harvard.edu

Hybrid Electromagnetic-Plasmonic Interferometer: Properties of Floating Gate

The electromagnetic propagation in the top branch in the interferometer from Figure 4 of the main text can be understood as follows. Since the gate over the top branch is floating, we can see this gate in conjunction with the adjacent CPW grounds as forming an *electromagnetic* transmission line. As the 2DEG below the floating gate is depleted, signals propagating through this 2DEG predominately couple into this fast electromagnetic transmission line. Even if the 2DEG beneath the floating gate is not completely depleted, propagation is still predominantly through the floating gate. To illustrate this more quantitatively, we can create a simple circuit model of this phenomenon. A gated 2DEG can be modeled using its kinetic inductance, scattering resistance, and capacitance to the gate, as in Figure S1 (a more sophisticated model of the plasmonic propagation is provided later with Figure S5). The floating metallic gate may be treated as a short circuit, given its small length relative to the electromagnetic wavelength. The small device length and the absence of ground connection for the floating gate allow us to treat this path simply as a lumped element with an impedance *Z across* it (not to be confused with the local port input impedance). This impedance can be approximated as $Z = 2Z_0 \cdot \tanh\{(\alpha + j\beta)l/2\}$,

where $Z_0 \sim \sqrt{L/C}$, $\alpha = R_s/2Z_0$, $\beta = \omega\sqrt{LC}$, with L, C, and R_s being the kinetic inductance, capacitance, and scattering resistance per unit length, respectively, and l being the length of the gate. When the 2DEG is depleted to very low electron density, R_s increases, and thus α becomes large enough that Z approaches $2Z_0$ which is real, like a lumped resistor. This indicates that this branch with the floating gate introduces unappreciable phase delay. It physically corresponds to the fact that excitations beneath one edge of the gate simply travel into the gate and to the other side.

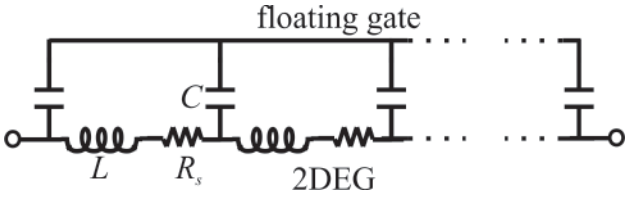

Figure S1. Circuit model of path with floating gate as a lumped element.

Additional Measurements

Measurements are performed with five additional 2D plasmonic crystals to further confirm the behavioral dependency on the planar geometry. Fig. S2a shows a pair of devices sharing the same square-lattice crystal structure of a 14-μm periodicity, but of different overall lengths and widths. As seen in the measured s_{21} magnitudes (Fig. S2b), the bandgap positions, which are governed by the distance between two adjacent vertical crystal planes, are very similar. The difference in the overall amplitude is well expected, as one device is longer, thus more lossy, than the other. The measured dispersions (Fig. S2c; obtained from the measured s_{21} phase; *y*-axis is frequency; *x*-axis is phase delay per unit cell) also indicate very similar bandgap positions (*y*-axis) and passband dispersions; the differing phase delays per unit cell (*x*-axis) within the bandgap are not surprising, as they in general spuriously arise from noise and calibration errors

when the signal transmission is low. In comparison to these results, the bandgap position of another square-lattice device with a larger periodicity of 16 μm (Fig. S2d) is expectedly moved to a lower frequency (Fig. S2e-f). Fig. R1g shows another pair of devices sharing exactly the same crystal structure (this time, hexagonal lattices with a periodicity of 14 μm), but with different overall lengths and widths. The measured s_{21} magnitudes and dispersions (Fig. S2h-i) again show very similar bandgap positions, along with the expected amplitude difference due to different lengths. These new measurements confirm that the key behaviors arise primarily from geometry.

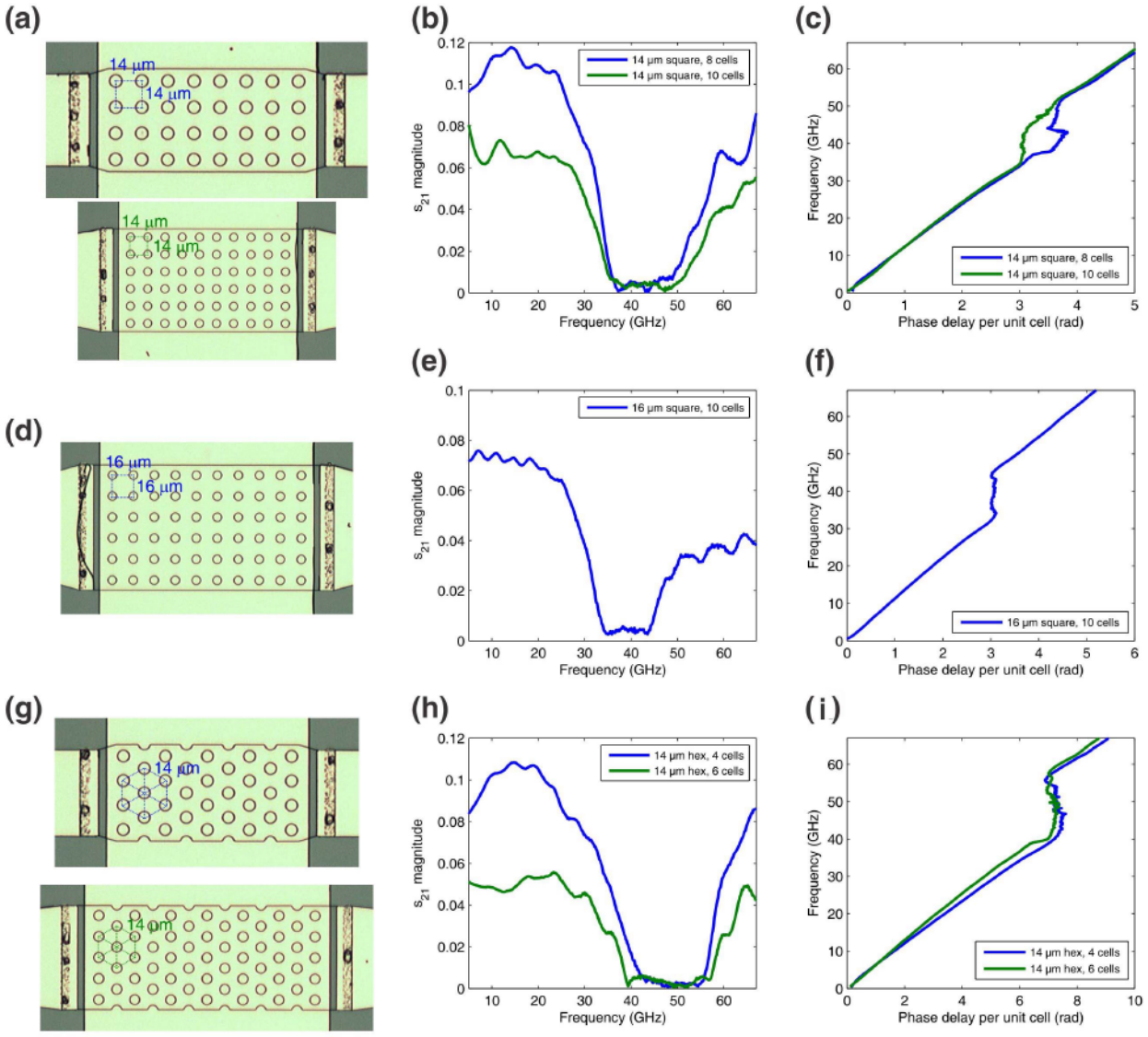

Figure S2. (a-c) Device images, measured s_{21} magnitudes, and s_{21} phase delays per unit cell, respectively, for a pair of devices that share the same square-lattice crystal structure of a 14-μm periodicity (the image scale is different between the two devices), but of different overall lengths and widths. (d-f) Device image and measurement results for another square-lattice device with a larger periodicity of 16 μm. (g-i) Device images and measurement results for another pair of devices that share the same hexagonal-lattice crystal structure of a 14-μm periodicity, but of different overall lengths and widths.

Methods

Fabrication

Our plasmonic circuits are fabricated on GaAs/AlGaAs 2DEG substrates grown by molecular beam epitaxy (MBE). From the 2DEG up, the layer structure is as follows: 48 nm $Al_{0.3}Ga_{0.7}As$, 26 nm Si-doped $Al_{0.3}Ga_{0.7}As$, and 6 nm GaAs cap, totaling to 80 nm. At 3.7 K, before processing, the carrier density is $1.54\times10^{11}/cm^2$, with mobility 2.5×10^6 cm^2/Vs in the dark, and the carrier density is $2.8\times10^{11}/cm^2$ with mobility 3.9×10^6 cm^2/Vs after illumination.

Mesas defining the shape of a plasmonic circuit are formed by photolithography followed by wet etching by >80 nm to beneath the 2DEG layer, using 240:8:1 H_2O:H_2O_2:H_2SO_4. The next photolithography step defines the Ni-Au-Ge ohmic contacts, which overlap the mesas by 6 μm from the ends. After an ammonia dip, the contact metals are deposited by thermal evaporation as follows: 5 nm Ni, 20 nm Au, 25 nm Ge, 10 nm Au, 5 nm Ni, 40 nm Au. After liftoff, the contacts are annealed at 420ºC for 50 seconds. The final photolithography step defines the coplanar waveguides (CPWs) as well as the gate that covers the plasmonic medium, which is connected to the ground (G) lines of the CPWs. The signal (S) lines of the CPWs extend just over the ohmic contacts. The gate is separated from the contacts and signal lines by roughly 4-μm wide ungated areas. The CPWs and gate metals, also deposited by thermal evaporation, consist of 8 nm Cr and 500 nm Au.

Measurements

Measurements take place in a Lakeshore cryogenic probe station, at its minimum achievable

temperature of 3 K. Inside the probe station are two microwave probe arms which we use to contact the CPWs that lie on the chip. Like the on-chip CPWs, the probes have ground-signal-ground tips. The separation between the signal tip and each of the two ground tips (the probe pitch) is 100 µm. Coaxial cables lead from the probes to the network analyzer, which generates the excitation signals (with –45dBm power reaching the devices), and measures the signals that are reflected from and transmitted through the device; the s-parameters s_{11} and s_{21} are the reflected and transmitted voltages relative to the voltage of the excitation, and their magnitude and phase allow us to fully characterize each device. We focus our analyses on s_{21} since s_{11} is sensitive to impedance mismatches into the plasmonic device and to contact resistances. Measurements were initially performed with an Agilent E8364A network analyzer, which supports frequencies up to 50 GHz; since some features of the devices in Figures 1-4 extended beyond 50 GHz, we then used an Agilent E8361A to extend those measurements up to 67 GHz.

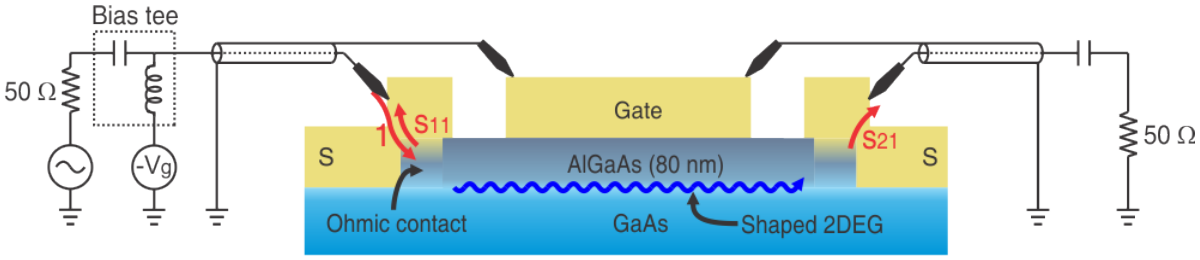

Figure S3. Schematic of measurement setup.

The plasmonic circuits are biased using bias-tees that are installed within the network analyzer (as in the schematic of Figure S3). Since the gate of each device is connected to the CPW ground lines, the gate voltage is fixed to the ground of the measurement system. So to attain a desired gate-source bias, it is actually the dc voltage of the signal lines that is controlled, while the gate remains at 0 volts. So when we refer to a positive [negative] gate bias, in practice we are setting the signal dc bias to a negative [positive] voltage.

The entire measurement setup, including the ports of the network analyzer, the cables, and the probes has an impedance of 50Ω. The on-chip CPWs that lead up to each device are also designed to have 50Ω impedance. For some devices, the CPWs are tapered from the ends where the probes are landed up to where they meet the plasmonic device, but the tapering is done in such a way that they retain 50Ω impedance along their entire length. We use the Sonnet electromagnetic field solver to design these CPWs.

At GHz frequencies, the cables are much longer than the electromagnetic wavelength, and the cables and various connectors exhibit phase delays and loss. As with any microwave measurement, in order for the measurements to reflect only the plasmonic circuits on the chip, we must calibrate the system up to the tips of the probes. We calibrate with the NIST-style TRL method [Marks, R. B. *IEEE Trans. Microwave Theory Tech.* **1991**, *39*, 1205-1215], which requires landing the probes on a separate set of CPWs of various lengths and taking s-parameter measurements. We fabricate calibration substrates, one of which is shown in Figure S4, using undoped GaAs samples; this matches the conditions on the 2DEG sample, since the device CPWs lie over the regions where the donors and 2DEG are etched away. The calibration CPWs are fabricated in the same way as those on the plasmonic circuits. The calibration must be performed at the same temperature when the plasmonic measurements are performed (3K), since the cables within the probe station expand and contract as temperature changes. Using these calibration CPWs, we calibrate up to the probe tips, and thus do not calibrate out the effect of the device CPWs that lead up to the plasmonic devices, because their length of 130 μm is much shorter than the millimeter electromagnetic wavelengths and contributes very little phase delay.

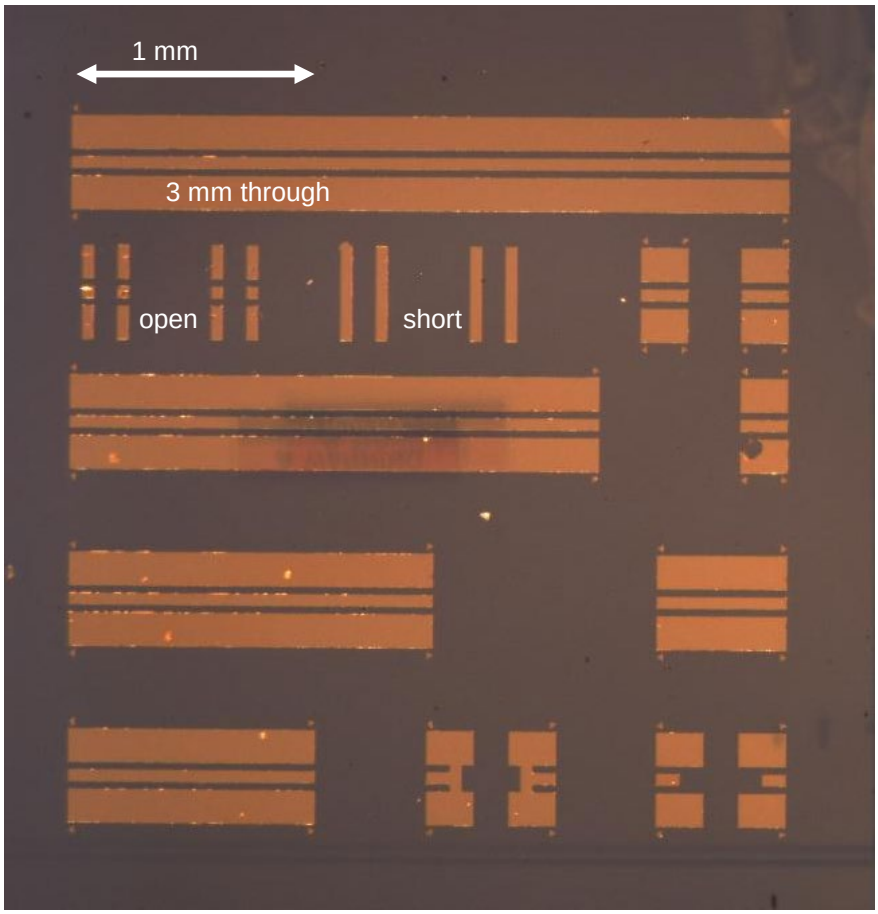

Figure S4. Optical image of calibration substrate, including CPWs of various lengths. The widths of the signal and ground lines, and the spacing between them, are the same as the widths and spacing at the ends of the CPWs on the plasmonic chip. The plasmonic devices in Figures 2d-e and 3 of the main text use larger CPW structures than Figures 1 and 2a-c, so two different calibration substrates are made (the former is shown here). Two longer lines with lengths 6 mm and 9 mm also on substrate are not shown.

De-Embedding of Parasitic Coupling

In addition to calibrating out the effects of the coaxial cables and probes, a parasitic electromagnetic transmission between the measurement probes with very small phase delay,

indicating purely electromagnetic origin, is subtracted from the transmission measurements. Since this parasitic coupling is present even when the plasma is depleted using a negative gate bias, we can measure the parasitic transmission in isolation at the negative bias and then subtract it from the overall transmission.

In measuring the parasitic transmission, to avoid drastically changing the matching conditions at the device terminals from those of the 'on' state, we measure the coupling at a bias where the plasma is not fully depleted, but where the total loss through the device is high enough that plasmonic transmission is negligible (e.g., -0.1 V). In measurements, the parasitic coupling changes proportionally to s_{11} as bias is varied, suggesting that this coupling arises when electromagnetic signals are reflected from impedances mismatches at the contacts. Therefore we subtract the coupling as follows: (s_{21} of plasmonic device at 0.5V) = (s_{21} at 0.5V) – (s_{21} at -0.1V)×(s_{11} at 0.5V)/(s_{11} at -0.1V). Figure S5 shows, as an example, transmission data for the device of Figure 1 before and after subtraction of the probe couplings.

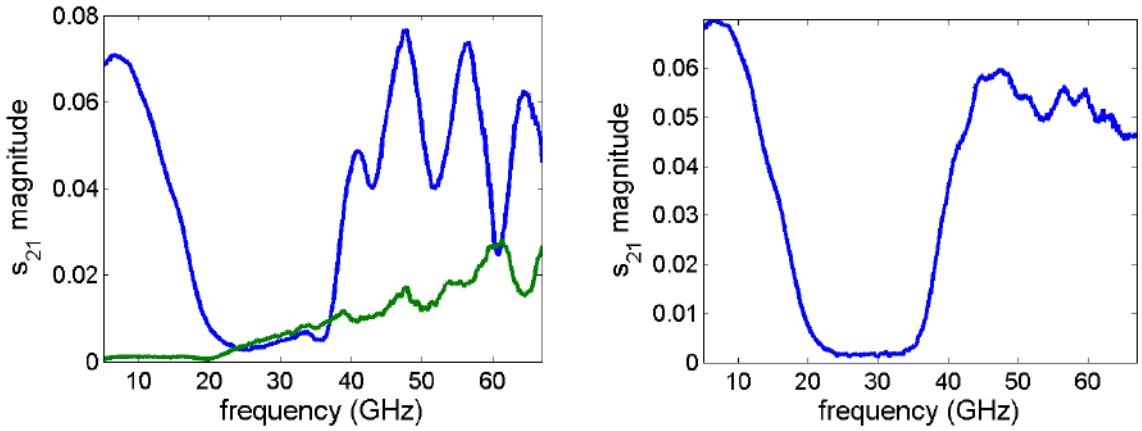

Figure S5. Left: Measured overall transmission for device from Figure 1 of main text, at 0.5V gate bias (blue) and the parasitic transmission measured at -0.1V (green). Right: After subtraction of parasitic coupling.

2DEG Plasma Circuit Modeling

In order to design the circuits presented in this work, it was necessary to find an accurate high-frequency circuit model for the plasma system's general behavior. One important component in such a model is the very large 'kinetic inductance' of the 2DEG [Meservey, R; Tedrow, P. M. *J. Appl. Phys.* **1969**, *40*, 2028-2034], which is a manifestation of the electrons' inertia, or equivalently, their acceleration, a defining component of a collective plasmonic mode. In our sample, at 0.5V gate bias for example, this kinetic inductance was measured as 0.9 nH/square, meaning that, for instance, a kinetic inductance of 3.5 nH could be produced from a 3.9×1 μm^2 2DEG strip, which is ~11,000 times smaller than a typical on-chip magnetic 3.5-nH inductor [Woo, K. *et al*. *IEEE J. Solid-State Circ*. **2008**, *43*, 379].

In addition to the various devices in Figures 1-4, we also characterize a plain gated 2DEG strip along with its ohmic contacts. We fabricate a gated 2DEG with a length of 67 μm and width of 24 μm, contacted in the same way as the devices in the main text. The gated region can be modeled as a *plasmonic transmission line* with distributed kinetic inductance and distributed capacitance. The capacitance represents the electrostatic capacitance resulting from Coulomb interactions, in series with the quantum capacitance resulting from Pauli pressure. (Figure S6, top) Note that although the plasmonic medium is modeled in a similar manner to an electromagnetic transmission line, the physics described by the model is vastly different. If the gate is placed sufficiently close to the 2DEG (in our case 80 nm above the 2DEG) the usual magnetic inductance between the 2DEG and the gate becomes negligibly small compared to the very large kinetic inductance of the 2DEG. This is responsible for the extreme plasma wave slowing of our work, as the plasma wave velocity derived from the transmission line model is $1/\sqrt{LC}$ where L and C are inductance (dominated by the kinetic inductance which is much

larger than the usual magnetic inductance) and capacitance per unit length. In a purely electromagnetic transmission line, wave velocity remains comparable to the speed of light.

The losses due to electron scattering within the 2DEG can be modeled as a resistance in series with the distributed inductance (Figure S6, top). We also find that in our sample, the donor layer underneath the gate is not fully depleted of conducting charges, and dissipation that occurs in this layer capacitively coupled to the 2DEG can be modeled with a conductance in parallel to the kinetic inductance (Figure S6, top). The ohmic contacts are more difficult to model because their behavior at GHz frequencies is not very well understood. The model that seems to best fit measurement data includes the dc contact resistance, followed by a short distributed *LC* network (Figure S6, top), which could represent a short length of 2DEG that is overhung by the signal line of the CPW, as it may not be perfectly aligned to the ohmic contact region. A phase delay for the on-chip CPW leads is also included in the model (though not shown in schematic).

The s-parameter values of the model can be made to fit the measured data very well when the values of all the circuit elements are optimized (Figure S6). The fitting is performed in MATLAB using the 'lsqcurvefit' function. The fitted values of the circuit elements allow us to characterize the 2DEG strip. For instance, the plasma wave velocity is $1/\sqrt{LC}$ as discussed above, where *L* and *C* are the extracted kinetic inductance and capacitance per unit length. As discussed in the main text, we can determine the plasma wave velocity just by a quick examination of the measured phase of s_{21}, but this phase also includes a delay from the ohmic contacts. Extracting the specific 2DEG circuit parameters from the measurement data characterizes the 2DEG strip more accurately because the 2DEG and ohmic contacts each have their own fitted parameters.

Using measurements of a gated 2DEG strip with 76 μm length and 24 μm width, we thus

extract the plasma velocities at various gate bias voltages. As mentioned in the main text, we extract velocities from *c*/187 at 0.8V, down to *c*/660 at -0.13 V. As the bias is varied, only the extracted kinetic inductance and loss parameters of the 2DEG are significantly changed, while the extracted parameters of the ohmic contacts remain roughly constant, which is one cogent indicator that the model is physically sound.

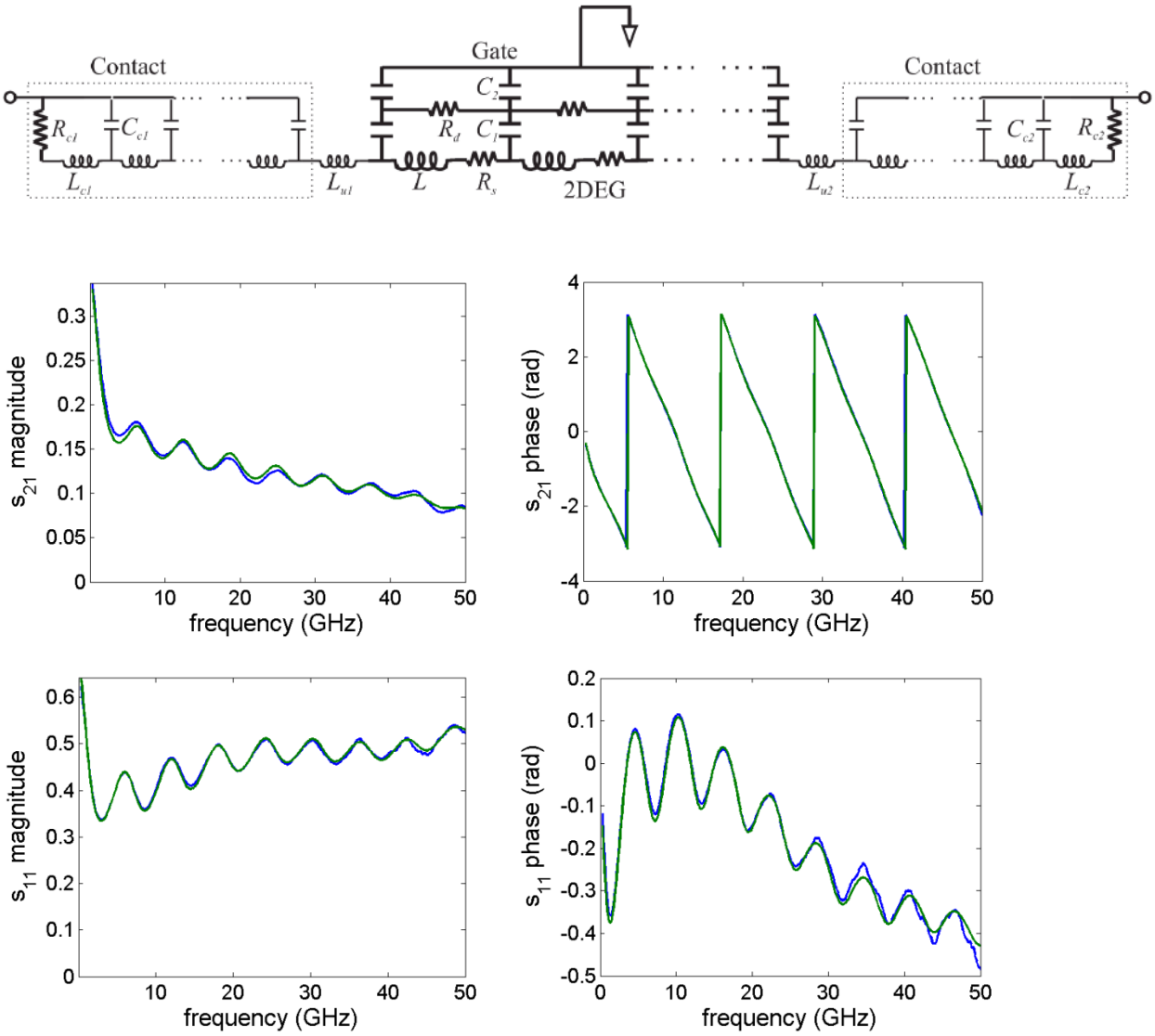

Figure S6. Top: Circuit model for gated 2DEG strip & ohmic contacts. Bottom: Blue & green curves are measured & calculated *s*-parameters for fitted values of model parameters, showing close match.

Ohmic Contact De-embedding

In the short-length stand-alone cavity device of Figure 2d, a weak Fabry-Perot resonance occurs between the two ohmic contacts, appearing as a broad dip around 34 GHz in Figure S6b, which displays the s_{21} magnitude through the device before ohmic contacts are de-embedded. Since we are interested only in the behaviors of the plasmonic cavity circuit, we can de-embed the ohmic contacts using the fitted parameters from the model discussed in the previous section. We calculate the transmission matrices of the ohmic contacts, and multiply their inverses by the transmission matrix of the plasmonic cavity structure (converted from its measured s-parameters). This process de-embeds the parasitic resonance from the measured data of Figure S7b, leading to Figure 2e. Fabry-Perot resonance is appreciable only in the short-length plasmonic cavity of Figure 2d, and it does not appear in any of the other devices presented, all of which are longer.

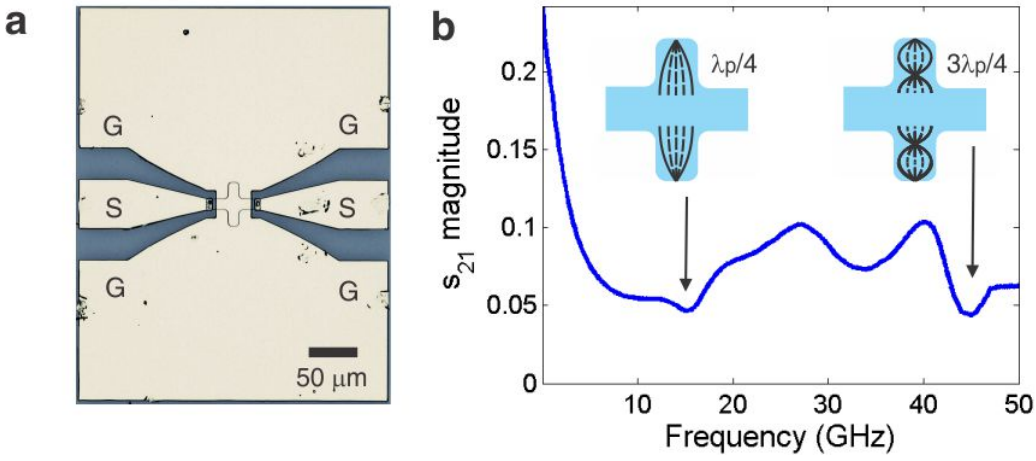

Figure S7. (a) Optical image of the stand-alone cavity structure from Figure 2d. (b) Transmission s_{21} magnitude through the cavity structure, without ohmic contact de-embedding. The broad dip around 34 GHz indicates weak Fabry-Perot resonance between the contacts. This de-embedding is not necessary in other, longer devices.