

## Simultaneous control of intensity, phase, and polarization in real time under a weak oscillation theory: supplement

RUNZHANG XIE,<sup>1,2</sup>  PENG WANG,<sup>1</sup> FANG WANG,<sup>1</sup> XIAOSHUANG CHEN,<sup>1</sup> WEI LU,<sup>1</sup> AND WEIDA HU<sup>1,2,\*</sup> 

<sup>1</sup>State Key Laboratory of Infrared Physics, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China

<sup>2</sup>University of the Chinese Academy of Sciences, Beijing 100049, China

\*Corresponding author: [wdu@mail.sitp.ac.cn](mailto:wdu@mail.sitp.ac.cn)

---

This supplement published with The Optical Society on 10 March 2021 by The Authors under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: <https://doi.org/10.6084/m9.figshare.13931450>

Parent Article DOI: <https://doi.org/10.1364/OL.412851>

# Simultaneous control of intensity, phase, and polarization in real-time under weak oscillation theory: supplemental document

5

## CONTENTS

S1. Two-dimensional oscillator system

1. The derivation of transmission and reflection fields

10

2. The derivation of equation 4 in the main text.

S2. Simulation methods

S3. Jones matrix expressions

S4. Optimization algorithm for full Jones matrix control

1. Gradual optimization algorithm without loss

15

2. Adiabatic method to include the loss into the gradual optimization algorithm

3. Gradient descent method

S5. Design configuration and feasibility of the structure

S6. Windowed Fourier transformation of a simple circuit

Reference

20

## S1. Two-dimensional Oscillator system

Let us first discuss the single-layered system, which is composed of a set of oscillators in xOy-plane being periodic along the x-direction and transverse symmetric along the y-direction. As a two-dimensional material system, the method dealing with the response of the system as a surface current distribution is used by many researchers. However, this method only changes our problem to another equivalent form, unless oversimplified assumptions are applied, such as assuming that the periodic system is uniform along the x-direction. Another candidate theory is the plasma model of the free electron gas, which calculate the dielectric function in the interior of the bulk material, and determine the unique solution by taking the boundary condition into consideration. It should be notice that the concept of interior introduced in the plasma model is to avoid dealing with the interaction between dipoles explicitly. A modified form of plasma model explicitly specifying the interaction is presented as follows, which applies to the periodic single-layered structure with no interior.

Assuming a harmonic external field,  $\mathbf{E}_0(t) = \mathbf{E}_0 e^{-ik_z z} e^{-i\omega t}$ , the equation of motion of an oscillator at  $x_0$  could be written as

$$m\ddot{x} + m\gamma\dot{x} + m\omega_0^2(x - x_0) = qE_0 e^{-i\omega t} + F_{inter},$$

where  $F_{inter}$  is the interaction term between oscillators. Because in the case of steady state,  $F_{inter}$  always has the form  $F e^{-i\omega t}$ , the explicit mechanism of interaction inducing  $F_{inter}$  is not important. Then, the homogeneity of Maxwell equations permits us to include the interaction term into the third term on the left-hand side with the resonance frequency  $\omega_0$  being substituted by  $\omega' \equiv \sqrt{\omega_0^2 + g\omega^2}$ , where  $g \propto (q/m)\tilde{G}_{x_0}^\omega(x_0; x_0)$ , and where  $\tilde{G}_x^\omega$  is the Green's function in the frequency domain of the system with the oscillator at  $x$  eliminated, which charactering the explicit mechanism of the interaction between oscillators. This decouples the interaction between oscillators, and permits us to define the cross-section of an individual oscillator at a specific position  $x_0$  by  $\sigma_{x_0} = P_{x_0} / \Phi_{inj}$ , where  $P_{x_0}$  is the power absorbed by the oscillator and  $\Phi_{inj}$  is the incident field flux density. By writing the solution of the equation of motion as  $x(t) = x_1 e^{i\varphi} e^{-i\omega t} + x_0$  with  $x_1$  a positive real number, the work done by the external field, i.e., the power absorbed by the oscillator, could be derived directly. Explicitly writing the form of incident field flux density, and substituting the power absorbed by the oscillators and the incident

field flux, the cross-section can be written as  $\sigma_{x_0} = (x_1/nc)(q\omega \sin \varphi / \epsilon_0 E_0)$ , where  $n$  is the refractive index around the given dipole.

Assuming  $n_x c \gg \omega / 2\pi$ , where  $c$  and  $n_x$  are the speed of light and number density of oscillators in x-direction, the radiation damping could utilize directly the result for continuous current density. In this case, we could directly write the radiation damping as  $\gamma_{rad} = c\mu_0 n_x q^2 / (2m)$  by averaging the power of radiation of current density corresponding to the oscillator system over periods both in time and in the x-direction, where  $n_x$  is the number density of oscillators along x-direction.

### 1. The derivation of transmission and reflection fields

It should be noticed that,  $\omega'$  is unknown unless the equation of motion is solved. Alternatively, we assume  $\omega'$  as an independent variable, and solve for the phase retardation  $\varphi(\omega')$  of the radiated field with the incident field numerically from Maxwell equation to avoid to use boundary condition explicitly. Assuming the incident field propagation along  $-\hat{e}_z$  with polarization along x-axis, all photons share the same quantum state. Define  $\lambda_a \equiv n_a / n_i$  as the part of photons absorbed by the oscillation system, where  $n_a$  is the number of photons absorbed by the periodic oscillator system per second, and  $n_i$  is the number of photons of incident field impinging onto the system per second. The field intensities before and after the absorption process are  $E_0$  and  $\sqrt{1-\lambda_a} E_0$ , respectively. Define  $\lambda_e \equiv n_e / n_i$  as the ratio between the number of emitted photons  $n_e$  and the number of impinging photons. If the phase of the emitted photons coincides with the incident, the field intensity is simply  $\sqrt{1-\lambda_a + \lambda_e} E_0$ . Denoting the phase difference between emitted photons and incident photons as  $\varphi$ , and assuming the field intensity has the form  $\sqrt{1-\lambda_a} E_0 + E_{emit} e^{i\varphi}$ , the relation of conservation of energy could be simplified into

$$(1-\lambda_a) E_0^2 + E_{emit}^2 + 2\sqrt{1-\lambda_a} \left( \frac{e^{i\varphi} - e^{-i\varphi}}{2} \right) E_0 E_{emit} = (1-\lambda_a + \lambda_e) E_0^2.$$

Solving this equation and comparing with the limiting case of  $\varphi = 0$ , we could derive the field intensity

$$E = E_0 \left( \sqrt{1-\lambda_a} + \sqrt{(1-\lambda_a) \cos^2 \varphi + \lambda_e} e^{i\varphi} - \sqrt{1-\lambda_a} e^{i\varphi} \cos \varphi \right).$$

In deriving the transmission and reflection expression, we followed the procedure discussed above. The symmetry between the radiated field in the  $-\hat{e}_z$  direction and  $\hat{e}_z$  direction is used.

There is no incident field in the  $\hat{e}_z$  direction, the energy density of the derived field coincides with the expression in classical electrodynamics. The different treatment of transmission and reflection caused by the asymmetric incident field is the basis of near transparent approximation.

2. The derivation of equation 4 in the main text.

To obtain lowest loss in the oscillation system supporting weak oscillations, we could find that  $\varphi_0$  should be chosen in the neighbourhood of the origin, as is shown in Fig. 1 in the main text. Thus, we could assume  $N \approx \varphi_{max}/\varphi_0$ . By Jensen's inequality, the maximum of the total transmission coefficient is  $t_{tot} = \prod_{i=1}^N t_i = t_c^N$ . Taking the logarithm of both sides of this relation, we have

$$\log t_{tot} = N \log t_c$$

Substituting  $N \approx \varphi_{max}/\varphi_0$  into this expression, we have

$$\frac{\varphi_{max}}{\varphi_0} \log t_c = \log t_{tot}$$

Representing transmission coefficient in decibels, and noticing that  $\varphi_{max}$  is not related to the oscillation system, which is decided by structure designers according to their own demand in designing, we should change this equality to proportionality to reveal the nature of the weak oscillation. Then, we have

$$t_c \text{ (in decibel)}/\varphi_0 \propto t_{tot} \text{ (in decibel)}.$$

## S2. Simulation methods

There are three algorithms and two implementations used in the optimization process in deriving bias voltages, as is shown in Fig. S3 and the discussion in section 4. The first and second algorithm can be called gradual optimization algorithms. In the gradual optimization algorithms, we removed the resonance state by eliminating a small interval in the strong oscillation mode containing the resonance state. This will divide the range of external quantity into two or many disjoint parts. In the gradual optimization algorithms, we try to decide which connected component should be chosen for each layer of the structure. Then, we discretize the range of the external quantities eliminating the small interval containing the resonance states, into  $N_s$  different points, which is denoted as the sample set  $\{\varphi_s\}$ . The partial order of the sample set is defined by inequality  $\varphi_i \leq \varphi_j$ . By choosing all  $\varphi_i$  in the sample set different, we could assign a positive index to every  $\varphi_i$  denoted by  $n(\varphi_i)$ , and an index vector could be constructed.

The initial value of the gradual optimization algorithms is a randomly generated series. In the first optimization program, we generate  $N_R$  random initial index vectors and optimize them using the gradual optimization algorithm without loss, which will be discussed in detail in the first subsection in section 4. Then, the result of the gradual optimization algorithm without loss is further optimized by an adiabatic method, of which the results are denoted as rough optimization results in Fig. S3. We compare all the rough optimization results, and select the result having the largest value of the target function, as the final result of the rough optimization. This result is further optimized by the gradient descent algorithm. In this case, we use the whole range of external quantity as the domain of the gradient descent algorithm.

The static electric field distribution with bias voltage is calculated by the finite element method (FEM) simulations. The carrier concentration is assumed to be changed with the bias voltage, and the chemical potential is changed with the carrier concentration. Phase change of the single-layered structure is simulated by the FEM, and near transparent approximation is used to get the transforming results of the multiple-layered structure. The evolution of the transmitted field is performed by FEM simulation. In the simulation of the evolution of the field, the incident

field entered into the simulation region is derived by multiplying the transforming results of the multiple-layered structure with the monochromatic wave of the given polarization state.

### S3. Jones matrix expressions

The general theory of the N-layered oscillator system can be constructed by the transfer-matrix method. To manipulate the polarization state, the  $i$ th layer is rotated by an angle  $\theta_i$  along the z-axis for every positive  $i \leq N$ . If we define the amplitude vector as  $(E_x^+ E_y^+ E_x^- E_y^-)^T$ , the transfer-matrix for the  $i$ th layer could be written as

$$T_i = \begin{pmatrix} R(\theta_i) & \\ & R(\theta_i) \end{pmatrix} S_{2,3} \begin{pmatrix} T_i^x & \\ & T_i^y \end{pmatrix} S_{2,3}^{-1} \begin{pmatrix} R(\theta_i) & \\ & R(\theta_i) \end{pmatrix}^{-1},$$

where  $T_i^x$  and  $T_i^y$  are the transfer-matrix with polarization along and perpendicular to the direction of oscillation.  $R(\theta_i)$  is again the rotation matrix along z-axis.  $S_{m,n}$  is the permutation matrix exchanging the  $m$ th and the  $n$ th column. Then, the total transfer-matrix can be written as  $T = (T_{ij}) = T_1 P(d_1) T_2 \cdots P(d_{N-1}) T_N$ , where  $P$  is the propagation matrix between adjacent layers with the separation  $d_i$ . The transmission can be described by a generalized form of the Jones matrix

$$J_U = P(D)^{-1} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}^{-1} = U \begin{pmatrix} t_1 e^{-i\varphi_1} & t'_c \\ t_c & t_2 e^{-i\varphi_2} \end{pmatrix} U^*$$

The phase change of the one-layered structure can be decomposed into three processes, which are schematically shown in Fig. S1. If the distance from the light source to the structure and the distance from the structure to the detector does not change with time, we could isolate the phase change of the metasurface from the total phase change. We denote the polarization components along and perpendicular to the oscillator as  $|\lambda_{\parallel}\rangle$  and  $|\lambda_{\perp}\rangle$  in this section. For the component  $|\lambda_{\parallel}\rangle$ , the extra phase change is given by the interaction between the external field and the system. We assume that there is no collective excitation supported by the oscillator in the direction perpendicular to it. The phase change of the  $|\lambda_{\perp}\rangle$  in the metasurface is only influenced by the dielectric function of the material in the metasurface. We denote the relative phase change between  $|\lambda_{\parallel}\rangle$  and  $|\lambda_{\perp}\rangle$  as  $\delta\varphi = \varphi_{\parallel} - \varphi_{\perp}$ , and write down the Jones matrix of a layer of structure as

$$\begin{aligned} J_{\lambda} &= R(\theta_{\lambda}) \begin{pmatrix} e^{i\delta\varphi} & \\ & 1 \end{pmatrix} \begin{pmatrix} e^{i\varphi_{\perp}} & \\ & e^{i\varphi_{\parallel}} \end{pmatrix} R(-\theta_{\lambda}) \\ &= e^{i\varphi_{\perp}} R(\theta_{\lambda}) \begin{pmatrix} e^{i\delta\varphi} & \\ & 1 \end{pmatrix} R(-\theta_{\lambda}) \end{aligned},$$

where  $\theta_\lambda$  is the angle between the direction of the x-axis and the direction of the oscillation.  $R(\theta_\lambda)$  is the corresponding rotation matrix. Assuming  $e^{i\varphi_\perp}$  does not change with the external quantity, the Jones matrix can be further simplified to

$$J_\lambda(\Omega) = R(\theta_\lambda) \begin{pmatrix} e^{i\delta\varphi(\Omega)} & \\ & 1 \end{pmatrix} R(-\theta_\lambda),$$

5 which is consistent with the definition of the Jones matrix in other work.

The Jones matrix of multi-layered structures can be obtained by multiplying the individual Jones matrices

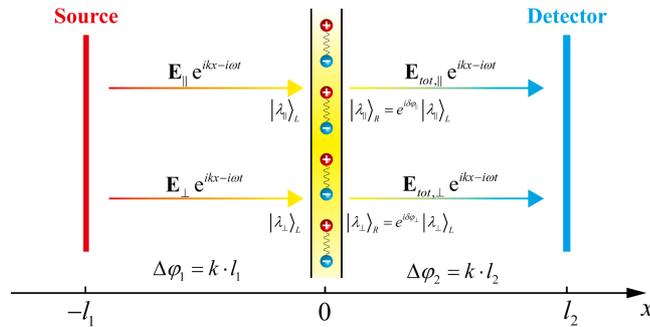
$$K(\{\theta_i\}, \{\Omega_i\}) = \prod_{\substack{\Omega_i \in \{\Omega_i\} \\ \theta_i \in \{\theta_i\}}} R(\theta_i) \begin{pmatrix} e^{i\delta\varphi(\Omega)} & \\ & 1 \end{pmatrix} R(-\theta_i),$$

10 where  $\theta_i$  and  $\Omega_i$  are the direction of oscillation and the external quantity of the  $i$ th layer. We label the bottom layer as the 1st layer, and the top layer as the  $N^{\text{th}}$  layer. Then, the multiplication runs from  $N$  to 1. This definition of the Jones matrix permits us to choose different directions of the oscillation in different layers. A simple case is to choose  $\{\theta_i\}$  such that all  $\theta_i$  have the same value. Without loss of generality, we assume  $\theta_i \equiv 0$  and the Jones matrix can be written as

$$K(\{\Omega_i\}) = \begin{pmatrix} e^{i\sum_{\Omega_i \in \{\Omega_i\}} \delta\varphi(\Omega)} & \\ & 1 \end{pmatrix},$$

15 which is the case to control the phase of the polarization component along the x-axis. Moreover, if the controlling of intensity is undesired, it is unnecessary to control all the phases independently. So we could simplify the Jones matrix as

$$K(\Omega) = \begin{pmatrix} e^{iN\delta\varphi(\Omega)} & \\ & 1 \end{pmatrix}.$$



20 **Fig. S1. Schematic diagram splitting the phase change of single-layered structure into three processes.**

Another simple case is to choose half of the layers along the y-axis and the other half of layers along the x-axis, which is denoted as the step type configuration in the main text. In this case, the Jones matrix can be written as

$$K(\Omega_x, \Omega_y) = \begin{pmatrix} e^{iN\delta\varphi(\Omega_x)/2} & \\ & e^{iN\delta\varphi(\Omega_y)/2} \end{pmatrix},$$

which is the case to control the x and y components independently. The last case is the linear type in the main text, where the angle varies linearly with the index

$$\theta_i = \delta\theta \times i.$$

In this case, the multiplication can be simplified as

$$K(\{\Omega_i\}) = R(\theta_{N+1}) \left[ \prod_{\Omega \in \{\Omega_i\}} R(\delta\theta) \begin{pmatrix} e^{i\delta\varphi(\Omega)} & \\ & 1 \end{pmatrix} \right] R(-\delta\theta).$$

The Jones matrix can be decomposed by solving the characteristic equation. Either the eigenvalues or the coefficients of the eigenvectors can be complex, which gives a higher flexibility. In the case that we include the loss of the structure into our theory, we could write the Jones matrix as

$$J_\lambda(\Omega) = R(\theta_\lambda) \begin{pmatrix} t_1(\Omega) e^{i\delta\varphi(\Omega)} & \\ & t_2(\Omega) \end{pmatrix} R(-\theta_\lambda),$$

where  $t_1$  and  $t_2$  are two real numbers denoting the ratio of the amplitude of the transmission field to the amplitude of the incident field. The Jones matrix for the linear type structure can be written as

$$K(\{\Omega_i\}) = R(\theta_{N+1}) \left[ \prod_{\Omega \in \{\Omega_i\}} R(\delta\theta) \begin{pmatrix} t_1(\Omega) e^{i\delta\varphi(\Omega)} & \\ & t_2(\Omega) \end{pmatrix} \right] R(-\delta\theta).$$

#### S4. Optimization algorithm for full Jones matrix control

In this section, we discuss the optimization algorithms used to derive the values of external quantities in detail assuming a linear type structure. A discussion of the pulling back projection of the Jones matrix is presented, which is used to construct the objective function of the algorithms.

There are two implementations of our algorithms to optimize the real Jones matrix type and the unitary Jones matrix type, separately. The real Jones matrix is the case where the eigenvectors form an orthogonal matrix, which is the case taken as the definition of the Jones matrix in many papers. The unitary Jones matrix is the case where the eigenvectors of the matrix form a unitary matrix, as we discussed in the main text. The only difference between these two implementations is the construction of the objective function, which will be discussed below.

For a specified  $\xi$ , the transformation  $\mathbf{K}$  is chosen. To produce the correct transformation result for this fixed  $\mathbf{K}$ , we may manipulate the incident field instead, in which has incorporated the information of the target Jones matrix  $\mathbf{J}(\theta, \varphi_1, \varphi_2)$  required in the optimization. As is shown in Fig. S2, this manipulated incident field could be taken as a correction and could be compared with the incident field for an ideal apparatus which is unnecessary to be corrected. By assuming that these two cases give the same transmitted field, we have

$$\mathbf{E}_{trans} = \mathbf{J}(\theta, \varphi_1, \varphi_2) \mathbf{E}_{inc}^J = \mathbf{K}(\{\Omega_i\}) \mathbf{E}_{inc}^K,$$

where  $\mathbf{J}(\theta, \varphi_1, \varphi_2)$  is the given target Jones matrix. Therefore, we could obtain  $\mathbf{E}_{inc}^J$  from this relation. Noticing that  $\mathbf{J}(\theta, \varphi_1, \varphi_2)$  is always invertible when  $t_1 \neq 0$  and  $t_2 \neq 0$ , we could write the projection of  $\mathbf{E}_{inc}^K$  onto  $\mathbf{E}_{inc}^J$  by the pulling back projection of the linear transformation

$$\langle \hat{\mathbf{E}}_{inc}^J | \hat{\mathbf{E}}_{inc}^K \rangle_{i,j} = \langle i | \mathbf{J}(\theta, \varphi_1, \varphi_2)^{-1} \mathbf{K}(\{\Omega_i\}) | j \rangle,$$

where  $|i\rangle$  and  $|j\rangle$  runs over two chosen linearly independent polarization components. If we choose the polarization components as  $|x\rangle$  and  $|y\rangle$ , we could use a simple form of pulling back projection directly

$$\mathbf{P} = \mathbf{J}(\theta, \varphi_1, \varphi_2)^{-1} \mathbf{K}(\{\Omega_i\}).$$

In the case of a unitary Jones matrix, we could substitute the diagonalized Jones matrix into the pulling back projection

$$\mathbf{P} = \mathbf{U} \begin{pmatrix} t_1 e^{-i\varphi_1} & \\ & t_2 e^{-i\varphi_2} \end{pmatrix} \mathbf{U}^{-1} \mathbf{K}(\{\Omega_i\}),$$

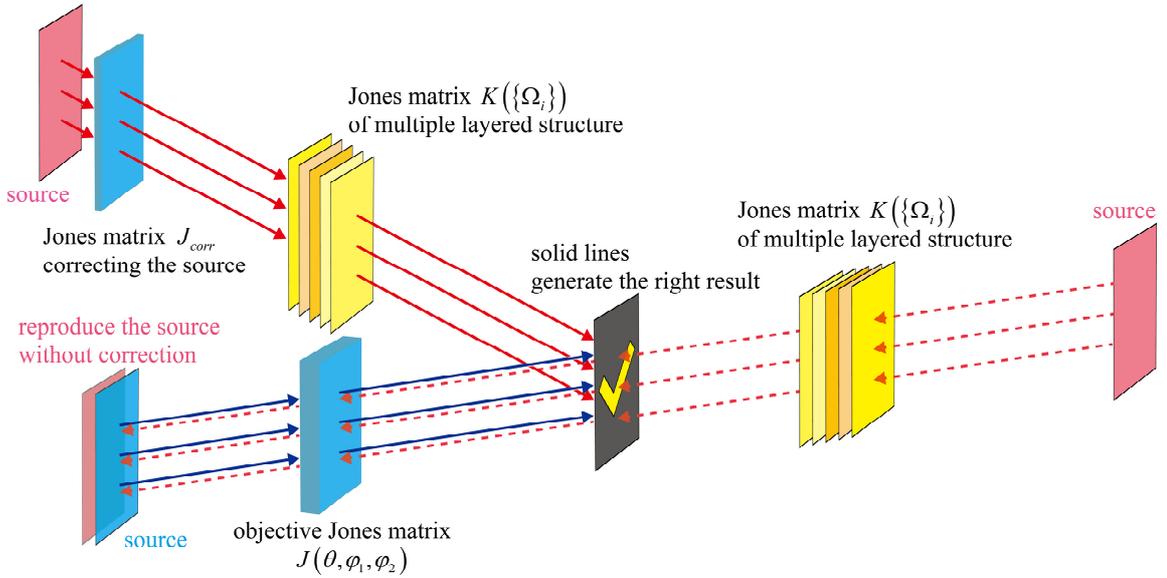
where  $\mathbf{U}$  is the unitary matrix defined in the main text.

The ideal case is that the pulling back projection is an identity matrix, which can be verified by substituting  $\mathbf{J}(\theta, \varphi_1, \varphi_2)$  into  $\mathbf{K}(\{\Omega_i\})$ . If we allow the metasurface to be dissipative, but the dissipation is independent of polarization, the ideal form of the pulling back projection is

$$\mathbf{P} = s\mathbf{I},$$

where  $s$  is a positive real number, which is the signal defined in the main text.  $\mathbf{I}$  is the identity matrix. The noise can be defined as

$$N = \left\| \mathbf{P} - \frac{1}{2} \text{Re}[\text{Tr}(\mathbf{P})] \mathbf{I} \right\|,$$

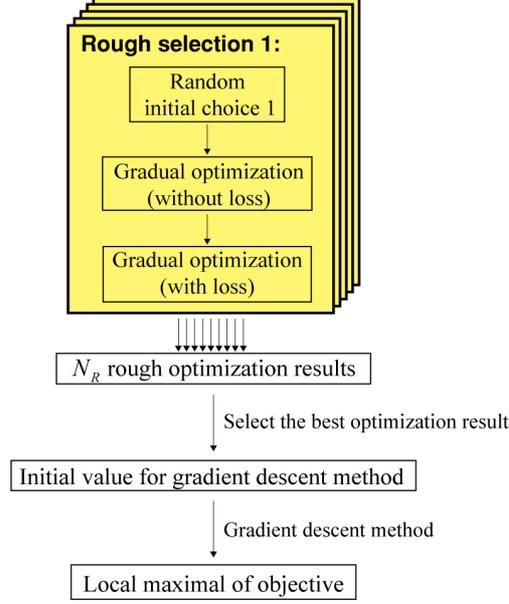


**Fig. S2. Physical significance of the pulling back projection of the Jones matrix.** The red solid lines correspond to the case where the source is properly corrected. The blue solid lines correspond to the ideal transformation of the source directly into the correct result. The dashed lines correspond to the transformation of the pulling back projection.

where  $\|\cdot\|$  is the norm of the matrix, and is defined as

$$\|M\| = \sum_i \sum_j |M_{i,j}|.$$

In the first two algorithms, we directly choose the signal as the optimization objective. In the fine adjustment process using the gradient descent method, we use  $s - \alpha N$  to emphasize the low noise character of the structure, where  $\alpha$  is a proportional coefficient.



**Fig. S3. Workflow of the optimization algorithms.**

### 1. Gradual optimization algorithm without loss

The loss in the Jones matrix will bring additional complexity into the optimization. In this algorithm, we neglect the loss in the Jones matrix. If we would like to optimize the  $j$ -th layer, the Jones matrix can be written as

$$K_j(\{\varphi_i\}, \beta) = M_1(j, \beta) R(\delta\theta_j) \begin{pmatrix} (t_1)^\beta e^{i\delta\varphi_j} & \\ & (t_2)^\beta \end{pmatrix} M_2(j, \beta),$$

where  $M_1(j)$  and  $M_2(j)$  are defined as

$$M_1(j, \beta) = R(\theta_{N+1}) \left[ \prod_{\varphi \in \{\varphi_i | i < j\}} R(\delta\theta) \begin{pmatrix} (t_1)^\beta e^{i\delta\varphi} & \\ & (t_2)^\beta \end{pmatrix} \right]$$

and

$$M_2(j, \beta) = \left[ \prod_{\varphi \in \{\varphi_i | i > j\}} R(\delta\theta) \begin{pmatrix} (t_1)^\beta e^{i\delta\varphi} & \\ & (t_2)^\beta \end{pmatrix} \right] R(-\delta\theta),$$

respectively. The coefficient  $\beta$  is used in the second algorithm to gradually include the loss into the optimization problem. In the current algorithm, to eliminate the influence of the loss, we directly set  $\beta = 0$ . The optimal  $\varphi_j$  can be derived by substituting every element of  $\{\varphi_s\}$  into the Jones matrix and compare the signal of every choice. The choice of the phase for other layers is given in  $\{\varphi_i\}$ . If the index of the optimal phase change  $n(\varphi'_j)$  is larger than that of the current phase change  $n(\varphi_j)$ , we will choose the new  $\varphi_j^{new}$  to satisfy the relation

$$n(\varphi_j) + 1 = n(\varphi_j^{new}).$$

If  $n(\varphi'_j) < n(\varphi_j)$  is satisfied, then the new  $\varphi_j^{new}$  should satisfy  $n(\varphi_j) - 1 = n(\varphi_j^{new})$ . When the equality  $n(\varphi'_j) = n(\varphi_j)$  holds, the phase change  $\varphi_j$  is already the optimal phase change at present, which means that we could directly set  $\varphi_j^{new} = \varphi_j$ .

In this algorithm, we start from the first layer and iteratively optimize every layer. The values of external quantities are repetitively optimized, until every layer has the optimal phase change. Because the index changes only one at most, every layer has much time to try many times that if this continuous branch contains an optimum choice. The number of elements in the sample set is not too large to sustain a slowly damping oscillation in the optimization process, therefore this algorithm converges quickly to a local maximum.

## 2. Adiabatic method to include the loss into the gradual optimization algorithm

The adiabatic method is a very common technique to solve a nonlinear problem. In the adiabatic method, the nonlinear problem reduces to a linear problem which is easy to solve. After the linear problem is solved, the nonlinear terms are added to the problem little by little, and the solution should be kept converging in this process, which is similar to the adiabatic process in physics. In our gradual optimization algorithm, we have already defined a coefficient  $\beta$  in the optimization to include the loss little by little into the algorithm. There are many routes varying from  $\beta=0$  to  $\beta=1$ . In the program deriving the result in this article,  $\beta$  is simply increased linearly to  $\beta=1$ . We start from the result derived in the gradual optimization algorithm without loss, and then increase  $\beta$  step by step. In each step, we optimize the external quantity using the same technique discussed in the gradual optimization algorithm without loss.

## 3. Gradient descent method

The choice of value of external quantity in the gradual optimization algorithms is in the discrete sample set proposed to eliminate the state near the resonance state and to speed up the optimization process. The choice of the small interval containing the resonance state is relatively arbitrary, and the discretization of the phase change is relatively random. Thus, the optimized result derived in the gradual optimization can be further optimized by considering all available phase changes. This is relatively simple and can be done by the gradient descent method.

By decreasing the signal to a smaller value in the optimization process, the intensity could be controlled by the external quantity. In this case, the noise could be defined as

$$N = \left\| \mathbf{P} - \begin{pmatrix} s'_1 & \\ & s'_2 \end{pmatrix} \right\|,$$

where  $s'_i$  are nonnegative numbers smaller than the optimum signal  $s$ .

## S5. Design configuration and feasibility of the structure

The *ab initio* calculations show that the energy gap of graphene nanoribbon decreases with the increasing of ribbon widths.[1, 2] In the case that the Fermi level is much larger than the energy gap of graphene nanoribbon, the optical conductivity of graphene could still be derived by the Kubo equation,[3, 4] which has already included the absorption of graphene as the real part of the optical conductivity. The temperature, wavelength, and scattering rate are set to 300K, 3.5  $\mu\text{m}$ , and  $2 \times 10^{-3} e_0/h$ , respectively. To support both weak and strong oscillation, the Fermi level of the upper layer is optimized to 0.17 eV.[5, 6] The Fermi level of the graphene layer is not easy to be controlled accurately in the fabrication process. A constant shift of bias voltage profiles may be needed to adjust the Fermi level of individual devices. The lower layer is highly doped and serves as the gate to adjust the Fermi energy of the upper layer. With the development of nanofabrication technology, especially the extreme ultraviolet (EUV) lithography,[7-10] we could design periodic structures with smaller artificial elements supporting SPPs. The period  $p$  of the structure is  $p=17.8$  nm. The width of the graphene strip is  $w=11.5$  nm, of which the half energy gap opened up,  $\Delta E_g/2$ , is around 0.034 eV in the case of zigzag graphene nanoribbons with hydrogen passivated zigzag edges. The separation between the upper and lower layers is  $h=21.3$  nm. The dielectric layer is chosen to be  $\text{CaF}_2$ . The thickness of the large scale and high-quality  $\text{CaF}_2$  could be accurately controlled by the molecular beam epitaxy (MBE).[11] According to the thickness-dependent theory of breakdown of Forlani and Minnaja,[12] the breakdown electric field for the  $\text{CaF}_2$  layer with a 21.3 nm thickness is approximately  $12.47 \times 10^6$  V/cm at room temperature,[13, 14] which well above the designed operation bias field. The side length of the pixel shown in Fig. 3 is 2  $\mu\text{m}$  and the number of layers is 43. The linear-type is used and  $\delta\theta = 6\pi/43$ . By controlling the bias voltages of each layer of each element, we could control the intensity, phase, and polarization state simultaneously. The set of all the bias voltages could be denoted as  $\xi = \{V_i^m\}$ , where  $V_i^m$  is the bias voltage of the  $m^{\text{th}}$  element in the  $i^{\text{th}}$  layer. A subsidiary problem of this structure is the influence of the circuit layout. The most straightforward case is to connect every piece of the periodic graphene structure by a wire separately. The precondition of a circuit responding to the monochromatic field of a given wavelength is that the spatial frequency of the structure satisfies the dispersion

relation to excite SPPs.[15] Thus, the principle to the design of the circuit layout is that the wire should not introduce any Fourier coefficient near the spatial frequency of SPPs.

## S6. Windowed Fourier transformation of a simple circuit

The surface plasmon polaritons supported by the two-dimensional material are characterized by the optical conductivity of the two-dimensional material. We could use the optical conductivity  $\sigma(\omega)$  of the specified materials to derive the dispersion relation by the same routinely process presented in other works. The dispersion relation can be written as

$$\frac{\varepsilon_1(\omega)}{k_{z_1}} + \frac{\varepsilon_2(\omega)}{k_{z_2}} + i \frac{\sigma(\omega)}{\omega \varepsilon_0} = 0,$$

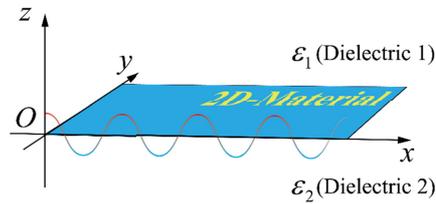
where  $\varepsilon_1(\omega)$  and  $\varepsilon_2(\omega)$  are the dielectric function of upper and lower materials encapsulating the two-dimensional material.  $k_{z_1}$  and  $k_{z_2}$  are the evanescent wave vector of SPPs in the ansatz is

$$\begin{aligned} E_j &= (E_{xj} \delta_{xj} + E_{zj} \delta_{zj}) e^{iqx} e^{-k_j|z|} \\ B_j &= B_{yj} \delta_{yj} e^{iqx} e^{-k_j|z|} \end{aligned},$$

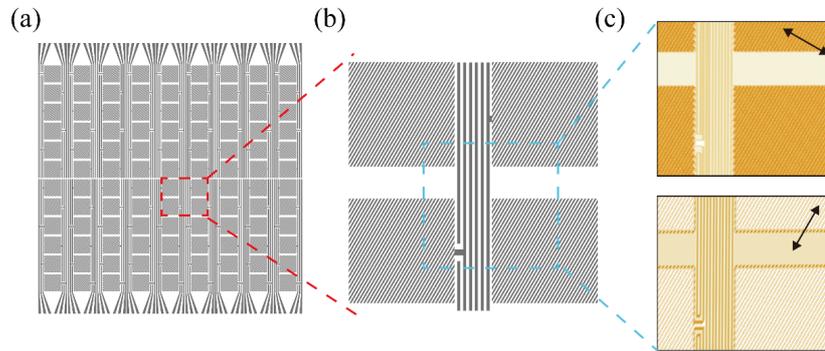
where  $\delta_{ij}$  is the Kronecker symbol.

The spatial frequency should satisfy the dispersion relation to excite SPPs. In the periodic graphene structure, the spatial frequency of the graphene slabs is fixed in the design process.

In the circuit layout shown in Fig. S5, the width of the wire and the separation between wires are equal to the period of the graphene structure. The window function of the windowed Fourier transformation is a Hamming window function. To show the wire clearly, the width of the wires in Fig. S5(a) is three times larger than the designed width. The area occupied by the wires is smaller than that it appears to be.



**Fig. S4. Surface plasmon polaritons supported by two-dimensional material encapsulated by dielectric 1 and dielectric 2.**



**Fig. S5. Influence of the circuit layout.** (a-b) A straightforward circuit layout design connected everything out. (c) Result of the windowed Fourier transformation of the structure. The black arrows show which polarization state of the windowed Fourier transformation is performed.

5

## Reference

1. Y.-W. Son, M. L. Cohen, and S. G. Louie, "Energy Gaps in Graphene Nanoribbons," *Physical Review Letters* **97**, 216803 (2006).
2. F. Jabbarzadeh, M. Heydari, and A. Habibzadeh-Sharif, "A comparative analysis of the accuracy of Kubo formulations for graphene plasmonics," *Materials Research Express* **6**, 086209 (2019).
3. L. A. Falkovsky, and A. A. Varlamov, "Space-time dispersion of graphene conductivity," *The European Physical Journal B* **56**, 281-284 (2007).
4. G. W. Hanson, "Dyadic Green's functions and guided surface waves for a surface conductivity model of graphene," *Journal of Applied Physics* **103**, 064302 (2008).
5. F. Schwabl, *Statistical mechanics* (Springer, 2002).
6. S. M. Sze, and K. K. Ng, *Physics of semiconductor devices* (John Wiley & sons, 2006).
7. Y. Jia, L. Guo, J. Lin, J. Yang, and X. Chen, "Direct growth of unidirectional spindle-shaped graphene nanoribbons on SiC," *Carbon* **114**, 585-590 (2017).
8. P. Gao, M. Pu, X. Ma, X. Li, Y. Guo, C. Wang, Z. Zhao, and X. Luo, "Plasmonic lithography for the fabrication of surface nanostructures with a feature size down to 9 nm," *Nanoscale* **12**, 2415-2421 (2020).
9. X. Wang, L.-T. Tseng, I. Mochi, M. Vockenhuber, L. van Lent-Protasova, R. Custers, G. Rispens, R. Hoefnagels, and Y. Ekinici, *Progress overview of EUV resists status towards high-NA EUV lithography* (SPIE, 2019).
10. J. Zhao, S. Yang, C. Xue, L. Wang, Z. Liang, L. Zhang, Y. Wang, Y. Wu, and R. Tai, "The recent development of soft x-ray interference lithography in SSRF," *International Journal of Extreme Manufacturing* **2**, 012005 (2020).
11. C. Wen, A. G. Banskchikov, Y. Y. Illarionov, W. Frammelsberger, T. Knobloch, F. Hui, N. S. Sokolov, T. Grasser, and M. Lanza, "Dielectric Properties of Ultrathin CaF<sub>2</sub> Ionic Crystals," *Advanced Materials* **32**, 2002525 (2020).
12. F. Forlani, and N. Minnaja, "Thickness Influence in Breakdown Phenomena of Thin Dielectric Films," *physica status solidi (b)* **4**, 311-324 (1964).
13. I. H. Malitson, "A Redetermination of Some Optical Properties of Calcium Fluoride," *Appl. Opt.* **2**, 1103-1107 (1963).
14. P. P. Budenstein, P. J. Hayes, J. L. Smith, and W. B. Smith, "Destructive Breakdown in Thin Films of SiO<sub>2</sub>, MgF<sub>2</sub>, CaF<sub>2</sub>, CeF<sub>3</sub>, CeO<sub>2</sub> and Teflon," *Journal of Vacuum Science and Technology* **6**, 289-303 (1969).
15. S. A. Maier, *Plasmonics: fundamentals and applications* (Springer Science & Business Media, 2007).

35