# Supplementary Information for Overcoming Barriers to Dynamic Phase-only Modulation in Transmissive Metasurfaces via Diffraction Control

Ju Young Kim†, Ruzan Sokhoyan†, Minkyoon Yi, Sangjun Han, Harry Atwater*, Min Seok Jang*

## 1 Temporal Coupled-Mode Theory

The model framework that we utilize to analyze and design the system is the temporal coupled-mode theory (TCMT) [1]. The validity of TCMT is described in detail in [2]. The dynamics of amplitude $a$ of a resonance mode (We will only deal with single-resonance systems in this paper) can be described by the following equations:

$$\frac{da}{dt} = -i\tilde{\omega}_0 a + \langle \kappa^* | s_+ \rangle \tag{S1}$$

$$|s_-\rangle = C|s_+\rangle + a|d\rangle \tag{S2}$$

where $\tilde{\omega}_0 = \omega_0 + i\gamma$ is the complex resonance frequency whose real ($\omega_0$) and imaginary ($\gamma$) parts are the center frequency and the overall decay rate of the resonance, respectively. Here $|a|^2$ is normalized to be the energy of the corresponding mode, $|s_+\rangle$ and $|s_-\rangle$ are the vectors of amplitudes of the incoming and outgoing waves, and their number of elements is equal to the number of input ($m_+$) and output ($m_-$) ports, respectively. $|d\rangle$ and $|d\rangle$ are $m_+$ and $m_-$ dimensional vectors describing the coupling from the input port to the resonance, and from the resonance to the output port, respectively. $C$ is the $m_- \times m_+$ direct scattering matrix from input to output ports. Note that there is freedom in choosing where the ports are placed from the resonator, which allows the complex parameters in the terms $|d\rangle$, $|\kappa\rangle$ and $C$ to be decided up to a factor $e^{ikl} = e^{i\theta}$, where $l$ is the port's distance from the resonator. This allows us to later adopt the convention used in the main paper, which is to place all the ports such that the coupling parameters $|d\rangle$ take on real values. This convention is possible for single-resonance systems, but is not guaranteed for systems with multiple resonances as choosing the set of port distances that gives real coupling parameters for one resonance does not generally make the coupling parameters real for the other resonances. Also, choosing the coupling parameters $|d\rangle$ to be real simultaneously affects the complex parameters of the direct scattering matrix $C$, and generally results in complex values for $C$.

For generality, we will start off with a four-port system and reduce to two-port systems when necessary. There can be many ways to make the system have four ports - by using different polarizations, nonzero incidence angles, or diffraction orders. We will assume to have two additional diffraction ports, but the mathematics for systems with more diffraction ports (such as the six-port system in the main paper) is similar. Imagine a four-port system as follows (Fig. S1):

*Notice that no assumptions on 'where' the ports 1, 2, 3, 4 are positioned are used in the following derivations, only the fact that port 3 = port 4 due to symmetry. The figure below is just for illustrative purposes. The analysis below is more general than what's shown in the figure, as the following formalism can also take into account systems with ports 3 and 4 on port 2's side, and systems with ports 1 and 2 'misaligned' or at an angle with one another, and many other configurations.*

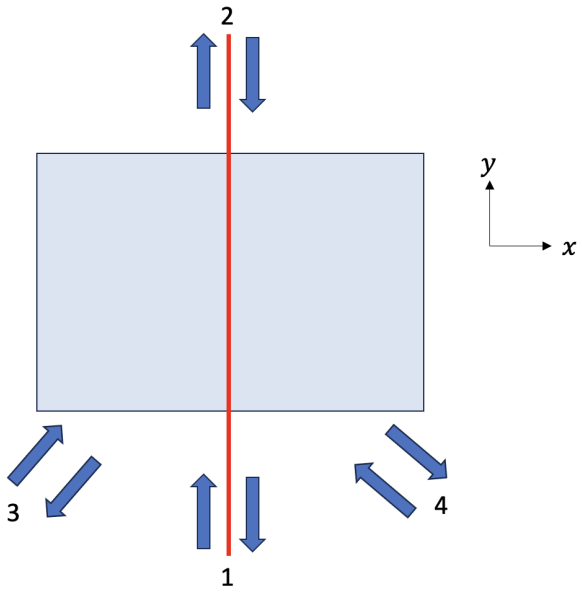


Figure S1: A four-port system with parity symmetry in the x-direction across the y-axis

Then along with Eq. S1 and S2 above, we have for the incoming and outgoing wave vectors and the coupling vectors:

$$|s_-\rangle = \begin{pmatrix} s_{1-} \\ s_{2-} \\ s_{3-} \\ s_{4-} \end{pmatrix} \quad |s_+\rangle = \begin{pmatrix} s_{1+} \\ s_{2+} \\ s_{3+} \\ s_{4+} \end{pmatrix} \quad |\kappa\rangle = \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_3 \\ \kappa_4 \end{pmatrix} \quad |d\rangle = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}$$

and the direct scattering matrix $C$ is given by:

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{pmatrix}$$

where $c_{ij}$ refers to the direct scattering term from port $i$ to port $j$.

Applying time reversal symmetry, energy conservation, and reciprocity [1] we get

$$\langle d|d\rangle = \frac{2}{\tau} \tag{S3}$$

$$|d\rangle = |\kappa\rangle \tag{S4}$$

$$C|d\rangle^* = -|d\rangle \tag{S5}$$

Also, assuming parity symmetry in the x-direction about the y-axis as shown in Fig. S1 above, and also assuming that the resonance field profile in question is also even about the y-axis (our mode belongs to this class), as well as applying time reversal symmetry, energy conservation, and reciprocity ([3]) we get the following scattering matrix $C$ :

$$C = \begin{pmatrix} r_1 & t & \alpha & \alpha \\ t & r_2 & \beta & \beta \\ \alpha & \beta & c & e \\ \alpha & \beta & e & c \end{pmatrix}$$

with the following norm equations:

$$|r_1|^2 + |t|^2 + 2|\alpha|^2 = 1 \tag{S6}$$

$$|t|^2 + |r_2|^2 + 2|\beta|^2 = 1 \tag{S7}$$

$$|\alpha|^2 + |\beta|^2 + |c|^2 + |e|^2 = 1 \tag{S8}$$

and the following orthogonality relations:

$$r_1 t^* + t r_2{}^* + 2\alpha\beta^* = 0 \tag{S9}$$

$$r_1\alpha^* + t\beta^* + \alpha c^* + \alpha e^* = 0 \tag{S10}$$

$$t\alpha^* + r_2\beta^* + \beta c^* + \beta e^* = 0 \tag{S11}$$

$$|\alpha|^2 + |\beta|^2 + ce^* + ec^* = 0 \tag{S12}$$

By assuming steady state excitation we have:

$$a = \frac{\langle \kappa^* | s_+ \rangle}{-i(\omega - \omega_0) + \gamma} \tag{S13}$$

where $\gamma = \frac{1}{\tau} = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_d$. Here $\gamma_i$ refers to the radiative coupling amplitudes (energy loss) for each port, and $\gamma_d$ refers to dissipative loss. The radiative coupling $d_i$, which is tied to the decay/gain in the field amplitude of the resonance (rather than the energy of the resonance, like the $\gamma_i$ terms) is related to the $\gamma_i$ terms through Eq.S1 above, and by turning 'on' only a single port while the other ports are turned 'off'. This gives rise to $d_i = \sqrt{2\gamma_i} e^{i\theta_i}$.

Now to get the relation between the coupling coefficients $d_i$ and the direct scattering coefficients of matrix $C$, we use Eq.S3 above:

$$\begin{pmatrix} r_1 & t & \alpha & \alpha \\ t & r_2 & \beta & \beta \\ \alpha & \beta & c & e \\ \alpha & \beta & e & c \end{pmatrix} \begin{pmatrix} d_1^* \\ d_2^* \\ d_3^* \\ d_4^* \end{pmatrix} = - \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}$$

Then we get:

$$r_1 d_1^* + t d_2^* + \alpha(d_3^* + d_4^*) = -d_1 \tag{S14}$$

$$t d_1^* + r_2 d_2^* + \beta(d_3^* + d_4^*) = -d_2 \tag{S15}$$

$$\alpha d_1^* + \beta d_2^* + c d_3^* + e d_4^* = -d_3 \tag{S16}$$

$$\alpha d_1^* + \beta d_2^* + e d_3^* + c d_4^* = -d_4 \tag{S17}$$

Although due to symmetry and the even field profile we have $d_3 = d_4$, we will still denote them as $d_3$ and $d_4$ separately.

Now applying $|s_-\rangle = C|s_+\rangle + a|d\rangle$:

$$\begin{pmatrix} s_{1-} \\ s_{2-} \\ s_{3-} \\ s_{4-} \end{pmatrix} = \begin{pmatrix} r_1 & t & \alpha & \alpha \\ t & r_2 & \beta & \beta \\ \alpha & \beta & c & e \\ \alpha & \beta & e & c \end{pmatrix} \begin{pmatrix} s_{1+} \\ s_{2+} \\ s_{3+} \\ s_{4+} \end{pmatrix} + a \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}$$

Because the incident light is only coming through port 1, $s_{2+} = s_{3+} = s_{4+} = 0$ and along with Eq.S13

$$\mathbf{r} = \frac{s_{1-}}{s_{1+}} = r_1 + \frac{i d_1^2}{(\omega - \omega_0) + i\gamma} = r_1 + \frac{2i\gamma_1 e^{2i\theta_1}}{(\omega - \omega_0) + i\gamma} \equiv \mathbf{A_r} + \mathbf{B_r} \tag{S18}$$

$$\mathbf{t} = \frac{s_{2-}}{s_{1+}} = t + \frac{i d_1 d_2}{(\omega - \omega_0) + i\gamma} = t + \frac{2i\sqrt{\gamma_1 \gamma_2} e^{i(\theta_1 + \theta_2)}}{(\omega - \omega_0) + i\gamma} \equiv \mathbf{A_t} + \mathbf{B_t} \tag{S19}$$

$$\mathbf{r_{10}} = \frac{s_{3-}}{s_{1+}} = \alpha + \frac{id_1d_3}{(\omega-\omega_0)+i\gamma} = \alpha + \frac{2i\sqrt{\gamma_1\gamma_3}e^{i(\theta_1+\theta_3)}}{(\omega-\omega_0)+i\gamma} \equiv \mathbf{A_{d1}} + \mathbf{B_{d1}}$$

Now we are in the position to examine how the $\mathbf{A_{t,r}}$ and $\mathbf{B_{t,r}}$ behave. Let's look at them one by one:

First, they were denoted in bold because on the complex plane, if you think of their real parts as the x-axis projections and their imaginary parts as the y-axis projections, we can think of them as 2-D vectors.

Second, the terms $\mathbf{A_{t,r}}$ are relatively constant as the changes in $\omega - \omega_0$ make a significant difference in the overall transmission and reflection curves, provided that the losses $\gamma_d$ in $\gamma$ don't render the 'widths' (FWHM in the case of a Lorentzian curve) of these curves to be too large (in which case the curves will likely be useless for phase modulation, which we shall soon see).

Third, the terms $\mathbf{B_{t,r}}$ draw a circle on the complex plane as the $\omega$ sweeps from 0 to $\infty$. We can see this as follows. Let's first isolate the common term from $\mathbf{B_{t,r}}$ and rewrite it by making the denominator real:

$$\frac{i}{(\omega-\omega_0)+i\gamma} = \frac{i}{\sqrt{(\omega-\omega_0)^2+\gamma^2}}e^{i\alpha}$$

where the term $\alpha$ was defined so that $\cos\alpha = \frac{\omega-\omega_0}{\sqrt{(\omega-\omega_0)^2+\gamma^2}}$ and $\sin\alpha = \frac{-\gamma}{\sqrt{(\omega-\omega_0)^2+\gamma^2}}$. Denoting $\omega - \omega_0 = \overline{\omega}$ for simplicity, we can see that as $\omega$ goes from 0 to $\infty$, $\overline{\omega}$ goes as roughly $-\infty \to \infty$ ($\omega_0$ is a very large number typically). Then from the definition of the $\cos\alpha$, it goes as $-1 \to 1$, and coupled with that fact that $\sin\alpha$ is negative due to $-\gamma < 0$ ($\alpha < 0$), we can see that $\alpha$ sweeps as $-\pi \to 0$ from the reference of $i$ (the imaginary axis). We can see this in Fig. S2 below.

Now, to see that the common term in $\mathbf{B_{t,r}}$ terms draw out a circle, we write it as:

$$\frac{-\sin\alpha}{\sqrt{(\omega-\omega_0)^2+\gamma^2}} + i\frac{\cos\alpha}{\sqrt{(\omega-\omega_0)^2+\gamma^2}} = \frac{\gamma}{(\omega-\omega_0)^2+\gamma^2} + i\frac{\omega-\omega_0}{(\omega-\omega_0)^2+\gamma^2} \equiv x + iy$$

Now, showing that $x$ and $y$ satisfy the equation of a circle:

$$x^2 + y^2 = \frac{\gamma^2}{((\omega-\omega_0)^2+\gamma^2)^2} + \frac{(\omega-\omega_0)^2}{((\omega-\omega_0)^2+\gamma^2)^2} = \frac{(\omega-\omega_0)^2+\gamma^2}{((\omega-\omega_0)^2+\gamma^2)^2} = \frac{1}{(\omega-\omega_0)^2+\gamma^2}$$

$$x^2 - \frac{1}{(\omega-\omega_0)^2+\gamma^2} + y^2 + \frac{1}{(2\gamma)^2} = \frac{1}{(2\gamma)^2}$$

$$x^2 - 2(\frac{\gamma}{(\omega-\omega_0)^2+\gamma^2})(\frac{1}{2\gamma}) + y^2 + \frac{1}{(2\gamma)^2} = \frac{1}{(2\gamma)^2}$$

Defining $R' = \frac{1}{2\gamma}$ we can see the equation becomes:

$$(x-R')^2 + y^2 = R'^2$$

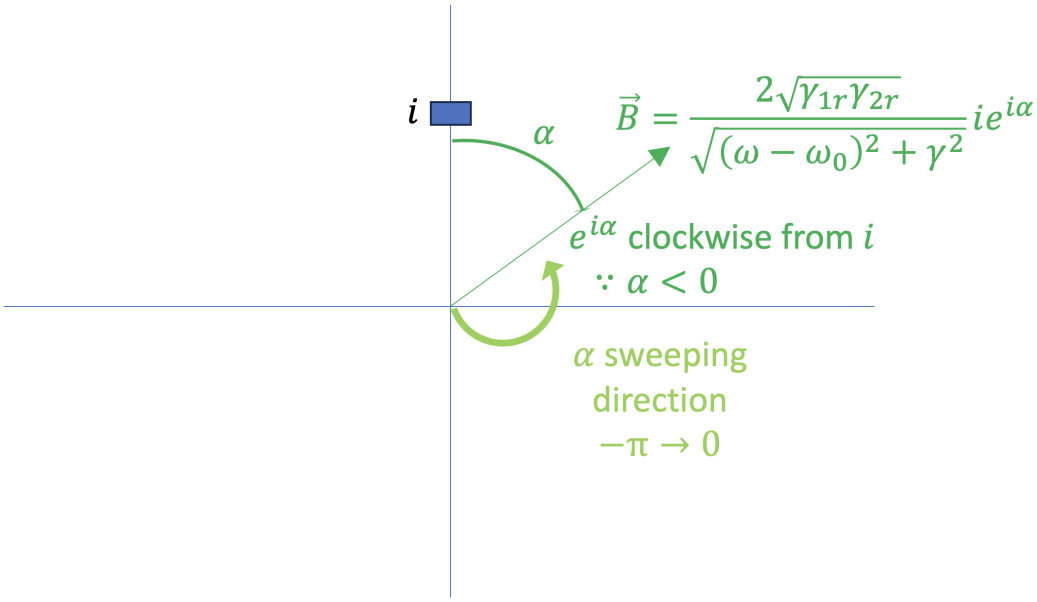


Figure S2: The sweeping of the B vector in the right side plane

This shows that the common term in $\mathbf{B_{t,r}}$ satisfies the equation of a circle defined on the complex plane, with the radius of $R'$ and the center $(R', 0)$. Then, with the corresponding factors on either $\mathbf{B_{t,r}}$, we can see that $\mathbf{B_{t,r}}$ draws out a circle like below in Fig. S3, with $\mathbf{B_t}$ having $R = \frac{\sqrt{\gamma_1\gamma_2}}{\gamma}$ and $\Theta = \theta_1 + \theta_2$, while $\mathbf{B_r}$ has $R = \frac{\gamma_1}{\gamma}$ and $\Theta = 2\theta_1$.

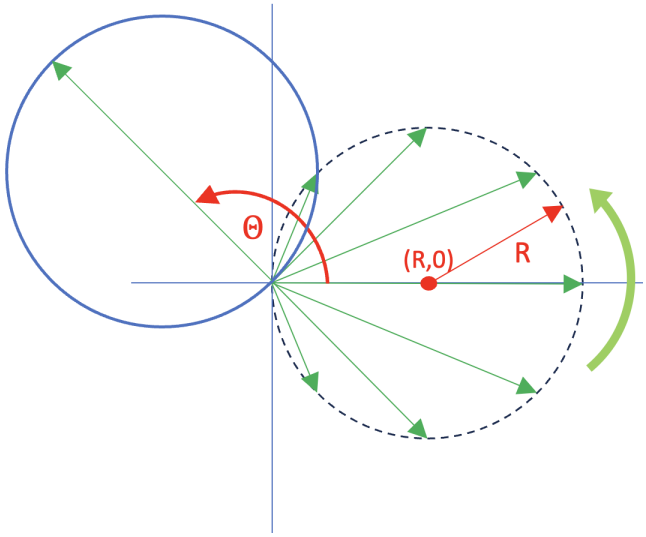


Figure S3: The $B_{t,r}$ vector drawing a circle on the complex plane

Now, before we see whether or not the transmission amplitude circle encircles around the origin on the complex plane, let's first rotate the complex transmission amplitude circle by $-\Theta = -(\theta_1 + \theta_2)$.

Then we have:

$$\mathbf{t}e^{-i(\theta_1+\theta_2)} = te^{-i(\theta_1+\theta_2)} + \frac{2i\sqrt{\gamma_1\gamma_2}}{(\omega-\omega_0)+i\gamma} \equiv \mathbf{A}+\mathbf{B} \tag{S20}$$

The reason for this is two-fold. One, we will now adopt the convention assumed in the main paper ($\theta_1 = \theta_2 = 0$, the real coupling condition), which, from the above, allows us to adopt the framework of displacing the term $\frac{2i\sqrt{\gamma_1\gamma_2}}{(\omega-\omega_0)+i\gamma}$ by the 'direct scattering parameter' $t_d = te^{-i(\theta_1+\theta_2)}$ in the main paper, and two, this illustrates the point made in the paper that changes in the radiative coupling values will naturally change $\theta_i$ and therefore will be reflected as 'compensated' changes in $\arg[t_{d,\mathbf{Im}[d_i]=0-convention}] = \arg[te^{-i(\theta_1+\theta_2)}]$.

Let's now adopt the real coupling condition where the $\theta_i = 0$, by placing the ports accordingly. Then we can observe that the circle is displaced by the $t$ vector, but its angular orientation is maintained, with the $t = \mathbf{A}$ vector being the circle's leftmost point and the rightmost point corresponding to the resonance frequency $\omega = \omega_0$ (the latter fact can be quickly checked by seeing that putting in $\omega = \omega_0$ in $\mathbf{B}$ gives $\mathbf{B} = \frac{2\sqrt{\gamma_1\gamma_2}}{\gamma}$, which is purely positively real). Fig.S4 nicely illustrates this.

Now, back to the matter of the encircling of the origin. Graphically, this can result in 3 scenarios: The transmission circle NOT encircling the origin (Fig. S4a), encircling the origin (Fig. S4b), and exactly going through the origin (Fig. S3). The mathematical determinant for the encirclement of the origin is given by $R^2 - |\mathbf{C}|^2$ and its sign, where $\mathbf{C}$ is the displacement vector of the center of the circle from the origin, and is given by $\mathbf{C} = \mathbf{A} + \frac{\sqrt{\gamma_1\gamma_2}}{\gamma} = (\frac{\sqrt{\gamma_1\gamma_2}}{\gamma} + \mathbf{Re}[t]) + i(\mathbf{Im}[t])$.

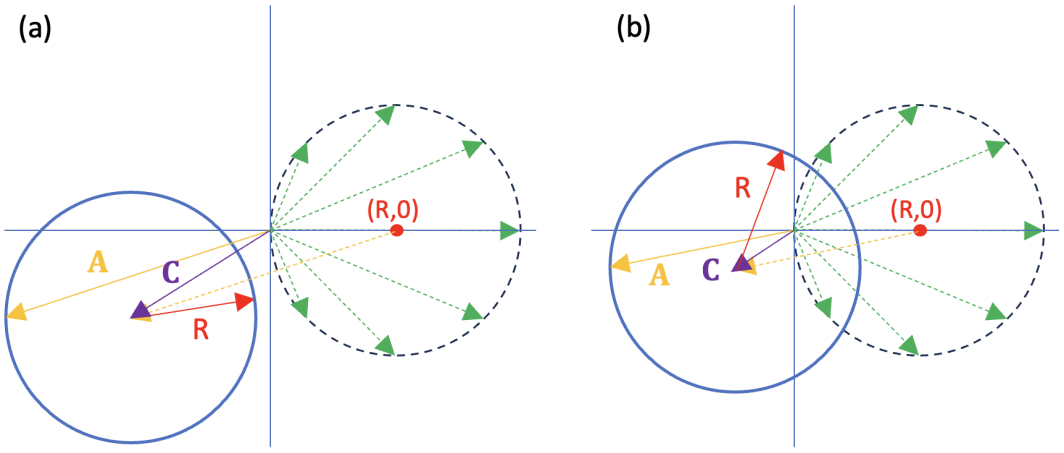


Figure S4: (a). The transmission circle not encircling the origin (b). The transmission circle encircling the origin

Now evaluating $R^2 - |\mathbf{C}|^2$ gives us

$$R^2 - |\mathbf{C}|^2 = (\frac{\sqrt{\gamma_1\gamma_2}}{\gamma})^2 - [(\frac{\sqrt{\gamma_1\gamma_2}}{\gamma} + \mathbf{Re}[t])^2 + (\mathbf{Im}[t])^2]$$

$$= (\frac{\sqrt{\gamma_1\gamma_2}}{\gamma})^2 - [(\frac{\sqrt{\gamma_1\gamma_2}}{\gamma})^2 + 2\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}\mathbf{Re}[t] + (\mathbf{Re}[t])^2 + (\mathbf{Im}[t])^2]$$

$$= -2\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}\mathbf{Re}[t] - |t|^2 \tag{S21}$$

Then rearranging Eq. S14 above and absolute-squaring it (now all the $d_i$ terms are purely real), we have:

$$|r_1 + \alpha\frac{d_3+d_4}{d_1}|^2 = |-t\frac{d_2}{d_1} - 1|^2 = |t\sqrt{\frac{\gamma_2}{\gamma_1}} + 1|^2$$

$$= 1 + \frac{\gamma_2}{\gamma_1}|t|^2 + 2\sqrt{\frac{\gamma_2}{\gamma_1}}\mathbf{Re}[t]$$

Then we have the following from above:

$$-2\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}\mathbf{Re}[t] = \frac{\gamma_1}{\gamma}(1 + \frac{\gamma_2}{\gamma_1}|t|^2 - |r_1 + \alpha\frac{\sqrt{2\gamma_3}+\sqrt{2\gamma_4}}{\sqrt{2\gamma_1}}|^2) \tag{S22}$$

Then inserting this into Eq. S21 above:

$$R^2 - |\mathbf{C}|^2 = \frac{\gamma_1}{\gamma}(1 + \frac{\gamma_2}{\gamma_1}|t|^2 - |r_1 + \alpha\frac{\sqrt{\gamma_3}+\sqrt{\gamma_4}}{\sqrt{\gamma_1}}|^2) - |t|^2$$

$$= \frac{\gamma_1}{\gamma}(1 - \frac{\gamma_1+\gamma_3+\gamma_4+\gamma_d}{\gamma_1}|t|^2 - |r_1 + \alpha\frac{\sqrt{\gamma_3}+\sqrt{\gamma_4}}{\sqrt{\gamma_1}}|^2) \tag{S23}$$

where we have used the fact that $\gamma = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_d$.

Now in a two-port system, we have $\gamma_3 = \gamma_4 = \alpha = 0$ and $|t|^2 + |r_1|^2 = 1$ so we get:

$$R^2 - |\mathbf{C}|^2 = \frac{\gamma_1}{\gamma}(1 - \frac{\gamma_1+\gamma_d}{\gamma_1}|t|^2 - |r_1|^2) = -\frac{\gamma_d}{\gamma}|t|^2 < 0 \tag{S24}$$

Equation S24 shows that in a two-port system, as long as there are dissipative losses $\gamma_d > 0$ you will always have a complex transmission circle that doesn't encircle the origin (shown in Fig. 1 of the main paper, with the best case having the circle go through the origin when $\gamma_d = 0$. This means for a two-port system, unless you add gain to the system ($\gamma_d < 0$), the phase variation around the resonance will always be limited as $< 180°$. Also, from Eq. S21 and from the fact that $R^2 - |\mathbf{C}|^2 = -2\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}\mathbf{Re}[t] - |t|^2 = 0$ when $\gamma_d = 0$, it follows that $\mathbf{Re}[t] < 0$

Now the idea behind using additional diffraction ports is this: Because $R^2 - |\mathbf{C}|^2 = 2\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}|\mathbf{Re}[t]| - |t|^2 = 0$ just before diffraction is added (two-port system), if we can decrease the magnitude of $|t|$ through the additional terms in the energy conservation equation $|r_1|^2 + |t|^2 + 2|\alpha|^2 = 1$ and engineer the complex valued $t$ so that $\mathbf{Re}[t] \approx t$, then we could have the second negative term $-|t|^2$ being smaller than the first positive term $2\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}|\mathbf{Re}[t]|$ and have $R^2 - |\mathbf{C}|^2 > 0$.

Since we show that adding diffraction channels can indeed make $R^2 - |\mathbf{C}|^2 > 0$ in Fig. 2 of the main paper, we will now outline our objectives for optimal phase-only modulation.

1. We want the largest possible complex transmission amplitude circle, aka. having the largest radius $R = \frac{\sqrt{\gamma_1\gamma_2}}{\gamma_1+\gamma_2+\gamma_3+\gamma_4+\gamma_d}$

2. We want the circle centered as close to the origin as possible, aka. having the smallest $|\mathbf{C}| = |(\frac{\sqrt{\gamma_1\gamma_2}}{\gamma} + \mathbf{Re}[t]) + i(\mathbf{Im}[t])|$

We will now show below the conditions for having the largest possible radius while having $|\mathbf{C}| = 0$, making it a perfect phase-only modulation scheme.

First, the largest $R$ is achieved when $\gamma_3 = \gamma_4 = \gamma_d = 0$ (restricting ourselves to the case of having no gain, $\gamma_d = 0$) and having $\gamma_1 = \gamma_2$, resulting in $R_{max} = 0.5$ (from the inequality $x + y \geq 2\sqrt{xy}$, with the equality holding when $x = y$). This makes sense intuitively, as we would expect the resonator spitting out equal number of photons per second into ports 1 and 2. While it is tempting to think that having $\gamma_2 > \gamma_1$ would be better, this would simply mean that not enough photons would come in through port 1 in the first place for port 2 to spit out (having a CAPACITY for port 2 to spit out 100 photons per second is meaningless if port 1 can ONLY take in 10 photons per second into the resonator).
Now, with that in mind, let's look at the condition of having $|\mathbf{C}| = 0$. This means having the following two conditions:

$$\mathbf{Re}[\mathbf{C}] = 0 \quad \rightarrow \quad \frac{\sqrt{\gamma_1\gamma_2}}{\gamma} = -\mathbf{Re}[t] \tag{S25}$$

$$\mathbf{Im}[\mathbf{C}] = 0 \quad \rightarrow \quad \mathbf{Im}[t] = 0 \tag{S26}$$

Because all the $\gamma's$ are defined to be positive, we can see from Eq. S25 that $\mathbf{Re}[t] < 0$, and from Eq. S26, we have the required condition that $t$ is negative and real. Secondly, because $\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}$ is the radius of the complex transmission amplitude circle, and our theoretical upper bound for the radius is 0.5, we can succinctly reduce the optimal conditions for transmissive phase-only modulation to be:

$$\arg[t] = \pi \tag{S27}$$

$$|t| = R = \frac{\sqrt{\gamma_1\gamma_2}}{\gamma} = 0.5 \tag{S28}$$

These are equations 2 and 3 in the main paper, and with these conditions the transmission circle starts at the baseline of $|t| = \frac{1}{2}$ from the negative real axis and draws a perfectly centered circle at the origin while maintaining the amplitude (radius) of 0.5.

## 2 Figure 2 Details

The figures S5, S6, S7 below show the spectral transmittance and phase line plots for three periods ($P_x = 1344$ nm, $P_x = 1401.2$ nm, and $P_x = 1751.5$ nm), as well as the spectral reflection/diffraction plots for the diffraction-regime periods $P_x = 1401.2$ nm and $P_x = 1751.5$ nm in Fig. S7 and S8, respectively. $R_{00}$ refers to normal reflection, and $R_{10}$ refers to first-order diffraction in the $x-$direction.

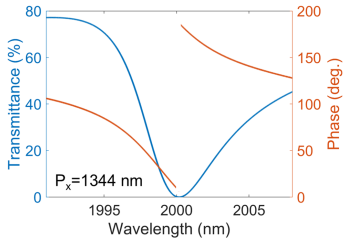


Figure S5: Transmittance and phase spectra for the metasurface depicted in Fig. 2 of the main manuscript. The period $P_x = 1344$ nm is smaller than the effective wavelength within the substrate (2000 nm/1.44 = 1388.9 nm). Hence, no diffraction is observed in reflection or transmission. Note that the phase exhibits discontinuity and spans 180 degrees.

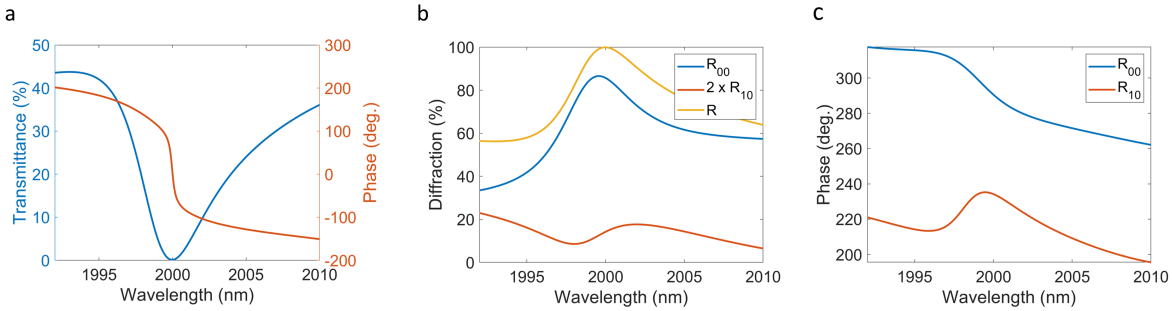


Figure S6: Transmittance and phase spectra for the metasurface with $P_x = 1401.2$ nm. The effective wavelength within the substrate is smaller than the metasurface period in the $x$-direction, and diffraction within the substrate is observed. (b) shows diffraction efficiencies in the normal direction $R_{00}$ and in the $x$-direction $R_{10}$. Note that we plot $2 \times R_{10}$ to account for $\pm 1^{\text{st}}$ diffraction orders. (c) plots the phases of each of the diffracted waves.

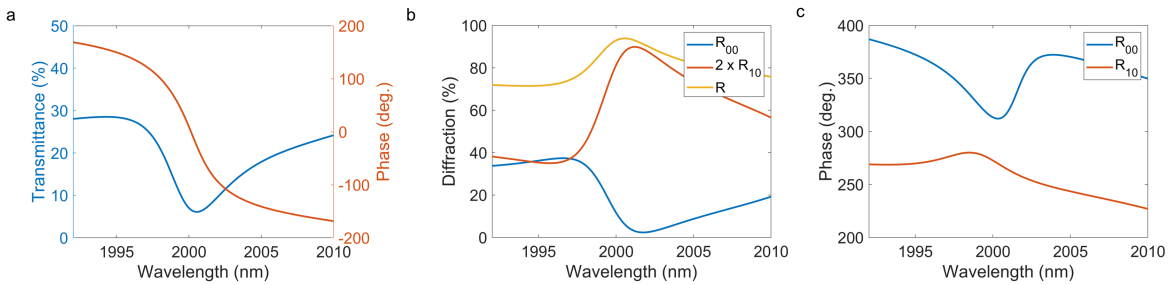


Figure S7: Transmittance and phase spectra for the metasurface with $P_x = 1751.5$ nm. The effective wavelength within the substrate is smaller than the metasurface period in the $x$-direction, and diffraction within the substrate is observed. (b) shows diffraction efficiencies in the normal direction $R_{00}$ and in the $x$-direction $R_{10}$. Note that we plot $2 \times R_{10}$ to account for $\pm 1^{\text{st}}$ diffraction orders. (c) plots the phases of each of the diffracted waves.

## 3 Zeros and poles on the complex frequency plane

Recent work [4] highlights the role of zeros and poles of the transmission amplitude in phase modulation. It illustrates that in order for the resonance to span a full $2\pi$ phase range, the zero should be above the real frequency line such that the branch cut between the zero and the pole forms an intersection with the real frequency line. We will now show that achieving the condition of the complex transmission amplitude circle encircling the origin ($R^2 - |\mathbf{C}|^2 > 0$) is mathematically equivalent to having the transmission zero move upwards into the upper complex plane from the real frequency axis. Equation 1 in the main paper can be rearranged as follows:

$$\mathbf{t} = t + \frac{2i\sqrt{\gamma_1\gamma_2}}{(\omega-\omega_0)+i\gamma} = \frac{t(\omega-\omega_0)+it\gamma+2i\sqrt{\gamma_1\gamma_2}}{(\omega-\omega_0)+i\gamma} = t\frac{\omega-(\omega_0-i(\gamma+\frac{2\sqrt{\gamma_1\gamma_2}}{t}))}{\omega-(\omega_0-i\gamma)} \equiv A\exp(iB\omega)\frac{\omega-\omega_z}{\omega-\omega_p}$$

where the Weierstrass expansion form is shown on the utmost right. From the above we can see that the direct transmission scattering parameter $t$ (complex-valued) corresponds to the term $A\exp(iB\omega)$, where the very slowly varying nature of $t$ with respect to the frequency $\omega$ (compared to the resonance term) allows us to assume it is constant in the vicinity of the resonance ($B\omega \approx B_0, \rightarrow t \equiv A\exp(iB_0)$). The location of the complex zero is given by $\omega_z \equiv \omega_0 - i(\gamma + \frac{2\sqrt{\gamma_1\gamma_2}}{t})$, and the location of the complex pole, similarly, is given by $\omega_p \equiv \omega_0 - i\gamma$.

If the transmission zero moves up into the positive imaginary half-plane in the complex frequency space, it means that the imaginary component of the complex zero becomes positive: $\mathbf{Im}[\omega_z] = -\mathbf{Re}[\gamma + \frac{2\sqrt{\gamma_1\gamma_2}}{t}] > 0$ (Notice that the **Re** part is necessary because the term $\gamma + \frac{2\sqrt{\gamma_1\gamma_2}}{t}$ itself is complex-valued due to the complex $t$). Then rearranging $-\mathbf{Re}[\gamma + \frac{2\sqrt{\gamma_1\gamma_2}}{t}] > 0$ gives:

$$\mathbf{Re}[\gamma + \frac{2\sqrt{\gamma_1\gamma_2}}{t}] = \mathbf{Re}[\frac{2t^*\sqrt{\gamma_1\gamma_2}}{|t|^2} + \gamma] = \frac{2\sqrt{\gamma_2\gamma_2}}{|t|^2}\mathbf{Re}[t^*] + \gamma = \frac{2\sqrt{\gamma_1\gamma_2}}{|t|^2}\mathbf{Re}[t] + \gamma < 0 \tag{S29}$$

where the relation $\mathbf{Re}[t^*] = \mathbf{Re}[t]$ was used. Then dividing Eq. S29 by $\gamma$ and multiplying by $-|t|^2$ gives us $-2\frac{\sqrt{\gamma_1\gamma_2}}{\gamma}\mathbf{Re}[t] - |t|^2 > 0$, which, from Eq. S21, is precisely the condition of $R^2 - |\mathbf{C}|^2 > 0$. Because the complex transmission pole is located below the real frequency axis due to $\mathbf{Im}[\omega_p] = -\gamma < 0$, the branch cut between the zero and the pole crosses the real frequency line, allowing for a full $2\pi$ phase change. It also naturally follows that having the condition of $R^2 - |\mathbf{C}|^2 = 0$ equates to having the transmission zero exactly on the real frequency axis, as in the case of two-port, single-resonance systems without loss or gain, and having the condition of $R^2 - |\mathbf{C}|^2 < 0$ indicates that the transmission zero is below the real frequency axis, failing to yield a full $2\pi$ phase change.

We will go further and qualitatively explain why the transmission zero cannot be on the real axis or exist as complex conjugate pairs, using the logical framework employed in the work [4].

The logic of the work was as follows. Suppose there is a transmissionless state for a two-port

system with its associated complex frequency, $\omega_{TZ}$. If there's time-reversal symmetry, you could apply the time-reversal operation on to this state $\omega_{TZ}$ and get $\omega_{TZ}^*$, which, in a two-port system, would also be a tranmissionless state. This means for a transmissionless state that either $\omega_{TZ} = \omega_{TZ}^*$ or $\omega_{TZ} \neq \omega_{TZ}^*$. Therefore, the transmissionless state complex frequencies $\omega_{TZ}$ exist either as real values or as complex conjugate pairs. If $\omega_{TZ}$ is a real value, there is zero transmission for that given real-frequency excitation, which prevents $2\pi$ phase range. If the transmissionless states exist as complex conjugate pairs, then one of the zero-pole pairs has its branch cut crossing the real frequency line, permitting $2\pi$ phase change. This spontaneous symmetry-broken state involves two resonances, the notable example being the Huygen metasurface involving multipoles of different symmetries satisfying the Kerker condition [5]. Additionally, there is the method of explicitly breaking the time-reversal symmetry to push the transmission zero away from the real axis. However, having intrinsic losses in the system pushes the zero downwards [6], which prevents the branch cut from crossing the real frequency line, so in the context of phase modulation, engineering the system to have a phase change of $2\pi$ is rather non-trivial unless sophisticated setups involving gain are employed.

Now let's apply the logical framework above to our system with additional diffraction ports. Suppose there is a transmissionless state $\omega_{TZ}$. Due to the additional diffraction channels, this state entails having an incident beam coming through the incident port, having no light go through the transmission port, and also have light radiate out through the diffration ports as well. Because there's no intrinsic system loss, we can apply the time-reversal operation. After time-reversal, the system would now have the incident and the diffraction ports having light converging back into the system/resonator (all with intensities less than unity), and light being emitted out through the incident port with unity intensity. Although technically the transmission port does not carry any incoming or outgoing light, it is difficult to call this a 'transmissionless' state because the system requires destructive interference of the beams coming from both the incident AND the diffraction ports at the transmission port. If we were to only consider the incident and the transmission port through projection onto the incident/transmissive port subspace and disregard the diffraction ports, energy conservation would be violated, which would prevent the use of time-reversal symmetry. Therefore, in our system, the transmission zeros do NOT have to exist either on the real axis or as complex conjugate pairs.

We support our claim above with complex frequency plots showing the locations of the zeros and poles for our resonance in Fig. 2 of the main paper. The results are shown in Fig S8 below. The complex frequency sweep was done with the S4 RCWA software. Fig. S8a shows the complex frequency plot for the structure in Fig. 2 without diffraction (having $P_x$ =1380 nm). As expected, the system is only a two-port system without diffraction ports, and therefore the transmission zero is indeed on the real frequency axis, with its exact location being $f_{\text{zero}} = 149.74 + 0i$ THz. This prevents the resonance from attaining a full $2\pi$ phase change. Fig. S8b shows the same system with a slightly larger period so that diffraction exists ($P_x$ =1400 nm, the threshold period for diffraction is $P_{x,\text{threshold}} = 1388.9$ nm). The zero is now slightly displaced above the real axis, with its location being $f_{\text{zero}} = 149.75 + 0.01i$ THz. Fig. S8c shows the same structure with a much larger period $P_x = 1650$ nm, so that the zero

is displaced even higher, with its location being $f_{\text{zero}} = 149.75 + 0.06i$ THz. As all three complex frequency plots show, there is only a single zero-pole pair, signifying that only a single resonance is being utilized. For the diffraction cases, the $\arg t$ plots show the branch cut between the zero and the pole crossing the real frequency line, allowing for a full $2\pi$ phase change.

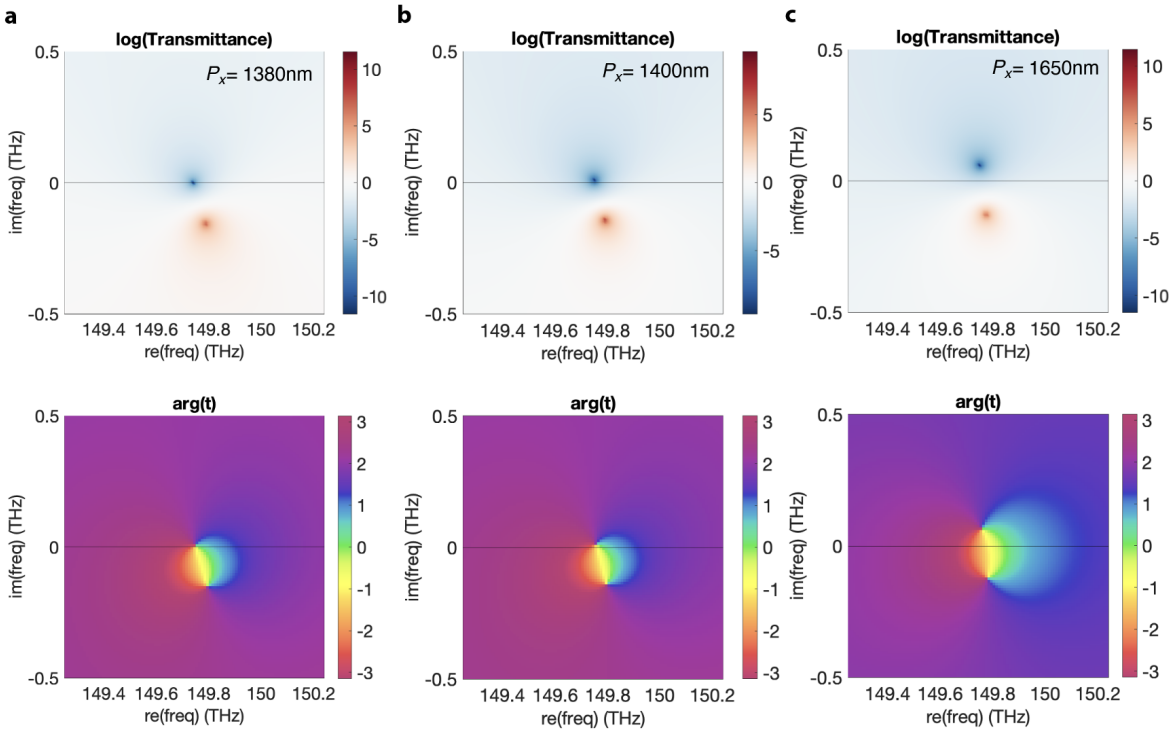


Figure S8: The complex frequency plot for $(\log |t|^2)$ and $\arg t$ for the structures (with differing $P_x$'s) shown in Fig. 2 of the main paper. The common geometric parameters for all three cases $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are: $P_y = 1328.6$ nm, Germanium ($n = 4$) pillar dimensions are 1167.7 nm (W) $\times$ 1167.7 nm (L) $\times$ 999 nm (H), and the substrate is silicon dioxide ($n = 1.44$). **a**. $P_x = 1380$ nm. **b**. $P_x = 1400$ nm **c**. $P_x = 1650$ nm.

# 4 Figure 3 Details

The resonance wavelengths $\lambda_0$ and the wavelength ranges used to plot Fig. 3 of the main manuscript were described in the caption of the figure. Here, we additionally the exact values for the complex direct transmission scattering parameters $t_d$, the radii $R$ of the complex transmission amplitude circles, and the quality factors $Q$, for Fig. 3b-3f:

| Figure | $t_d$ | $R$ | $Q$ |
|---|---|---|---|
| 3b. | $-0.6205 + 0.1513i$ | 0.329793 | 363.214 |
| 3c. | $-0.4561 + 0.1427i$ | 0.3352 | 375.949 |
| 3d. | $-0.5871 + 0.0608i$ | 0.370913 | 2943.99 |
| 3e. | $-0.4282 - 0.0154i$ | 0.448322 | 3254.32 |
| 3f. | $-0.4682 + 0.0205i$ | 0.455155 | 19864.6 |

Table S1: $t_d$, $R$, and $Q$ values for the circles in Fig. 3

The coloring scheme of the circles in Fig. 3 is as follows. For the circles in Fig. 3b and 3c, the points on the circles were plotted with each point having an angular frequency difference of ($\Delta\omega = 3.64147 \times 10^9$ rad/s), and for the circles in Fig. 3d, 3e, and 3f, and angular frequency difference of ($\Delta\omega = 3.74177 \times 10^9$rad/s). The magnitude of the color change in terms of its RGB values from one point on the circle to its nearest neighboring point was the same for all the circles in Fig. 3b to 3f. This means that if a resonance has a higher quality factor, fewer points would be required to sweep around the circle (due to the narrow spectral linewidth), and therefore visually the circle would have a more spread-out rainbow coloring around the resonance point (the dark red point with a black outline). As can be confirmed in Table S1 above, the quality factor keeps increasing as the circles progress from Fig. 3b to 3f, with an order of magnitude difference between 3b, 3c and 3d, 3e (the spreading-out of rainbow coloring), and another big jump in the order of magnitude from 3d, 3e to 3f (indicated by the 3f circle becoming discretized, as the consecutive points become too far in distance from one another).

Figures S9 to S13 below show the spectral transmittance and phase line plots for the complex transmission amplitude curve/circles in Fig. 3a to 3f of the main manuscript, as well as the spectral reflection ($R_{00}$) and diffraction ($R_{10}$ for $x-$directional first-order diffraction and $R_{01}$ for $y-$directional first-order diffraction) line plots and the corresponding phase plots for the diffraction-enabled cases of Fig. 3c to 3f.

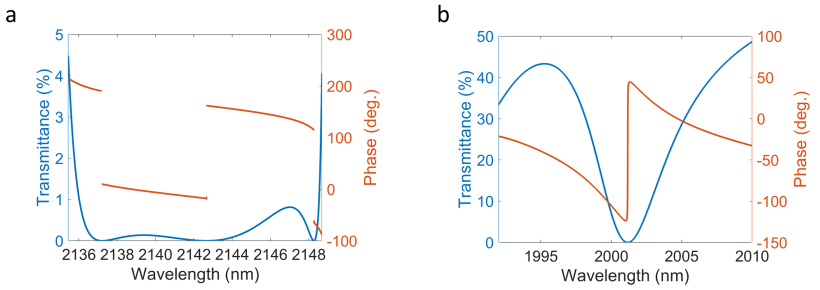


Figure S9: Phase and transmittance spectra corresponding to the complex amplitude plots shown in Fig. 3a and 3b of the manuscript. In (a), the transmission phase shows discontinuity. In (b), the phase extracted from full-wave simulations is continuous but spans from $-120°$ to $50°$, instead of the desired $0°$–$360°$.

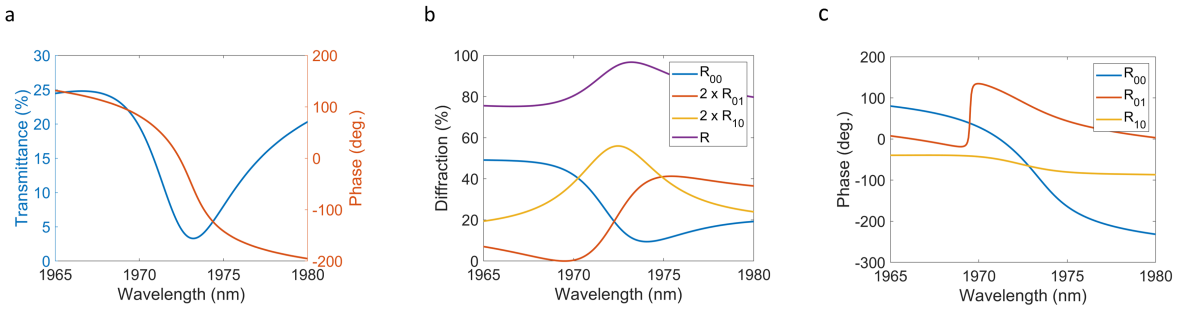


Figure S10: Phase and transmittance spectra corresponding to Fig. 3c.The effective wavelength within the substrate is now smaller than the metasurface period and the diffraction within the substrate is observed. (b) shows diffraction efficiencies in the normal direction $R_{00}$, $y$-direction $R_{01}$, and $x$-direction $R_{10}$. We plot $2 \times R_{01}$ and $2 \times R_{10}$ to account for $\pm 1^{\text{st}}$ diffraction orders. (c) plots the phases of the diffracted waves. $R$ stands for the total Reflectance.

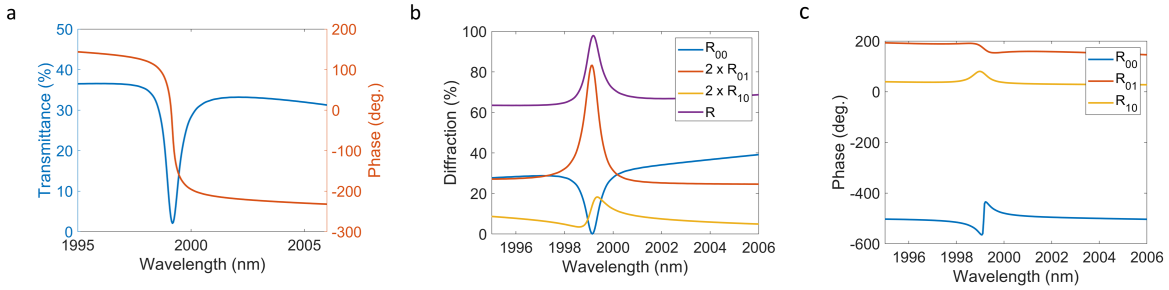


Figure S11: Same as above, corresponding to Fig. 3d.

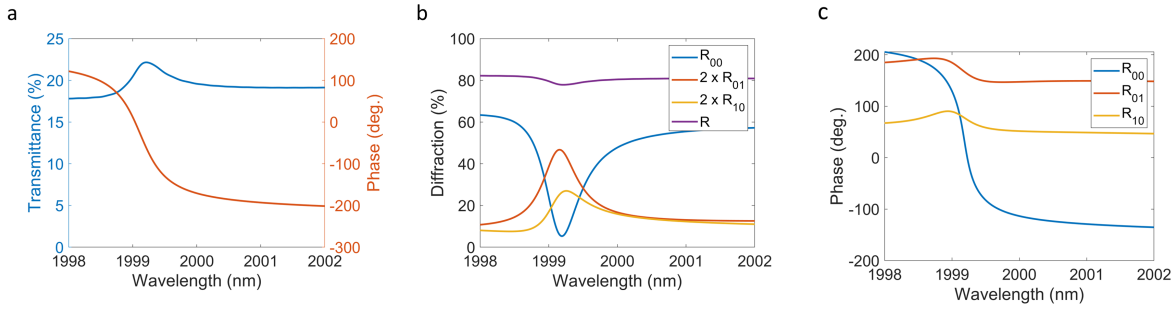


Figure S12: Same as above, corresponding to Fig. 3e.

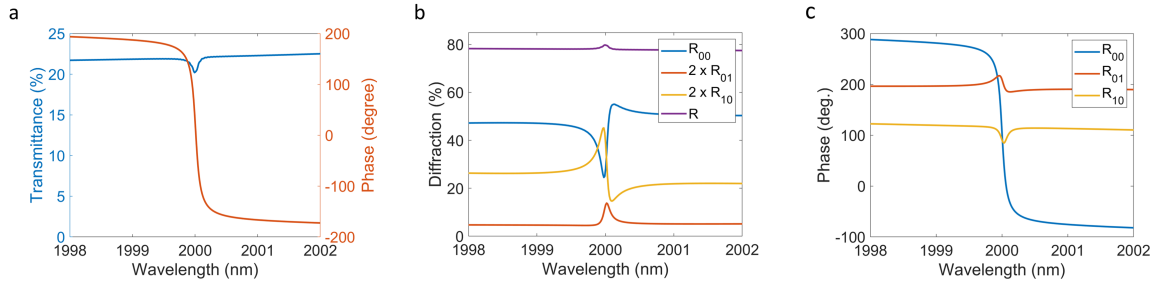


Figure S13: Same as above, corresponding to Fig. 3f.

# 5 Identifying Optimal Parameters

Here, we outline the process of identifying optimal parameters for centralized complex transmission amplitude circle with the largest radius. We consider our metasurface configuration excluding the ITO electrodes, not included in the simulation below. The assumed LN thickness is 339.6 nm, and the period in the $x$-direction is $P_x = 1959.2$ nm. The dimensions of the Ge pillar are $1175.5 \times 1175.5 \times 1005.7$ nm, where 1005.7 nm is the pillar height (along the $z$-direction). We simultaneously vary the width of the LN stripe and the period in the $y$-direction $P_y$. The wavelength range considered (1997–2001 nm) supports only one resonance. We plot the average transmittance ($(T_{\max} + T_{\min})/2$), transmittance variation ( $T_{\max} - T_{\min}$), and a figure of merit (FOM):

$$\text{FOM} = \frac{T_{\max} - T_{\min}}{[(T_{\max} + T_{\min})/2]^{2.5}}$$

as a function of LN width and $P_y$ below in Fig. S14. We choose geometric parameters that minimize the FOM. When minimized, the transmission spectra exhibit nearly constant transmission (phase-only control), while simultaneously optimizing for high transmittance (large circular radius for the complex transmission amplitude circle). As can be seen in Fig. S14c, there are multiple parameter sets that result in a high FOM, allowing us to select the one with the largest circular radius. Notice that the geometric parameters of the plot don't exactly match that of the parameters in Fig. 3 or Fig. 4 of the main manuscript. These plots are here to merely illustrate the process of optimal parameter selection.

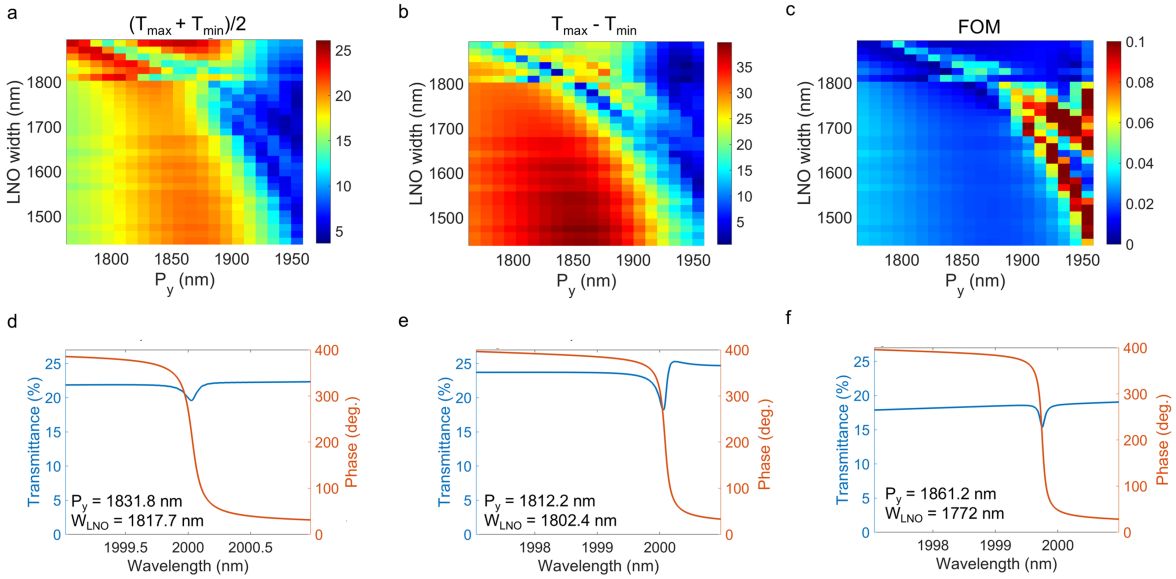


Figure S14: (a)–(c) plot average transmittance, transmittance variation, and the FOM as functions of LN width and $P_y$. In (a)–(c), the wavelength range is 1997–2001 nm. (d)–(f) show the transmission and phase spectra corresponding to the minimum FOM.

# 6 Metasurfaces with continuous (unetched) LN films

Figure S15 below shows spectral transmittance and phase line plots for a metasurface configuration with a continuous LN film underneath the Ge pillars. These curves were among the best of the phase-only modulation results with such an unetched configuration. This is to illustrate that although the continuous LN film configuration below performs decently well in terms of transmissive phase-only modulation, etching of the LN to create the pedestal structure is necessary to achieve a near-optimal (close to the theoretical upper bound) results.

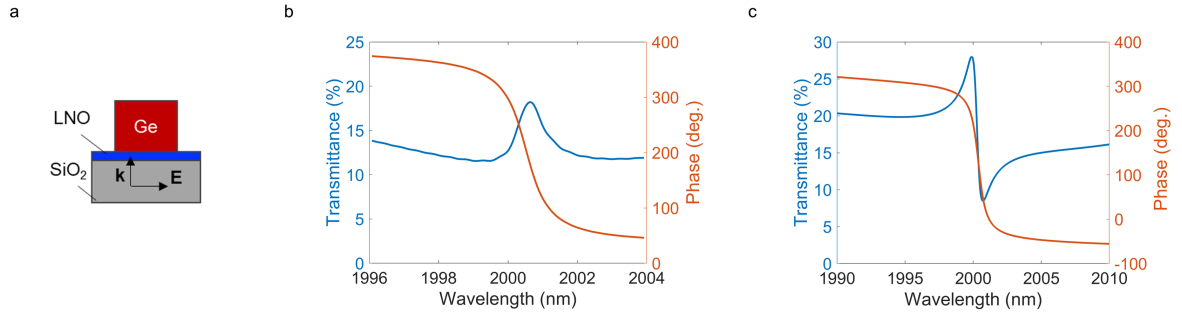


Figure S15: (a) Schematic of the metasurface unit cell with a continuous (unetched) LN film. (b) and (c) show transmittance and phase spectra for the following two geometries. For (b), the pillar height is 1005.9 nm, the pillar width and length are both 1175.8 nm, the LN thickness is 522.5 nm, $P_x$ = 1806.7 nm, and $P_y$ = 1858.8 nm. In c), the pillar height is 1009.2 nm, the pillar width and length are both 1179.5 nm, the LN thickness is 131 nm, $P_x$ = 1793.4 nm, and $P_y$ = 1914.2 nm. Simulations exclude electrodes, but we envision placing lateral transparent conducting electrodes on top of the LN films for practical realization.

# References


[1] Shanhui Fan, Wonjoo Suh, and John D Joannopoulos. Temporal coupled-mode theory for the fano resonance in optical resonators. *Journal of the Optical Society of America A*, 20(3):569–572, 2003.

[2] Tong Wu and Philippe Lalanne. Exact maxwell evolution equation of resonator dynamics: Temporal coupled-mode theory revisited. *Optics Express*, 32(12):20904–20914, 2024.

[3] Hermann A Haus. *Waves and fields in optoelectronics*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[4] Rémi Colom, Elena Mikheeva, Karim Achouri, Jesus Zuniga-Perez, Nicolas Bonod, Olivier JF Martin, Sven Burger, and Patrice Genevet. Crossing of the branch cut: The topological origin of a universal $2\pi$-phase retardation in non-hermitian metasurfaces. *Laser & Photonics Reviews*, 17(6): 2200976, 2023.

[5] Wei Liu and Yuri S Kivshar. Generalized kerker effects in nanophotonics and meta-optics. *Optics express*, 26(10):13085–13105, 2018.

[6] Yuhao Kang and Azriel Z Genack. Transmission zeros with topological symmetry in complex systems. *Physical Review B*, 103(10):L100201, 2021.