Scattering concentration bounds: brightness theorems for waves

HANWEN ZHANG,1,2 CHIA WEI HSU,1,3 AND OWEN D. MILLER1,2,*

1Department of Applied Physics, Yale University, New Haven, Connecticut 06511, USA
2Energy Sciences Institute, Yale University, New Haven, Connecticut 06511, USA
3Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, California 90089, USA
*Corresponding author: owen.miller@yale.edu

Received 23 May 2019; revised 30 August 2019; accepted 31 August 2019 (Doc. ID 368204); published 2 October 2019

The brightness theorem—brightness is nonincreasing in passive systems—is a foundational conservation law, with applications ranging from photovoltaics to displays, yet it is restricted to the field of ray optics. For general linear wave scattering, we show that power per scattering channel generalizes brightness, and we derive power-concentration bounds for systems of arbitrary coherence. The bounds motivate a concept of “wave étendue” as a measure of incoherence among the scattering-channel amplitudes and which is given by the rank of an appropriate density matrix. The bounds apply to nonreciprocal systems that are of increasing interest, and we demonstrate their applicability to maximal control in nanophotonics, for metasurfaces and waveguide junctions. Through inverse design, we discover metasurface elements operating near the theoretical limits.

The “brightness theorem” states that optical radiance cannot increase in passive ray-optical systems [1]. It is a consequence of a phase-space conservation law for optical étendue, which is a measure of the spatial and angular spread of a bundle of rays and has had a wide-ranging impact: it dictates the upper bounds to solar-energy concentration [2, 3] and fluorescent-photovoltaic efficiency [3], it is a critical design criterion for projectors and displays [4], and it undergirds the theory of nonimaging optics [5]. Yet a generalization to electromagnetic radiance is not possible, as coherent wave interference can yield dramatic enhancements. A natural question is whether Maxwell’s equations, and more general wave-scattering physics, exhibit related conservation laws.

In this paper, we develop analogous conservation laws for power flow through the scattering channels that comprise the bases of linear scattering matrices. By a density-matrix framework more familiar to quantum settings, we derive bounds on power concentration in scattering channels, determined by the coherence of the incident field. The ranks of the density matrices for the incoming and outgoing fields play the role of étendue, and maximal eigenvalues dictate maximum possible power concentration. For the specific case of a purely incoherent excitation of $N$ incoming channels, power cannot be concentrated onto fewer than $N$ outgoing channels, which in the ray-optical limit simplifies to the classical brightness theorem. In resonant systems described by temporal coupled-mode theory, the number of coupled resonant modes additionally restricts the flow of wave étendue through the system. The bounds require only passivity and apply to nonreciprocal systems. We discuss their ramifications in nanophotonics—for the design of metasurfaces, waveguide multiplexers, random-media transmission, and more—while noting that the bounds apply more generally to scattering in acoustic, quantum, and other wave systems.

Background: Optical rays exist in a four-dimensional phase space determined by their position and momentum values in a plane transverse to their propagation direction. Optical étendue [5] denotes the phase-space volume occupied by a ray bundle. In ideal optical systems, phase-space evolution is governed by Liouville’s theorem, and thus radiance and étendue are invariants of the propagation. A differential ray bundle propagating through area $dA$ and solid angle $dΩ$, in a medium of refractive index $n$ and tilted at an angle $θ$, has an étendue of $n^2 \cos θ dA dΩ$. Figure 1(a) depicts étendue conservation in ray-optical systems and the consequent trade-off between spatial ($dA$) and angular ($dΩ$) concentration. Electromagnetic radiance is intensity per unit area per unit solid angle, which in ray optics is proportional to the flux per unit étendue. By étendue invariance, in tandem with energy conservation, ray-optical brightness cannot increase. In nonideal systems, étendue can decrease when rays are reflected or absorbed, but any such reduction is accompanied by power loss, and the theorem still applies.

Extending radiometric concepts such as radiance into wave systems with coherence, beyond ray optics, has been the subject of considerable study [6–15]. Wigner functions can represent generalized phase-space distributions in such settings, and they are particularly useful for “first-order optics,” i.e., paraxial approximations, spherical waves, etc. Yet the Wigner function and similar approaches cannot simultaneously satisfy all necessary conditions.
The rank of the respective density matrix in Refs. [17–21]. As is well established in classical and quantum scattering theory [22–25], the operator $\mathcal{O}$ comprises two contributions: a “direct” (background) contribution from waves that travel from input to output without the scatterer present, which we denote with the operator $\mathcal{D}$, and a “scattered-field” contribution from waves that are scattered from input to output only in the presence of the scatterer, which we denote with the operator $\mathcal{T}$ (as in “$T$ matrix” approaches [26–29]). A key insight of Refs. [17–21] is that the $\mathcal{T}$ operator is compact (in fact, it is a Hilbert–Schmidt operator, by the integrability of the squared Frobenius norm of its kernel), which means that one can accurately represent it by a finite-dimensional singular-value decomposition

$$\mathcal{T} = \mathcal{U} \Sigma \mathcal{V}^\dagger,$$  

(1)

where $\mathcal{U}$ and $\mathcal{V}$ define orthonormal bases under an appropriate inner product $(\cdot, \cdot)$, and the restriction to finite dimensions is possible by retaining only those singular vectors corresponding to nonzero singular values, i.e., “well-coupled channels” [21]. The direct-process operator $\mathcal{D}$ is not necessarily compact—for example, $\mathcal{D}$ for scattering within a spherical domain is the identity operator [23,24,26]—and thus does not have the same natural decomposition. Nevertheless, we can still project the input and output states onto $\mathcal{V}$ and $\mathcal{U}$, respectively. Such a representation will necessarily miss an infinite number of input states with a non-trivial direct-process contribution, but by definition those states will have no interaction with the scatterer, and thus they have no consequence on power-concentration bounds or on the definition of a wave étendue. We include the direct process at all in order to naturally incorporate interference effects between the direct and scattering processes. Thus, for any scattering problem, the columns of $\mathcal{V}$ and $\mathcal{U}$ define our scattering channels, within which our input and output waves can be decomposed, as follows:

$$\psi_{\text{in}} = \mathcal{V} \psi_{\text{c}},$$  

(2a)

$$\psi_{\text{out}} = \mathcal{U} \psi_{\text{c}},$$  

(2b)

where $\psi_{\text{c}}$ and $\psi_{\text{c}}$ are the vector coefficients of the excitations on these channels as shown in Fig. 1(b). The scattering matrix connects $\psi_{\text{c}}$ to $\psi_{\text{c}}$ and can be found by starting with the definition of the $\mathcal{S}$ operator.

$$\psi_{\text{out}} = \mathcal{U} \psi_{\text{c}} = \mathcal{S} \psi_{\text{c}} = \mathcal{S} \psi_{\text{c}},$$

and then taking the inner product with $\mathcal{U}$ to find

$$\psi_{\text{out}} = \mathcal{U} \psi_{\text{c}} = \mathcal{S} \psi_{\text{c}} = \mathcal{S} \psi_{\text{c}} = \mathcal{S} \psi_{\text{c}}.$$  

(3)

We take our inner product to be a power normalization, such that $\psi_{\text{c}} \psi_{\text{c}}$ and $\psi_{\text{c}} \psi_{\text{c}}$ represent the total incoming and outgoing powers, respectively.

Perfectly coherent excitations allow for arbitrarily large modal concentration (e.g., through phase-conjugate optics [30,31]), but the introduction of incoherence incurs restrictions. To describe the coherence of incoming waves, we use a density matrix $\rho_{\text{in}}$ [32] that is the ensemble average (hereafter denoted by $(\cdot)$, over the source of incoherence) of the outer product of the incoming wave amplitudes, written as

$$\rho_{\text{in}} = \langle \psi_{\text{in}} \rho_{\text{in}} \psi_{\text{in}} \rangle.$$

(4)

The incoherence of the outgoing channels is represented in the corresponding outgoing-wave density matrix

$$\rho_{\text{out}} = \langle \psi_{\text{out}} \rho_{\text{out}} \psi_{\text{out}} \rangle = \mathcal{S} \rho_{\text{in}} \mathcal{S}^\dagger.$$

(5)

The matrices $\rho_{\text{in}}$ and $\rho_{\text{out}}$ represent density operators projected onto the $\mathcal{U}$ and $\mathcal{V}$ bases. Both matrices are Hermitian and positive semidefinite.

For inputs defined by some $\rho_{\text{in}}$, how much power can flow into a single output channel, or more generally into a linear combination given by a unit vector $\hat{u}$? If we denote $\hat{u} \rho_{\text{out}}$ as $\rho_{\text{out}, \hat{u}}$, then the power through $\hat{u}$ is $(\langle |c_{\text{out}, \hat{u}}|^2 \rangle) = \hat{u}^\dagger \rho_{\text{out}, \hat{u}} \hat{u} = \hat{u}^\dagger \mathcal{S} \rho_{\text{in}} \mathcal{S}^\dagger \hat{u}$. The quantity $(\langle |c_{\text{out}, \hat{u}}|^2 \rangle)$ is a quadratic form in $\rho_{\text{in}}$, such that its maximum value is dictated by its largest eigenvalue [33], $\lambda_{\text{max}}$, leading to the inequality

$$\langle |c_{\text{out}, \hat{u}}|^2 \rangle \leq \lambda_{\text{max}}(\rho_{\text{in}})(\hat{u}^\dagger \mathcal{S}^\dagger \mathcal{S} \hat{u}).$$

(6)

To bound the term in parentheses, $\hat{u}^\dagger \mathcal{S}^\dagger \mathcal{S} \hat{u}$, we consider coherent scattering for a new input: $c_{\text{in}} = \mathcal{S} \hat{u}$. For this input field, the incoming power is $\hat{u}^\dagger \mathcal{S}^\dagger \mathcal{S} \hat{u}$, while the outgoing power in the unit vector $\hat{u}$ is $| \langle c_{\text{out}, \hat{u}} \rangle |^2 = | \langle c_{\text{in}} \rangle |^2 = | \langle \mathcal{S} \hat{u} \rangle |^2$. Enforcing the inequality that the outgoing power in $\hat{u}$ must be no larger than the (coherent) total incoming power, we immediately have the identity $\hat{u}^\dagger \mathcal{S}^\dagger \mathcal{S} \hat{u} \leq 1$. (We provide an alternative proof in Supplement 1.) Inserting into Eq. (6), we arrive at the bound

$$\langle |c_{\text{out}, \hat{u}}|^2 \rangle \leq \lambda_{\text{max}}(\rho_{\text{in}}).$$

(7)
Equation (7) is a key theoretical result of this paper. It states that for a system whose incoming power flow and coherence are described by a density matrix $\rho_{in}$, the maximum concentration of power is the largest eigenvalue of that density matrix. For a coherent input (akin to quantum-mechanical “pure states” [34]), there is a single nonzero eigenvalue, equal to 1, such that all of the power can be concentrated into a single channel. For equal incoherent excitation of $N$ independent incoming states, the density matrix is diagonal with all nonzero eigenvalues equal to $1/N$, in which case
\[
\langle |c_{out,\hat{a}}|^2 \rangle \leq \frac{1}{N}.
\] (8)

Equation (8) is less general than Eq. (7) but provides intuition and is a closer generalization of the ray-optical brightness theorem. Since the average output power per independent state must be less than or equal to $1/N$, at least $N$ independent outgoing states must be excited, or a commensurate amount of power must be lost to dissipation. In reciprocal systems, this bound follows from reversibility. In Supplement 1, we prove that Eq. (8) simplifies to the ray-optical brightness theorem for continuous plane-wave channels in homogeneous media.

Just like the ray-optical brightness theorem [1,35], our scattering-channel bounds can alternatively be understood as a consequence of the second law of thermodynamics. If it were possible to concentrate incoherent excitations of multiple channels, then one could filter out all other channels and create a scenario in which a cold body on net sending energy to a warm body [20]. The partially coherent case is not as physically intuitive, but the application of such thermodynamic reasoning could be applied to the modes that diagonalize $\rho_{in}$, and then a basis transformation would yield Eq. (7).

Wave étendue: Equations (7) and (8) imply that the incoherent excitation of $N$ inputs cannot be fully concentrated to fewer than $N$ outputs. This motivates the identification of “wave étendue” as the number of incoherent excitations on any subset of channels (incoming, outgoing, etc.). For a density matrix $\rho$, one can count independence by the matrix rank and define étendue = rank($\rho$).

To understand the evolution of wave étendue through the scattering process, we reconsider the singular-value decomposition (SVD) of Eq. (1). The matrix $\Sigma$ is a square matrix with dimensions $N \times N$, where $N$ is the number of well-coupled pairs of input and output scattering channels. Since the singular values are nonzero, we know that $\Sigma$ is full rank. The scattering matrix $S$ is the sum of $\Sigma$ and the direct-process matrix, and its rank will be $N$ minus the number of coherent perfect absorber (CPA) states, $N - N_{CPA}$, where the CPA states arise if the direct process exactly cancels a scattered wave, yielding perfect absorption [36,37]. (Technically, these may be partial-CPA states, exhibiting perfect cancellation of the direct fields only on the range of the $T$ operator.) The density matrix $\rho_{in}$ is a representation of the incoming excitations on the basis $\mathcal{V}$ of Eq. (1), and thus it cannot have rank greater than $N$ itself. By the relation $\rho_{out} = S\rho_{in}S^\dagger$ and the matrix-product inequality $\text{rank}(AB) \leq \text{min}(\text{rank}(A), \text{rank}(B))$ (Ref. [33]), the rank of $\rho_{out}$ must lie within bounds given by the rank of $\rho_{in}$ minus the number of CPA states and the rank of $\rho_{in}$ itself as follows:
\[
\text{rank}(\rho_{in}) - N_{CPA} \leq \text{rank}(\rho_{out}) \leq \text{rank}(\rho_{in}).
\] (9)

Equation (9) defines the maximum diversity possible in the evolution of wave étendue in linear scattering systems. For lossless systems—or more generally any system without CPA states—we must have $N_{CPA} = 0$, in which case Eq. (9) is a conservation law stating that the density-matrix rank is always conserved. (In Supplement 1 we show that this simplifies to the classical wave-étendue conservation law in the ray-optics limit.) Figure 1(b) depicts this rank-defined (channel-counting) definition of wave étendue. In wave-scattering systems, phase space is defined by distinct scattering channels, without recourse to the position and momentum unique to free-space states.

**Metasurface designs:** To probe the channel-concentration bounds, we consider control of diffraction orders through complex metasurfaces for potential applications such as augmented-reality optics [38,39] and photovoltaic concentrators [40–42]. Figure 2(a) depicts a designable gradient refractive-index profile with a period of $2\lambda$ and a thickness of 0.5$\lambda$. (Such an element could be one unit cell within a larger, non-periodic...

---

**Fig. 2.** (a) Periodic metasurface element to be designed for maximal power in the +1 transmission diffraction order (yellow). We consider incoherent excitations among the four incident orders, with a diagonal density matrix, as well as partially coherent excitations between the 0 and -1 order, represented by an off-diagonal term with coherence parameter $c$. Inverse-designed metasurfaces closely approaching the coherence- and channel-dependent bounds are shown in (b) for incoherent excitations among up to four channels, and in (c) for partially coherent excitations between two channels. [Designs in (c) are all optimal for the fully incoherent case because $\rho_{in}$ is a constant multiple of the identity matrix. This should not be considered a generic phenomenon when excitation powers are unevenly distributed.]
metasurface [43–46].) For incoherent excitation of $N$ diffraction orders, Eq. (8) dictates that the maximum average efficiency of concentrating light into a single output order $\left(+1\right)$ cannot be greater than $1/N$ [dashed lines in Fig. 2(c)]. For $s$-polarized light incoherently incident from orders $0$ (red); $-1$, 0 (green); $-1$, 0, $+1$ (blue); and $-2$, $-1$, 0, $+1$ (purple) (20 deg angle of incidence for the zeroth order), we use adjoint-based “inverse design” [47–54] (Supplement 1) to discover optimal refractive-index profiles of the four metasurfaces shown in Fig. 2(b). (Broader angular control and binary refractive-index profiles could be generated through standard optimization augmentations [50,51], but here we emphasize the brightness-theorem consequences.) The transmission spectrum was computed by the Fourier modal method [55] with a freely available software package [56]. In Fig. 2(b), as the number of incoherent channels increased from 1 to 4, the average efficiency of the optimal structures decreases from $95.5\%$ to $24.9\%$. (In Supplement 1 we show that optimal blazed gratings fall far short of the bounds.) We also probe the effects of partial coherence by varying the coherence between two input orders, per the density matrix in Fig. 2(a). By Eq. (7), maximum concentration is determined by the largest eigenvalue of $\rho_{\text{int}}$, which is $1-c/2$, where $c$ is the coherence parameter. Figure 2(c) shows inverse-designed structures for $c=0.2, 0.4, 0.6, 0.8, 1$, with unique structures optimizing the response depending on the coherence of the excitation. All of the structures maximize efficiency in the incoherent $c=0$ case because the eigenvalues of the density matrix are degenerate, and thus transmission of any state is optimal.

Étendue transmission: An important related scenario to consider is one in which the direct (background) process is ignored, with the focus solely on interactions with scatterers. Instead of the input and output fields considered above, the relevant decomposition is instead into incident and scattered fields. Using the same terminology as in the input-output scattering operator, we can connect the incident and scattered fields by a $T$ operator [26–29] $\psi_{\text{scat}} = T \psi_{\text{inc}}$. Again, as shown in Refs. [17–21], the $T$ operator is compact and defines incoming- and scattered-wave bases by its SVD, Eq. (1). Furthermore, to align with various applications described below, we will specify a set of $N_{\text{trans}} \leq \text{rank } U$ desirable “transmission” channels that are a subset of the scattered-field channels defined by $U$. To understand transmission flow into these channels, we will define our finite-dimensional $T$ matrix as the restriction of $T$ onto this subset of scattered-field channels. The matrix $T$ connects the incoming-field channels to the transmission channels. The “transmission” terminology, partially meant to avoid further overload of the word “scattering,” is intended simply to represent the flow of energy through a system, enabled by interactions with a scatterer. For a planar or periodic scatterer, both reflected and transmitted waves would be part of this generalized “transmission” process, as long as they differ from the direct free-space process.

We define “étendue transmission” as the number of incoherent excitations that can successfully be transmitted through scatterer interactions onto the transmission channels. Equation (8) dictates that at least $N$ output channels are excited for $N$ orthogonal inputs, and indeed this result is proven in the incoherent case in Ref. [20] through an SVD of the $T$ matrix (denoted therein by “S”). If the number of transmission channels, $N_{\text{trans}}$, is less than $\text{rank}(\rho_{\text{inc}})$, where $\rho_{\text{inc}}$ is the incident-wave density matrix, then the incoherent excitations cannot all be concentrated onto the transmission channels, and some power must necessarily be scattered into other scattering channels.

Resonance-assisted transmission, in which resonances couple the incident and transmission channels, introduces an additional constraint: the number of resonant modes (resonances) $M$ coupled to the relevant channels. Resonant modes are not scattering channels; instead, they are the quasi-normal modes (QNM) of the scatterer, subject to outgoing boundary conditions. (Quasi-normal modes have been extensively studied and applied to various scattering systems for the last decade [57–59], and in the limit of closed systems and self-adjoint Maxwell operators they reduce to conventional guided and standing-wave modes [60].) We consider systems where the interaction with resonant modes can be described by temporal coupled mode theory (TCMT) [61–63], wherein the scattering process is encoded in an $M \times M$ matrix $\Omega$, comprising the real and imaginary parts of the resonant-mode resonant frequencies, and a matrix $K$, denoting channel–mode coupling. In TCMT, the $T$ matrix for the resonance-assisted transmission component is (Supplement 1) $\mathbf{T} = -i \mathbf{K}_{\text{trans}} (\Omega - \omega)^{-1} \mathbf{K}_{\text{inc}}^T$, where $\mathbf{K}_{\text{trans}}$ and $\mathbf{K}_{\text{inc}}$ are the $N_{\text{trans}} \times M$ and $N_{\text{inc}} \times M$ submatrices of $\mathbf{K}$ denoting modal couplings to the transmission and incident channels, respectively.

The maximum (average) power flow into a single transmission output channel is subject to the bounds of Eqs. (7) and (8), now in terms of the density matrix $\rho_{\text{inc}}$. The matrix $\rho_{\text{trans}}$ equals $T \rho_{\text{inc}} T^\dagger$. By recursive application of the matrix-rank inequality used above, we can see that

$$\text{rank}(\rho_{\text{trans}}) \leq \text{min}(\text{rank}(\rho_{\text{inc}}), M, N_{\text{trans}}).$$

(10)

The number of orthogonal outputs is less than or equal to the minimum of the numbers of incident inputs, resonant modes, and transmission channels. As depicted in Fig. 3, transmission channels and resonant modes act like apertures [35] in restricting the flow of étendue through a system.

We may also consider total transmission onto all $N_{\text{trans}}$ transmission channels, i.e., $\sum_i \langle |c_{\text{trans},i}|^2 \rangle$. Since the transmission onto a single output is bounded above by $\lambda_{\text{max}}(\rho_{\text{inc}})$, the total power is bounded above by the sum of the first $\text{rank}(\rho_{\text{inc}})$ eigenvalues (Supplement 1) as follows:

$$\sum_i \langle |c_{\text{trans},i}|^2 \rangle \leq \sum_{i=1}^{\text{min}(\text{rank}(\rho_{\text{inc}}), M, N_{\text{trans}})} \lambda_i,$$

(11)

where the eigenvalues are indexed in descending order. For incoherent excitation of the $N_{\text{inc}}$ channels, $\lambda_i(\rho_{\text{inc}}) = 1/N_{\text{inc}}$ for all $i$, and the term on the right of Eq. (11) simplifies to $\min(N_{\text{inc}}, M, N_{\text{trans}})/N_{\text{inc}}$. In resonance-assisted transmission

![Fig. 3. Étendue, defined as the rank of wave-scattering density matrices, is restricted in resonance-assisted transmission processes by the number of transmission channels and channel-coupled resonances in the process.](image-url)
Fig. 4. Robustness of waveguide junctions is susceptible to étendue restrictions. For two input channels, we consider (a) one output, (b) one mode, and (c) no restrictions. (c)–(g) Transmission for (a)–(c) with input phase angles in $\theta = [0, \pi/2]$. (d) Transmission as a function of phase, on resonance. Case (c) is designed to be almost perfectly insensitive to phase; such designs are impossible in cases (a) and (b).

Our work shows that fully coherent excitations are optimal for maximal concentration on the fewest possible states. A related question that remains open is whether partial coherence might be optimal for total transmission. The framework developed herein may lead to fundamental limits to control in such systems.

**Funding.** Air Force Office of Scientific Research (FA9550-17-1-0093).

**Acknowledgment.** We thank Erik Shipton and Scott McEldowney for helpful discussions regarding étendue in display systems. H. Z. and O. D. M. were supported by the Air Force Office of Scientific Research.

See Supplement 1 for supporting content.

**REFERENCES**

1. R. W. Boyd, *Radiometry and the Detection of Optical Radiation* (Wiley, 1983).
2. H. Ries, “Thermodynamic limitations of the concentration of electromagnetic radiation,” *J. Opt. Soc. Am.* 72, 380–385 (1982).
3. G. Smestad, H. Ries, R. Winston, and E. Yablonovitch, “The thermodynamic limits of light concentrators,” *Sol. Energy Mater.* 21, 99–111 (1990).
4. M. S. Brennesholtz and E. H. Stupp, *Projection Displays*, 2nd ed. (Wiley, 2008).
5. R. Winston, J. C. Miñano, and P. G. Benitez, *Nonimaging Optics* (Elsevier, 2005).
6. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University, 1995).
7. A. Walther, “Radiometry and coherence,” *J. Opt. Soc. Am.* 58, 1256–1259 (1968).
8. A. T. Friberg, “On the existence of a radiance function for finite planar sources of arbitrary states of coherence,” *J. Opt. Soc. Am.* 69, 192–198 (1979).
9. R. G. Littlejohn and R. Winston, “Corrections to classical radiometry,” *J. Opt. Soc. Am. A* 10, 2024–2037 (1993).
10. R. G. Littlejohn and R. Winston, “Generalized radiance and measurement,” *J. Opt. Soc. Am. A* 12, 2736–2743 (1995).
11. M. A. Alonso, “Radiometry and wide-angle wave fields. I. Coherent fields in two dimensions,” *J. Opt. Soc. Am. A* 18, 902–909 (2001).
12. M. A. Alonso, “Radiometry and wide-angle wave fields III: partial coherence,” *J. Opt. Soc. Am. A* 18, 2502–2511 (2001).
13. M. E. Testorf, B. M. Hennelly, and J. Ojeda-Castañeda, *Phase-space Optics: Fundamentals and Applications* (McGraw-Hill, 2010).
72. L. D. Tzuang, K. Fang, P. Nussenzveig, S. Fan, and M. Lipson, “Non-reciprocal phase shift induced by an effective magnetic flux for light,” Nat. Photonics 8, 701–705 (2014).
73. D. L. Sounas and A. Alù, “Non-reciprocal photonics based on time modulation,” Nat. Photonics 11, 774–783 (2017).
74. S. A. Cummer, J. Christensen, and A. Alù, “Controlling sound with acoustical metamaterials,” Nat. Rev. Mater. 1, 16001 (2016).
75. A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, “Controlling waves in space and time for imaging and focusing in complex media,” Nat. Photonics 6, 283–292 (2012).
76. C. W. Hsu, S. F. Liew, A. Goetschy, H. Cao, and A. D. Stone, “Correlation-enhanced control of wave focusing in disordered media,” Nat. Phys. 13, 497–502 (2017).
77. C. W. Hsu, A. Goetschy, Y. Bromberg, A. D. Stone, and H. Cao, “Broadband coherent enhancement of transmission and absorption in disordered media,” Phys. Rev. Lett. 115, 223901 (2015).
78. F. Bigourdan, R. Pierrat, and R. Carminati, “Enhanced absorption of waves in stealth hyperuniform disordered media,” Opt. Express 27, 8666–8682 (2019).
Scattering concentration bounds: brightness theorems for waves: supplementary material

Hanwen Zhang1,2, Chia Wei Hsu1,3, and Owen D. Miller1,2,*

1 Department of Applied Physics, Yale University, New Haven, Connecticut 06511, USA
2 Energy Sciences Institute, Yale University, New Haven, Connecticut 06511, USA
3 Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, California 90089, USA
*Corresponding author: owen.miller@yale.edu

Published 2 October 2019

This document provides supplementary information to "Scattering concentration bounds: brightness theorems for waves," https://doi.org/10.1364/OPTICA.6.001321.

CONTENTS

1 Deriving the classical brightness theorem from its wave-scattering generalization

2 Alternative proof for the brightness-concentration inequality

3 Coupled-mode theory for étendue transmission

4 Coupling constants of couple-mode models of waveguide combiners

5 Permittivity data of optimized metasurface unit cells and inverse design methodology

1. DERIVING THE CLASSICAL BRIGHTNESS THEOREM FROM ITS WAVE-SCATTERING GENERALIZATION

Here we show that the classical ray-optical brightness theorem follows from our wave-scattering generalization. Consider a ray-optical system with an entrance plane and an exit plane. Let us consider the power flow from within a differential area ΔA on the entrance plane through a differential area ΔA2 on the exit plane. In a wave-scattering framework, what are the power-normalized channels? And how many channels are there in a differential area ΔA?

In ray optics, the wavelength is taken to zero, such that even an infinitesimal area is arbitrarily large relative to the wavelength. Thus we can consider the differential area as the “box” (actually square) into which the states must fit with periodic boundary condition. If we take ΔA = ΔxΔy (with z as the propagation direction), the states that satisfy periodic boundary conditions on ΔA are plane waves with kx and ky taking integral multiples of 2π/Δx and 2π/Δy, and kz fixed by the frequency and specific values of kx and ky. Thus we can write non-normalized electric-field states as

\[ E_i = e^{jk_i \cdot r} \]  

where k_i is the corresponding wavevector for state i. In our manuscript, we choose for simplicity a channel definition such that the total power flowing through a given channel is 1. Since the intensity of a plane wave of amplitude E is |E|^2 / 2Z_0, where Z_0 is the impedance of the medium, we can choose our properly normalized channel states as

\[ E_i = \sqrt{\frac{2Z_0}{\Delta A}} e^{jk_i \cdot r} \]  

Taking H_i to be the magnetic field of channel i, one can then verify that the real part of (1/2) \( \int_{\Delta A} E_i \times H_j \) is indeed δ_ij, and we have a basis of power-orthonormal channels.

Now we can answer the question about the number of channels (states) within ΔA in the range from k_⊥ to k_⊥ + Δk_⊥ and k_z to k_z + Δk_z (which determine the propagating direction), so that k_1^2 + k_2^2 = k_⊥^2. A state occupies a region (2π/Δx)(2π/Δy) = 4π^2/ΔA in k_xk_y-space, then the total number of states is (including an extra factor of two for polarization)

\[ N = 2 \frac{\Delta k_\perp}{4\pi^2/\Delta A} = \frac{1}{2\pi^2} \Delta A k_\perp \Delta k_\perp \Delta \phi \]  

Since k_1^2 = k^2 sin θ^2, where θ is the angle between k and z axis, we have k_⊥ Δk_⊥ = k^2 cos θ sin θ Δθ. Substitute this relation into
Eq. (S3), we obtain

\[ N = \frac{1}{2\pi^2} k^2 \cos \theta \Delta \sin \theta \Delta \varphi = \frac{1}{2\pi^2} \frac{\omega^2}{c^2} \frac{\varepsilon^2 \cos \theta \Delta \Delta \Delta}{\Delta \varphi}. \]  

This function can be represented by its Fourier transform as

\[ f(x) = \sum_{n=-N}^{N} c_n e^{i k_n x}. \]  

Our claim is that the error between \( f(x) \) and \( f_N(x) \), measured in the \( L^2 \) norm, goes to zero in the ray optics limit \( (L/\lambda \to \infty, \) or \( N \to \infty) \), i.e.

\[ \epsilon_N = \| f(x) - f_N(x) \|_2 = \int_{-L/2}^{L/2} f(x) - f_N(x) \|^2 \to 0. \]  

**Proof:** Given Eq. (S7), it is straightforward to find the coefficients \( c_n \) that are the optimal Fourier-series coefficients of \( f_N(x) \) that most closely approximate \( f(x) \):

\[ c_n = \frac{1}{L} \int_{-L/2}^{L/2} f(x) e^{-i \frac{\pi}{2} n x} \, dx = \int_{-L/2}^{L/2} f(x) e^{-i k_n x} \]  

where \( k_n = \frac{2\pi n}{L} \). Then we can insert the Fourier-transform definition of \( f(x) \) into Eq. (S11) and integrate to find that the \( c_n \) are given by

\[ f_N(x) = \int dk_n \left\{ \sum_{n=-N}^{N} \sin \left( \frac{L x - k_n L}{2} \right) \right\} f(k_n). \]  

We can directly insert this expression for \( f_N(x) \) into the error expression, Eq. (S10), and carry out the integration over \( x \) to obtain

\[ \epsilon_N = L \int dk_n \int dk_n f(k_n) f^*(k_n) \left\{ \sin \left( \frac{L x - k_n L}{2} \right) \right\} \]  

with the help of the orthogonality of Fourier series,

\[ \int_{-L/2}^{L/2} e^{i (k_n - k_n') x} \, dx = \delta_{nn'} L. \]  

What remains to be shown is that the whole expression in the curly bracket goes to zero,

\[ \sin \left( \frac{L x - k_n L}{2} \right) \]  

\[ - \sum_{n=-N}^{N} \sin \left( \frac{L x - k_n L}{2} \right) \sin \left( \frac{L x - k_n L}{2} \right) \]  

\[ = \sin \left( \frac{L x - k_n L}{2} \right) \]  

\[ - \sum_{n=-N}^{N} \sin \left( \frac{L x - n \pi}{2} \right) \sin \left( \frac{L x - n \pi}{2} \right) \]  

\[ \to 0, \]  

as \( N \) goes to \( \infty \). This can be proven by invoking the Mittag-Leffler expansion of \( \cot(z) \),

\[ \cot(z) = \sum_{n=-\infty}^{\infty} \frac{1}{z - n \pi}, \]  

which can be found in complex-analysis textbooks [6].

Supplementary Material
In order to use the expansion, we rewrite the double sinc sum by partial fraction. Hence,

\[
\sum_{n=-N}^{N} \sin \left( \frac{k_x L}{2} - n\pi \right) \sin \left( \frac{k_y L}{2} - n\pi \right) = \sum_{n=-N}^{N} \sin \left( \frac{k_x L}{2} - n\pi \right) \sin \left( \frac{k_y L}{2} - n\pi \right) \frac{1}{k_x - k_y} \frac{1}{k_x \pm k_y - n\pi} = \sin \left( \frac{k_x L}{2} \right) \sin \left( \frac{k_y L}{2} \right) \frac{1}{k_x - k_y} \frac{1}{k_x \pm k_y - n\pi}
\]

\[
\sum_{n=-N}^{N} \sin \left( \frac{k_x L}{2} - n\pi \right) \sin \left( \frac{k_y L}{2} - n\pi \right) \frac{1}{k_x - k_y} \frac{1}{k_x \pm k_y - n\pi} = \sin \left( \frac{k_x L}{2} \right) \sin \left( \frac{k_y L}{2} \right) \frac{1}{k_x - k_y} \frac{1}{k_x \pm k_y - n\pi}
\]

as \( N \) goes to \( \infty \), which is the ray optics limit that \( \frac{L}{x} \) goes to \( \infty \). It is worth mentioning that the convergence is uniform, as convergence of Eq. (S16) is uniformly for all complex \( z \). Consequently, for any bounded \( f(k) \), we can bring the \( N \) limit into the \( k_x \) and \( k_y \) integral in Eq. (S14), proving that

\[
\lim_{N \to \infty} c_N = 0.
\]

To conclude, we have proven that the discrete plane wave basis in the ray-optics limit is capable of representing arbitrary wave forms. More abstractly, what we have proven is a well-known fact — that a Fourier series is able to represent any integrable function on a closed interval in the \( L^2 \) norm sense [5]. The \( L^2 \) norm is the physically meaningful one here, since observable power/momentum flows are defined by inner products of the appropriate wave/field functions.

2. ALTERNATIVE PROOF FOR THE BRIGHTNESS-CONCENTRATION INEQUALITY

Here we provide a more compact—but with less physical intuition and non-constructive—proof of the fact that \( u_i^\dagger SS u_i \leq 1 \). For any matrix \( A \), the matrices \( A^\dagger A \) and \( AA^\dagger \) have the same eigenvalues, as can be proven by inserting a singular valued decomposition of \( A \) into the expressions. By energy conservation, \( S \) is unitary, and the eigenvalues of \( SS^\dagger \) and therefore of \( SS^\dagger \), must be smaller than 1. By the variational principle (i.e. Rayleigh quotient [1]), we therefore must have \( u_i^\dagger SS^\dagger u_i \leq 1 \) for any unit vector \( u_i \).

3. COUPLED-MODE THEORY FOR ÉTENDUE TRANSMISSION

In this section we specify the step-by-step procedure to identify the rank of the transmission rank in coupled-mode theory, which plays the critical role in restricting étendue transmission through the system.

We start with standard coupled-mode theory: in a scattering system with \( N \) channels, we have \( N \times 1 \) incoming and outgoing coefficients \( c_i \) and \( c_{out} \), respectively, and an \( M \times 1 \) vector (for \( M \) resonances) of resonant-mode amplitudes \( a \). The three vectors are related by the coupled-mode equations:

\[
i(\Omega - \omega) a = K^T c_{in}
\]

\[
c_{out} = C_{in} + Ka,
\]

where \( a \) is a matrix whose Hermitian part is diagonal, comprising the real parts of the resonant frequencies, and whose anti-Hermitian part encodes dissipation via external coupling. The matrix \( K \) denotes coupling between the modes and the incoming/outgoing channels, while \( C \) denotes direct-scattering processes, independent of the resonances. Typically one can take \( C = -I \) (as the waveguide combiner cases in the main text), where \( I \) is the \( N \times N \) identity matrix, essentially as a normalization stating that in the absence of the scatterer, all energy flows back into the channel it came in on, with a negative amplitude.

Next one can solve for \( a \) to get \( c_{out} \) in terms of \( c_{in} \):

\[
c_{out} = [\frac{C}{K^T} (\Omega - \omega)^{-1]}^T] c_{in}.
\]

where the term in square brackets is the scattering matrix.

We are interested in the resonance assisted transmission from \( N_{inc} \) to an orthogonal subset of \( N_{trans} \) scattering channels. So we are invited to write Eq. (S22) as

\[
c_{out} - C_{in} = -iK^T (\Omega - \omega)^{-1}K c_{in}.
\]

Because there are only \( N_{inc} \) non-zero excitations, it is convenient to work with a \( N_{inc} \times 1 \) vector \( c_{in} \), a subset of \( c_{in} \). Similarly, as only transmissions from \( N_{trans} \) channels are collected, we can now work with a \( N_{trans} \times 1 \) vector \( c_{trans} \), a subset of \( c_{out} \). Accordingly, only a submatrix of \( K \) of size \( N_{inc} \times M \), denoted by \( K_{trans} \), contains the coupling information between \( c_{in} \) to resonance modes; another submatrix of size \( N_{trans} \times M \), denoted by \( K_{trans} \), connects the resonance modes with \( c_{trans} \). Following the above decompositions, for the \( T \)-matrix (or transmission matrix) from \( c_{inc} \) to \( c_{trans} \), namely

\[
c_{trans} = Tc_{inc},
\]

can be written as

\[
T = -iK_{trans}^T (\Omega - \omega)^{-1}K_{trans}.
\]

With the \( T \)-matrix definition of Eq. (S24), we can analyze resonance assisted power flow into the transmission channels due to incoherent excitations of the incident channels with the same matrix-trace approach used for \( c_{inc} \) and \( c_{trans} \) and the scattering matrix in the main text. As a result, the average power in transmission channel \( i \) is given by

\[
\langle |c_{trans,i}|^2 \rangle = u_i^\dagger T_{inc}^T u_i.
\]

Let us consider the total average power (equal to the average total power) transmitted onto all transmission channels, which
We can still do better by bounding the rank of $T$, where eigenvalues are placed in decreasing order with $v_i$ (we label those eigenvalues in decreasing order, i.e. $\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots$). The summation over $i$ includes all necessary channels so that $\sum_i u_i^T u_i$ and $\sum_i v_i v_i^T$ are both identity. However, the rank of $T$ restricts the meaningful inclusion of $v_i$ — the number of $v_i$ such that $Tv_i \neq 0$ is at most rank $(T)$ as a consequence of the rank-nullity theorem[2]. Besides, since there are at most rank($\rho_{inc}$) number of $v_i$s are relevant, we can thus bound Eq. (S31) by

$$\sum_i |\langle k_{\text{trans},i} \rangle|^2 \leq \sum_{i=1}^{U} \lambda_i v_i^T T^i T v_i,$$  \hspace{1cm} (S32)

where $U = \min(\text{rank}(\rho_{inc}), \text{rank}(T))$.

In the main text, we used a coherent-scattering example to show that for each individual $i$, $u_i^T S S^T u_i \leq 1$. Exactly the same proof, but with the inc/trans channels replacing the in/out channels, leads to the same inequality for $T$, i.e. $u_i^T T^i T u_i \leq 1$ and $v_i^T T^i T v_i \leq 1$. Thus all of the eigenvalues of $TT^i$ and $T^i T$ thus both less than or equal to one (and they are greater than or equal to zero because $TT^i$ and $T^i T$ are positive semidefinite). This reduce Eq. (S31) into

$$\sum_i |\langle k_{\text{trans},i} \rangle|^2 \leq \sum_{i=1}^{U} \lambda_i.$$  \hspace{1cm} (S33)

We can still do better by bounding the rank of $T$ using the fact that rank($AB$) $\leq \min(\text{rank}(A), \text{rank}(B))$. We write again the coupled-mode expression for $T$, Eq. (S25), now enumerating the number of rows and columns of each matrix:

$$T = -i K_{\text{trans}} \frac{(\Omega - \omega)^{-1}}{N_{\text{trans}} \times M} K_{\text{inc}}^T \frac{M \times M}{M \times N_{\text{inc}}}.$$  \hspace{1cm} (S34)

By recursive application of the rank inequality for matrix products, we have

$$\text{rank}(T) \leq \min(N, M, N_{\text{trans}}).$$  \hspace{1cm} (S35)

Finally, combining Eq. (S33), Eq. (S35) and the fact rank($\rho_{inc}$) $\leq N_{\text{inc}}$ gives

$$\sum_i |\langle k_{\text{trans},i} \rangle|^2 \leq \sum_{i=1}^{\min(\text{rank}(\rho_{inc}), M, N_{\text{trans}})} \lambda_i,$$  \hspace{1cm} (S36)

where eigenvalues are placed in decreasing order with $\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots$. In the case of equal, incoherent excitations (for which all of the eigenvalues of $\rho_{inc}$ are $1/N_{\text{inc}}$):

$$\sum_i |\langle k_{\text{trans},i} \rangle|^2 \leq \frac{\min(N_{\text{inc}}, M, N_{\text{trans}})}{N_{\text{inc}}}.$$  \hspace{1cm} (S37)

### 4. COUPLED CONSTANTS OF COUPLE-MODE MODELS OF WAVEGUIDE COMBINERS

In this section, we list the coupling constants engineered to maximize transmission of waveguide combiners in Fig. 4. A random sweep on these parameters are done and those achieving a full transmission are selected. As aforementioned, $C$ is chosen to be $\Omega$ in Eq. (S21). The $\Omega$ in Eq. (S20) can be further split into a diagonal Hermitian part $\Omega_{H}$ and an anti-Hermitian part $-i\Gamma$, i.e.

$$\Omega = \Omega_{H} - i\Gamma.$$  \hspace{1cm} (S38)

Due to the constraint,

$$2\Gamma = K^T K,$$  \hspace{1cm} (S39)

it is sufficient to specify only $\Omega_{H}$ and $K$ to fully determine the coupled mode equations in Eqs. (S20,S21). For simplicity, in all three cases, $K$ is chosen so that $\Gamma$ is diagonal, meaning the resonance modes are orthogonal.

The numerical value of $K$ and $\Omega_{H}$ are as follows:

Waveguide combiner (a): two channels to one channel with two resonances:

$$\Omega_{H} = \begin{pmatrix} \omega_0 & 0 \\ 0 & \omega_1 \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} 3.58 & 2.81 \\ 2.81 & -3.58 \end{pmatrix}.$$  \hspace{1cm} (S40)

Waveguide combiner (b): two channels to two channels with one resonance:

$$\Omega_{H} = \omega_0 = 1 \quad \text{and} \quad K = \begin{pmatrix} 3.58 \\ 2.81 \end{pmatrix}.$$  \hspace{1cm} (S41)

Waveguide combiner (c): two channels to two channels with two resonances:

$$\Omega_{H} = \begin{pmatrix} \omega_0 & 0 \\ 0 & \omega_1 \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} 3.58 & 2.81 \\ 2.81 & -3.58 \end{pmatrix}.$$  \hspace{1cm} (S42)

### 5. PERMITTIVITY DATA OF OPTIMIZED METASURFACE UNIT CELLS AND INVERSE DESIGN METHODOLOGY

In this section we provide the permittivity $\epsilon$ data of the unit cells of metasurfaces in Fig. 2 of the main text. All structures have unit cell of size $2 \lambda$ and thickness $0.5 \lambda$, where $\lambda$ is the wavelength of incident excitations. The permittivity $\epsilon$ is constant in the vertical dimension and are discretized into 100 equal-sized grids in the horizontal dimension. The leftmost grid is labelled as 1 and the rightmost one is labelled as 100. In the following table, we tabulate the numerical value of $\epsilon$ from grid 1 to grid 100 for all nine structures, represented by the color they appear in Fig. 2(b) and (c). A blazed grating of the same unit cell dimension, shown in Fig. 7, is optimized as a comparison to metasurfaces in Fig.
We need to first solve the “direct” problem and obtain the current density for the case of metasurfaces. We define
\[ \Theta = \left( \begin{array}{c} -n \times \\ n \times \end{array} \right), \]
so that the modal orthogonality relation can be written as
\[ \int dS \cdot \left( \frac{E_j}{H_j^\dagger} \right)^\dagger \Theta \left( \frac{E_i}{H_i} \right) = 2\text{Re} \left\{ \int dS n \cdot E_i \times H_i^\dagger \right\} = 4P_i \delta_{ij}, \]
where \( P_i \) carries the meaning of time averaged power of the \( i \)th mode. The integral is over the surface through which the energy flow and \( n \) is the unit normal vector to that surface. In this basis, we have
\[ \left( \begin{array}{c} E \\ H \end{array} \right) = \sum_k a_k \left( \begin{array}{c} E_k \\ H_k \end{array} \right), \]
with
\[ a_k = \frac{1}{4P_k} \int dS \cdot \left( \frac{E_k}{H_k^\dagger} \right)^\dagger \Theta \left( \frac{E}{H} \right) = \frac{1}{4P_k} \int dS \cdot \left( \frac{E}{H} \right)^T \Theta \left( \frac{E_k}{H_k^\dagger} \right)^*. \]  
If the geometry in the region where our figure of merit is evaluated is not altered, the modes in the expansion Eq. (S48) remains the same. The only thing changes is the expansion amplitude \( a_k \). Hence, it is more convenient to write sources \( P \) and \( M \) as
\[ P_{\text{adj}} = \sum_k \frac{\partial f}{\partial a_k} \frac{\partial a_k}{\partial E} \] and \[ M_{\text{adj}} = -\sum_k \frac{\partial f}{\partial a_k} \frac{\partial a_k}{\partial H} - \frac{1}{\pi} n \times H_k^\dagger, \]
where the \( E \) and \( H \) derivative follows from Eq. (S49). Combine Eq. (S45) and Eq. (S50), we have
\[ \left( \begin{array}{c} J_{\text{adj}}^E \\ J_{\text{adj}}^M \end{array} \right) = \sum_k \frac{i \omega}{4\pi} \frac{\partial f}{\partial a_k} \left( n \times H_k^\dagger \right). \]
By the equivalence principle [4],
\[ \left( \begin{array}{c} J^E \\ J^M \end{array} \right) = \left( \begin{array}{c} n \times H_{\text{eqv}} \\ -n \times H_{\text{eqv}} \end{array} \right), \]
we conclude that the incident fields of the “adjoint” problem is
\[ \left( \begin{array}{c} E_{\text{adj}} \\ H_{\text{adj}} \end{array} \right) = -\sum_k \frac{i \omega}{4\pi} \frac{\partial f}{\partial a_k} \left( \begin{array}{c} E_k^\dagger \\ -H_k^\dagger \end{array} \right). \]

Recover its correct dimensions, we obtain the final expression:
\[ \delta f = \frac{\omega}{2} \text{Im} \left\{ \int dV \delta e(r) E_{\text{dir}} \cdot \Theta_{\text{adj}} + \delta \mu(r) H_{\text{dir}} \cdot \Theta_{\text{adj}} \right\}, \]
with incident fields for the “adjoint” problem as
\[ \left( \begin{array}{c} E_{\text{adj}} \\ H_{\text{adj}} \end{array} \right) = \sum_k \frac{\partial f}{\partial a_k} \frac{1}{P_k} \left( \begin{array}{c} E_k^\dagger \\ -H_k^\dagger \end{array} \right). \]
When the structure is discretized into small grids, where the \( i \)th grid has volume \( V_i \) and material constant denoted by \( \epsilon_i \) and \( \mu_i \). If the variation of \( \delta \epsilon \) and \( \delta \mu \) is constant for each grid, Eq. (S54) reduces to a grid-wise expression

\[
\frac{\delta f}{\delta \epsilon_i} = \frac{\omega}{2} \text{Im} \left\{ \int_{V_i} dV E_{\text{dir}} \cdot E_{\text{adj}} \right\},
\]

(S56)

\[
\frac{\delta f}{\delta \mu_i} = \frac{\omega}{2} \text{Im} \left\{ \int_{V_i} dV H_{\text{dir}} \cdot H_{\text{adj}} \right\}.
\]

(S57)

These two quantities serve as the gradient of the figure of merit \( f \) with respect to material constants, which can be used in gradient-based optimization algorithms.

REFERENCES

1. Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, 2nd ed. (Cambridge University Press, New York, NY, 2013) p. 235.
2. Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, 2nd ed. (Cambridge University Press, New York, NY, 2013) p. 6.
3. Owen Dennis Miller, *Photonic Design: From Fundamental Solar Cell Physics to Computational Inverse Design*, Ph.D. thesis, University of California, Berkeley (2012).
4. Allen Taflove, Ardavan Oskooi, and Steven G Johnson, *Advances in FDTD computational electrodynamics: photonics and nanotechnology* (Artech house, 2013).
5. Elias M Stein, and Rami Shakarchi, *Fourier analysis: an introduction* (Princeton University Press, 1996).
6. Edmund T Whittaker, and George N Watson, *A course of modern analysis* (Cambridge university press, 1996).
| Grid Number | Red  | Green | Blue  | Purple | Grid Number | Red  | Green | Blue  | Purple |
|-------------|------|-------|-------|--------|-------------|------|-------|-------|--------|
| 1           | 11.89| 7.922 | 6.601 | 11.39  | 41          | 9.231| 4.691 | 3.92  | 1      |
| 2           | 11.89| 9.241 | 5.928 | 8.304  | 42          | 9.231| 1.011 | 1.797 | 1.319  |
| 3           | 6.309| 2.162 | 9.373 | 12     | 43          | 1.082| 1.152 | 6.734 | 5.255  |
| 4           | 6.309| 2.866 | 2.437 | 9.75   | 44          | 1.082| 1.072 | 3.442 | 7.487  |
| 5           | 11.92| 1     | 2.641 | 10.39  | 45          | 1.103| 1.008 | 1.475 | 2.789  |
| 6           | 11.92| 1.752 | 3.947 | 8.941  | 46          | 1.103| 1.006 | 9.828 | 8.16   |
| 7           | 1.122| 9.385 | 8.295 | 7.522  | 47          | 1.128| 2.284 | 4.833 | 7.361  |
| 8           | 1.122| 7.499 | 4.863 | 2.762  | 48          | 1.128| 1.877 | 7.704 | 8.3    |
| 9           | 1.102| 11.8  | 3.239 | 4.076  | 49          | 1.131| 1.055 | 10.98 | 5.845  |
| 10          | 1.102| 5.583 | 5.46  | 6.249  | 50          | 1.131| 2.286 | 7.713 | 2.966  |
| 11          | 1.088| 11.9  | 5.003 | 9.985  | 51          | 1.115| 2.537 | 10.44 | 8.221  |
| 12          | 1.088| 8.991 | 6.793 | 11.97  | 52          | 1.115| 2.906 | 8     | 2.168  |
| 13          | 1.086| 12    | 6.307 | 9.947  | 53          | 1.14 | 11.87 | 10.85 | 2.438  |
| 14          | 1.086| 9.933 | 8.678 | 6.409  | 54          | 1.14 | 5.203 | 10.67 | 2.178  |
| 15          | 1.08  | 11.27 | 6.4   | 10.81  | 55          | 6.249| 4.319 | 7.427 | 2.906  |
| 16          | 1.08  | 10.97 | 11.46 | 5.345  | 56          | 6.249| 1.7   | 2.668 | 6.846  |
| 17          | 1.06  | 11.68 | 2.831 | 5.306  | 57          | 9.584| 3.795 | 6.294 | 2.725  |
| 18          | 1.06  | 11.84 | 1.311 | 2.27   | 58          | 9.584| 2.538 | 7.808 | 8.729  |
| 19          | 11.86 |12    | 1.081 | 1.036  | 59          | 11.87| 3.437 | 6.719 | 6.055  |
| 20          | 11.86 |3.006 | 1.214 | 6.295  | 60          | 11.87| 5.222 | 6.74  | 1.75   |
| 21          | 1.083 |3.88  | 3.766 | 9.55   | 61          | 11.92| 3.329 | 11.19 | 3.582  |
| 22          | 1.083 |1.151 | 3.966 | 7.194  | 62          | 11.92| 1.498 | 10.84 | 9.64   |
| 23          | 1.12  |2.929 | 2.799 | 4.837  | 63          | 11.94| 2.441 | 9.482 | 6.714  |
| 24          | 1.12  |2.164 | 6.935 | 7.654  | 64          | 11.94| 5.191 | 10.84 | 2.886  |
| 25          | 1.142 |2.22  | 5.21  | 6.807  | 65          | 1.083| 7.84  | 7.965 | 4.687  |
| 26          | 1.142 |4.144 | 6.708 | 8.086  | 66          | 1.083| 4.523 | 11.55 | 7.332  |
| 27          | 11.94 |9.396 | 8.646 | 9.314  | 67          | 11.71| 4.739 | 9.329 | 2.485  |
| 28          | 11.94 |10.45 | 9.447 | 9.668  | 68          | 11.71| 9.656 | 9.427 | 9.137  |
| 29          | 11.93 |12    | 1.272 | 8.637  | 69          | 1.093| 3.601 | 12    | 3.6    |
| 30          | 11.93 |11.59 | 1.016 | 4.561  | 70          | 1.093| 1.047 | 9.771 | 11.16  |
| 31          | 11.95 |11.08 | 7.303 | 9.999  | 71          | 1.118| 5.856 | 11.51 | 3.705  |
| 32          | 11.95 |8.987 | 1.188 | 1.809  | 72          | 1.118| 1     | 11.69 | 8.737  |
| 33          | 11.96 |10.81 | 5.516 | 7.839  | 73          | 1.138| 1.139 | 8.158 | 2.504  |
| 34          | 11.96 |7.072 | 5.624 | 11.81  | 74          | 1.138| 3.608 | 11.52 | 3.502  |
| 35          | 11.95 |2.775 | 9.345 | 9.258  | 75          | 11.84| 1.52  | 9.547 | 1.809  |
| 36          | 11.95 |4.248 | 5.329 | 10.3   | 76          | 11.84| 4.929 | 12    | 7.815  |
| 37          | 1.094 |9.242 | 4.408 | 7.411  | 77          | 10.04| 1.283 | 12    | 9.838  |
| 38          | 1.094 |6.303 | 7.479 | 9.25   | 78          | 10.04| 4.892 | 11.3  | 9.231  |
| 39          | 11.84 |4.283 | 8.235 | 7.464  | 79          | 1.064| 2.854 | 12    | 8.865  |
| 40          | 11.84 |6.691 | 5.801 | 1.981  | 80          | 1.064| 1.812 | 12    | 10.38  |
Table S1. Permittivity value of every grid point of the four metasurfaces in Fig. 2(b).

| Grid Number | Red   | Blue  | Purple | Yellow | Green |
|-------------|-------|-------|--------|--------|-------|
| 1           | 7.548 | 11.76 | 10.33  | 5.319  | 2.667 |
| 2           | 6.51  | 11.72 | 10.6   | 2.436  | 1.902 |
| 3           | 5.003 | 8.872 | 8.305  | 3.825  | 1.941 |
| 4           | 3.994 | 10.95 | 8.504  | 9.406  | 2.273 |
| 5           | 3.516 | 11.14 | 4.964  | 9.629  | 2.575 |
| 6           | 3.279 | 11.47 | 6.554  | 11.69  | 3.744 |
| 7           | 3.361 | 11.43 | 7.313  | 11.54  | 4.91  |
| 8           | 5.218 | 11.65 | 4.648  | 10.3   | 6.271 |
| 9           | 7.594 | 11.72 | 4.625  | 11.08  | 7.473 |
| 10          | 9.32  | 11.75 | 3.695  | 11.22  | 9.204 |
| 11          | 10.39 | 11.62 | 3.419  | 11.21  | 9.069 |
| 12          | 10.63 | 11.75 | 3      | 10.45  | 8.726 |
| 13          | 10.9  | 10.44 | 2.565  | 11.19  | 7     |
| 14          | 10.79 | 6.259 | 1.627  | 10.88  | 6.062 |
| 15          | 10.35 | 5.46  | 2.419  | 10.98  | 4.065 |
| 16          | 8.859 | 1.405 | 1.327  | 8.252  | 5.774 |
| 17          | 6.678 | 1.627 | 1.432  | 1.511  | 2.957 |
| 18          | 5.141 | 1.531 | 1.866  | 7.184  | 3.153 |
| 19          | 4.571 | 1.272 | 1.622  | 4.697  | 3.562 |
| 20          | 4.363 | 1.323 | 1.067  | 7.794  | 3.969 |
| 21          | 5.412 | 1.395 | 1.049  | 9.355  | 4.899 |
| 22          | 7.32  | 2.085 | 1.082  | 11.43  | 5.808 |
| 23          | 9.546 | 2.75  | 2.212  | 11.04  | 6.244 |
| 24          | 10.62 | 3.388 | 1.08   | 11.41  | 7.843 |
| 25          | 11.12 | 4.359 | 1.566  | 10.74  | 7.186 |
| 26          | 11.25 | 5.939 | 5.879  | 6.003  | 6.844 |
| 27          | 11.24 | 7.008 | 7.23   | 3.399  | 6.006 |
| 28          | 11.13 | 5.632 | 5.488  | 1.66   | 4.507 |
| 29          | 10.88 | 7.903 | 6.349  | 1.488  | 2.733 |
| 30          | 10.48 | 4.988 | 7.377  | 1.744  | 2.723 |
| 31          | 9.931 | 8.198 | 10.4   | 1.409  | 5.077 |
| 32          | 7.629 | 5.222 | 9.879  | 2.492  | 5.116 |
| 33          | 6.943 | 4.609 | 7.52   | 3.651  | 7.973 |
| 34          | 8.429 | 4.359 | 6.119  | 8.259  | 9.132 |
| 35          | 9.335 | 3.637 | 2.806  | 5.748  | 8.091 |
| 36          | 10.13 | 7.12  | 4.074  | 7.071  | 8.269 |
| 37          | 10.31 | 6.297 | 3.358  | 8.606  | 6.003 |
| 38          | 10.55 | 3.424 | 1.256  | 8.709  | 4.619 |
| 39          | 10.52 | 5.281 | 1.609  | 10.07  | 3.256 |
| 40          | 10.57 | 5.094 | 1.643  | 10.34  | 2.414 |
| Grid Number | Red   | Blue  | Purple | Yellow | Green |
|-------------|-------|-------|--------|--------|-------|
| 41          | 10.6  | 3.568 | 2.134  | 11.83  | 2.485 |
| 42          | 10.58 | 3.08  | 5.434  | 11.81  | 3.315 |
| 43          | 10.42 | 4.86  | 7.515  | 11.78  | 6.19  |
| 44          | 9.998 | 5.014 | 9.496  | 11.8  | 7.499 |
| 45          | 9.851 | 2.855 | 11.42  | 11.73  | 8.852 |
| 46          | 9.685 | 3.787 | 9.191  | 11.86  | 6.373 |
| 47          | 9.555 | 3.951 | 11.59  | 11.78  | 9.789 |
| 48          | 9.546 | 6.04  | 7.952  | 8.847  | 9.597 |
| 49          | 9.776 | 7.333 | 6.995  | 7.929  | 8.868 |
| 50          | 9.774 | 8.408 | 6.26   | 4.926  | 9.366 |
| 51          | 9.415 | 5.725 | 6.841  | 5.954  | 9.978 |
| 52          | 8.933 | 8.773 | 11.06  | 3.835  | 10.5  |
| 53          | 8.586 | 8.53  | 11.26  | 3.844  | 11.24 |
| 54          | 8.816 | 5.195 | 11.58  | 5.161  | 11.19 |
| 55          | 9.194 | 4.942 | 11.89  | 3.807  | 10.59 |
| 56          | 10.05 | 6.663 | 10.94  | 6.555  | 9.29  |
| 57          | 10.75 | 7.651 | 10.86  | 4.949  | 7.496 |
| 58          | 11.08 | 5.35  | 11.95  | 3.269  | 6.432 |
| 59          | 11.14 | 7.022 | 11.97  | 4.007  | 4.887 |
| 60          | 11.11 | 4.497 | 11.97  | 5.665  | 4.246 |
| 61          | 10.79 | 7.725 | 8.733  | 3.005  | 5.592 |
| 62          | 9.882 | 4.187 | 2.17   | 1.672  | 5.903 |
| 63          | 7.491 | 5.22  | 2.979  | 5.38   | 7.234 |
| 64          | 4.765 | 5.925 | 6.16   | 3.873  | 7.658 |
| 65          | 2.963 | 5.684 | 5.086  | 2.422  | 10.03 |
| 66          | 2.321 | 3.596 | 10.26  | 1.918  | 9.757 |
| 67          | 2.221 | 4.134 | 11.61  | 2.036  | 11.07 |
| 68          | 2.056 | 2.415 | 11.59  | 3.571  | 10.92 |
| 69          | 1.961 | 2.357 | 9.007  | 2.757  | 11.09 |
| 70          | 1.83  | 2.799 | 11.59  | 4.709  | 11.19 |
| 71          | 1.76  | 6.662 | 11.93  | 2.843  | 11.16 |
| 72          | 1.718 | 4.448 | 11.68  | 4.454  | 11.53 |
| 73          | 1.694 | 9.081 | 11.84  | 3.486  | 11.32 |
| 74          | 1.676 | 7.638 | 11.81  | 3.824  | 11.37 |
| 75          | 1.677 | 10.68 | 11.86  | 4.051  | 11.39 |
| 76          | 1.686 | 10.62 | 11.91  | 3.41   | 10.96 |
| 77          | 1.714 | 10.33 | 11.95  | 3.051  | 11.41 |
| 78          | 1.75  | 11.04 | 11.86  | 3.028  | 11.32 |
| 79          | 1.803 | 10.66 | 11.91  | 2.075  | 10.9  |
| 80          | 1.895 | 11.06 | 11.95  | 10.31  | 10.37 |

| Grid Number | Red   | Blue  | Purple | Yellow | Green |
|-------------|-------|-------|--------|--------|-------|
| 81          | 1.979 | 10.95 | 5.021  | 6.687  | 9.682 |
| 82          | 2.1   | 11.26 | 4.714  | 7.621  | 9.162 |
| 83          | 2.195 | 11.7  | 1.117  | 6.089  | 8.026 |
| 84          | 2.423 | 11.54 | 2.396  | 6.826  | 5.468 |
| 85          | 2.549 | 11.67 | 2.281  | 8.962  | 4.681 |
| 86          | 2.769 | 11.35 | 3.909  | 6.953  | 3.953 |
| 87          | 2.897 | 11.15 | 4.568  | 11.58  | 3.826 |
| 88          | 2.941 | 8.333 | 3.977  | 11.35  | 3.519 |
| 89          | 2.952 | 3.502 | 7.863  | 10.73  | 3.396 |
| 90          | 2.884 | 3.468 | 8.549  | 10.7   | 4.025 |
| 91          | 2.887 | 1.864 | 7.358  | 11.82  | 3.814 |
| 92          | 3.153 | 1.978 | 11.06  | 11.84  | 4.737 |
| 93          | 3.389 | 2.869 | 10.58  | 10.25  | 5.439 |
| 94          | 4.322 | 4.798 | 9.657  | 11.1   | 6.352 |
| 95          | 5.968 | 3.351 | 10.15  | 10.62  | 8.625 |
| 96          | 7.621 | 3.65  | 9.488  | 10.04  | 8.412 |
| 97          | 8.799 | 7.751 | 11.96  | 9.177  | 7.233 |
| 98          | 9.374 | 8.295 | 10.58  | 4.901  | 6.993 |
| 99          | 9.461 | 7.933 | 9.879  | 2.688  | 4.756 |
| 100         | 9.068 | 9.613 | 9.359  | 3.652  | 3.318 |

Table S2. Permittivity value of every grid point of the five metasurfaces in Fig. 2(c).