Nanoscale

PAPER

Cite this: *Nanoscale*, 2017, **9**, 9034

Hyperbolic metamaterials for dispersion-assisted directional light emission

Lorenzo Ferrari,[a] Joseph Stephen Thomas Smalley,[b] Yeshaiahu Fainman[b] and Zhaowei Liu *[a,b,c]

Received 10th February 2017,
Accepted 13th June 2017

DOI: 10.1039/c7nr00980a

rsc.li/nanoscale

A novel method is presented to outcouple high spatial frequency (large-k) waves from hyperbolic metamaterials (HMMs) without the use of a grating. This approach relies exclusively on dispersion engineering, and enables preferential power extraction from the top or from the side of a HMM. Multilayer (ML) HMMs are shown to be better suited for lateral outcoupling, while nanowire HMMs are the most convenient choice for top outcoupling. A 6-fold increase in laterally extracted power is predicted for a dipole–HMM system with a Ag/Si ML operating at λ = 530 nm, when metallic filling ratio is changed from an unoptimized to the optimized one. This new design concept supports the cost-effective mass production of high-speed HMM optical transmitters.

1 Introduction

Hyperbolic metamaterials (HMMs) are a class of optical metamaterials characterized by a uniaxial effective permittivity tensor, with components parallel and perpendicular to the optical axis that exhibit opposite signs.[1–6] They are fabricated predominantly in two configurations, as either a multilayer stack (ML) or a nanowire array (NW). The first consists of alternating metallic and dielectric (or semiconducting) layers of deep sub-wavelength thickness, while in the second metallic rods of deep sub-wavelength diameter and high aspect ratio are embedded into a dielectric matrix. The value of the effective permittivity components depends on the constituent permittivities and the metal filling ratio, ρ, which is the volumetric percentage of metal in a unit cell of the ML or the NW. The optical anisotropy of HMMs translates into hyperbolic dispersion, capable of supporting propagating waves with very large wavevectors (or k-vectors): a unique feature that enables several applications, including high-resolution imaging and lithography,[7–11] broadband absorption engineering,[12,13] thermal control at the nanoscale,[14] enhanced nonlinear processes[15,16] and spontaneous emission engineering[17–26]

The latter in particular is technologically relevant to the field of high-speed optical communications. As pointed out in a recent work,[27] incoherent light sources with a spontaneous emission rate enhanced *via* plasmonic[28–30] or hyperbolic nanostructures can achieve modulation speeds comparable to or larger than those of coherent sources, at a lower manufacturing and operational cost. The radiative spontaneous decay rate of a quantum emitter (molecule, electron–hole pair in quantum wells (QWs) or quantum dots (QDs)) $\gamma_r = 1/\tau_r$, where τ_r represents the radiative spontaneous emission lifetime,[31] defines the upper bound to the 3dB electrical modulation bandwidth, f_{3dB}, of a light-emitting diode (LED), according to the formula $f_{3dB} = (2\pi\tau_r)^{-1} = (2\pi)^{-1}\gamma_r$.[32] When such a quantum light source is brought within the near field of a HMM, the waves with large k-vectors contained in its emission spectrum couple to the hyperbolic medium – which supports their propagation – instead of decaying evanescently in the surrounding environment (usually air). The local photonic density of states (PDOS) accessible to the emitter is therefore much larger in HMMs than in conventional media. Fermi's Golden Rule states that the enhancement in PDOS, quantified by the Purcell Factor (PF), is proportional to the enhancement in radiative spontaneous decay rate:[6] because γ_r is proportional to f_{3dB} as shown above, we conclude that the emitter–HMM coupling enhances the electrical modulation bandwidth. The contextual increase in the data transmission rate of the emitter (proportional to f_{3dB} according to the Shannon Sampling Theorem[32]) opens up a host of opportunities in high-speed wireless (Light-Fidelity (Li-Fi),[33] underwater[34]) and fiber-optic[35] communication.

Despite this high potential, two major challenges hinder the practical usage of fast optical transmitters based on HMM technology. Firstly, the discontinuous decrease in PDOS from HMMs to their surrounding environment implies that their interface with air suffers a remarkable impedance mismatch.

[a] *Materials Science and Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0418, USA. E-mail: zhaowei@ucsd.edu*
[b] *Department of Electrical and Computer Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0407, USA*
[c] *Center for Memory and Recording Research, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0401, USA*

As a consequence, large-k waves remain trapped inside hyperbolic media, unless a suitable mechanism is provided that mediates the HMM-to-air transition. The traditional solution described in the literature consists of milling through or depositing on top of the HMM a sub-wavelength grating, whose periodicity provides the matching k-vector required by momentum conservation.[22,23,36–39] Secondly, it is not trivial to shape the far-field emission pattern of the light outcoupled from HMMs. Directional control of emission is beneficial for applications such as on-chip photonics, where light must be switched in-plane and out-of-plane with respect to the chip surface, and fiber-optic communication, where the optical signal must be effectively channeled into the numerical aperture of a multi-mode or single-mode fiber. To date, the only outcoupling structure shown to control the emission pattern is a bullseye grating.[36,38]

The existing approaches to address such challenges are inadequate from a manufacturing standpoint. Adding a sub-wavelength grating typically means extra time and cost in fabrication: gratings are defined *via* focused ion-beam (FIB)[22] or electron-beam lithography,[37,38] which are not economically sustainable for mass production because of their limited throughput. Furthermore, the optimum grating geometry determined *via* analytical or numerical simulations is often hardly achievable in practice with the above-mentioned techniques. 1D gratings with rectangular cross-section have not shown good directional control properties;[22] bullseye gratings extract radiation into a conical pattern,[38] but it is not clearly understood how to control and shape the emission pattern arbitrarily *via* design parameters. A common issue to all grating types is that the extraction efficiency and the directional control depend on the relative emitter-grating position: a quantum light source contained in a horizontal plane below the grating displays a different behavior whether it is located adjacent to a crest or to a trough.[22] As a consequence, the outcoupling performance of the grating is not univocally determined, but results from an average over the spatial distribution of the emitters.

In this paper, we propose a novel paradigm based on dispersion engineering to outcouple large-k waves from HMMs. With a suitable selection of the HMM filling ratio, we extract high k-vectors into the far field by compressing their component parallel or perpendicular to the HMM optical axis, thereby enhancing the overall power routed along the corresponding Cartesian direction (see Fig. 1). This method was previously suggested by West *et al.*[40] for the case of a ML HMM of Type II; the present work extends the tractation, to include both Type I and Type II dispersion and both ML and NW geometries. By lifting the requirement for a grating, our approach makes the fabrication of fast optical transmitters based on quantum emitter–HMM coupling more practical and versatile. The extraction mechanism relies on the bulk properties of the HMM, rather than on spatially varying features of a superimposed structure: as such, it affects equally all the emitters contained in the same plane parallel or orthogonal to the optical axis. This enables the effective channeling of emission

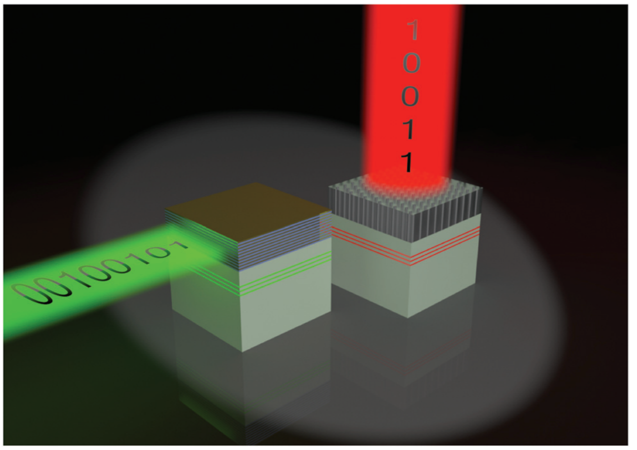

Fig. 1 Artistic representation of high-speed optical transmission *via* unpatterned HMMs. Light emitted from a multiple quantum well, depicted by 3 green (left) and red (right) luminescent layers, couples to a ML (left) and a NW (right) HMM block. The hyperbolic dispersions are designed to extract radiation respectively from the side or from the top interface.

from a QW, which for practical purposes can be thought of as a plane of quantum sources (electron–hole pairs). We first discuss the theory of dispersion-assisted directional outcoupling in the ideal case of zero optical losses, analyzing the four HMM configurations that induce this phenomenon. We then assess two metal/dielectric material systems for ML HMMs in the visible range, and observe how their loss restricts the practically achievable configurations and the band of outcoupled k-vectors. The developed model is finally applied to the study of a colloidal QD–HMM system: by means of finite element electromagnetic simulations, we determine how the light emitted from a point dipole (representing the QD) into a block of HMM is outcoupled by the latter into directional far-field radiation, polarized along the optical axis independently of the dipole orientation. After evaluating the performance of the QD–HMM system, we conclude our study by suggesting guidelines for its practical implementation and future development.

2 Results and discussion

2.1 Principle of large-k waves outcoupling *via* dispersion engineering

In the following we explain how large-k waves can be directionally extracted from lossless HMMs *via* a filling ratio optimization procedure. The results presented hold true for any medium with hyperbolic dispersion regardless of how its effective parameters are retrieved, and therefore indistinctly apply to ML and to NW configurations. Since ML HMMs offer a wider constituent materials choice and are more versatile from a fabrication standpoint, we will focus on this category in the subsequent analysis.

A hyperbolic medium is described by a uniaxial permittivity tensor of the form:[6]

$$\bar{\bar{\varepsilon}} = \begin{bmatrix} \varepsilon_\perp & 0 & 0 \\ 0 & \varepsilon_\perp & 0 \\ 0 & 0 & \varepsilon_{zz} \end{bmatrix}, \tag{1}$$

in a Cartesian frame of reference $\{\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{z}}\}$ where the unit vector $\hat{\mathbf{z}}$ is parallel to the optical axis. For a periodic ML HMM with layer interfaces orthogonal to $\hat{\mathbf{z}}$, the effective parameters $\varepsilon_\perp$ and ε_{zz} are obtained through the homogenization formulae:[6]

$$\varepsilon_\perp(\omega,\rho) = \rho\varepsilon_\mathrm{m}(\omega) + (1-\rho)\varepsilon_\mathrm{d}(\omega), \tag{2}$$

$$\varepsilon_{zz}(\omega,\rho) = \left(\frac{\rho}{\varepsilon_\mathrm{m}(\omega)} + \frac{1-\rho}{\varepsilon_\mathrm{d}(\omega)}\right)^{-1}. \tag{3}$$

$\varepsilon_\mathrm{m}(\omega)$ and $\varepsilon_\mathrm{d}(\omega)$ are respectively the permittivities of the metallic and the dielectric layers, which in the absence of spatial dispersion and optical loss depend solely on the angular frequency ω and are real quantities, and $0 < \rho < 1$ is the volumetric filling ratio of metal. The dispersion of the effective medium is hyperbolic only at those frequencies and at those filling ratios at which $\varepsilon_\perp\varepsilon_{zz} < 0$. This requirement classifies the behavior of HMMs into two distinct types:

Type I. When $\varepsilon_\perp(\omega,\rho) > 0$ and $\varepsilon_{zz}(\omega,\rho) < 0$;

Type II. When $\varepsilon_\perp(\omega,\rho) < 0$ and $\varepsilon_{zz}(\omega,\rho) > 0$.

Let us consider the interface, contained in the xy plane of the real space, between a nonmagnetic HMM in the $z < 0$ region and a lossless isotropic dielectric medium in the $z > 0$ region. A plane wave with wavevector $\mathbf{k}_\mathrm{HMM}$ of arbitrary magnitude is subject inside the hyperbolic medium to the dispersion:

$$\left(\frac{k_{\mathrm{HMM},\perp}}{k_0}\right)^2 \frac{1}{\varepsilon_{zz}} + \left(\frac{k_{\mathrm{HMM},z}}{k_0}\right)^2 \frac{1}{\varepsilon_\perp} = 1, \tag{4}$$

where, by virtue of the cylindrical symmetry of the permittivity, we introduced the cylindrical coordinate $k_{\mathrm{HMM},\perp} = \sqrt{k_{\mathrm{HMM},x}{}^2 + k_{\mathrm{HMM},y}{}^2}$, and $k_0 = \omega/c$ (c = speed of light in vacuum). In the dielectric medium, dispersion takes instead the form

$$\left(\frac{k_{\mathrm{diel},\perp}}{k_0}\right)^2 + \left(\frac{k_{\mathrm{diel},z}}{k_0}\right)^2 = n^2, \tag{5}$$

where $k_{\mathrm{diel},\perp} = \sqrt{k_{\mathrm{diel},x}{}^2 + k_{\mathrm{diel},y}{}^2}$ and n is the refractive index, and the magnitude of the wavevector $\mathbf{k}_\mathrm{diel}$ is equal to $k_0 n$. Eqn (4) and (5) define respectively a hyperboloid and a sphere in the space of k-vectors at a given frequency $\bar{\omega}$. Their cross-section in the $k_x k_z$ plane is represented in Fig. 2(a) for the case of a Type II HMM and air. To simplify the notation and make the discussion clearer, we restrict the following analysis to the ($k_x > 0$, $k_z > 0$) quadrant; analogous considerations can be extended by symmetry to the remaining 3 quadrants. We define "large-k waves" those plane waves with k-vector $|\mathbf{k}_\mathrm{HMM}| > |\mathbf{k}_\mathrm{air}|$, and "short-$k$ waves" those with k-vector $|\mathbf{k}_\mathrm{HMM}| \le |\mathbf{k}_\mathrm{air}|$. The HMM supports the propagation of large-k waves of components $k_{\mathrm{HMM},x} > k_0 n_\mathrm{air} = k_0$ and $k_{\mathrm{HMM},z} \lessgtr k_0$. When such waves reach the HMM–air boundary, the conservation of the k-vector component parallel to the interface mandates that $k_{\mathrm{air},x} = k_{\mathrm{HMM},x}$. Since $k_{\mathrm{HMM},x} > k_0$, $k_{\mathrm{air},z} = \sqrt{k_0{}^2 - k_{\mathrm{HMM},x}{}^2}$ is a purely

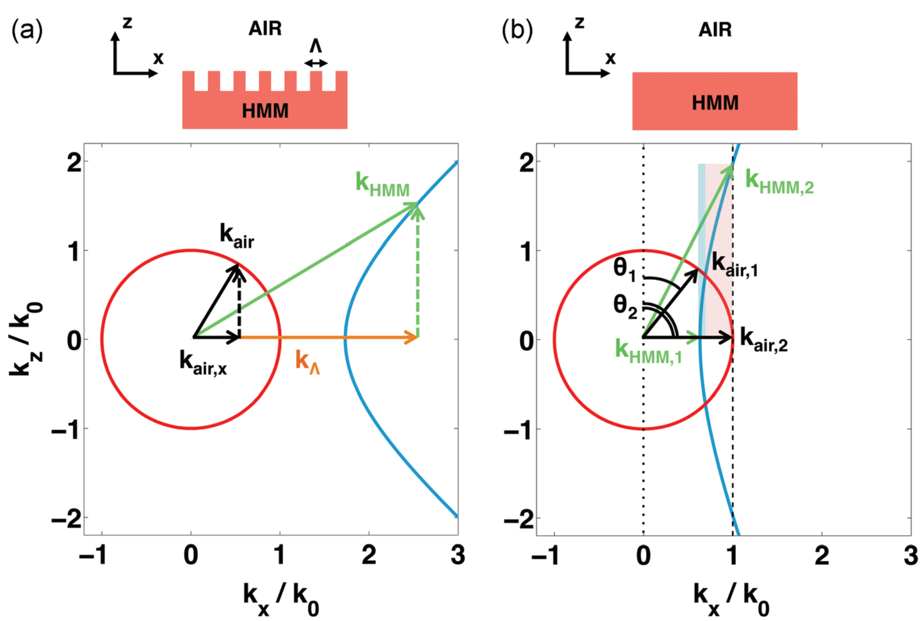

Fig. 2 Iso-frequency curve of Type II HMM (blue hyperbola) showing large k-vector outcoupling upwards into air (red circle) *via* (a) grating and (b) dispersion engineering. The two configurations are schematically represented above the plots (Λ = grating pitch). For clarity, only the right branch of the HMM iso-frequency curve is shown, and the normalization by k_0 is omitted in the vector nomenclature. In (b), the portion of outcoupled k-vector band falling within the ($k_x > 0$, $k_z > 0$) quadrant is highlighted in light blue (short-k waves) and pink (large-k waves).

imaginary number. This turns a propagating wave in the HMM into an evanescent wave in air, preventing its successful outcoupling into the far-field.

Impedance mismatch is traditionally bridged by patterning a grating onto the HMM (inset of Fig. 2(a)). A 1D grating with pitch Λ and periodicity along $\hat{\mathbf{x}}$ provides an extra k-vector $\mathbf{k}_\Lambda = k_\Lambda\hat{\mathbf{x}} = (2\pi/\Lambda)\hat{\mathbf{x}}$,[41] outcoupling into air the band of k-vectors with components $k_{\mathrm{HMM},x} > k_0$ such that $k_{\mathrm{HMM},x} = k_\Lambda + k_{\mathrm{air},x}$ is verified for some real $k_{\mathrm{air},x} \le k_0$. At that point indeed the plane wave becomes propagating in air, as eqn (5) forces its component $k_{\mathrm{air},z}$ to be a real quantity $\le k_0$.

The same goal can be achieved with the alternative approach schematized in Fig. 2(b). If the filling ratio ρ of the HMM is properly designed, the hyperbolic iso-frequency curve gets "straightened" along the z direction and "squeezed" along the x direction within the circular iso-frequency curve of air. In this case, there exists in the HMM a band of k-vectors which possess a conserved component $k_{\mathrm{HMM},x} \le k_0$, and therefore retain their propagating nature across the HMM–air boundary. The band is delimited by two extremes $\mathbf{k}_{\mathrm{HMM},1}$ and $\mathbf{k}_{\mathrm{HMM},2}$, and contains both short- and large-k waves, separated by a vector $\mathbf{k}_{\mathrm{HMM,sep}}$ (not drawn in the figure). $\mathbf{k}_{\mathrm{HMM},1}$, $\mathbf{k}_{\mathrm{HMM,sep}}$ and $\mathbf{k}_{\mathrm{HMM},2}$ are defined as the intersections of the hyperbolic iso-frequency curve respectively with the x axis, the circular iso-frequency curve of air and the straight line $k_x \equiv k_0$. The relative contribution of large-k waves, $(k_{\mathrm{HMM},2x} - k_{\mathrm{HMM,sep}\ x})/(k_{\mathrm{HMM},2x} - k_{\mathrm{HMM},1x})$, increases as the hyperbolic iso-frequency curve gets more squeezed; the z-bandwidth $k_{\mathrm{HMM},2z} - k_{\mathrm{HMM,sep}\ z}$ and the z-density $dk_{\mathrm{HMM}\ z}/dk_{\mathrm{HMM}\ x}$ of large-k waves increase as the hyperbolic iso-frequency curve gets straighter. Refraction into air occurs within an angular range $\theta_1 \le \theta \le \theta_2$, where θ is the angle between the refracted k-vector and the optical axis z, $\theta_1 = \arctan(k_{\mathrm{HMM},1x}/\sqrt{k_0{}^2 - k_{\mathrm{HMM},1x}{}^2})$ and $\theta_2 = 90°$. The limit case of infinite z-bandwidth and refraction normal to the interface ($\theta_1 = \theta_2 = 0°$) is reached when the hyperbola branches become straight lines and collapse onto the optical axis, as mathematically detailed in Appendix A. Although such situation appears ideal for applications, before drawing conclusions we must include energy propagation in our analysis. We recall that the power flux within a medium is perpendicular to the isofrequency surfaces.[6] When the two branches of the hyperbola become parallel to the z axis, namely the curvature at vertices becomes minimal, energy within the HMM flows parallel to the HMM–air interface and never reaches it, making outcoupling impossible. Therefore a tradeoff is required: the curvature of the hyperbola at vertices must be small enough to allow a sufficient z-bandwidth, but large enough to avoid that energy travels too long within the HMM – with the risk of being dissipated by loss – before touching the boundary with air.

2.2 Design guidelines and outcoupling configurations

We now derive comprehensive guidelines to optimize large-k wave extraction from a Type II HMM into air through a flat interface contained in the xy plane (top outcoupling). Our considerations are again restricted for simplicity to a cross-section of the iso-frequency curve in the ($k_x > 0$, $k_z > 0$) quadrant. As concluded in the previous subsection, the target dispersion will be the one that enables infinite bandwidth and normal emission; once the optimum ε_{zz} and $\varepsilon_\perp$ are obtained, we will have to arbitrarily (the tradeoff depends on the application) relax the infinite bandwidth condition to achieve successful outcoupling.

The target iso-frequency curve corresponds to a limit hyperbolic curve with maximal straightening and maximal squeezing; these simultaneous requests are formalized in system (13) of Appendix A. When optimization is performed at a given emission frequency $\bar{\omega}$, we have two equations, involving two functions $\varepsilon_\perp(\bar{\omega},\rho)$ and $\varepsilon_{zz}(\bar{\omega},\rho)$, and only one variable ρ. The resulting overdetermined system admits a solution $\bar{\rho}$ such that $\varepsilon_{zz}(\rho) \to 0^+$ as $\rho \to \bar{\rho}$, provided that $|\varepsilon_\perp(\rho)|$ approaches a finite value or approaches 0 less quickly than $\varepsilon_{zz}(\rho)$ does. Therefore, we optimize dispersion as follows: we choose $\bar{\rho}$ such that $\varepsilon_{zz}(\bar{\rho}) \sim 0^+$ (but not exactly $\varepsilon_{zz}(\bar{\rho}) = 0$, to avoid the energy propagation issue discussed above), making sure that the related condition on $\varepsilon_\perp(\bar{\rho})$ is verified. Fig. 3(c) shows an iso-frequency curve where ε_{zz} does not take the limit value 0, but is small enough to possess a band of outcoupled k-vectors, highlighted in light blue (short-k waves) and pink (large-k waves).

Large wavevectors can also be extracted when the HMM–air interface lies in the yz plane (side outcoupling). In this case, one equation suffices to request both maximal straightening and maximal squeezing (Appendix A). Its solution $\bar{\rho}$ satisfies $|\varepsilon_\perp(\rho)| \to 0^+$ as $\rho \to \bar{\rho}$, provided that $\varepsilon_{zz}(\rho)$ approaches a finite value or approaches 0 less quickly than $|\varepsilon_\perp(\rho)|$ does. In analogy to the previous reasoning, we do not select $\bar{\rho}$ such that $|\varepsilon_\perp(\bar{\rho})| = 0$, but rather $|\varepsilon_\perp(\bar{\rho})| \sim 0^+$; a situation where $\varepsilon_\perp$ is small enough to possess an outcoupled band, but not yet approaching the limit value 0, is represented in Fig. 3(d). While in the case of top extraction the outcoupled band always contains both short- and large-k waves, in the case of side extraction only the latter are present if the condition $\varepsilon_{zz}(\bar{\rho}) > 1$ is verified (like in Fig. 3(d)).

Type I dispersion is mathematically a 90° rotation of the Type II one (Appendix A). With arguments and caveats similar to those hitherto discussed, top outcoupling is achieved at $\bar{\rho}$ such that $|\varepsilon_{zz}(\rho)| \sim 0^+$ (an iso-frequency curve with small ε_{zz} is shown in Fig. 3(a)), while side outcoupling requires $\varepsilon_\perp(\rho) \sim 0^+$ (an iso-frequency curve with small $\varepsilon_\perp$ is shown in Fig. 3(b)). Both short- and large-k waves are always present in a side-outcoupled band, while a top-outcoupled one features exclusively large-k components if the condition $\varepsilon_\perp(\bar{\rho}) > 0$ is verified (Fig. 3(a)).

The design guidelines traced in the present subsection, which relate large-k extraction to the epsilon-near-zero (ENZ) behavior of the effective permittivity components, are summarized in Table 1.

2.3 Influence of loss on dispersion

Thus far we have assumed that the constituent materials of the ML or the NW are lossless. This implies that their permittivities ε_m and ε_d, and therefore the effective permittivity com-

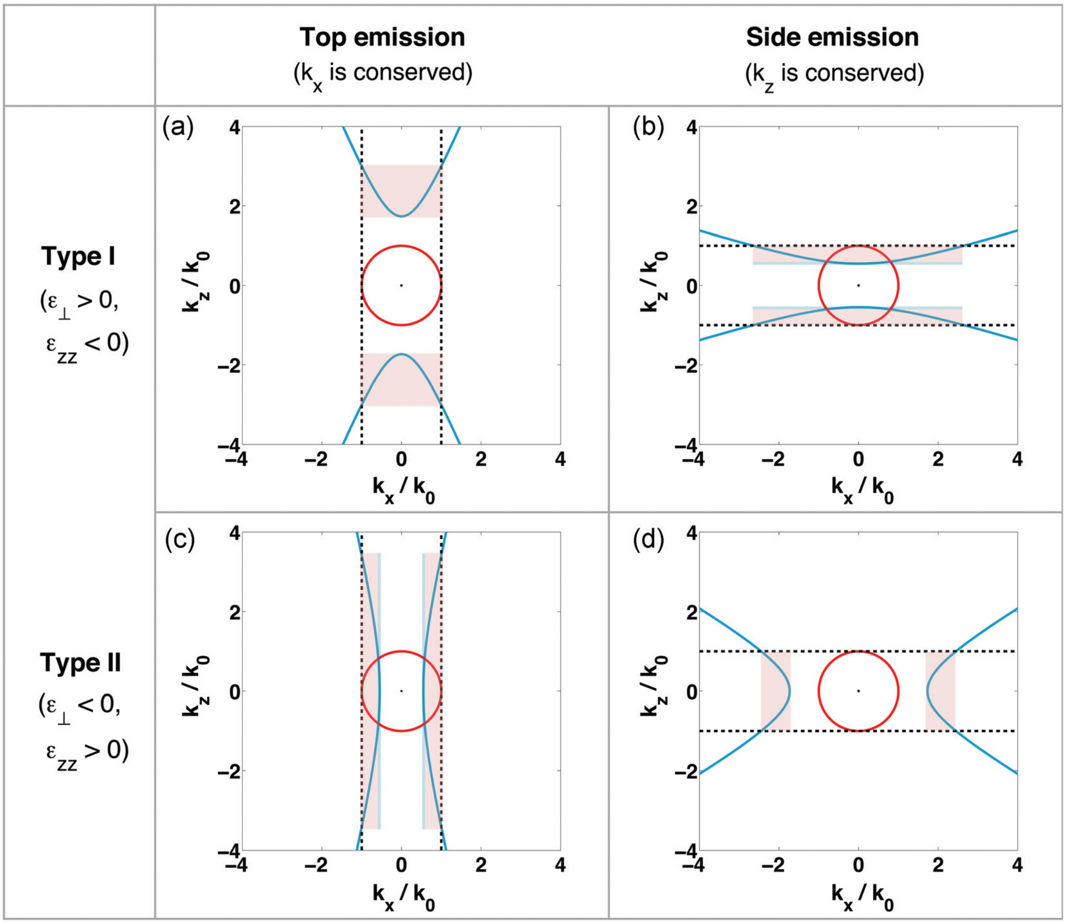

Fig. 3 Cases of dispersion-assisted directional outcoupling from HMMs (blue hyperbola) into air (red circle). The effective parameters of the hyperbolic iso-frequency curves are: (a) $\varepsilon_\perp = 3$, $\varepsilon_{zz} = -0.5$; (b) $\varepsilon_\perp = 0.3$, $\varepsilon_{zz} = -3$; (c) $\varepsilon_\perp = -5$, $\varepsilon_{zz} = 0.3$; (d) $\varepsilon_\perp = -1$, $\varepsilon_{zz} = 3$. The band of outcoupled *k*-vectors is highlighted in light blue (short-*k* waves, only present in (b) and (c)) and in pink (large-*k* waves).

Table 1 Design guidelines for dispersion-assisted outcoupling in HMMs. The notation $f(x) = o(g(x))$ as $x \to \bar{x}$, read "*f* is little-*o* of *g* as *x* approaches $\bar{x}$," means that $\lim_{x\to\bar{x}} f(x)/g(x) = 0$

HMM Type	Outcoupling	Guideline (pick ρ such that.)	Caveat (verify that.)
Type I	Top	$\lvert\varepsilon_{zz}(\rho)\rvert \sim 0^+$	$\lvert\varepsilon_{zz}(\rho)\rvert = o(\varepsilon_\perp(\rho))$
	Side	$\varepsilon_\perp(\rho) \sim 0^+$	$\varepsilon_\perp(\rho) = o(\lvert\varepsilon_{zz}(\rho)\rvert)$
Type II	Top	$\varepsilon_{zz}(\rho) \sim 0^+$	$\varepsilon_{zz}(\rho) = o(\lvert\varepsilon_\perp(\rho)\rvert)$
	Side	$\lvert\varepsilon_\perp(\rho)\rvert \sim 0^+$	$\lvert\varepsilon_\perp(\rho)\rvert = o(\varepsilon_{zz}(\rho))$

ponents $\varepsilon_\perp$, ε_{zz}, are real quantities. Loss in actual materials, mathematically represented by a nonzero imaginary part of permittivity, distorts the iso-frequency curve. In eqn (4), the *k*-vector component conserved at the HMM-dielectric interface is the independent variable and takes only real values. Since now $\varepsilon_\perp$ and ε_{zz} are complex quantities, the *k*-vector component orthogonal to the interface becomes complex, and it is its real part that appears in the iso-frequency curve.

Effective permittivities with a small imaginary part do not alter excessively the iso-frequency curve, compared with the lossless case. Our analysis showed that top outcoupling is achieved for $\mathrm{Re}\{\varepsilon_{zz}\}$-near-zero ($\mathrm{Re}\{\varepsilon_{zz}\} \sim 0^-$ for Type I, $\mathrm{Re}\{\varepsilon_{zz}\} \sim 0^+$ for Type II), and side outcoupling for $\mathrm{Re}\{\varepsilon_\perp\}$-near-zero ($\mathrm{Re}\{\varepsilon_\perp\} \sim 0^+$ for Type I, $\mathrm{Re}\{\varepsilon_\perp\} \sim 0^-$ for Type II). We subsequently investigate whether any coupling behavior is associated with a resonance, and therefore incurs a penalty for large losses (a large imaginary permittivity).

In ML HMMs, eqn (2) prescribes that $\mathrm{Re}\{\varepsilon_\perp\}$ as a function of ω have a zero induced by the zero of $\mathrm{Re}\{\varepsilon_m\}$; $\mathrm{Re}\{\varepsilon_{zz}\}$, according to eqn (3), can have a zero as well, but also crosses zero in correspondence of a resonant pole.[42] The zero of $\mathrm{Re}\{\varepsilon_\perp\}$ and the pole of $\mathrm{Re}\{\varepsilon_{zz}\}$ can be continuously tuned with ρ: and since $\mathrm{Im}\{\varepsilon_\perp\}$ may be low, but $\mathrm{Im}\{\varepsilon_{zz}\}$ is always large, we conclude that for ML HMMs side outcoupling is more viable than top outcoupling. The situation is reversed for NW HMMs: in that geometry, the direction of free electron motion (parallel to the wires) is orthogonal to the one in MLs (parallel to the layers), so the role of the $\varepsilon_\perp$ and ε_{zz} components is exchanged. NW HMMs are therefore a better choice for top outcoupling, and a poor one for side outcoupling.

We now consider a hyperbolic medium with effective parameters $\varepsilon_\perp$, ε_{zz} calculated for a ML geometry, and optimize the side extraction of large-*k* waves at the wavelengths $\lambda_1 = 530$ nm and λ_2 = 470 nm relevant for optical communication. We achieve this goal by determining for a selected material system (Ag/Si for λ_1, Ag/SiO_2 for λ_2) the filling ratio that maximizes the bandwidth of outcoupled *k*-vectors. The SiO_2 permittivity is assumed to be 2.25, while the Ag and the Si permittivities are taken respectively from (ref. 43 and 44). At λ_1, a Ag/Si effective medium displays a Type I behavior with $\mathrm{Re}\{\varepsilon_\perp\} \sim 0^+$ at $\rho = 0.58$ (the exact $\mathrm{Re}\{\varepsilon_\perp\} \to 0^+$ condition is achieved at decimal digits of ρ that are meaningless from a fabrication standpoint) (Fig. 11(a and c) of Appendix C). The material permittivities $\varepsilon_{Ag} = -11.66 - 0.36i$ and $\varepsilon_{Si} = 17.23 - 0.44i$ result in effective parameters $\varepsilon_\perp = 0.5 - 0.4i$ and $\varepsilon_{zz} = -39.03 - 3.34i$. Restricting our considerations to the ($k_x > 0$, $k_z > 0$) quadrant, we obtain a band including both short-*k* ($0 < k_z/k_0 \leq 0.53$, $0 < |\mathrm{Re}\{k_x/k_0\}| \leq 0.84$) and large-*k* waves ($0.53 < k_z/k_0 \leq 1$, $0.84 < |\mathrm{Re}\{k_x/k_0\}| \leq 4.72$). The total outcoupled band, including the remaining 3 quadrants, is showed in Fig. 4(a). At λ_2 instead, a Ag/SiO_2 effective medium achieves a Type II regime with $\mathrm{Re}\{\varepsilon_\perp\} \sim 0^-$ at $\rho = 0.22$ (Fig. 11(b and d) of Appendix C). The material permittivities $\varepsilon_{Ag} = -8.15 - 0.28i$ and $\varepsilon_{SiO_2} = 2.25$ generate effective parameters $\varepsilon_\perp = -0.04 - 0.06i$ and $\varepsilon_{zz} = 3.13-0.01i$. Again restricting our considerations to the ($k_x > 0$, $k_z > 0$) quadrant, we obtain a band entirely composed of large-*k* waves ($0 < k_z/k_0 \leq 1$, $1.77 < |\mathrm{Re}\{k_x/k_0\}| \leq 5.98$). The total outcoupled band, including the remaining 3 quadrants, is showed in Fig. 4(b). Compared to the ideal behavior of lossless HMMs for side outcoupling (Fig. 3(b and d)), the Ag/SiO_2 iso-frequency curve does not present major alterations, whereas the Ag/Si iso-frequency curve exhibits distortion for values of $\mathrm{Re}\{k_x/k_0\}$ close to 0. While $\mathrm{Im}\{\varepsilon_\perp\}$ and $\mathrm{Im}\{\varepsilon_{zz}\}$ for the Ag/SiO_2 system are low enough to preserve the hyperbolicity of the iso-frequency curve, the correspondent parameters for the Ag/Si system, respectively one and two orders of magnitude larger, alter the iso-frequency curve by conferring it a hybrid hyperbolic-elliptical character.

2.4 Influence of finite periodicity on dispersion

Effective medium theory (EMT), leading to eqn (2) and (3), assumes that the length of the ML period is much smaller than the effective wavelength of light within the medium. As the momentum k_x increases, the effective wavelength $\lambda_{\mathrm{eff}} = 2\pi/k_x$ decreases. The accuracy of EMT is thus circumscribed to a finite range of supported large-*k* waves. Bloch's theorem,[45,46] unlike EMT, explicitly incorporates the thickness of the ML components, and lets us estimate the maximum of the large-*k* range as $\max[\mathrm{Re}\{k_x/k_0\}] = \lambda_0/(2L)$. In this formula, $\lambda_0/(2L)$ is the normalized Brillouin zone boundary, L is the period length and λ_0 vacuum wavelength. If, for instance, a practically achievable period length L = 20 nm is assumed for both the Ag/Si and the Ag/SiO_2 MLs, the first supports a maximum $\mathrm{Re}\{k_x/k_0\} = 13.25$ at the vacuum wavelength λ_1, while the second supports a maximum $\mathrm{Re}\{k_x/k_0\} = 11.75$ at the vacuum wavelength λ_2. The upper extremes of the respective outcoupled bands, determined in subsection 3, fall below these maxima: therefore, our results based on the effective medium are consistent for real structures with the above-mentioned period length.

To further assess the accuracy of EMT in a quantitative fashion, we compute by means of Bloch's theorem the isofre-

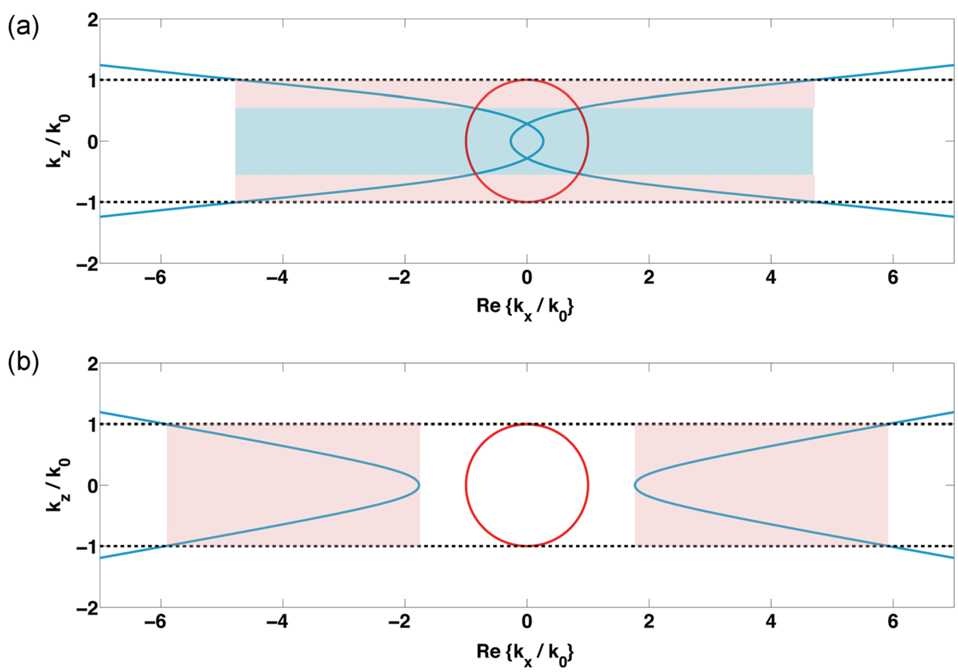

Fig. 4 Iso-frequency curves of Type I and Type II effective media (blue curve) for lateral outcoupling into air (red curve), taking into account the actual loss of the constituent materials: (a) Ag/Si effective medium at λ_1 = 530 nm and filling ratio ρ = 0.58 (Type I), (b) Ag/SiO_2 effective medium at λ_2 = 470 nm and filling ratio ρ = 0.22 (Type II). The band of outcoupled *k*-vectors is highlighted in light blue (short-*k* waves, only present in (a)) and in pink (large-*k* waves).

quency surfaces of the Ag/Si and the Ag/SiO_2 systems. To simplify the notation we restrict our considerations to the k_xk_z plane. Assuming a ML of infinite lateral and vertical extent, finite periodicity L, and with layers orthogonal to the z axis, the dispersion for TM waves is governed by[47]

$$k_{\mathrm{B},z} = L^{-1}\cos^{-1}\left(\frac{A+D}{2}\right), \tag{6}$$

where

$$A, D = \exp(\pm \mathrm{i}k_{\mathrm{m},z}d_\mathrm{m}) \left[\cos(k_{\mathrm{d},z}d_\mathrm{d}) \pm \mathrm{i}\frac{1}{2}\left(\frac{\varepsilon_\mathrm{d}k_{\mathrm{m},z}}{\varepsilon_\mathrm{m}k_{\mathrm{d},z}} + \frac{\varepsilon_\mathrm{m}k_{\mathrm{d},z}}{\varepsilon_\mathrm{d}k_{\mathrm{m},z}}\right)\sin(k_{\mathrm{d},z}d_\mathrm{d})\right]. \tag{7}$$

Eqn (6) expresses the effective k_z component for the entire ML, $k_{\mathrm{B},z}$, as a function of the conserved component k_x, treated as an independent variable and implicitly contained in eqn (7): here, the + and − signs correspond respectively to A and D, d_m (d_d) is the thickness of the metallic (dielectric) layers, and $k_{\mathrm{m},z} = \sqrt{\varepsilon_\mathrm{m}k_0{}^2 - k_x{}^2}$ ($k_{\mathrm{d},z} = \sqrt{\varepsilon_\mathrm{d}k_0{}^2 - k_x{}^2}$) is the z component of the wavevector in the metallic (dielectric) layers.

Fig. 5(a) shows the Bloch iso-frequency curve of TM-polarized waves in a periodic Ag/Si system at λ_1, with L = 20 nm and ρ = 0.58 (corresponding to d_{Ag} = 11.6 nm and d_{Si} = 8.4 nm), and compares it with the EMT iso-frequency curve of Fig. 4(a). The maximum outcoupled component Re$\{k_x/k_0\}$ = 2.93 is 38% lower than the corresponding value obtained with EMT. An analogous comparison is shown in Fig. 5(b) for a periodic Ag/SiO_2 system at λ_2, with L = 20 nm and ρ = 0.22 (corresponding to d_{Ag} = 4.4 nm and $d_{\mathrm{SiO_2}}$ = 15.6 nm). The maximum outcoupled component Re$\{k_x/k_0\}$ = 4.66 is 22% lower than the corresponding value obtained with EMT. This analysis shows that, although reduced, the outcoupled bandwidth remains significant when accounting for finite periodicity. Therefore, the general method we propose retains its validity beyond EMT.

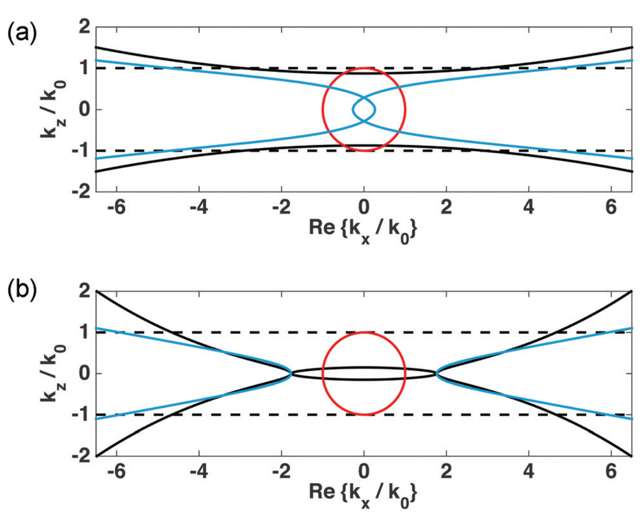

Fig. 5 Influence of finite periodicity on outcoupling. Black curve: Bloch iso-frequency curve; blue curve: EMT iso-frequency curve; red curve: iso-frequency curve of air. (a) Ag/Si ML with period L = 20 nm and filling ratio ρ = 0.58 at λ_1 = 530 nm. (b) Ag/SiO_2 ML with period L = 20 nm and filling ratio ρ = 0.22 at λ_2 = 470 nm.

2.5 Quantum dot–HMM coupling

We apply the results of subsection 2.3 to an elemental light-emitting system, composed of a quantum dot and a block of HMM in air. The system is designed to operate at λ_1 = 530 nm, and to be implemented in a ML geometry: the effective parameters $\varepsilon_\perp$, ε_{zz} are thus obtained from a Ag/Si ML HMM with ρ = 0.58, and enable large-k extraction from the lateral faces of the block. By means of full wave 2D COMSOL simulations, we study a HMM block infinitely extended along the y direction, and with rectangular cross-section l_xl_z. The height l_z = 200 nm corresponds to a ML of 10 periods of length L = 20 nm; a well-controlled structure of this kind has been grown *via* sputtering by our group in the past.[22] The width l_x = 3 μm is easily achievable with standard photolithographic techniques. We model the QD with a point dipole emitting at λ_1, located 5 nm below the bottom surface and 150 nm left of the right face of the HMM block (Fig. 6). Such asymmetric positioning allows a simultaneous evaluation of the extraction efficiency from the HMM lateral faces when the emitter-face distance is either within (150 nm) or outside of (2850 nm) the near-field.

The large-k waves contained in the spectrum of the dipole are coupled into the HMM block, through which they travel until they reach the lateral boundary with air. To quantify how dispersion inhibits or favors their extraction into the far-field, we compare the performance of an unoptimized filling ratio, ρ_A = 0.54, with that of the optimized one, ρ_B = 0.58. The respective isofrequency contours in the k_xk_z are shown in Fig. 12 of Appendix C. At ρ_A, the HMM exhibits Type I behavior, and only an extremely narrow band of dispersion-outcoupled large-k waves is supported; as ρ is increased, the HMM still retains Type I nature, but the bandwidth of the extracted large-k waves grows, assuming its largest value at ρ_B (*cfr*. subsection 2.3). The effectiveness of outcoupling through the lateral faces is determined by the amount of power emitted into a circular arc coaxial with $\hat{\mathbf{x}}$, with vertex in the center of the face and aperture θ = 30° Fig. 6 illustrates the different response of the two filling ratios to a dipole oriented along $\hat{\mathbf{x}}$ ("X-dipole") and to one oriented along $\hat{\mathbf{z}}$ ("Z-dipole"). For an X-dipole, the power outcoupled through the right face of the HMM block at ρ_B results in 7 times larger than that outcoupled at ρ_A, while for a Z-dipole the power extracted at ρ_B exceeds by 39 times that extracted at ρ_A. The dipole–HMM interaction is stronger for a Z-dipole, which outcouples at ρ_B almost twice as much power as an X-dipole. Due to the material loss of the hyperbolic medium, large-k extraction efficiency decreases as the emitter-face distance is increased: the power outcoupled at ρ_B through the left face is 3 orders of magnitude lower than that outcoupled through the right one, both for an X- and for a Z-dipole. We notice that at ρ_A some radiation, guided by the HMM–air interface, reaches the left side of the HMM and gets scattered by the edges; but power propagation through the

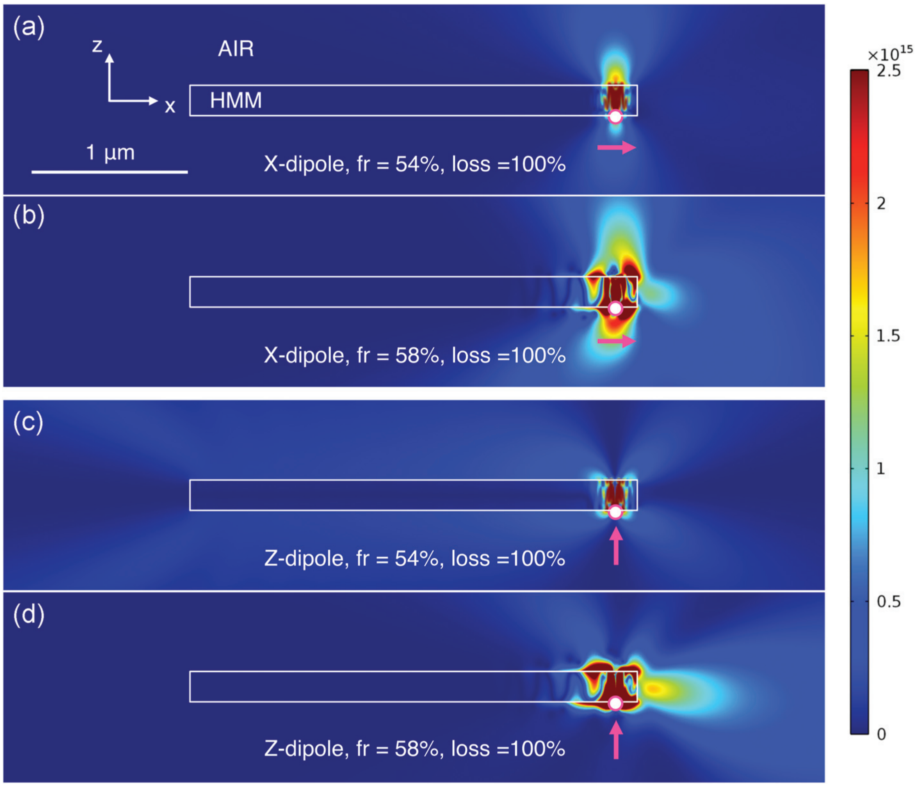

Fig. 6 Spatial power distribution (magnitude of the Poynting vector) at λ_1 = 530 nm of a dipole 5 nm below a block of Ag/Si effective medium in air. The dipole location and orientation are respectively indicated by a pink-bordered white dot and a pink arrow below it. The material loss of the HMM corresponds to 100% of its original value. (a) X-dipole, filling ratio ρ_A = 0.54; (b) X-dipole, filling ratio ρ_B = 0.58; (c) Z-dipole, filling ratio ρ_A = 0.54; (d) Z-dipole, filling ratio ρ_B = 0.58.

bulk of the HMM, which is what our outcoupling mechanism leverages, is again prevented by loss.

The radiation extracted into the far-field through the right face is polarized parallel to the optical axis $\hat{\mathbf{z}}$, regardless of the dipole orientation (Fig. 7). Such behavior has been experimentally observed in a luminescent hyperbolic metasurface, composed of alternating Ag layers and InGaAsP multiple QWs (MQWs), where both parallel- and perpendicular-polarized pumping of the MQWs with respect to the optical axis result in parallel-polarized emission.[48] It is attributed to the fact that only modes with an electric field component parallel to the optical axis are allowed to propagate in a hyperbolic medium.[49]

The potential of dispersion-assisted outcoupling becomes apparent if we artificially reduce the imaginary parts $\mathrm{Im}\{\varepsilon_\perp\}$ and $\mathrm{Im}\{\varepsilon_{zz}\}$, representing loss, to 1% of their original value

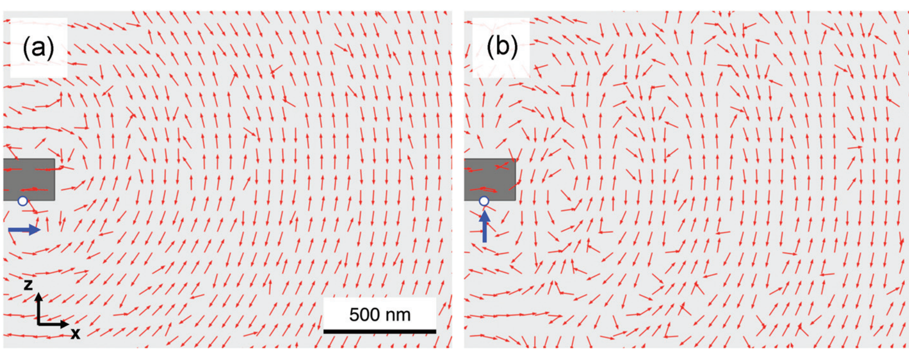

Fig. 7 Electric field polarization of the dipole radiation outcoupled through the right face of the HMM block (whose rightmost portion is shown in grey on the left of the image) at the optimized filling ratio ρ_B = 0.58. The dipole location and orientation are respectively indicated by a white dot and a blue arrow below it. Whether the dipole is polarized along $\hat{\mathbf{x}}$ (a) or along $\hat{\mathbf{z}}$ (b), the radiation leaving the HMM face is predominantly polarized along $\hat{\mathbf{z}}$.

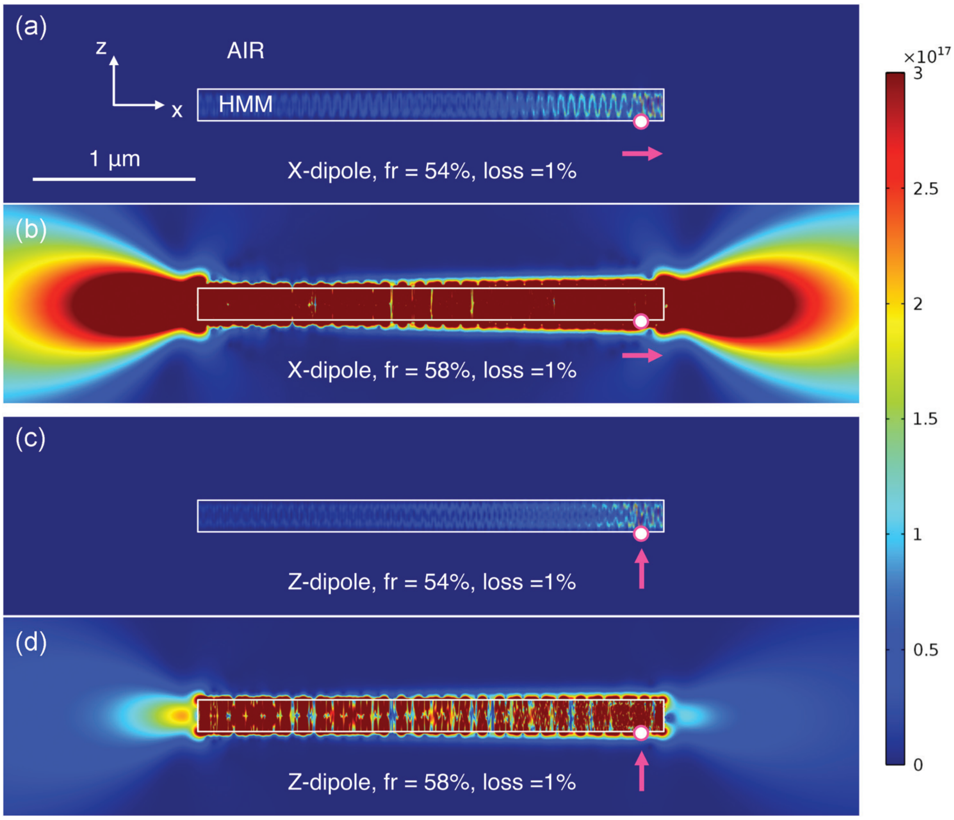

Fig. 8 Spatial power distribution (magnitude of the Poynting vector) at λ_1 = 530 nm of a dipole 5 nm below a block of Ag/Si effective medium in air. The dipole location and orientation are respectively indicated by a pink-bordered white dot and a pink arrow below it. The material loss of the HMM has been reduced to 1% of its original value. (a) X-dipole, filling ratio ρ_A = 0.54; (b) X-dipole, filling ratio ρ_B = 0.58; (c) Z-dipole, filling ratio ρ_A = 0.54; (d) Z-dipole, filling ratio ρ_B = 0.58.

(Fig. 8). In this ideal case, power propagates inside the hyperbolic medium along the characteristic resonance cones, reflected off the top and bottom surfaces of the HMM, until it reaches the left and right boundaries with air. Here, large-*k* waves are efficiently extracted only when the filling ratio is optimized: the power outcoupled through the right face at ρ_B is 3 orders of magnitude larger than the one outcoupled at ρ_A, both for an X- and for a Z-dipole. The strong imbalance between the left and the right face is also removed, as for an X-dipole the power outcoupled at ρ_B through either boundary is almost the same, while for a Z-dipole the left face outcouples at ρ_B 1.7 times as much power as the right one. Which face outcouples the largest power depends on its distance from the dipole. The radiation reaching one face is the superposition of the waves directly traveling to said face and those reflected off the other, whose propagation is not any longer suppressed by loss. For certain dipole locations, the contribution of interference selects the face farther from the dipole as the main outcoupling gateway. Radiation polarized along the optical axis $\hat{\mathbf{z}}$ irrespective of the dipole orientation is emitted into the far-field through both of the lateral faces.

The simultaneous approximation of a ML HMM with an effective medium description and of the 3D space with a 2D environment enables the study of extended structures (size ≥ few µm) with an accurate mesh (mesh element size of the order of 1 nm) and in a reasonable computational time (not exceeding few hours). To test the predictions of our outcoupling method in the absence of approximations, in Appendix B we perform a 3D analysis of the coupling between a quantum dot (point dipole) and a block of HMM of reduced size (base 300 nm × 300 nm, height 200 nm). The Si/Ag ML HMM is first modeled as an effective medium, and then as a periodically layered structure. The results obtained are consistent with the general findings described here.

2.6. Discussion

With the due distinctions in terms of light coherence, a system formed by a suitably designed HMM block coupled to solid-state quantum emitters (QWs, QDs) can become the equivalent of an edge-emitting laser or a vertical-cavity surface-emitting laser (VCSEL), based on spontaneous rather than stimulated emission. "Hyperbolically-enhanced" LEDs with a modulation speed comparable to that of lasers could replace the latter in very-short-haul optical interconnects (on-chip or chip-to-chip communications), where their limitations related to pulse broadening and dispersion do not outweigh their advantages in energy budget, thermal management, reliability and manufacturing cost.[27]

Our analysis showed that ML HMMs naturally support side emission, but are not suited for top emission. In addition, as observed in the previous subsection, the intrinsic loss of constituent materials restricts efficient outcoupling to those emitters in the near field of the side interface. These two constraints can be relaxed simultaneously by fabricating the ML into vertical lamellae, where the non-metallic constituent is a gain medium such as a semiconductor MQW. A configuration of this kind was recently demonstrated with a Ag/InGaAsP system.[48,50–53] Operating in the 1200 nm–1600 nm spectral range, the structure exhibits narrow bands of large-k waves that outcouple to air without the need for a grating, for the reasons theoretically explained in the present work. Another design that meets the top emission and efficient outcoupling requirements exploits a hyperlens-like geometry: the layers here are arranged in a concentric semi-spherical or semi-cylindrical stack, contained in the $z < 0$ region and comprised between the external and internal radii r_{ext} and r_{int}, and the surface of the innermost layer is coated with QDs. Provided that r_{ext} is within the near field of the emitters, radiation from the QDs couples to the HMM, shortly travels along the direction tangential to the layers without experiencing excessive attenuation and gets extracted along the z axis. An alternative approach to achieve top emission consists in growing the ML on the side walls of a vertical nanowire light-emitting structure that incorporates one or more MQWs; the geometry of the nanowires and of the MQWs can be adjusted, as shown in (ref. 54) for blue/green-emitting III-nitride semiconductor nanowires, to boost the emitter–ML interaction and optimize the outcoupling efficiency. Top emission is convenient from a manufacturing standpoint: while edge-emitting devices grown on a wafer must be first diced in sub-units and then tested individually, surface-emitting ones can be tested all at once on the wafer where they are fabricated. Side emission however is desirable for on-chip or chip-to-chip communication, and enables direct light coupling into waveguide-like planar devices. In order to maintain a HMM feature size compatible with photolithography, without limiting efficient outcoupling to the emitters closer to the side faces, the imaginary component of the constituent permittivities can be minimized with low-loss materials[55,56] or compensated for with gain. van der Waals heterostructures, built with single-atomic layer materials ("2D crystals") including graphene, hexagonal boron nitride (hBN) and 2D oxides,[57] can incorporate atomic monolayers of transition metal dichalcogenides such as tungsten diselenide (WSe_2) and tungsten disulfide (WS_2) as active media.[58,59] Heterostructures based on hBN, a natural HMM,[60] or on ultra-thin MLs with 2D active layers, can greatly reduce loss and fully leverage our light extraction method, since their size forbids grating inscription as a practical option. Finally, hybrid ML-NW geometries, or MLs with atomically thin metallic films subject to quantum confinement, might elude with their dispersion the guidelines traced in subsection 2.2.

The feasibility of our approach relies on the exact control of the filling ratio. In the visible range, the ML period must not exceed few tens of nanometers in order to be sub-wavelength (and therefore justify the effective medium approximation). Metallic films for operations at visible wavelengths are typically grown by DC sputtering or e-beam evaporation. In either case, at thicknesses of 10 nm or less, the metal forms islands, rather than a uniform layer. While the size and shape of the grains can be controlled to some extent by tuning the deposition parameters, the resulting morphology intrinsically yields a space-dependent filling ratio. Therefore, the effective HMM parameters need to have a good tolerance with ρ variations, to preserve the applicability of our model. This represents less of a concern for HMMs in the infrared (IR) range: NW and ML systems can be grown with sub-nanometer accuracy *via* atomic layer deposition (ALD) and chemical vapor deposition (CVD),[56,61] and periods of several tens of nm are already well sub-wavelength. IR HMMs hence constitute a promising candidate for early experimental testing. A dynamic fine-tuning of the filling ratio from non-outcoupling to outcoupling, detectable as an enhanced directionality of the emission pattern, can be performed in voltage-controlled HMMs.[60,62,63]

An emitter–HMM system suitable for optical communication should simultaneously maximize, at a given wavelength, the far-field power extracted into a preferential direction and the PF, which determines the modulation bandwidth enhancement. However, the filling ratios that optimize both quantities in general do not coincide. A recent work suggests tapering a hyperbolic ML block to adiabatically outcouple large-k waves from its side.[40] According to that approach, we could design a HMM with filling ratio varying from ρ_1 (in the proximity of the emitter) to ρ_2 (away from the emitter), to first maximize the PF and then the outcoupling. Such transition however must happen over a space scale of few microns, which imposes a constraint on the device size. In addition, power extraction is limited to the tapering direction, making the mechanism inefficient. Finally, the shadowed deposition technique utilized seems impractical in an industrial context. We propose instead to keep a single-ρ design, that achieves a balanced performance between outcoupling and PF. Optimization methods can be developed to tune this trade-off, which we believe is the most viable solution in a mass production perspective.

3 Conclusions

The present work introduced a systematic approach to extract large-k waves from HMMs without the use of a grating. This novel method relies on dispersion engineering, and is applicable to any medium, natural or artificial, described by a hyperbolic permittivity tensor. Extraction of a large-k wave band from the top or from the side faces of the HMM is achieved, as summarized in Table 1, in the ENZ regime of either permittivity component. Loss in the effective medium, dictated both by the loss and the ML or NW arrangement of its constituent materials, selects the preferential outcoupling configuration: side outcoupling for ML HMMs, top outcoupling for NW HMMs. We provided guidelines to maximize the extent of the large-k wave band extracted through the top or side faces, and we applied them to the optimization of a Ag/Si and a Ag/SiO$_2$ ML HMM, respectively at λ_1 = 530 nm and λ_2 = 470 nm. We further discussed how the effective medium description moderately overestimates the outcoupled bandwidth, when the finite periodicity of practical HMM realizations is taken into account. We finally studied a QD-Ag/Si ML HMM system at λ_1, modeling the hyperbolic medium both as an effective medium and as a multilayered structure. For a 3D block of Ag/Si layered medium, we observed a 6-fold increase in lateral power extraction at an optimized filling ratio compared to an unoptimized one.

Future work will explore different material combinations for ML HMMs, with the goal of optimizing lateral extraction at standard IR wavelengths for fiber-optic communication. In parallel, NW HMMs will be studied in view of designing an emitter–HMM system for top outcoupling.

Appendix

A. Limit cases of hyperbolic dispersion

A hyperbola in the xz plane with center in the origin and vertices on the x axis is described by the equation

$$\frac{x^2}{a^2} - \frac{z^2}{b^2} = 1, \tag{8}$$

where a (semi-major axis) is the distance between a vertex and the origin, and b (semi-minor axis) is the distance between a vertex and the asymptote above (or below) it. While a alone determines the separation of the hyperbola branches from the z axis, by defining the coordinates of the vertices $(x, z) = (\pm a, 0)$, the ratio of b to a controls the curvature of the branches at the vertices. The latter is expressed in terms of the eccentricity

$$e = \sqrt{1 + \left(\frac{b}{a}\right)^2}, \tag{9}$$

a positive quantity with limiting values of 1 to $+\infty$. These extreme values are reached by

$$\lim_{\frac{b}{a}\to 0^+} e = 1, \quad \lim_{\frac{b}{a}\to +\infty} e = +\infty, \tag{10, 11}$$

corresponding to cases of maximal and minimal curvature, respectively.

Let us consider a lossless HMM of Type II ($\varepsilon_{zz} > 0$ and $\varepsilon_\perp < 0$). Its dispersion is described by eqn (8), provided that the following substitutions are made:

$$\begin{array}{ll} x \to k_x/k_0 & a^2 \to \varepsilon_{zz} \\ z \to k_z/k_0 & b^2 \to |\varepsilon_\perp|, \end{array} \tag{12}$$

For top outcoupling, k_x is conserved, and large $k_{\text{HMM},z}$ components can be extracted (Fig. 3(c) of main text). Following the main text, we restrict our analysis to the ($k_x > 0$, $k_z > 0$) quadrant. The limit case in which an infinite band of waves propagating within the HMM is converted to propagating waves in air occurs when two conditions are simultaneously verified: the branches of the hyperbola must be "straightened" until their curvature at the vertices vanishes and they align with the z axis, and "squeezed" along the x axis, until the vertex coordinate a becomes <1. The first requirement ensures the infinite extension of the band, the second forces all the k_x components of the band to be identical to a certain $\bar{k}_x < k_0$. This, by virtue of eqn (5), preserves the propagating nature of all the k-vectors in the band across the HMM–air interface. It also implies a vanishing angular spread in the emission, as all the waves in the HMM are refracted into air only at one specific angle with the z axis, $\bar{\theta}_{\text{air}} = \arctan\ (\bar{k}_x/\sqrt{k_0{}^2 - \bar{k}_x{}^2})$. Squeezing is maximized when $\bar{k}_x = 0$: in that case, the two straight branches collapse onto the z axis, $\bar{\theta}_{\text{air}} = 0$ and the emission is orthogonal to the interface. The band edges $\mathbf{k}_{\text{HMM},1}$ and $\mathbf{k}_{\text{HMM},2}$ and the vector $\mathbf{k}_{\text{HMM,sep}}$ separating the short- and large-k constituents of the band (*cfr.* main text) have zero x components, and z components $k_{\text{HMM},1z} = 0$, $k_{\text{HMM},2z} = +\infty$ and $k_{\text{HMM,sep}\ z} = k_0$. The simultaneous requests of infinite bandwidth (*i.e.* maximal straightening, or vanishing curvature at vertices) and normal emission (*i.e.* maximal squeezing, or vertices coinciding with the origin), are formulated by means of eqn (9) and (11) as follows:

$$\begin{cases} \dfrac{b}{a} \to +\infty \\ a \to 0^+ \end{cases} \overset{\Leftrightarrow}{(12)} \begin{cases} \sqrt{\dfrac{|\varepsilon_\perp|}{\varepsilon_{zz}}} \to +\infty \\ \sqrt{\varepsilon_{zz}} \to 0^+. \end{cases} \tag{13}$$

We look for physical solutions of the system, namely finite or vanishing values of $|\varepsilon_\perp|$ and ε_{zz}. The second equation demands that, at a fixed frequency $\bar{\omega}$, the filling ratio ρ take a value $\bar{\rho}$ such that $\varepsilon_{zz}(\rho) \to 0^+$ as $\rho \to \bar{\rho}$. The first equation is simultaneously satisfied if, as $\rho \to \bar{\rho}$, $|\varepsilon_\perp(\rho)|$ either approaches a finite value $|\varepsilon_\perp(\bar{\rho})|$ or approaches 0 less quickly than $\varepsilon_{zz}(\rho)$ does: in mathematical terms, $\varepsilon_{zz}(\rho) = o(|\varepsilon_\perp(\rho)|)$ as $\rho \to \bar{\rho}$.

For side outcoupling, k_z is conserved, and large $k_{\text{HMM},x}$ components can be extracted (Fig. 3(d) of main text). Restricting our analysis to the ($k_x < 0$, $k_z > 0$) quadrant, the band edges $\mathbf{k}_{\text{HMM},1}$ and $\mathbf{k}_{\text{HMM},2}$ and the separation vector $\mathbf{k}_{\text{HMM,sep}}$ are redefined as the intersections of the hyperbolic iso-frequency curve respectively with the x axis, the straight line $k_z \equiv k_0$ and the circular iso-frequency curve of air. In the considered configuration, both maximal straightening and maximal squeezing are accomplished by solely requesting that the curvature at vertices be infinity: by means of eqn (9) and (10), this reads

$$\frac{b}{a} \to 0^+ \quad \overset{\Leftrightarrow}{(12)} \quad \sqrt{\frac{|\varepsilon_\perp|}{\varepsilon_{zz}}} \to 0^+. \tag{14}$$

Eqn (14) is satisfied when, at a fixed frequency $\bar{\omega}$, the filling ratio ρ takes a value $\bar{\rho}$ such that $|\varepsilon_\perp(\rho)| \to 0^+$ and $|\varepsilon_\perp(\rho)| = o(\varepsilon_{zz}(\rho))$ as $\rho \to \bar{\rho}$ (we discarded the unphysical solution $\varepsilon_{zz}(\rho) \to +\infty$, and $|\varepsilon_\perp(\rho)| \to |\varepsilon_\perp(\bar{\rho})|$ finite or $|\varepsilon_\perp(\rho)| \to +\infty$ slower than $\varepsilon_{zz}(\rho)$). The latter condition implies that $\varepsilon_{zz}(\rho)$ either approaches 0 less quickly than $|\varepsilon_\perp(\rho)|$ does, or approaches a finite value $\varepsilon_{zz}(\bar{\rho})$. In the first case, $\mathbf{k}_{\text{HMM},1}$, $\mathbf{k}_{\text{HMM,sep}}$ and $\mathbf{k}_{\text{HMM},2}$ have zero z components, and x components $k_{\text{HMM},1x} = 0$, $k_{\text{HMM,sep}\ x} = k_0$ and $k_{\text{HMM},2x} = +\infty$. In the second case, all the z components remain zero and it is still $k_{\text{HMM},2x} = +\infty$, but $k_{\text{HMM},1x} = \sqrt{\varepsilon_{zz}(\bar{\rho})}k_0$ and $k_{\text{HMM, sep}\ x}$ depends on $\varepsilon_{zz}(\bar{\rho})$: if $\varepsilon_{zz}(\bar{\rho})$

≤ 1, $k_{\text{HMM,sep }x} = k_0$, while if $\varepsilon_{zz}(\bar{\rho}) > 1$ $k_{\text{HMM,sep }x}$ and therefore $\mathbf{k}_{\text{HMM,sep}}$ do not exist, as the hyperbolic and circular iso-frequency curves do not intersect (outcoupled band exclusively composed of large-k waves).

Let us now consider a lossless HMM of Type I ($\varepsilon_{zz} < 0$ and $\varepsilon_{\perp} > 0$). By means of the substitutions

$$\begin{aligned} x &\rightarrow k_x/k_0 \quad a^2 \rightarrow -|\varepsilon_{zz}| \\ z &\rightarrow k_z/k_0 \quad b^2 \rightarrow -\varepsilon_{\perp}, \end{aligned} \tag{15}$$

we can rewrite eqn (8) as

$$-\frac{x^2}{a'^2} + \frac{z^2}{b'^2} = 1, \tag{16}$$

where we renamed

$$|\varepsilon_{zz}| \rightarrow a'^2 \quad \varepsilon_{\perp} \rightarrow \mathrm{b}'^2. \tag{17}$$

Eqn (15)–(17) describe a dispersion of Type I. The change of signs in the left side of eqn (16) corresponds to a 90° rotation of the hyperbola, whose vertices now lie on the z axis. Identical considerations to those just discussed for Type II therefore apply to Type I, keeping in mind that the behavior of top and side outcoupling is now swapped by cause of the rotation. For top outcoupling, an infinite bandwidth of k-vectors is orthogonally extracted through the top HMM–air boundary when, at fixed $\bar{\omega}$, the filling ratio ρ takes a value $\bar{\rho}$ such that $|\varepsilon_{zz}(\rho)| \rightarrow 0^+$ and $|\varepsilon_{zz}(\rho)| = o(\varepsilon_{\perp}(\rho))$ as $\rho \rightarrow \bar{\rho}$. For side outcoupling, an infinite bandwidth of k-vectors is orthogonally extracted through the side HMM–air interface, for a given $\bar{\omega}$, at $\bar{\rho}$ such that $\varepsilon_{\perp}(\rho) \rightarrow 0^+$ and $\varepsilon_{\perp}(\rho) = o(|\varepsilon_{zz}(\rho)|)$ as $\rho \rightarrow \bar{\rho}$.

B. Quantum dot–HMM coupling: 3D study

In the present appendix we extend the analysis of the QD–HMM system performed in subsection 2.5 to a full 3D simulation environment. As the dimensionality increases, so do memory requirements, imposing a concomitant decrease of the domain size if the computational time is to remain finite. We therefore consider a block of HMM with identical height to the 2D structure (l_z = 200 nm), but with shortened width coincident with the length $l_x = l_y$ = 300 nm. To maximize the analogy with the 2D case, we locate a point dipole emitting at λ_1 5 nm below the center of the bottom surface, so that its distance from each lateral surface is 150 nm. The power emitted through a lateral face is collected into a spherical cap coaxial with $\hat{\mathbf{x}}$ (for the faces parallel to the yz plane) or $\hat{\mathbf{y}}$ (for the faces parallel to the xz plane), with vertex in the center of the face and aperture θ = 30°. Symmetry reduces all the dipole-lateral face permutations to 3 geometries: X-dipole, power outcoupled through face yz ($P_{\text{X},\ yz}$); X-dipole, power outcoupled through face xz ($P_{\text{X},\ xz}$); and Z-dipole, power outcoupled through any of the lateral faces (P_{Z}).

We first model the Si/Ag HMM as an effective medium. The powers $P_{\text{X},\ yz}$ and P_{Z} at the optimized filling ratio ρ_{B} are respectively 4 and 31 times larger than at the unoptimized filling ratio ρ_{A}, a proportion close to that of the correspondent 2D enhancements (subsection 2.5). The power $P_{\text{X},xz}$ which lacks of a 2D counterpart, results comparable for the two filling ratios. The lateral power extraction for a Z-dipole coupled to an EMT block with filling ratio ρ_{B} is visualized in Fig. 9.

In order to further investigate the dipole–HMM interaction for practical HMM realizations, we replace the effective medium with a ML structure. The layered stack contains 10 Ag/Si periods of length L = 20, for a total height l_z = 200 nm, and has a width coincident with the length $l_x = l_y$ = 300 nm, so that the block size remains unchanged. In this case, both the collected powers $P_{\text{X},\ yz}$ and $P_{\text{X},\ xz}$ are comparable for ρ_{A} and ρ_{B}, whereas P_{Z} is 6 times larger at ρ_{B} than at ρ_{A}. The quantitative difference with the effective medium result can be understood by recalling that EMT overestimates the power coupled from a dipole into a HMM.[64] Lateral extraction, both for the effective and the layered medium, is best achieved with a Z-dipole, which compared to an X-dipole exhibits a stronger coupling to a HMM.[24,65] Fig. 10 compares the far-field radiation patterns of a Z-dipole-ML block system at unoptimized and optimized filling ratios. At ρ_{B}, a dominant horizontal emission is generated by side outcoupling (upper lobe), accompanied by scatter-

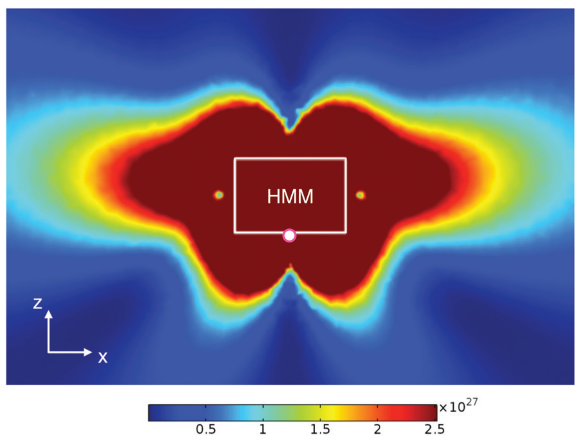

Fig. 9 Spatial power distribution (magnitude of the Poynting vector) in the cross-sectional plane y = 0 of a Z-dipole 5 nm below a 3D block of Ag/Si effective medium in air. The dipole, whose location is indicated by a pink-bordered white dot, emits at λ_1 = 530 nm. The effective medium parameters correspond to the optimized filling ratio ρ_{B} = 0.58.

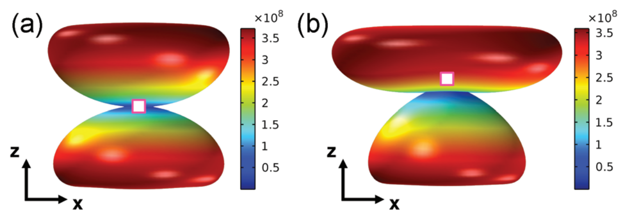

Fig. 10 Far-field emission pattern (electric far-field norm) at λ_1 = 530 nm of a Z-dipole 5 nm below a 3D block of Ag/Si layered medium in air. (a) Filling ratio ρ_{A} = 0.54; (b) filling ratio ρ_{B} = 0.58. The pink-bordered white square indicates the position of the dipole-ML block system with respect to the emission pattern.

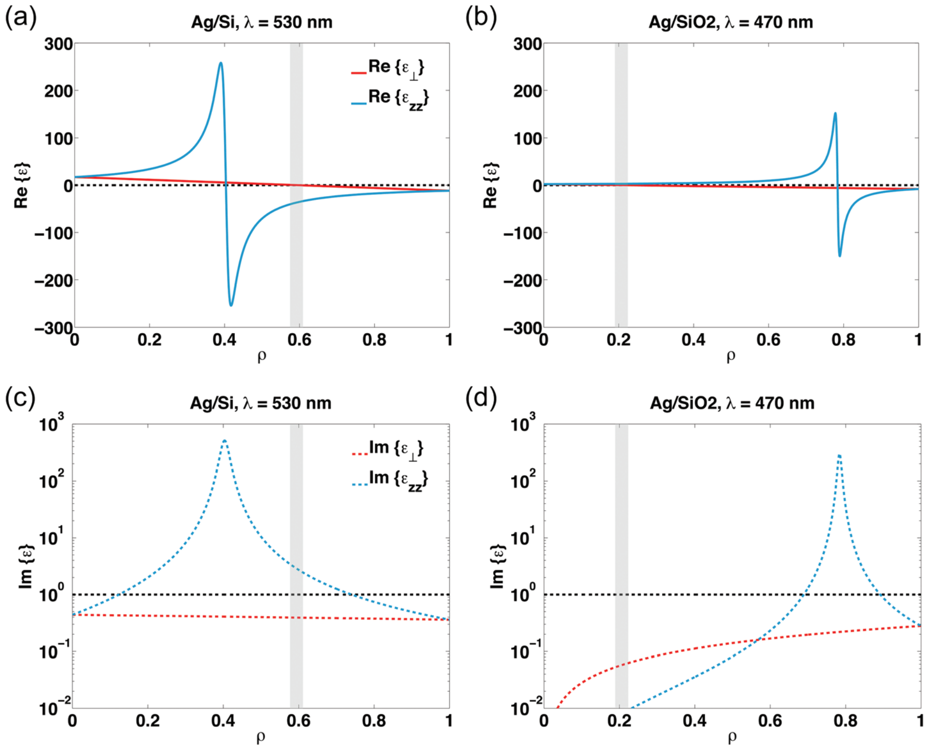

Fig. 11 Real (a, b) and imaginary (c, d, logarithmic scale) parts of the effective permittivity components as a function of the filling ratio ρ in (a, c) Ag/Si ML at λ = 530 nm, and in (b, d) Ag/SiO_2 ML at λ = 470 nm. The grey vertical stripe denotes the Re{$\varepsilon_\perp$}-near-zero region.

ing from the lower block edges closer to the dipole (lower lobe). At ρ_A instead, two lobes symmetrically departing from the dipole-ML block system indicate that the main light extraction mechanism is scattering from the upper and lower edges of the block, rather than outcoupling from the lateral faces.

C. Additional plots for the Ag/Si and Ag/SiO_2 systems

Fig. 11 shows the real and imaginary parts of the effective components for the Ag/Si and Ag/SiO_2 MLs discussed in the main text. Since the losses associated to the Re{$\varepsilon_\perp$}-near-zero region are low, it is possible to achieve side outcoupling. A Re{ε_{zz}}-near-zero region is also present, but it corresponds to a resonant pole of eqn (3): the related high loss prevents effective top outcoupling.

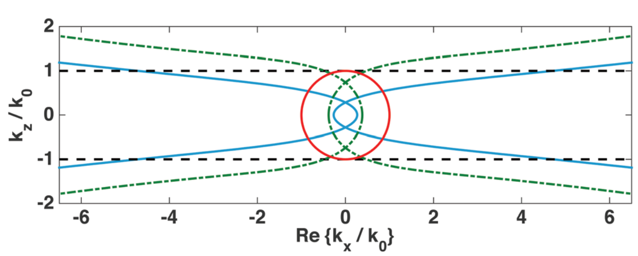

Fig. 12 Iso-frequency curve of the Ag/Si effective medium at λ_1 = 530 nm for different filling ratios ρ. Dash-dotted green curve: ρ_A = 0.54; solid blue curve: ρ_B = 0.58; red curve: iso-frequency curve of air.

Fig. 12 compares the effective medium iso-frequency curve of the Ag/Si ML discussed in the main text at two filling ratios ρ_A = 0.54 and ρ_B = 0.58. In both cases the effective medium exhibits a Type I behavior, distorted from the ideal case of Fig. 3(b) by the presence of loss; however, at ρ_A the portion of laterally outcoupled large-*k* waves is negligible, whereas at ρ_B, optimized for lateral extraction, the outcoupled large-*k* bandwidth becomes significant.

References

1 D. R. Smith and D. Schurig, *Phys. Rev. Lett.*, 2003, **90**, 077405.
2 A. J. Hoffman, L. Alekseyev, S. S. Howard, K. J. Franz, D. Wasserman, V. A. Podolskiy, E. E. Narimanov, D. L. Sivco and C. Gmachl, *Nat. Mater.*, 2007, **6**, 946–950.
3 J. Yao, Z. Liu, Y. Liu, Y. Wang, C. Sun, G. Bartal, A. M. Stacy and X. Zhang, *Science*, 2008, **321**, 930–930.
4 Y. Liu, G. Bartal and X. Zhang, *Opt. Express*, 2008, **16**, 15439–15448.
5 T. Xu, A. Agrawal, M. Abashin, K. J. Chau and H. J. Lezec, *Nature*, 2013, **497**, 470–474.
6 L. Ferrari, C. Wu, D. Lepage, X. Zhang and Z. Liu, *Prog. Quantum Electron.*, 2015, **40**, 1–40.
7 Z. Jacob, L. V. Alekseyev and E. Narimanov, *Opt. Express*, 2006, **14**, 8247–8256.

8 Z. Liu, H. Lee, Y. Xiong, C. Sun and X. Zhang, *Science*, 2007, **315**, 1686–1686.
9 J. Sun, M. I. Shalaev and N. M. Litchinitser, *Nat. Commun.*, 2015, **6**, 7201.
10 Y. Xiong, Z. Liu and X. Zhang, *Appl. Phys. Lett.*, 2009, **94**, 203108.
11 J. Sun, T. Xu and N. M. Litchinitser, *Nano Lett.*, 2016, **16**, 7905–7909.
12 D. Ji, H. Song, X. Zeng, H. Hu, K. Liu, N. Zhang and Q. Gan, *Sci. Rep.*, 2014, **4**, 4498.
13 C. T. Riley, J. S. T. Smalley, J. R. J. Brodie, Y. Fainman, D. J. Sirbuly and Z. Liu, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 1264–1268.
14 J. Shi, B. Liu, P. Li, L. Y. Ng and S. Shen, *Nano Lett.*, 2015, **15**, 1217–1221.
15 A. D. Neira, G. A. Wurtz, P. Ginzburg and A. V. Zayats, *Opt. Express*, 2014, **22**, 10987–10994.
16 Y. Sun, Z. Zheng, J. Cheng, G. Sun and G. Qiao, *Opt. Express*, 2015, **23**, 6370–6378.
17 Z. Jacob, J.-Y. Kim, G. Naik, A. Boltasseva, E. Narimanov and V. Shalaev, *Appl. Phys. B: Lasers Opt.*, 2010, **100**, 215–218.
18 M. A. Noginov, H. Li, Y. A. Barnakov, D. Dryden, G. Nataraj, G. Zhu, C. E. Bonner, M. Mayy, Z. Jacob and E. E. Narimanov, *Opt. Lett.*, 2010, **35**, 1863–1865.
19 O. Kidwai, S. V. Zhukovsky and J. E. Sipe, *Opt. Lett.*, 2011, **36**, 2530–2532.
20 J. Kim, V. P. Drachev, Z. Jacob, G. V. Naik, A. Boltasseva, E. E. Narimanov and V. M. Shalaev, *Opt. Express*, 2012, **20**, 8100–8116.
21 H. N. S. Krishnamoorthy, Z. Jacob, E. Narimanov, I. Kretzschmar and V. M. Menon, *Science*, 2012, **336**, 205–209.
22 D. Lu, J. J. Kan, E. E. Fullerton and Z. Liu, *Nat. Nanotechnol.*, 2014, **9**, 48–53.
23 L. Ferrari, D. Lu, D. Lepage and Z. Liu, *Opt. Express*, 2014, **22**, 4301–4306.
24 C. Guclu, T. S. Luk, G. T. Wang and F. Capolino, *Appl. Phys. Lett.*, 2014, **105**, 123101.
25 M. Y. Shalaginov, V. V. Vorobyov, J. Liu, M. Ferrera, A. V. Akimov, A. Lagutchev, A. N. Smolyaninov, V. V. Klimov, J. Irudayaraj, A. V. Kildishev, A. Boltasseva and V. M. Shalaev, *Laser Photonics Rev.*, 2015, **9**, 120–127.
26 T. Li and J. B. Khurgin, *Optica*, 2016, **3**, 1388–1396.
27 K. L. Tsakmakidis, R. W. Boyd, E. Yablonovitch and X. Zhang, *Opt. Express*, 2016, **24**, 17916–17927.
28 K. J. Russell, T.-L. Liu, S. Cui and E. L. Hu, *Nat. Photonics*, 2012, **6**, 459–462.
29 G. M. Akselrod, C. Argyropoulos, T. B. Hoang, C. Ciracì, C. Fang, J. Huang, D. R. Smith and M. H. Mikkelsen, *Nat. Photonics*, 2014, **8**, 835–840.
30 M. S. Eggleston, K. Messer, L. Zhang, E. Yablonovitch and M. C. Wu, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 1704–1709.
31 L. Novotny and B. Hecht, *Principles of Nano-Optics*, Cambridge University Press, Cambridge, 2nd edn., 2012.
32 G. P. Agrawal, *Fiber-Optic Communication Systems*, Wiley, Hoboken, 4th edn, 2010.
33 H. Chun, S. Rajbhandari, G. Faulkner, D. Tsonev, E. Xie, J. J. D. McKendry, E. Gu, M. D. Dawson, D. C. O'Brien and H. Haas, *J. Lightwave Technol.*, 2016, **34**, 3047–3052.
34 J. A. Simpson, B. L. Hughes and J. F. Muth, *IEEE J. Sel. Area. Comm.*, 2012, **30**, 964–974.
35 J. Vinogradov, R. Kruglov, K. L. Chi, J. W. Shi, M. Bloos, S. Loquai and O. Ziemann, *IEEE Photonics Technol. Lett.*, 2014, **26**, 2473–2475.
36 W. D. Newman, C. L. Cortes and Z. Jacob, *J. Opt. Soc. Am. B*, 2013, **30**, 766–775.
37 K. V. Sreekanth, K. H. Krishna, A. De Luca and G. Strangi, *Sci. Rep.*, 2014, **4**, 6340.
38 T. Galfsky, H. N. S. Krishnamoorthy, W. Newman, E. E. Narimanov, Z. Jacob and V. M. Menon, *Optica*, 2015, **2**, 62–65.
39 T.-H. Kao and Y.-C. Hung, *Proc. SPIE*, 2016, **9883**, 98830U.
40 P. R. West, N. Kinsey, M. Ferrera, A. V. Kildishev, V. M. Shalaev and A. Boltasseva, *Nano Lett.*, 2015, **15**, 498–505.
41 S. A. Maier, *Plasmonics: Fundamentals and Applications*, Springer, New York, 1st edn, 2007.
42 P. Shekhar, J. Atkinson and Z. Jacob, *Nano Convergence*, 2014, **1**, 14.
43 P. B. Johnson and R. W. Christy, *Phys. Rev. B: Solid State*, 1972, **6**, 4370–4379.
44 D. E. Aspnes and A. A. Studna, *Phys. Rev. B: Condens. Matter*, 1983, **27**, 985–1009.
45 A. Orlov, I. Iorsh, P. Belov and Y. Kivshar, *Opt. Express*, 2013, **21**, 1593–1598.
46 S. Campione, T. S. Luk, S. Liu and M. B. Sinclair, *J. Opt. Soc. Am. B*, 2015, **32**, 1809–1815.
47 A. Yariv and P. Yeh, *Photonics: Optical Electronics in Modern Communications*, Oxford University Press, Oxford, 6th edn, 2006.
48 J. S. T. Smalley, F. Vallini, S. A. Montoya, L. Ferrari, S. Shahin, C. T. Riley, B. Kanté, E. E. Fullerton, Z. Liu and Y. Fainman, *Nat. Commun.*, 2017, **8**, 13793.
49 I. Avrutsky, I. Salakhutdinov, J. Elser and V. Podolskiy, *Phys. Rev. B: Condens. Matter*, 2007, **75**, 241402.
50 J. S. T. Smalley, F. Vallini, B. Kanté and Y. Fainman, *Opt. Express*, 2014, **22**, 21088–21105.
51 J. S. T. Smalley, F. Vallini, S. Shahin, B. Kanté and Y. Fainman, *Opt. Mater. Express*, 2015, **5**, 2300–2312.
52 J. S. T. Smalley, F. Vallini, S. Montoya, E. E. Fullerton and Y. Fainman, *Proc. SPIE*, 2015, **9544**, 95440S.
53 J. S. Smalley, F. Vallini, S. Montoya, L. Ferrari, C. Riley, S. Shahin, B. Kante, E. E. Fullerton, Z. Liu and Y. Fainman, *Advanced Photonics 2016 (IPR, NOMA, Sensors, Networks, SPPCom, SOF)*, 2016, p. NoM3C.3.
54 S. Zhao, H. P. Nguyen, M. G. Kibria and Z. Mi, *Prog. Quantum Electron.*, 2015, **44**, 14–68.
55 A. Boltasseva and H. A. Atwater, *Science*, 2011, **331**, 290–291.

56 C. T. Riley, J. S. T. Smalley, K. W. Post, D. N. Basov, Y. Fainman, D. Wang, Z. Liu and D. J. Sirbuly, *Small*, 2016, **12**, 892–901.
57 A. K. Geim and I. V. Grigorieva, *Nature*, 2013, **499**, 419–425.
58 G. Clark, J. R. Schaibley, J. Ross, T. Taniguchi, K. Watanabe, J. R. Hendrickson, S. Mou, W. Yao and X. Xu, *Nano Lett.*, 2016, **16**, 3944–3948.
59 A. K. Geim and I. V. Grigorieva, *Nature*, 2013, **499**, 419–425.
60 S. Dai, Q. Ma, M. K. Liu, T. Andersen, Z. Fei, M. D. Goldflam, M. Wagner, K. Watanabe, T. Taniguchi, M. Thiemens, F. Keilmann, G. C. A. M. Janssen, S.-E. Zhu, P. Jarillo-Herrero, M. M. Fogler and D. N. Basov, *Nat. Nanotechnol.*, 2015, **10**, 682–686.
61 Y.-C. Chang, C.-H. Liu, C.-H. Liu, S. Zhang, S. R. Marder, E. E. Narimanov, Z. Zhong and T. B. Norris, *Nat. Commun.*, 2016, **7**, 10568.
62 G. T. Papadakis and H. A. Atwater, *Phys. Rev. B: Condens. Matter*, 2015, **92**, 184101.
63 S. Prayakarao, B. Mendoza, A. Devine, C. Kyaw, R. B. van Dover, V. Liberman and M. A. Noginov, *Appl. Phys. Lett.*, 2016, **109**, 061105.
64 C. Guclu, S. Campione and F. Capolino, *Phys. Rev. B: Condens. Matter*, 2012, **86**, 205130.
65 W. L. Barnes, *J. Mod. Opt.*, 1998, **45**, 661–699.