SCIENTIFIC REPORTS

Origin of Shape Resonance in Second-Harmonic Generation from Metallic Nanohole Arrays

Ben-Li Wang*, Rui Wang*, R. J. Liu, X. H. Lu, Jimin Zhao & Zhi-Yuan Li

Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China.

Second-harmonic generation (SHG) from periodic arrays of subwavelength rectangular air hole with various aspect ratios perforated in gold thin films can get resonantly enhanced for some specific geometric shapes. Here we clarify the physical origin of this shape resonance effect. A nonlinear coupled-mode theory is set up to solve energy conversion from fundamental wave (FW) mode to second-harmonic wave (SHW) mode within the nanoscale air hole. It reveals that several physical mechanisms, including the FW mode excitation amplitude, FW-SHW modal spatial overlap, FW-SHW mode phase mismatch, and SHW mode attenuation, are all geometric shape sensitive and altogether act to induce the SHG shape resonance effect. The theory agrees well with experimental observations and provides an accurate and complete explanation for the long-emphasized but elusive shape effect. The study may stimulate deeper insights to visualize general nonlinear nanophotonic processes and pave the way to engineering high-efficiency nonlinear nanophotonic devices.

Light propagation through subwavelength structures has exhibited a variety of interesting properties that are otherwise not observable in macroscopic materials and structures[1–13]. Experiments and theories have shown that the linear transmission intensities through sub-wavelength air-hole arrays perforated in metal thin films are sensitive to many factors, such as array periodicity, hole shape[8–11], hole size[14], incidence angle[15], and film thickness[16,17]. The physical mechanisms behind these linear-optical shape effect have been well established, where the excitation, transport, and coupling of surface plasmon polaritons have played the key role.

In addition to the fascinating properties of linear light transport, nonlinear optical properties, in particular, the second-harmonic generation (SHG) from metallic nanostructures has also been investigated extensively, such as those in single nanoparticle[18], bowtie nano-apertures and nano-antennaes[19,20], L-shape arrays[21], and the silver film with triangular air holes arrays[22]. Methods of enhancing the SHG were also studied widely, such as those by double plasmonic resonance[23], Fano resonance[24], and even nanocup structure[25]. Asymmetry and circular dichro-ism in optical SHG emission of gold nanostructures have also been studied[26–28]. Recently, it was further shown that geometric shape factors such as spacing between the double holes[29,30] and the aspect ratio (AR) of rectangular holes[31] critically determine the resonant enhancement of second harmonic generation (SHG) in these metallic nanostructures[29–33]. For instance, Ref. 31 reported an experimental observation that the SHG signal could vary by several orders of magnitude when the AR of rectangular holes was tuned and a remarkable resonance emerged at a specific value of AR. These studies clearly showed the essential role of shape effect on nonlinear optical interaction in metallic nanostructures. However, the fundamental mechanism of the resonantly enhanced SHG signal is so far not clear, as the physics is more complicated than in the linear-optical regime. To understand why shape effect occurs, slow light effect has been suggested in Ref. 31. However, further investigation showed that the SHG signal does not enhance in proportion to the reduction of the group velocity[34]. Therefore a clarified microscopic picture about the shape resonance effect is still lack and highly desired.

In this paper we investigate the explicit origin of shape effect in these nonlinear optical nanostructures both experimentally and theoretically. In our experiment, we measured the SHG from a periodic array of subwave-length rectangular air holes using femtosecond laser pulses and observed the dependence of SHG signal intensity on the AR of the holes. Particularly we measured the propagation time duration of the pulses through the air hole with an aim to clarify the slow light effect. In our theory, we present a nonlinear coupled-mode theory analysis based on Maxwell's equations that describes the interaction between fundamental wave (FW) and the second harmonic wave (SHW) within the nanoscale air holes. The theory shows that the transverse modal coupling

and conversion between FW and SHW modes and their transport characteristics play an essential role in the shape-resonant effect of SHG observed in our experiment and previous reports[31]. Our theory is of great help to draw a clear picture about how several physical effects, including the modal profile similarity, the local field enhancement, the phase mismatch, and modal attenuation, cooperate and compete in ultimately shaping the SHG features in these nonlinear metallic nanostructures.
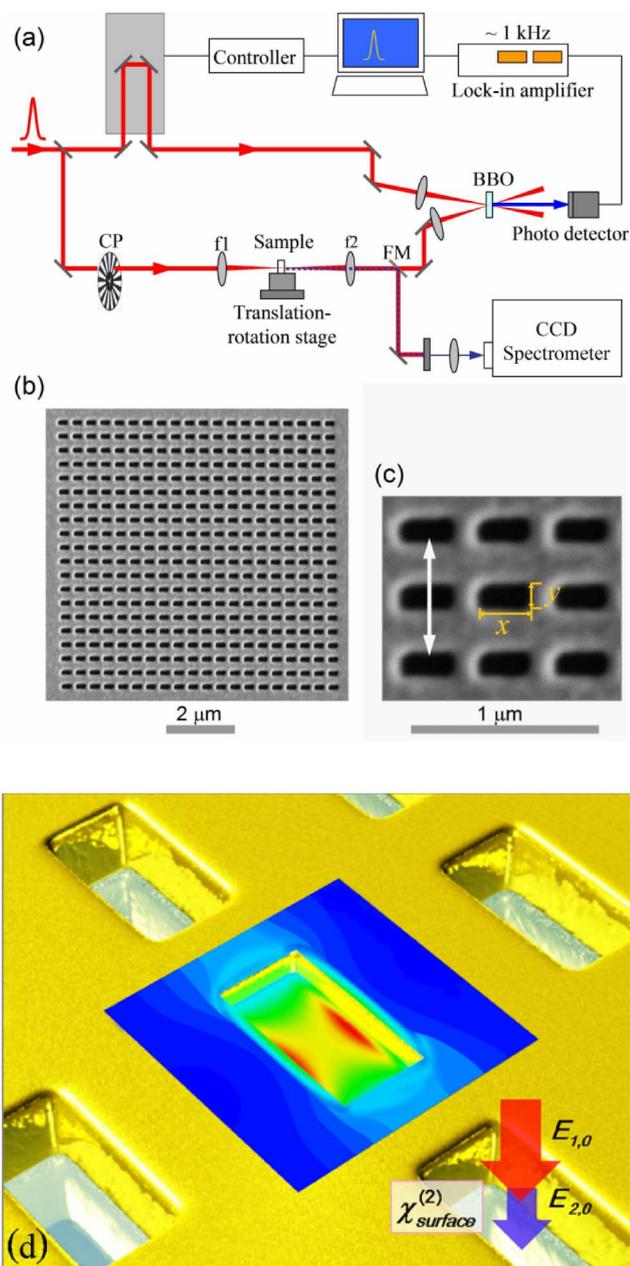

Figure 1 | Experiment on SHG from nanostructured gold thin film.
(a) The schematic experimental setup. FM: flip mirror; CP: chopper; BBO: nonlinear crystal; f1 and f2: focusing lens. (b) The scanning electron microscope image of a typical sample of the rectangular hole array, whereby AR = 2. The AR of the rectangular hole is defined as the ratio between the y-axis edge length (260 nm here) and the x-axis edge length (130 nm here). (c) A high resolution image, with the x-y direction and the laser polarization (parallel to the y-axis) marked on the figure. (d) The schematic physical picture illustrating generation and transport of SHW mode $\mathbf{E}_{2,0}(x,y)$ from surface gold atoms of the thin film under excitation of the FW mode $\mathbf{E}_{1,0}(x,y)$ within the air hole (which can be seen as a channel).

Results

The experimental setup for SHG measurement is depicted in Fig. 1(a). In our experiments, a series of rectangular hole array samples of different AR were prepared in a gold thin film of 160 nm thickness, where the 20×20 square-lattice array composes of rectangular holes with a spatial periodicity of 410 nm and a total area of $8 \ \mu m \times 8 \ \mu m$ (Fig. 1) (see Method, Sample Preparation). All the samples are designed to have the same area as $S = 260 \ nm \times 130 \ nm$. The scanning electron microscopy (SEM) picture of a typical sample (with AR = 2) is displayed in Figs. 1(b) and 1(c). The light spot on the sample had a diameter of about 30 μm. As depicted in Fig. 1(d), under the excitation of high-peak power FW laser pulse, the surface of gold thin film would generate SHW, which will propagate through the air hole and eventually radiate out from the thin film. When needed, each sample was rotated by 90° around the z axis to facilitate taking two AR-dependent data. The measurements of SHG waves (see Method, Measurement of SHG signals) and of time delay can be easily switched.

In Fig. 2(a) we show the observed SHG signal intensity as a function of the aspect ratio. A prominent peak is observed for an AR value around 2. In the inset we show a typical spectrum of the SHG signal for the AR = 2.037 sample. Our data are consistent with those reported in Ref. 31. The time delay in group velocity was measured
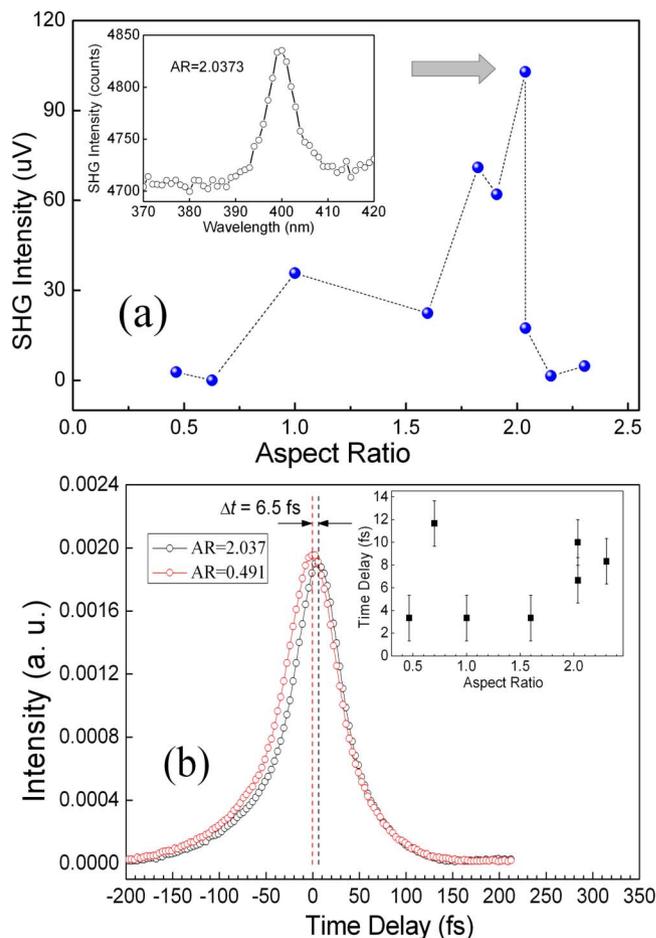

Figure 2 | Experimental data of SHG from different gold thin film samples. (a) The measured SHG signal intensity versus the aspect ratio of the rectangular hole. The inset is a typical spectrum of the SHG signal. A resonance peak is clearly observed for the AR = 2.037 sample.
(b) Convolution temporal profile of the light pulse after propagating through the AR = 2.037 and AR = 0.491 patterns, with a relative time-delay readout of 6.5 fs. The inset is the measured time delay versus aspect ratio, obtained by taking square vacant areas as references.

by cross-correlating the transmitted fundamental pulse with a reference pulse in a 1 mm-thick beta barium borate (BBO) crystal[35]. We averaged for 15 scans and the overall precision obtained became exceptionally high, which was about 2 fs (see Methods, Method to achieve ultrahigh temporal resolution).

To determine the absolute time delay, we compared the temporal traces through patterns with that through a large square hole on the Au film. A 8×8 μm² vacant area were milled on each Au film. In such a way additional effect caused by the spatial pulse tail is minimized. The results are shown in the inset of Fig. 2(b), where we have subtracted the peak positions for the patterns by those for the vacant areas, respectively. Note that the two data points around AR = 2 are for two different samples. For all the AR samples discernable time delay was clearly observed for the laser pulses after traversing through the metal nanohole array, as compared with those through the vacant window, indicating the existence of light slowdown. However, overall there does not exist a dramatic slowdown for any of the samples. Comparison of the value for AR = 0.5 with that for AR = 2.0 best tests the slow light effect, since most of the other factors such as the laser tail effect are absent for a rotated sample. Also for this reason, to elucidate a pure shape effect, it is better to compare the AR = 2 and AR = 0.5 cases, rather than to compare the AR = 2 case with air.

In Fig. 2(b) we show explicitly the temporal traces for the AR = 2.037 and AR = 0.491 patterns, respectively. Note that there is only a 6.5 fs delay between these two cases. The delay for the AR = 2.037 sample is 2 ∼ 3 times larger than that for the AR = 0.491 pattern [see inset of Fig. 2(b)]. However the bit rates of slowing-down in group velocity, $\Delta t_{delay}/\Delta t_{pulse\_duration}$, are 0.065 and 0.208 respectively, both of which are much less than 1. Therefore the nonlinear interaction time for the AR = 2.037 sample is not dramatically longer than that for the AR = 0.491 sample. Apparently the slowing-down exists [inset of Fig. 2(b)], however, it does not appear to be the major factor leading to the shape-dependent resonance. An alternative explanation for the fundamental mechanism of the nonlinear shape resonance effect must be found.

Maxwell's equations, being scale invariant, provide a secure basis for optical problems ranging from classical optics to fashioned nanophotonics, from linear optics to nonlinear optics. Therefore, they are naturally the starting point for accurately describing various nonlinear nanophotonic problems and surely will provide the ultimate answer to our puzzling problems of shape resonant SHG from metallic nanostructures. In conventional nonlinear optics for macroscopic materials two factors are crucial for generating nonlinear optical signals: energy conservation and phase matching (thus momentum conservation). However, for nanostructures, there might be other crucial factors that dictate the nonlinear optical processes but have been overlooked in conventional nonlinear optics.

Before we go into the details of the theoretical analysis, let us first draw a general physical picture about the SHG from the metallic nanostructure, which has been depicted in Fig. 1(d). The first thing is the source of SHG. SHG comes from a second-order nonlinear process originated from the surface atoms of nanostructured gold thin film, where the broken spatial inversion symmetry will lead to non-zero second-order nonlinear susceptibility. The illumination of high peak power femtosecond pulse laser upon these surface gold atoms will induce second-order nonlinear polarization, which will radiate SHW. The second thing is that in principle all the surfaces, including the outer surfaces (the front and back side of gold film) and the inner surfaces (the wall of each rectangular air hole), will contribute to SHG. However, as the linear electric field will be subject to a much larger enhancement effect within cavity-like nanoscale air holes than in the front and back side of the gold film, it is expected that the SHG within the air holes will make a dominant contribution to observed SHG signals, and this is the first stage (the generation process of SHW) where the geometric shape of air hole can come in

to play a role. The third thing is that the radiated SHG signals must be subject to the modulation by the air hole before they can ultimately transport and transmit through the nanostructure gold thin film. This is the second stage (the transport process of SHW) where the geometric shape of air hole comes in to play a role. At this stage, both the transverse modal profile and longitudinal modal dispersion of the FW and SHW within the air hole can make significant contributions. The final thing is that the detected SHG signal is just the summed contribution from all the SHW exiting the air hole and enduring zero-order Bragg diffraction. The above analysis has decomposed the whole complicated nanoscale SHG problem into several processes with the corresponding underlying physical mechanism being easy to understand. By this, the problem becomes much simpler to yield a much clarified picture, allowing the major physical mechanism of the puzzled shape resonance effect to be disclosed. However, the ultimate answer to this question again relies on quantitative solution and analysis based on rigorous Maxwell's equations.

According to the above argument, the SHG within each air hole may play the key role. Besides, as adjacent air holes do not have direct optical coupling due to strong attenuation of light within bulk gold materials, we can consider SHG within a single air hole. By the way, it is worthy to mention that the film thickness (160 nm) is comparable with that of the lateral dimension (260 nm × 130 nm) of the metallic hole. Thus it is safe to take the hole as a channel or waveguide for the light propagation. For these reasons, we start by applying Maxwell's equations to a single unit air hole taken as a waveguide (although with a very short length). It is assumed that the rectangular air hole has its edges parallel to the x- and y-axis, respectively, and the light transports along the z-axis. The transport of linear light through the air holes and the corresponding three-dimensional (3D) electric field distribution can be readily solved by 3D finite-difference time-domain (FDTD) method (see Method, Numerical simulations). The wave equation for the SHW field within the air hole can be written as

$$\nabla^2 \mathbf{E}_2 - \frac{\varepsilon(2\omega)}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{E}_2 = \frac{1}{\varepsilon_0 c^2}\frac{\partial^2}{\partial t^2}\mathbf{P}^{(2)}, \tag{1}$$

where the nonlinear polarization $\mathbf{P}^{(2)} = \chi^{(2)} : \mathbf{E}_1\mathbf{E}_1$ acts as the source term of SHW, and $\mathbf{E}_1$ and $\mathbf{E}_2$ denote FW and SHW fields, respectively. The nonlinear susceptibility $\chi^{(2)}$ is a second-rank tensor in general, which is non-zero only at the surfaces where spatial symmetry is broken. As the outer and inner surfaces of our gold nanostructure samples have an amorphous morphology (see Method, Sample preparation), the general second-rank tensor $\chi^{(2)}$ reduces to an isotropic quantity that contains only non-zero diagonal elements $\chi^{(2)}_{xxx} = \chi^{(2)}_{yyy} = \chi^{(2)}_{zzz} = \chi^{(2)}_a$, which is assumed to be uniform for all the surfaces. Thus we have $\mathbf{P}^{(2)} = \chi^{(2)}_a|\mathbf{E}_1|^2$. For none-depletion nonlinear optical processes, this can be seen as a perturbation term in the inhomogeneous differential equation Eq. (1).

In our solution we only consider the contribution from the dominant metallic-waveguide mode (with the smallest attenuation strength) for both FW and SHW fields. Other higher order modes have much larger attenuation strengths and are expected to make negligible influences to the solutions. With such approximation, the FW and SHW fields are written as

$$\mathbf{E}_1(x,y,z) = A_0\mathbf{E}_{1,0}(x,y)e^{i(k_{1,0}z - \omega t)}, \tag{2}$$

and

$$\mathbf{E}_2(x,y,z) = B_0(z)\mathbf{E}_{2,0}(x,y)e^{i(k_{2,0}z - 2\omega t)}, \tag{3}$$

respectively. It is assumed that the FW field is non-depleted and has a constant amplitude A_0, as in practice the fraction of FW power transformed to SHW power is negligible. $B_0(z)$, which is the amplitude of the SHW field to be solved, is a slowly-varying quantity with respect to the interaction length z. $\mathbf{E}_{1,0}(x,y)$ and $\mathbf{E}_{2,0}(x,y)$ are the

dimensionless transverse modal profiles of the lowest-attenuation FW and SHW waveguide modes with complex propagation wave vectors $k_{1,0} = k_{1,0}^r + i k_{1,0}^i$ and $k_{2,0} = k_{2,0}^r + i k_{2,0}^i$, respectively. These two modal functions are normalized in practical operation with its maximum modulus set to unity. Substituting $\mathbf{E}_1$, $\mathbf{E}_2$, and $\mathbf{P}^{(2)}$ into Eq. (1), and using the well-established slowly-varying amplitude approximation, we arrive at the following relation

$$2ik_{2,0}\mathbf{E}_{2,0}(x,y)\frac{dB_0(z)}{dz} = \frac{-4\omega^2\chi_a^{(2)}}{\varepsilon_0 c^2}\delta(x-x_S, y-y_S)A_0^2\mathbf{E}_{1,0}^2(x,y)e^{i\Delta kz}, \quad (4)$$

where $\Delta k = 2k_{1,0} - k_{2,0}$ denotes the phase mismatch of the lowest-attenuation FW and SHW waveguide modes. The function $\delta(x-x_S, y-y_S)$ means that only the inner wall surfaces of the air hole channel, whose positions are denoted by (x_S, y_S), make non-zero contributions to SHG signals.

The method of moment is employed to solve Eq. (4). Both sides of Eq. (4) are multiplied by the transverse modal function $\mathbf{E}_{2,0}^*(x,y)$ and then integrated in the transverse dimensions where the nonlinear susceptibility is non-zero. The following equation is obtained:

$$\frac{dB_0}{dz} = iD(\omega)A_0^2\Xi e^{i\Delta kz}, \quad (5)$$

where $D(\omega) = (2\omega^2\chi_a^{(2)})/(k_{2,0}\varepsilon_0 c^2)$, and Ξ is the physical quantity spatial overlap factor, which is defined as

$$\Xi = \iint_S (\mathbf{E}_{1,0}\mathbf{E}_{1,0}) : \mathbf{E}_{2,0}^* dxdy / \iint_S |\mathbf{E}_{2,0}|^2 dxdy. \quad (6)$$

Here S denotes the inner wall surface area of the air hole. The integrand is the product between the tensor polarization and the vector field of SHW. The normalized integration represents the degree of effective spatial overlap between the FW and SHW modal fields. Assuming the initial SHW field at $z = 0$ to be zero, we obtain the solution for Eq. (5) as

$$|B_0(z)| = \frac{2\omega^2\chi_a^{(2)}}{\varepsilon_0 c^2}\left|\frac{\sin(\Delta kz/2)}{k_{2,0}\Delta k/2}\right||A_0|^2|\Xi|. \quad (7)$$

Having obtained the amplitude of SHW waveguide mode, the intensity of SHW at the exit surface (the back surface of the gold thin film, located at $z = h$, where h is the gold film thickness) is then given by Eq. (3) and Eq. (7).

$$I_{SHG} \propto \iint_{rect} dxdy|\mathbf{E}_2(x,y,h)|^2 = |B_0(h)|^2 e^{-2k_{2,0}^i h}\iint_{rect} dxdy|\mathbf{E}_{2,0}(x,y)|^2$$

$$= \left[\frac{2\omega^2\chi_a^{(2)}}{\varepsilon_0 c^2}\right]^2\left|\frac{\sin(\Delta kh/2)}{k_{2,0}\Delta k/2}\right|^2|A_0|^4|\Xi|^2 e^{-2k_{2,0}^i h}\iint_{rect} dxdy|\mathbf{E}_{2,0}(x,y)|^2. \quad (8)$$

The integration of Eq. (8) is upon the whole cross section area of the rectangular air hole. This average intensity is proportional to the observed far-field SHG signal radiating from the periodic array of subwavelength air hole array.

Discussion

According to Eq. (8), the SHG signal is determined by several factors. The first factor is the local field enhancement effect described by the amplitude A_0 of the FW mode $\mathbf{E}_{1,0}(x,y)$. Note that this value does not directly correspond to the transmission power of FW, it just reflects the field within the hole channel. The second factor is the spatial overlap factor Ξ, which is closely relevant with the modal profile similarity between the FW and SHW waveguide modes that take part in the nonlinear optical process of SHG. The third factor is the phase mismatch between the FW and SHW modes, which can also make a strong influence due to the strong dispersion nature of the metallic-waveguide mode, even if the thickness of the gold films is several times smaller than the wavelength. The fourth one is the

attenuation strength of the SHW mode described by the quantity $k_{2,0}^i$. In essence, all these four factors can be closely dependent on the geometric shape of the air hole. The nonlinear coupled-wave theory has confirmed the validity of our above general description of the physical picture of SHG from these nonlinear nanophotonic structures, and has quantified the contribution of each relevant physical mechanism.

It is helpful to compare our theory with previous investigations in Ref. 31. In that work, the effective nonlinear response from the air-hole array is enhanced by more than one order of magnitude for a certain particular value of aspect ratio. This enhancement was merely attributed to the influence of the hole shape on the linear transmission properties, in particular, the slow propagation of the fundamental wavelength through the holes which occurs close to the hole cutoff. Other important characteristics on the nonlinear modal conversion between FW and SHW, as have been clearly analyzed in our theory, were not accounted for explicitly in Ref. 31.

We first determined the characteristics of the lowest-attenuation FW and SHW waveguide mode, including their transverse modal profile function [$\mathbf{E}_{1,0}(x,y)$ and $\mathbf{E}_{2,0}(x,y)$] and their dispersion, namely, the complex propagation wave vectors $k_{1,0}$ and $k_{2,0}$ as a function of the air hole AR. This was done via the 3D FDTD numerical calculations. The dispersive nature of gold material was taken into account[36]. The simulation results are displayed in Fig. 3. In Fig. 3(a) and 3(b), the real and imaginary part of $k_{1,0}$ ($k_{1,0} = k_{1,0}^r + i k_{1,0}^i$) and $k_{2,0}(k_{2,0} = k_{2,0}^r + i k_{2,0}^i)$ are plotted for the 800 nm FW and 400 nm SHW. For the FW, $k_{1,0}^r$ first keeps at a low level when AR increases from 0.5, then grows rapidly when it reaches around AR = 2.2. As $k_{1,0}^r$ describes the phase factor and is related to the phase index, the result might be a signature of the ignition of slow transport mode at AR = 2.2, which has been assumed in Ref. 31 as the major reason responsible for the shape resonance of SHG. However, $k_{1,0}^i$ keeps decreasing (almost linearly) when AR grows. Thus the variation feature of $k_{1,0}^i$ does not match the picture of slow transport mode because in this dissipative material the ignition of slow mode usually means a greatly enhanced dissipation of the mode. The simulation result of linear light dispersion properties, together with the measurement result of light pulse delay depicted in Fig. 2(b), strongly indicates that there is not obvious slow transport mode for FW in the rectangular hole with various AR. The dispersion curves of SHW, as depicted in Fig. 3(b) shows a very different behavior from that of FW. When the hole AR grows, $k_{2,0}^r$ continuously increases without a distinct jump, while $k_{2,0}^i$ decreases very rapidly. The results clearly show that the geometric shape, reflected in the aspect ratio, has a significant influence on the transport properties of both FW and SHW. The attenuation is much reduced for large AR than for small AR. This is not difficult to understand. As the incident FW is parallel to the y-axis, a larger AR means a larger x-axis space of the rectangular hole witnessed by the light, and therefore a smaller attenuation due to waveguide cutoff.

Our numerical calculations show that the prominent field components of the waveguide mode $\mathbf{E}_1(x,y,z)$ and $\mathbf{E}_2(x,y,z)$[see Eqs. (2) and (3)] under the excitation of the y-axis polarization incident light is the E_y field, while other two field components are much smaller, similar to the incident light. As a result, for simplicity we may consider only the E_y field to describe $\mathbf{E}_1(x,y,z)$ and $\mathbf{E}_2(x,y,z)$. Figures 3(c)–(f) display the E_y-field amplitude pattern of waveguide mode with wavelengths 800 nm and 400 nm at the entrance of the air hole for two specific hole shapes with AR = 2.3 and AR = 1.6. Note that a transverse mode has small variation of phase within the air hole area. The incident light has unity amplitude. These field patterns should correspond to $\mathbf{E}_1(x,y,0)$ and $\mathbf{E}_2(x,y,0)$and they involve the information of both the transverse mode profile functions [$\mathbf{E}_{1,0}(x,y)$ and $\mathbf{E}_{2,0}(x,y)$] and the FW mode amplitude A_0. It can be seen that the modal profile for the two hole shapes are similar for
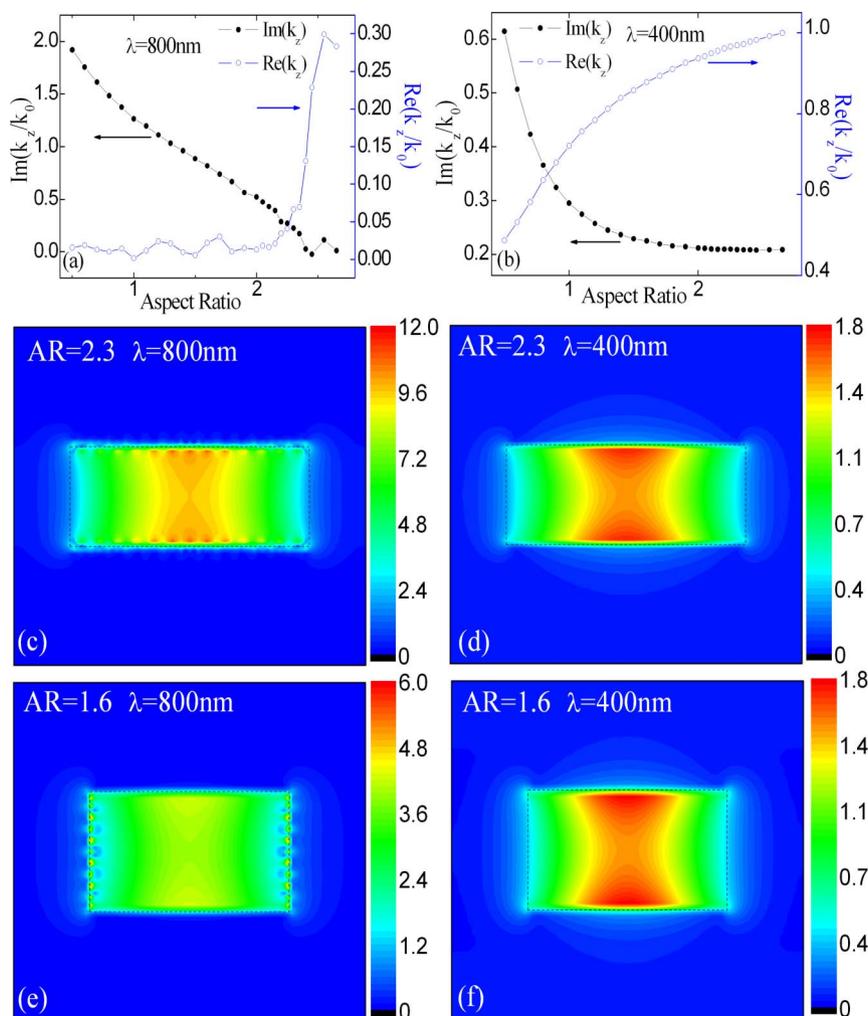
Figure 3 | **Physical properties of the FW and SHW waveguide modes.** (a) Calculated imaginary part and the real part of the propagation constant at the wavelength of (a) 800 nm and (b) 400 nm, respectively. (c)–(f) E_y-field amplitude pattern of at the wavelength of 800 nm [for the FW mode $\mathbf{E}_{1,0}(x,y)$] and 400 nm [for the SHW mode $\mathbf{E}_{2,0}(x,y)$] for air hole with AR = 2.3 and 1.6. The dotted rectangular denotes the air hole boundary. The incident electric field is y-polarized and the dominant component of electric field is E_y. All the field intensities are in unit of the incident field intensity.

both the FW and SHW modes. Besides, the modal profile for FW and SHW does not differ much at the same AR value either, although FW and SHW seem to have slightly higher similarity for AR = 2.3 than for AR = 1.6. All the four modes have the maximum field intensity located at the center of the upper and lower x-axis edges of the hole. However, there is a prominent difference: The FW field amplitude is much stronger for AR = 2.3 than for AR = 1.6, indicating a much higher field enhancement of the FW mode for the AR = 2.3 hole.

The above knowledge on the FW and SHW waveguide mode properties allows us to evaluate quantitatively the SHG signal intensity from various hole array samples. The calculation results based on Eq. (7) and Eq. (8) are summarized in Fig. 4 for a series of nanostructured gold thin films with various air hole AR. Figure 4(e) illustrates the amplitude $B_0(h)$ of SHW waveguide mode at the exit plane of the thin film, which originated from the excitation by FW and evolution of SHW within the hole. Figure 4(f) illustrates the corresponding intensity of SHG signal emitting from the gold thin film. In these two figures, the results have been normalized to the value at AR = 0.5 (which is set to 1). The calculated amplitude of SHW mode and SHG signal intensity all show a prominent resonant peak at AR = 2.3. And there is a three-orders-of-magnitude variation of the SHG signal intensity for different AR samples. The theoretical result agrees well with our experimental data as shown in Fig. 2, as well as with those in Ref. 31.

The theory developed in the last section has pointed out several factors that contribute to the overall SHG signals. The calculation results on the contribution from the following physical factors: (a) The amplitude of FW mode $|A_0|$; (b) the spatial overlap factor $|\Xi|$; (c) the FW-SHW phase mismatch factor $|\sin(\Delta k h/2)/(k_{2,0}\Delta k/2)|$; and (d) the SHW attenuation factor $\exp(-k_{2,0}^i h)$ for various hole AR values are displayed in Figs. 4(a)–4(d), respectively. Figure 4(a) clearly shows that the amplitude of FW mode $|A_0|$ almost grows monotonously (accompanied with some fluctuations) with the increasing AR of air hole, and the peak takes place at AR = 2.3 with the magnitude of 6.7 (normalized with the amplitude at AR = 0.5). It is clear from Eq. (7) and Eq. (8) that the SHW mode amplitude is $B_0 \propto |A_0|^2$, while the SHG signal intensity is $I_{SHG} \propto |A_0|^4$. As a result for this fourth power law, the modest enhancement on the excited FW mode amplitude $|A_0|$ will lead to three orders of magnitude enhancement in SHG signal intensity.

The FW-SHW spatial overlap factor $|\Xi|$ follows some complicated variation but overall it fluctuates around a high level of 0.7 for all the hole shapes (AR values). It can be said that this spatial overlap factor does not play an important role in shaping the SHG signal intensity for different hole shapes in the current nonlinear nanophotonic system. The reason is because that the modal profile of FW and SHW modes does not change much for the current air hole geometry. However, this factor might play a dominant role in other systems
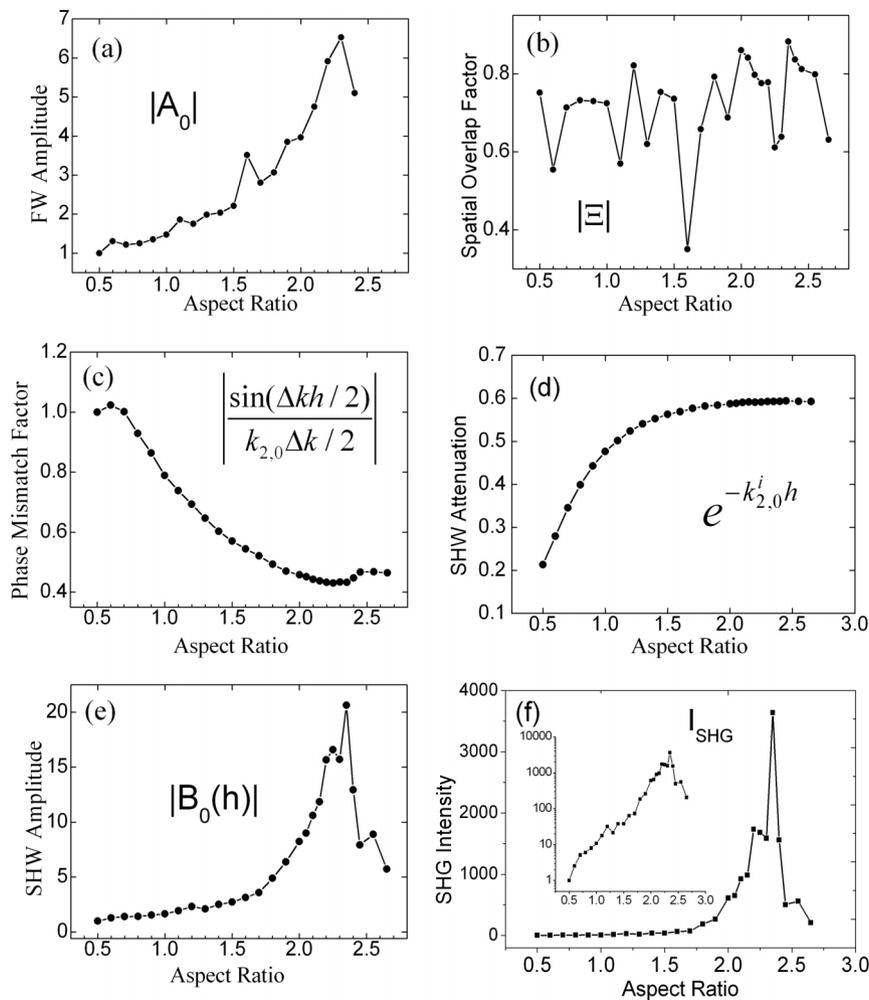
Figure 4 | Theoretical results of SHG signals as a function of the air-hole aspect ratio. Panels (a)–(d) show the contribution from various physical factors to SHG: (a) The amplitude of FW mode A_0; (b) the spatial overlap factor $|\Xi|$; (c) the FW-SHW phase mismatch factor; and (d) the SHW attenuation factor. Panels (e) and (f) display the calculated amplitude of SHW mode and SHG intensity (linear scale, with the corresponding semi-log scale plot depicted in the inset), respectively, which show a prominent resonant peak at AR = 2.3 and a three orders of magnitude variation of the SHG intensity for different AR samples with a total area of $S = 260$ nm $\times$ 130 nm.

where the FW and SHW modes have a very different modal profile and each of them depends sensitively on the geometry of the non-linear nanostructures. The above two factors shown in Fig. 4(a) and 4(b) can be attributed to the transverse modal profile of FW mode and SHW mode, with $|A_0|$ being related to the coupling efficiency of FW mode with the incident light, while $|\Xi|$ being related to the conversion efficiency from FW to SHW. It can be said that these two factors reflect the local interaction property of FW with metallic nanostructures.

In contrast, the other two factors shown in Fig. 4(c) and 4(d) are closely related with the dispersion nature of FW and SHW, namely with the subtle way how these two modes transport along the air hole channel. The FW-SHW phase mismatch factor $|\sin(\Delta kh/2)/(k_{2,0}\Delta k/2)|$ [Fig. 4(c), all normalized to that at AR = 0.5], which is well known to play a key role in conventional macroscopic nonlinear optics, e.g., in nonlinear crystals, is also very important in current nonlinear nanophotonic system. The phase mismatch factor almost decreases monotonously with increasing value of AR, with the value at AR = 2.3 about 2 times smaller than the value at AR = 0.5. The SHW attenuation factor $\exp(-k_{2,0}^i h)$ show an opposite variation behavior to that of the phase mismatch factor. It grows monotonously with increasing AR value, and gradu-ally reaches a saturation plateau when AR reaches 2.0. The value at

AR = 2.3 is about 2 times larger than the value at AR = 0.5. Obviously the SHW attenuation factor also plays an important role in SHG from the air hole channel. It can be said that the above two factors reflect the global transport property of FW and SHW. Although they are all important for SHG, they counteract and cancel with each other, leaving the overall influence from the global trans-port property not a significant factor for the ultimate intensity of the observed SHG signal in the current nonlinear nanophotonic system. However, this might be an accidental event. In many other systems they altogether can be a significant factor in determining the intens-ity of SHG.

There is some discrepancy between theory and experiment (see Fig. 2 and Fig. 4) in regard to the hole shape, i.e., AR value (2.3 in theory and 2.0 in experiment) at which the maximum SHG signal occurs. We attribute this discrepancy to the inconsistence in the geometric shape of practical samples with those adopted in numer-ical simulation. In simulation, we assumed perfect rectangular air holes with strictly vertical walls etched into a strictly perfect gold thin film. In practice, there exists considerable surface roughness (gold islands) at the inner and outer surfaces, the size of air hole cannot be perfectly matched with preset value due to the inaccuracy of FIB milling, there exists over-etching of gold in the front and back exit of the air hole, and the wall of air hole can be slightly inclined with

respect to the surface of gold thin film. All these deviations in the geometric shape of practical hole from the modeled one would modify the waveguide mode properties and change the way how the SHG signal depends on the AR. The above assumptions have been confirmed by our further numerical calculations for different values of rectangular hole area, which are illustrated in Fig. 5. Three values of air hole area are further considered, $S = 260$ nm $\times$ 140 nm, 270 nm $\times$ 140 nm, and 260 nm $\times$ 120 nm. Comparison with the situation of $S = 260$ nm $\times$ 130 nm (Fig. 4) shows that the maximum SHG intensity takes place at smaller AR for larger air hole area, and at larger AR for smaller air hole area. For $S = 270$ nm $\times$ 140 nm, the SHG resonance happens at AR = 2.0, which is very close to our experimental data shown in Fig. 2(a). In addition, the maximum SHG signal intensity also depends on the hole area. These results again strongly support the ideology that geometric shape plays a key role in SHG from nonlinear nanophotonic structures.

Our above analysis has clarified the contribution from several relevant physical factors, in which the local FW field enhancement overall dominates the SHG signals, It is worthwhile to compare our findings with previous literatures, in particular Ref. 31. To account
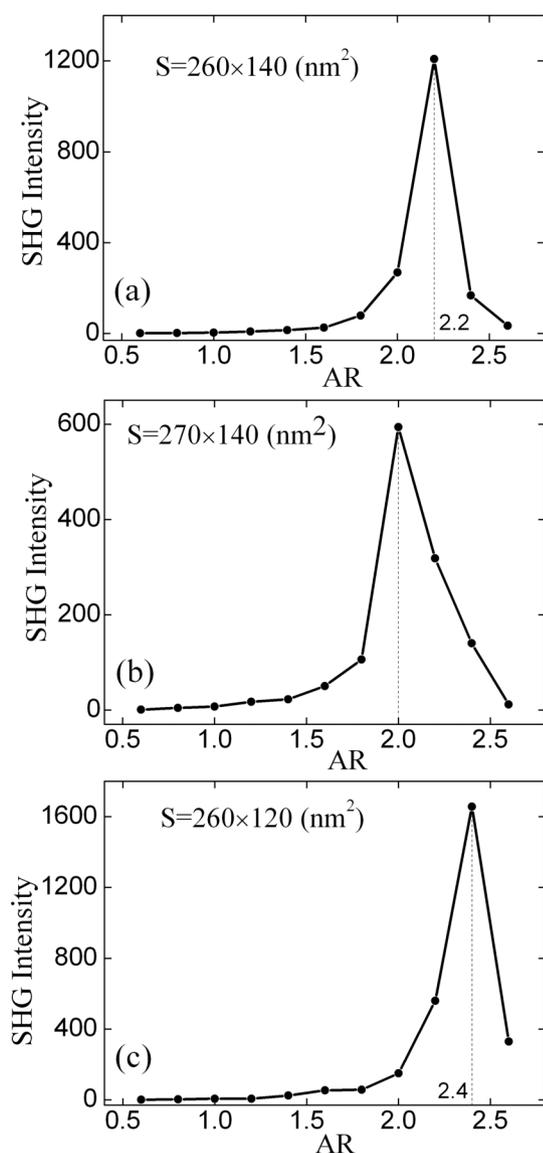
for the beautiful experimental data of enhanced SHG, the slow light effect was employed, which leads to an enhanced effective nonlinear susceptibility of the hole array gold thin film that is adopted to describe phenomenologically the overall SHG response of the sample. Although the slow light effect can contribute to local field enhancement, local field enhancement can be induced by many other mechanisms. As both the FW mode band diagram calculation and the pulse delay experiment have excluded the existence of a significant slow light effect, we assume that the higher coupling efficiency of external FW light energy into the FW waveguide mode plays the dominant role. In addition, the spatial overlap factor $|\Xi|$ and the SHW attenuation factor were not involved in the model considered in Ref. 31, and the FW-SHW phase mismatch factor was considered, but deemed to negligible due to the small film thickness compared with the wavelength. According to our analysis, these three factors each can make considerable effect on the SHG.

In summary, we have theoretically and experimentally investigated the physical origin of SHG shape resonance effect in nonlinear nanophotonic structures made from periodic arrays of rectangular air hole with various aspect ratios perforated in gold thin film. We have developed a nonlinear coupled mode theory to account for the energy conversion from FW mode to SHG mode involved in the SHG process within the nanoscale channel. The theory is complete and accurate as it is based on Maxwell's equations for nonlinear nanophotonic structures. The complete theory reveals that several physical mechanisms, including the FW mode excitation amplitude, the FW-SHW modal spatial overlap, and FW-SHW mode phase mismatch, and the SHW mode attenuation, can make significant contribution to the SHG signal intensity. They are all geometric shape sensitive and altogether act to induce the shape resonance effect of SHG from these nanostructured gold thin films. In short, the SHG shape resonance can be ascribed to the strongly shape dependent intrinsic phase and attenuation feature of the FW and SHW waveguide mode and the external excitation efficiency of the FW mode.

The theory agrees well with our experimental observations of shape dependent SHG and also with those reported in previous literatures that the SHG signal intensity can be three orders of magnitude larger at a specific hole shape than at other shapes. Both of our ultrafast pulse delay experiment and guided mode band diagram calculation exclude the long-held model that slow light effect would play a key role in such shape resonance of SHG. Our theory provides an accurate explanation for the long-emphasized but elusive shape effect in this puzzling problem. It may stimulate new angles and deeper insights to visualize general nonlinear nanophotonic processes and pave the way to engineering high-efficiency nonlinear nanophotonic structures. Our finding would inspire a very useful insight into other general nanoscale nonlinear optical problems that have been inconspicuous in conventional macroscopic nonlinear optics.

Methods

Sample preparation. Gold thin film of 160 nm thickness was deposited onto BK7 glass substrates, with a 5 nm thick Ti adhesion layer, by means of magnetron sputtering method. The gold thin film is thus not a single crystal, but composed of poly-crystalline gold where domain boundaries exist on and right below all surfaces. A 20 $\times$ 20 square-lattice array of rectangular holes were milled on these films by means of focused ion beam milling method, with a spatial periodicity of 410 nm and a total area of 8 μm $\times$ 8 μm (Fig. 1). In all the samples each hole is designed to have the same area. Due to the practical limitation of nanofabrication accuracy, the outer and inner surfaces of the nanostructured samples are not perfect atomic-level flat surface patterned with single-crystalline gold atom lattice, but involve many nanometer-scale roughened gold islands with different crystalline orientations (Fig. 1c and 5e). Thus all surfaces of the film all possess the necessary spatial inversion symmetry breaking for SHG. In addition, all the surfaces have an amorphous morphology, leading to a uniform and isotropic second-order susceptibility for the surface gold atoms.

Measurement of SHG signals. We used femtosecond laser pulse to investigate the linear and nonlinear light propagation. Laser (Coherent Inc.) pulses of 800 nm central wavelength, 80 MHz repetition rate, and 45 fs pulse duration were focused onto the sample with a f = 50.8 mm lens. The light spot on the sample had a diameter

Figure 5 | Calculated SHG intensity as a function of air-hole aspect ratio for different hole areas. The SHG intensity has been normalized to the value at AR = 0.6.

of about 30 μm. The transmitted FW and SHW passed through a $f = 63.5$ mm focusing lens. The samples were mounted onto a manual translation-rotation stage, with the y direction always parallel to the laser polarization. A flip mirror was employed to switch between the SHG and time delay measurements. For the SHG measurement a 20 mW average laser power was used, with the flip mirror in use. A low-pass edge filter with a transmission of $T = 6 \times 10^{-9}$ for the FW beam and T = 98% for the SHG signal was used.

Method to achieve ultrahigh temporal resolution. By averaging multiple scans, the temporal position of the cross-correlation peaks became very stable. The temporal resolving capability, or the experimental precision Δ, is explicitly determined by

$$\Delta_{precision} = \max\left(\delta_{resolution}, \sqrt{\left(\delta_{accuracy}\right)^2 + \left(\delta_{repeatability}\right)^2 / N} \right),$$ where δ are various systematic errors, and N is the number of scanning. Our linear stage has a resolution of 0.02 μm, accuracy of $+/-$ 0.3 μm, and repeatability of $+/-$ 0.1 μm. The unevenness of the substrate thickness was also considered and measured carefully, which does not bring in a noticeable system error in time delay. Furthermore the method is insensitive to the pulse-to-pulse jittering and can survive the pulse duration broadening. These are curial for obtaining a decisive conclusion for such a precise experiment. Overall the total experimental error is about ± 2 fs.

Numerical simulations. The 3D-FDTD method was used for the calculation of electric fields. The period of the structure was kept as 410 nm. The source was set as y-polarized in all calculations. The wavelength of FW is 800 nm. We used periodic boundary conditions along x- and y-directions, and perfectly matched layers (PML) were added along z-direction. The grid size is set as 5 nm. The lowest-attenuation waveguide mode of FW and SHW was determined by allowing light transporting within a deep rectangular hole. The complex propagation wave vector ($k_{1,0}$ and $k_{2,0}$) could be retrieved by comparing the phase and amplitude between two appropriately selected adjacent cross-sectional planes (both sufficiently far away from the entrance of the hole) with a small separation distance, while the transverse modal profile function $[\mathbf{E}_{1,0}(x,y)$ and $\mathbf{E}_{2,0}(x,y)]$ could be obtained by recording the field profile in one plane.

1. Ebbesen, T. W., Lezec, H. J., Ghaemi, H. F., Thio, T. & Wolff, P. A. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature* **391**, 667–669 (1998).
2. Merlin, R. Radiationless electromagnetic interference: evanescent-field lenses and perfect focusing. *Science* **317**, 927–929 (2007).
3. Oulton, R. F., Sorger, V. J., Genov, D. A., Pile, D. F. P. & Zhang, X. A hybrid plasmonic waveguide for subwavelength confinement and long-range propagation. *Nat. Photonics* **2**, 496–500 (2008).
4. Kinkhabwala, A. *et al.* Large sing-molecule fluorescence enhancements produced by a bowtie nanoantenna. *Nat. Photonics* **3**, 654–657 (2009).
5. Chang, C. W. *et al.* Optical möbius symmetry in metamaterials. *Phys. Rev. Lett.* **105**, 235501 (2010).
6. Sandtke, M. & Kuipers, L. Slow guided surface plasmons at telecom frequencies. *Nat. Photonics* **1**, 573–576 (2007).
7. Palomba, S. *et al.* Optical negative refraction by four-wave mixing in thin metallic nanostructures. *Nat. Materials* **11**, 34–38 (2012).
8. Genet, C. & Ebbesen, T. W. Light in tiny holes. *Nature* **445**, 39–46 (2007).
9. Gordon, R. *et al.* Strong polarization in the optical transmission through elliptical nanohole arrays. *Phys. Rev. Lett.* **92**, 037401 (2004).
10. Koerkamp, K. J. K., Enoch, S., Segerink, F. B., van Hulst, N. F. & Kuipers, L. Strong influence of hole shape on extraordinary transmission through periodic arrays of subwavelength holes. *Phys. Rev. Lett.* **92**, 183901 (2004).
11. Molen van der, K. L. *et al.* Role of shape and localized resonances in extraordinary transmission through periodic arrays of subwavelength holes: experiment and theory. *Phys. Rev. B* **72**, 045421 (2005).
12. Novotny, L. & Hecht, B. *Principles of Nano-Optics* (Cambridge University Press, 2006).
13. Linden, S. *et al.* Collective effects in second-harmonic generation from split-ring-resonator arrays. *Phys. Rev. Lett.* **109**, 015502 (2012).
14. Degiron, A. & Ebbesen, T. W. The role of localized surface plasmon modes in the enhanced transmission of periodic subwavelength apertures. *J. Opt. A: Pure Appl. Opt.* **7**, S90–S96 (2005).
15. Altewischer, E., van Exter, M. P. & Woerdman, J. P. Plasmon-assisted transmission of entangled photons. *Nature* **418**, 304–306 (2002).
16. Degiron, A., Lezec, H. J., Barnes, W. L. & Ebbesen, T. W. Effects of hole depth on enhanced light transmission through subwavelength hole arrays. *Appl. Phys. Lett.* **81**, 4327 (2002).
17. Martín-Moreno, L. *et al.* Theory of extraordinary optical transmission through subwavelength hole Arrays. *Phys. Rev. Lett.* **86**, 1114–1117 (2001).
18. Butet, J. *et al.* Optical second harmonic generation of single metallic nanoparticles embedded in a homogeneous medium. *Nano Lett.* **10**, 1717–1721 (2010).
19. Park, S., Hahn, J. W. & Lee, J. Y. Doubly resonant metallic nanostructure for high conversion efficiency of second harmonic generation. *Opt. Express* **20**, 4856–4870 (2012).
20. Ko, K. D. *et al.* Nonlinear optical response from arrays of Au bowtie nanoantennas. *Nano Lett.* **11**, 61–65 (2011).
21. Husu, H. *et al.* Metamaterials with tailored nonlinear optical response. *Nano Lett.* **12**, 673–677 (2012).
22. Salomon, A., Zielinski, M., Kolkowski, R., Zyss, J. & Prior, Y. Size and shape resonances in second harmonic generation from silver nanocavities. *J. Phys. Chem. C* DOI: 10.1021/jp403010q (2013).
23. Thyagarajan, K., Rivier, S., Lovera, A. & Martin, O. J. F. Enhanced second-harmonic generation from double resonant plasmonic antennae. *Opt. Express* **20**, 12860–12865 (2012).
24. Thyagarajan, K., Butet, J. & Martin, O. J. F. Augmenting second harmonic generation using fano resonances in plasmonic systems. *Nano Lett.* **13**, 1847–1851 (2013).
25. Zhang, Y., Grady, N. K., Ayala-Orozco, C. & Halas, N. J. Thress-dimensional nanostructures as highly efficient generators of second harmonic light. *Nano Lett.* **11**, 5519–5523 (2011).
26. Valev, V. K. *et al.* Asymmetric optical second-harmonic generation from chiral g-shaped gold nanostructures. *Phys. Rev. Lett.* **104**, 127401 (2010).
27. Valev, V. K. *et al.* Plasmonic ratchet wheels: switching circular dichroism by arranging chiral nanostructures. *Nano Lett.* **9**, 3945–3948 (2009).
28. Belardini, A., Larciprete, M. C., Centini, M., Fazio, E. & Sibilia, C. Circular dichroism in the optical second-harmonic emission of curved gold metal nanowires. *Phys. Rev. Lett.* **107**, 257401 (2011).
29. Lesuffleur, A., Kumar, L. K. S. & Gordon, R. Enhanced second harmonic generation from nanoscale double-hole arrays in a gold film. *Appl. Phys. Lett.* **88**, 261104 (2006).
30. Lesuffleur, A., Kumar, L. K. S. & Gordon, R. Apex-enhanced second-harmonic generation by using double-hole arrays in a gold film. *Phys. Rev. B* **75**, 045423 (2007).
31. van Nieuwstadt, J. A. H. *et al.* Strong modification of the nonlinear optical response of metallic subwavelength hole arrays. *Phys. Rev. Lett.* **97**, 146102 (2006).
32. Airola, M., Liu, Y. & Blair, S. Second-harmonic generation from an array of sub-wavelength metal apertures. *J. Opt. A: Pure Appl. Opt.* **7**, S118–S123 (2005).
33. Xu, T., Jiao, X., Zhang, G. P. & Blair, S. Second-harmonic emission from sub-wavelength apertures: effects of aperture symmetry and lattice arrangement. *Opt. Express* **15**, 13894–13906 (2007).
34. Prangsma, J. C., van Oosten, D., Moerland, R. J. & Kuipers, L. Increase of group delay and nonlinear effects with hole shape in subwavelength hole arrays. *New J. Phys.* **12**, 013005 (2010).
35. Prineas, J. P., Johnston, W. J., Yildirim, M., Zhao, J. & Smirl, A. L. Tunable slow light in bragg-spaced quantum wells. *Appl. Phys. Lett.* **89**, 241106 (2006).
36. Palik, E. D. *Handbook of Optical Constants of Solids* (Academic, Orlando, Florida, USA, 1985).

Acknowledgements

Author contributions

J.Z. and Z.L. conceived the idea, supervised the project and wrote the paper. R.W. and J.Z. did the experiment. Z.L. developed the theory. B.W. and Z.L. did the calculation. X.L. helped on experiment. R.L. helped on calculation. J.Z. and Z.L. together analyzed data and explained results. All authors discussed the results and commented on the manuscript.

Additional information