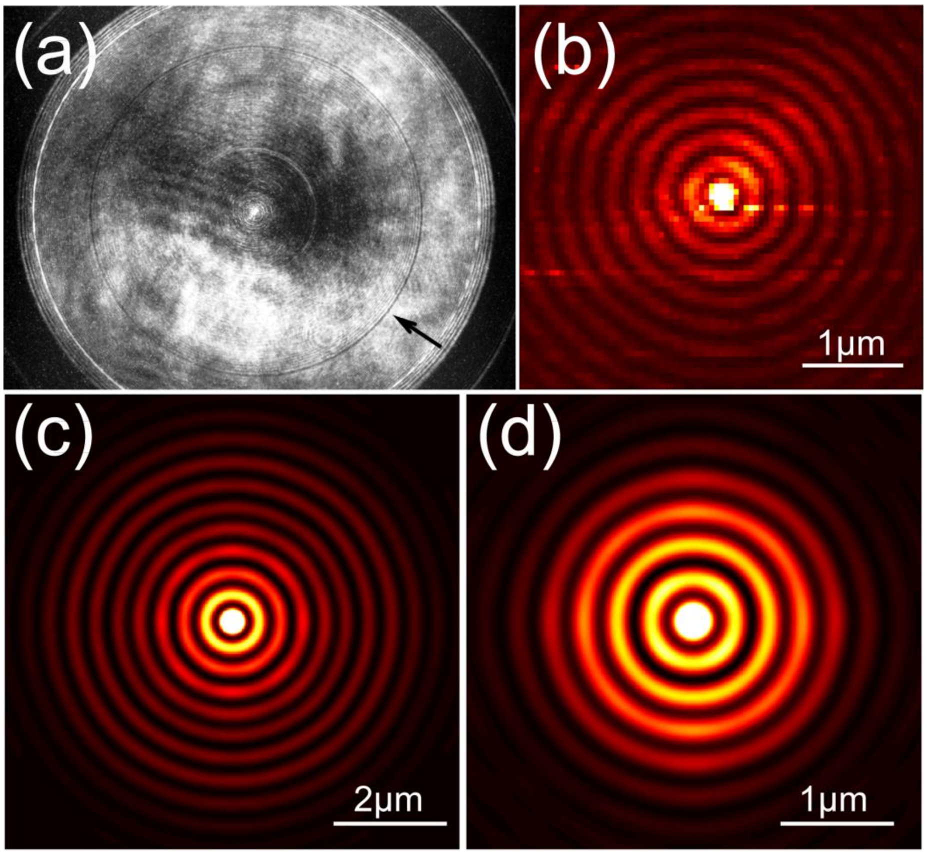
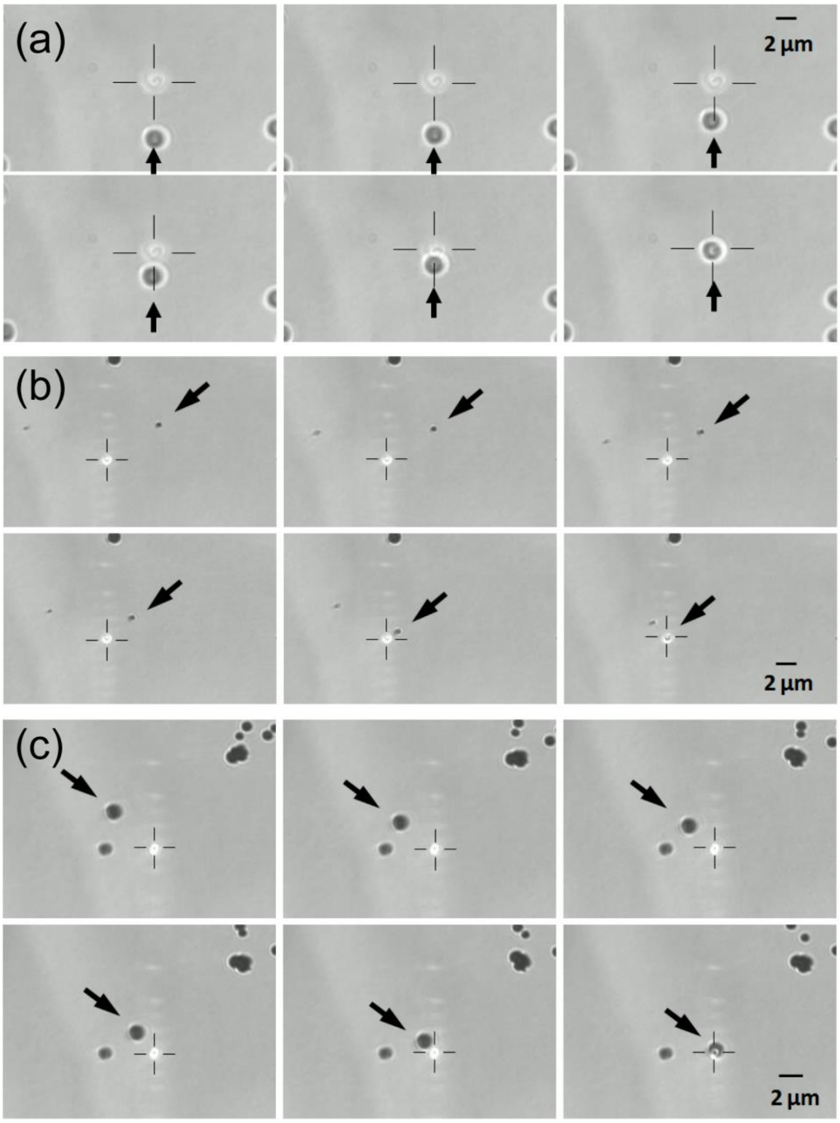
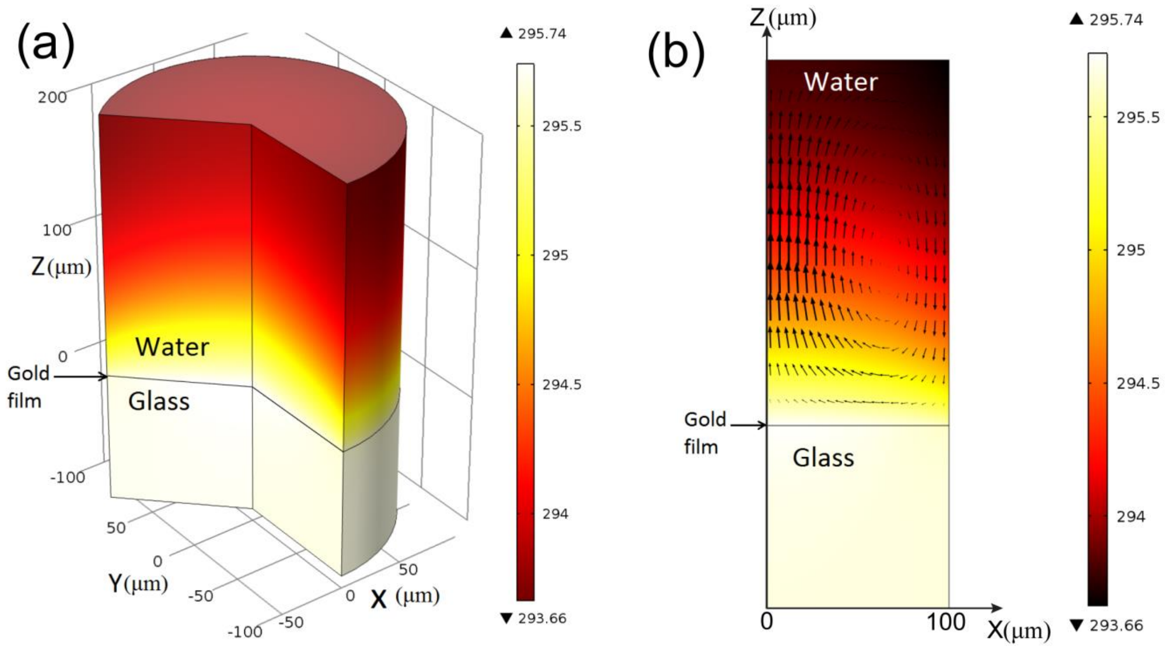
Supplementary Information



Supplementary Figure S1 | The experimental characterization and simulation of the SPP virtual probe. (a) The image on the back focal plane of the objective lens was obtained using a CCD camera at the wavelength of 1064 nm; the dark ring (indicated by a black arrow) corresponds to low reflectance resulting from SPP excitation. All parameters are as in Fig. 1 of main text. (b) Experimentally measured near-field intensity distribution of the SPP virtual probe on metal film (*x*-*y* plane) obtained using a new method referred to as surface-enhanced Raman scattering (SERS) mapping technique[38]. Here, without loss of generality, we chose the incident wavelength of 532 nm in the experiment, because our SERS mapping setup is dedicated to the 532 nm light source. Nevertheless, except for pattern size, the actual pattern of the virtual probe at 1064 nm should be very similar to that at 532 nm. We employ a 3D FDTD method (Rsoft Fullwave v8.1) to show both simulated results of the virtual probe field at 1064 nm and 532 nm in (c) and (d), respectively, to confirm that they have very similar patterns.



Supplementary Figure S2 | Successive images of experimental trapping of dielectric and metallic particles in plasmonic tweezers. The sample particles include (a) dielectric particle (silica) with a diameter of 2.47 µm, and gold particles with a diameter of about (b) 0.5 µm and (c) 1.5 µm. Black arrows indicate the trapped particle, and the black crosses indicate the position of the plasmonic virtual probe. These experimental results verify that our plasmonic tweezers can trap both dielectric and metallic particles over an effective large size range from nanometer to micrometer.



Supplementary Figure S3 | Temperature distribution and thermal convection in the focused plasmonic tweezers system. (a) 3D Temperature distribution around the central plasmonic virtual probe (x=0, y=0) during illumination calculated using the FEM method (COMSOL Multiphysics v4.3). Gold film with a thickness of 45 nm is located at z=0. Although the power of the laser is about 100 mW, the metal film is usually located 2~3 μm below the focal plane of the high NA objective lens, hence the light spot on the metal film is large and the average intensity is about 0.2 mW/μm^2. Thus, in the FEM simulation, we chose an illumination intensity of 0.2 mW/μm^2. The calculated result shows that maximum temperature increase is only about 2.1°K, because most of the power is reflected and only a small amount of light is coupled to the SPP (shown as the dark ring in Supplementary Figure S1(a)) and contributed to heating, and the heat is rapidly conducted to the whole gold film due to the high thermal conductivity. (b) 2D distribution of thermal convection velocity near the plasmonic virtual probe (x=0). The background shows the temperature distribution, and the black arrows indicate the direction of convection in water. The result of thermal convection currents demonstrates that the circulating fluid could produce a convection force to the central plasmonic virtual probe near the gold film. This convection force could further strengthen the trapping force along the horizontal dimension.

Supplementary Note 1 | Derivation of the total electromagnetic force on the particle

Lorentz force: $\mathbf{f}=\rho\mathbf{E}+\mathbf{J}\times\mathbf{B}$. (S1)

Dynamical Coulomb force (gradient or polarization force):

$$\mathbf{F}_{\mathrm{grad}}(t)=\int_v \boldsymbol{\rho}(\mathbf{r},t)\mathbf{E}(\mathbf{r},t)\,\mathrm{d}v\,. \tag{S2}$$

Dynamical Laplace force (scattering force):

$$\mathbf{F}_{\mathrm{scat}}(t)=\frac{1}{c}\int_v \mathbf{j}(\mathbf{r},t)\wedge\mathbf{H}(\mathbf{r},t)\,\mathrm{d}v\,. \tag{S3}$$

According to the inhomogeneous Maxwell equations:

$$\begin{cases}\nabla\cdot\mathbf{E}=\rho/\varepsilon \\ \nabla\cdot\mathbf{B}=\mu\mathbf{J}+\mu\varepsilon\dfrac{\partial\mathbf{E}}{\partial t}\end{cases}, \tag{S4}$$

then equation (S1) becomes

$$\mathbf{f}=\varepsilon(\nabla\cdot\mathbf{E})\mathbf{E}+\frac{1}{\mu}(\nabla\times\mathbf{B})\times\mathbf{B}-\varepsilon\frac{\partial\mathbf{E}}{\partial t}\cdot\mathbf{B}\,. \tag{S5}$$

Because $\dfrac{\partial(\mathbf{E}\times\mathbf{B})}{\partial t}=\dfrac{\partial\mathbf{E}}{\partial t}\times\mathbf{B}+\mathbf{E}\times\dfrac{\partial\mathbf{B}}{\partial t}=\dfrac{\partial\mathbf{E}}{\partial t}\times\mathbf{B}+\mathbf{E}(-\nabla\cdot\mathbf{E})$, then equation (S5) yields:

$$\mathbf{f}=\varepsilon(\nabla\cdot\mathbf{E})\mathbf{E}+\frac{1}{\mu}(\nabla\times\mathbf{B})\times\mathbf{B}-\varepsilon\frac{\partial(\mathbf{E}\times\mathbf{B})}{\partial t}-\varepsilon\times\mathbf{E}(\nabla\times\mathbf{E})\,,$$

$$\mathbf{f}=\varepsilon[(\nabla\cdot\mathbf{E})\mathbf{E}-\mathbf{E}(\nabla\times\mathbf{E})]+\frac{1}{\mu}(\nabla\times\mathbf{B})\times\mathbf{B}-\varepsilon\frac{\partial(\mathbf{E}\times\mathbf{B})}{\partial t}\,,$$

$$\mathbf{f}=\varepsilon[(\nabla\cdot\mathbf{E})\mathbf{E}-\mathbf{E}\times(\nabla\times\mathbf{E})]+\frac{1}{\mu}[(\nabla\cdot\mathbf{B})\mathbf{B}-\mathbf{B}\times(\nabla\times\mathbf{B})]-\varepsilon\frac{\partial(\mathbf{E}\times\mathbf{B})}{\partial t}\,,$$

$$\mathbf{f}=\varepsilon[(\nabla\cdot\mathbf{E})\mathbf{E}+(\mathbf{E}\cdot\nabla)\mathbf{E}-\frac{1}{2}\nabla\mathbf{E}^2]+\frac{1}{\mu}[(\nabla\cdot\mathbf{B})\mathbf{B}+(\mathbf{B}\cdot\nabla)\mathbf{B}-\frac{1}{2}\nabla\mathbf{B}^2]-\varepsilon\frac{\partial(\mathbf{E}\times\mathbf{B})}{\partial t}\,,$$

$$\mathbf{f}=\varepsilon[\nabla(\mathbf{EE})-\frac{1}{2}\nabla(\mathbf{IE}^2)]+\frac{1}{\mu}[\nabla(\mathbf{BB})-\frac{1}{2}\nabla(\mathbf{IB}^2)]-\varepsilon\frac{\partial(\mathbf{E}\times\mathbf{B})}{\partial t}\,. \tag{S6}$$

Because $\mathbf{EE}$ and $\mathbf{BB}$ are tensors and $\mathbf{I}$ an unit tensor, we obtain

$$\mathbf{f}=\nabla[\varepsilon(\mathbf{EE})-\frac{\varepsilon}{2}\mathbf{IE}^2+\frac{1}{\mu}(\mathbf{BB})-\frac{1}{2\mu}\mathbf{IB}^2]-\varepsilon\frac{\partial(\mathbf{E}\times\mathbf{B})}{\partial t}\,.$$

Using the Poynting vector $\mathbf{S} = \mathbf{E} \times \mathbf{H} = \frac{1}{\mu} \mathbf{E} \times \mathbf{B}$, we have

$$\mathbf{f} = \nabla \cdot \mathbf{T} - \varepsilon \frac{\partial(\mathbf{E} \times \mathbf{B})}{\partial t} = \nabla \cdot \mathbf{T} - \varepsilon\mu \frac{\partial \mathbf{S}}{\partial t}, \tag{S7}$$

where $\mathbf{T} = \varepsilon\varepsilon_0 \mathbf{E}\mathbf{E} + \mu\mu_0 \mathbf{H}\mathbf{H} - \frac{\mathbf{I}}{2}(\varepsilon\varepsilon_0 |\mathbf{E}|^2 + \mu\mu_0 |\mathbf{H}|^2)$ represents the Maxwell stress tensor matrix. Thus, the total electromagnetic force (Lorentz force) on the particle can be described as

$$\mathbf{F} = \int_v \mathbf{f} \, \mathrm{d}v = \int_v \nabla \cdot \mathbf{T} \, \mathrm{d}v - \varepsilon\mu \int_v \frac{\partial \mathbf{S}}{\partial t} \mathrm{d}v. \tag{S8}$$

According to the divergence theorem $\int_v \nabla \cdot \mathbf{A} \, \mathrm{d}v = \oint_s \mathbf{A} \cdot \mathbf{n} \, \mathrm{d}s$, equation (S8) becomes

$$\mathbf{F} = \oint_s \mathbf{T} \cdot \mathbf{n} \, \mathrm{d}s - \varepsilon\mu \frac{\partial}{\partial t} \int_v \mathbf{S} \, \mathrm{d}v. \tag{S9}$$

In a static electromagnetic field, $\frac{\partial \mathbf{S}}{\partial t} = 0$ after time-averaging, the total averaged force is

$$\langle \mathbf{F} \rangle = \left\langle \oint_s \mathbf{T} \cdot \mathbf{n} \, \mathrm{d}s \right\rangle, \tag{S10}$$

where $\mathbf{n}$ is the unit normal perpendicular to the small area $\mathrm{d}s$ on the particle surface.

Supplementary Note 2 | Derivation of the gradient and scattering forces

The gradient force is essentially a Coulomb force that depends on the local electric field and the charge density induced in particle. The scattering force is a Laplace force related to local magnetic field[39]. Thus, the electrical field force (gradient force) component of the Lorentz force can be expressed as in equation (S2). According to equation (S5), we obtain:

$$\mathbf{f}_{\mathrm{grad}} = \varepsilon(\nabla\cdot\mathbf{E})\mathbf{E} = \varepsilon(\frac{\partial E_x}{\partial x}+\frac{\partial E_y}{\partial y}+\frac{\partial E_z}{\partial z})\begin{pmatrix} E_x\mathbf{e}_x \\ E_y\mathbf{e}_y \\ E_z\mathbf{e}_z \end{pmatrix}, \tag{S11}$$

$$\mathbf{f}_{\mathrm{grad}} = \varepsilon\begin{pmatrix} \frac{\partial E_x}{\partial x}E_x+\frac{\partial E_y}{\partial y}E_x+\frac{\partial E_z}{\partial z}E_x \\ \frac{\partial E_z}{\partial x}E_y+\frac{\partial E_y}{\partial y}E_y+\frac{\partial E_z}{\partial z}E_y \\ \frac{\partial E_x}{\partial x}E_z+\frac{\partial E_y}{\partial y}E_z+\frac{\partial E_z}{\partial z}E_z \end{pmatrix}\begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \\ \mathbf{e}_z \end{pmatrix}, \tag{S12}$$

where $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$ are the unit vectors in *x*, *y*, and *z* directions. Thus $\mathbf{f}_{\mathrm{grad}}$ is from the first term of $\varepsilon\nabla\cdot(\mathbf{EE}) = \varepsilon[(\nabla\cdot\mathbf{E})\mathbf{E}+(\mathbf{E}\cdot\nabla)\mathbf{E}]$, shown in equation (S6), and the second term $\varepsilon(\mathbf{E}\cdot\nabla)\mathbf{E}$ is a part of the magnetic force. To obtain $\mathbf{f}_{\mathrm{grad}}$, we also need to calculate the second term $\varepsilon(\mathbf{E}\cdot\nabla)\mathbf{E}$, which can be derived as follows:

$$\varepsilon(\mathbf{E}\cdot\nabla)\mathbf{E} = \varepsilon(E_x\frac{\partial}{\partial x}+E_y\frac{\partial}{\partial y}+E_z\frac{\partial}{\partial z})\begin{pmatrix} E_x\mathbf{e}_x \\ E_y\mathbf{e}_y \\ E_z\mathbf{e}_z \end{pmatrix}, \tag{S13}$$

then $$(\mathbf{E}\cdot\nabla)\mathbf{E} = \begin{pmatrix} E_x\frac{\partial}{\partial x}E_x+E_y\frac{\partial}{\partial y}E_x+E_z\frac{\partial}{\partial z}E_x \\ E_x\frac{\partial}{\partial x}E_y+E_y\frac{\partial}{\partial y}E_y+E_z\frac{\partial}{\partial z}E_y \\ E_x\frac{\partial}{\partial x}E_z+E_y\frac{\partial}{\partial y}E_z+E_z\frac{\partial}{\partial z}E_z \end{pmatrix}\begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \\ \mathbf{e}_z \end{pmatrix}. \tag{S14}$$

When a mesoscopic or Mie metallic particle is in an external electrical field, the electrical field vanishes inside the metal particle, and the charge only distributes on the surface of the particle. This is observed from the electric field distributions in Fig. 3(c). Thus, for the slope (or differential) of electric field component from the surface to an arbitrary point on the surface of metal sphere, we obtain

$$\frac{\partial E_x}{\partial x} = \frac{E_x(x,y,z) - E_x(x-\Delta x, y, z)}{\Delta x} = \frac{E_x(x,y,z) - 0}{\Delta x} = \frac{E_x(x,y,z)}{\Delta x}.$$

In a similar way, we find

$$\frac{\partial E_a}{\partial b} = \frac{E_a(x,y,z)}{\Delta b}, \tag{S15}$$

where $a = x, y, z$, $b = x, y, z$; note that these differential results are approximate and only valid for large metallic particles (mesoscopic or Mie metallic particles). For Rayleigh metallic nanoparticles smaller than the skin depth of the metal, equation (S15) is invalid because the electrical field does not vanish inside the particle.

Substituting equation (S15) into the expressions for $(\nabla \cdot \mathbf{E})\mathbf{E}$ and $(\mathbf{E} \cdot \nabla)\mathbf{E}$, we then have

$$(\nabla \cdot \mathbf{E})\mathbf{E} = \mathrm{Real}\begin{pmatrix} \frac{E_x E_x^*}{\Delta x} + \frac{E_y E_x^*}{\Delta y} + \frac{E_z E_x^*}{\Delta z} \\ \frac{E_x E_y^*}{\Delta x} + \frac{E_y E_y^*}{\Delta y} + \frac{E_z E_y^*}{\Delta z} \\ \frac{\partial E_x E_z^*}{\Delta x} + \frac{E_y E_z^*}{\Delta y} + \frac{E_z E_z^*}{\Delta z} \end{pmatrix} \begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \\ \mathbf{e}_z \end{pmatrix}, \tag{S16}$$

$$(\mathbf{E} \cdot \nabla)\mathbf{E} = \mathrm{Real}\begin{pmatrix} \frac{E_x E_x^*}{\Delta x} + \frac{E_y E_x^*}{\Delta y} + \frac{E_z E_x^*}{\Delta z} \\ \frac{E_x E_y^*}{\Delta x} + \frac{E_y E_y^*}{\Delta y} + \frac{E_z E_y^*}{\Delta z} \\ \frac{\partial E_x E_z^*}{\Delta x} + \frac{E_y E_z^*}{\Delta y} + \frac{E_z E_z^*}{\Delta z} \end{pmatrix} \begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \\ \mathbf{e}_z \end{pmatrix}. \tag{S17}$$

As equations (S16) and (S17) are equal, we conclude

$$(\nabla \cdot \mathbf{E})\mathbf{E} = (\mathbf{E} \cdot \nabla)\mathbf{E} = \frac{1}{2}\nabla \cdot (\mathbf{EE}). \tag{S18}$$

Hence, we find from equations (S11) and (S18) the gradient force component of the Lorentz force to be:

$$\mathbf{f}_{\mathrm{grad}} = \varepsilon(\nabla \cdot \mathbf{E})\mathbf{E} = \frac{\varepsilon}{2}\nabla \cdot (\mathbf{EE}) \tag{S19}$$

As $\mathbf{EE}$ is a tensor, the gradient force can also be expressed in tensor form similar to the total force given in equation (S10). The total electrical field force (gradient force) becomes

$$\left\langle \mathbf{F}_{\text{grad}} \right\rangle = \left\langle \int_v \mathbf{f}_{\text{grad}} \, \mathrm{d}v \right\rangle = \left\langle \int_s \mathbf{T}_{\text{grad}} \cdot \mathbf{n} \, \mathrm{d}s \right\rangle, \tag{S20}$$

where the tensor of gradient force:

$$\mathbf{T}_{\text{grad}} = \frac{\varepsilon}{2} \mathbf{EE} = \frac{\varepsilon}{2} \begin{pmatrix} E_x E_x & E_x E_y & E_x E_z \\ E_y E_x & E_y E_y & E_y E_z \\ E_z E_x & E_z E_y & E_z E_z \end{pmatrix}. \tag{S21}$$

Because the Maxwell stress tensor matrix $\mathbf{T}$ consists of two components: $\mathbf{T} = \mathbf{T}_{\text{grad}} + \mathbf{T}_{\text{scat}}$,

the total magnetic field force (scattering force)

$$\left\langle \mathbf{F}_{\text{scat}} \right\rangle = \left\langle \int_v \mathbf{f}_{\text{scat}} \, \mathrm{d}v \right\rangle = \left\langle \int_s \mathbf{T}_{\text{scat}} \cdot \mathbf{n} \, \mathrm{d}s \right\rangle, \tag{S22}$$

where the tensor of the scattering gradient force is:

$$\mathbf{T}_{\text{scat}} = \mathbf{T} - \mathbf{T}_{\text{grad}} = \begin{bmatrix} \frac{1}{2}\varepsilon E_x E_x - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_x H_x - \frac{H^2}{2}) & \frac{1}{2}\varepsilon E_x E_y - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_x H_y - \frac{H^2}{2}) & \frac{1}{2}\varepsilon E_x E_z - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_x H_z - \frac{H^2}{2}) \\ \frac{1}{2}\varepsilon E_y E_x - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_y H_x - \frac{H^2}{2}) & \frac{1}{2}\varepsilon E_y E_y - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_y H_y - \frac{H^2}{2}) & \frac{1}{2}\varepsilon E_y E_z - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_y H_z - \frac{H^2}{2}) \\ \frac{1}{2}\varepsilon E_z E_x - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_z H_x - \frac{H^2}{2}) & \frac{1}{2}\varepsilon E_z E_y - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_z H_y - \frac{H^2}{2}) & \frac{1}{2}\varepsilon E_z E_z - \frac{\varepsilon}{2}E^2 + \frac{\mu}{2}(H_z H_z - \frac{H^2}{2}) \end{bmatrix}$$

Finally, the gradient force and scattering force can be calculated by substituting $\mathbf{T}_{\text{grad}}$ and $\mathbf{T}_{\text{scat}}$.

Supplementary References

38. L. P. Du, D. Y. Lei, G. H. Yuan, H. Fang, X. Zhang, Q. Wang, D. Y. Tang, C. J. Min, S. A. Maier, and X. C. Yuan, Mapping plasmonic near-field profiles and interferences by surface-enhanced Raman scattering. *Scientific Reports* **3**, 3064 (2013).
39. R. Quidant and C. Girard, Surface-plasmon-based optical manipulation. *Laser & Photon. Rev.* **2**, 47-57 (2008).