ARTICLE

Non-Abelian gauge field optics

Yuntian Chen[1,2,6], Ruo-Yang Zhang[3,6], Zhongfei Xiong[1], Zhi Hong Hang[4], Jensen Li[3], Jian Qi Shen[5] & C.T. Chan[3]

The concept of gauge field is a cornerstone of modern physics and the synthetic gauge field has emerged as a new way to manipulate particles in many disciplines. In optics, several schemes of Abelian synthetic gauge fields have been proposed. Here, we introduce a new platform for realizing synthetic SU(2) non-Abelian gauge fields acting on two-dimensional optical waves in a wide class of anisotropic materials and discover novel phenomena. We show that a virtual non-Abelian Lorentz force arising from material anisotropy can induce light beams to travel along Zitterbewegung trajectories even in homogeneous media. We further design an optical non-Abelian Aharonov–Bohm system which results in the exotic spin density interference effect. We can extract the Wilson loop of an arbitrary closed optical path from a series of gauge fixed points in the interference fringes. Our scheme offers a new route to study SU(2) gauge field related physics using optics.

[1] School of Optical and Electronic Information, Huazhong University of Science and Technology, 430074 Wuhan, China. [2] Wuhan National Laboratory of Optoelectronics, Huazhong University of Science and Technology, 430074 Wuhan, China. [3] Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. [4] School of Physical Science and Technology and Institute for Advanced Study, Soochow University, 215006 Suzhou, China. [5] Centre for Optical and Electromagnetic Research, State Key Laboratory of Modern Optical Instrumentations, Zhejiang University, 310058 Hangzhou, China. [6] These authors contributed equally: Yuntian Chen, Ruo-Yang Zhang. Correspondence and requests for materials should be addressed to J.Q.S. (email: jqshen@zju.edu.cn) or to C.T.C. (email: phchan@ust.hk)

G auge fields originated from classical electromagnetism, and have become the kernel of fundamental physics after being extended to non-Abelian by Yang and Mills[1]. Apart from real gauge bosons, emergent gauge fields in either real[2] or parameter spaces[3,4] have recently been widely used to elucidate the complicated dynamics in a variety of physical systems[5], including electronic[6,7], ultracold atom[8–10], and photonic[11–31] systems. The geometric nature[32] of gauge theory makes it a powerful tool for studying the topological phases of matter[33–36].

The concept of emergent gauge fields has offered us new insights in optics and photonics, such as the manifestation of the gauge structure (Berry connection and curvature) in momentum space[11–16]. Artificial gauge fields realized by breaking time reversal symmetry with magnetic effects[17–19] or dynamic modulation[21–23] have given rise to new paradigms for controlling light trajectories in real space. Even for time-reversal-invariant systems, a pair of virtual magnetic fields—each being the time-reversed partner of the other—can be generated using methods, such as coupled optical resonators[20], engineering lattices with strain[24,25], or reciprocal metamaterials[26–30]. However, except for a few works revealing the non-Abelian gauge structure in momentum space[13,14,16], all of these schemes of synthetic gauge fields in real space are restricted to the Abelian type.

Recently, anisotropic metamaterials were used to manipulate light through artificial Abelian gauge fields[27–30]. It was demonstrated that the off-diagonal components of permittivity and permeability appear as a pair of "spin-dependent" vector potentials in the two-dimensional (2D) wave equation for certain anisotropic media. Though the material parameters are subjected to strong restriction in this scheme, the internal pseudo-spin degree of freedom implies the possible generalization to a synthetic non-Abelian gauge field theory for light by coupling the spin-up and spin-down states.

In this work, we discover that the transport of optical waves in a wide class of anisotropic media can be associated with an emergent 2D non-Abelian SU(2) gauge interaction in real space, enabling us to obtain the first scheme for realizing synthetic non-Abelian gauge field for classical waves. Contrary to intuition, we show that a more exotic general SU(2) gauge framework can manifest in 2D optical dynamics, provided the restriction on the material parameters employed in refs. [27–30] is relaxed. Our platform presents broader applicability and allows the study of novel optical phenomena not found in Abelian synthetic gauge field systems. We illustrate our idea with two examples. The first example is the Zitterbewegung (ZB) of light in homogeneous non-Abelian media, which refers to the trembling motion of wave packets[37]. ZB has been realized in systems possessing Dirac dispersion[38–43], but we will see that ZB of light can arise from a distinctly different mechanism: emergent non-Abelian Lorentz force. In the second example, we propose for the first time a concrete design of a genuine non-Abelian Aharonov–Bohm (AB) system[44] using two synthetic non-Abelian vortices, and reveal that the noncommutativity of winding around the two vortices gives rise to nontrivial interference results. In particular, we show that there exists a series of fixed points in the interference fringes invariant under gauge transformation, from which we can obtain the Wilson loops of the closed path concatenated by the two interfering optical paths. As evidenced by the examples, our scheme offers a fresh angle to understand the dynamic effects of light in anisotropic media, and also suggests an optical approach to probe new physics accompanied by SU(2) gauge fields.

Results

Non-Abelian gauge fields acting on light. Our scheme focuses on 2D propagating optical waves in nondissipative anisotropic media characterized by the permittivity and permeability tensors:

$$\overleftrightarrow{\varepsilon}/\varepsilon_0 = \left(\begin{array}{c|c} \overleftrightarrow{\varepsilon}_T & \mathbf{g}_1 \\ \hline \mathbf{g}_1^\dagger & \varepsilon_z \end{array} \right), \quad \overleftrightarrow{\mu}/\mu_0 = \left(\begin{array}{c|c} \overleftrightarrow{\mu}_T & \mathbf{g}_2 \\ \hline \mathbf{g}_2^\dagger & \mu_z \end{array} \right). \quad (1)$$

Here, all of the parameters depend on x, y; the diagonal blocks $\overleftrightarrow{\varepsilon}_T$, $\overleftrightarrow{\mu}_T$, ε_z, μ_z are real numbers, while the off-block-diagonal components $\mathbf{g}_i = (g_{ix}, g_{iy})^\top = g_{ix}\mathbf{e}_x + g_{iy}\mathbf{e}_y$ ($i = 1$, 2) are in-plane complex vectors whose imaginary parts could be induced by the gyrotropic effect with in-plane gyration vectors. The only constraint on the media is the "in-plane duality", $\overleftrightarrow{\varepsilon}_T = \alpha\overleftrightarrow{\mu}_T$, where α is a positive constant. For simplicity, we set $\alpha = 1$ in the following, and $\alpha \neq 1$ results can be obtained directly by redefining $\varepsilon_0 \to \alpha\varepsilon_0$. Under this constraint, the in-plane monochromatic wave equation of frequency ω can be written as

$$\hat{H}|\psi\rangle = \left[\frac{1}{2}(\hat{\mathbf{p}} - \hat{\mathcal{A}}) \cdot \overleftrightarrow{\mathbf{m}}^{-1} \cdot (\hat{\mathbf{p}} - \hat{\mathcal{A}}) - \hat{\mathcal{A}}_0 + V_0\right]|\psi\rangle = 0. \quad (2)$$

Here $|\psi\rangle = (E_z, \eta_0 H_z)^\top$ $\left(\eta_0 = \sqrt{\mu_0/\varepsilon_0}\right)$ serves as a two-component wave function, and $\hat{H}$ resembles the Hamiltonian of a non-relativistic spin-1/2 particle traveling in SU(2) non-Abelian gauge potentials[45], where $\hat{\mathbf{p}} = -i\hat{\sigma}_0\partial_i\mathbf{e}^i$ ($i = 1$, 2) is the canonical momentum operator with $\hat{\sigma}_0$ being the 2D identity matrix, $\overleftrightarrow{\mathbf{m}} = \overleftrightarrow{\varepsilon}_T^{-1}\det(\overleftrightarrow{\varepsilon}_T)/2$ represents an effective anisotropic mass, in particular, $\hat{\mathcal{A}} = \mathcal{A}^1\hat{\sigma}_1 + \mathcal{A}^2\hat{\sigma}_2$ and $\hat{\mathcal{A}}_0 = \mathcal{A}_0^a\hat{\sigma}_a$ ($\hat{\sigma}_a$ ($a = 1$, 2, 3) are Pauli matrices) can be interpreted as emergent non-Abelian vector and scalar potentials, respectively, and V_0 is an additional Abelian scalar potential. As shown in Table 1, the emergent gauge potentials are determined by the material parameters, especially, the vector potential directly corresponds to the off-diagonal terms of $\overleftrightarrow{\varepsilon}$ and $\overleftrightarrow{\mu}$. This correspondence can be intuitively understood from the SU(2) gauge covariance of 2D Maxwell equations (see the "Methods" section), and the detailed derivation of Eq. (2) is given in the Supplementary Note 1. Thereby, in this broad class of anisotropic media, the materials' influence on the 2D optical waves imitates a SU(2) gauge interaction. Furthermore, if the background media are extended to be bi-anisotropic materials, a complete construction of U(2) = SU(2)U(1) gauge fields for light can be achieved (see Supplementary Note 1).

The emergent SU(2) gauge potential $\{\hat{\mathcal{A}}_\mu\} = \{\hat{\mathcal{A}}_0, \hat{\mathcal{A}}\}$ induces a synthetic SU(2) gauge field acting on light:

$$\hat{\mathcal{F}}_{\mu\nu} = i[\hat{\mathcal{D}}_\mu, \hat{\mathcal{D}}_\nu] = \partial_\mu\hat{\mathcal{A}}_\nu - \partial_\nu\hat{\mathcal{A}}_\mu - i[\hat{\mathcal{A}}_\mu, \hat{\mathcal{A}}_\nu], \quad (3)$$

where $\hat{\mathcal{D}}_\mu = \hat{\sigma}_0\partial_\mu - i\hat{\mathcal{A}}_\mu$ ($\mu = 0$, 1, 2) is the covariant derivative. Analogous to real electromagnetic (EM) fields, the synthetic SU

Table 1 The expressions of the synthetic SU(2) and U(1) gauge potentials in terms of the constitutive parameters of the non-Abelian media

SU(2)	Vector potential $\hat{\mathcal{A}} = \mathcal{A}^a\hat{\sigma}_a$	$\mathcal{A}^1 = k_0 Re(g_-) \times e_z$ [a]
		$\mathcal{A}^2 = k_0 Im(g_-) \times e_z$
		$\mathcal{A}^3 = 0$
	Scalar potential $\hat{\mathcal{A}}_0 = \mathcal{A}_0^a\hat{\sigma}_a$	$\mathcal{A}_0^1 = k_0 e_z \cdot \left[\nabla \times \left(\overleftrightarrow{\varepsilon}_T^{-1} \cdot Im(g_+)\right)\right]$
		$\mathcal{A}_0^2 = -k_0 e_z \cdot \left[\nabla \times \left(\overleftrightarrow{\varepsilon}_T^{-1} \cdot Re(g_+)\right)\right]$
		$\mathcal{A}_0^3 = k_0^2\left[\frac{\varepsilon_z - \mu_z}{2} - 2 Re\left(g_-^\dagger \cdot \overleftrightarrow{\varepsilon}_T^{-1} \cdot g_+\right)\right]$
U(1)	Scalar potential	$V_0 = k_0^2\left[\left(g_+^\dagger \cdot \overleftrightarrow{\varepsilon}_T^{-1} \cdot g_+\right) - \frac{\varepsilon_z + \mu_z}{2}\right]$

[a] Here $k_0 = \omega/c$ is the vacuum wave number, and $g_\pm = (g_1 \pm g_2^*)/2$.

(2) gauge field can be separated into a non-Abelian magnetic field $\hat{\mathcal{B}} = \frac{1}{2}\epsilon^{ij}\hat{\mathcal{F}}_{ij}\mathbf{e}_z$ along the z-axis and a non-Abelian in-plane electric field $\hat{\mathcal{E}} = -\hat{\mathcal{F}}_{0i}\mathbf{e}_i$, which are associated with the gauge potential as

$$\hat{\mathcal{B}} = \nabla\times\hat{\mathcal{A}} - \mathrm{i}\hat{\mathcal{A}}\times\hat{\mathcal{A}}, \quad \hat{\mathcal{E}} = \nabla\hat{\mathcal{A}}_0 + \mathrm{i}[\hat{\mathcal{A}}_0, \hat{\mathcal{A}}]. \quad (4)$$

The second terms of $\hat{\mathcal{B}}, \hat{\mathcal{E}}$ cannot be found in the Abelian case since they are induced entirely by the noncommutativity of the non-Abelian gauge potential. Indeed, a matrix-valued gauge potential would not be regarded as (apparently) non-Abelian, unless some of its components do not commute with each other $[\hat{\mathcal{A}}_\mu, \hat{\mathcal{A}}_\nu] \neq 0$[10]. For instance, the scheme in ref. [27] is actually a specific reduction of ours with the strict constraints on the media that (i) $\mathbf{g}_1 = -\mathbf{g}_2$ being real and (ii) $\varepsilon_z = \mu_z$. In this case, the vector potential only has $\hat{\sigma}_1$ component $\hat{\mathcal{A}} = \mathcal{A}^1\hat{\sigma}_1$ and the scalar potential $\hat{\mathcal{A}}_0$ vanishes. As such, $[\hat{\mathcal{A}}_i, \hat{\mathcal{A}}_j] \equiv 0$, and the gauge group is reduced to the Abelian subgroup U(1) of SU(2). In general, if Eq. (2) has any U(1) spin rotation symmetry, which means $\hat{U}\hat{H}\hat{U}^\dagger = \hat{H}$ for $\hat{U} = \exp\left(\mathrm{i}\phi\vec{n}\cdot\vec{\sigma}\right)$ with a parameter ϕ, the gauge potential would be reducible. Hence, only for those materials that can imitate irreducible SU(2) gauge potentials, we call them non-Abelian media.

The two-component wave function of light $|\psi\rangle$ behaves like a spin-1/2 spinor with the pseudo-spin at a local point

$$\vec{s} = \langle\psi|\vec{\hat{\sigma}}|\psi\rangle/|\psi|^2, \quad (5)$$

where the overhead arrow indicates a vector in the pseudo-spin space, and $\langle\psi|\vec{\hat{\sigma}}|\psi\rangle$ gives the local spin density. The frame $\{\vec{e}_a\}$ in the pseudo-spin space can be chosen arbitrarily. The rotation of the frame corresponds to a gauge transformation of spinor $|\psi'\rangle = \hat{U}(\mathbf{r})|\psi\rangle$, where in general $\hat{U}(\mathbf{r})$ is a space-varying SU(2) matrix. By substituting $|\psi'\rangle$ into Eq. (2), one can easily check that the wave equation is gauge covariant as long as the material is transformed accordingly (see Supplementary Note 2), while synthetic gauge potentials and fields obey the gauge transformations

$$\hat{\mathcal{A}}'_\mu = \hat{U}\hat{\mathcal{A}}_\mu\hat{U}^\dagger + \mathrm{i}\hat{U}\partial_\mu\hat{U}^\dagger, \quad (6)$$

$$\hat{\mathcal{B}}' = \hat{U}\hat{\mathcal{B}}\hat{U}^\dagger, \quad \hat{\mathcal{E}}' = \hat{U}\hat{\mathcal{E}}\hat{U}^\dagger. \quad (7)$$

In addition, it is worth comparing the present idea of non-Abelian gauge field optics (NAGFO) with the transformation optics (TO)[46–50]. When TO is applied to design invisibility cloaks, it results in anisotropic media whose permittivity and permeability are real and equal $\overleftrightarrow{\varepsilon} = \overleftrightarrow{\mu}$[46,47]. Due to the equivalence of the constitutive tensor and the metric of a curved spacetime for light, such kind of duality symmetric materials can also be used to mimic gravitational effects[49–53]. In contrast to TO, NAGFO involves a more general class of complex-valued media respecting in-plane duality symmetry. The in-plane block $\overleftrightarrow{\varepsilon}_T$ of permittivity, which determines the effective mass in Eq. (2), can alternatively be equated to the metric of a virtual 2D curved space as with TO, whereas, apart from $\overleftrightarrow{\varepsilon}_T$, all the remaining components of $\overleftrightarrow{\varepsilon}$ and $\overleftrightarrow{\mu}$ contribute to the synthetic SU(2) gauge potentials. Therefore, NAGFO proposes an optical way to simulate the 2D spinor systems under both a SU(2) gauge interaction and the influence of a curved space. To highlight the effects stemming purely from the non-Abelian gauge interaction, we will hereinafter concentrate on the simplified scenario that $\overleftrightarrow{\varepsilon}_T = \varepsilon_T\overleftrightarrow{\mathbf{I}}_{2\times2}$ is isotropic and homogeneous. As such, the virtual 2D background space is trivialized to be flat, and the effective mass is reduced to $m = \varepsilon_T/2$.

Zitterbewegung of optical beams. The wave packet dynamics in homogeneous media can give the most straightforward effect distinguishing the non-Abelian media from the Abelian type. The effective Abelian electric and magnetic fields vanish in homogeneous media[27], whereas the non-Abelian fields persist due to the noncommutativity of $\hat{\mathcal{A}}_\mu$. In our case, $\hat{\mathcal{B}} = \mathcal{B}\hat{\sigma}_3$ with $\mathcal{B} = \mathrm{i}k_0^2(\mathbf{g}_-\times\mathbf{g}_-^*)$, and $\hat{\mathcal{E}} = 2\mathcal{A}_0^3(\mathcal{A}^2\hat{\sigma}_1 - \mathcal{A}^1\hat{\sigma}_2)$. We consider the propagation of 2D optical beams in homogeneous non-Abelian media. In general, there are two non-degenerate branches of plane wave eigenstates. Because the two eigenstates of a certain direction of wave vector $\mathbf{k}$ are always orthogonal, their pseudo-spins correspond to a pair of antipodal points on the Bloch sphere. Generally speaking, the non-degenerate eigenmodes would evolves independently along different semiclassical trajectories. However, if the two eigenstates for a particular direction of $\mathbf{k}$ are quasi-degenerate, in the overlapped region, their superposed wave can be viewed as an intact "semiclassical particle" with an internal spin degree of freedom, whose centroid trajectory follows the Hamilton's canonical equations (see the "Methods" section)

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\langle\hat{\mathbf{p}}\rangle = \mathrm{i}\langle[\hat{H},\hat{\mathbf{p}}]\rangle \equiv 0 \quad \Rightarrow \quad \langle\hat{\mathbf{p}}\rangle \equiv \mathbf{k}, \quad (8)$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\langle\hat{\mathbf{r}}\rangle = \langle\hat{\mathbf{v}}\rangle = \frac{1}{m}(\mathbf{k} - \mathcal{A}^a\langle\hat{\sigma}_a\rangle). \quad (9)$$

Here $\hat{\mathbf{v}} = \frac{\mathrm{d}}{\mathrm{d}\tau}\hat{\mathbf{r}} = \mathrm{i}[\hat{H},\hat{\mathbf{r}}] = (\hat{\mathbf{p}} - \hat{\mathcal{A}})/m$ is the velocity operator, τ represents path parameter along the beam, and $\langle\hat{a}\rangle(\mathbf{r}_0) = \int\mathrm{d}\mathbf{r}_\perp \psi^\dagger(\mathbf{r}_0 + \mathbf{r}_\perp)\hat{a}(\mathbf{r}_0 + \mathbf{r}_\perp)\psi(\mathbf{r}_0 + \mathbf{r}_\perp)$ means the expectation value of an operator $\hat{a}$ averaged over the transverse cross section of a point $\mathbf{r}_0$ along an optical beam, differing from the local expectation value $\langle\psi|\hat{a}|\psi\rangle(\mathbf{r}) = \psi^\dagger(\mathbf{r})\hat{a}(\mathbf{r})\psi(\mathbf{r})$. According to Eq. (8), the canonical momentum along the beam is conserved, and is equal to the quasi-degenerate eigen wave vector $\mathbf{k}$ (see Eq. (39)). Moreover, it turns out that the in-plane projection of the total energy flux over the transverse cross section of the beam is always parallel to the velocity given by Eq. (9) (see proof in the section "Methods"), therefore the canonical equations do describe the path of energy propagation. Along the optical beam, the pseudo-spin $\vec{s} = \langle\vec{\hat{\sigma}}\rangle$ undergoes precession as follows:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\vec{s} = \mathrm{i}\langle[\hat{H},\vec{\hat{\sigma}}]\rangle = \vec{\Omega}\times\vec{s}, \quad (10)$$

where $\vec{\Omega} = -2(\mathcal{A}_0^a + \frac{1}{m}\mathbf{k}\cdot\mathcal{A}^a)\vec{e}_a$ is the precession angular velocity. During precession, the pseudo-spin component parallel to $\vec{\Omega}$ is conserved.

In terms of Eqs. (8–10), we arrive at the Newton-type equation of motion where a virtual non-Abelian Lorentz force[10,45] associated with the non-Abelian fields emerges

$$m\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\langle\hat{\mathbf{r}}\rangle = \frac{1}{2}\langle\hat{\mathbf{v}}\times\hat{\mathcal{B}} + \hat{\mathcal{B}}\times\hat{\mathbf{v}}\rangle + \langle\hat{\mathcal{E}}\rangle$$
$$= \langle\hat{\mathbf{j}}_{\hat{\sigma}_3}\rangle\times\mathcal{B} + \mathcal{E}^a\langle\hat{\sigma}_a\rangle, \quad (11)$$

Here, $\hat{\mathbf{j}}_{\hat{\sigma}_3} = \frac{1}{2}(\hat{\mathbf{v}}\hat{\sigma}_3 + \hat{\sigma}_3\hat{\mathbf{v}}) = \frac{1}{m}\hat{\mathbf{p}}\hat{\sigma}_3$ represents the $\hat{\sigma}_3$-component of the linear spin current operator[54], thus the non-Abelian Lorentz force can also be regarded as a spin-induced force with a magnetic part acting on the spin current and an electric part acting on the average spin over the transverse cross sections of the beam. In particular, the magnetic part of the force, $\mathbf{f}_{\hat{\sigma}_3} = \langle\hat{\mathbf{j}}_{\hat{\sigma}_3}\rangle\times\mathcal{B}$, duplicates the "spin transverse force" in electronics which acts on an electronic spin current exerted by a vertical electric field[54].

3

The integration of either the canonical equations or Eq. (11) yields the intensity centroid trajectory of the beam

$$\langle \widehat{\mathbf{r}} \rangle = \frac{1}{m}\left[\mathbf{k} - \mathcal{A}^a s_{0a} + \frac{1}{\Omega^2}\mathbf{F}^a \epsilon_{abc}\Omega^b s_0{}^c \right]\tau$$
$$- \frac{\mathbf{F}^a}{m\Omega^2}\left[(\cos(\tau\Omega) - 1)\delta_{ac} + \frac{\sin(\tau\Omega)}{\Omega}\epsilon_{abc}\Omega^b \right]s_0{}^c, \tag{12}$$

where $\mathbf{F}^a = (\vec{\Omega}\times\hat{\mathcal{A}})^a = \mathcal{E}^a + \mathbf{k}\times\mathcal{B}^a/m$, $\Omega = |\vec{\Omega}|$, $\vec{s}_0$ represents the initial spin, ϵ_{abc} is the Levi–Civita symbol, and the initial position of the beam is assumed at $\langle\widehat{\mathbf{r}}\rangle_0 = 0$. The first line of the equation refers to a straight path, while the second line shows that the beam oscillates around the equilibrium path periodically. As a result, the emergent non-Abelian Lorentz force may lead to wavy trajectories for optical beams propagating in the non-Abelian media. This phenomenon resembles the ZB effect of Dirac particles[37]. According to Eq. (12), the trembling motion of light depends not only on the non-Abelian gauge fields but also on the initial spin $\vec{s}_0$ of the beam. If the initial state is purely one of the eigenmodes with the wave vector in $\mathbf{k}$ direction, i.e., $\vec{s}_0$ is along $\vec{\Omega}(\mathbf{k})$, the trembling term in Eq. (12) will vanish. This implies the present ZB effect stems from the interference of the two quasi-degenerate eigenmodes just as electronic ZB is caused by the superposition of positive and negative energy components (see Supplementary Note 3).

In recent years, ZB has been investigated for spin–orbit coupled atoms[38,39] and photons[40–43]. However, unlike most schemes of ZB for light realized in periodic systems[40–42], our result shows that light can travel along curved paths even if the background medium is homogeneous. At first glance, this counterintuitive curved trajectory seems to violate the momentum conservation in translation invariant systems. However, it is well known that the kinetic momentum associated with centroid movement can be different from the canonical momentum for a particle traveling in a background vector potential. This conclusion is also valid for our situation. As shown in Eqs. (8) and (9), the semiclassical canonical momentum $\langle\widehat{\mathbf{p}}\rangle$ is always conserved in homogeneous media, while the kinetic momentum $m\langle\widehat{\mathbf{v}}\rangle$ deviates from $\langle\widehat{\mathbf{p}}\rangle$ and can change along the path by virtue of the synthetic non-Abelian potential $\hat{\mathcal{A}}$. A more rigorous analysis shows that the conserved quantity protected by space translational symmetry in generic non-Abelian media is the time-averaged Minkowski-type momentum $\int\mathrm{d}^3x\,\mathrm{Re}(\mathbf{D}^*\times\mathbf{B})$, while the centroid motion corresponds to the Abraham-type momentum $\int\mathrm{d}^3x\,\mathrm{Re}(\mathbf{E}^*\times\mathbf{H})/c^2$.

Example I: ZB induced by non-Abelian magnetic field. According to the theory, the ZB effect for monochromatic beams can be generated by either non-Abelian magnetic fields or non-Abelian electric fields. In Fig. 1a–e, we first show an example of ZB induced solely by a non-Abelian magnetic field. To realize nonzero $\hat{\mathcal{B}}$ but vanishing $\hat{\mathcal{E}}$, we let the medium satisfy $\varepsilon_z = \mu_z$, $\mathbf{g}_1 = -\mathbf{g}_2^* = (-\mathrm{i}\mathcal{A}_y^2/k_0, \mathcal{A}_x^1/k_0)^\top$, then the synthetic SU(2) magnetic field in this medium is given by $\hat{\mathcal{B}} = 2\mathcal{A}_x^1\mathcal{A}_y^2\mathbf{e}_z\hat{\sigma}_3$. The iso-frequency surfaces of eigenmodes are illustrated in Fig. 1a. Along the k_x direction, the two eigenstates are $|\rightarrow\rangle = (1,1)^\top/\sqrt{2}$ and $|\leftarrow\rangle = (1,-1)^\top/\sqrt{2}$ with the wave vectors $\mathbf{k}_\pm = \left[\sqrt{k_0^2\varepsilon_T\varepsilon_z - (\mathcal{A}_y^2)^2} \pm \mathcal{A}_x^1\right]\mathbf{e}_x$, and their pseudo-spins are polarized along the $\hat{\sigma}_1$-axis, as labeled on the Bloch sphere in Fig. 1c. As long as $|\mathcal{A}_x^1| \ll k = \sqrt{k_0^2\varepsilon_T\varepsilon_z - (\mathcal{A}_y^2)^2}$, the quasi-degenerate approximation is valid for beams incident from x direction. In this case, the precession angular velocity is $\vec{\Omega} = -4k\mathcal{A}_x^1/\varepsilon_T\vec{e}_1$, so the pseudo-spin will precess about the

$\hat{\sigma}_1$-axis. For the initial spin $\vec{s}_0 = (\cos\theta_0, \sin\theta_0\cos\phi_0, \sin\theta_0\sin\phi_0)^\top$ at an angle θ_0 from $\hat{\sigma}_1$-axis, we can obtain the centroid trajectory of the beam by eliminating the ray parameter τ in Eq. (12),

$$y(x) = Y_{\mathrm{ZB}}\left[\sin(k_{\mathrm{ZB}}(x - x_0) - \phi_0) + \sin\phi_0\right], \tag{13}$$

where x, y are the coordinates of centroid. The ZB amplitude

$$Y_{\mathrm{ZB}} = \frac{-\mathcal{A}_y^2\sin\theta_0}{2k\mathcal{A}_x^1} = \frac{-\mathcal{A}_y^2\sin\theta_0}{2\mathcal{A}_x^1\sqrt{k_0^2\varepsilon_T\varepsilon_z - (\mathcal{A}_y^2)^2}} \tag{14}$$

is proportional to $\sin\theta_0$, so ZB reaches the maximum when the initial spin $\vec{s}_0$ is perpendicular to $\vec{\Omega}$, corresponding to the equal-weighted superposition of the two eigenmodes. Meanwhile, the ZB wave number

$$k_{\mathrm{ZB}} = \frac{2k\mathcal{A}_x^1}{k - \mathcal{A}_x^1\cos\theta_0} \approx 2\mathcal{A}_x^1 = k_+ - k_- \tag{15}$$

is equal to the difference of the two eigen wave vectors, showing that ZB originates from the beating between the two eigenstates. Yet we should emphasize the phase beating is not a sufficient condition to realize ZB, and the ZB amplitude cannot be obtained without the knowledge of the non-Abelian dynamics. For instance, if $\mathcal{A}_y^2 = 0$ in the present medium, the beat of the two states persists, however, as the medium is relegated to the Abelian-type with $\hat{\mathcal{B}} = 0$, ZB just disappears.

We have performed the full-wave simulation of a Gaussian beam propagating in this medium using COMSOL Multiphysics. The beam is emitted along x-direction and the angle between its initial spin and $\hat{\sigma}_1$-axis is set as $\theta_0 = 0.43\pi$. Figure 1b shows the k-space Fourier amplitude of the simulated wave function ψ, the two peaks in the spectrum manifest that the beam is mainly comprised of the two eigenstates $|\rightarrow\rangle$ and $|\leftarrow\rangle$. The numerical time-averaged energy densities plotted in Fig. 1d show clearly a transverse tremor along the beam. As shown in Fig. 1e, the centroid trajectory extracted from the full-wave result agrees perfectly with the analytical expression in Eq. (13). And according to the numerical data of the pseudo-spins in one ZB period shown in Fig. 1c, the spin precession about the $\hat{\sigma}_1$-axis is also demonstrated.

**Example II: ZB induced by non-Abelian electric field.** In the previous example, the non-Abelian medium contains both gyroelectric and gyromagnetic components. In fact, the synthetic non-Abelian gauge fields as well as ZB can be simply realized with reciprocal media without gyrotropy. Here, we elaborate on synthesizing non-Abelian electric field with a biaxial non-magnetic material and the ZB effect in it.

We consider a non-magnetic material with the biaxial permittivity $\widetilde{\varepsilon}/\varepsilon_0 = \mathrm{diag}(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ ($\varepsilon_1 < \varepsilon_2 < \varepsilon_3$) along the principal axis and the permeability $\mu/\mu_0 = 1$. If the second principal axis of $\widetilde{\varepsilon}$ is fixed along the y-axis, while the first principal axis is rotated by an angle φ with respect to the x-axis such that $\cos^2\varphi\,\varepsilon_1 + \sin^2\varphi\,\varepsilon_3 = \varepsilon_2$ ($|\varphi| < \pi/2$) as shown in Fig. 1f, the permittivity tensor in the xyz coordinate system is given by

$$\overleftrightarrow{\varepsilon}/\varepsilon_0 = \begin{pmatrix} \varepsilon_2 & 0 & g \\ 0 & \varepsilon_2 & 0 \\ g & 0 & \varepsilon_z \end{pmatrix}, \tag{16}$$

with $\varepsilon_z = \varepsilon_1 + \varepsilon_3 - \varepsilon_2$ and $g = \mathrm{sgn}(\varphi)\sqrt{(\varepsilon_2 - \varepsilon_1)(\varepsilon_3 - \varepsilon_2)}$. Since the in-plane duality condition is satisfied as $\varepsilon_T = \varepsilon_2\mu_T$ ($\mu_T = 1$), by rescaling the vacuum permittivity $\varepsilon_0' = \varepsilon_2\varepsilon_0$, we obtain the
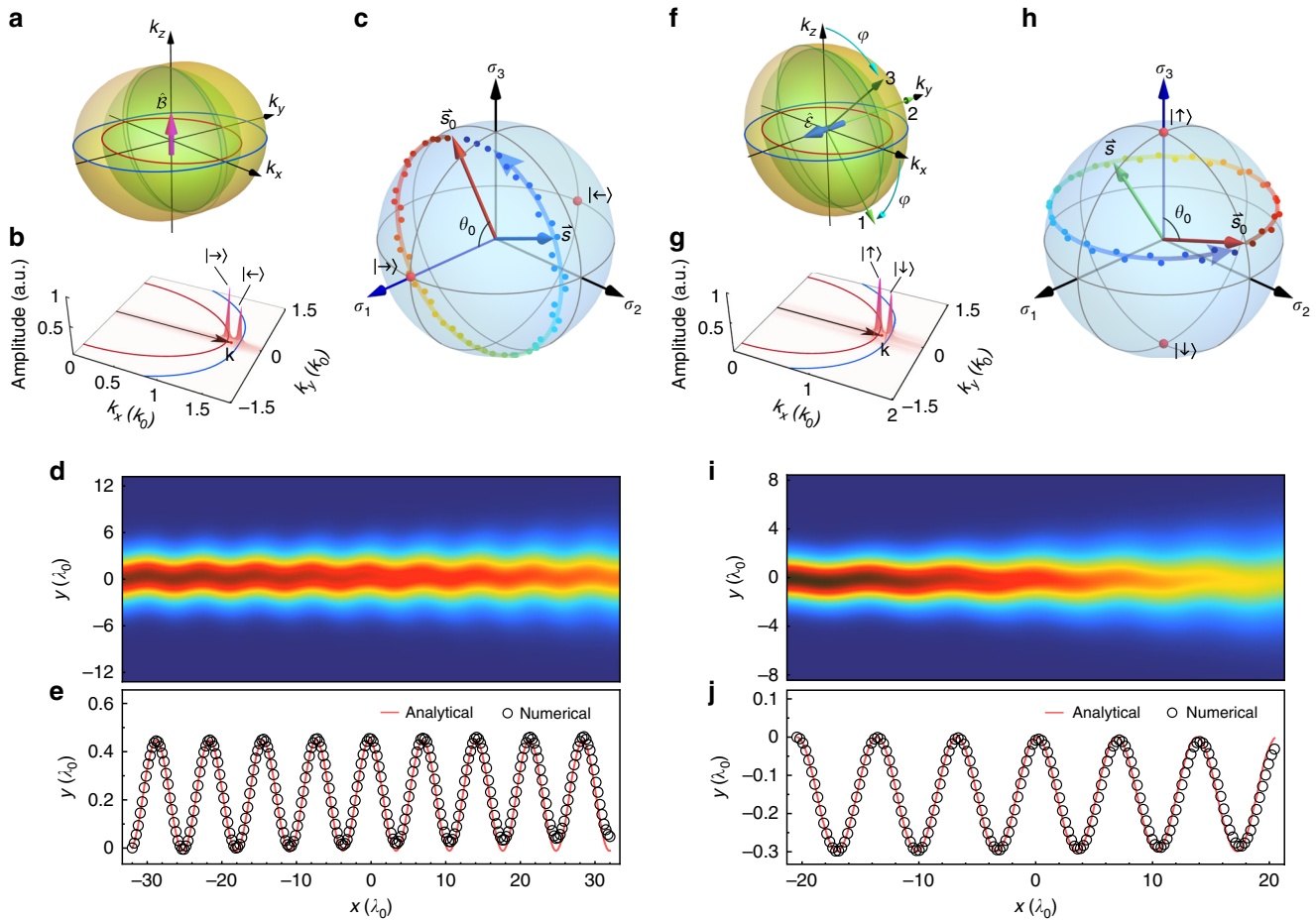
Fig. 1 Zitterbewegung effect in homogeneous non-Abelian media. **a–e** ZB induced by a synthetic non-Abelian magnetic field in a gyrotropic medium with the parameters $\overleftrightarrow{\varepsilon}_T = \overleftrightarrow{\mu}_T = 1.5\, \mathbf{I}_{2\times 2}$, $\varepsilon_z = \mu_z = 1.5$, $\mathbf{g}_1 = -\mathbf{g}_2^* = (0.3\mathrm{i}, -0.07)^\top$. This medium produces a synthetic SU(2) magnetic field along z direction, $\hat{\mathcal{B}} = -k_0^2\, 0.042\mathbf{e}_z\hat{\sigma}_3$, with a null SU(2) electric field $\hat{\mathcal{E}} = 0$. **f–j** ZB induced by a synthetic non-Abelian electric field in a biaxial non-magnetic medium with the parameters $\varepsilon_1 = 1.65$, $\varepsilon_2 = 2.45$, $\varepsilon_3 = 3$, and $\mu/\mu_0 = 1$. The synthetic SU(2) electric field, $\hat{\mathcal{E}} = -k_0^3\, 0.08919\, \mathbf{e}_y\hat{\sigma}_2$, is along the y-axis, while the SU(2) magnetic field vanishes $\hat{\mathcal{B}} = 0$. **a, f** The isofrequency surfaces and their xy cross sections (red and blue curves) of both cases. The green arrows in **f** are the three principal axes 1, 2, 3 of permittivity tensor. **b, g** Fourier spectra in k-space of the beams in the two media. In each case, the two peaks in the spectrum correspond to the two eigenmodes with wave vectors in the x direction. And the average wave vectors **k** are marked by the black arrows. **c, h** The spin precession along each beam on the Bloch sphere. The colored dots are the numerical data within one ZB period. **d, i** Full-wave simulated intensity distributions, where the beam waists equal $4.4\lambda_0$ and $6.2\lambda_0$, respectively ($\lambda_0 = 2\pi/k_0$ is the wavelength in vacuum). **e, j** Numerical (black circles) and analytical (red curves) trajectories of the intensity centroid

synthetic gauge potentials

$$\hat{\mathcal{A}} = -\frac{k_0 g}{2\sqrt{\varepsilon_2}}\mathbf{e}_y\hat{\sigma}_1, \qquad \hat{\mathcal{A}}_0 = k_0^2 \frac{\varepsilon_1\varepsilon_3 - \varepsilon_2^2}{2\varepsilon_2}\hat{\sigma}_3, \qquad (17)$$

and a uniform non-Abelian electric field polarized along the second principal axis

$$\hat{\mathcal{E}} = k_0^3 \frac{g\,(\varepsilon_1\varepsilon_3 - \varepsilon_2^2)}{2\,\varepsilon_2^{3/2}}\mathbf{e}_y\hat{\sigma}_2. \qquad (18)$$

The two eigenstates in the x-direction are $|\uparrow\rangle = (1,0)^\top$ and $|\downarrow\rangle = (0,1)^\top$ corresponding to the two poles along the $\hat{\sigma}_3$-axis on the Bloch sphere (see Fig. 1h), the eigen wave vectors are $\mathbf{k}_\uparrow = \sqrt{\varepsilon_1\varepsilon_3/\varepsilon_2}\,k_0\mathbf{e}_x$ and $\mathbf{k}_\downarrow = \sqrt{\varepsilon_2}\,k_0\mathbf{e}_x$ respectively. Providing that $|\sqrt{\varepsilon_1\varepsilon_3}/\varepsilon_2 - 1|$ is small enough, the centroid trajectory of a beam mainly consisting of these two states satisfies

$$y(x) = Y_{ZB}\left[\sin(k_{ZB}(x - x_0) + \phi_0) - \sin\phi_0\right], \qquad (19)$$

where θ_0, ϕ_0 are the Euler angles of the initial spin

$\vec{s}_0 = (\sin\theta_0\cos\phi_0, \sin\theta_0\sin\phi_0, \cos\theta_0)^\top$, the ZB amplitude is

$$Y_{ZB} = \frac{\mathcal{A}_y^1 \sin\theta_0}{\mathcal{A}_0^3} = \frac{\sqrt{\varepsilon_2}\,g\,\sin\theta_0}{k_0(\varepsilon_2^2 - \varepsilon_1\varepsilon_3)}, \qquad (20)$$

and the ZB wave number

$$k_{ZB} = \frac{k_0^2(\varepsilon_2 - \varepsilon_1\varepsilon_3/\varepsilon_2)}{2k} = \frac{k_\downarrow^2 - k_\uparrow^2}{2k} \approx k_\downarrow - k_\uparrow \qquad (21)$$

is still determined by the beating of the two eigenstates. In the full-wave simulation of Fig. 1i, we obtained a trembling beam (also see Fig. 1g for its Fourier spectrum) where the decay of intensity along the beam is due to the beam divergence, the extracted centroid trajectory faithfully reproduces the analytic prediction of Eq. (19), shown by Fig. 1j. In Fig. 1h, the numerical spin trajectory on the Bloch sphere also verifies that the pseudo-spin precesses about the $\hat{\sigma}_3$-axis.

In principle, the ZB effect induced by non-Abelian electric field can be observed in any natural and artificial biaxial non-magnetic materials. Here, we designed a simple metamaterial structure with the unit cell shown in Fig. 2a for realizing ZB in
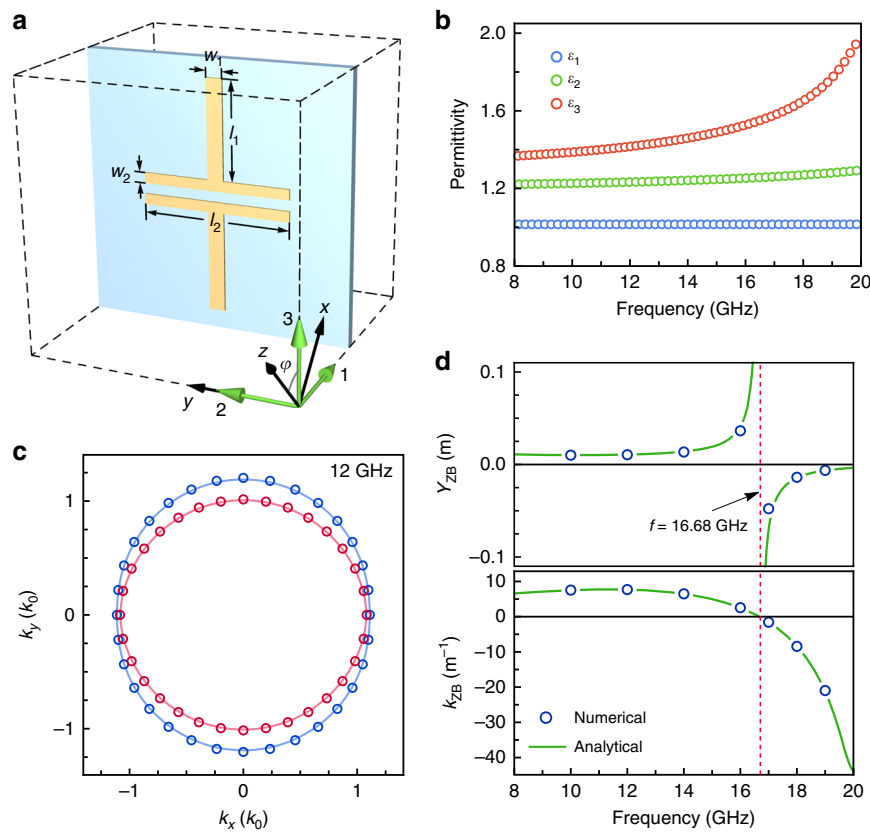
Fig. 2 Design of a biaxial metamaterial. **a** Unit cell of the simple cubic structure with lattice constants $d = 5$ mm, where a FR4 PCB slab (light blue) with thickness 0.2 mm and relative permittivity $\varepsilon_{pcb} = 3.3$ fills the coronal plane, and a copper structure of thickness 0.035 mm is printed on the PCB slab with the geometric parameters $l_1 = 2$ mm, $l_2 = 2.8$ mm, $w_1 = 0.3$ mm, and $w_2 = 0.2$ mm. **b** Dispersions of the retrieved relative permittivities along the three principal axes. **c** Isofrequency contours of the metamaterial in the xy-plane at 12 GHz, where the circular markers and solid curves correspond, respectively, to the real structure in **a** and the retrieved homogeneous medium. **d** Full-wave simulated (blue circles) and analytical (green curves) ZB amplitude Y_{ZB} and ZB wave number k_{ZB} in the homogenized media varying with frequency

microwave regime. The copper strips on printed circuit board (PCB) layers support strong and anisotropic electric dipole resonances along principal axes labeled as 1, 2. Consequently, all the three principal values ε_i ($i = 1, 2, 3$) of the effective permittivity are different, and their dispersions obtained by S-parameter retrieval approach[55] are plotted in Fig. 2b. According to our theory, the ZB beams should travel in the xy plane whose orientation is determined by ε_i and thus is frequency-dependent. As an example, we compared in Fig. 2c the isofrequency contours in xy-plane of the real structure and that of the homogenized medium at 12 GHz. Their perfect consistency confirms the retrieval result. To test the ZB effect in the metamaterial, we numerically simulated the ZB beams with a constant waist of 0.2 m propagating along x-direction in the retrieved media at some discrete frequencies and extracted the ZB amplitudes Y_{ZB} and ZB wave numbers k_{ZB}. We find good agreement with the theoretical predictions given by Eqs. (20) and (21) as shown in Fig. 2d. Notably, both of the ZB amplitude and period tend to infinity at a singular frequency $f = 16.68$ GHz, due to the fact that $\varepsilon_1\varepsilon_3 = \varepsilon_2^2$ is accidentally satisfied at the frequency such that the material is reduced to Abelian type with $\hat{\mathcal{E}} = 0$, and the beam splits into two branches[30]. We have also analyzed the finite width effect of the beam in the z-direction, and the analysis demonstrates that the 2D theory works well in the region where the two eigenmodes do not split away along the z-axis (see Supplementary Note 4).

Non-Abelian Aharonov–Bohm system for light. ZB discussed in the previous section can be viewed as the interference between two eigenmodes, each of which evolves with Abelian dynamics. In this sense, ZB is an apparent non-Abelian effect. Next, we will introduce the genuine non-Abelian AB effect, which cannot be reduced to Abelian subsystems.

The AB effect covers a group of phenomena associated with the path-dependent phase factors for particles traveling in a field-free region, but with irremovable gauge potential $\hat{A}_\mu$, the discovery of which confirmed the physical reality of gauge potentials and the nonlocality of gauge interactions[56,57]. The AB effect was first generalized to non-Abelian by Wu and Yang[32], who showed that the scattering of nucleons (isospinors) around a non-Abelian flux tube (vortex) can generate peculiar phenomena. However, their governing Hamiltonian can be globally diagonalized into two decoupled Abelian subsystems under a proper gauge[58], and all relevant phenomena can be interpreted from the superposition of the two subsystems. Hence, Wu and Yang's proposal is now viewed as an apparent non-Abelian effect[10,44]. According to a rigorous definition[44], a genuine non-Abelian AB system requires its holonomy group Hol($\hat{A}$) to be non-Abelian (see the "Methods" section and Supplementary Note 5). As such, there should exist such loops based at a fixed point that their non-Abelian AB phase factors (holonomies) are noncommutable, i.e. if a particle travels along two such loops in opposite sequences, the obtained AB phase factors would be different. This implies that at least two vortices exist in a genuine non-Abelian system[44].
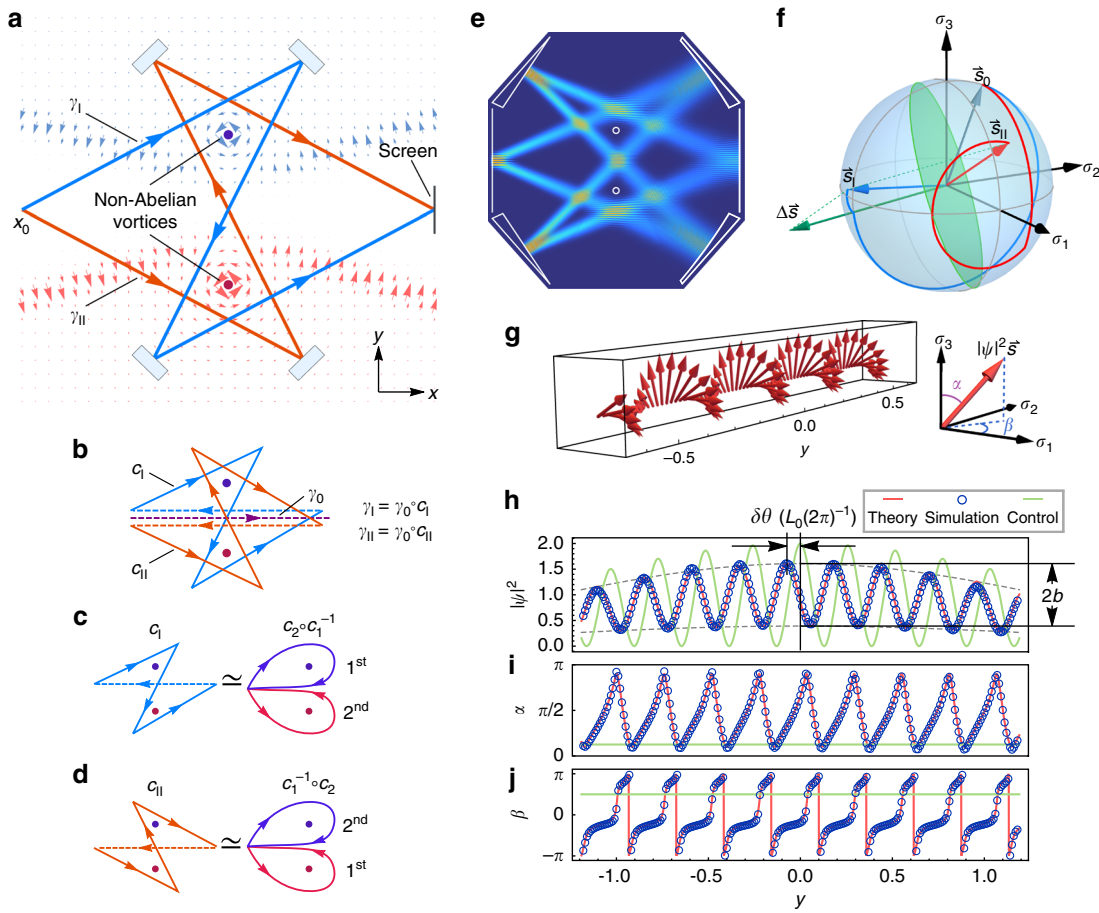
Fig. 3 Genuine non-Abelian AB effect for light. **a** Sketch of the non-Abelian AB system with two optical paths γ_I, γ_{II} interfering on the screen, where the background light blue (red) arrows denote the $\hat{\sigma}_1$ ($\hat{\sigma}_2$) component $\mathcal{A}^1$ ($\mathcal{A}^2$) of the non-Abelian vector potential. **b** γ_I (γ_{II}) can be divided into a closed loop c_I (c_{II}) and a common path γ_0. **c**, **d** c_I and c_{II} can, respectively, deform continuously into a closed path that winds around the two vortices successively but in opposite sequences. **e** Snapshot of the simulated field intensity for the proposed non-Abelian optical interferometer with incident spinor $(1, i1/5)^{\top}$ for both beams and the vortex fluxes $\Phi_1 = -2\pi/3$, $\Phi_2 = -\pi/3$. **f** Spin evolution on the Bloch sphere along two beams γ_I, γ_{II}, which share the same initial spin $\vec{s}_0$ but achieve different final spins $\vec{s}_I$ and $\vec{s}_{II}$. **g** Spin density interference corresponding to **e**, where each arrow denotes the local pseudo-spin density $|\psi|^2\vec{s}$ at a point on the screen. All of the local spins $\vec{s}(y)$ are perpendicular to $\Delta\vec{s} = \vec{s}_I - \vec{s}_{II}$, and thus fall on the green circle in **f**. The corresponding intensity interference $|\psi|^2(y)$ and the two Euler angles α, β of the local spins $\vec{s}(y)$ on the screen are shown in **h–j**, where blue circles and red curves indicate simulated and theoretical results, respectively, and $\delta\theta$, b are the phase shift and relative amplitude relative to the case of $\hat{\mathcal{A}} = 0$ (L_0 is the period of $\Delta\theta(y) \bmod 2\pi$). The green lines correspond to the "control experiment"

Indeed, we can use anisotropic and gyrotropic materials (see Table 1) to synthesize two vortices of SU(2) vector potential $\hat{\mathcal{A}} = \mathcal{A}^1\hat{\sigma}_1 + \mathcal{A}^2\hat{\sigma}_2$ $\left(\hat{\mathcal{A}}_0 = 0\right)$ with vanishing field $\hat{\mathcal{B}} = 0$ in the whole space except for two small domains, taken as point singularities for simplicity. Here, we provide the synopsis of our scheme, and more details are given in Supplementary Note 6 (also see Supplementary Note 8 for an alternative design). As illustrated in Fig. 3a, we demand $\hat{\mathcal{A}} = \mathcal{A}^1\hat{\sigma}_1$ $\left(\mathcal{A}^2 = 0\right)$ in the upper half-space, while $\hat{\mathcal{A}} = \mathcal{A}^2\hat{\sigma}_2$ $\left(\mathcal{A}^1 = 0\right)$ in the lower half-space. We also require that $\mathcal{A}^1$, $\mathcal{A}^2$ smoothly tend to zero in the middle region without overlap. In the vicinity of the upper (lower) singularity, $\mathcal{A}^1$ $\left(\mathcal{A}^2\right)$ forms an irrotational vortex carrying the flux Φ_1 (Φ_2) (see Supplementary Eq. (44) for the concrete expression of $\hat{\mathcal{A}}$ fulfilling these requirements). For a closed loop with a fixed base-point, its non-Abelian holonomy is invariant against continuous deformation of the path within the $\hat{\mathcal{B}} = 0$ region. As a consequence, for the two homotopy classes of loops $[c_1]$ and $[c_2]$ (where $[c_i]$ denote the path homotopy classes; see the "Methods" section), based at $\mathbf{x}_0$ and encircling the upper (for $[c_1]$)

or lower (for $[c_2]$) vortex once, their holonomies are $\hat{U}_i = \hat{U}_{[c_i]}[\mathbf{x}_0] = \exp[i\Phi_i\hat{\sigma}_i]$ $(i = 1, 2,)$ respectively. As $\hat{U}_1$ and $\hat{U}_2$ do not commute with each other, this double-vortex system is a genuine non-Abelian AB system.

In order to realize the vector potential shown in Fig. 3a, the background media are set up as $\mathbf{g}_1 = -\mathbf{g}_2^*$ (i.e. $\mathbf{g}_+ = 0$) and $\varepsilon_T = \varepsilon_z = \mu_T = \mu_z = \text{const.}$ to guarantee $\hat{\mathcal{A}}_0 \equiv 0$ and $V_0 = \text{const.}$ Also, we use reciprocal anisotropic materials with purely real off-block-diagonal components $\mathbf{g}_1 = -\mathbf{g}_2$ to build the vector potential $\hat{\mathcal{A}} = \mathcal{A}^1\hat{\sigma}_1$ in the upper half plane but gyrotropic materials with purely imaginary $\mathbf{g}_1 = \mathbf{g}_2$ to build $\hat{\mathcal{A}} = \mathcal{A}^2\hat{\sigma}_2$ in the lower half plane (see Supplementary Note 6 for details). As a result, we have designed a genuine non-Abelian AB system for light. Then, we will show how the genuine non-Abelian nature of the system can be detected from interference effects.

Non-Abelian AB interference. Consider two coherent light beams with the same initial spin $\vec{s}_0$ propagating separately along the two folded paths γ_I and γ_{II}, and finally superposing on the

screen (Fig. 3a). For the trivial situation of $\hat{A} = 0$, the two beams are uniformly polarized along the whole paths, thus their final states are given by $|\psi_i(y)\rangle = a(y)e^{i\theta_i(y)}|s_0\rangle$ ($i = $ I, II), where $a(y)$ is the envelope of both beams on the screen, $|s_0\rangle$ is the normalized initial spinor at $\mathbf{x}_0$, and $\theta_i(y)$ denote the dynamic phases which have included the initial phases. The dynamic phase difference, $\Delta\theta(y) = \theta_I(y) - \theta_{II}(y)$, determines the interference pattern: $|\psi_I + \psi_{II}|^2(y) = 2a(y)[1 + \cos(\Delta\theta(y))]$.

In the presence of the two non-Abelian vortices of $\hat{A}$, the two optical paths are unchanged thanks to the null gauge field. However, the gauge potential drives the pseudo-spins to rotate along the paths, and the two final states convert to

$$|\psi_i(y)\rangle = a(y)\hat{U}_{\gamma_i}e^{i\theta_i(y)}|s_0\rangle, \quad (i = \text{I, II}) \tag{22}$$

where an additional non-Abelian AB phase factor $\hat{U}_{\gamma_i} = \mathcal{P}\exp\left[i\int_{\gamma_i}\hat{A}\cdot d\mathbf{r}\right]$ appears in each state. The optical path of each beam can be regarded as a concatenation of a closed loop c_i and a common path γ_0, i.e., $\gamma_i = \gamma_0 \circ c_i$ ($i = $ I, II), as illustrated in Fig. 3b. The closed loop c_I can be further deformed continuously into two successive loops $c_2 \circ c_1^{-1}$, which winds around the upper vortex (clockwise) first and subsequently the lower vortex (anticlockwise) (Fig. 3c). Likewise, c_{II} is homotopic to $c_1^{-1} \circ c_2$, namely c_{II} winds around the lower vortex first before it does the upper vortex (Fig. 3d). Because of the noncommutativity of the sequences of winding around the two vortices, the AB phase factors of the two beams are different:

$$\hat{U}_{\gamma_I} = \hat{U}_{\gamma_0}\hat{U}_2\hat{U}_1^{-1} \neq \hat{U}_{\gamma_{II}} = \hat{U}_{\gamma_0}\hat{U}_1^{-1}\hat{U}_2. \tag{23}$$

Consequently, the two beams will end up with distinct spins $\vec{s}_I$ and $\vec{s}_{II}$ on the screen (Fig. 3f), and they will interfere with each other in a nontrivial way. The term spin density interference was coined for this phenomenon and it can be calculated as follows:

$$\langle\psi_I(y) + \psi_{II}(y)|\vec{\hat{\sigma}}|\psi_I(y) + \psi_{II}(y)\rangle = |\psi|^2(y)\vec{s}(y). \tag{24}$$

Here, the angular bracket denotes the spinor inner product at a local position y on the screen, the obtained result describes the spin density distribution on the screen. The spin density can be further decomposed into two parts: the intensity interference $|\psi|^2(y)$ and the spin orientation interference $\vec{s}(y)$. The intensity interference part can be derived as

$$\begin{aligned}|\psi|^2(y) &= 2[a(y)^2 + \text{Re}\langle\psi_{II}|\psi_I\rangle(y)] \\ &= 2a(y)^2\left[1 + \text{Re}\left(e^{i\Delta\theta(y)}\langle s_0|\hat{\mathcal{U}}_{[c_0]}|s_0\rangle\right)\right] \\ &= 2a(y)^2[1 + b\cos(\Delta\theta(y) + \delta\theta)],\end{aligned} \tag{25}$$

where $\hat{\mathcal{U}}_{[c_0]} = \hat{U}_{\gamma_{II}}^{-1}\hat{U}_{\gamma_I} = \hat{U}_2^{-1}\hat{U}_1\hat{U}_2\hat{U}_1^{-1}$ is the non-Abelian holonomy of the closed path $c_0 = \gamma_{II}^{-1} \circ \gamma_I$. The nontrivial expectation value of the holonomy of c_0, $\langle s_0|\hat{\mathcal{U}}_{[c_0]}|s_0\rangle = be^{i\delta\theta} \neq 1$, leads to a phase shift $\delta\theta$ and a change of the relative amplitude b (≤ 1) in comparison with the interference result of $\hat{A} = 0$. In the mean time, the interfering spin orientation $\vec{s}(y)$ turns out to be always perpendicular to $\Delta\vec{s} = \vec{s}_I - \vec{s}_{II}$, namely lying on the green great circle of $\vec{s}(y)\cdot\Delta\vec{s} \equiv 0$ in Fig. 3f, and fluctuates around it (see Supplementary Note 7).

We have performed a full-wave simulation of this non-Abelian AB interference as shown in Fig. 3e. In the simulation, the envelope $a(y)$ of each beam is set to be Gaussian type with a central amplitude $a(0) = 1/\sqrt{2}$. The spin density interference is shown in Fig. 3g, with the intensity interference $|\psi|^2(y)$ in Fig. 3h, and the spin orientation given by Euler angles in Fig. 3i, j. In Fig. 3h–j, the blue circles are the simulated results, which are

fairly consistent with the theoretical results (red curves) obtained from Eq. (24).

To demonstrate that the non-Abelian feature of the above design is indeed genuine, we consider a control experiment with an almost identical system except that the vector potential is $\hat{A} \propto \hat{\sigma}_1$ in the whole space. In this case, $\hat{U}_i = \exp[i\Phi_i\hat{\sigma}_1]$ ($i = 1, 2$) commute with each other, and their winding around the two vortices in opposite sequences gives the same AB phase factor $\hat{U}_{\gamma_I} = \hat{U}_{\gamma_{II}} = \exp[i(\Phi_2 - \Phi_1)\hat{\sigma}_1]$. Thus, the interfering spin density is uniformly orientated, and there is no phase shift ($\delta\theta \equiv 0$) and amplitude contraction ($b \equiv 1$) compared with the case of $\hat{A} = 0$ (see green lines in Fig. 3h–j).

Measurement of Wilson loops. In Abelian AB systems, the AB phase factor (holonomy) of a closed loop only depends on the flux inside the loop but independent of the choice of gauge. However, in non-Abelian systems, the holonomy $\hat{\mathcal{U}}_{[c]}[\mathbf{x}_0]$ of a closed path c based at $\mathbf{x}_0$ varies as $\hat{\mathcal{U}}'_{[c]}[\mathbf{x}_0] = \hat{U}(\mathbf{x}_0)\hat{\mathcal{U}}_{[c]}[\mathbf{x}_0]\hat{U}^\dagger(\mathbf{x}_0)$, under a gauge transformation $\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger + i\hat{U}\nabla_T\hat{U}^\dagger$. Nevertheless, the trace of holonomy is an important gauge invariant observable, called the Wilson loop of the closed path c:

$$W(c) = \text{Tr}\left(\mathcal{P}\exp\left[i\oint_c\hat{A}\cdot d\mathbf{r}\right]\right) = \text{Tr}\hat{\mathcal{U}}_{[c]} = \text{Tr}\hat{\mathcal{U}}'_{[c]}. \tag{26}$$

In what follows, we show how to extract the Wilson loop of an arbitrary closed path via interferometry.

In order to obtain the Wilson loop of a homotopy class $[c]$ in a non-Abelian AB system, we consider the interference of two beams along any two paths γ_1 and γ_2 as long as $\gamma_2^{-1} \circ \gamma_1 = c$ forms a closed loop in the class $[c]$ as sketched in Fig. 4a. As we deduced in Eq. (25), the holonomy of c, together with the initial spinor $|s_0\rangle$, determines the phase shift and the relative amplitude through the term $\langle s_0|\hat{\mathcal{U}}_{[c]}|s_0\rangle = be^{i\delta\theta}$. In fact, its real part depends solely on the Wilson loop of c (see proof in the "Methods" section):

$$W(c) = 2\text{Re}\langle s_0|\hat{\mathcal{U}}_{[c]}|s_0\rangle = 2b\cos\delta\theta. \tag{27}$$

Thus, at certain positions y_n satisfying $\Delta\theta(y_n) = n\pi$ (n belongs to integers), the intensities only depend on the Wilson loop of c and hence are fixed under gauge transformation:

$$|\psi|^2(y_n) \equiv a(y_n)^2[2 + (-1)^n W(c)], \tag{28}$$

where the two beams are supposed to share the same envelope $a(y)$ on the screen, and the locations y_n correspond to the crests and troughs in the interference fringes of $\hat{A} = 0$. These particular points in the intensity fringes are termed the gauge fixed points for the closed path c. Since the change of incident spin at $\mathbf{x}_0$ is equivalent to a global gauge transformation, the interference fringes for different incident spins should intersect at the gauge fixed points.

Using the above method, we examine the two optical paths γ_I, γ_{II} in Fig. 3a to extract the Wilson loop of $c_0 = \gamma_{II}^{-1} \circ \gamma_I \simeq c_2^{-1} \circ c_1 \circ c_2 \circ c_1^{-1}$. Figure 4b shows the intensity interference curves corresponding to four different incident spins. Indeed, they intersect exactly at the gauge fixed points (red targets in Fig. 4b) whose locations y_n coincide with the crests and troughs of the interference fringe pattern for $\hat{A} = 0$. By fitting the even and odd subsequences of the gauge fixed points, we obtain two curves $a(y)^2[2 \pm W(c_0)]$ corresponding to the two red dashed lines in Fig. 4b. Thus, the Wilson loop $W(c_0)$ can be identified from the difference of the two dashed curves.
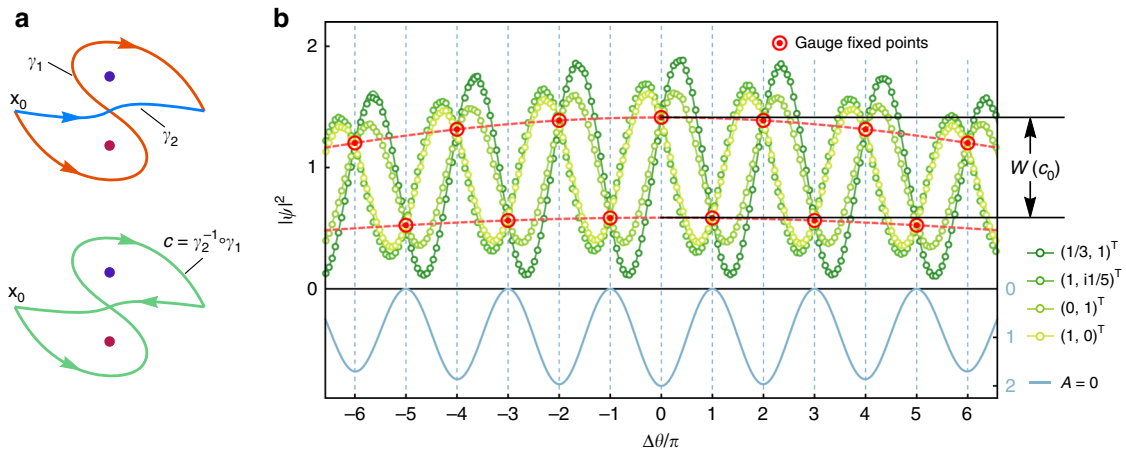
Fig. 4 Extracting Wilson loops from gauge fixed points. **a** For two arbitrary beams γ_1, γ_2 interfering on the screen, the Wilson loop of the concatenate path $c = \gamma_2^{-1} \circ \gamma_1$ can be extracted from the interference fringes of the two beams. **b** Four intensity interference curves corresponding to four different incident spinors $(1,0)^\top$, $(0,1)^\top$, $(1,i1/5)^\top$, $(1/3,1)^\top$ for the non-Abelian AB system shown in Fig. 3a with vortex fluxes $\Phi_1 = 0.22\pi$, $\Phi_2 = -0.33\pi$, where circles and solid curves represent numerical and analytical results, respectively. Their intersections, marked by red targets, are the gauge fixed points, which are located at the crests and troughs of the interference fringes (light blue curve) of $\hat{\mathcal{A}} = 0$. The maximal difference between the envelops of even and odd gauge fixed points gives the Wilson loop $W(c_0)$ of $c_0 = \gamma_{\mathrm{II}}^{-1} \circ \gamma_{\mathrm{I}}$

Discussion

We have shown that the dynamics of 2D optical waves in a broad class of anisotropic media can be understood through an emergent SU(2) gauge interaction in real space. We predicted that the Zitterbewegung effect of light can be realized even in homogeneous anisotropic media, and we proposed a biaxial metamaterial to achieve synthetic non-Abelian electric field and ZB in microwave regime. We have also designed a genuine non-Abelian AB system with two synthetic non-Abelian vortices, and suggested a spin density interferometry to demonstrate the non-commutative feature of non-Abelian holonomies. Our scheme opens the door to the colorful non-Abelian world for light. In addition to inspiring new ideas to manipulate the flow and polarization of light, the scheme offers an optical platform to study physical effects relevant to SU(2) gauge fields, such as synthetic spin–orbit coupling[59] and topological band structures in periodic non-Abelian gauge fields[60–63]. Furthermore, since the SU(2) gauge field description is valid for photons down to quantum scale, this approach might be applicable to the design of geometric gates for realizing non-Abelian holonomic quantum computation[64,65] with photons.

Methods

Notations. In this paper, vectors in real space and in pseudo-spin space are indicated, respectively, by bold letters and letters with an overhead arrow "→". Letters with an overhead bidirectional arrow "↔" denote two-order tensors in real space. Symbols with an overhead hat "∧" denote operators acting on the spinor wave functions. We use Greek letters, e.g. μ, ν, to denote indices of (2+1)-dimensional spacetime. Latin letters i, j denote 2D spatial coordinate indices, and Latin letters a, b, c denote indices in pseudo-spin space. We follow the Einstein summation convention for repeated indices. The orthonormal coordinate bases in real space and pseudo-spin space are expressed as $\mathbf{e}_i$ and $\vec{e}_a$, respectively.

SU(2) gauge covariance of 2D Maxwell equations. In block-diagonalized duality symmetric media, $\overleftrightarrow{\varepsilon}/\varepsilon_0 = \overleftrightarrow{\mu}/\mu_0 = \mathrm{diag}(\overleftrightarrow{\varepsilon}_T, \varepsilon_z)$, the Maxwell's equations for 2D waves can be rearranged as

$$\left(\begin{array}{c|c} 0 & i\hat{\sigma}_2 (\mathrm{i}\nabla_T \times) \mathbf{e}_z \\ \hline i\hat{\sigma}_2 \mathbf{e}_z \cdot (\mathrm{i}\nabla_T \times) & 0 \end{array} \right) \Psi = k_0 \underbrace{\left(\begin{array}{c|c} \hat{\sigma}_0 \overleftrightarrow{\varepsilon}_T & 0 \\ \hline 0 & \hat{\sigma}_0 \varepsilon_z \end{array} \right)}_{\mathcal{N}} \Psi, \tag{29}$$

with $\Psi = (\mathbf{E}_T, \eta_0\mathbf{H}_T, E_z, \eta_0 H_z)^\top$. For an arbitrary (global) transformation $\hat{U} \in$ SU(2) acting on $|\psi\rangle = (E_z, \eta_0 H_z)^\top$, the corresponding transformation for Ψ is

defined as

$$\tilde{U} = \hat{U}_T \oplus \hat{U} = (\hat{\sigma}_2 \hat{U} \hat{\sigma}_2) \oplus \hat{U}, \tag{30}$$

which belongs to a 4D representation of SU(2). It turns out that $\mathcal{M}$ and $\mathcal{N}$ defined in Eq. (29) transform according to

$$\tilde{U}\mathcal{M}\tilde{U}^\dagger = \left(\begin{array}{c|c} 0 & i\hat{U}_T \hat{\sigma}_2 U^\dagger (\mathrm{i}\nabla_T \times)\mathbf{e}_z \\ \hline i\hat{U}\hat{\sigma}_2 \hat{U}_T^\dagger \mathbf{e}_z \cdot (\mathrm{i}\nabla_T \times) & 0 \end{array} \right) = \mathcal{M}, \tag{31}$$

$$\tilde{U}\mathcal{N}\tilde{U}^\dagger = \mathcal{N}. \tag{32}$$

Hence, the 2D Maxwell equations are invariant under this SU(2) transformation. As the EM duality transformation $\hat{R} \in$ SO(2) is a special case of $\hat{U}$, the emergent SU(2) symmetry for the 2D Maxwell equations in block-diagonalized duality symmetric materials is indeed the generalization of the original EM duality symmetry.

If $\hat{U}(x,y)$ is dependent on the x, y coordinates, the transformation of $\mathcal{M}$ changes to

$$\tilde{U}(x,y)\mathcal{M}\tilde{U}^\dagger(x,y) = \mathcal{M} + \Delta\mathcal{M} \tag{33}$$

with an additional term

$$\begin{aligned}\Delta\mathcal{M} &= \left(\begin{array}{c|c} 0 & i\hat{\sigma}_2 (i\hat{U}\nabla_T \hat{U}^\dagger) \times \mathbf{e}_z \\ \hline -i(i\hat{U}\nabla_T \hat{U}^\dagger) \times \mathbf{e}_z \hat{\sigma}_2 & 0 \end{array} \right) \\ &= \left(\begin{array}{c|c} 0 & (\hat{\sigma}_3 \mathcal{A}^1 + \mathrm{i}\hat{\sigma}_0 \mathcal{A}^2) \times \mathbf{e}_z \\ \hline (\hat{\sigma}_3 \mathcal{A}^1 - \mathrm{i}\hat{\sigma}_0 \mathcal{A}^2) \times \mathbf{e}_z & 0 \end{array} \right),\end{aligned} \tag{34}$$

where $\mathcal{A}^a \hat{\sigma}_a = \mathrm{i}\hat{U}\nabla_T \hat{U}^\dagger$ is precisely the vector potential induced purely by the gauge transformation, and only the components $\mathcal{A}^1$, $\mathcal{A}^2$ are supposed to exist. If we move the term $\Delta\mathcal{M}$ to the right side of the Maxwell equation (29), it can be alternatively interpreted as a part of the constitutive tensor. By rotating Ψ to the ordinary basis of EM field,

$$\left(\begin{array}{c} \mathbf{E} \\ \eta_0 \mathbf{H} \end{array} \right) = U_0 \Psi = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) \left(\begin{array}{c} \mathbf{E}_T \\ \eta_0 \mathbf{H}_T \\ \hline E_z \\ \eta_0 H_z \end{array} \right), \tag{35}$$

we obtain explicitly the contribution of $\Delta\mathcal{M}$ to the material tensors

$$\left(\begin{array}{cc} \varepsilon & 0 \\ 0 & \mu \end{array} \right) = U_0(\mathcal{N} - \Delta\mathcal{N}/k_0)U_0^\dagger, \tag{36}$$

which shows that the effective SU(2) vector potential $\hat{\mathcal{A}}$ emerging in $\Delta\mathcal{M}$ just

corresponds to the off-diagonal terms of $\vec{\bar{\varepsilon}}, \vec{\bar{\mu}}$:

$$\mathbf{g}_1 = -\mathbf{g}_2^* = \mathbf{e}_z \times (\mathcal{A}^1 + i\mathcal{A}^2)/k_0. \qquad (37)$$

Indeed, this relation is valid for arbitrary $\hat{\mathcal{A}} = \mathcal{A}^1\hat{\sigma}_1 + \mathcal{A}^2\hat{\sigma}_2$ but not limited to the pure gauge case $\hat{\mathcal{A}} = i\hat{U}\nabla_T\hat{U}^\dagger$. Furthermore, this correspondence can be generalized to any media satisfying in-plane duality condition $\vec{\bar{\varepsilon}}_T = \alpha\vec{\bar{\mu}}_T$ where SU(2) scalar potential may also appear (Supplementary Note 1).

Quasi-degenerate approximation for ZB. Eq. (2) is essentially the stationary wave equation describing spin-1/2 particles coupling to the background SU(2) gauge fields without any approximation. However, the semiclassical trajectories of non-degenerate eigenmodes often split away from each other. To manifest the coupling effects of different eigenmodes in the geometric optics, the media of concern are usually assumed to be weakly anisotropic[14]. Nevertheless, if the eigenmodes are approximately degenerate in a particular direction of wave vector but not necessarily in all directions, it turns out that an intact wave composed of modes in the vicinity of the quasi-degenerate direction can be described adequately by the semiclassical approach including the interaction between eigenmodes in their interfering region[66].

In homogeneous non-Abelian media, we separate the effective Hamiltonian into two parts:

$$\hat{H}(\mathbf{k}) = \underbrace{\left[\frac{1}{2m}\left(\mathbf{k}^2 + (\hat{\mathcal{A}})^2\right) + V_0\right]\hat{\sigma}_0}_{\hat{H}_0(\mathbf{k})} + \underbrace{\left(\frac{-1}{m}\mathbf{k}\cdot\hat{\mathcal{A}} - \hat{\mathcal{A}}_0\right)}_{\delta\hat{H}(\mathbf{k}) = \vec{\Omega}\cdot\vec{\sigma}/2}. \qquad (38)$$

If only $\hat{H}_0$ is present, the isofrequency surface is a doubly degenerate sphere with the radius $k = \sqrt{-2mV_0 - (\hat{\mathcal{A}})^2}$. When $\delta\hat{H}(\mathbf{k})$ is taken into account, as long as it is sufficiently small for a given direction $\mathbf{e}_k$, the two eigenstates can be regarded as quasi-degenerate at the wave vector

$$\mathbf{k} = k\mathbf{e}_k = \sqrt{-2mV_0 - (\hat{\mathcal{A}})^2}\,\mathbf{e}_k, \qquad (39)$$

and we can implement the eikonal approximation to the wave function mainly superposed by the two quasi-degenerate modes[66]: $|\psi\rangle = \widetilde{\psi}(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{r})$ with a slowly varying envelope $\widetilde{\psi}(\mathbf{r})$ (i.e. $|\nabla\widetilde{\psi}/\widetilde{\psi}| \ll k$). Substituting $|\psi\rangle$ into the wave equation (2), we obtain the equation of $\widetilde{\psi}$ with accuracy up to the first order of k:

$$i\widehat{\mathbf{v}}\cdot\nabla\widetilde{\psi} = \hat{H}(\mathbf{k})\widetilde{\psi}. \qquad (40)$$

By adopting the ansatz that the velocity operator $\widehat{\mathbf{v}} = \partial\hat{H}(\mathbf{k})/\partial\mathbf{k} = (\mathbf{k} - \hat{\mathcal{A}})/m$ can be replaced by its averaged value $\langle\widehat{\mathbf{v}}\rangle$ over the transverse cross section of an optical beam, we find that the operator $\widehat{\mathbf{v}}\cdot\nabla \to \langle\widehat{\mathbf{v}}\rangle\cdot\nabla = d/d\tau$ corresponds to the total derivative with respect to the ray parameter τ along the beam. Therefore, Eq. (40) is reformulated into a time-dependent Schrödinger equation

$$i\frac{d}{d\tau}\widetilde{\psi} = \hat{H}(\mathbf{k})\widetilde{\psi}. \qquad (41)$$

Consequently, Eqs. (8-10) can be directly obtained in terms of Ehrenfest theorem.

Relation between Poynting vector and velocity operator. The in-plane projection of the time-averaged Poynting vector $\bar{\mathbf{S}}_T$ for monochromatic waves can be written as

$$\begin{aligned}\bar{\mathbf{S}}_T &= \tfrac{1}{2}\mathrm{Re}\left[E_z^*\times\mathbf{H}_T + \mathbf{E}_T^*\times H_z\right] \\ &= \tfrac{1}{2}\mathrm{Re}\left[(E_z^*, \mathbf{H}_z^*)\begin{pmatrix} 0 & \overleftrightarrow{\mathbf{I}}\times\overleftrightarrow{\mathbf{I}} \\ -\overleftrightarrow{\mathbf{I}}\times\overleftrightarrow{\mathbf{I}} & 0 \end{pmatrix}\begin{pmatrix} \mathbf{E}_T \\ \mathbf{H}_T \end{pmatrix}\right] \\ &= \tfrac{1}{2\eta_0}\mathrm{Re}\left[(E_z^*, \eta_0 H_z^*)(i\hat{\sigma}_2\varepsilon^{izj}\mathbf{e}_i)\begin{pmatrix} \mathbf{E}_T \\ \eta_0\mathbf{H}_T \end{pmatrix}_j\right].\end{aligned} \qquad (42)$$

Substituting Maxwell's equations into Eq. (42) yields

$$\begin{aligned}\bar{\mathbf{S}}_T &= \tfrac{1}{2\eta_0}\mathrm{Re}\langle\psi|(i\hat{\sigma}_2\varepsilon^{izj}\mathbf{e}_i)\tfrac{-i\hat{\sigma}_2\,\varepsilon_{jkz}}{k_0\,\varepsilon_T}(\widehat{\mathbf{p}} - \hat{\mathcal{A}}_{(c)})^k|\psi\rangle \\ &= \tfrac{1}{2\eta_0 k_0}\mathrm{Re}\langle\psi|\tfrac{1}{2m}(\widehat{\mathbf{p}} - \hat{\mathcal{A}}_{(c)})|\psi\rangle \\ &= \tfrac{1}{4\mu_0\omega_0}\langle\psi|\tfrac{1}{m}(\mathbf{k} - \hat{\mathcal{A}})|\psi\rangle = \tfrac{1}{4\mu_0\omega_0}\langle\psi|\widehat{\mathbf{v}}|\psi\rangle,\end{aligned} \qquad (43)$$

where $\hat{\mathcal{A}}_{(c)} = \frac{k_0}{2}\left\{[(\mathbf{g}_1 - \mathbf{g}_2)\times\mathbf{e}_z]\hat{\sigma}_1 - i[(\mathbf{g}_1 + \mathbf{g}_2)\times\mathbf{e}_z]\hat{\sigma}_2\right\}$. In the third step, we replaced $\widehat{\mathbf{p}}$ with $\mathbf{k}$ according to the eikonal approximation. As a result, the total in-plane energy flux over a transverse cross section of the optical beam is proportional to the expectation value of the velocity operator:

$$\langle\bar{\mathbf{S}}_T\rangle = \frac{1}{4\mu_0\omega_0}\int d\mathbf{r}_\perp\langle\psi|\widehat{\mathbf{v}}|\psi\rangle = \frac{1}{4\mu_0\omega_0}\langle\widehat{\mathbf{v}}\rangle. \qquad (44)$$

And it shows that the time-averaged Poynting vector $\bar{\mathbf{S}}_T$ is invariant under the gauge transformation Eq. (30) for EM fields (Supplementary Note 2).

Holonomy and genuine non-Abelian AB system. From a geometric viewpoint, gauge potential and field in the physical space M can be described as the connection and curvature in a G-principle fiber bundle[32], where the physical space serves as the base manifold, and G denotes the gauge group, in our case $G = \mathrm{SU}(2)$. Consider a particle (wave packet) travels in the physical space. Along its trajectory γ, the gauge potential engenders a matrix-valued geometric phase factor $\mathcal{P}\exp\left[i\int_\gamma\hat{\mathcal{A}}_\mu dx^\mu\right] \in G$ ($\mathcal{P}$ denotes path-ordering) on the state vector, corresponding to the parallel transport of the state in the bundle space. In particular, for a closed path c starting and ending at the same point $c(0) = c(1) = \mathbf{x}_0$, the phase factor of c,

$$\hat{\mathcal{U}}_c(\hat{\mathcal{A}}) = \mathcal{P}\exp\left[i\oint_c\hat{\mathcal{A}}_\mu dx^\mu\right], \qquad (45)$$

is called the holonomy of the closed path c with respect to the gauge $\hat{\mathcal{A}}$. The collection of the holonomies corresponding to all those closed paths based at the same point $\mathbf{x}_0$ forms a subgroup of the gauge group G:

$$\mathrm{Hol}(\hat{\mathcal{A}}) = \left\{\hat{\mathcal{U}}_c(\hat{\mathcal{A}})|c(0) = c(1) = \mathbf{x}_0\right\} \subseteq G, \qquad (46)$$

which is the holonomy group for the gauge $\hat{\mathcal{A}}$. In the literature, a gauge system is regarded as genuinely non-Abelian if and only if the holonomy group is a non-Abelian group, namely the holonomies of some loops are noncommutable with each other[10,44]. If the base manifold is simply a Euclidean space, the non-commutativity of holonomies can be traced back to noncommutable gauge fields $[\hat{\mathcal{F}}_{\mu\nu}, \hat{\mathcal{F}}_{\mu'\nu'}]\neq 0$. However, if the base manifold possesses nontrivial topology, non-commutative holonomies can be achieved even though the gauge field vanishes everywhere (i.e. AB systems).

For an AB system, the corresponding fiber bundle is a flat bundle, since the curvature (field) $\hat{\mathcal{F}}_{\mu\nu} = 0$ in the whole base manifold M (flux regions are excluded from M). Here, the topology of the base manifold is characterized by its first fundamental group,

$$\pi_1(M) = \{[c]\,|\,c(0) = c(1) = \mathbf{x}_0\}, \qquad (47)$$

which is the set of path homotopy equivalent classes $[c]$ of closed paths based at $\mathbf{x}_0$. Path homotopy is a topologically equivalent relation "$\simeq$" for paths. If two paths c_1, c_1 with the same fixed base-point $\mathbf{x}_0$ can deform into each other continuously, they are said to be path homotopic $c_1 \simeq c_2$ and to belong to the same homotopy class $[c_1]$. In flat bundles, the holonomies (AB phase factors) of all loops in the same homotopy class $[c]$ are identical: $\hat{\mathcal{U}}_{[c]}$ (see proof in Supplementary Note 5). Based on this property, two necessary conditions for genuine non-Abelian AB systems can be obtained[44]:

1. The gauge group G is non-Abelian;
2. The first fundamental group $\pi_1(M)$ is non-Abelian.

According to the second criterion, the Wu–Yang AB system is not genuinely non-Abelian, because the fundamental group of its base manifold (a punctured plane $\mathbb{R}^2 - \mathbf{0}$) is an Abelian group $\pi_1(\mathbb{R}^2 - \mathbf{0}) = \mathbb{Z}$. However, for a twice-punctured plane as shown in Fig. 3a, its fundamental group is the free group on two generators, $\mathbb{Z} * \mathbb{Z}$ (where $*$ denotes a free product), which is non-Abelian[67]. Therefore, a twice-punctured plane is a qualified prototype of a genuine non-Abelian AB system.

Gauge fixed points. The derivation of the intensity interference given by Eq. (25) is in fact valid for two arbitrary interfering beams γ_1, γ_1 with the same initial spin $\vec{s}_0$ and final envelop $a(y)$: $|\psi|^2 = 2a(y)[1 + \mathrm{Re}(e^{i\Delta\theta(y)}\langle s_0|\hat{\mathcal{U}}_{[c]}|s_0\rangle)]$, where $\hat{\mathcal{U}}_{[c]}$ is the holonomy of the closed path $c = \gamma_2^{-1}\circ\gamma_1$. Since $\hat{\mathcal{U}}_{[c]} \in \mathrm{SU}(2)$, it can be generically expressed as

$$\hat{\mathcal{U}}_{[c]} = \begin{pmatrix} u_1 & u_2 \\ -u_2^* & u_1^* \end{pmatrix}, \quad \text{with}\quad |u_1|^2 + |u_2|^2 = 1. \qquad (48)$$

The Wilson loop reads $W(c) = \mathrm{Tr}\hat{\mathcal{U}}_{[c]} = 2\,\mathrm{Re}(u_1)$. For an arbitrary spinor state $|s_0\rangle = \left(\cos\frac{\alpha}{2}e^{-i\beta/2}, \sin\frac{\alpha}{2}e^{i\beta/2}\right)^\top$, we have

$$\begin{aligned}\langle s_0|\hat{\mathcal{U}}_{[c]}|s_0\rangle &= be^{i\delta\theta} \\ &= \mathrm{Re}(u_1) + i\left[\cos\alpha\,\mathrm{Im}(u_1) + \sin\alpha\,\mathrm{Im}(u_2 e^{i\beta})\right].\end{aligned} \qquad (49)$$

Therefore, the identity in Eq. (27) is established for any $|s_0\rangle$.

In fact, different incident spinors can interconvert through a global gauge transformation: $|s_0'\rangle = \hat{U}|s_0\rangle$. Hence, the relation in Eq. (27) is straightforward

$$\begin{aligned}2\,\mathrm{Re}\langle s_0'|\hat{\mathcal{U}}_{[c]}|s_0'\rangle &= 2\,\mathrm{Re}\langle s_0|\hat{U}^{-1}\hat{\mathcal{U}}_{[c]}\hat{U}|s_0\rangle \\ &= \mathrm{Tr}\left(\hat{U}^{-1}\hat{\mathcal{U}}_{[c]}\hat{U}\right) \equiv W(c).\end{aligned} \qquad (50)$$

As a result, at the positions such that $\Delta\theta(y_n) = n\pi$, i.e. at the crests and troughs of the original interference fringes when $\hat{\mathcal{A}} = 0$, the intensities given in Eq. (28) are fixed for arbitrary incident spins, yet they are only determined by the Wilson loop $W(c)$, provided that the dynamic phases of γ_1, γ_2 are unchanged.

For the two optical path γ_{I}, γ_{II} in Fig. 3, the Wilson loop of $c_0 = \gamma_{\mathrm{II}}^{-1} \circ \gamma_{\mathrm{I}}$ is determined by the fluxes of the two vortices as $W(c_0) = 2 - 4\sin^2\Phi_1\sin^2\Phi_2$. Therefore, if $\sin^2\Phi_1\sin^2\Phi_2 = 1/2$, $W(c_0)$ will be reduced to zero, and the two dashed curves in Fig. 4 will completely overlap (see Supplementary Note 7 for details).

Simulation of non-Abelian AB interference. The full-wave results of the non-Abelian AB interference shown in Figs. 3 and 4 are simulated using the commercial software COMSOL Multiphysics. In order to avoid spin–flip after reflection, the mirrors shown in Fig. 3a, e are made of an impedance-matched material, namely $\varepsilon_m/\mu_m = 1$, with a lower refractive index than the surrounding media to achieve total reflection at their surfaces. Meanwhile, the two mirrors on the right-hand side in Fig. 3e are slightly concave, so that the reflected beams with reduced widths can bypass the two singularities.

Data availability

The authors declare that all data supporting the findings of this study are available from the corresponding authors upon reasonable request.

References

1. Yang, C. N. & Mills, R. L. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.* **96**, 191–195 (1954).
2. Mead, C. A. & Truhlar, D. G. On the determination of Born–Oppenheimer nuclear motion wave functions including complications due to conical intersections and identical nuclei. *J. Chem. Phys.* **70**, 2284–2296 (1979).
3. Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A* **392**, 45–57 (1984).
4. Wilczek, F. & Zee, A. Appearance of gauge structure in simple dynamical systems. *Phys. Rev. Lett.* **52**, 2111–2114 (1984).
5. Aidelsburger, M., Nascimbene, S. & Goldman, N. Artificial gauge fields in materials and engineered systems. *C. R. Phys.* **19**, 394–432 (2018).
6. Xiao, D., Chang, M.-C. & Niu, Q. Berry phase effects on electronic properties. *Rev. Mod. Phys.* **82**, 1959–2007 (2010).
7. Fujita, T., Jalil, M., Tan, S. & Murakami, S. Gauge fields in spintronics. *J. Appl. Phys.* **110**, 121301 (2011).
8. Lin, Y., Compton, R., Jiménez-García, K., Porto, J. & Spielman, I. Synthetic magnetic fields for ultracold neutral atoms. *Nature* **462**, 628–632 (2009).
9. Dalibard, J., Gerbier, F., Juzeliūnas, G. & Öhberg, P. Colloquium: artificial gauge potentials for neutral atoms. *Rev. Mod. Phys.* **83**, 1523–1543 (2011).
10. Goldman, N., Juzeliūnas, G., Öhberg, P. & Spielman, I. B. Light-induced gauge fields for ultracold atoms. *Rep. Prog. Phys.* **77**, 126401 (2014).
11. Onoda, M., Murakami, S. & Nagaosa, N. Hall effect of light. *Phys. Rev. Lett.* **93**, 083901 (2004).
12. Bliokh, K. Y. & Bliokh, Y. P. Modified geometrical optics of a smoothly inhomogeneous isotropic medium: the anisotropy, Berry phase, and the optical magnus effect. *Phys. Rev. E* **70**, 026605 (2004).
13. Onoda, M., Murakami, S. & Nagaosa, N. Geometrical aspects in optical wave-packet dynamics. *Phys. Rev. E* **74**, 066610 (2006).
14. Bliokh, K. Y., Frolov, D. Y. & Kravtsov, Y. A. Non-Abelian evolution of electromagnetic waves in a weakly anisotropic inhomogeneous medium. *Phys. Rev. A* **75**, 053821 (2007).
15. Bliokh, K. Y., Niv, A., Kleiner, V. & Hasman, E. Geometrodynamics of spinning light. *Nat. Photon.* **2**, 748–753 (2008).
16. Ma, L. et al. Spin–orbit coupling of light in asymmetric microcavities. *Nat. Commun.* **7**, 10983 (2016).
17. Sawada, K. & Nagaosa, N. Optical magnetoelectric effect in multiferroic materials: evidence for a Lorentz force acting on a ray of light. *Phys. Rev. Lett.* **95**, 237402 (2005).
18. Wang, Z., Chong, Y. D., Joannopoulos, J. D. & Soljačić, M. Reflection-free one-way edge modes in a gyromagnetic photonic crystal. *Phys. Rev. Lett.* **100**, 013905 (2008).
19. Fang, K. & Fan, S. Effective magnetic field for photons based on the magneto-optical effect. *Phys. Rev. A* **88**, 043847 (2013).
20. Hafezi, M., Demler, E. A., Lukin, M. D. & Taylor, J. M. Robust optical delay lines with topological protection. *Nat. Phys.* **7**, 907–912 (2011).
21. Fang, K., Yu, Z. & Fan, S. Photonic Aharonov–Bohm effect based on dynamic modulation. *Phys. Rev. Lett.* **108**, 153901 (2012).
22. Fang, K., Yu, Z. & Fan, S. Realizing effective magnetic field for photons by controlling the phase of dynamic modulation. *Nat. Photon.* **6**, 782–787 (2012).
23. Fang, K. & Fan, S. Controlling the flow of light using the inhomogeneous effective gauge field that emerges from dynamic modulation. *Phys. Rev. Lett.* **111**, 203901 (2013).
24. Rechtsman, M. C. et al. Strain-induced pseudomagnetic field and photonic Landau levels in dielectric structures. *Nat. Photon.* **7**, 153–158 (2013).
25. Jia, H. et al. Observation of chiral zero mode in inhomogeneous three-dimensional Weyl metamaterials. *Science* **363**, 148–151 (2019).
26. Chen, W.-J. et al. Experimental realization of photonic topological insulator in a uniaxial metacrystal waveguide. *Nat. Commun.* **5**, 5782 (2014).
27. Liu, F. & Li, J. Gauge field optics with anisotropic media. *Phys. Rev. Lett.* **114**, 103902 (2015).
28. Liu, F., Wang, S., Xiao, S., Hang, Z. H. & Li, J. Polarization-dependent optics using gauge-field metamaterials. *Appl. Phys. Lett.* **107**, 241106 (2015).
29. Liu, F., Horsley, S. A. R. & Li, J. Invisibility cloaking using pseudomagnetic field for photon. *Phys. Rev. B* **95**, 075157 (2017).
30. Liu, F., Xu, T., Wang, S., Hang, Z. H. & Li, J. Polarization beam splitting with gauge field metamaterials. *Adv. Opt. Mater.* **7**, 1801582 (2019).
31. Schine, N., Ryou, A., Gromov, A., Sommer, A. & Simon, J. Synthetic Landau levels for photons. *Nature* **534**, 671–675 (2016).
32. Wu, T. T. & Yang, C. N. Concept of nonintegrable phase factors and global formulation of gauge fields. *Phys. Rev. D* **12**, 3845–3857 (1975).
33. Qi, X.-L. & Zhang, S.-C. Topological insulators and superconductors. *Rev. Mod. Phys.* **83**, 1057–1110 (2011).
34. Senthil, T. Symmetry-protected topological phases of quantum matter. *Annu. Rev. Condens. Matter Phys.* **6**, 299–324 (2015).
35. Lu, L., Joannopoulos, J. D. & Soljačić, M. Topological photonics. *Nat. Photon.* **8**, 821–829 (2014).
36. Ozawa, T. et al. Topological photonics. *Rev. Mod. Phys.* **91**, 015006 (2019).
37. Zawadzki, W. & Rusin, T. M. Zitterbewegung (trembling motion) of electrons in semiconductors: a review. *J. Phys.: Condens. Matter* **23**, 143201 (2011).
38. Vaishnav, J. Y. & Clark, C. W. Observing Zitterbewegung with ultracold atoms. *Phys. Rev. Lett.* **100**, 153002 (2008).
39. Gerritsma, R. et al. Quantum simulation of the Dirac equation. *Nature* **463**, 68–71 (2010).
40. Zhang, X. Observing Zitterbewegung for photons near the Dirac point of a two-dimensional photonic crystal. *Phys. Rev. Lett.* **100**, 113903 (2008).
41. Dreisow, F. et al. Classical simulation of relativistic Zitterbewegung in photonic lattices. *Phys. Rev. Lett.* **105**, 143902 (2010).
42. Fan, Y. et al. Plasmonic zitterbewegung in binary graphene sheet arrays. *Opt. Lett.* **40**, 2945–2948 (2015).
43. Guo, R.-P., Guo, Q.-H., Wu, L.-T., Chen, J. & Fan, D. Optical spin-sensitive Zitterbewegung in bianisotropic metamaterials. *Opt. Express* **24**, 13788–13799 (2016).
44. Sundrum, R. & Tassie, L. J. Non-Abelian Aharonov–Bohm effects, Feynman paths, and topology. *J. Math. Phys.* **27**, 1566–1570 (1986).
45. Wong, S. K. Field and particle equations for the classical Yang–Mills field and particles with isotopic spin. *Il. Nuovo Cimento A (1965–1970)* **65**, 689–694 (1970).
46. Pendry, J. B., Schurig, D. & Smith, D. R. Controlling electromagnetic fields. *Science* **312**, 1780–1782 (2006).
47. Schurig, D., Pendry, J. B. & Smith, D. R. Calculation of material properties and ray tracing in transformation media. *Opt. Express* **14**, 9794–9804 (2006).
48. Chen, H., Chan, C. T. & Sheng, P. Transformation optics and metamaterials. *Nat. Mater.* **9**, 387–396 (2010).
49. Leonhardt, U. & Philbin, T. G. General relativity in electrical engineering. *New J. Phys.* **8**, 247 (2006).
50. Leonhardt, U. & Philbin, T. G. Transformation optics and the geometry of light. *Prog. Opt.* **53**, 69–152 (2009).
51. Genov, D. A., Zhang, S. & Zhang, X. Mimicking celestial mechanics in metamaterials. *Nat. Phys.* **5**, 687–692 (2009).
52. Chen, H., Miao, R.-X. & Li, M. Transformation optics that mimics the system outside a schwarzschild black hole. *Opt. Express* **18**, 15183–15188 (2010).
53. Sheng, C., Liu, H., Wang, Y., Zhu, S. & Genov, D. Trapping light by mimicking gravitational lensing. *Nat. Photon.* **7**, 902–906 (2013).
54. Shen, S.-Q. Spin transverse force on spin current in an electric field. *Phys. Rev. Lett.* **95**, 187203 (2005).
55. Liu, R., Cui, T. J., Huang, D., Zhao, B. & Smith, D. R. Description and explanation of electromagnetic behaviors in artificial metamaterials based on effective medium theory. *Phys. Rev. E* **76**, 026606 (2007).
56. Aharonov, Y. & Bohm, D. Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.* **115**, 485–491 (1959).
57. Batelaan, H. & Tonomura, A. The Aharonov–Bohm effects: variations on a subtle theme. *Phys. Today* **62**, 38–43 (2009).
58. Horváthy, P. A. Non-Abelian Aharonov–Bohm effect. *Phys. Rev. D* **33**, 407–414 (1986).
59. Galitski, V. & Spielman, I. B. Spin–orbit coupling in quantum gases. *Nature* **494**, 49–54 (2013).

60. Osterloh, K., Baig, M., Santos, L., Zoller, P. & Lewenstein, M. Cold atoms in non-Abelian gauge potentials: from the Hofstadter "moth" to lattice gauge theory. *Phys. Rev. Lett.* **95**, 010403 (2005).
61. Goldman, N. et al. Non-Abelian optical lattices: anomalous quantum Hall effect and Dirac fermions. *Phys. Rev. Lett.* **103**, 035301 (2009).
62. Goldman, N., Kubasiak, A., Gaspard, P. & Lewenstein, M. Ultracold atomic gases in non-Abelian gauge potentials: The case of constant Wilson loop. *Phys. Rev. A* **79**, 023624 (2009).
63. Lepori, L., Fulga, I. C., Trombettoni, A. & Burrello, M. Double Weyl points and Fermi arcs of topological semimetals in non-Abelian gauge potentials. *Phys. Rev. A* **94**, 053633 (2016).
64. Zanardi, P. & Rasetti, M. Holonomic quantum computation. *Phys. Lett. A* **264**, 94–99 (1999).
65. Duan, L.-M., Cirac, J. & Zoller, P. Geometric manipulation of trapped ions for quantum computation. *Science* **292**, 1695–1697 (2001).
66. Kravtsov, Y. A. & Orlov, Y. I. *Geometrical Optics of Inhomogeneous Media*. (Springer-Verlag, Berlin, 1990).
67. Massey, W. S. *Algebraic Topology: An Introduction*. (Springer-Verlag, New York, 1977).

Acknowledgements

Author contributions

Y.C. conceived the original idea. Y.C., J.Q.S. developed the standard framework and discovered ZB effect. R.-Y.Z., Y.C. extended the framework to more general cases. R.-Y.Z. performed analytic study of ZB effect, and designed the scheme of non-Abelian AB effect. Y.C., Z.X. R.-Y.Z. carried out the numerical simulations. Z.H.H., J.L. designed the metamaterial. C.T.C. supervised the whole project. R.-Y.Z., Y.C., C.T.C. wrote the manuscript. And all authors were involved in the analysis and discussion of the results.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-10974-8.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Peer review information: *Nature Communications* thanks Simon Horsley and other anonymous reviewer(s) for their contribution to the peer review of this work.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.