

Inverse Hamiltonian design by automatic differentiation

Koji Inui ^{1,2}✉ & Yukitoshi Motome ¹

An ultimate goal of materials science is to deliver materials with desired properties at will. Solving the inverse problem to obtain an appropriate Hamiltonian directly from the desired properties has the potential to reach qualitatively new principles, but most research to date has been limited to quantitative determination of parameters within known models. Here, we develop a general framework that can automatically design a Hamiltonian with desired physical properties by using automatic differentiation. In the application to the quantum anomalous Hall effect, our framework can not only construct the Haldane model automatically but also generate Hamiltonians that exhibit a six-times larger anomalous Hall effect. In addition, the application to the photovoltaic effect gives an optimal Hamiltonian for electrons moving on a noncoplanar spin texture, which can generate $\sim 700 \text{ Am}^{-2}$ under solar radiation. This framework would accelerate materials exploration by automatic construction of models and principles.

¹Department of Applied Physics, The University of Tokyo, Hongo, Tokyo 113-8656, Japan. ²RIKEN Center for Quantum Computing (RQC), Hirosawa 2-1, Wako, Saitama 351-0198, Japan. ✉email: koji.inui@riken.jp

A conventional theoretical approach to materials exploration is to search for Hamiltonians that produce physical properties of interest (Fig. 1a). This is not only tedious but also nontrivial since the parameter space to be explored is usually unknown a priori. Therefore, most of the research to date has been conducted for the known Hamiltonians and their extensions. However, these approaches make it difficult to reach qualitatively new models and principles. In contrast, the inverse approach to find appropriate Hamiltonians directly from the desired properties is not only efficient but also has the potential to unveil qualitatively new physics (Fig. 1a). Many proposals have been made for the inverse approach^{1–13}. Since the early stage, the perturbation theory^{2,7,14}, the potential interpolation^{15,16}, and the eigenstate-to-Hamiltonian construction¹⁷ have been employed, but their applications were limited to the objective functions in terms of energy. In recent years, machine learning-based methods, such as the generative models using neural networks^{1,18,19}, the Bayesian optimization using Gaussian processes^{5,20,21}, and the genetic algorithms^{22,23} have been developed, but they require numerous data and computational resources for training. In particular, the Bayesian optimization and the genetic algorithms do not necessarily improve the objective function after parameter update, and the generative models would fail in the parameter space where data is insufficient. For these reasons, the previous research has been limited to the quantitative estimation of a few parameters within known Hamiltonians. Thus, it is still challenging to explore new models and principles by taking full advantage of the inverse problem.

To address these issues, we develop a framework that can automatically design a Hamiltonian with desired physical properties by using automatic differentiation. Automatic differentiation enables us to compute the analytic derivatives of any functions by adapting chain rules, which have been widely used in the field of deep learning in the process of backpropagation²⁴, even for over a trillion parameters²⁵. In recent years, automatic differentiation has been applied to physics, such as computing physical quantities represented by derivatives^{26,27}, calculating

conditions for solar cells²⁸, applications to quantum gate control^{29–31}, non-equilibrium steady state³², numerical renormalization group³³, Hartree–Fock calculation^{34,35}, molecular dynamics³⁶, and density functional theory^{37,38}. However, the application to the inverse design of a Hamiltonian has not been fully explored thus far to the best of our knowledge.

In this article, we first describe the framework and its advantages over previous methods. Then, we demonstrate a proof of concept of this framework by applying it to two problems: the anomalous Hall effect (AHE) and the photovoltaic effect (PVE). We show that our framework can automatically construct the Haldane model with the quantum AHE on the honeycomb lattice. Moreover, by applying the framework to a model on the triangular lattice, we find a Hamiltonian that exhibits a six-time larger AHE than that of the Haldane model. For the PVE, we are able to automatically generate a spin-charge-coupled Hamiltonian with electrons moving over an umbrella-shaped spin configuration, which can produce a photocurrent of about 700 A m⁻². Our framework is applicable to a wide range of systems and physical properties, including first-principles Hamiltonians, strongly correlated electron systems, and interacting bosonic systems.

Results and discussion

Framework. The flowchart of our framework is shown in Fig. 1b. First, we prepare a Hamiltonian $\mathcal{H}(\theta)$ with a set of parameters θ . We also define the objective function $L(\theta)$ to be minimized for achieving the desired properties; for instance, if the objective is to maximize the expectation value of a physical quantity P , we can take $L(\theta) = -\langle P(\theta) \rangle$. Next, we compute the derivative $\frac{\partial L}{\partial \theta}$ by automatic differentiation. Then, we update the Hamiltonian by changing the parameters θ according to $\frac{\partial L}{\partial \theta}$. By repeating this procedure until θ converge, we end up with the Hamiltonian $\mathcal{H}(\theta_{\text{opt}})$ that optimizes the desired properties, where θ_{opt} are the parameters after the convergence, as commonly done in machine learning.

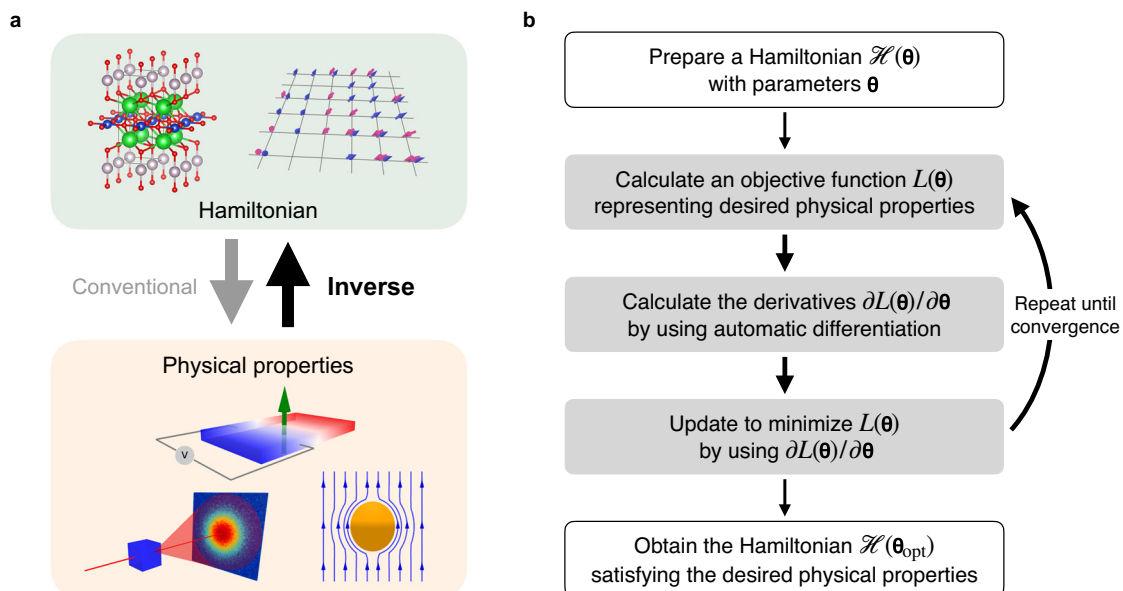


Fig. 1 Inverse design of Hamiltonian. **a** In the conventional approach, the Hamiltonian is first constructed based on phenomenology or first principles, and then, the optimal parameters of the Hamiltonian are explored through physical properties calculated from the Hamiltonian. In contrast, in the inverse approach, the desired physical properties are prepared first, and then, the Hamiltonian to realize them is obtained directly. **b** Flowchart proposed in the present study to solve the inverse problem by using automatic differentiation. First, we prepare a Hamiltonian $\mathcal{H}(\theta)$ that depends on parameters θ . Next, we calculate the objective function $L(\theta)$, which represents the desired physical properties. By optimizing θ to minimize $L(\theta)$, we obtain a Hamiltonian $\mathcal{H}(\theta_{\text{opt}})$ that satisfies the desired physical properties, where θ_{opt} are the parameters after the optimization.

Our framework has the following advantages in comparison with the existing methods^{1,2,5,7,14–23}: (i) It does not require training, hence, there is no need to collect data or consume computational resources on the training. (ii) It performs the optimization by using the analytical derivatives, which can achieve higher accuracy than the approximations based on neural networks even for large parameter space. (iii) It is applicable to a wide range of objective functions, unlike the perturbation theory. Therefore, our framework is able to deal with a large number of parameters in the Hamiltonian, which may lead to the findings of Hamiltonians that have not been reported thus far.

Automatic construction of the Haldane model showing spontaneous quantum AHE. First, we demonstrate that our framework can automatically find the Haldane model with a spontaneous quantum AHE³⁹. We consider a tight-binding model on a honeycomb lattice with two sublattices, whose Hamiltonian reads

$$\mathcal{H} = \sum_{i, a_i \in \{A, B\}} M^{a_i} c_i^\dagger c_i + \sum_{\langle i, j \rangle} t_1 c_i^\dagger c_j + \sum_{\langle\langle i, j \rangle\rangle} t_2^{d_{ij}} c_i^\dagger c_j, \quad (1)$$

where c_i^\dagger (c_i) is the creation (annihilation) operator of a spinless fermion at site i ; the first term describes an on-site staggered potential with real coefficients M^{a_i} ($a_i = A$ or B denotes the sublattice), and the second and third terms represent the hopping of fermions to nearest- and second-neighbor sites, respectively.

Here, we set $t_1 = 1$ as an energy unit and parametrize $t_2^{d_{ij}}$ as $t_2^{d_{ij}} = \sigma(r^{d_{ij}}) \exp(i\phi^{d_{ij}})$ with real variables $r^{d_{ij}}$ and $\phi^{d_{ij}}$, where $\sigma(x) = 1/(1 + e^{-x})$ is the sigmoid function to avoid the divergence of the absolute value of $t_2^{d_{ij}}$, and d_{ij} denotes the direction of the second-neighbor hopping, $d_{ij} \in \{A1, A2, A3, B1, B2, B3\}$ (see Fig. 2a). Thus, the model includes 14 parameters in total represented by $\theta = \{M^A, M^B, \{r^{d_{ij}}\}, \{\phi^{d_{ij}}\}\}$. The Haldane model is given by taking $M^A = +M$, $M^B = -M$, and $t_2^{d_{ij}} = t_2 \exp(i\phi)$ regardless of d . The phase diagram is shown in Fig. 2b, which has two topologically nontrivial phases with a spontaneous quantum AHE corresponding to the nonzero Chern numbers $C = \pm 1$.

With this setup of $\mathcal{H}(\theta)$, we try to obtain a Hamiltonian that maximizes the AHE by the framework in Fig. 1b. For this aim, we take the objective function as $L(\theta) = -\sigma_{xy}(\theta)$, where σ_{xy} is the Hall conductivity. Details of the calculations are described in the “Methods” section. We find that σ_{xy} increases monotonically through the optimization, as shown in Fig. 2d. Note that we introduce temperature and control it as shown in Fig. 2c to avoid that $\frac{\partial L}{\partial \theta}$ becomes zero due to the quantization (β is the inverse temperature). In contrast to the continuous change of σ_{xy} , the Chern numbers of the two bands, which are separated by the band gap shown in the inset of Fig. 2d, converge quickly to $C \simeq \pm 1$ in the very early stage of the optimization, as shown in Fig. 2e. The evolution of each parameter is plotted in Figs. 2f–h. We find that both M^A and M^B converge to zero, and $|t_2^{d_{ij}}| \rightarrow 1$ and $\phi^{d_{ij}} \rightarrow \pi/2$ for all d_{ij} . These values correspond to the center of the topological phase with $C=1$ in the Haldane model, indicated by the star in Fig. 2b. We confirm that different initial conditions converge to the same state (see Supplementary Note 1). Thus, our framework automatically constructs the Haldane model with a spontaneous quantum AHE under the condition of maximizing σ_{xy} . The reason why the optimal state is always at the center of the $C=1$ phase is due to the introduction of temperature; at nonzero

temperature, σ_{xy} becomes largest at the center where the band gap becomes largest in the topological phase. We note that the value of σ_{xy} in Fig. 2d is considerably smaller than the quantized value $+1$, which is also due to the finite temperature.

Finding a Hamiltonian with large quantum AHE on a triangular lattice. To demonstrate that our framework can find more complex models automatically, we apply it to a triangular lattice assuming a four-sublattice unit cell (Fig. 3a). The Hamiltonian reads

$$\mathcal{H} = \sum_{\langle i, j \rangle} t_1^{ij} c_i^\dagger c_j + \sum_{\langle\langle i, j \rangle\rangle} t_2^{ij} c_i^\dagger c_j + \sum_{\langle\langle\langle i, j \rangle\rangle\rangle} t_3^{ij} c_i^\dagger c_j. \quad (2)$$

We take $t_1^{ij} = \exp(i\phi_1^{ij})$ and $t_m^{ij} = \sigma(r_m) \exp(i\phi_m^{ij})$ for $m=2$ and 3 (see the arrows in Fig. 3a). Thus, the model includes 38 parameters in total represented by $\theta = \{r_2, r_3, \{\phi_1^{ij}\}, \{\phi_2^{ij}\}, \{\phi_3^{ij}\}\}$. As in the previous calculation, we take $L(\theta) = -\sigma_{xy}(\theta)$ to maximize the AHE. We optimize the parameters with a schedule of temperature shown in Fig. 3b. At each optimization step, the fermion density is fixed at half filling by tuning the chemical potential using the bisection method.

We find that the Chern numbers for four bands converge to $C = 5, 1, -3$, and -3 from the lower band, as shown in Fig. 3d. This indicates that σ_{xy} reaches 6 at half filling, which is six times larger than that in the Haldane model, although σ_{xy} in Fig. 3c is much smaller due to the finite temperature similar to the previous case. The band structure is shown in Fig. 3e with the Berry curvature Ω (see the “Methods” section). Note that the system recovers (approximately) threefold rotational symmetry after the convergence (see Supplementary Note 2). Ω of the lowest energy band is positive at all wave numbers, whose sum gives the largest $C = 5$, while the other bands include negative contributions. This indicates that our framework tries to maximize C for the lowest energy band. We note that the same conclusion is obtained for many other initial conditions, while some cases converge to $C = 3, 3, -1$, and -5 from the lower band, which gives the same value of $\sigma_{xy} = 6$. The reason why the solution in Fig. 3 is rather preferred is the finite temperature introduced in the optimization process, for the same reason as in the honeycomb lattice model for which the center of the topological phase was obtained (see Supplementary Note 2).

Let us discuss the optimized parameters. We find that both $|t_2|$ and $|t_3|$ converge to $\simeq 1$, while the phases take the various values shown by colors in Fig. 3a. We show, however, that their sums along closed loops in the counter-clockwise direction, $\Phi_m = \sum \phi_m^{ij}$, representing the fictitious magnetic fluxes, take some regular values: $\Phi_1 \simeq 7\pi/4$ for the smallest triangles composed of t_1 (Fig. 3f), and Φ_2 takes $\simeq 0.91\pi$ and $\simeq 1.59\pi$ for larger triangles of t_2 facing right and left, respectively (Fig. 3g), while Φ_3 is always $\simeq \pi$ (ϕ_3^{ij} is either $\simeq 0$ or π). Although ϕ_m^{ij} take different values for different initial conditions, Φ_m converges to the same values. These results indicate that our framework automatically finds a model whose complex hoppings realize spontaneous fictitious magnetic fluxes to maximize σ_{xy} , which is hard to obtain by intuition. Based on the results, we can also refine the Hamiltonian by taking more regular values of the phases (multiples of $\pi/4$) (see Supplementary Note 2).

Maximizing photovoltaic current generation in a spin-charge-coupled system. Finally, we apply our framework to optimize the PVE in a bulk system with broken spatial inversion symmetry^{40–44}. An example is the shift current, which is understood as a shift in the real space of electron wave functions excited by light. For simplicity, here we focus on (quasi-)

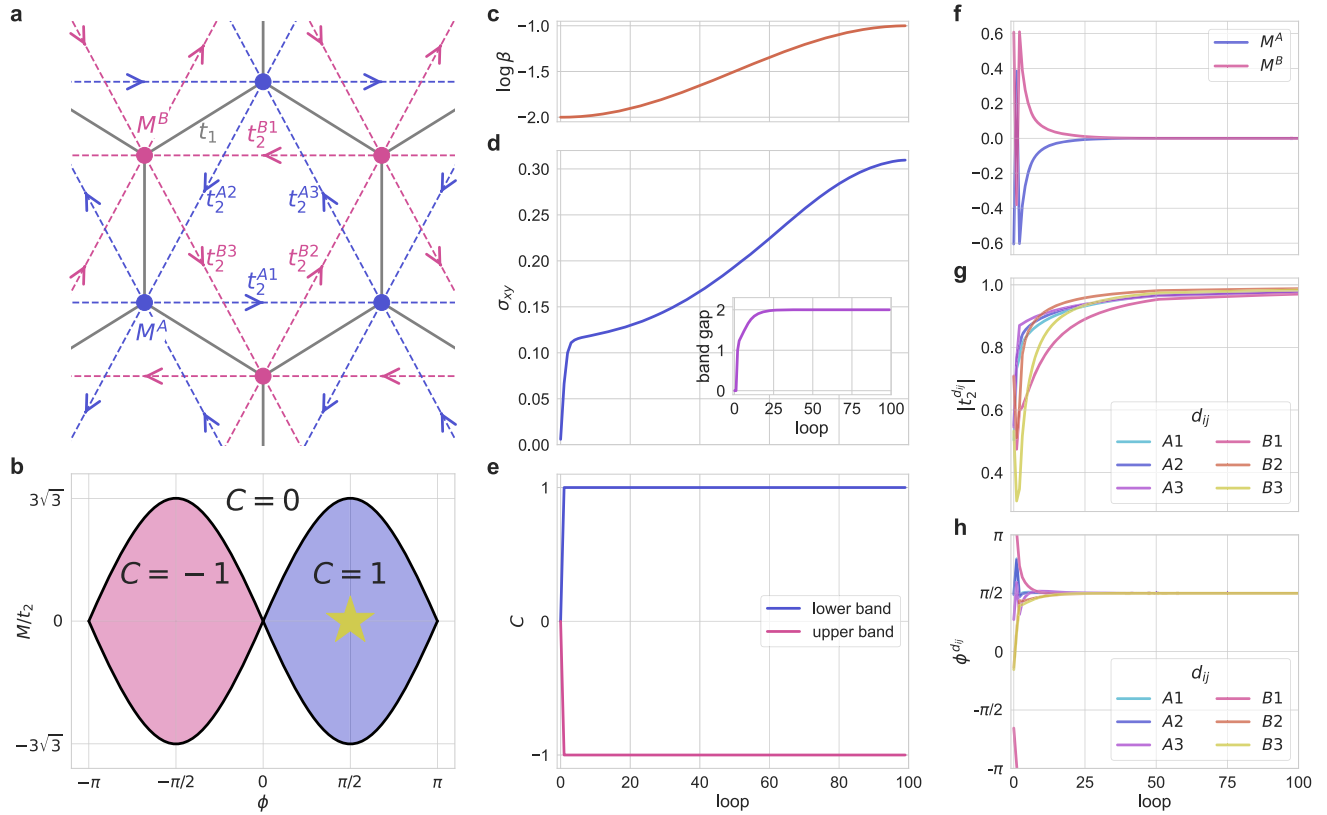


Fig. 2 Automatic construction of the Haldane model. **a** A tight-binding model on a honeycomb lattice in Eq. (1). There are 14 parameters including on-site potential M^a , and the amplitudes and phases of the second-neighbor hopping $t_2^{d_j}$, where $a_j \in \{A, B\}$ is the index of the sub-lattice and $d_j \in \{A1, A2, A3, B1, B2, B3\}$ is the direction of the hopping. The nearest-neighbor hopping t_1 is fixed to 1. **b** Phase diagram of the Haldane model, where $M^A = +M$, $M^B = -M$, $t_1 = 1$, and $t_2^{d_j} = t_2 \exp(i\phi)$. There are two topologically nontrivial phases with nonzero Chern numbers $C = \pm 1$. The yellow star represents where our framework reaches after the convergence. **c**, Schedule of $\log \beta$, where β is the inverse temperature. **d**, **e** Changes of the Hall conductivity σ_{xy} (**d**) and the Chern numbers C for the two bands (**e**) through the optimization process. The inset in **d** shows the change of the band gap. **f-h** Changes of the parameters: M^a (**f**), $|t_2^{d_j}|$ (**g**), and ϕ^{d_j} (**h**), where ϕ^{d_j} is the phase of $t_2^{d_j}$.

one-dimensional spin-charge-coupled systems where the spin configurations break spatial inversion symmetry⁴⁵. The schematic is shown in Fig. 4a. Note that the model approximately describes chiral magnetic metals, such as CrNb_3S_6 ⁴⁶ and $\text{Yb}(\text{Ni}_{1-x}\text{Cu}_x)_3\text{Al}_9$ ⁴⁷. The Hamiltonian reads

$$\mathcal{H} = \sum_{i,\alpha} \left(t_1 c_{i\alpha}^\dagger c_{i+1\alpha} + t_2 c_{i\alpha}^\dagger c_{i+2\alpha} + \text{H.c.} \right) + J \sum_{i,\alpha,\beta} c_{i\alpha}^\dagger \sigma_{\alpha\beta} c_{i\beta} \cdot \mathbf{S}_i, \tag{3}$$

where $c_{i\alpha}^\dagger$ ($c_{i\alpha}$) denotes the creation (annihilation) operator of an electron at site i with spin α . Here, we take $t_1 = \sqrt{2} \tanh(r_t) \cos(\theta_t) \times 0.1$ [eV], $t_2 = \sqrt{2} \tanh(r_t) \sin(\theta_t) \times 0.1$ [eV], and $J = \log(1 + \exp(r_j))$ [eV]; the spins are treated as classical and their configurations are parametrized as $\mathbf{S}_i = (\sin \theta_i \cos \phi_i, \sin \theta_i \sin \phi_i, \cos \theta_i)$, with $\theta_i = \pi \sigma(\eta_i)$. $|t_1|$ and $|t_2|$ are represented by the hyperbolic tangent functions to be bounded, otherwise, they will become too large through the optimization since the shift current increases with increasing momentum derivatives of the band dispersions. We set $|t_1|$ and $|t_2|$ to be within about 0.1 eV, considering the situation in the real materials. J is set to be positive without loss of generality. We set the number of sublattice sites to $N = 12$. Thus, the model includes $3 + 2N = 27$ parameters in total represented by $\theta = \{r_t, \theta_t, r_j, \{\eta_i\}, \{\phi_i\}\}$. The quantity of our interest is the photocurrent under solar radiation, defined as $I = \int d\omega \sigma_{\text{PVE}}(\omega) |E(\omega)|^2$ [A m⁻²], where $\sigma_{\text{PVE}}(\omega)$ is the nonlinear optical conductivity^{48,49}, and $|E(\omega)|^2$ denotes the intensity of

the linearly polarized solar light with frequency ω , approximately given by blackbody radiation at $T = 5500$ K (the inset of Fig. 4a) (see the “Methods” section); we take $L(\theta) = -I$. We consider a three-dimensional system in which the one-dimensional chains are arranged in a square lattice fashion for simplicity, taking the lattice constants $a_z = 9 \text{ \AA}$ in the chain direction and $a_x = a_y = 4 \text{ \AA}$ in the orthogonal directions, referring to a chiral magnet⁴⁷. The fermion density is fixed at half filling as for the previous model.

Figure 4c shows the optimization process of the photocurrent I under the schedule of temperature shown in Fig. 4b. We obtain $I \sim 700 \text{ A m}^{-2}$ after the convergence. This value is comparable to or larger than those for Ge semiconductors⁵⁰ and perovskites substances^{51,52}. Changes in the parameters t_1 , t_2 , and J are plotted in the inset of Fig. 4c. The optimized spin configuration is an umbrella-shaped chiral state with a three-site period, as shown in Fig. 4d-f. We also note that other noncoplanar spin configurations are also obtained for different initial conditions, but they generate smaller I (see Supplementary Note 3).

To elaborate the mechanism behind the optimization of the photocurrent, we plot the ω dependence of $I(\omega) = \sigma_{\text{PVE}}(\omega) |E(\omega)|^2$ in Fig. 4g, together with $\sigma_{\text{PVE}}(\omega) \omega^2$ and $|E(\omega)|^2$ in the inset. We find that $I(\omega)$ has a sharp peak at $\omega \sim 7.15 \times 10^{14}$ [rad s⁻¹], due to the peak of $\sigma_{\text{PVE}}(\omega) \omega^2$ located at the frequency where $|E(\omega)|^2$ becomes large. We show that dominant contributions to the peak come from the interband processes between the conduction and valence bands split by $2J \approx 0.5$ [eV] $\approx 7.15 \times 10^{14}$ [rad s⁻¹], as

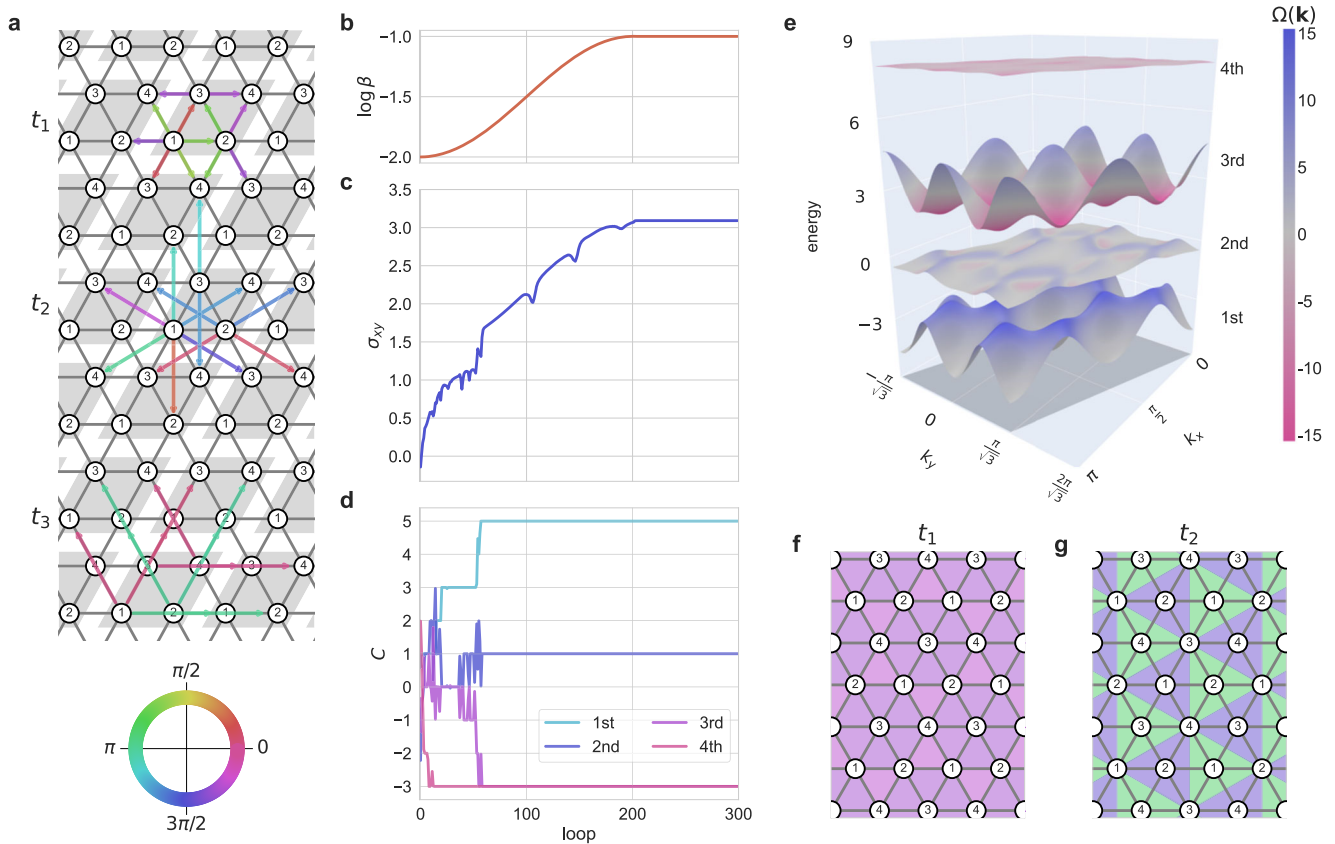


Fig. 3 Automatic construction of a Hamiltonian showing a six-times larger quantum anomalous Hall effect than the Haldane model. **a** A tight-binding model on a triangular lattice with 38 parameters, including the nearest-neighbor (t_1), the second-neighbor (t_2) and the third-neighbor (t_3) hoppings. The shades denote four-sublattice unit cells. The color of the arrows represents the optimum phase of each hopping, ϕ_m^j , after the convergence, according to the inset below. **b** Schedule of $\log \beta$, where β is the inverse temperature. **c, d** Changes of the Hall conductivity σ_{xy} (**c**) and the Chern number C for four bands (**d**). **e** The band structure after the convergence plotted with the Berry curvature $\Omega(\mathbf{k})$ at each wavenumber $\mathbf{k} = (k_x, k_y)$. **f, g** Fictitious magnetic fluxes defined by the sum of phases along the counter-clockwise direction as $\Phi_m = \sum_{ij} \phi_m^j$ on the smallest triangles by the nearest-neighbor hopping t_1 (**f**) and larger ones by the second-neighbor hopping t_2 (**g**), which are indicated by the same color code as the inset of (**a**).

shown in Fig. 4h (see the “Methods” section). The results indicate that the enhanced photocurrent of $\sim 700 \text{ A m}^{-2}$ under solar radiation is generated by band engineering with automatic optimization of t_1 , t_2 , J , and the spin configurations. We note that the peak value of $\sigma_{\text{PVE}}(\omega) \sim 0.06 \text{ A V}^{-2}$ is considerably large compared to existing materials, such as BaTiO_3 ^{40,53} and TaAs ⁵⁴, and is also even an order of magnitude larger than the value obtained in the previous theoretical study⁴⁵, while we may need substantially large competing magnetic interactions to stabilize the umbrella spin configuration at room temperature.

Conclusions

Through the applications to AHE and PVE, our framework has proven capable of automatically finding Hamiltonians that optimize the physical properties of interest. The key aspect is in the use of automatic differentiation in the inverse problem, which provides the derivatives of the objective function in terms of a large number of parameters; although the current studies are limited to several tens of parameters, we can practically deal with a million or more. Since automatic differentiation is a versatile technique, our framework has a wide range of applicability, such as first-principles Hamiltonians computed by the Kohn–Sham equations, strongly correlated electron systems, quantum spin systems, and interacting bosonic systems, as long as the forward computation can be performed efficiently. In addition, it is applicable to a wide

range of physical properties to be optimized, including the reproduction of experimental raw data. Thus, our findings will be useful for the exploration of new models and principles in materials science.

Methods

Application to the AHE. The Hall conductivity is calculated by using the Kubo formula as

$$\sigma_{xy} = -\frac{e^2}{h} \frac{V}{2\pi N_{\mathbf{k}}} \sum_{m,n,\mathbf{k}} (f(E_{\mathbf{k}n}, \beta) - f(E_{\mathbf{k}m}, \beta)) \Omega(\mathbf{k}), \quad (4)$$

where e is the elementary charge, h is the Planck constant, V is the volume of the Brillouin zone, $N_{\mathbf{k}}$ is the number of \mathbf{k} points, $f(E, \beta)$ is the Fermi distribution function at inverse temperature β , $E_{\mathbf{k}n}$ is the energy at \mathbf{k} in n th band; $\Omega(\mathbf{k})$ is the Berry curvature given by

$$\Omega(\mathbf{k}) = \text{Im} \frac{\langle \mathbf{k}n | \frac{\partial \mathcal{H}}{\partial k_y} | \mathbf{k}m \rangle \langle \mathbf{k}m | \frac{\partial \mathcal{H}}{\partial k_x} | \mathbf{k}n \rangle}{(E_{\mathbf{k}n} - E_{\mathbf{k}m})^2 + i\delta}, \quad (5)$$

where $|\mathbf{k}n\rangle$ is an eigenstate at \mathbf{k} in n th band. We take $e = h = 1$, $N_{\mathbf{k}} = 100^2$, and $\delta = 10^{-5}$.

The optimization starts from initial parameters randomly chosen as $M^A, M^B \in (-1, 1)$, $r^{d_j} \in (0, 1)$, and $\phi^{d_j} \in (-\pi, \pi)$ for the honeycomb lattice model, and $r_2, r_3 \in (0, 1)$ and $\phi_1^j, \phi_2^j, \phi_3^j \in (-\pi, \pi)$ for the triangular lattice model. Automatic differentiation is implemented using JAX⁵⁵. Note that $\frac{\partial \mathcal{H}}{\partial k_x}$ and $\frac{\partial \mathcal{H}}{\partial k_y}$ in Eq. (5) are also calculated by using automatic differentiation. We employ RMSPROP⁵⁶ as an optimization method, in which we take the learning rate, the decay factor, and the infinitesimal as 0.1, 0.99, and 10^{-8} , respectively.

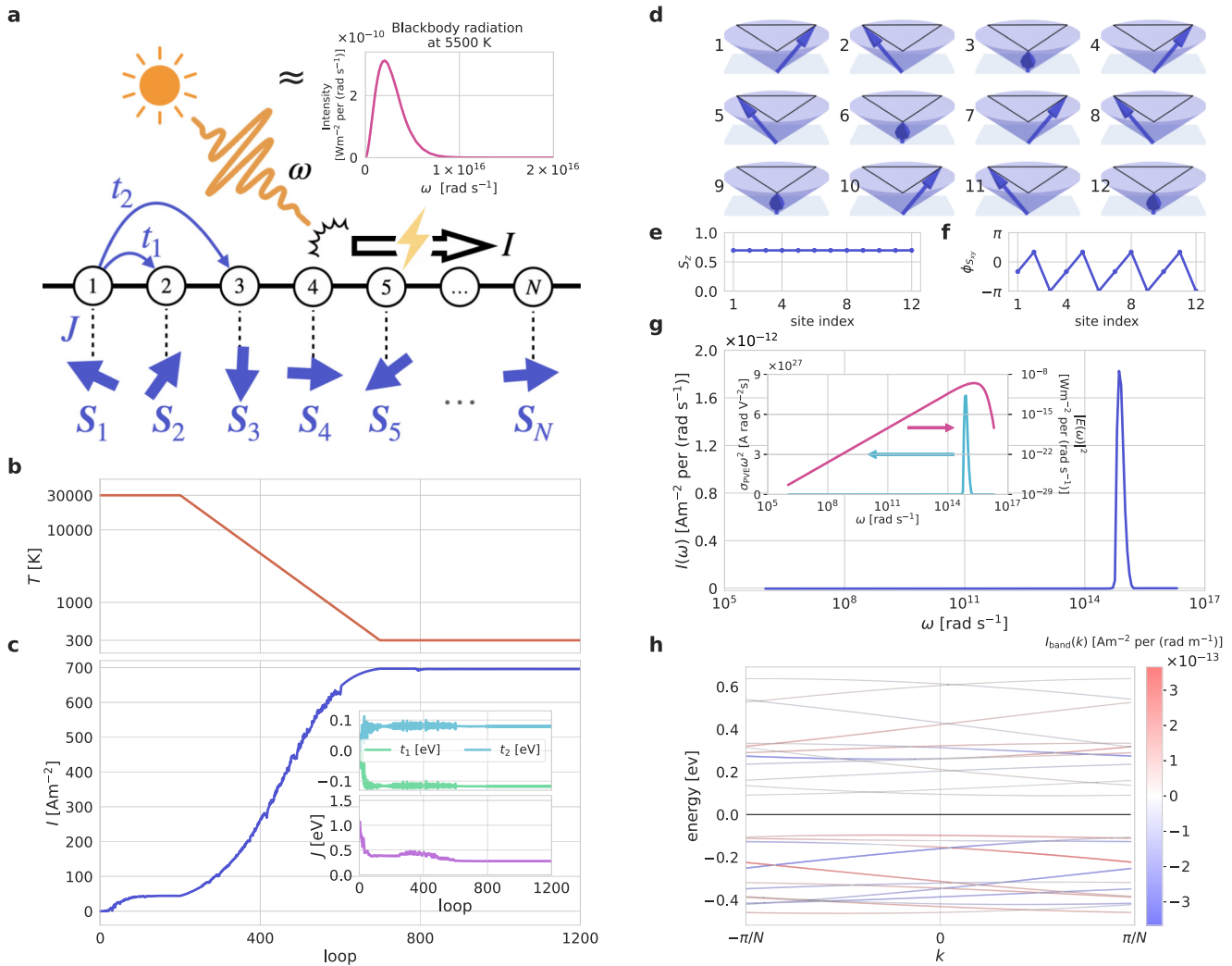


Fig. 4 Automatic construction of a Hamiltonian for electrons moving on a noncoplanar spin texture, which can generate $\sim 700 \text{ A m}^{-2}$ under solar radiation. **a** Schematic of the system. A photocurrent is generated by solar radiation (blackbody radiation at 5500 K in the inset) onto the one-dimensional spin-charge coupled system. **b** Schedule of temperature T [K]. **c** Change of the photocurrent I [A m^{-2}]. The insets show the changes of the nearest-neighbor hopping t_1 [eV], the second-neighbor hopping t_2 [eV], and the coupling constant J [eV]. **d–f** Spin configurations after the convergence (**d**), plotted with the z components S_z (**e**) and the angles of spins projected onto the xy plane, $\phi_{s_{xy}}$ (**f**). The S_z axis is taken in the direction of the total magnetization. **g** ω dependence of $I(\omega) = \sigma_{\text{PVE}} \epsilon(\omega) |E(\omega)|^2$, where $\sigma_{\text{PVE}}(\omega)$ is the nonlinear optical conductivity and $|E(\omega)|^2$ is the intensity of solar light (inset). **h** The band structure of electrons. $I_{\text{band}}(k)$ shown in color bar represents the contribution to I from each band.

Application to the PVE. According to the second-order optical response theory^{44,45}, a nonlinear electric current produced by electric fields $E(\omega_1)$ and $E(\omega_2)$ with two frequencies ω_1 and ω_2 , respectively, is given by

$$I(\omega_1 + \omega_2; \omega_1, \omega_2) = \sigma_{\text{opt}}(\omega_1 + \omega_2; \omega_1, \omega_2) E(\omega_1) E(\omega_2), \quad (6)$$

with the second-order optical conductivity $\sigma_{\text{opt}}(\omega_1 + \omega_2; \omega_1, \omega_2)$. In the case of $\omega_1 = -\omega_2$, a DC current is generated as

$$I(\omega) = \sigma_{\text{PVE}}(\omega) |E(\omega)|^2, \quad (7)$$

where $I(\omega) = I(0; \omega, -\omega)$ and $\sigma_{\text{PVE}}(\omega) = \sigma_{\text{opt}}(0; \omega, -\omega)$. The ω integral $I = \int d\omega I(\omega)$ gives a photocurrent generated by the shift current mechanism^{42,43,45}, which is used for the objective function in the main text. We approximate solar radiation by blackbody radiation $B(\omega, T)$ at 5500 K as

$$|E(\omega)|^2 = 2\mu_0 c C_{\text{solar}} \int d\omega B(\omega, T = 5500 \text{ K}), \quad (8)$$

where μ_0 , c , and C_{solar} are the magnetic constant, speed of light, and solar constant, respectively;

$$B(\omega, T) = \frac{\hbar \omega^3}{4\pi^2 c^2} \frac{1}{\exp(\frac{\hbar \omega}{k_B T}) - 1}, \quad (9)$$

where \hbar and k_B are the reduced Planck constant and the Boltzmann constant,

respectively. In Eq. (7), $\sigma_{\text{PVE}}(\omega)$ is computed as^{44,45}

$$\sigma_{\text{PVE}}(\omega) = -\frac{V e^3}{(2\pi)^2 N_i \omega^2} (\sigma_{\text{PVE},1} + \sigma_{\text{PVE},2} + \sigma_{\text{PVE},3} + \sigma_{\text{PVE},4}), \quad (10)$$

where

$$\sigma_{\text{PVE},1} = -\sum_{k,a} f(E_k, \beta) J_{aa}^{(3)}, \quad (11)$$

$$\sigma_{\text{PVE},2} = \sum_{k,a,b} \left(\frac{f_{ab} J_{ab}^{(1)} J_{ba}^{(2)}}{\omega + i\gamma/2 - E_{ab}} + \frac{f_{ab} J_{ab}^{(1)} J_{ba}^{(2)}}{-\omega + i\gamma/2 - E_{ab}} \right), \quad (12)$$

$$\sigma_{\text{PVE},3} = \sum_{k,a,b} \frac{f_{ab} J_{ab}^{(2)} J_{ba}^{(1)}}{i\gamma - E_{ab}}, \quad (13)$$

$$\sigma_{\text{PVE},4} = -\sum_{k,a,b,c} \frac{J_{ab}^{(1)} J_{bc}^{(1)} J_{ca}^{(1)}}{i\gamma - E_{ca}} \left(\frac{f_{ab}}{\omega + i\gamma/2 - E_{ba}} + \frac{f_{cb}}{\omega + i\gamma/2 - E_{cb}} + \frac{f_{ab}}{-\omega + i\gamma/2 - E_{ba}} + \frac{f_{cb}}{-\omega + i\gamma/2 - E_{cb}} \right). \quad (14)$$

Here, a , b , and c denote the bands; $E_{ab} = E_{ka} - E_{kb}$, $f_{ab} = f(E_{ka}, \beta) - f(E_{kb}, \beta)$, and $J_{ab}^{(n)} = \langle ka | \frac{\partial^n \mathcal{H}}{\partial k^n} | kb \rangle$. We use $V = \frac{(2\pi)^3}{a_x a_y a_z}$, $N_k = 100$, and $\gamma = 2\pi \times 10^{13} \text{ [rad s}^{-1}\text{]}$. $\frac{\partial^n \mathcal{H}}{\partial k^n}$ in $J_{ab}^{(n)}$ are calculated by using automatic differentiation. We also calculate the contribution to I from each k point in each band, $I_{\text{band}}(k)$, by calculating I without taking the summations of k and the band indices in Eqs. (11)–(14). The optimization

starts from initial parameters randomly chosen as $r_t \in (-1, 1)$, $\theta_t \in (-\pi, \pi)$, $r_j \in (0, 0.5)$, $\eta_j \in (-1, 1)$, and $\phi_j \in (-\pi, \pi)$.

Data availability

All the data can be generated from the code below.

Code availability

We have published the code to reproduce all the results on <https://github.com/koji-inui/automatic-hamiltonian-design.git>.

Received: 10 May 2022; Accepted: 11 January 2023;

Published online: 01 March 2023

References

- Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning: generative models for matter engineering. *Science* **361**, 360–365 (2018).
- Weymuth, T. & Reiher, M. Inverse quantum chemistry: concepts and strategies for rational compound design. *Int. J. Quantum Chem.* **114**, 823–837 (2014).
- Kuhn, C. & Beratan, D. N. Inverse strategies for molecular design. *J. Phys. Chem.* **100**, 10595–10599 (1996).
- Zunger, A. Inverse design in search of materials with target functionalities. *Nat. Rev. Chem.* **2**, 0121 (2018).
- Tamura, R. & Hukushima, K. Method for estimating spin–spin interactions from magnetization curves. *Phys. Rev. B* **95**, 064407 (2017).
- Yu, S., Gao, Y., Chen, B.-B. & Li, W. Learning the effective spin hamiltonian of a quantum magnet. *Chin. Phys. Lett.* **38**, 097502 (2021).
- Fujita, H., Nakagawa, Y. O., Sugiura, S. & Oshikawa, M. Construction of hamiltonians by supervised learning of energy and entanglement spectra. *Phys. Rev. B* **97**, 075114 (2018).
- Franceschetti, A. & Zunger, A. The inverse band-structure problem of finding an atomic configuration with given electronic properties. *Nature* **402**, 60–63 (1999).
- Hart, G. L. W., Blum, V., Walorski, M. J. & Zunger, A. Evolutionary approach for determining first-principles hamiltonians. *Nat. Mater.* **4**, 391–394 (2005).
- Mertz, T. & Valenti, R. Engineering topological phases guided by statistical and machine learning methods. *Phys. Rev. Res.* **3**, 013132 (2021).
- Ajoy, A. & Cappellaro, P. Quantum simulation via filtered hamiltonian engineering: application to perfect quantum transport in spin networks. *Phys. Rev. Lett.* **110**, 220503 (2013).
- Greiter, M., Schnells, V. & Thomale, R. Method to identify parent Hamiltonians for trial states. *Phys. Rev. B* **98**, 081113 (2018).
- Pakrouski, K. Automatic design of Hamiltonians. *Quantum* **4**, 315 (2020).
- Kosman, W. M. & Hinze, J. Inverse perturbation analysis: improving the accuracy of potential energy curves. *J. Mol. Spectrosc.* **56**, 93–103 (1975).
- Ho, T., Rabitz, H., Choi, S. E. & Lester, M. I. An inverse method for obtaining smooth multidimensional potential energy surfaces: application to $\text{ar}+\text{oh}$ $\text{a}2+(\nu=0)$. *J. Chem. Phys.* **102**, 2282–2285 (1995).
- Zhang, D. H. & Light, J. C. Potential inversion via variational generalized inverse. *J. Chem. Phys.* **103**, 9713–9720 (1995).
- Chertkov, E. & Clark, B. K. Computational inverse method for constructing spaces of quantum models from wave functions. *Phys. Rev. X* **8**, 031029 (2018).
- Yao, Z. et al. Inverse design of nanoporous crystalline reticular materials with deep generative models. *Nat. Mach. Intell.* **3**, 76–86 (2021).
- Liu, Z., Zhu, D., Raju, L. & Cai, W. Tackling photonic inverse design with machine learning. *Adv. Sci.* **8**, 2002923 (2021).
- von Toussaint, U. Bayesian inference in physics. *Rev. Mod. Phys.* **83**, 943–999 (2011).
- Ikebata, H., Hongo, K., Isomura, T., Maezono, R. & Yoshida, R. Bayesian molecular design with a chemical language model. *J. Comput.-Aided Mol. Des.* **31**, 379–391 (2017).
- Supady, A., Blum, V. & Baldauf, C. First-principles molecular structure search with a genetic algorithm. *J. Chem. Inf. Model.* **55**, 2338–2348 (2015).
- Yoshikawa, N. et al. Population-based de novo molecule generation, using grammatical evolution. *Chem. Lett.* **47**, 1431–1434 (2018).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
- Fedus, W., Zoph, B. & Shazeer, N. Switch transformers: scaling to trillion parameter models with simple and efficient sparsity. *J. Mach. Learn. Res.* **23**, 120 (2022).
- Xie, H., Liu, J.-G. & Wang, L. Automatic differentiation of dominant eigensolver and its applications in quantum physics. *Phys. Rev. B* **101**, 245139 (2020).
- Liao, H.-J., Liu, J.-G., Wang, L. & Xiang, T. Differentiable programming tensor networks. *Phys. Rev. X* **9**, 031041 (2019).
- Mann, S. et al. ∂pv : an end-to-end differentiable solar-cell simulator. *Comput. Phys. Commun.* **272**, 108232 (2022).
- Leung, N., Abdelhafez, M., Koch, J. & Schuster, D. Speedup for quantum optimal control from automatic differentiation based on graphics processing units. *Phys. Rev. A* **95**, 042318 (2017).
- Abdelhafez, M., Schuster, D. I. & Koch, J. Gradient-based optimal control of open quantum systems using quantum trajectories and automatic differentiation. *Phys. Rev. A* **99**, 052327 (2019).
- Torlai, G., Carrasquilla, J., Fishman, M. T., Melko, R. G. & Fisher, M. P. A. Wavefunction positization via automatic differentiation. *Phys. Rev. Res.* **2**, 032060 (2020).
- Vargas-Hernández, R. A., Chen, R. T. Q., Jung, K. A. & Brumer, P. Fully differentiable optimization protocols for non-equilibrium steady states. *New J. Phys.* **23**, 123006 (2021).
- Rigo, J. B. & Mitchell, A. K. Automatic differentiable numerical renormalization group. *Phys. Rev. Res.* **4**, 013227 (2022).
- Tamayo-Mendoza, T., Kreisbeck, C., Lindh, R. & Aspuru-Guzik, A. Automatic differentiation in quantum chemistry with applications to fully variational Hartree–Fock. *ACS Cent. Sci.* **4**, 559–566 (2018).
- Yoshikawa, N. & Sumita, M. Automatic differentiation for the direct minimization approach to the Hartree–Fock method. *The J. Phys. Chem. A* **126**, 8487–8493 (2022).
- Schoenholz, S. & Cubuk, E. D. Jax md: a framework for differentiable physics. In *Advances in Neural Information Processing Systems*, Vol. 33 (eds Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. & Lin, H.) 11428–11441 (Curran Associates, Inc., 2020).
- Li, L. et al. Kohn-sham equations as regularizer: building prior knowledge into machine-learned physics. *Phys. Rev. Lett.* **126**, 036401 (2021).
- Kasim, M. F., Lehtola, S. & Vinko, S. M. Dqc: a python program package for differentiable quantum chemistry. *J. Chem. Phys.* **156**, 084801 (2022).
- Haldane, F. D. M. Model for a quantum Hall effect without landau levels: condensed-matter realization of the “parity anomaly”. *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
- Miller, R. C. Optical harmonic generation in single crystal BaTiO_3 . *Phys. Rev.* **134**, A1313–A1319 (1964).
- Glass, A. M., von der Linde, D. & Negran, T. J. High voltage bulk photovoltaic effect and the photorefractive process in LiNbO_3 . *Appl. Phys. Lett.* **25**, 233–235 (1974).
- von Baltz, R. & Kraut, W. Theory of the bulk photovoltaic effect in pure crystals. *Phys. Rev. B* **23**, 5590–5596 (1981).
- Young, S. M., Zheng, F. & Rappe, A. M. First-principles calculation of the bulk photovoltaic effect in bismuth ferrite. *Phys. Rev. Lett.* **109**, 236601 (2012).
- Parker, D. E., Morimoto, T., Orenstein, J. & Moore, J. E. Diagrammatic approach to nonlinear optical response with application to Weyl semimetals. *Phys. Rev. B* **99**, 045121 (2019).
- Okumura, S., Morimoto, T., Kato, Y. & Motome, Y. Quadratic optical responses in a chiral magnet. *Phys. Rev. B* **104**, L180407 (2021).
- Togawa, Y. et al. Magnetic soliton confinement and discretization effects arising from macroscopic coherence in a chiral spin soliton lattice. *Phys. Rev. B* **92**, 220412 (2015).
- Matsumura, T. et al. Chiral soliton lattice formation in monoaxial helimagnet $\text{Yb}(\text{Ni}_{1-x}\text{Cu}_x)\text{Al}_9$. *J. Phys. Soc. Jpn.* **86**, 124702 (2017).
- Boyd, R. W. *Nonlinear Optics* (Academic Press, 2020).
- Hanamura, E., Kawabe, Y. & Yamanaka, A. *Quantum Nonlinear Optics* (Springer Science & Business Media, 2007).
- Singh, P. & Ravindra, N. Temperature dependence of solar cell performance-an analysis. *Sol. Energy Mater. Sol. Cells* **101**, 36–45 (2012).
- Snaith, H. J. et al. Anomalous hysteresis in perovskite solar cells. *J. Phys. Chem. Lett.* **5**, 1511–1515 (2014).
- Commandeur, D., Morrissey, H. & Chen, Q. Solar cells with high short circuit currents based on cspbbr3 perovskite-modified ZnO nanorod composites. *ACS Appl. Nano Mater.* **3**, 5676–5686 (2020).
- Young, S. M. & Rappe, A. M. First principles calculation of the shift current photovoltaic effect in ferroelectrics. *Phys. Rev. Lett.* **109**, 116601 (2012).
- Osterhoudt, G. B. et al. Colossal mid-infrared bulk photovoltaic effect in a type-i Weyl semimetal. *Nat. Mater.* **18**, 471–475 (2019).
- Bradbury, J. et al. JAX: Composable Transformations of Python+NumPy Programs. <http://github.com/google/jax> (2018).
- Hinton, G., Srivastava, N. & Swersky, K. *Neural Networks for Machine Learning Lecture 6a Overview of Mini-batch Gradient Descent* <http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture-slides-lec6.pdf> (2012).

Acknowledgements

The authors thank Y. Kato, S. Okumura, R. Pohle, and K. Shimizu for fruitful discussions. This work is supported by KAKENIH Grant No. 20H00122, and a

Grant-in-Aid for Scientific Research on Innovative Areas “Quantum Liquid Crystals” (KAKENHI Grant No. JP19H05825) from JSPS of Japan. It is also supported by JST CREST Grant No. JPMJCR18T2.

Author contributions

K.I. conceived and implemented the algorithm through the discussion with Y.M. K.I. and Y.M. conceived the models, interpreted the results, and wrote the manuscript.

Competing interests

K.I. has filed a patent based on the algorithm reported in this paper. Y.M. has no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42005-023-01132-0>.

Correspondence and requests for materials should be addressed to Koji Inui.

Peer review information *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023