Supplementary Material:

Nonlinear Optical Control of Chiral Charge Pumping in a Topological Weyl Semimetal

M. Mehdi Jadidi,^{1, a} Mehdi Kargarian,^{2, b} Martin Mittendorff,^{3, 4} Yigit

Aytac,^{3,5} Bing Shen,⁶ Jacob C. Konig-Otto,⁷ Stephan Winnerl,⁷ Ni Ni,⁶

Alexander L. Gaeta,¹ Thomas E. Murphy,³ and H. Dennis Drew^{8, c}

¹Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027 USA

 ²Department of Physics, Sharif University of Technology, Tehran 14588-89694, Iran
 ³Institute for Research in Electronics & Applied Physics, University of Maryland, College Park, MD 20742, USA

⁴Universität Duisburg-Essen, Fakultät für Physik, 47057 Duisburg, Germany

⁵Science Systems and Applications, Inc. Lanham, MD 20706, USA ⁶Department of Physics and Astronomy,

University of California, Los Angeles, CA 90095, USA ⁷Helmholtz-Zentrum Dresden-Rossendorf, PO Box 510119, D-01314 Dresden, Germany

⁸Center for Nanophysics and Advanced Materials,

University of Maryland, College Park, Maryland 20742, USA

 $^{^{\}rm a}$ mehdi.jadi
di@columbia.edu

^b kargarian@physics.sharif.edu

 $^{^{\}rm c}$ hdrew@umd.edu

CONTENTS

I. Introduction	2
II. Weyl Semimetals: general remarks	3
III. Optical responses in the extreme quantum limit	4
A. Linear magneto-conductivity	5
B. Nonlinear optical response	6
IV. A simple lattice model for Weyl semimetals	7
V. Discussion on the observed pump-probe signal	9
A. more details on experimenatl measurments	9
B. Sign of reflection pump-probe signal	9
C. Contribution of Weyl bands to the pump-probe response	10
1. Hot carriers response	11
2. Nonlinear chiral charge pumping response	11
D. Estimation of the pump-induced chiral charge pumping signal	11
E. Chiral charge pumping process in the pump-probe measurements	12
VI. Supplemental pump-probe data for $E_{\text{pump}} \perp B$	12
VII. Supplemental pump-probe data for $E_{\rm pump} \parallel B$	13
VIII. XRD spectrum and DC magneto-transport of TaAs	13
References	15

I. INTRODUCTION

In the main text, we introduced the concept of the dynamical chiral charge pumping and its optical control in a Weyl semimetal. We presented the pump-probe measurement results at various magnetic fields applied parallel to pump/probe polarization, and for different pump fluences. The lifetime associated with this nonlinearity is the chiral charge relaxation rate, and is measured to be $\gg 1$ ns. We further confirm our results by showing that for the case in which the pump polarization is perpendicular to the applied magnetic field, no long-lived response associated with nonlinear dynamical chiral charge pumping is observed.

This supplemental materails contains the details of the theory we developed to understand the main findings of our optical measurements and extra experimetal data in various cases. We begin by a brief introduction of Weyl nodes appearing in the band structure of solids in Sec. II. Since we are interested in the optical transport of the system in the extreme quantum limit, we discuss the transport of the extreme quantum limit in Sec. III, where only zeroth Landau level (LL0) is occupied. Then, a lattice model for Weyl semimetals and the formation of LLs are presented in Sec. IV. Details of the pump-probe signal from TaAs at various cases alongside with the calculations of the reflection are given in Sec. V. More data for perpendicular and parallel field geometries are presented in Secs. VI-VII. The XRD and DC measurements on smaples are presented in Sec. VIII.

II. WEYL SEMIMETALS: GENERAL REMARKS

We consider a doped Weyl semimetal whose low-energy excitations near the nodes are described by the following Hamiltonian:

$$H_{\eta} = \eta \hbar v \boldsymbol{\sigma} \cdot \mathbf{k} - \varepsilon_F,\tag{1}$$

where the wave vector $\mathbf{k} = (k_x, k_y, k_z)$ is measured from the node and $\boldsymbol{\sigma}$ is a vector of Pauli matrices. Here, v is velocity and ε_F is the Fermi energy. For simplicity we assume that the system is composed of two Weyl nodes with opposite chiralities $\eta = \pm$. In the momentum space the Weyl nodes act as source and sinks of Berry curvature $\Omega_{\mathbf{k}}^{\eta} = \boldsymbol{\nabla} \times \mathbf{A}_{\mathbf{k}}^{\eta}$, where $\mathbf{A}_{\mathbf{k}}^{\eta} = i \langle u_{\mathbf{k}}^{\eta} | \boldsymbol{\nabla} u_{\mathbf{k}}^{\eta} \rangle$ is the Berry connection of a Bloch wave function $u_{\mathbf{k}}^{\eta}$ of a given node satisfying $H_{\eta}u_{\mathbf{k}}^{\eta} = \varepsilon_{\mathbf{k}}u_{\mathbf{k}}^{\eta}$. It yields $\Omega_{\mathbf{k}}^{\eta} = \eta \hat{\mathbf{k}}/2k^2$. The chirality of a given node is given by the total flux of the Berry curvature as $\eta = 1/2\pi \oint \Omega_{\mathbf{k}}^{\eta} \cdot d\mathbf{S}_k$, where the integral is taken over a closed surface enclosing the nodes in the momentum space.

III. OPTICAL RESPONSES IN THE EXTREME QUANTUM LIMIT

The Weyl cones are replaced by dispersive Landau levels in a strong magnetic field. The carriers are transported through the LLs, indexed by n, in an electric field directed along the z axis. The transport is described by the linearized Boltzmann kinetic equation for the node with chirality η as

$$\partial_t f_n^\eta(p_z) + eE\partial_{p_z} f_n^\eta(p_z) = I\{f_n^\eta(p_z)\},\tag{2}$$

where $f_n^{\eta}(p_z)$ is the distribution function for electron states with momentum p_z in the *n*-th LL and we take e = -|e|. Here *I* on the right-hand side denotes the collision integrals. We assume that the momentum relaxation rate within each valley τ_{intra}^{-1} is much larger than inter-valley relaxation rate τ^{-1} , i.e., $\tau \gg \tau_{intra}$. Thus, the distribution function $f_n^{\eta}(p_z)$ becomes isotropic in momentum and depends on energy as $f_n^{\eta}(p_z) = f^{\eta}(\varepsilon_n(p_z))$, where $\varepsilon_n(p_z) = \pm v \sqrt{2n\hbar|e|B + p_z^2}$ for $n = 1, 2, \cdots$ and $\varepsilon_0(p_z) = -\eta v p_z$ for n = 0. Therefore (2) can be rewritten as

$$\partial_t f^{\eta}(\varepsilon_n(p_z)) + eE\partial_{p_z} f^{\eta}(\varepsilon_n(p_z)) = I\{f^{\eta}(\varepsilon_n(p_z))\}.$$
(3)

Multiplying by $\sum_{n} \delta(\varepsilon - \varepsilon_n(p_z))$ and integrating over momentum, the above Boltzmann equation can be cast in the form

$$\partial_t f^{\eta}(\varepsilon) + \frac{\eta}{\nu^{\eta}(\varepsilon)} \frac{e^2 \mathbf{E} \cdot \mathbf{B}}{h^2} \partial_{\varepsilon} f^{\eta}(\varepsilon) = I\{f^{\eta}(\varepsilon)\},\tag{4}$$

where

$$\nu^{\eta}(\varepsilon) = \frac{1}{2\pi l_B^2} \sum_n \int \frac{dp_z}{h} \delta(\varepsilon - \varepsilon_n(p_z)) = \frac{1}{2\pi l_B^2} \frac{1}{hv} \left(1 + 2\sum_{n=1}^{\infty} \frac{\epsilon}{\sqrt{\epsilon^2 - 2n}} \right)$$
(5)

is the density of states with $l_B = \sqrt{\hbar/|e|B}$ as the magnetic length and $\epsilon = \varepsilon/\hbar v l_B^{-1}$. We use the relaxation time approximation and write the collision integral as

$$\partial_t f + \frac{\eta}{\nu(\varepsilon)} \frac{e^2 \mathbf{E} \cdot \mathbf{B}}{h^2} \partial_{\varepsilon} f = -\frac{f - f_0}{\tau(\varepsilon)}.$$
 (6)

To study the response of the system, we assume a monochromatic incident electric field with frequency ω , i.e. $\mathbf{E}(t) = \mathbf{E}e^{-i\omega t} + \mathbf{E}^*e^{i\omega t}$, disturbs the distribution function as

$$f = \sum_{n=0}^{\infty} f_n e^{-in\omega t},\tag{7}$$

where the terms f_n 's $(n \ge 1)$ are induced by the electric field. We truncate the series up to second order as [1]

$$f = f_0 + f_1 e^{-i\omega t} + f_2 e^{-2i\omega t}.$$
 (8)

The (6) then becomes

$$\begin{bmatrix} -i\omega f_1 e^{-i\omega t} - 2i\omega f_2 e^{-2i\omega t} \end{bmatrix} + \frac{\eta}{\nu(\varepsilon)} \frac{e^2 \mathbf{E}(t) \cdot \mathbf{B}}{h^2} \begin{bmatrix} \partial_{\varepsilon} f_0 + \partial_{\varepsilon} f_1 e^{-i\omega t} + \partial_{\varepsilon} f_2 e^{-2i\omega t} \end{bmatrix}$$
$$= -\frac{1}{\tau(\varepsilon)} \begin{bmatrix} f_1 e^{-i\omega t} + f_2 e^{-2i\omega t} \end{bmatrix}.$$
(9)

To measure the response to the probe field we consider a configuration for the electric fields as $\mathbf{E} = \mathbf{E}_{pu} + \mathbf{E}_{pr}$, where the lowercase indices denote the pump and probe components, respectively. We calculate the response linear in \mathbf{E}_{pr} , while the amplitude itself might have been modulated by the pump field \mathbf{E}_{pu} yielding a non-linear signal as discussed below. Plugging the electric field $\mathbf{E}(t)$ in the above equation and equating the terms proportional to $e^{-i\omega t}$ and $e^{-2i\omega t}$, we obtain the following expressions for $f_1 = f_1^{(1)} + f_1^{(3)}$, where linear and non-linear terms read as

$$f_1^{(1)} = -\frac{\tau(\varepsilon)}{1 - i\omega\tau(\varepsilon)} \frac{\eta}{\nu(\varepsilon)} \frac{e^2 \mathbf{E}_{\rm pr} \cdot \mathbf{B}}{h^2} \partial_{\varepsilon} f_0 \tag{10}$$

and

$$f_{1}^{(3)} = -\frac{3e^{2}\mathbf{E}_{\mathrm{pr}} \cdot \mathbf{B}}{h^{2}} \left(\frac{e^{2}\mathbf{E}_{\mathrm{pu}} \cdot \mathbf{B}}{h^{2}}\right)^{2} \frac{\tau(\varepsilon)}{1 - i\omega\tau(\varepsilon)} \frac{\eta}{\nu(\varepsilon)} \partial_{\varepsilon} \left(\frac{\tau(\varepsilon)}{1 - 2i\omega\tau(\varepsilon)} \frac{1}{\nu(\varepsilon)} \partial_{\varepsilon} \left(\frac{\tau(\varepsilon)}{1 - i\omega\tau(\varepsilon)} \frac{1}{\nu(\varepsilon)} \partial_{\varepsilon} f_{0}\right)\right).$$
(11)

A. Linear magneto-conductivity

At zero temperature the linear term $f_1^{(1)}$ gives rise to the chiral charge density at the node η

$$n_{\eta} = \int d\varepsilon \ \nu(\varepsilon) f_1^{(1)}(\varepsilon) = \frac{\tau_{\rm ch}}{1 - i\omega\tau_{\rm ch}} \frac{\eta e^2 \tilde{\mathbf{E}}_{\rm pr} \cdot \mathbf{B}}{h^2},\tag{12}$$

where $\tau_{\rm ch} = \tau(\varepsilon_F)$. It then follows that

$$\tilde{n}_{\rm ch}(\omega) = n_+ - n_- = \frac{2e^2}{h^2} \frac{\tau_{\rm ch}}{\sqrt{1 + \omega^2 \tau_{\rm ch}^2}} \tilde{\mathbf{E}}_{\rm pr}(\omega) \cdot \mathbf{B}.$$
(13)

For $\mathbf{E}_{pr} \parallel \mathbf{B}$ the current density reads as

$$J^{(1)} = -e\eta \int d\varepsilon \ v(\varepsilon)\nu(\varepsilon)f_1^{(1)}(\varepsilon) = \frac{\tau_{\rm ch}}{1 - i\omega\tau_{\rm ch}}\frac{e^2v}{4\pi^2\hbar l_B^2}E,\tag{14}$$

yielding the following expression for the optical magneto-conductivity

$$\sigma_{\rm ch}(\omega) = \frac{\tau_{\rm ch}}{1 - i\omega\tau_{\rm ch}} \frac{e^2 v}{4\pi^2 \hbar l_B^2} \tag{15}$$

which, at the DC limit, reduces to the expression obtained in Ref. [2].

B. Nonlinear optical response

The nonlinear response results from the $f_1^{(3)}$ in the distribution function. The linear dispersion of the chiral mode $\varepsilon_0(p_z) = -\eta v p_z$ leads to constant density of states $\nu_0 = 1/2\pi l_B^2 h v$, and the non-linear response vanishes identically. Therefore we have to depart from the linearly dispersed chiral mode and, specifically, we supplement it with quadratic dispersion. In Sec. IV we introduce a lattice model for Weyl semimetals where the lattice effects yield a quadratic correction to the linear dispersion with the following expression for the density of states:

$$\nu(\varepsilon) \approx \nu_0 (1 - \alpha \varepsilon/2 + 3\alpha^2 \varepsilon^2/4), \tag{16}$$

where $\alpha^{-1} = mv^2/2$. For our cases $\alpha \varepsilon_F \ll 1$ holds. We further assume that the scattering from the impurities relaxes the momentum. That is, the scattering rate is proportional to the density of states $\tau^{-1}(\varepsilon) = A\nu(\varepsilon)$, where the constant A depends on the scattering strength. In the regime of interest $\omega \tau \gg 1$, the chiral number density $\delta n_{\eta} = \int d\varepsilon \ \nu(\varepsilon) f_1^{(3)}(\varepsilon)$ reads as

$$\delta n_{\eta} = \eta \frac{15\alpha^2}{8\nu_0^2} \frac{1}{\omega^4 \tau_{\rm ch}} \frac{e^2 \mathbf{E}_{\rm pr} \cdot \mathbf{B}}{h^2} \left(\frac{e^2 \mathbf{E}_{\rm pu} \cdot \mathbf{B}}{h^2}\right)^2,\tag{17}$$

and for chiral density pumping we have

$$\delta \tilde{n}(\omega) = \delta n_{+} - \delta n_{-} = \frac{15\alpha^{2}}{4\nu_{0}^{2}} \frac{1}{\omega^{4}\tau_{ch}} \frac{e^{2}\tilde{\mathbf{E}}_{pr} \cdot \mathbf{B}}{h^{2}} \left(\frac{e^{2}\tilde{\mathbf{E}}_{pu} \cdot \mathbf{B}}{h^{2}}\right)^{2}.$$
 (18)

We also obtain the following expression for the non-linear conductivity

$$\delta\sigma^{\rm NL}(\omega) = \frac{|e|^7 v}{h^6} \frac{3\alpha^2 \tau_{\rm ch}^3}{2\nu_0^2} \frac{-1 + 7i\omega\tau_{\rm ch} + 10\omega^2 \tau_{\rm ch}^2 - 3i\omega^3 \tau_{\rm ch}^3}{(1 - i\omega\tau_{\rm ch})^4 (1 - 2i\omega\tau_{\rm ch})^2} B\left(\tilde{\mathbf{E}}_{\rm pu} \cdot \mathbf{B}\right)^2,$$
(19)

which clearly shows that the signal vanishes for perpendicular field alignments $\mathbf{E}_{pu} \perp \mathbf{B}$. Therefore the enhancement observed in the reflection results from the above nonlinear conductivity. In particular we see that the imaginary part is positive as

$$\operatorname{Im}[\delta\sigma^{\mathrm{NL}}(\omega)] = \frac{|e|^7 v}{h^6} \frac{9\alpha^2}{8\nu_0^2} \frac{\tau_{\mathrm{ch}}^3}{(\omega\tau_{\mathrm{ch}})^3} B\left(\tilde{\mathbf{E}}_{\mathrm{pu}} \cdot \mathbf{B}\right)^2,\tag{20}$$

which is required for the enhancement of reflection. This equation is the same as equation (4) in the main text.

IV. A SIMPLE LATTICE MODEL FOR WEYL SEMIMETALS

We present a simple lattice model for a Weyl semimetal to simulate the essential features such as the isolated touching points between non-degenerate valence and conduction bands in the Brillouin zone. The lattice model reads as [3, 4]

$$H(\mathbf{k}) = t[\cos(k_x a) + \cos(k_y a) - 2]\sigma^z - t_z[\cos(k_z c) - \cos(Qc)]\sigma^z + t_{xy}\left[\sin(k_x a)\sigma^x + \sin(k_y a)\sigma^y\right],$$
(21)

where a and c denote the lattice constants in the ab plane and along the c axis. The Weyl nodes are located at $k_z = \pm Q$:

$$H_W = \hbar v (k_x \sigma^x + k_y \sigma^y) - t_z [\cos(k_z c) - \cos(Qc)] \sigma^z, \qquad (22)$$

where $v = t_{xy}a/\hbar$. In the presence of the quantizing field $\mathbf{B} = B\hat{z}$ the Landau levels are obtained via the substitution $\mathbf{\Pi} = \mathbf{p} + |e|\mathbf{A}$ for canonical momentum $\mathbf{p} = \hbar \mathbf{k}$, where in the Landau gauge $\mathbf{A} = (0, Bx, 0)$. We use the following commutation relation for $\Pi_{\pm} = \Pi_x \pm i \Pi_y$

$$[\Pi_{-},\Pi_{+}] = 2e\hbar B \tag{23}$$

to define lowering and raising boson operators

$$b = \frac{l_B}{\sqrt{2\hbar}} \Pi_{-}, \qquad b^{\dagger} = \frac{l_B}{\sqrt{2\hbar}} \Pi_{+}. \tag{24}$$

The Weyl Hamiltonian (22) is written as

$$H_W = \begin{pmatrix} \varepsilon_0(k_z) & \frac{\sqrt{2\hbar v}}{l_B} b\\ \frac{\sqrt{2\hbar v}}{l_B} b^{\dagger} & -\varepsilon_0(k_z) \end{pmatrix}, \quad \varepsilon_0(k_z) = -t_z [\cos(k_z c) - \cos(Qc)], \quad (25)$$

and the energy spectrum reads as

$$\varepsilon_n(k_z) = \begin{cases} \pm \sqrt{\varepsilon_0^2(k_z) + \frac{2\hbar^2 v^2}{l_B^2}n} & n = 1, 2, \cdots \\ \varepsilon_0(k_z) & n = 0. \end{cases}$$

Near the node we have

$$\varepsilon_0(k_z) \approx \frac{\hbar^2 k_z^2}{2m} + \hbar v_z k_z, \qquad (26)$$

where $m^{-1} = 2t_z c^2 \cos(Qc)/\hbar^2$, $v_z = t_z c \sin(Qc)/\hbar$, and k_z is measured from the node Q.

V. DISCUSSION ON THE OBSERVED PUMP-PROBE SIGNAL

A. more details on experimenatl measurments

In the presented pump-probe measurements on TaAs, the bulk response is mainly being excited and monitored, as the admittance associated with the surface response is very small compared with the high index of refraction of TaAs at 14 meV. According to (18) and (19) (equations (3) and (4) in the main text), the signature of pump-induced nonlinearity in the chiral charge pumping is a long-lived positive change in the probe reflection as discussed in the following. We note that other nonlinearities are present, such as hot carriers effects, that can contribute to a pump-induced change in the probe reflection. To distinguish those, we carry out measurements at different magnetic fields, pump fluences, and pump polarizations. Furthermore, since the nonlinearities have different time scales, the time delay scan helps differentiate between various pump-induced nonlinearities.

B. Sign of reflection pump-probe signal

The permittivity ϵ can be written as follows in terms of the conductivity $\sigma(\omega)$,

$$\epsilon = \epsilon_{\infty} + i \frac{\sigma}{\omega \epsilon_0},\tag{27}$$

and reflection is calculated as

$$R(T) = \left| \frac{1 - \sqrt{\epsilon}}{1 + \sqrt{\epsilon}} \right|^2 \tag{28}$$

For $|\Delta R| \ll R$ and $|\Delta \epsilon| \ll |\epsilon_b|$ (ϵ_b : background permittivity), one can obtain the following relation for the change in reflection,

$$\Delta R \approx 2|\epsilon_b|^{-3/2} |\Delta\epsilon| \cos(3\theta/2 - \gamma) \tag{29}$$

where $\gamma = \angle \Delta \epsilon$ and $\theta = \angle \epsilon_b$.

From Fig. 1(a) in Ref.[5], we can estimate the permittivity of TaAs at $\omega = 14$ meV to be $\epsilon_b = -530 + i450 = 695 \angle 140^\circ$, and thus $\epsilon_b = 695$ and $\theta = 140^\circ$. From (29), in order to get a positive pump-probe signal ($\Delta R > 0$), we need have a pump-induced change in reflection with $30^\circ < \gamma < 210^\circ$. This means that nonlinear chiral charge pumping that gives rise to a large imaginary change in conductivity ($\gamma = 90^{\circ}$) should lead to a positive change in reflection. Similarly, for hot carriers effects with a positive change in Drude weight ($\gamma = 45^{\circ}$), ΔR will be positive.

C. Contribution of Weyl bands to the pump-probe response

There are three sets of energy dispersion bands in TaAs: electron-doped W1 Weyl bands, electron-doped W2 Weyl bands, and topologically-trivial hole-doped bands with a bandgap of $\approx 50 \text{ meV}$ [5, 14]. From optical reflection spectroscopy, the Fermi energy of W1, W2, and hole-doped bands are estimated to respectively be 14 meV, 2 meV, and 20 meV for the samples considered in our study [5]. Based on the numbers from TaAs bandstrucutre calculations [15], we estimate that for $B \geq 2$ T, W2 Weyl bands go to extreme quantum limit (EQL) where only the zeroth Landau level is occupied: From the ARPES and band structure, we can estimate the Landau level energies [15]. For the W2 pockets the n=1 Landau level at 2 T is E1=8.5 meV and 9.7 meV for the two orientations. The Femi level is estimated at 4 meV so that the EQL occurs at very low B in W2. This is supported by the linear optical response [5]. The W1 and hole pockets have higher Fermi energies and have smaller E1 so that they are not necessarily in the EQL at 2 T.

For $E \parallel B \parallel z$, the principle axis of ellipsoidal pocket, the cyclotron orbit, is in the plane perpendicular to B so that the cyclotron motion, v_{perp} , is perpendicular to B and so $\mathbf{E} \cdot \mathbf{v}_{perp}$ = 0. Therefore, cyclotron resonance is not excited. Similarly, for the curved carrier Fermi surfaces, the W1 electrons and the holes in TaAs, the argument is more complex. A tilted ellipsoidal Fermi surface cyclotron resonance can be excited when the principle axis is tilted by angle θ with respect to $E \parallel B \parallel z$. In this case the cyclotron currents in the xy plane are proportional to $\sin(\theta)$. By symmetry, however, there are Fermi surface sections tilted by $-\theta$ which cancel the currents from the $+\theta$ sections. Therefore, we do not expect to observe cyclotron resonance in the $E \parallel B$ geometry as was reported in [5].

In the following, we examine the contribution of different bands to the measured pumpprobe response:

1. Hot carriers response

The initial hot carriers response exhibits a positive change in reflection that is a signature of Drude weight increase. In the gapped hole band, Sommerfeld expansion implies that the chemical potential and hence the Drude weight of carriers decreases with temperature. This behavior is similar to Drude weight dependence on temperature of semiconductors. Therefore, the observed Drude weight increase due to the hot carriers has to be from W1 and W2 Weyl bands where hot carrier electron-hole pair production can increase the Drude weight. As we increase the magnetic field, W2 carriers go to extreme quantum limit where they no longer contribute to a temperature-dependent Drude weight. Therefore, for high magnetic fields, W1 carriers only are responsible for the initial hot carriers response. This explains the behavior of Δ_1 with magnetic field in Fig. 3(c) of the main part of the manuscript.

2. Nonlinear chiral charge pumping response

In principle, both W1 and W2 Weyl nodes can contribute to a pump-induced chiral charge pumping process discussed in our study. For W1 Weyl bands where several Landau levels are occupied, the pump-induced chiral quasiparticles (yellow carriers in Fig. 1c of the main text) can relax back to equilibrium via intra-Weyl-node relaxations. On the other hand, for W2 carriers residing in the extreme quantum limit, the pump-induced chiral quasiparticles can only be relaxed back to the equilibrium chemical potential by an chiral charge pumping (inter-Weyl-node) relaxation process. Therefore, only nonlinear chiral charge pumping in W2 bands can only contribute to the observed metastable state persisting with a long time constant of τ_{ch} .

D. Estimation of the pump-induced chiral charge pumping signal

From equations (2) and (4) in the main text, the ratio of nonlinear to linear chiral pumping conductivity can be written as

$$r_{\rm NL} \equiv \left| \frac{\delta \sigma_{\rm ch}^{\rm NL}}{\sigma_{\rm ch}} \right| = \frac{9}{8} \left(\frac{\alpha e v}{\omega} \right)^2 |\tilde{E}_{\rm pump}|^2 \tag{30}$$

For W2 Weyl pocket, we have $v = 2 \times 10^5$ m/s, and $\alpha = (50 \text{meV})^{-1}$ [5, 15]. Therefore, for $\tilde{E}_{\text{pump}}=50$ kV/cm, and our measured frequency of $\hbar\omega = 14$ meV, $r_{\text{NL}} \approx 1$.

We now present an order of magnitude estimation for the nonlinear chiral charge pumping pump-probe signal characterized by Δ_2 . We take the peak terahertz electric field inside TaAs to be on average around 15 kV/cm in our measurements, and therefore $r_{\rm NL} \approx 0.1$. The linear magneto-optical conductivity associated with the chiral charge pumping is about 5 percent of the background conductivity at 14 meV photon energy and magnetic field of order of 7T, $\sigma_{ch}/\sigma_{\rm b} \approx 0.05$ [5]. This is because the Drude weight associated with the W2 pockets which is the only contributing part to the chiral chrge pumping conductivity is few percent of the total Drude weight [15]. Therefore, $|\delta\sigma_{\rm ch}^{\rm NL}/\sigma_{\rm b}| \approx 0.005$, consistent with our experimental results in Fig. 3(c) in main text.

E. Chiral charge pumping process in the pump-probe measurements

In order to provide insight into the dynamics of chiral carriers in the described pumpprobe measurements, in Fig. S1, we sketch a cartoon image of the chiral charge density \tilde{n}_{ch} produced by the probe pulse ($\mathbf{E}_{probe} \parallel \mathbf{B}$) as a function of the pump-probe time delay Δt . For $\Delta t < 0$, i.e. before the pump pulse impinges on TaAs, the chiral charge imbalance is $\tilde{n}_{ch} \propto \tilde{\sigma}_{ch} E_{probe}$ (from equation 2 in main text) and hence oscillates synchronously with E_{probe} (similar to Fig. 1b). At $\Delta t \sim 0$, when the pump pulse illuminates the TaAs, the chiral anomaly conductivity seen by the probe pulse increases by $\delta \sigma_{ch}$ (equation 3 in the main text), causing \tilde{n}_{ch} to enhance. The conductivity enhancement, which appears as an increase in the probe reflection, continues for $\Delta t > 0$ (i.e. after the pump pulse passed) as long as the quasiparticle excitations produced by the pump-induced chiral charge pumping persist. The picture is that the optical pump creates excited quasiparticles and that these excited quasiparticles modify the optical conductivity measured by the probe. As described before, this process is governed by the chiral pumping relaxation with time constant of τ_{ch} . For $\Delta t \gg \tau_{ch}$, \tilde{n}_{ch} relaxes back to the equilibrium oscillations similar to the case for $\Delta t < 0$.

VI. SUPPLEMENTAL PUMP-PROBE DATA FOR $E_{pump} \perp B$

In Fig. S2a, we plot the measured pump probe response at different magnetic fields along with the exponential fits (zero offset) to each pump probe trace. Data at lower magnetic fields exhibit a relaxation of 300 ± 50 ps. The peak of the pump-probe response exhibits an

oscillatory behavior as a function of magnetic field (Fig. S2b). This is due to the splitting of the plasma edge with applied magnetic field when $E_{pump} \perp B$ [16]. The edge for one sense of circular polarization shifts to higher frequencies by ω_c while that for the opposite polarization shifts to lower frequency by $-\omega_c$. These are the cyclotron shifted plasma edges. As a result, the reflectance/absorbance of linearly polarized 14 meV light is non monotonic. It oscillates and the peak absorption (peak electron temperature rise/peak pump-probe response) occurs when the downshifted mode has its minimum reflectance at Re(n) = 1.

Fig. S2c shows the measured pump-probe data at B = 7 T at variety of pump fluences. The black curves in Fig. S2c are exponential fits with zero offset, exhibiting similar relaxations of 400 ± 50 ps just slightly slower than the data at low magnetic fields in Fig. S2a. Therefore, all the pump-probe signals for $E_{\text{pump}} \perp E_{\text{probe}} \parallel B$ at different pump-fluences exhibit similar time dynamics, and no long-lived response.

VII. SUPPLEMENTAL PUMP-PROBE DATA FOR $E_{pump} \parallel B$

For $E_{\text{pump}} \parallel B$, for B = 4 T to 7 T, the peak of the pump probe response exhibits a slight oscillatory behavior as a function of the applied magnetic field. This behavior could be also related to weak excitation of the same cyclotron resonance that cannot be excited for a usual ellipsoidal-shaped Fermi pockets in $E_{\text{pump}} \parallel B$ geometry, but can be weakly excited due to the banana-like shape of Fermi pockets in TaAs [15]. More details of Fig.3(c) in main text are supplemented in Fig. S3, where we show exponential fits to right after the zero-delay pump-probe peak at different magnetic fields for the $E_{\text{pump}} \parallel E_{\text{probe}} \parallel B$ geometry. Finally, in Fig. S4 we show data for both geometries for a better comparison of relaxations.

VIII. XRD SPECTRUM AND DC MAGNETO-TRANSPORT OF TAAS

The XRD spectrum in the top panel of Fig. S5 illustrates the high quality of our sample TaAs. Moreover, we also measured the dc transport response in the presence of the magnetic field. Our dc transport measurements presented in in the lower panel of Fig. S5 show that our samples are consistent with the measurements presented in Ref. [17] and therefore demonstrate the high quality of the sample used in our work. These data and those of Ref. [17] also show that current jetting effects dominate the dc magneto resistance of TaAs.

This observation validates the point that terahertz magneto-optical measurements are better suited for the study of chiral pumping than dc transport.

- [1] T. Morimoto, S. Zhong, J. Orenstein, and J. E. Moore, Physical Review B 94 (2016), 10.1103/PhysRevB.94.245121.
- [2] D. T. Son and B. Z. Spivak, Physical Review B 88, 104412 (2013).
- [3] M. Kargarian, M. Randeria, and N. Trivedi, Scientific Reports 5 (2015), 10.1038/srep12683.
- [4] P. Goswami, J. H. Pixley, and S. Das Sarma, Physical Review B 92, 075205 (2015).
- [5] A. L. Levy, A. B. Sushkov, F. Liu, B. Shen, N. Ni, H. D. Drew, and G. S. Jenkins, Physical Review B, 101, 125102 (2020).
- [6] P. Hosur, S. A. Parameswaran, and A. Vishwanath, Physical Review Letters 108, 046602 (2012).
- [7] B. Xu, Y. M. Dai, L. X. Zhao, K. Wang, R. Yang, W. Zhang, J. Y. Liu, H. Xiao, G. F. Chen, A. J. Taylor, D. A. Yarotski, R. P. Prasankumar, and X. G. Qiu, Physical Review B 93 (2016), 10.1103/PhysRevB.93.121110.
- [8] G. S. Jenkins, C. Lane, B. Barbiellini, A. B. Sushkov, R. L. Carey, F. Liu, J. W. Krizan, S. K. Kushwaha, Q. Gibson, T.-R. Chang, H.-T. Jeng, H. Lin, R. J. Cava, A. Bansil, and H. D. Drew, Physical Review B 94 (2016), 10.1103/PhysRevB.94.085121.
- [9] R. Lundgren and G. A. Fiete, Physical Review B 92 (2015), 10.1103/PhysRevB.92.125139.
- [10] J. C. Song, M. Y. Reizer, and L. S. Levitov, Physical Review Letters 109, 106602 (2012).
- M. M. Jadidi, R. J. Suess, C. Tan, X. Cai, K. Watanabe, T. Taniguchi, A. B. Sushkov,
 M. Mittendorff, J. Hone, H. D. Drew, et al., Physical Review Letters 117, 257401 (2016).
- [12] M. W. Graham, S.-F. Shi, D. C. Ralph, J. Park, and P. L. McEuen, Nature Physics 9, 103 (2013).
- [13] M. M. Jadidi, J. C. Konig-Otto, S. Winnerl, A. B. Sushkov, H. D. Drew, T. E. Murphy, and M. Mittendorff, Nano Letters 16, 2734 (2016).
- [14] F. Arnold, M. Naumann, S.-C. Wu, Y. Sun, M. Schmidt, H. Borrmann, C. Felser, B. Yan, and E. Hassinger, Physical Review Letters 117, 146401 (2016).
- [15] C.-C. Lee, S.-Y. Xu, S.-M. Huang, D. S. Sanchez, I. Belopolski, G. Chang, G. Bian, N. Alidoust, H. Zheng, M. Neupane, *et al.*, Physical Review B **92**, 235104 (2015).
- [16] J. Mavroides, Optical properties of solids. North Holland, Amsterdam (1972).

[17] S. Liang, J. Lin, S. Kushwaha, J. Xing, N. Ni, R. J. Cava, and N. P. Ong, Phys. Rev. X 8, 031002 (2018).



FIG. S1. Schematic of the chiral charge density induced by the probe pulse at different pumpprobe time delays. For negative time delays ($\Delta t < 0$), the chiral current is linear with E_{probe} and sinusoidally oscillates. For positive time delays, the strong pump pulse induce changes to the probeproduced chiral current. The nonlinear pump-induced change in the chiral current goes to zero at time delays much larger than the chiral pumping relaxation time $\Delta t \gg \tau_{ch}$. We note that the conductivity measured in the presented pump-probe experiments is the nonlinear pump-induced change in the conductivity; the linear part of optical conductivity is normalized out in pump-probe traces. The schematic illustrates the charge pumping oscillations and its nonlinear pump-induced change that is being measured by pump-probe experiments.



FIG. S2. $E_{\text{pump}} \perp E_{\text{probe}} \parallel B(a)$ Pump-probe signals at different magnetic fields. The grey dashed lines are exponential fits (zero offset) to the pump-probe data, showing similar relaxation of 300 ± 50 ps. (b) Peak of the pump-probe response as a function of the applied magnetic field. (c) Relative pump-induced change in probe reflection for variety of pump fluences as a function of pump-probe time delay at B = 7 T. The black curves are exponential fits with time constant in the 400 ± 50 ps range and zero offset, that is slightly higher than for data at low magnetic fields. All the pumpprobe signals for $E_{\text{pump}} \perp E_{\text{probe}} \parallel B$ at different pump fluences and magnetic field exhibit similar time dynamics, and no long-lived response.



FIG. S3. This figure shows exponential fits to right after the zero-delay pump-probe peak at different magnetic fields for the $E_{\text{pump}} \parallel E_{\text{probe}} \parallel B$ geometry (same data as in Fig. 3(c) in the main text). At non-zero applied magnetic field, the measurements exhibit similar fast pulsewidth-limited relaxation around zero time delay, followed by a slower tail with time constant of ≈ 55 ps associated with electron-phonon cooling, and finite long-lived chiral charge pumping response increasing with magnetic field. For higher magnetic fields of 6 T and 7 T, it is harder to observe the 55 ps thermal relaxation as the response at long time delays is strongly dominated by the long-lived chiral charge pumping response.



FIG. S4. This figure combines the data set for 7T in Fig.3(a) in main text (green line) with data set (blue and black) in Fig.4(a) for a compasion between different field geometries.



FIG. S5. Top panel: XDR spectrum of TaAs measurements. Bottom panel: in-plane dc tranposrt measurement of TaAs in the presence of a coalligned magnetic field. The plot shows that the current jetting effects obscure the chiral anomaly.