Supplementary information

REDUCTION OF THE KUBO FORMULA

In this section we give the derivation of the electric-field induced nonlinear correction via a Kubo formula. The details of the diagrammatic approach are listed in Refs. [30]. We specifically the insertion of finite lifetimes according to the prescription of Ref. [32], and employ the notation therein. We expand to order τ^0, which represents the dissipation-less correction to the linear conductivity. In the diagrammatic picture, 4 diagrams contribute (may be insert these later). For simplicity, we consider here the case of $\sigma^{xx;y}$. An analogous expression can be derived for $\sigma^{yy;x}$. For compactness, we omit the Fermi occupation factor throughout. The total conductivity $\sigma^{xx;y}$ at order τ^0 reads,

$$\sigma^{xx;y} = \mathcal{W}^{xx;y} + \mathcal{V}^{xx;y}. \tag{S1}$$

$\mathcal{W}^{xx;y}$ contains contributions from two-photon vertices, while $\mathcal{V}^{xx;y}$ are three-legged diagrams. Parsing $\mathcal{W}^{xx;y}$,

$$\mathcal{W}^{xx;y}_{\tau^0} = -2\left[\varepsilon^{-3}w^{xx}, v_y\right] - \left[\varepsilon^{-3}v^{x}, w^{xy}\right] = -2\left[\varepsilon^{-3}v_y, w^{xx}\right] - \left[\varepsilon^{-3}v^{x}, w^{xy}\right]. \tag{S2}$$

Recall that $w^{xx} = 2i\Delta^x\mathcal{A}^x - [\varepsilon\mathcal{A}^x, \mathcal{A}^x] + i\varepsilon S^{xx}$. Furthermore, $w^{xy} = i\Delta^x\mathcal{A}^y + i\Delta^y\mathcal{A}^x - \frac{1}{2}\left[\varepsilon\mathcal{A}^x, \mathcal{A}^y\right] - \frac{1}{2}\left[\varepsilon\mathcal{A}^y, \mathcal{A}^x\right] + i\varepsilon S^{xy}$. $v^{x,y} = i\varepsilon\mathcal{A}^{x,y}$ (since only off-diagonal components are involved).

$$\begin{aligned}\mathcal{W}^{xx;y}_{\tau^0} = &-2i\left[\varepsilon^{-2}\mathcal{A}^y, 2i\Delta^x\mathcal{A}^x - [\varepsilon\mathcal{A}^x, \mathcal{A}^x] + i\varepsilon S^{xx}\right] - \\ &i\left[\varepsilon^{-2}\mathcal{A}^x, i\Delta^x\mathcal{A}^y + i\Delta^y\mathcal{A}^x - (1/2)[\varepsilon\mathcal{A}^x, \mathcal{A}^y] - (1/2)[\varepsilon\mathcal{A}^y, \mathcal{A}^x] + i\varepsilon S^{xy}\right].\end{aligned} \tag{S3}$$

Terms may now be rearranged given the transposition properties of objects inside the commutator. For example, $\left[\varepsilon^{-2}\mathcal{A}^x, \Delta^x\mathcal{A}^y\right] = \left[\varepsilon^{-2}\mathcal{A}^y, \Delta^x\mathcal{A}^x\right]$. This stems from the fact that $\varepsilon^2_{nm} = \varepsilon^2_{mn}$, but $\Delta^x_{nm} = -\Delta^x_{mn}$. Thus,

$$\begin{aligned}\mathcal{W}^{xx;y}_{\tau^0} = &\,5\left[\varepsilon^{-2}\mathcal{A}^y, \Delta^x\mathcal{A}^x\right] + \left[\varepsilon^{-2}\mathcal{A}^x, \Delta^y\mathcal{A}^x\right] + 2i\left[\varepsilon^{-2}\mathcal{A}^y, [\varepsilon\mathcal{A}^x, \mathcal{A}^x]\right] - 2\left[\varepsilon^{-1}\mathcal{A}^y, S^{xx}\right] + \\ &\frac{i}{2}\left[\varepsilon^{-2}\mathcal{A}^x, [\varepsilon\mathcal{A}^x, \mathcal{A}^y] + [\varepsilon\mathcal{A}^y, \mathcal{A}^x]\right] - \left[\varepsilon^{-1}\mathcal{A}^x, S^{xy}\right].\end{aligned} \tag{S4}$$

Next we turn our to $\mathcal{V}^{xx;y}$. As the expressions contain denominators ε_{nl} which depend on an intermediate index, we first isolate two cases of interest: $l = n$, $l = m$. The remainder are pieces for which $l \neq n, m$ and therefore are amenable to being written as proper commutators, as noted in the introduction. Since this section will involve explicit diagonal parts of the velocity operators $v^{x,y}_{nn}$ we restore the Fermi occupation factors. Firstly,

$$\mathcal{V}^{xx;y}_{l=n,\tau^0} = \frac{f_{nm}}{2\varepsilon^4_{nm}}v^x_{nm}v^x_{mn}v^y_{nn} + \frac{f_{nm}}{2\varepsilon^4_{nm}}v^y_{nn}v^x_{mn}v^x_{nm} \tag{S5}$$

Interchanging the summation on the second term ($n \leftrightarrow m$), gives,

$$\frac{f_{nm}}{2\varepsilon^4_{nm}}\left(v^x_{nm}v^x_{mn}(v^y_{nn} - v^y_{mm})\right) = \frac{1}{2}\left[\frac{v^x}{\varepsilon^4}\Delta^y, v^x\right] = \frac{1}{2}\left[\frac{\mathcal{A}^x}{\varepsilon^2}\Delta^y, \mathcal{A}^x\right]. \tag{S6}$$

Next is the case of $l = m$,

$$\mathcal{V}^{xx;y}_{l=m,\tau^0} = -7\frac{f_{nm}}{\varepsilon^4_{nm}}v^x_{nm}v^x_{mm}v^y_{mn} - 7\frac{f_{nm}}{\varepsilon^4_{nm}}v^y_{nm}v^x_{mm}v^x_{mn} = 7\left[\varepsilon^{-4}v^y\Delta^x, v^x\right] = 7\left[\varepsilon^{-2}\mathcal{A}^y\Delta^x, \mathcal{A}^x\right] = -7\left[\varepsilon^{-2}\mathcal{A}^y, \Delta^x\mathcal{A}^x\right]. \tag{S7}$$

The remaining terms, such that $l \neq n, m$ are,

$$\mathcal{V}^{xx;y,(1)}_{l\neq n,m,\tau^0} = -4\frac{f_{nm}}{\varepsilon_{nm}\varepsilon^3_{nl}}v^x_{nm}v^x_{ml}v^y_{ln} - 2\frac{f_{nm}}{\varepsilon^2_{nm}\varepsilon^2_{nl}}v^x_{nm}v^x_{ml}v^y_{ln} - \frac{f_{nm}}{\varepsilon^3_{nm}\varepsilon_{nl}}v^x_{nm}v^x_{ml}v^y_{ln}, \tag{S8}$$

while $\mathcal{V}^{xx;y,(2)}_{l\neq n,m}$ is the complex conjugate of $\mathcal{V}^{xx;y,(1)}_{l\neq n,m}$. Let us treat each term separately. After substituting $v^{x,y}$, and adding the complex conjugate the first term yields $-4\frac{f_{nm}}{\varepsilon_{nm}\varepsilon^3_{nl}}v^x_{nm}v^x_{ml}v^y_{ln} = -4i\left[\mathcal{A}^x, \left[\varepsilon\mathcal{A}^x, \varepsilon^{-2}\mathcal{A}^y\right]\right]$. We break this term up into two pieces of equal prefactor, 2, and the first piece is replaced using a Jacobi identity. That is, $-2i\left[\mathcal{A}^x, \left[\varepsilon\mathcal{A}^x, \varepsilon^{-2}\mathcal{A}^y\right]\right] = 2i\left[\varepsilon^{-2}\mathcal{A}^y, [\mathcal{A}^x, \varepsilon\mathcal{A}^x]\right] + 2i\left[\varepsilon\mathcal{A}^x, \left[\varepsilon^{-2}\mathcal{A}^y, \mathcal{A}^x\right]\right]$. The other piece is written down differently.

Using the fact that $\varepsilon_{ml} = \varepsilon_{mn} + \varepsilon_{nl}$, $-2i\left[\mathcal{A}^x, \left[\varepsilon\mathcal{A}^x, \varepsilon^{-2}\mathcal{A}^y\right]\right] = -2i\left[\mathcal{A}^x, \varepsilon\left[\mathcal{A}^x, \varepsilon^{-2}\mathcal{A}^y\right]\right] + 2i\left[\mathcal{A}^x, \left[\mathcal{A}^x, \varepsilon^{-1}\mathcal{A}^y\right]\right]$. Adding this up once more gives,

$$-4i\left[\mathcal{A}^x, \left[\varepsilon\mathcal{A}^x, \varepsilon^{-2}\mathcal{A}^y\right]\right] = 2i\left[\varepsilon^{-2}\mathcal{A}^y, \left[\mathcal{A}^x, \varepsilon\mathcal{A}^x\right]\right] + 2i\left[\mathcal{A}^x, \left[\mathcal{A}^x, \varepsilon^{-1}\mathcal{A}^y\right]\right]. \tag{S9}$$

The next term in Eq. (S8) is decomposed analogously.

$$-2\frac{f_{nm}}{\varepsilon_{nm}^2\varepsilon_{ln}^2} v_{nm}^x v_{ml}^x v_{ln}^y = 2i\left[\frac{\mathcal{A}^x}{\varepsilon}, \left[\varepsilon\mathcal{A}^x, \varepsilon^{-1}\mathcal{A}^y\right]\right] = -2i\left[\mathcal{A}^x, \left[\mathcal{A}^x, \varepsilon^{-1}\mathcal{A}^y\right]\right] - 2i\left[\varepsilon^{-1}\mathcal{A}^x, \left[\mathcal{A}^x, \mathcal{A}^y\right]\right] \tag{S10}$$

The last term in Eq. (S8) is given by,

$$-\frac{f_{nm}}{\varepsilon_{nm}^3\varepsilon_{nl}} v_{nm}^x v_{ml}^x v_{ln}^y = -i\left[\varepsilon^{-2}\mathcal{A}^x, \left[\varepsilon\mathcal{A}^x, \mathcal{A}^y\right]\right]. \tag{S11}$$

We are now ready to assemble all the pieces we have. We combine,

$$\begin{aligned}
&\mathcal{V}^{xx;y} + \mathcal{W}^{xx;y} = \\
&5\left[\varepsilon^{-2}\mathcal{A}^y, \Delta^x\mathcal{A}^x\right] + \left[\varepsilon^{-2}\mathcal{A}^x, \Delta^y\mathcal{A}^x\right] + 2i\left[\varepsilon^{-2}\mathcal{A}^y, \left[\varepsilon\mathcal{A}^x, \mathcal{A}^x\right]\right] - 2\left[\varepsilon^{-1}\mathcal{A}^y, S^{xx}\right] + \\
&\frac{i}{2}\left[\varepsilon^{-2}\mathcal{A}^x, \left[\varepsilon\mathcal{A}^x, \mathcal{A}^y\right] + \left[\varepsilon\mathcal{A}^y, \mathcal{A}^x\right]\right] - \left[\varepsilon^{-1}\mathcal{A}^x, S^{xy}\right] + \frac{1}{2}\left[\frac{\mathcal{A}^x}{\varepsilon^2}\Delta^y, \mathcal{A}^x\right] - 7\left[\varepsilon^{-2}\mathcal{A}^y, \Delta^x\mathcal{A}^x\right] + 2i\left[\varepsilon^{-2}\mathcal{A}^y, \left[\mathcal{A}^x, \varepsilon\mathcal{A}^x\right]\right] \\
&+ 2i\left[\mathcal{A}^x, \left[\mathcal{A}^x, \varepsilon^{-1}\mathcal{A}^y\right]\right] - 2i\left[\mathcal{A}^x, \left[\mathcal{A}^x, \varepsilon^{-1}\mathcal{A}^y\right]\right] - 2i\left[\varepsilon^{-1}\mathcal{A}^x, \left[\mathcal{A}^x, \mathcal{A}^y\right]\right] - i\left[\varepsilon^{-2}\mathcal{A}^x, \left[\varepsilon\mathcal{A}^x, \mathcal{A}^y\right]\right].
\end{aligned} \tag{S12}$$

We obtain,

$$\begin{aligned}
&\mathcal{V}^{xx;y} + \mathcal{W}^{xx;y} = \\
&-2\left[\varepsilon^{-2}\mathcal{A}^y, \Delta^x\mathcal{A}^x\right] + \frac{1}{2}\left[\varepsilon^{-2}\mathcal{A}^x, \Delta^y\mathcal{A}^x\right] - 2\left[\varepsilon^{-1}\mathcal{A}^y, S^{xx}\right] + \frac{i}{2}\left[\varepsilon^{-2}\mathcal{A}^x, -\left[\varepsilon\mathcal{A}^x, \mathcal{A}^y\right] + \left[\varepsilon\mathcal{A}^y, \mathcal{A}^x\right]\right] + \\
&-\left[\varepsilon^{-1}\mathcal{A}^x, S^{xy}\right] - 2i\left[\varepsilon^{-1}\mathcal{A}^x, \left[\mathcal{A}^x, \mathcal{A}^y\right]\right].
\end{aligned} \tag{S13}$$

This term contains several total derivatives, which are Fermi surface terms, and give rise to the recently proposed gravitational anomaly and intrinsic non-dissipative Hall effects [40, 56]. We remove them by observing that,

$$\partial_y\left[\varepsilon^{-1}\mathcal{A}^x, \mathcal{A}^x\right] = 2\left[\varepsilon^{-1}\mathcal{A}^x, S^{xy}\right] + \left[\varepsilon^{-2}\mathcal{A}^x, \Delta^y\mathcal{A}^x\right] + \left[\varepsilon^{-1}\mathcal{A}^x, i[\mathcal{A}^y, \mathcal{A}^x]\right] \tag{S14}$$

$$\partial_x\left[\varepsilon^{-1}\mathcal{A}^y, \mathcal{A}^x\right] = \left[\varepsilon^{-1}\mathcal{A}^x, S^{xy}\right] + \left[\varepsilon^{-1}\mathcal{A}^y, S^{xx}\right] - \left[\varepsilon^{-2}\mathcal{A}^y\Delta^x, \mathcal{A}^x\right] + \left[\varepsilon^{-1}\mathcal{A}^x, (i/2)[\mathcal{A}^x, \mathcal{A}^y]\right]. \tag{S15}$$

By combining these identities with the symmetrization condition one finds,

$$\begin{aligned}
\sigma^{xx;y} &= 2\partial_y\left[\varepsilon^{-1}\mathcal{A}^x, \mathcal{A}^x\right] - \partial_x\left[\varepsilon^{-1}\mathcal{A}^y, \mathcal{A}^x\right] + \\
&\frac{1}{2}\left[\varepsilon^{-2}\mathcal{A}^y, \Delta^x\mathcal{A}^x\right] - \frac{1}{2}\left[\varepsilon^{-2}\mathcal{A}^x, \Delta^y\mathcal{A}^x\right] + \left[\varepsilon^{-1}\mathcal{A}^x, S^{xy}\right] - \left[\varepsilon^{-1}\mathcal{A}^y, S^{xx}\right] + \left[\varepsilon^{-1}\mathcal{A}^x, (i/2)[\mathcal{A}^x, \mathcal{A}^y]\right].
\end{aligned} \tag{S16}$$

In the above, we employed the identity that $\tilde{\Omega}^{xy,1} - \tilde{\Omega}^{yx,1} = \varepsilon\Omega^{ab}$, otherwise proven here [32]. We note that this form is fully compatible with the consistent separation of nonlinear response conductivity into Hall components carried out in Ref. [57]. We also note that the positional shift S^{xy} described here is related (while more general) to the quantum metric dipole, also recently shown to present an in-gap Hall conductivity [58].

GAUGE INVARIANCE OF THE DERIVED CORRECTION

The introduction of derivatives of the wavefunction in the expressions for $I_1 - I_3$ Eqs. (4)-(6) requires delicate handling of gauge transformations. The Bloch manifold of cell-periodic states is characterized by an invariance to the $U(1)^N$ transformation,

$$|n\mathbf{k}\rangle \to e^{-i\theta_n(k)}\,|n\mathbf{k}\rangle\,. \tag{S17}$$

For notational ease, we suppress below the label $\mathbf{k}$, and refer to $|n\rangle$, as $|n\rangle = |n\mathbf{k}\rangle$. The Bloch periodic part of the Hamiltonian commutes with this gauge transformation since $[H(\mathbf{k}), f(k)] = 0$. The observables of optical response are modified due to the gauge covariance of the states. The velocity operator $v^\alpha \rightarrow Uv^\alpha$, where $U_{nm} = e^{i\theta_{nm}}$, $\theta_{nm} = \theta_n - \theta_m$. Clearly, only diagonal components, such as those comprising $\Delta^\alpha_{nm} = v^\alpha_{nn} - v^\alpha_{mm}$, are automatically gauge invariant. Products such as $(v^\alpha)^\dagger v^\beta \rightarrow v^\alpha U^\dagger U v^\beta = v^\alpha v^\beta$ are gauge invariant. In this respect, the equations of optical response in the velocity gauge are manifestly gauge invariant, as they combine products of gauge covariant operators derived from the Hamiltonian. The issue of local $U(1)$ gauge invariance in the context of electromagnetism and optical response has recently been addressed, with gauge invariance formally proven [59–62]. Our primary focus, therefore, is to show that Eqs. (4)-(6) which involve *derivatives* of the Bloch periodic part of the electronic wavefunctions are also gauge invariant. The principle derivatives defining 2nd order optical response are [32]:

$$\mathcal{A}^\alpha_{nm} = \langle n\,|i\partial_\alpha|\, m\rangle \rightarrow \partial_\alpha\theta_n\delta_{nm} + e^{i\theta_{nm}}\langle n\,|i\partial_\alpha|\, m\rangle = e^{i\theta_{nm}}\mathcal{A}^\alpha_{nm} + \partial_\alpha\theta_n\delta_{nm} \tag{S18}$$

$$\lambda^{\alpha\beta}_{nm} = \frac{1}{2}\langle n\,|i\partial_\alpha\partial_\beta|\, m\rangle + (\text{c.c.}, n \leftrightarrow m) \rightarrow e^{i\theta_{nm}}\lambda^{\alpha\beta}_{nm} + \frac{ie^{i\theta_{nm}}}{2}\left(\mathcal{A}^\alpha_{nm}\partial_\beta\theta_{nm} + \mathcal{A}^\beta_{nm}\partial_\alpha\theta_{nm}\right),\ n \neq m. \tag{S19}$$

We define objects which are gauge *covariant* as those which transform as $A_{nm} \rightarrow A_{nm}e^{i\theta_{nm}}$. Consequently, combinations of the form $A_{nm}A_{mn}$ are manifestly gauge *invariant* since they transform like the velocity operator, as shown above, or $e^{i\theta_{nm}}A_{nm}e^{i\theta_{mn}}A_{mn} = A_{nm}A_{mn}$. Generally, for *any* two covariant objects A, B, the commutator of the two satisfies,

$$[A, B]_{nm} \rightarrow e^{i\theta_{nm}}[A, B]_{nm}, \tag{S20}$$

rendering its diagonal part gauge invariant. This follows from the definition introduced in the main text, $[A, B]_{nm} = \sum_{l\neq n,m} A_{nl}B_{lm} - (A \leftrightarrow B)$. We now prove that the quantity $S^{\alpha\beta}_{nm}, n \neq m$ is gauge covariant, which appears in I_2, of Eq. (5). We note that $S^{\alpha\beta}$ consists of two portions: $\lambda^{\alpha\beta}_{nm}$ and the Hadamard product, $\frac{i}{2}\mathcal{A}^\alpha_{nm}\delta^\beta_{nm} + (\alpha \leftrightarrow \beta)$. Tackling the latter first,

$$\frac{i}{2}\mathcal{A}^\alpha_{nm}\delta^\beta_{nm} \rightarrow \frac{i}{2}e^{i\theta_{nm}}\mathcal{A}^\alpha_{nm}\partial_\beta\theta_{nm}. \tag{S21}$$

To uncover the transformation properties of $\lambda^{\alpha\beta}_{nm}$ for $n \neq m$, we first observe that under $|m\rangle \rightarrow e^{i\theta_m}|m\rangle$, $\partial_\alpha\partial_\beta|m\rangle \rightarrow e^{i\theta_m}\left(i\partial_\alpha\theta_m|\partial_\beta m\rangle + i\partial_\beta\theta_m|\partial_\alpha m\rangle + |\partial_\alpha\partial_\beta m\rangle\right)$. We explicitly remove the term $e^{i\theta_m}\partial_\alpha\partial_\beta\theta_m|m\rangle$ since we assume that $n \neq m$ and this contribution must vanish when projected back onto the Bloch states. Multiplying on the left with $i\langle n|$, we find that,

$$\lambda^{\alpha\beta}_{nm} \rightarrow \frac{e^{i\theta_{nm}}}{2}\lambda^{\alpha\beta}_{nm} + \frac{i}{2}e^{i\theta_{nm}}\mathcal{A}^\alpha\partial_\beta\theta_{mn} + (\alpha \leftrightarrow \beta). \tag{S22}$$

It follows that the sum of the objects,

$$\frac{1}{2}\lambda^{\alpha\beta}_{nm} + \frac{i}{2}A^\alpha_{nm}\delta^\beta_{nm} + (a \leftrightarrow b) \rightarrow e^{i\theta_{nm}}(\frac{1}{2}\lambda^{\alpha\beta}_{nm} + \frac{i}{2}A^\alpha_{nm}\delta^\beta_{nm}) + (a \leftrightarrow b), \tag{S23}$$

With the θ dependent part explicitly cancelling as it appears with opposite indices θ_{nm} vs θ_{mn}. Lastly, the triple commutator introduced in Eq. (6) transforms using the rules outlined above for commutators. The product reads,

$$[\mathcal{A}^x, i[\mathcal{A}^x, \mathcal{A}^y]]_{nn} = i\mathcal{A}^x_{nm}\left[\mathcal{A}^x, \mathcal{A}^y\right]_{mn} - i\left[\mathcal{A}^x, \mathcal{A}^y\right]_{nm}\mathcal{A}^x_{mn} \rightarrow e^{i\theta_{nm}}\mathcal{A}^x_{nm}e^{i\theta_{mn}}[\mathcal{A}^x, \mathcal{A}^y]_{mn} - (\text{c.c.}) = [\mathcal{A}^x, i[\mathcal{A}^x, \mathcal{A}^y]]_{nn}. \tag{S24}$$

Here, we used the fact that commutator as defined does not sum over any diagonal contributions and off-diagonal parts of the Berry connection transform according to Eq. (S18).

QUANTIZATION OF THE LINEAR KUBO FORMULA FOR THE IQHE

The linear response contribution at zero frequency reads [63],

$$\sigma^{xy} = \frac{ie^2}{L_xL_y\hbar}\sum_{n,m} f_{nm}\frac{v^x_{nm}v^y_{mn}}{\varepsilon^2_{nm}}. \tag{S25}$$

Here the sum n, m runs over all bands (occupied and empty) including the degenerate manifold of each occupied Landau level. Based on the identities defined in the main text for the velocity matrix elements, we have,

$$v^x_{nm}v^y_{mn} = \frac{i\hbar\omega_c}{2M}\left(\sqrt{m+1}\delta_{n,m+1} - \sqrt{m}\delta_{n,m-1}\right)\left(\sqrt{m+1}\delta_{n,m+1} + \sqrt{m}\delta_{n,m-1}\right). \tag{S26}$$

The energy difference (measured in frequency units) is $\varepsilon_{nm} = \omega_c(n-m)$. The only surviving terms above are products of equal delta functions. Thus,

$$\frac{v^x_{nm} v^y_{mn}}{\varepsilon^2_{nm}} = \frac{i\hbar\omega_c}{2M} \frac{((m+1)\delta_{n,m+1} - m\delta_{n,m-1})}{\omega_c^2(n-m)^2}. \tag{S27}$$

Assume now that there are ν occupied bands, which are each $N = \frac{eBL_xL_y}{h}$-fold degenerate. The linear conductivity Eq. (S25) has a global Fermi occupation factor difference $f_{nm} = f_n - f_m$. In the bulk of the quantum Hall fluid, a fully flat Landau band is either completely occupied or completely empty. Therefore, f_{nm} is non-zero only in the cases $n = \nu, m = \nu + 1, f_{nm} = 1$, and $n = \nu + 1, m = \nu, f_{nm} = -1$. By further accounting for the degeneracy, we have,

$$\begin{aligned}\sigma^{xy} &= \frac{ie^2}{L_xL_y\hbar} \sum_{n,m} f_{nm} \frac{v^x_{nm} v^y_{mn}}{\varepsilon^2_{nm}} = \\ &\frac{eBL_xL_y}{h} \times \left[\frac{ie^2}{L_xL_y\hbar} \frac{i\hbar\omega_c}{2M} \left(\frac{-(\nu+1)}{\omega_c^2} - \frac{(\nu+1)}{\omega_c^2}\right)\right] = \frac{e^2}{h}(\nu+1).\end{aligned} \tag{S28}$$

σ^{xy} is therefore quantized by the number of filled Landau levels.

TWO BAND MODELS

The simplest approximation that can be made consists of a two band topological model with non-vanishing Berry curvature. We show here that a fundamental condition for the emergence of our derived correction is the presence of nonlinear terms. A minimal Hamiltonian for the which produces a non-zero correction reads,

$$H = (M + 2 - \cos(k_x) - \cos(k_y))\sigma_z \ + (\sin(k_x) + \alpha L_x)\, \sigma_x + (\sin(k_y) + \beta L_y)\, \sigma_y \tag{S29}$$

$$L_x = \sin(k_x)\sin(k_y),\ L_y = \sin(k_x)\cos(k_y), \tag{S30}$$

which is a model for a single Dirac cone with nonlinear terms represented by L_x, L_y. Here, σ are Pauli matrices in an orbital basis. When $\alpha = \beta = 0$, the model retains inversion symmetry, which is defined by $P = \sigma_z(k \to -k)$. This remains true when $\alpha = 0$ alone, since L_y is odd under $(k \to -k)$. The model is topological for all $-2 < M < 0$ with Chern number $C_N = -1$, and is trivial otherwise. The nonlinearities L_x, L_y do not merely break inversion and mirror symmetries remaining mirror but add higher derivatives of the Hamiltonian, i.e., $w^{ab} = \partial_a\partial_b H_0$, and induce a correction to the current operator, according to Eq. (1). In the absence of non-linearities, sum rules [64], notably that the fact that $\partial_a v^b_{nm} = i[r^a, v^b]_{nm} + iv^b_{nm}\delta^a_{nm} + ir^a_{nm}\Delta^b_{nm}$ emerge, enforcing cancellations between terms and nulling the correction. In Fig. S1(b), we show the magnitude of the correction as a function of of the mass parameter M of the model. The response diverges as $M \to 0$ due to the presence of ε^{-n}_{nm}, $n > 1$ in all terms. Unlike the Berry curvature, the flux of this expression, i.e. the integral $\sim \int \mathrm{d}k\varepsilon^{-n}k$ does not equal a constant but decays as k^{-3} in the leading order. This makes the correction fundamentally different from the Berry curvature, as it is highly singular in the limit $k \to 0$ for a vanishing mass. The corrections are sensitive to the topology of the system, as shown in Fig. S1(b). In the non-topological case ($M > 0$), with vanishing total Berry phase, the corrections decay roughly as $M^{-3/2}$ while for $M < 0$, the decrease is of the form M^{-1}, indicating a slower suppression. The momentum space distribution is also different for the two cases. In Figs. S1(c-d) we plot the momentum space density of the terms $I_1 - I_3$ (Eqs. (4)-(6)). In the trivial case ($M > 0$, Fig. S1(d)), the divergence around the Γ point of the model is clearly apparent, while when $M < 0$ the density is diffused across a broader region in momentum space, with the peak way from the Γ point. As stated above, in the 2-band limit, I_3 (Eq. (6)) cannot contribute, and the non-linearity is most significantly encountered for the velocity shift I_1 (Eq. (4)).

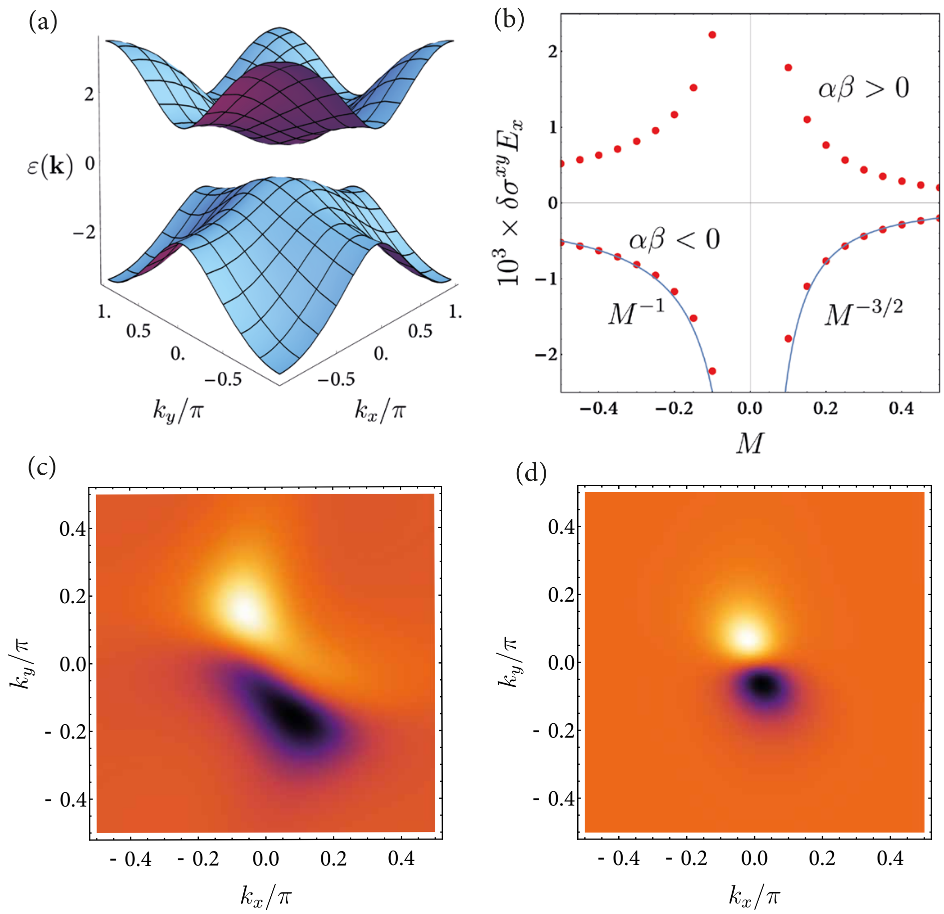

Figure S1. Nonlinear correction to the anomalous Hall conductivity of a gapped Dirac cone. (a) A two band model (Eq. (S30)) in the topological phase, when $M = -0.5$. (b) Magnitude of the correction $\delta\sigma^{xy}$ to the Hall conductivity, for $E_x = 1$. The sign of the correction depends on the sign of the product $\alpha\beta$. In the topological phase ($M < 0$), the correction decays like $M{-1}$, while in the trivial phase ($M > 0$) it decreases with $M^{-3/2}$. (c-d) Momentum space distribution of $I_1 - I_3$ (Eqs. (4)-(6)) for the topological phase (c) and the trivial phase (d)