

Hot-Carrier Distribution Spectroscopy by Transconductance in Two-Dimensional Field-Effect Transistors

Katsunori Wakabayashi

Research Center for Materials Nanoarchitectonics (MANA),
National Institute for Materials Science (NIMS), Namiki 1-1, Tsukuba 305-0044, Japan

(Dated: September 10, 2026)

The transconductance $g_m = dI_D/dV_G$ of a field-effect transistor (FET) is conventionally read as a proxy for carrier density. We show that it is instead a spectroscopic probe of the carrier distribution: because g_m weights the spectral current $j(E)$ by the gate-voltage derivative $\partial f(E)/\partial V_G$ and integrates over energy, it is sensitive to the *shape* of $f(E)$, not merely its integrated weight n . We develop an energy-resolved transport framework for two-dimensional (2D) FETs and, within a gate-independent spectral-kernel approximation, derive the decomposition $g_m = g_m^{(n)} + g_m^{(\alpha)}$ into the conventional density-modulation term $g_m^{(n)}$ and a distribution-shape-driven term $g_m^{(\alpha)}$. The latter, obtained as the residual after subtracting the smooth density-modulation background from the measured g_m , exhibits a characteristic anomalous peak at a gate voltage V_G^{pk} . This peak has no counterpart in equilibrium transport and *cannot be explained by carrier density modulation alone*. With the spectral kernel calibrated, the peak position and height—extracted from standard DC/lock-in g_m sweeps—constrain the hot-carrier energy E_0 , spectral width σ , and generation threshold n_c , realizing a steady-state, all-electrical spectroscopy of the carrier distribution. An optional time-resolved extension further recovers the carrier relaxation time τ from the transient response following a pump excitation, establishing the 2D FET as a distribution-function spectrometer that requires no optical readout.

I. INTRODUCTION

Two-dimensional (2D) semiconductors such as transition metal dichalcogenides (TMDs) [1–3] and their van der Waals heterostructures have emerged as a rich arena for exploring electronic transport in reduced dimensions [4, 5]. Field-effect transistors (FETs) fabricated from these materials simultaneously serve as testbeds for fundamental physics and as building blocks of next-generation devices [6–10].

In the standard description of FET transport, the drain current is written as

$$I_D = \frac{qW}{L} n \mu V_D, \quad (1)$$

where q is the elementary charge, W and L are the channel width and length, n is the areal carrier density, μ is the mobility, and V_D is the source–drain voltage [11]. This relation implicitly assumes that carriers obey an equilibrium or quasi-equilibrium distribution, so that μ is a single scalar that summarizes all microscopic scattering. Under strong source–drain fields, carrier injection, or optical pumping, however, the distribution function acquires a hot-carrier tail whose energy is well above the lattice temperature. In such nonequilibrium states, the shape of the distribution—not merely its integrated weight n —governs transport, rendering Eq. (1) inadequate [12, 13].

Nonequilibrium carriers in 2D materials have been studied extensively by ultrafast optical spectroscopy [14]. Pump–probe experiments on graphene [15] and MoS₂ [16–19] reveal hot-carrier relaxation times in the range $\tau \sim 0.1$ –10 ps, reflecting carrier–phonon scattering dynamics. Yet such optical probes are

rarely combined with operando electrical gating, and they do not directly access the transport-relevant, velocity-weighted distribution that governs I_D .

Here we develop an *energy-resolved* transport theory for 2D FETs that connects the shape of $f(E)$ to measurable device characteristics. Our starting point is a spectroscopic reformulation of transport: within an energy-resolved transport description with a gate-independent spectral current kernel $j(E)$, the transconductance takes the form

$$g_m = \frac{qW}{L} \int_0^\infty j(E) \frac{\partial f(E; V_G)}{\partial V_G} dE, \quad (2)$$

which makes explicit that *the transconductance does not merely reflect carrier density modulation but probes the energy distribution of carriers*. Equation (2) is itself a natural consequence of writing the current as a spectral integral and differentiating with respect to V_G . It is a close relative of the Landauer–Büttiker formula [13, 20] and of the transport distribution function of thermoelectric transport [21]. More broadly, it instantiates a general principle—that a derivative of the conductance resolves spectral structure—which underlies scanning-tunneling [22] and quantum-dot transconductance [23] spectroscopies as well as the electrical measurement of the quasiparticle distribution in mesoscopic wires [24].

Our contribution is to apply this principle to the *shape of a nonequilibrium distribution* in a 2D FET, where it yields the specific, measurable anomaly developed below. In a related line of work, the same transconductance-as-probe principle is applied to the valley degree of freedom in WSe₂, using g_m to read out valley thermodynamics in multilayer films [25] and to guide valley engineering in bilayer gate-all-around transistors [26]. A dy-

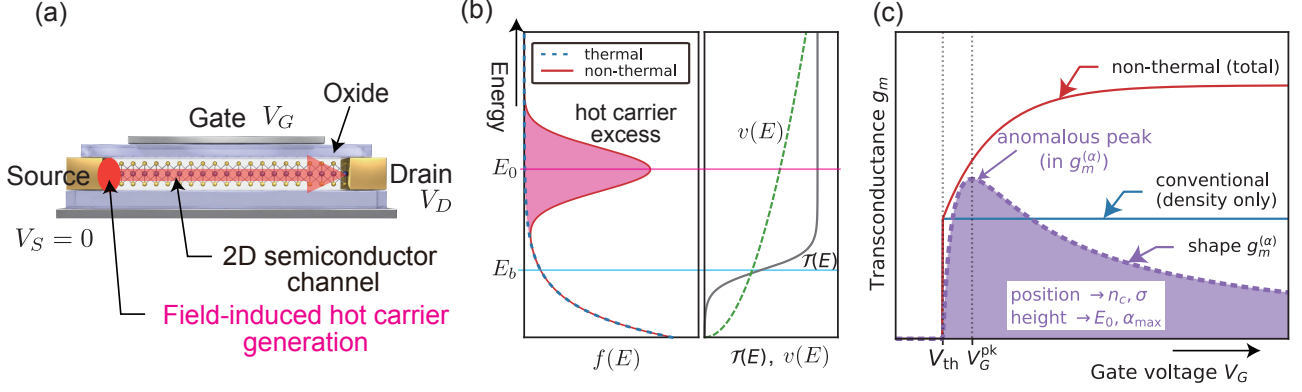


FIG. 1. Concept, device, and measurement principle (schematic). (a) A gated monolayer 2D-semiconductor field-effect transistor: a gate voltage tunes the carrier density in the channel between source and drain contacts, while field-induced hot-carrier generation drives the distribution out of equilibrium. (b) Physical mechanism (energy on the vertical axis). *Left*: the carrier distribution $f(E)$ is either thermal (dotted) or non-thermal (red), the latter carrying a localized hot-carrier excess near E_0 . *Right*: the transport kernel—an energy-selective transmission $\mathcal{T}(E)$ that turns on near the onset E_b and a velocity $v(E)$ that grows with energy—weights high-energy carriers more heavily, so each carrier in the excess near $E_0 > E_b$ contributes more current. (c) Measurement principle (schematic): a conventional, density-only response is featureless in V_G (blue), and the total non-thermal g_m (red) rises monotonically. The non-thermal signature is the shape-driven contribution $g_m^{(\alpha)}$ (purple, shaded), which produces an anomalous peak whose position and height read out the spectral parameters of $f(E)$ —transconductance spectroscopy. (In the representative case shown here the total measured g_m itself remains monotone; the peak is isolated in $g_m^{(\alpha)}$, as quantified in Fig. 3.) The response turns on at the threshold voltage V_{th} [Eq. (12)], where the channel begins to populate, and the anomaly peaks at V_G^{pk} [Eq. (20)].

namical extension of the same idea uses the transconductance to extract the intervalley relaxation time of WSe₂ transistors [27]. Corrections from a V_G -dependent kernel $j(E, V_G)$ (e.g. barrier lowering, contact modulation, or band bending) are additive and discussed in Sec. IV.

For a Boltzmann distribution, $\partial f_{eq}/\partial V_G$ is featureless in energy, and g_m is flat—the conventional result. For a nonequilibrium distribution with a hot-carrier peak at E_0 , $\partial f_{neq}/\partial V_G$ acquires structure at E_0 and g_m reports it as a measurable anomaly in the shape-resolved residual. The Gaussian hot-carrier peak adopted below is only a minimal model: the mechanism applies to any localized nonthermal excess $\delta f(E)$ above the transport onset, not to a specific line shape. Figure 1(a) shows the device geometry, (b) the spectral mechanism, and (c) the resulting measurement principle.

Our main results are threefold. (i) The drain current takes the compact form $I_D = (qW/L) n \bar{v}(\alpha)$, with an effective velocity $\bar{v}(\alpha)$ set by a single parameter α that quantifies the hot-carrier population. (ii) Within the same spectral-kernel approximation, g_m admits a transparent decomposition $g_m = g_m^{(n)} + g_m^{(\alpha)}$. The new term $g_m^{(\alpha)}$ produces an anomalous peak in the shape-resolved residual—obtained after calibrating and subtracting the smooth density-modulation background—that *cannot be explained by carrier density modulation* and directly constrains the spectral parameters $\{E_0, \sigma, n_c\}$. This peak carries a decisive drain-voltage fingerprint—its position shifts to lower V_G and its height saturates with V_D —that

no separable, quasi-equilibrium mobility-rolloff model can reproduce. (iii) A time-resolved electrical measurement after a pump excitation yields the carrier relaxation time τ , completing an all-electrical spectroscopy of the nonequilibrium distribution.

The remainder of this paper is organized as follows. Section II develops the energy-resolved transport framework and derives the decomposition $g_m = g_m^{(n)} + g_m^{(\alpha)}$. Section III presents the anomalous $g_m^{(\alpha)}$ peak, its drain-voltage fingerprint, and the resulting extraction of the spectral parameters $\{E_0, \sigma, n_c\}$. Section IV places the framework in the context of prior hot-carrier work and addresses corrections from a gate-dependent kernel, and Sec. V concludes. Appendices A–C present the time-resolved extension, a fit to synthetic data, and the robustness of the anomaly against modeling choices.

II. MODEL AND THEORY

A. Energy-resolved drain current

For definiteness, we formulate the model for an n-type 2D FET, where $E \geq 0$ denotes the electron kinetic energy measured from the conduction-band minimum. The same formalism applies to p-type devices by interpreting E as the hole kinetic energy measured from the valence-band maximum.

For a 2D semiconductor of width W and length L , we

write the drain current as an integral over carrier energy [12, 13]:

$$I_D = \frac{qW}{L} \int_0^\infty D(E) v(E) \mathcal{T}(E) f(E) dE, \quad (3)$$

where $D(E)$ is the density of states, $v(E)$ is the group velocity, $\mathcal{T}(E)$ is an energy-selective transmission probability, and $f(E)$ is the carrier distribution function. For a 2D system with parabolic dispersion, $D(E) = D_0 = gm^*/(\pi\hbar^2)$ is constant (with g the spin-valley degeneracy and m^* the effective mass), and the group velocity is

$$v(E) = v_T \sqrt{E/k_B T}, \quad (4)$$

where $v_T = \sqrt{k_B T/m^*}$ is a reference (thermal) velocity scale.[28] The spectral current $j(E) \equiv v(E)\mathcal{T}(E)$ acts as an energy filter: it weights high-energy carriers more heavily in proportion to their velocity.

We model the energy selectivity of the transmission by a smooth onset,

$$\mathcal{T}(E) = \left[1 + \exp\left(\frac{E_b - E}{\Delta}\right) \right]^{-1}, \quad (5)$$

a smooth step centered at an effective onset energy E_b with a broadening parameter Δ . We do *not* identify E_b with a unique microscopic barrier height. Rather, E_b is an effective crossover energy of the transport kernel that folds contact injection (e.g. a Schottky barrier), electrostatic bottlenecks (channel pinch-off, i.e. the saddle point of the channel potential), and energy-dependent scattering into a single energy-filtering scale. Carriers with $E < E_b$ contribute weakly to the drain current, while those with $E > E_b$ are efficiently transmitted [12]. Both E_b and Δ are thus effective transport-kernel parameters, to be calibrated from the equilibrium response $\bar{v}_{\text{eq}}(T)$ rather than identified with an independent microscopic barrier. Importantly, none of our central results depend on this specific functional form. They require only that the spectral current $j(E) = v(E)\mathcal{T}(E)$ be a smooth, monotonically increasing filter that favors high-energy carriers, so that a hot-carrier population at $E_0 > E_b$ is preferentially transmitted. Any transmission with these properties yields the same anomalous $g_m^{(\alpha)}$ peak, in the same spirit as the robustness against the choice of $\alpha(n)$ demonstrated in Appendix C.

B. Carrier distribution functions

We compare two classes of distribution, both normalized to the same areal carrier density n via $\int_0^\infty f(E) dE = n$. Here $f(E)$ is a carrier spectral density (carriers per unit area and energy), not a dimensionless occupation probability; accordingly it carries dimensions of $[n]/[\text{energy}]$, with the density of states D_0 absorbed into the current prefactor as noted below.

a. Equilibrium (Boltzmann).

$$f_{\text{eq}}(E) = A_{\text{eq}} e^{-E/k_B T}, \quad (6)$$

where $A_{\text{eq}} = n/N_{\text{eq}}$ and $N_{\text{eq}} = \int_0^\infty e^{-E/k_B T} dE = k_B T$. We adopt the Boltzmann (nondegenerate) form for analytical transparency. It is the leading approximation to the Fermi-Dirac distribution $f_{\text{FD}}(E) = \{1 + \exp[(E - E_F)/k_B T]\}^{-1}$ when the Fermi level lies below the transport states by more than a few $k_B T$, i.e. in the moderately doped, nondegenerate regime relevant to the density scale $n_c \sim 10^{12} \text{ cm}^{-2}$ considered here. For strongly degenerate channels the Boltzmann factor $e^{-E/k_B T}$ in Eqs. (6)–(7) should be replaced by $f_{\text{FD}}(E)$. This changes only the numerical values of the spectral moments N_{eq} and J_{eq} , leaving the spectroscopic identity Eq. (15) and the decomposition $g_m = g_m^{(n)} + g_m^{(\alpha)}$ unchanged in form. The hot-carrier component, centered well above E_F at $E_0 \gg k_B T$, is nondegenerate in either description.

b. Nonequilibrium (Boltzmann + hot-carrier Gaussian).

$$f_{\text{neq}}(E) = A_{\text{neq}} \left[e^{-E/k_B T} + \alpha e^{-(E-E_0)^2/\sigma^2} \right], \quad (7)$$

where $A_{\text{neq}} = n/(N_{\text{eq}} + \alpha N_{\text{neq}})$, $N_{\text{neq}} = \int_0^\infty e^{-(E-E_0)^2/\sigma^2} dE \approx \sigma\sqrt{\pi}$ (for $E_0 \gg \sigma$), and $\alpha \geq 0$ controls the amplitude of the Gaussian hot-carrier bump centered at energy $E_0 > E_b$ with width σ . The constraint $\int f(E) dE = n$ —where the constant D_0 is absorbed into the current prefactor, so that it cancels in $\bar{v}$ and in every g_m ratio below—is satisfied by construction via A_{neq} . The Gaussian is a minimal representation of a localized nonthermal excess at E_0 .

Physically, such a peaked nonthermal population can be established in several ways: high-field acceleration in the drain-side region of a short-channel FET followed by energy-selective optical-phonon or intervalley scattering; carrier injection over a contact or barrier bottleneck that admits a narrow energy window; impact-ionization-like generation near a characteristic energy; or optical pumping at a fixed photon energy. In each case carriers accumulate around a characteristic energy E_0 set by the dominant generation and relaxation bottleneck rather than by the lattice temperature. We do not commit to a specific microscopic mechanism; the Gaussian models only the resulting localized excess. The spectroscopic mechanism developed below relies only on the presence of such an excess well above the transport onset E_b . The same $g_m^{(\alpha)}$ anomaly therefore arises for any nonthermal excess $\delta f(E)$ that is sharply peaked in this energy range, irrespective of its detailed shape.

The physical meaning of α is the relative number of hot carriers in the Gaussian component: at $\alpha = 0$ the distribution is purely thermal, while increasing α builds up a high-energy tail at $E_0 \gg k_B T$.

C. Effective drift velocity and current decomposition

Substituting Eqs. (6)–(7) into Eq. (3) and using the linearity of the integral, we define spectral current moments

$$J_{\text{eq}} = \int_0^\infty j(E) e^{-E/k_B T} dE, \quad (8)$$

$$J_{\text{neq}} = \int_0^\infty j(E) e^{-(E-E_0)^2/\sigma^2} dE. \quad (9)$$

The drain current then takes the compact form

$$I_D = \frac{qW}{L} n \bar{v}(\alpha), \quad (10)$$

where the *effective drift velocity* is

$$\bar{v}(\alpha) = \frac{J_{\text{eq}} + \alpha J_{\text{neq}}}{N_{\text{eq}} + \alpha N_{\text{neq}}}. \quad (11)$$

Equivalently, $\bar{v}(\alpha) = \int_0^\infty j(E) f(E) dE / \int_0^\infty f(E) dE$ is simply the carrier-weighted average of the spectral current $j(E)$: the normalization A_{neq} and the density n cancel between numerator and denominator, leaving each moment split into a thermal and an α -weighted hot-carrier part. Equation (10) preserves the form of the conventional relation $I_D \propto n v_{\text{drift}}$, but now $\bar{v}(\alpha)$ is distribution-shape-sensitive: it interpolates between the pure-Boltzmann limit $\bar{v}(0) = J_{\text{eq}}/N_{\text{eq}} \equiv \bar{v}_{\text{eq}}$ and the Gaussian limit $\bar{v}(\infty) = J_{\text{neq}}/N_{\text{neq}} \equiv \bar{v}_{\text{neq}}$. Because hot carriers at $E_0 \gg E_b$ are above the barrier ($\mathcal{T}(E_0) \approx 1$) and have higher velocity ($v(E_0) \propto \sqrt{E_0}$), we have $\bar{v}_{\text{neq}} \gg \bar{v}_{\text{eq}}$, so even a small α significantly boosts I_D .

D. Gate-controlled carrier density and nonequilibrium parameter

Above threshold the carrier density follows the linear gate relation,

$$n(V_G) = \frac{C_{\text{ox}}}{q} (V_G - V_{\text{th}}), \quad (12)$$

where C_{ox} is the oxide capacitance per unit area and V_{th} is the threshold voltage. This is an idealized sharp threshold ($n \equiv 0$ below V_{th} , linear above), so dn/dV_G jumps discontinuously at V_{th} . Through the decomposition derived below, this produces two turn-on features in g_m exactly at V_{th} : a step in the density term $g_m^{(n)} \propto dn/dV_G$ (which jumps from zero to a constant), and—in the non-thermal case—a kink in the shape term $g_m^{(\alpha)} \propto n$ (whose slope breaks where n begins to rise from zero). Both are artifacts of the idealized threshold: in a real device, subthreshold conduction ($n \sim e^{qV_G/k_B T}$ below V_{th}) smooths them into a gradual turn-on. In any case, these features sit at V_{th} and are well separated from—and do not affect—the anomalous peak at V_G^{pk} .

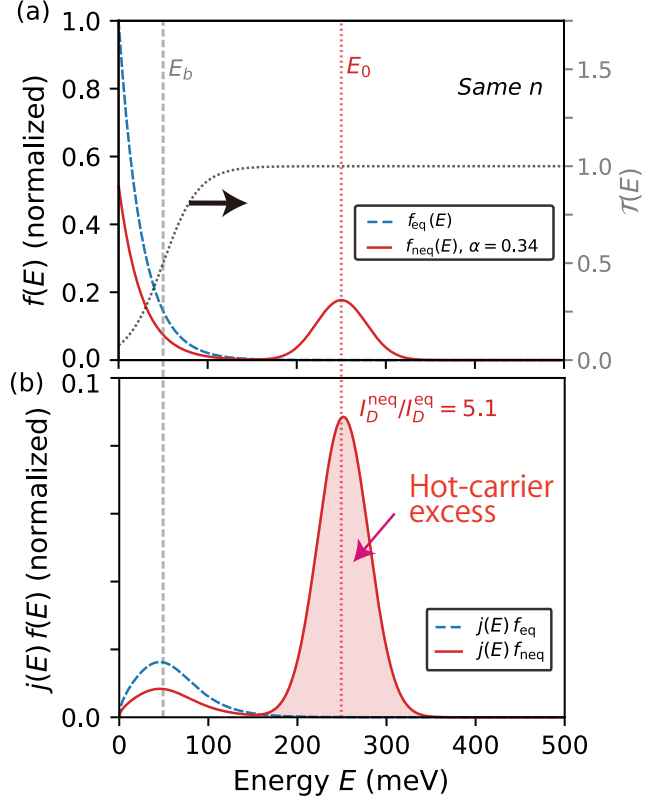


FIG. 2. (a) Carrier distribution functions $f(E)$ for the equilibrium (Boltzmann, blue dashed) and nonequilibrium (Boltzmann + Gaussian bump, red solid, $\alpha = 0.34$) cases, both normalized to the same carrier density n . The gray dotted curve on the right axis shows the transmission $\mathcal{T}(E)$; vertical dashed lines mark the onset energy E_b and the hot-carrier peak energy E_0 . (b) Spectral current integrand $j(E)f(E)$ for both distributions (shared energy axis with (a)). The shaded region is the hot-carrier excess contribution to the current. Despite identical carrier densities, the nonequilibrium current is enhanced by $I_D^{\text{neq}}/I_D^{\text{eq}} \approx 5.1$ at this density, reflecting the larger per-carrier current of the hot-carrier excess near E_0 ($\bar{v}_{\text{neq}}/\bar{v}_{\text{eq}} \approx 9.4$ for the pure Gaussian component).

The hot-carrier amplitude α depends on both V_G and V_D . The key qualitative property required for an anomalous peak in $g_m^{(\alpha)}$ is that $\alpha(n)$ rises from zero at small n and saturates as $n \rightarrow \infty$. A broad class of such monotone-and-saturating functions generically produces a localized peak in the distribution-shape contribution $g_m^{(\alpha)}$ derived below, because $d\alpha/dn$ inherits the bell shape of the derivative of a saturating function. The qualitative peak position, height, and V_D -dependence are robust fingerprints of this mechanism, only weakly dependent on the precise functional form of $\alpha(n)$. Alternative analytic forms (hyperbolic-tangent, exponential saturation, and a Hill function in n) yield quantitatively similar $g_m^{(\alpha)}(V_G)$ peaks, differing only in the prefactor of the peak density, as verified explicitly in Appendix C (Fig. 8).

For analytical tractability we adopt the representative

phenomenological model

$$\alpha(V_G, V_D) = \alpha_{\max} \frac{V_D^2}{V_D^2 + V_c^2} \frac{n(V_G)}{n(V_G) + n_c}, \quad (13)$$

which is proportional to the power density (through V_D) and to the density of impact-ionizing or inelastically scattered carriers (through n). Here $\alpha_{\max}$ is the saturation amplitude, V_c is the drain-voltage crossover, and n_c is the carrier-density crossover for hot-carrier generation. The high-field crossover at V_c parallels the onset of velocity saturation observed in monolayer MoS₂ [29, 30]. In real devices n_c is not a universal material constant but an effective crossover density set by the channel length, contact field concentration [31], dielectric environment, heat dissipation, and the dominant inelastic-scattering pathway [32]. The value adopted below should therefore be viewed as a representative effective scale for strongly gated TMD FETs. For $n \ll n_c$, $\alpha \propto n$ (few carriers, few collisions); for $n \gg n_c$, $\alpha \rightarrow \alpha_{\max} h(V_D)$ (saturation).

We stress that Eq. (13) is a transparent phenomenological interpolation, not a microscopic derivation, and that the two factors are chosen for their limiting behavior rather than for a specific functional form. The drain-voltage factor $h(V_D) = V_D^2/(V_D^2 + V_c^2)$ represents field-driven hot-carrier generation that switches on with the power input at low bias and saturates once the high-field region spans the channel. Any monotonic function rising from zero and saturating at large V_D yields the same qualitative $g_m^{(\alpha)}(V_D)$. The density factor $n/(n + n_c)$ expresses that at low density the hot-carrier population grows in proportion to the number of available carriers (and hence collisions), whereas at high density it saturates as phase-space restriction, screening, and energy-relaxation bottlenecks limit further accumulation. What matters for the anomalous peak is only that $\alpha(n)$ rise from zero and saturate, so that $d\alpha/dn$ is bell-shaped. We verify explicitly in Appendix C (Fig. 8) that hyperbolic-tangent, exponential, and Hill forms all produce the same peak, with only minor quantitative differences. A first-principles kinetic treatment of $\alpha(n, V_D)$ from the coupled hot-carrier generation–relaxation balance is beyond the scope of the present work and is left for future study.

E. Transconductance as a spectral probe

Before decomposing g_m , we establish its fundamental spectroscopic content. Differentiating Eq. (3) with respect to V_G gives, by the product rule, two contributions,

$$g_m = \frac{qW}{L} \int_0^\infty \left[\frac{\partial j(E; V_G)}{\partial V_G} f(E; V_G) + j(E; V_G) \frac{\partial f(E; V_G)}{\partial V_G} \right] dE, \quad (14)$$

a *kernel-modulation* term and a *distribution-modulation* term. Within the present transport description the spectral current kernel $j(E) = v(E)\mathcal{T}(E)$ [Eqs. (4)–(5)] has no explicit gate dependence, so $\partial j/\partial V_G = 0$, the first

term vanishes, and we obtain

$$g_m = \frac{qW}{L} \int_0^\infty j(E) \frac{\partial f(E; V_G)}{\partial V_G} dE. \quad (15)$$

The kernel-modulation term is nonzero when the gate lowers the barrier or modulates the contacts [$\mathcal{T} = \mathcal{T}(E, V_G)$], induces band bending, or otherwise renders j gate-dependent. This term is additive and is estimated in Sec. IV. Equation (15) shows that the distribution-function contribution to g_m is exactly the spectral current $j(E)$ weighted by $\partial f/\partial V_G$: a *spectrally resolved* response function of the distribution. For f_{eq} , the derivative $\partial f_{\text{eq}}/\partial V_G \propto e^{-E/k_B T}$ is featureless, and g_m depends only on the total carrier density. For f_{neq} , $\partial f_{\text{neq}}/\partial V_G$ acquires additional structure at E_0 whenever α depends on V_G (i.e. on n), and g_m directly reports this spectral feature—even though no optical measurement is performed. This is the central spectroscopic principle of the paper.

Decomposing $g_m = dI_D/dV_G$ explicitly via Eq. (10) with respect to V_G :

$$g_m = \underbrace{\frac{WC_{\text{ox}}}{L} \bar{v}(\alpha)}_{g_m^{(n)}} + \underbrace{\frac{WC_{\text{ox}}}{L} n \frac{d\bar{v}}{d\alpha} \frac{d\alpha}{dn}}_{g_m^{(\alpha)}}, \quad (16)$$

where we used $dn/dV_G = C_{\text{ox}}/q$ from Eq. (12). The two terms have distinct physical origins:

- $g_m^{(n)} = (WC_{\text{ox}}/L) \bar{v}(\alpha)$: the conventional density-modulation term. At equilibrium ($\alpha = 0$), this equals the constant value $(WC_{\text{ox}}/L) \bar{v}_{\text{eq}}$, giving a flat g_m – V_G curve in the linear regime.
- $g_m^{(\alpha)}$: a new distribution-shape-modulation term that arises because the gate voltage also controls α through $n(V_G)$. Its magnitude is proportional to $d\bar{v}/d\alpha = N_{\text{eq}} N_{\text{neq}} (\bar{v}_{\text{neq}} - \bar{v}_{\text{eq}})/(N_{\text{eq}} + \alpha N_{\text{neq}})^2$, which is positive whenever $\bar{v}_{\text{neq}} > \bar{v}_{\text{eq}}$.

The gate-voltage dependence of $g_m^{(\alpha)}$ is governed by

$$\frac{d\alpha}{dn} = \frac{\alpha_{\max} h(V_D) n_c}{(n + n_c)^2}, \quad (17)$$

where $h(V_D) = V_D^2/(V_D^2 + V_c^2)$. To find the peak of $g_m^{(\alpha)}$ analytically, we substitute the $\alpha(n)$ dependence into the expression for $d\bar{v}/d\alpha$ and simplify. Defining $\beta = \alpha_{\max} h(V_D)$, one finds

$$\begin{aligned} g_m^{(\alpha)} &\propto \frac{n}{(N_{\text{eq}} + \beta N_{\text{neq}} \frac{n}{n + n_c})^2 (n + n_c)^2} \\ &= \frac{n}{(A n + B)^2}, \end{aligned} \quad (18)$$

where $A = N_{\text{eq}} + \beta N_{\text{neq}}$ and $B = N_{\text{eq}} n_c$. Setting $d/dn[n/(A n + B)^2] = 0$ gives the exact peak density

$$n_{\text{pk}} = \frac{B}{A} = \frac{N_{\text{eq}} n_c}{N_{\text{eq}} + \alpha_{\max} h(V_D) N_{\text{neq}}}, \quad (19)$$

and the corresponding peak gate voltage

$$V_G^{\text{pk}} = V_{\text{th}} + \frac{q n_{\text{pk}}}{C_{\text{ox}}}. \quad (20)$$

In the limit $\alpha_{\text{max}} \rightarrow 0$, Eq. (19) gives $n_{\text{pk}} \rightarrow n_c$ and $V_G^{\text{pk}} \rightarrow V_{\text{th}} + q n_c / C_{\text{ox}} \equiv V_G^*$. For finite α_{max} , the peak shifts to lower gate voltages by the factor $N_{\text{eq}} / (N_{\text{eq}} + \alpha_{\text{max}} h(V_D) N_{\text{neq}})$. Since $N_{\text{eq}} = k_B T$ and $N_{\text{neq}} \approx \sigma \sqrt{\pi}$, this shift encodes the hot-carrier energy width σ . Consequently, $g_m^{(\alpha)}$ exhibits a peak at V_G^{pk} —an anomalous feature absent in equilibrium transport—and its position jointly reports n_c (the density crossover) and σ (the hot-carrier spectral width).

III. RESULTS

We evaluate Eqs. (10)–(16) numerically for parameters representative of a monolayer MoS₂-like 2D semiconductor ($m^* = 0.5 m_0$, spin–valley degeneracy $g = 4$, $T = 300$ K). The gate dielectric is taken as 30 nm HfO₂ ($\varepsilon_r \approx 25$), giving $C_{\text{ox}}/q = 1.4 \times 10^{13} \text{ cm}^{-2} \text{V}^{-1}$. The transport-kernel onset parameters are $E_b = 50$ meV and $\Delta = 20$ meV. Hot carriers are characterized by $E_0 = 250$ meV, $\sigma = 40$ meV, $\alpha_{\text{max}} = 0.80$, $n_c = 2.0 \times 10^{12} \text{ cm}^{-2}$, and $V_c = 0.30$ V. Without loss of generality we set $V_{\text{th}} = 0$, so that all gate voltages are measured relative to threshold. A finite V_{th} merely translates the V_G axis and leaves every result unchanged.

A. Carrier distributions at fixed density

Figure 2(a) compares $f_{\text{eq}}(E)$ and $f_{\text{neq}}(E)$ at the same carrier density n . The equilibrium distribution decays exponentially with thermal energy $k_B T = 26$ meV, while the nonequilibrium distribution is depressed at low energies (due to normalization) but acquires a prominent Gaussian bump at $E_0 = 250$ meV.

Figure 2(b) shows the spectral current integrand $j(E)f(E)$, which determines the contribution to I_D at each energy. The equilibrium integrand is concentrated at low energies where most carriers reside, but the barrier $\mathcal{T}(E)$ strongly suppresses the contribution below E_b . The nonequilibrium integrand acquires a large additional peak at E_0 , because those hot carriers have both higher velocity $v(E_0) = v_T \sqrt{E_0/k_B T}$ and unit transmission $\mathcal{T}(E_0) \approx 1$. The ratio of integrated spectral currents is $\bar{v}_{\text{neq}}/\bar{v}_{\text{eq}} \approx 9.4$, meaning each hot carrier contributes $9.4 \times$ more current than a thermal carrier at $T = 300$ K. This value 9.4 quantifies the per-carrier current contribution in the *pure* Gaussian component (i.e., $\bar{v}_{\text{neq}}/\bar{v}_{\text{eq}}$), whereas the actual mixed-distribution current enhancement at $\alpha = \alpha(n_0, V_D = 0.5 \text{ V})$, weighted by the relative populations of the two components, is the smaller value $I_D^{\text{neq}}/I_D^{\text{eq}} \approx 5.1$ displayed in Fig. 2(b).

B. Steady-state transport prediction

Figure 3(a) shows $I_D(V_G)$ at $V_D = 0.5$ V. Both curves increase linearly with V_G , as expected from $I_D \propto n$, but the nonequilibrium current is enhanced by a factor that grows with V_G up to $n \sim n_c$ and then saturates. This saturation is traced in Fig. 3(b): as $n \gg n_c$, $\alpha(n)$ saturates at $\alpha_{\text{max}} h(V_D)$, and $\bar{v}(\alpha)$ approaches a constant. The vertical line at V_G^* marks the crossover density $n = n_c$. At this gate voltage, the current enhancement is maximal in a differential sense—the nonequilibrium I_D increases faster with V_G than the equilibrium baseline, a hallmark of the $g_m^{(\alpha)}$ contribution.

The central experimental prediction is shown in Fig. 3(c), which plots $g_m(V_G)$ together with its decomposition. The equilibrium transconductance is constant at $(WC_{\text{ox}}/L) \bar{v}_{\text{eq}}$ above threshold, as expected from conventional theory. For the representative parameters used here, the total nonequilibrium g_m itself rises monotonically with V_G and saturates at a higher plateau $((WC_{\text{ox}}/L) \bar{v}(\alpha_{\text{max}} h(V_D)))$ than the equilibrium value, because the density term $g_m^{(n)}$ is monotone by construction and dominates the magnitude. The non-thermal signature is therefore carried not by the total g_m but by its *shape-driven contribution* $g_m^{(\alpha)} = g_m - g_m^{(n)}$, which exhibits a pronounced *anomalous peak* at V_G^{pk} [Eq. (20)] and vanishes in any density-only description (where $g_m^{(\alpha)} \equiv 0$). Because the density background $g_m^{(n)} = (WC_{\text{ox}}/L) \bar{v}(\alpha)$ is smooth and monotone, this localized peak appears as a shoulder on $g_m(V_G)$ and is isolated by subtracting the calibrated monotone background. It is this shape term—the gate-voltage derivative of the distribution shape—that directly encodes spectral information beyond n . This is the key signature of transconductance spectroscopy. With the present parameters ($T = 300$ K, $\alpha_{\text{max}} = 0.80$, $\sigma = 40$ meV, $V_D = 0.5$ V), Eq. (19) gives $n_{\text{pk}}/n_c \approx 0.38$, so the peak appears at $V_G^{\text{pk}} \approx 55$ mV, well below the reference scale $V_G^* = q n_c / C_{\text{ox}} \approx 143$ mV.

The same panel decomposes g_m into its two components. The density-driven term $g_m^{(n)}$ is a smoothly increasing function that inherits the V_G dependence of $\bar{v}(\alpha)$, which rises monotonically as α increases. The shape-driven term $g_m^{(\alpha)}$ is localized near V_G^* and vanishes both below threshold ($n \rightarrow 0$) and for $n \gg n_c$ ($d\alpha/dn \rightarrow 0$). The anomalous peak thus resides entirely in $g_m^{(\alpha)}$, while the total g_m remains monotone—underscoring that the signature is a feature of the distribution-shape contribution, isolated from the smooth density background, rather than a maximum of the raw transconductance.

Figure 3(d) shows the peak height of $g_m^{(\alpha)}$ as a function of V_D . It follows $h(V_D) = V_D^2 / (V_D^2 + V_c^2)$ (gray dashed), confirming Eq. (17). For $V_D \lesssim V_c$, $g_m^{(\alpha)}$ grows quadratically with V_D , while for $V_D \gg V_c$ it saturates. This V_D -tunability provides an additional experimental signature: measuring $g_m(V_G)$ at several V_D values and fitting the peak height to $h(V_D)$ extracts V_c , which encodes the

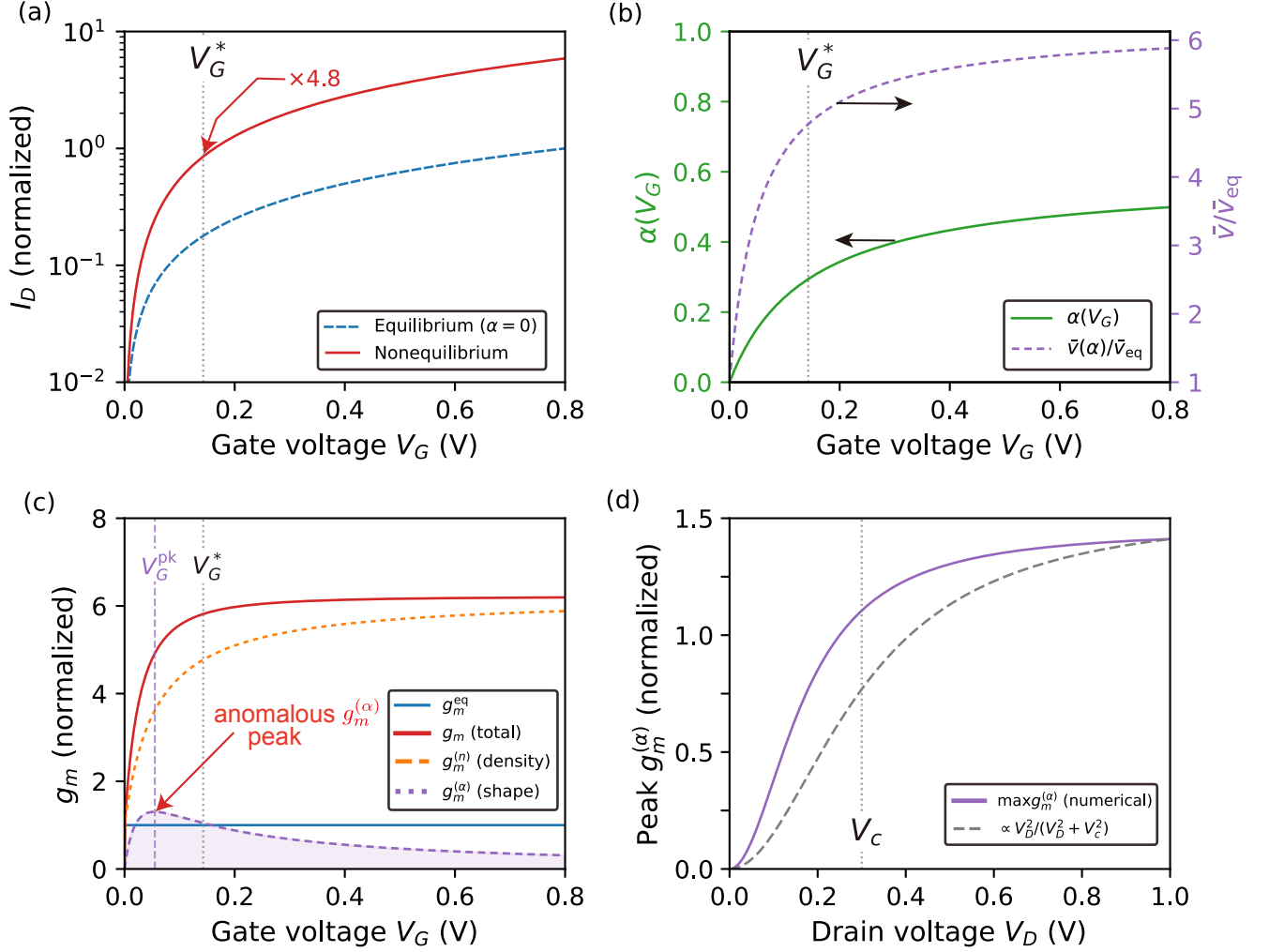


FIG. 3. Main steady-state prediction, combining the current-voltage and transconductance characteristics. (a) Drain current I_D vs gate voltage V_G at $V_D = 0.5$ V (logarithmic scale), normalized to the equilibrium saturation current. The nonequilibrium current (red solid) exceeds the equilibrium baseline (blue dashed) due to the hot-carrier velocity enhancement $\bar{v}(\alpha)$; the vertical dotted line marks V_G^* where $n = n_c$. (b) Left axis: nonequilibrium amplitude $\alpha(V_G)$, showing saturation for $V_G \gg V_G^*$. Right axis: effective velocity ratio $\bar{v}(\alpha)/\bar{v}_{eq}$, which reaches its maximum as $\alpha \rightarrow \alpha_{max} h(V_D)$. (c) Transconductance $g_m(V_G)$ for the equilibrium (g_m^{eq} , blue solid) and nonequilibrium (total, red solid) cases, decomposed into the density-driven term $g_m^{(n)}$ (orange dashed) and the distribution-shape-driven term $g_m^{(\alpha)}$ (purple dotted, filled). The equilibrium g_m is flat (conventional) and the total nonequilibrium g_m rises monotonically; the anomalous peak near V_G^{pk} resides in the shape contribution $g_m^{(\alpha)}$ (purple), isolated from the smooth monotone density background $g_m^{(n)}$. (d) Peak height of $g_m^{(\alpha)}$ vs drain voltage V_D , following $h(V_D) = V_D^2/(V_D^2 + V_c^2)$ (gray dashed); the vertical dotted line marks $V_c = 0.3$ V.

electric-field scale for hot-carrier generation.

C. Discrimination from quasi-equilibrium models

The anomalous peak in the shape contribution $g_m^{(\alpha)}$ is the central signature of the non-thermal mechanism, but it is not, by itself, unique to it. Two quasi-equilibrium descriptions must be excluded before a peak extracted from $g_m(V_G)$ can be attributed to the distribution shape: a gate-dependent field-effect mobility $\mu(V_G)$,

and an effective-temperature (heated-Boltzmann) ensemble. Both fail, and the way they fail provides the decisive, falsifiable signature of the mechanism.

a. Exclusion of a mobility-rolloff peak. A gate-dependent mobility $\mu(V_G)$ —from surface-roughness or phonon scattering that strengthens with the vertical field [32]—also produces a g_m peak. The two are separated by their drain-voltage dependence. In the linear regime any quasi-equilibrium mobility description gives

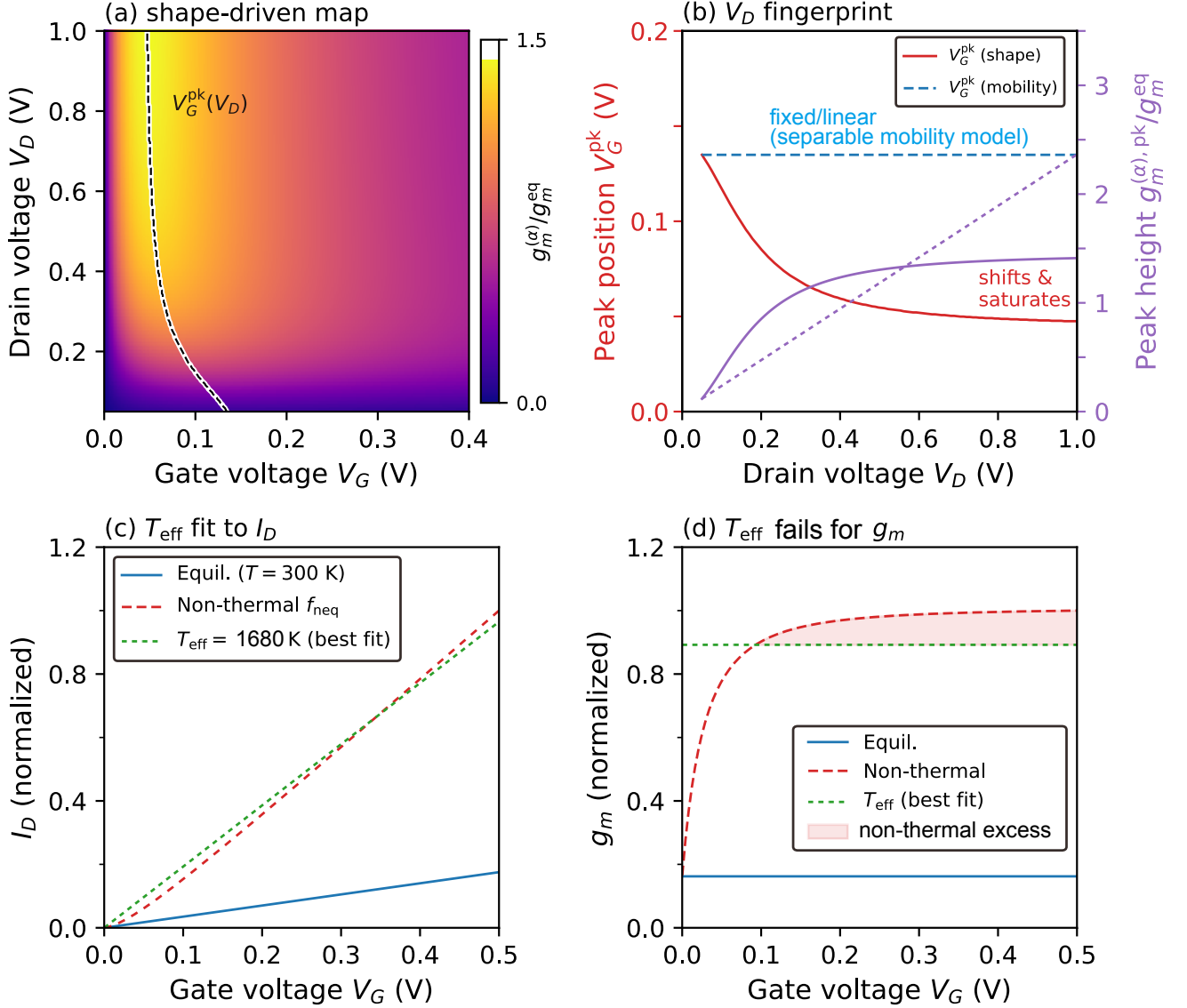


FIG. 4. Discrimination of the non-thermal distribution-shape peak from quasi-equilibrium descriptions. *Top row—exclusion of a mobility-rolloff peak.* (a) Density map of the shape-driven transconductance $g_m^{(\alpha)}(V_G, V_D)$, normalized to the equilibrium plateau g_m^{eq} ; the overlaid curve is the peak trajectory $V_G^{\text{pk}}(V_D)$, which shifts to lower V_G (≈ 70 mV from $V_D = 0.1$ to 1.0 V) while the ridge intensity saturates at large V_D . (b) The two fingerprints versus V_D : the peak position V_G^{pk} (red, left axis) shifts and saturates, and the peak height (purple, right axis) saturates as $h(V_D)$; a separable mobility model [Eq. (21)] instead predicts a V_D -independent position (blue dashed) and a height linear in V_D (purple dotted). *Bottom row—exclusion of an effective-temperature model ($V_D = 0.5$ V).* (c) The best-fit heated-Boltzmann $T_{\text{eff}} \approx 1680$ K (green dash-dot) reproduces $I_D(V_G)$ reasonably well, but only by invoking an electronic temperature far above any quasi-equilibrium heating regime. (d) The same T_{eff} model gives a flat $g_m(V_G)$, whereas the non-thermal g_m rises with V_G ; the shaded region is the non-thermal excess over the flat T_{eff} prediction, unaccounted for by any effective-temperature model.

$I_D = (qW/L) n \mu(V_G) V_D$, so

$$g_m(V_G, V_D) = V_D \frac{qW}{L} \frac{d[n \mu(V_G)]}{dV_G} \equiv V_D G(V_G), \quad (21)$$

which is *separable* in V_G and V_D : the peak position is V_D -independent and the peak height scales linearly with V_D . The shape mechanism violates both. By

Eqs. (19)–(20) the peak position shifts to lower gate voltage as V_D grows, while the peak height saturates as $h(V_D) = V_D^2/(V_D^2 + V_c^2)$ [Fig. 3(d)]. For the present parameters the position shifts by $\Delta V_G^{\text{pk}} \approx 70$ mV between $V_D = 0.1$ and 1.0 V [Fig. 4(a)], and the height saturates above $V_D \sim V_c$ [Fig. 4(b)]. The *simultaneous* position shift and height saturation are not reproduced by

any separable, quasi-equilibrium mobility-rolloff model [Eq. (21)]. We note that a contrived contact or mobility response with strong, nonseparable V_D dependence could in principle mimic one of the two signatures, so the discriminating power rests on their joint appearance rather than on either alone.

b. Exclusion of an effective-temperature model. Previous treatments of hot carriers in FETs typically introduce an effective electron temperature $T_{\text{eff}} > T$ and replace μ by $\mu(T_{\text{eff}})$ [11, 12]. Such models assume that the heated distribution remains a displaced Maxwellian, and therefore cannot capture the Gaussian hot-carrier component in Eq. (7), which represents a localized non-thermal excess generated, e.g., by impact ionization, optical pumping at a specific photon energy, or injection over a discrete barrier. We find the optimal T_{eff} by minimizing the least-squares residual in $I_D(V_G)$. For the present MoS₂-like parameters at $V_D = 0.5$ V this gives $T_{\text{eff}} \approx 1680$ K, far beyond the regime in which a simple quasi-equilibrium heating picture is physically transparent. Yet even at this best-fit T_{eff} , which reproduces $I_D(V_G)$ reasonably well [Fig. 4(c)], the model fails for g_m [Fig. 4(d)]: because $g_m^{T_{\text{eff}}} = (WC_{\text{ox}}/L)\bar{v}(T_{\text{eff}})$ is a constant in V_G , the model predicts a flat transconductance, whereas the non-thermal g_m rises with V_G and carries the localized shape contribution $g_m^{(\alpha)}$ that a heated Boltzmann distribution cannot generate. The shaded region in Fig. 4(d) is the resulting non-thermal excess of g_m over the flat T_{eff} prediction—a direct fingerprint of the non-thermal character of f_{neq} that no effective-temperature description can reproduce.

c. Feasibility. The anomaly is large and the discriminator is undemanding. The shape-driven peak reaches $g_m^{(\alpha),\text{pk}} \sim (0.4\text{--}1.4)g_m^{\text{eq}}$ for $V_D = 0.1\text{--}1.0$ V—an order-unity modulation of an already routinely measured quantity, far above the $\lesssim 1\%$ resolution of standard DC/lock-in transconductance. The diagnostic position shift of several tens of mV is resolved by sub-mV gate steps, and because $E_0 \gg k_B T$ the signature survives to room temperature. The decisive measurement is thus a set of $g_m(V_G)$ sweeps at several V_D : a peak that shifts and saturates identifies the shape mechanism, whereas a fixed peak with V_D -linear height identifies conventional mobility rolloff.

D. Two-time-scale separation

Beyond the steady state, the framework predicts a specific transient response to pulsed excitation. In the simplest case a pump injects hot carriers at $t = 0$ and only the hot-carrier amplitude relaxes, $\alpha(t) = \alpha_0 e^{-t/\tau_E}$, while the carrier density stays fixed. This single-time-scale pump-probe limit, which yields the energy relaxation time τ_E directly from the decay of $g_m(t)$, is treated in Appendix A. In a more general scenario—such as a pulsed gate voltage or optical pump that transiently injects both extra carriers and hot-carrier excitations—two

distinct transients appear simultaneously:

$$I_D(t) = [n_0 + \Delta n e^{-t/\tau_n}] \bar{v}(\alpha_0 e^{-t/\tau_E}), \quad (22)$$

where τ_n is the carrier-density relaxation time (set by the device RC time or dielectric relaxation) and τ_E is the energy relaxation time (set by carrier-phonon scattering). For $\tau_n \ll \tau_E$, these two time scales are separated and can be independently extracted from a single $I_D(t)$ transient.

Figure 5(a) shows $I_D(t)/I_D^{\text{eq}}$ on a linear scale for $\tau_n = 0.2$ ps and $\tau_E = 1.0$ ps. Two regimes are apparent: a fast initial decay (duration $\sim \tau_n$) where the excess density Δn equilibrates, followed by a pronounced intermediate plateau where the density has returned to n_0 but α is still large. The current then decays slowly from the plateau to the equilibrium baseline over the time scale τ_E . The height of the plateau encodes the amplitude α_0 through $I_D^{\text{plateau}} = n_0 \bar{v}(\alpha_0)$, providing an independent determination of the initial hot-carrier injection level.

Figure 5(b) presents the same transient on a semi-logarithmic scale, where the two exponential components appear as two linear segments with distinct slopes $1/\tau_n$ and $1/\tau_E$. This is the standard fingerprint of a bi-exponential decay, familiar from optical pump-probe studies [15, 33], now accessible by all-electrical measurements. Fitting the long-time slope directly gives τ_E without contamination from the fast density dynamics, provided measurements are taken at $t \gtrsim 3\tau_n$.

Figure 5(c) shows the normalized anomaly $\Delta I_D(t)/\Delta I_D(0)$ for a range of separation ratios $\tau_n/\tau_E = 0.05\text{--}0.50$. In all cases the long-time tail converges to a common e^{-t/τ_E} behavior (gray dotted reference), confirming that τ_E can be extracted cleanly once $\tau_n/\tau_E \lesssim 0.5$. In optimized high-bandwidth FET geometries the device-limited density relaxation can in principle be made substantially shorter than the carrier-phonon energy relaxation time $\tau_E \sim 1\text{--}5$ ps expected for TMDs, placing this system within the resolvable regime. The specific bandwidth requirements are discussed in Sec. IV.

IV. DISCUSSION

a. Why transconductance spectroscopy is distinct from previous hot-carrier work. One might ask whether the present results are simply a restatement of known hot-carrier physics in different notation. They are not, for the following reasons. First, Eq. (15) holds for *any* $f(E)$, not just the Boltzmann-plus-Gaussian model used here, provided the spectral kernel $j(E)$ is gate-independent (corrections from a V_G -dependent kernel are addressed below). This means that $g_m(V_G)$ curves measured in any device, under any excitation, are spectral transforms of the underlying distribution. Second, the anomalous peak in $g_m^{(\alpha)}$ cannot be reproduced by any density-only or quasi-equilibrium description without invoking an ad hoc nonmonotonic V_G -dependence in $\mu(n)$

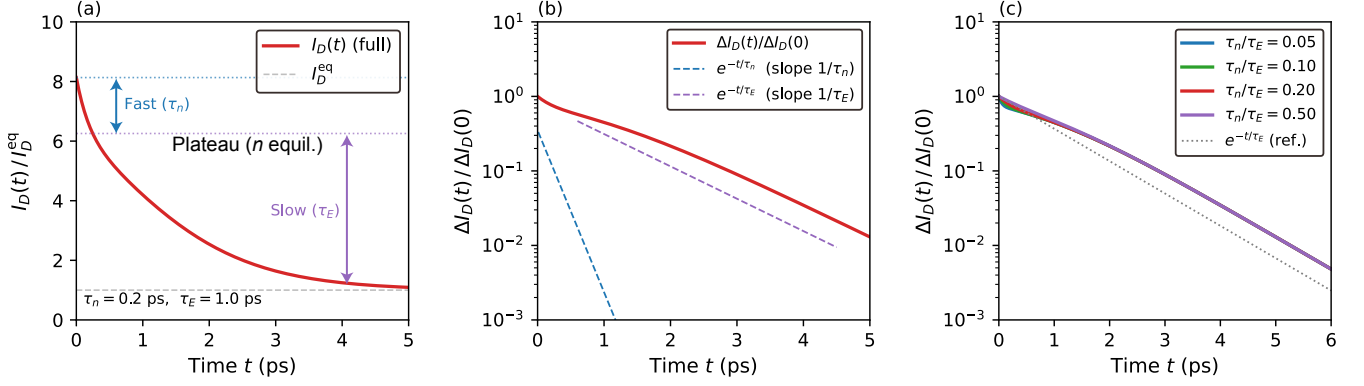


FIG. 5. Two-time-scale separation in the transient drain current after a pulsed excitation ($\Delta n/n_0 = 0.30$, $\alpha_0 = 0.60$). (a) $I_D(t)/I_D^{\text{eq}}$ on a linear scale ($\tau_n = 0.2$ ps, $\tau_E = 1.0$ ps). The fast density decay (τ_n) and slow spectral decay (τ_E) are separated by an intermediate plateau where n has equilibrated but hot carriers persist. (b) $\Delta I_D(t)/\Delta I_D(0)$ on a semi-logarithmic scale, revealing two linear segments with slopes $1/\tau_n$ (fast, blue dashed) and $1/\tau_E$ (slow, purple dashed). (c) Normalized anomaly for $\tau_n/\tau_E = 0.05$ – 0.50 . All curves converge to the common e^{-t/τ_E} behavior (gray dotted) at long times, demonstrating clean extraction of τ_E when $\tau_n/\tau_E \lesssim 0.5$.

or $T_{\text{eff}}(n)$. Indeed, a monotone $\mu(n)$ or monotone $T_{\text{eff}}(n)$ necessarily yields a monotone $g_m \propto \bar{v}(\alpha(n)) (C_{\text{ox}}/q)$ in the linear regime. By contrast, the distribution-shape mechanism predicts the peak position, height, and V_D -dependence simultaneously from a single unified parameter set $\{E_0, \sigma, n_c\}$, providing a tightly constrained, falsifiable prediction. Third, the measurement is all-electrical: it requires only a standard lock-in g_m sweep and is compatible with any gated 2D device, with no need for pump lasers, time-resolved optics, or sub-kelvin cooling.

b. Robustness against gate-dependent kernels. The spectroscopic identity Eq. (15) was derived assuming a gate-independent spectral kernel $j(E) = v(E)\mathcal{T}(E)$. In real devices, V_G can also modulate E_b (barrier lowering), the contact transmission [31, 34, 35], or the band bending, giving an additional kernel-modulation contribution

$$g_m^{(j)} = \frac{qW}{L} \int_0^\infty f(E, V_G) \frac{\partial j(E, V_G)}{\partial V_G} dE, \quad (23)$$

so that in general $g_m = g_m^{(n)} + g_m^{(\alpha)} + g_m^{(j)}$. For the Fermi barrier of Eq. (5), $\partial j/\partial V_G = -v(E)(\partial \mathcal{T}/\partial E)(\partial E_b/\partial V_G)$. Since $\partial \mathcal{T}/\partial E$ is sharply peaked at $E = E_b$ (width Δ) with unit area, the kernel term reduces to

$$g_m^{(j)} \simeq -\frac{qW}{L} \frac{\partial E_b}{\partial V_G} v(E_b) f(E_b; V_G). \quad (24)$$

Two consequences follow. First, $g_m^{(j)}$ samples the distribution at the barrier edge E_b , not at the hot-carrier energy $E_0 \gg E_b$. Because the bump contributes $\propto e^{-(E_b - E_0)^2/\sigma^2} \approx e^{-25}$ there, $g_m^{(j)}$ carries *no* peak structure from f_{neq} and adds only a smooth background to $g_m(V_G)$ —it cannot reproduce the anomalous peak. Second, its magnitude is set by the barrier-lowering rate $\partial E_b/\partial V_G$, an electrostatic/contact property that for a

well-gated channel in the on-state is small ($\partial E_b/\partial V_G \ll 1$). The distinguishing signature of $g_m^{(\alpha)}$ from $g_m^{(j)}$ is their qualitatively different V_D dependence: $g_m^{(\alpha)}$ scales as $h(V_D) = V_D^2/(V_D^2 + V_c^2)$ [Fig. 3(d)] and saturates at large V_D , while $g_m^{(j)}$ contributions (barrier-lowering, contact modulation) at fixed V_G generally scale linearly or only weakly in V_D . Measuring the V_D -dependence of the anomalous peak thus provides an in-situ separation of distribution-shape and kernel-modulation mechanisms. In practice, since contact and barrier-modulation effects can themselves exhibit nonlinear V_D behavior, the most robust criterion is not a single V_D -dependence alone but the *simultaneous* consistency of peak position, peak height, and the joint fit to I_D and g_m (Appendix B, Fig. 7)—the same over-constrained signature that excludes the T_{eff} description.

c. Other device-related mechanisms and the over-constrained criterion. We emphasize that the appearance of a single peak in $g_m(V_G)$ is *not*, by itself, claimed as a unique fingerprint of hot carriers. Contact resistance and its bias dependence, interface traps, self-heating, field- or density-dependent mobility degradation, and gate-induced barrier modulation can each produce non-monotonic features in the raw transconductance. What discriminates the distribution-shape mechanism is not the peak alone but the *joint, over-constrained* behavior of several observables that these parasitic effects do not reproduce together: (i) the peak position V_G^{pk} shifts to lower gate voltage with increasing V_D [Eqs. (19)–(20)], whereas a separable mobility or trap response gives a V_D -independent position; (ii) the peak height saturates as $h(V_D)$ rather than growing linearly in V_D ; (iii) a single spectral parameter set $\{E_0, \sigma, \alpha_{\text{max}}, n_c\}$ must reproduce $I_D(V_G)$ and $g_m(V_G)$ simultaneously (Appendix B); and (iv) the low-bias limit $V_D \ll V_c$ must recover the flat, cal-

ibrated equilibrium transconductance with $V_G^{\text{pk}} \rightarrow V_G^*$. Contact, trap, and self-heating artifacts are, moreover, suppressible or separable by standard controls—four-probe or transfer-length contact de-embedding, gate-bias- and frequency-dependent trap characterization, and pulsed (low-duty-cycle) measurements that eliminate steady-state self-heating. The mechanism should therefore be regarded as identified by this diagnostic protocol as a whole, not by any single peak.

d. Role of inelastic scattering and the transmission kernel. The energy-selective transmission $\mathcal{T}(E)$ in Eq. (3) is an effective, elastic-like energy filter calibrated from the equilibrium response; it does not resolve individual inelastic collisions. Inelastic scattering enters our description not through $\mathcal{T}(E)$ but through the distribution itself—its steady-state hot-carrier amplitude α and, in the transient case, its relaxation law $\alpha(t) = \alpha_0 e^{-t/\tau_E}$. The steady-state spectroscopy is therefore independent of the microscopic collision integral: it reads out whatever nonthermal $f(E)$ is present, however generated. The time-domain extension (Sec. III; Appendix A) treats energy relaxation as a slow evolution of this distribution parameter relative to the instantaneous transport response, which is justified when the transit/RC time is short compared with τ_E . It is not a microscopic transport theory of energy-nonconserving events during a single traversal, and it would break down if inelastic mean free paths became comparable to the channel length, so that $\mathcal{T}$ itself acquired a strong, energy-redistributing bias dependence. That regime lies beyond the present kernel approximation.

e. Over-constrained simultaneous fit. Beyond excluding the effective-temperature description (Sec. III C), the non-thermal model provides an over-constrained demonstration: a *single* parameter set $\{E_0, \sigma, \alpha_{\text{max}}, n_c\}$ simultaneously reproduces both $g_m(V_G)$ and $I_D(V_G)$, whereas the T_{eff} model adds one free parameter yet already fails at the first observable (g_m). A quantitative simultaneous fit to synthetic data, confirming that the spectral shape of $f(E)$ —not its integrated density or effective temperature—governs transport, is given in Appendix B (Fig. 7).

f. Spectroscopic information from the g_m anomaly. Equations (19)–(20) show that the peak position V_G^{pk} constrains two quantities: (i) n_c , extracted from V_G^{pk} in the weak hot-carrier limit ($V_D \ll V_c$) where $n_{\text{pk}} \rightarrow n_c$; and (ii) the Gaussian width σ (via $N_{\text{neq}} \approx \sigma\sqrt{\pi}$), extracted from the shift $\Delta V_G^{\text{pk}}(V_D) = V_G^* - V_G^{\text{pk}}(V_D)$ as a function of V_D . Independently, the peak height $g_m^{(\alpha), \text{pk}} \propto J_{\text{neq}} N_{\text{eq}} - J_{\text{eq}} N_{\text{neq}} \propto \bar{v}_{\text{neq}} - \bar{v}_{\text{eq}}$ is sensitive to E_0 through $v(E_0) = v_T \sqrt{E_0/k_B T}$, providing the hot-carrier peak energy. By performing $g_m(V_G)$ measurements at multiple V_D values one can further extract V_c , and temperature-dependent sweeps give E_b and Δ from $\bar{v}_{\text{eq}}(T)$. Taken together, a single set of g_m - V_G - V_D measurements at varied T constrains the parameters $\{n_c, \sigma, E_0, V_c, E_b, \Delta\}$, complementary to optical techniques that typically ac-

cess only one or two of these simultaneously. These extractions are model-assisted: they presuppose the calibrated spectral-kernel form (Step 1) and a localized excess, so the inferred values are best regarded as effective parameters of that description rather than uniquely model-independent quantities. In particular, E_0 enters the peak height only together with α_{max} and the kernel, and is not fixed by the height alone.

g. Two-time-scale separation as the third experimental pillar. Equation (22) establishes a hierarchy of electrical measurements for complete spectroscopy of $f(E)$. The steady-state $g_m(V_G)$ sweep (Pillar I: Fig. 3) constrains the spectral parameters $\{E_0, \sigma, n_c\}$ through the peak position and height. The over-constrained consistency between g_m and I_D (Pillar II) excludes any effective-temperature description in the main text [Fig. 4(c,d)]. A quantitative simultaneous fit to synthetic data, which verifies that the non-thermal model is strongly constrained by the joint g_m - I_D data, is provided as a supplementary demonstration in Appendix B (Fig. 7). The two-time-scale transient (Pillar III: Fig. 5) then isolates the energy relaxation time τ_E from the density relaxation time τ_n in a single $I_D(t)$ measurement by exploiting the bi-exponential structure of Eq. (22). Together these three observables constrain the full parameter set $\{E_0, \sigma, n_c, \tau_E\}$ within the calibrated model from electrical measurements alone, without relying on optical spectroscopy for readout. Whereas Pillars I and II rely on standard DC/lock-in sweeps, Pillar III requires GHz-to-THz-bandwidth time-domain instrumentation and is therefore best viewed as a specialized extension that complements the steady-state observables rather than replacing them.

h. Experimental measurement protocol. We outline a self-contained protocol by which the spectroscopic parameters can be extracted sequentially from standard electrical measurements. The minimum experimental dataset for verifying the non-thermal mechanism and extracting $\{n_c, \sigma, V_c, E_0, \alpha_{\text{max}}\}$ is a set of $g_m(V_G)$ traces at several V_D values (Steps 2–4 below), which requires only a standard DC/lock-in FET sweep. The time-domain extension (Step 5) is optional and yields the additional parameter τ_E .

Step 1 — equilibrium calibration. At low drain voltage $V_D \ll V_c$, measure $I_D(V_G)$ and $g_m(V_G)$ at several temperatures $T = 200$ – 350 K. The transconductance is flat and equals $(WC_{\text{ox}}/L) \bar{v}_{\text{eq}}(T)$, giving the equilibrium velocity $\bar{v}_{\text{eq}}(T)$, which follows the thermally averaged spectral kernel and increases with temperature. Fitting this temperature dependence to Eqs. (8)–(11) with $\alpha = 0$ extracts the transport-kernel onset parameters E_b and Δ .

Step 2 — hot-carrier signature. Increase V_D to values comparable to or exceeding V_c . Subtract the smooth equilibrium-calibrated density background $g_m^{(n)}$ from the measured $g_m(V_G)$. If the residual shape contribution $g_m^{(\alpha)}$ develops a localized peak, the effective-temperature description has broken down and the non-thermal model is required. The absence of such a peak at the same

V_D and T would instead confirm quasi-equilibrium conditions, providing a built-in null test. Repeat at several V_D values ($V_D = 0.2\text{--}0.7$ V recommended) to track the peak height as a function of V_D .

Step 3 — V_D -dependent peak height. Plot the peak height $g_m^{(\alpha),\text{pk}}(V_D)$ against V_D . A fit to $h(V_D) = V_D^2/(V_D^2 + V_c^2)$ [Eq. (17)] extracts V_c , the electric-field crossover for hot-carrier generation. The peak position $V_G^{\text{pk}}(V_D)$ simultaneously yields $n_{\text{pk}}(V_D)$ via Eq. (12); fitting to Eq. (19) with V_c known then extracts n_c and the spectral width σ jointly.

Step 4 — simultaneous fit and falsification. Using the parameters $\{E_b, \Delta, V_c, n_c, \sigma\}$ from Steps 1–3, compute $I_D(V_G)$ and $g_m(V_G)$ at all measured V_D values. The non-thermal model should reproduce both observables with strongly reduced parameter freedom: only E_0 and α_{max} remain as global fit parameters, both already constrained by Step 2. As a falsification test, confirm that the best-fit T_{eff} model gives a flat $g_m(V_G)$ and fails to capture the shape contribution $g_m^{(\alpha)}$ [Fig. 4(d)].

Step 5 — time-domain extraction of τ_E . Apply a fast electrical pulse (gate or drain) with a sub-nanosecond to picosecond-scale rise time and record $I_D(t)$ with high-bandwidth (GHz-to-THz) instrumentation. Fit the bi-exponential transient [Eq. (22)] to extract τ_n (the fast, density-relaxation component, set by device RC) and τ_E (the slow, spectral-relaxation component, set by carrier-phonon scattering). The full parameter set $\{E_0, \sigma, n_c, V_c, E_b, \Delta, \tau_E\}$ is then constrained within the calibrated model from all-electrical measurements.

i. Applicability and experimental scales. The method applies whenever a localized nonthermal excess is separated from the transport onset by more than the thermal broadening, $E_0 - E_b \gtrsim \text{few} \times k_B T$, so that the hot-carrier feature is not washed out by the Boltzmann tail. For the representative TMD parameters used here ($E_b = 50$ meV, $E_0 = 250$ meV) this is comfortably satisfied at room temperature ($k_B T \approx 26$ meV), and the signature therefore survives to 300 K and above. Lowering the temperature sharpens the spectral resolution but can enhance competing trap and contact effects; raising it broadens both the thermal background and the hot-carrier excess. In gate voltage, the relevant density scale is $n \sim n_c \sim 10^{12} \text{ cm}^{-2}$, reached routinely in strongly gated, high- κ TMD FETs, and short channels ($L \sim 10\text{--}100$ nm) maximize the high-field region that generates the excess. The steady-state observables (Pillars I and II) require only standard DC/lock-in instrumentation, whereas the optional time-domain extraction of τ_E (Pillar III) assumes energy-relaxation times $\tau_E \sim 0.1\text{--}10$ ps typical of TMDs and hence GHz-to-THz measurement bandwidth. Outside these ranges—strongly degenerate channels, E_0 comparable to E_b , or dominant elastic disorder or spatial inhomogeneity (e.g. threshold variation or contact nonuniformity) that smears the onset—the shape contribution persists in principle but its peak may be too weak or too broad to resolve.

j. Material and geometry considerations. The signal-to-noise ratio for detecting $g_m^{(\alpha)}$ is maximized when $\bar{v}_{\text{neq}}/\bar{v}_{\text{eq}}$ is large, which requires $E_0 \gg E_b$ and $E_0 \gg k_B T$. TMDs with high- κ gate dielectrics (large C_{ox}) shift V_G^* to more accessible gate voltages. Short-channel devices ($L \sim 10\text{--}50$ nm) increase the high-field region and hence α at a given V_D . For the pump-probe time-domain experiment, the bandwidth requirement is $\sim 1/\tau \sim 1$ THz for TMD materials [36], suggesting that THz near-field probing or GHz on-chip circuitry could realize this measurement.

V. CONCLUSION

We have established that the transconductance g_m of a 2D FET is fundamentally a spectroscopic quantity. It is the spectral current $j(E)$ weighted by the gate-voltage derivative of the carrier distribution $f(E)$ and integrated over energy. As such, it encodes the *shape* of $f(E)$, not merely its integrated density. This spectroscopic identity [Eq. (15)] is exact for a gate-independent spectral kernel. It gives rise to an anomalous peak in the distribution-shape contribution $g_m^{(\alpha)}$, isolated from the smooth monotone density background of the measured g_m . This peak (i) has no counterpart in equilibrium transport, (ii) cannot be reproduced by any quasi-equilibrium description without invoking ad hoc nonmonotonic V_G -dependence in transport coefficients, and (iii) constrains the spectral parameters of the hot-carrier distribution through its position, height, and drain-voltage dependence.

Two steady-state observables, accessible with standard DC/lock-in instrumentation, provide the primary experimental verification of the non-thermal character of the distribution. Pillar I is the anomalous peak in the shape contribution $g_m^{(\alpha)}$ (Fig. 3). Its decisive drain-voltage fingerprint—a peak that shifts to lower V_G and saturates in height with V_D —distinguishes it from a conventional mobility-rolloff peak, which would instead remain fixed in position and grow linearly with V_D [Fig. 4(a,b)]. Pillar II is the exclusion of any effective-temperature description, which gives a flat $g_m(V_G)$ and fails to reproduce the shape contribution $g_m^{(\alpha)}$ even at its best fit [Fig. 4(c,d)]. The over-constrained simultaneous fit of both I_D and g_m with a single parameter set $\{E_0, \sigma, \alpha_{\text{max}}, n_c\}$ provides a supplementary quantitative confirmation (Appendix B, Fig. 7). Together these constrain the effective spectral parameters $\{E_0, \sigma, n_c, V_c, E_b, \Delta\}$ without optical spectroscopic readout. The energy relaxation time τ_E becomes accessible through a more demanding time-domain extension—the bi-exponential transient $I_D(t)$ (Pillar III, Fig. 5)—which separates the energy relaxation from the density relaxation τ_n but requires GHz-to-THz-bandwidth instrumentation.

The transconductance provides a spectroscopic probe of carrier distributions, revealing non-thermal features that lie beyond the reach of any quasi-equilibrium description

without *ad hoc nonmonotonic gate-voltage dependence*. This capability positions gated 2D semiconductors as on-chip distribution-function spectrometers, opening pathways to (i) systematic mapping of hot-carrier dynamics across phonon channels, valley degeneracies, and temperatures in van der Waals monolayers; (ii) engineering of $g_m^{(\alpha)}$ peaks via dielectric and channel geometry for energy-selective carrier amplification; and (iii) extending transconductance spectroscopy to multi-band and spin-orbit-coupled 2D systems where distribution-shape effects couple to valley and spin degrees of freedom.

ACKNOWLEDGMENTS

The author thanks S. Adhikary and T. Kameda for useful discussions. This work was supported by JSPS KAKENHI (Grants No. JP25K01609, No. JP22H05473, and No. JP21H01019).

DATA AVAILABILITY

The data that support the findings of this study, and the code used to generate the figures, are available from the corresponding author upon reasonable request.

Appendix A: Pump-probe transient at fixed carrier density

The two-time-scale analysis of the main text reduces, in the limit of a fixed carrier density, to a single-exponential pump-probe transient that isolates the energy relaxation time τ_E . Consider a pump (optical or electrical) that injects hot carriers at $t = 0$, establishing $\alpha(t=0) = \alpha_0$. For $t > 0$ the hot carriers relax via carrier-phonon scattering, $\alpha(t) = \alpha_0 e^{-t/\tau_E}$, with τ_E the energy relaxation time. In a pump-probe geometry, the gate is held fixed at $V_G = V_G^*$ and a small AC modulation δV_G probes the instantaneous transconductance. Because the pump-generated α is externally imposed and does not depend on V_G , only the density-driven term survives:

$$g_m(t) = \frac{WC_{\text{ox}}}{L} \bar{v}(\alpha_0 e^{-t/\tau_E}). \quad (\text{A1})$$

Figures 6(a) and (b) show the resulting transients in $I_D(t)$ and $g_m(t)$ for $\tau_E = 1$ ps, normalized to their equilibrium values. Both decay monotonically from their pump-excited values back to equilibrium. The decay is exponential in $\alpha(t)$ but not strictly exponential in I_D or g_m , because $\bar{v}(\alpha)$ is nonlinear in α . Fitting the transient to Eq. (A1) therefore provides τ_E and simultaneously $\bar{v}_{\text{neq}}/\bar{v}_{\text{eq}}$, a ratio that encodes E_0 (the hot-carrier energy) via Eqs. (8)–(11).

Figure 6(c) shows normalized transients $\Delta g_m(t)/\Delta g_m(0)$ for several relaxation times τ_E .

The curves are distinguishable over a decade in τ_E within a 6 ps window, demonstrating that the measurement discriminates phonon-scattering channels with sub-picosecond resolution, consistent with time-resolved Raman measurements of longitudinal-optical (LO) phonon populations in III–V semiconductors [33]. Compared to optical pump-probe techniques [15, 16] and recent electrical photoresponse measurements in van der Waals heterostructures [37], the present approach directly accesses the *transport-weighted* carrier distribution rather than the optical spectral weight or the transient photocurrent, and can be combined with electrostatic gating to tune n and V_G^* in situ.

Appendix B: Simultaneous fit to synthetic data

Figure 7 demonstrates quantitatively that the non-thermal model f_{neq} [Eq. (7)] with the *single* parameter set $\{E_0, \sigma, \alpha_{\text{max}}, n_c\}$ simultaneously reproduces both $g_m(V_G)$ [panel (a)] and $I_D(V_G)$ [panel (b)]. The two observables probe fundamentally different moments of $f(E)$: g_m is sensitive to $\partial f/\partial V_G$ and therefore to the *shape* of f , while I_D tracks the total spectral weight. A model that fits both with a single parameter set is therefore over-constrained, providing unambiguous evidence that the spectral shape of $f(E)$ governs transport, not merely its integrated density or effective temperature. The T_{eff} model cannot achieve this: it adds only one free parameter yet already fails at the first observable (g_m).

Appendix C: Robustness against modeling choices

The two phenomenological ingredients of the model—the saturation function $g(n)$ in the hot-carrier amplitude $\alpha(n)$ and the energy-selective transmission $\mathcal{T}(E)$ —are not derived from first principles. We show here that the anomalous $g_m^{(\alpha)}$ peak is insensitive to both choices.

1. Saturation function $\alpha(n)$

The phenomenological amplitude of Eq. (13) uses the saturating form $g(n) = n/(n + n_c)$. The anomalous peak in the shape-driven transconductance does not rely on this specific choice: any monotonically saturating $g(n)$ gives a bell-shaped $d\alpha/dn$ and hence a localized $g_m^{(\alpha)}$ peak. Figure 8(a,b) compares four representative forms—the baseline hyperbola $n/(n + n_c)$, a hyperbolic tangent $\tanh(n/n_c)$, an exponential saturation $1 - e^{-n/n_c}$, and a Hill function $n^2/(n^2 + n_c^2)$ —all normalized to rise from zero and saturate on the scale $n \sim n_c$ [Fig. 8(a)]. The resulting $g_m^{(\alpha)}(V_G)$ curves [Fig. 8(b)], each normalized to its own peak, all exhibit a localized peak near V_G^* with comparable width. The forms with

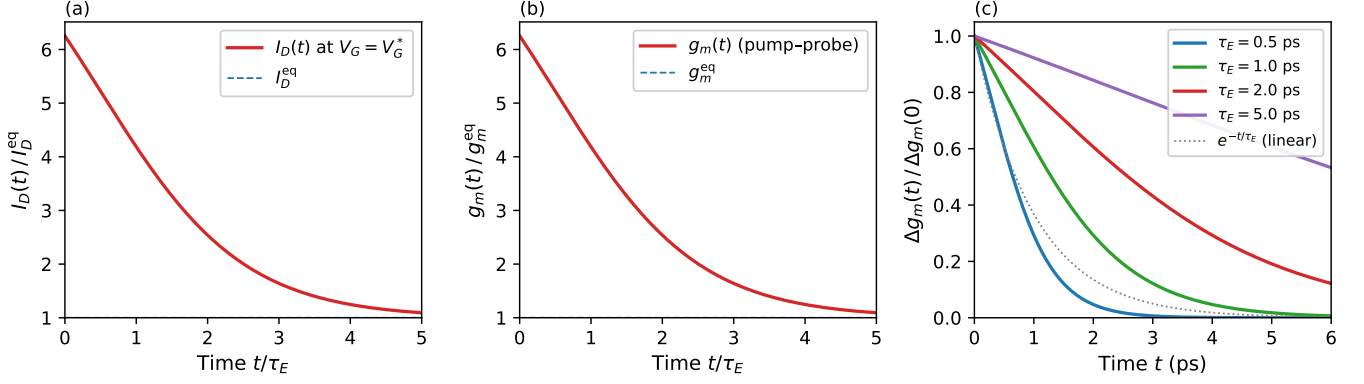


FIG. 6. Time-resolved response after pump excitation that injects hot carriers with $\alpha_0 = 0.6$ at $t = 0$, for $\tau_E = 1$ ps. (a) Drain current transient $I_D(t)/I_D^{\text{eq}}$ at $V_G = V_G^*$. (b) Transconductance transient $g_m(t)/g_m^{\text{eq}}$, which decays from the nonequilibrium value back to equilibrium as hot carriers relax. (c) Normalized transconductance anomaly $\Delta g_m(t)/\Delta g_m(0)$ for relaxation times $\tau_E = 0.5, 1.0, 2.0$, and 5.0 ps (blue, green, red, purple). The gray dotted line shows the pure exponential e^{-t/τ_E} for comparison, illustrating the slight nonlinearity due to $\bar{v}(\alpha)$.

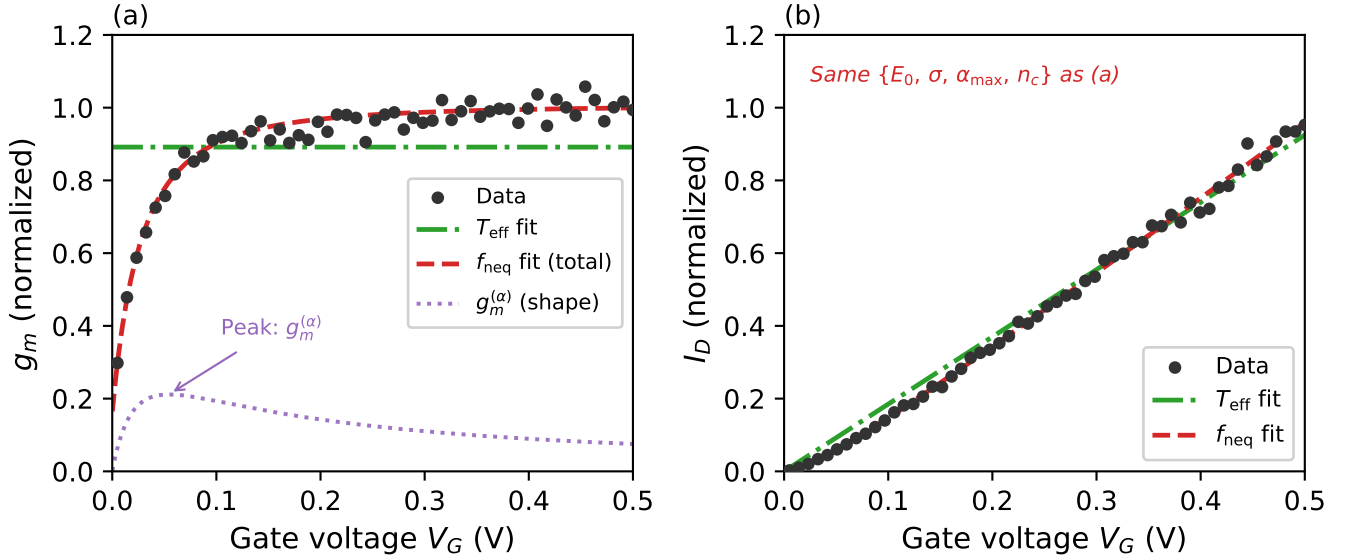


FIG. 7. Simultaneous fit of $g_m(V_G)$ and $I_D(V_G)$ using the non-thermal model f_{neq} (red dashed) versus the best-fit $T_{\text{eff}} \approx 1680$ K model (green dash-dot). Black circles are synthetic data generated from the full non-thermal model with 3% Gaussian noise added (realistic noise level for transport measurements). (a) $g_m(V_G)$: the non-thermal fit reproduces the data, including the shape contribution $g_m^{(\alpha)}$ (purple dotted) that carries the anomalous peak; the T_{eff} model gives a flat response with no such shape term. (b) $I_D(V_G)$: both models approximate the current, but only f_{neq} uses the *same* four parameters as panel (a). The simultaneous fit over-constrains the model and demonstrates that spectral shape—not density or temperature alone—governs transport.

a steeper onset (tanh, Hill) place the peak slightly lower in V_G and decay faster at large V_G , but the qualitative signature—a localized $g_m^{(\alpha)}$ peak driven purely by the distribution shape—is common to all of them. The peak *position* is therefore a robust fingerprint of the saturating hot-carrier generation mechanism rather than an artifact of the particular analytic model adopted in the main text.

2. Transmission $\mathcal{T}(E)$

The main text models the transmission by the Fermi-function barrier of Eq. (5). The anomaly does not rely on this form either. The shape-driven term enters through the spectral current moments $J_{\text{eq,neq}} = \int j(E) f_{\text{eq,neq}}(E) dE$ with $j(E) = v(E)\mathcal{T}(E)$, and its sign and localization require only that the hot-carrier component at $E_0 > E_b$ be transmitted more efficiently per

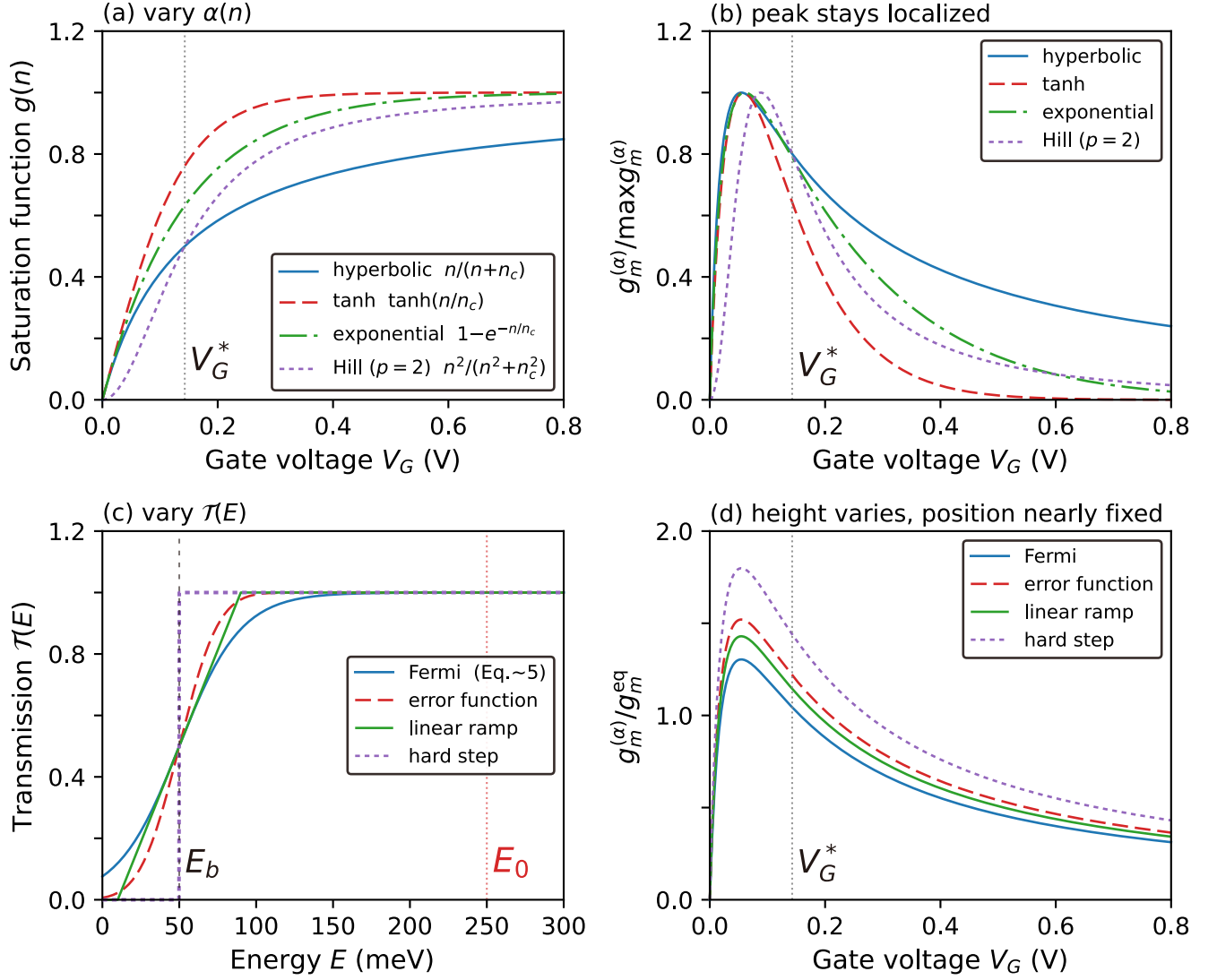


FIG. 8. Robustness of the $g_m^{(\alpha)}$ peak against the two phenomenological modeling choices. *Top row*—saturation function $\alpha(n)$. (a) Four saturating forms of $g(n)$ versus gate voltage: hyperbolic $n/(n+n_c)$ (blue), $\tanh(n/n_c)$ (red dashed), $1 - e^{-n/n_c}$ (green dash-dot), and the Hill function $n^2/(n^2 + n_c^2)$ (purple dotted). (b) Corresponding $g_m^{(\alpha)}(V_G)$, each normalized to its own peak: the peak *position* stays near V_G^* for every form. *Bottom row*—transmission $\tau(E)$. (c) Four transmission forms rising from 0 to 1 across E_b : Fermi step (blue), error-function (red dashed), linear ramp (green), and the discontinuous hard step $\Theta(E - E_b)$ (purple dotted); E_0 marks the hot-carrier energy. (d) Corresponding $g_m^{(\alpha)}(V_G)$, normalized to the equilibrium plateau g_m^{eq} : the peak occurs at the *same* gate voltage for all forms (position set by $\alpha(n)$, not τ); only the *height* varies.

carrier than the thermal component, i.e. $\bar{v}_{\text{neq}} > \bar{v}_{\text{eq}}$. Figure 8(c,d) compares four transmission forms that all rise from 0 to 1 across E_b —the baseline Fermi step $[1 + e^{(E_b - E)/\Delta}]^{-1}$, an error-function step, a linear ramp, and the discontinuous hard step $\Theta(E - E_b)$ [Fig. 8(c)]. The resulting $g_m^{(\alpha)}(V_G)$, normalized to the equilibrium plateau g_m^{eq} of the same form [Fig. 8(d)], peaks at the *same* gate voltage for every transmission—the peak position is fixed by $\alpha(n)$, not by $\tau(E)$ —while its height varies only mildly, remaining of order the equilibrium response (1.3–1.8 g_m^{eq}) even for the discontinuous hard step.

The transmission shape thus sets the magnitude of the anomaly but neither its existence nor its location. Taken together, the two robustness checks are complementary: the peak *position* is set by $\alpha(n)$ and is insensitive to $\tau(E)$, while the peak *height* is set by $\tau(E)$ and the hot-carrier energy—so neither phenomenological input controls the existence of the signature.

Appendix D: Extension to linear (Dirac) dispersion

The framework does not rely on the constant density of states of a parabolic 2D band. For a linear (Dirac) dispersion $E = \hbar v_F |\mathbf{k}|$, the density of states is $D(E) = g E / (2\pi \hbar^2 v_F^2) \propto |E|$ and the band velocity is energy-independent, $v(E) = v_F$. Writing the drain current in the same energy-resolved form [Eq. (3)], the transport now enters through the spectral kernel

$$j_D(E) = D(E) v_F \mathcal{T}(E) \propto E \mathcal{T}(E), \quad (\text{D1})$$

which replaces the parabolic-band kernel $j(E) = v(E) \mathcal{T}(E) \propto \sqrt{E} \mathcal{T}(E)$. In both cases $j(E)$ is a smooth, monotonically increasing energy filter that favors high-

energy carriers, so a localized nonthermal excess at $E_0 > E_b$ is still weighted more heavily than the thermal population. Consequently the transconductance retains the spectroscopic form Eq. (15) and the decomposition $g_m = g_m^{(n)} + g_m^{(\alpha)}$, and the anomalous $g_m^{(\alpha)}$ peak persists. Only the quantitative details change. Because the Dirac kernel grows linearly rather than as $\sqrt{E}$, it weights the hot-carrier excess even more strongly, so the per-carrier enhancement $\bar{v}_{\text{neq}}/\bar{v}_{\text{eq}}$, and hence the peak height, is generally larger. The peak *position*, set by $\alpha(n)$ (Appendix C), is unchanged. The same reasoning applies to any monotonic product $D(E) v(E)$, so the mechanism is generic to gapped and gapless 2D materials alike.

-
- [1] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Atomically thin MoS₂: a new direct-gap semiconductor, *Phys. Rev. Lett.* **105**, 136805 (2010).
 - [2] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, Electronics and optoelectronics of two-dimensional transition metal dichalcogenides, *Nat. Nanotechnol.* **7**, 699 (2012).
 - [3] M. Chhowalla, H. S. Shin, G. Eda, L.-J. Li, K. P. Loh, and H. Zhang, The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets, *Nat. Chem.* **5**, 263 (2013).
 - [4] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, Single-layer MoS₂ transistors, *Nat. Nanotechnol.* **6**, 147 (2011).
 - [5] K. Kang, S. Xie, L. Huang, Y. Han, P. Y. Huang, K. F. Mak, C.-J. Kim, D. Muller, and J. Park, High-mobility three-atom-thick semiconducting films with wafer-scale homogeneity, *Nature* **520**, 656 (2015).
 - [6] L. Liao, Y.-C. Lin, M. Bao, R. Cheng, J. Bai, Y. Liu, Y. Qu, K. L. Wang, Y. Huang, and X. Duan, High-speed graphene transistors with a self-aligned nanowire gate, *Nature* **467**, 305 (2010).
 - [7] L. Li, Y. Yu, G. J. Ye, Q. Ge, X. Ou, H. Wu, D. Feng, X. H. Chen, and Y. Zhang, Black phosphorus field-effect transistors, *Nat. Nanotechnol.* **9**, 372 (2014).
 - [8] G. Fiori, F. Bonaccorso, G. Iannaccone, T. Palacios, D. Neumaier, A. Seabaugh, S. K. Banerjee, and L. Colombo, Electronics based on two-dimensional materials, *Nat. Nanotechnol.* **9**, 768 (2014).
 - [9] W. Li, H. Shen, H. Qiu, Y. Shi, and X. Wang, Two-dimensional semiconductor transistors and integrated circuits for advanced technology nodes, *Natl. Sci. Rev.* **11**, nwae001 (2024).
 - [10] J. Chen, M.-Y. Sun, Z.-H. Wang, Z. Zhang, K. Zhang, S. Wang, Y. Zhang, X. Wu, T.-L. Ren, H. Liu, and L. Han, Performance limits and advancements in single 2d transition metal dichalcogenide transistor, *Nano-Micro Lett.* **16**, 264 (2024).
 - [11] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. (Wiley-Interscience, Hoboken, NJ, 2007).
 - [12] M. Lundstrom, *Fundamentals of Carrier Transport*, 2nd ed. (Cambridge University Press, Cambridge, 2000).
 - [13] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995).
 - [14] M. Massicotte, G. Soavi, A. Principi, and K.-J. Tielrooij, Hot carriers in graphene—fundamentals and applications, *Nanoscale* **13**, 8376 (2021).
 - [15] D. Sun, Z.-K. Wu, C. Divin, X. Li, C. Berger, W. A. de Heer, P. N. First, and T. B. Norris, Ultrafast relaxation of excited Dirac fermions in epitaxial graphene using optical differential transmission spectroscopy, *Phys. Rev. Lett.* **101**, 157402 (2008).
 - [16] E. A. A. Pogna, M. Marsili, D. De Fazio, S. Dal Conte, C. Manzoni, D. Sangalli, D. Yoon, A. Lombardo, A. C. Ferrari, A. Marini, G. Cerullo, and D. Prezzi, Photo-induced bandgap renormalization governs the ultrafast response of single-layer MoS₂, *ACS Nano* **10**, 1182 (2016).
 - [17] H. Shi, R. Yan, S. Bertolazzi, J. Brivio, B. Gao, A. Kis, D. Jena, H. G. Xing, and L. Huang, Exciton dynamics in suspended monolayer and few-layer MoS₂ 2D crystals, *ACS Nano* **7**, 1072 (2013).
 - [18] W. Y. Wang, N. Sui, X. C. Chi, Z. H. Kang, Q. Zhou, L. Li, H. Z. Zhang, J. B. Gao, and Y. H. Wang, Investigation of hot carrier cooling dynamics in monolayer mos₂, *J. Phys. Chem. Lett.* **12**, 861 (2021).
 - [19] W. Lee, Y. Lin, L.-S. Lu, W.-C. Chueh, M. Liu, X. Li, W.-H. Chang, R. A. Kaindl, and C.-K. Shih, Time-resolved arpes determination of a quasi-particle band gap and hot electron dynamics in monolayer mos₂, *Nano Lett.* **21**, 7363 (2021).
 - [20] M. Büttiker, Four-terminal phase-coherent conductance, *Phys. Rev. Lett.* **57**, 1761 (1986).
 - [21] G. D. Mahan and J. O. Sofo, The best thermoelectric, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7436 (1996).
 - [22] J. Tersoff and D. R. Hamann, Theory of the scanning tunneling microscope, *Phys. Rev. B* **31**, 805 (1985).
 - [23] L. P. Kouwenhoven, D. G. Austing, and S. Tarucha, Few-electron quantum dots, *Rep. Prog. Phys.* **64**, 701 (2001).
 - [24] H. Pothier, S. Guéron, N. O. Birge, D. Esteve, and M. H. Devoret, Energy distribution function of quasiparticles in mesoscopic wires, *Phys. Rev. Lett.* **79**, 3490 (1997).
 - [25] K. Wakabayashi, S. Adhikary, and T. Kameda, Transconductance as a probe of valley thermodynamics in multi-layer WSe₂, *Phys. Rev. B* **114**, 115403 (2026).

- [26] K. Wakabayashi, S. Adhikary, and K. Tsukagoshi, Valley-landscape engineering in bilayer WSe_2 gate-all-around transistors (2026), arXiv:2606.08955 [cond-mat.mes-hall].
- [27] K. Wakabayashi, Electrical spectroscopy of intervalley relaxation in WSe_2 transistors, *Phys. Rev. B* **114**, 105421 (2026).
- [28] The exact group velocity of the 2D parabolic band is $v(E) = \sqrt{2E/m^*}$; here we absorb the $O(1)$ numerical prefactor (the $\sqrt{2}$ and the angular average of the transport-direction component) into the definition of v_T , since only the $\sqrt{E}$ energy dependence enters the spectral shape—overall constants cancel in the velocity ratios $\bar{v}(\alpha)$ and in g_m .
- [29] K. K. H. Smithe, C. D. English, S. V. Suryavanshi, and E. Pop, High-field transport and velocity saturation in synthetic monolayer mos_2 , *Nano Lett.* **18**, 4516 (2018).
- [30] M. A. Wang and E. Pop, Monte carlo simulation of electrical transport with joule heating and strain in monolayer mos_2 devices, *Nano Lett.* **25**, 6841 (2025).
- [31] A. Allain, J. Kang, K. Banerjee, and A. Kis, Electrical contacts to two-dimensional semiconductors, *Nat. Mater.* **14**, 1195 (2015).
- [32] K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, Phonon-limited mobility in n-type single-layer mos_2 from first principles, *Phys. Rev. B* **85**, 115317 (2012).
- [33] J. A. Kash, J. C. Tsang, and J. M. Hvam, Subpicosecond time-resolved Raman spectroscopy of LO phonons in GaAs, *Phys. Rev. Lett.* **54**, 2151 (1985).
- [34] P.-C. Shen, C. Su, Y. Lin, A.-S. Chou, C.-C. Cheng, J.-H. Park, M.-H. Chiu, A.-Y. Lu, H.-L. Tang, M. M. Tavakoli, G. Pitner, X. Ji, Z. Cai, N. Mao, J. Wang, V. Tung, J. Li, J. Bokor, A. Zettl, C.-I. Wu, T. Palacios, L.-J. Li, and J. Kong, Ultralow contact resistance between semimetal and monolayer semiconductors, *Nature* **593**, 211 (2021).
- [35] M. Du, W. Li, G. Xiong, C. Zhao, F. Hou, W. Gan, X. Gong, N. Zou, L. Liu, X. Zou, T. Li, W. Sun, D. Fan, Z. Yu, X. Tu, Y. Gao, H. Shen, H. Qiu, L. Ma, J. Wang, Y. Nie, L. Tao, J.-B. Xu, J. Lin, J. Xu, Y. Shi, and X. Wang, Scaled crystalline antimony ohmic contacts for two-dimensional transistors, *Nat. Electron.* **8**, 1191 (2025).
- [36] N. Kumar, J. He, D. He, Y. Wang, and H. Zhao, Charge carrier dynamics in bulk MoS_2 crystal studied by transient absorption microscopy, *J. Appl. Phys.* **113**, 133702 (2013).
- [37] M. Massicotte, P. Schmidt, F. Vialla, K. G. Schädler, A. Reserbat-Plantey, K. Watanabe, T. Taniguchi, K. J. Tielrooij, and F. H. L. Koppens, Picosecond photoreponse in van der Waals heterostructures, *Nat. Nanotechnol.* **11**, 42 (2016).