Valley-polarization in biased bilayer graphene using circularly polarized light

A. Friedlan, M. M. Dignam
Department of Physics, Engineering Physics & Astronomy, Queen’s University, Kingston, Ontario, Canada, K7L 3N6

Achieving a population imbalance between the two inequivalent valleys is a critical first step for any valleytronic device. A valley-polarization can be induced in biased bilayer graphene using circularly polarized light. In this paper, we present a detailed theoretical study of valley-polarization in biased bilayer graphene. We show that a nearly perfect valley-polarization can be achieved with the proper choices of external bias and pulse frequency. We find that the optimal pulse frequency \( \omega \) is given by \( \hbar \omega = 2a \), where \( 2a \) is the potential energy difference between the graphene layers. We also find that the valley-polarization originates not from the Dirac points themselves, but rather from a ring of states surrounding each. Intervalley scattering is found to significantly reduce the valley-polarization at high frequencies, and thermal populations are found to reduce the valley-polarization at small biases. This work provides insight into the origin of valley-polarization in bilayer graphene and will aid experimentalists seeking to study valley-polarization in the lab.

I. INTRODUCTION

Since its first realization in 2004 \[1\], graphene has promised to revolutionize electronics with its high electron mobility \[2, 3\], impressive mechanical strength \[4\], and tunable Fermi level \[5\]. There exist two inequivalent local minima in graphene’s band structure known as valleys or Dirac points which we label \( K \) and \( K' \). In analogy with spintronics \[6\], the valley index is binary and the concept of using this two-state system to perform logical operations is known as valleytronics \[7\]. To realize such a system, we require a way to induce a valley-polarization, that is, a differential electron population between the \( K \) and \( K' \) valleys.

There have been many proposals for valleytronic devices based on monolayer graphene, however most have relied on configurations that may be difficult to realize in the lab \[8–12\]. Intrinsically, the \( K \) and \( K' \) valleys are indistinguishable from one another. This means that it is difficult to selectively populate the valleys, say, using an optical field. Inversion symmetry breaking is necessary for graphene-based valleytronics \[13, 14\]. One solution is to use a staggered sublattice potential, for instance, by growing graphene on a substrate of hBN \[15\]. Another option is to consider materials with intrinsically broken inversion symmetry. TMDs such as monolayer MoS\(_2\) have gained significant interest recently, in part due to the presence of an intrinsic band gap at the Dirac points \[16–18\]. In this work we consider bilayer graphene, which consists of two graphene sheets stacked in an AB/Bernal stacking arrangement \[19\]. Biasing the bilayer by applying a potential difference across the two graphene sheets breaks the inversion symmetry and opens a band gap \[20, 23\]. Not only that, but the band gap can be tuned continuously from zero to the mid-infrared by adjusting the strength of the external bias \[24, 26\]. For an excellent review of the electronic properties of both monolayer and bilayer graphene, please see McCann \[27\].

It has been proposed that circularly polarized light can be used preferentially inject carriers into the \( K \) and \( K' \) valleys of bilayer graphene \[14\]. Right-hand circularly polarized light couples strongly to the \( K \) valley, while light of the opposite helicity couples strongly to \( K' \). There has been significant work towards inducing valley-polarized currents in bilayer graphene with broken inversion symmetry \[28, 29\], but very few studies have focused on using circularly polarized light to induce a valley-polarization \[30, 31\]. To the best of our knowledge, no studies have yet sought to maximize the optically-induced valley-polarization, leaving experimentalists ill-equipped to study this phenomenon in the laboratory. Several important questions remain unanswered: What is the optimal operating external bias? What is the optimal operating pulse frequency? And what pulse duration should be used? It is also to date unknown as to which scattering processes fundamentally limit performance of bilayer-graphene-based valleytronic devices: How clean a sample is required? Can a valley-polarization be observed at room temperature? In this paper, we seek to answer these questions as well as offer valuable insight into the underlying physics of valley-polarization in bilayer graphene.

Our findings can be summarized as follows. At low temperatures, and in the absence of scattering, the strongest valley-polarization is achieved when \( \hbar \omega = 2a \), where \( \omega \) is the frequency of the optical field, and \( 2a \) is the potential energy difference between the graphene layers. This result originates from a \( k \)-dependent valley-contrasting optical selection rule which becomes exact when \( \hbar \omega = 2a \). This finding is qualitatively consistent with some previous calculations \[14, 30\], but so far seems to have gone unnoticed in the literature. Our results also indicate that intervalley scattering via optical phonons can significantly reduce the valley-polarization when operating at high frequencies. In addition, we find that thermal populations can significantly reduce the valley-polarization when operating at low biases. With scattering and thermal populations taken into account, we find that a valley-polarization of up to 70% can be achieved at room temperature. At low temperatures \((< 150 \text{ K})\), the valley-polarization can be as large as 97%. 
This paper is organized as follows. In Sec. II we present our theoretical model. We construct the graphene Hamiltonian and solve for the energy bands and eigenstates. We then develop our density matrix equations of motion, solving them perturbatively for excitation by circularly polarized light. In Sec. III we study the resulting valley-polarization, primarily as a function of the frequency of the exciting field and of the external bias between the graphene layers. We first examine the effects of varying the pulse duration and decoherence time before concluding in Sec. IV.

II. THEORY

We employ a four-band nearest-neighbor tight-binding model to calculate the low-energy electron bands and Bloch eigenstates. We perturb the system with an optical field, treating the interaction within the length gauge. We develop density matrix equations of motion and solve them up to second-order. We calculate the electron populations in the K and K' valleys that result from the linear absorption of a circularly polarized Gaussian pulse.

A. Tight-binding

We use as our basis the single-atom Bloch functions

$$\Phi_i(k, r) = \sum_{j=1}^{N} e^{i k R_{j,i}} \phi(r - R_{j,i}),$$

where \( k \) is the Bloch wave vector, \( r \) is the position vector, and \( \phi(r) \) is a carbon 2p\(_z\) orbital \[27\]. The sum is over the \( N \) different unit cells, and \( R_{j,i} \equiv R_j + r_i \), where \( R_j \) is a Bravais lattice vector and \( r_i \) is a basis vector, which denotes the position of one of the four atoms in the unit cell. Following the coordinate conventions of Ref. [32], we have \( r_{A_1} = d\hat{z}, r_{B_1} = a_0 \hat{x} + d\hat{z}, r_{A_2} = -a_0 \hat{x} - d\hat{z}, \) and \( r_{B_2} = -d\hat{z} \), where \( a_0 = 1.42 \text{ Å} \) is the interatomic distance, and \( 2d \) is the interlayer spacing (see Fig. [1]).

Using these basis states, we construct our eigenstates

$$\Psi_{nk}(r) = \langle r|nk\rangle = A_n(k) \sum_i c_{ni}(k) \Phi_i(k, r),$$

where \( A_n(k) \) is a normalization factor, the \( c_{ni}(k) \) are expansion coefficients, \( i \) indexes the atoms, and \( n \) labels the band. In the basis \( i = \{ A_1, B_1, A_2, B_2 \} \), including hopping between nearest-neighbors within each layer, and between the overlapping A1 and B2 atoms in opposite layers, we obtain the nearest-neighbor tight-binding Hamiltonian \[27\]

$$\mathcal{H}_0 = \begin{pmatrix}
  a & f(k) t \parallel & 0 & t \perp \\
  f^*(k) t \parallel & a & 0 & 0 \\
  0 & 0 & -a & f(k) t \parallel \\
  t \perp & f^*(k) t \parallel & -a & -a
\end{pmatrix},$$

where \( (by \ convention) 2a \geq 0 \) is the potential energy difference between the graphene layers, and where \( t \parallel = 3.3 \text{ eV} \) and \( t \perp = 0.42 \text{ eV} \) are, respectively, the intra- and inter-layer hopping energies \[19\]. The function \( f(k) = 1 + e^{-i k \cdot a_0} + e^{-i k \cdot a_2} \) describes hopping between nearest-neighbor sites, where \( a_1 = a_0 (3\hat{x} + \sqrt{3}\hat{y})/2 \) and \( a_2 = a_0 (3\hat{x} - \sqrt{3}\hat{y})/2 \) are the primitive translation vectors.

In what follows, we will focus on the dynamics in the vicinities of the Dirac points \( \mathbf{K} = 4\pi \hat{y}/3\sqrt{3}a_0 \) and \( \mathbf{K}' = -4\pi \hat{y}/3\sqrt{3}a_0 \). To obtain the Hamiltonian for electrons close to these points, we expand \( f(k) \) about \( \mathbf{K} \) and \( \mathbf{K}' \) and obtain \( f(k) \approx \frac{1}{2} \epsilon_0 k^2 \hat{e} \cdot \hat{e} \), with the plus and minus signs corresponding to the \( \mathbf{K} \) and \( \mathbf{K}' \) valleys respectively. Note that we have transformed to a polar coordinate system with origin at \( \mathbf{K} \) or \( \mathbf{K}' \), where \( k = |k| = (k_x^2 + k_y^2)^{1/2} \), and \( \theta_k \) is the angle \( k \) makes with the \( k_x \) axis. The function \( f(k) \) may also be expressed in terms of the graphene Fermi velocity \( v_f = 3a_0 t_1/2\hbar \approx 10^6 \text{ m/s}^{-1} \) according to \( f(k) = i\hbar v_f ke^{\pm i\theta_k}/t \parallel \).

Neglecting overlap between inequivalent atoms, we solve for the energies \( E_n(k) \) and eigenvectors \( \Psi_{nk}(r) \) of \( \mathcal{H}_0 \). The dynamics of the system will be studied using only the two lowest-energy bands whose (dimensionless) energies are given by \[21\]

$$\tilde{E}_n(k) = \frac{\tilde{E}_n(k)}{E_n(k)} = \pm \sqrt{2/|f(k)|^2 + \tilde{a}^2 + \frac{1}{2} - \tilde{\epsilon}(k)},$$

where \( \tilde{E}_n(k) \equiv E_n(k)/t_\perp \), and where \( n = \{ c, v \} \) labels the conduction band and valence band, and where we have introduced the dimensionless quantities \( \tilde{a} \equiv a/t_\perp \), \( t \equiv t_\parallel/t_\perp \), and \( \tilde{\epsilon}(k) \equiv ((4\tilde{a}^2 + 1)|f(k)|^2 + 1/4)^{1/2} \). The conduction and valence bands are shown in Fig. [2] for four
The expansion coefficients of $\Psi_{nk}$ emphasize the symmetry. The dispersion is electron-hole symmetric. At the Dirac points ($k = 0$), $E_c(k) = a$ and $E_v(k) = -a$. The band gap for a particular bias $a$ is given by

$$\Delta E(a) = \frac{2at_1}{\sqrt{4a^2 + t_1^2}}. \quad (5)$$

The expansion coefficients of $\Psi_{nk}(r)$ are

$$c_{1n}(k) = \tilde{E}_n(k) - \tilde{a},$$
$$c_{2n}(k) = i\tilde{f}^* \overline{c_{1n}(k)},$$
$$c_{3n}(k) = \frac{\tilde{c}_{2n}(k)c_{4n}(k)}{\tilde{E}_n(k) - \tilde{a}},$$
$$c_{4n}(k) = c_{1n}(k)^2 - |c_{2n}(k)|^2. \quad (6)$$

### B. Connection elements

We treat the carrier-field interaction using the length gauge Hamiltonian $H_L = -e\mathbf{E}(t) \cdot \mathbf{r}$, where $e = -|e|$ is the electron charge, $\mathbf{E}(t)$ is the (classical) electric field of the optical pulse at the graphene, and $\mathbf{r}$ is the electron position operator. The density matrix equations of motion will derive in the following section require matrix elements of $\mathbf{r}$. Following Blount[33], we have

$$\langle n| \mathbf{r} | mk \rangle = i\delta_{nm} \nabla_k \delta(|k - k'|) + \delta(|k - k'|) \xi_{nm}(k), \quad (7)$$

where we have defined the connection elements

$$\xi_{nm}(k) \equiv \frac{(2\pi)^2}{\Omega} \int d^3\mathbf{r} u^*_{nk}(\mathbf{r}) \nabla_k u_{mk}(\mathbf{r}), \quad (8)$$

where $\Omega$ is the area of a real-space unit cell, and the integration is over $\Omega$, and over $-\infty < z < \infty$ perpendicular to the plane. The cell-periodic function $u_{nk}(\mathbf{r})$ is defined by

$$\Psi_{nk}(\mathbf{r}) = e^{ik \cdot \mathbf{r}} u_{nk}(\mathbf{r}), \quad (9)$$

upon which we impose the orthonormality condition

$$\int d^3\mathbf{r} u^*_{nk}(\mathbf{r}) u_{mk}(\mathbf{r}) = \frac{\Omega}{(2\pi)^2} \delta_{nm}, \quad (10)$$

from which we obtain, in the nearest-neighbor tight-binding approximation, the normalization factor

$$A_n(k) = \sqrt{\frac{\Omega}{2\pi}} \left( \sum_i |c_{in}(k)|^2 \right)^{-1/2}. \quad (11)$$

The connection element for transitions between the conduction band and valence band is [34]

$$\xi_{cv}(k) = i \frac{(2\pi)^2}{\Omega} A^*_n(k) A_c(k) \times \sum_i (c^*_{ic}(k) \nabla_k c_{iv}(k) - i\mathbf{r}_i c^*_{ic}(k) c_{iv}(k)). \quad (12)$$

In what follows, we neglect the term proportional to $\mathbf{r}_i$ in Eq. (12), as we neglect terms of similar magnitude when we expand $f(k)$ to first-order in $k$ about the Dirac points. Eq. (12) can be evaluated analytically, but as the expression is very long we do not present it here explicitly. We find that when $f(k)$ is taken to first-order, the connection element takes the form

$$\xi_{cv}(k) = iA(a,k)\hat{k} + B(a,k)\hat{\theta}_k, \quad (13)$$

where $\hat{k}$ and $\hat{\theta}_k$ are the standard unit vectors in our polar coordinate system, and where the plus and minus signs correspond to the $K$ and $K'$ valleys respectively. Here, $A$ and $B$ are real, positive functions that depend only on the magnitude of $k$ and the external bias $a$. We plot $A$ and $B$ as functions of $k$ in Fig. [5] for an example bias of $a = 100$ meV. As can be seen, $A = B = 0$ at $k = 0$ (i.e. at the Dirac points) [35]. There is second $k$ for which $A = B$, whose value depends on the external bias. In the large-$k$ limit, $B \sim 1/k$, just as it does in monolayer graphene [30]. Biased bilayer graphene is unlike monolayer graphene or unbiased bilayer graphene ($a = 0$) in that $A$ is nonzero [32, 36, 37]. In the limit of large $k$, $A \sim 1/k^3$. Under $a \rightarrow -a$, $A$ is unchanged but $B \rightarrow -B$. We will return these functions in Sec. [11] when we discuss the results of our calculations. As we shall see, it is the non-vanishing of $A$, and the crucial sign difference between the $K$ and $K'$ valleys that unlocks the possibility of valley-polarization.

### C. Equations of motion

The Heisenberg equations of motion for the reduced density operator $\rho$ in the basis of Bloch states $|n\mathbf{k}\rangle$ in
the relaxation time approximation are \[36\]

\[
\frac{d\rho_{nm}(k)}{dt} = -i\omega_{nm}(k)\rho_{nm}(k) - \frac{e}{\hbar}E(t) \cdot \nabla_k \rho_{nm}(k)
\]

\[
i\frac{e}{\hbar}E(t) \cdot \sum_l \left[ \xi_{nl}(k)\rho_{lm}(k) - \xi_{lm}(k)\rho_{nl}(k) \right] - \gamma_{nm}(k)\left[\rho_{nm}(k) - \rho_{nm}^{eq}(k)\right],
\] (14)

where \(\hbar \omega_{nm}(k) = E_n(k) - E_m(k)\), \(\rho_{nm}^{eq}(k) = \rho_{nm}^{eq}\) is the equilibrium density matrix, and \(\gamma_{nm}(k)\) is a matrix of scattering rates. If \(n = m\), we refer to \(\gamma_{nm}(k)\) as the intraband scattering rate. If \(n \neq m\), we refer to \(\gamma_{nm}(k)\) as the interband decoherence rate. Using first-order perturbation theory, we solve Eq. (14) subject to the initial conditions

\[
\begin{cases}
\rho_{nm}^{eq}(k) = f_n(k) \\
\rho_{nm}^{(1)}(k) = 0
\end{cases}
\] at \(t = -\infty\),

(15)

where \(f_n(k) = f_n\) is the Fermi-Dirac distribution. To zeroth-order we obtain \(\rho_{nm}^{(0)}(k) = f_n(k)\). Using standard techniques, the first-order result for the off-diagonal density matrix element at time \(t\) is found to be

\[
\rho_{cc}^{(1)}(k) = e^{-(i\omega_{cc}(k)+\gamma_{cc}(k))t} \int_{-\infty}^{t} i\frac{e}{\hbar}E(t')
\cdot \xi_{cc}(k)\left[f_c(k) - f_e(k)\right]e^{(i\omega_{cc}(k)+\gamma_{cc}(k))t'} dt'.
\] (16)

We are interested in the total electron populations around the \(K\) and \(K'\) valleys, not the \(k\)-space distributions. Therefore, when calculating the populations to second-order, we neglect the second term in Eq. (14) because this term simply redistributes carrier momentum within each valley. We find that the second-order contribution to the conduction band population is

\[
\rho_{cc}^{(2)}(k) = e^{-\gamma_{cc}(k)t} \int_{-\infty}^{t} i\frac{e}{\hbar}E(t')
\cdot \left[\xi_{cc}(k)\rho_{cc}^{(1)}(k) - \xi_{cc}(k)\rho_{cc}^{(1)}(k)\right]e^{\gamma_{cc}(k)t'} dt',
\] (17)

where \(\xi_{cc}(k) = \xi_{cc}^*\). The \(m\)th-order contribution to the carrier density in the conduction band about each Dirac point is given by

\[
n_{cc}^{(m)}(t) = \frac{2}{(2\pi)^2} \int \rho_{cc}^{(m)}(k) d^2k,
\] (18)

where \(V = \{K, K'\}\) labels the valley, the factor of two accounts for spin degeneracy, and the integration is only in the vicinity of the particular Dirac point. The zeroth-order contribution gives the thermal carrier density, while the second-order contribution gives the injected carrier density.

D. Calculating the injected carrier density

To evaluate the injected carrier density, we must specify the electric field of the optical pulse. It is well known that a valley-polarization can be induced in biased bilayer graphene using circularly polarized light \[14\]. In Appendix A, we discuss the possibility of using light of a more general elliptical polarization, but here we limit our discussion to circularly polarized light as we find circularly polarized light to always be optimal. We take \(E(t)\) to be a right-hand circularly polarized Gaussian pulse with central frequency \(\omega\), amplitude \(E_0\), and pulse duration \(t_p\). It will be convenient to express the field in polar coordinates. With the origin of our \(k\)-space coordinate system at either \(K\) or \(K'\),

\[
E(t) = E(t)(\hat{k} + i\theta) e^{-i(\omega t - \theta \frac{\pi}{2})} + c.c.,
\] (19)

where \(E(t) = E_0 e^{-t^2/t_p^2}\). With Eqs. (17) and (18), one has, for the injected carrier density,

\[
n_{cc}^{(2)}(t) = \frac{2}{(2\pi)^2} \int\int kdk d\theta_k e^{-\gamma_{cc}(k)t} \int_{-\infty}^{t} i\frac{e}{\hbar}E(t')
\cdot \left[\xi_{cc}(k)\rho_{cc}^{(1)}(k) - \xi_{cc}(k)\rho_{cc}^{(1)}(k)\right]e^{\gamma_{cc}(k)t'} dt' (20)

Since the energy bands are isotropic, we assume that the scattering rates are as well so that we may write \(\gamma_{nm}(k) = \gamma_{nm}(\omega)\). We may then pull \(e^{-\gamma_{cc}(k)t}\) through the angular integral, and interchange the order of the temporal and angular integration. Performing the angular integral first, one can show (after significant work)
where we have suppressed explicit \( k \)-dependence, and where the upper and lower operations correspond to the \( K \) and \( K' \) valleys respectively. The Greek letters represent the quantities

\[
\lambda_\mp \equiv \frac{t_p}{2} (\gamma_{cc} + i(\omega_{cc} + \omega)),
\]

\[
\delta_\mp \equiv 2\lambda_\mp + \gamma_{cc}t_p,
\]

where the subscripts indicate whether the term is resonant (−) or anti-resonant (+) with the optical field.

We are faced with the following integral:

\[
J(t) = \int_{-\infty}^{t} dt' \left[ \text{erf}\left( \frac{t'}{t_p} + \lambda \right) + 1 \right] \times \exp\left( \lambda^2 + \frac{t'}{t_p} \delta - \frac{t'^2}{t_p^2} \right).
\]

Making the replacement \( \eta = t'/t_p + \lambda \), and completing the square in the exponential, Eq. (24) may be re-expressed as

\[
J(t) = t_p \exp(\beta) \int_{-\infty}^{t/t_p + \lambda} d\eta \left[ \text{erf}(\eta) + 1 \right] \times \exp\left( -(\eta + c)^2 \right),
\]

where

\[
\beta \equiv \lambda^2 + \frac{1}{4} \delta^2,
\]

\[
c \equiv -\lambda - \frac{1}{2} \delta.
\]

To the best of our knowledge, the integral in Eq. (25) cannot be performed analytically. However, an analytic result is provided by Ref. [38] in the limit \( t \to \infty \). In Appendix B, we address the subtleties of taking this limit given that \( \lambda \) is in general complex. We find

\[
J(t \to \infty) = \sqrt{\pi} t_p \exp(\beta) \text{erfc}\left( \frac{c}{\sqrt{2}} \right).
\]

For times \( t = t_f \) by which the integrand of Eq. (25) has decayed essentially to zero, we may make the approximation \( J(t_f) \approx J(t \to \infty) \). This approximation limits the applicability of this analytic result to times \( t_f \) that are at least several times the pulse duration \( t_p \). This should not pose a problem in practice, because one would likely only wish to manipulate the valley-polarized carriers after the exciting pulse has passed. The pulse durations we will be considering are only on the order of 10s of femtoseconds, so this delay is very small. Using Eq. (27) in Eq. (21), we obtain the key result of this section: the carrier density injected into valley \( V \) at time \( t = t_f \),

\[
n_V^{(2)}(t) = \frac{1}{2} \frac{e^2}{\hbar^2} E_0^2 t_p \int_0^{\infty} dk \, k e^{-\gamma_{cc}k} \left[ (A \pm B)^2 \exp(\beta_-) \text{erfc}\left( \frac{c_-}{\sqrt{2}} \right) + (A \mp B)^2 \exp(\beta_+) \text{erfc}\left( \frac{c_+}{\sqrt{2}} \right) \right] + c.c.
\]
III. RESULTS

We wish to compare the carrier densities in the conduction bands of the $K$ and $K'$ valleys shortly after excitation by a pulse of circularly polarized light. Depending on the temperature, a significant contribution to the carrier density can come from the zeroth-order thermal population $n_0^{(0)}$. We will consider this in more detail in Sec. III D, but for now we focus solely on the second-order response. To this end, we define the second-order valley-polarization $\mathcal{P}^{(2)}$ to be the difference between the carrier densities injected around the $K$ and $K'$ valleys at time $t = t_f$, normalized by their sum:

$$\mathcal{P}^{(2)} = \frac{n_{K}^{(2)} - n_{K'}^{(2)}}{n_{K}^{(2)} + n_{K'}^{(2)}}.$$ (29)

When the system is completely valley-polarized in favor of $K$ ($K'$) electrons, $\mathcal{P}^{(2)} = 1$ ($-1$). If the system is not valley-polarized, $\mathcal{P}^{(2)} = 0$. Throughout this section, we consider a (Gaussian) right-hand circularly polarized pulse, so we expect the system to be valley-polarized in favor of $K$ electrons.

We vary the external bias and pulse frequency to examine how the valley-polarization depends on these two parameters. We consider the frequency-bias pairs that result in the strongest valley-polarization to be the optimal operating parameters for valleytronic devices. Unless otherwise stated, we take the temperature to be 300 K and the chemical potential to be at the charge neutrality point ($\mu = 0$). Because of electron-hole symmetry, this choice of $\mu$ leads to identical results for both electron and hole populations. For this reason, we discuss only electron populations. In what follows, we restrict ourselves to external biases greater than 50 meV, because for lower energies, the pulse duration can be shorter than a single period. In Fig. 4 we plot $1 - \mathcal{P}^{(2)}$, that is, the deviation of the valley-polarization from perfect polarization ($\mathcal{P}^{(2)} = 1$), on a logarithmic scale as a function of the external bias $a$ and the central photon energy $h\omega$ of the exciting right-hand circularly polarized Gaussian pulse [40]. Lighter colors correspond to stronger valley-polarizations. The strongest valley-polarizations are concentrated in the low-frequency—low-bias regime, along the line $h\omega = 2a$ (indicated by a straight dashed line). The valley-polarization falls off on either side of $h\omega = 2a$. The valley-polarization degrades and broadens with increasing frequency and bias. The innermost contour corresponds to $\mathcal{P}^{(2)} > 0.97$. The two next-to-innermost contours correspond to $\mathcal{P}^{(2)} > 0.95$ and 0.90 respectively. Over the parameter space considered in Fig. 4, the valley-polarization ranges from 0.10 $\leq \mathcal{P}^{(2)} \leq 0.98$. The optimal operating frequency-bias pair occurs for $(h\omega, a) \approx (241, 127)$ meV. Note however that both the optimal operating pair and the corresponding valley-polarization depend on the pulse duration and decoherence time.

To help develop the main ideas, in Sec. III A we work with a simplified model in which we neglect intraband scattering, accounting only for interband decoherence. In Sec. III B we will introduce intraband scattering via optical phonons. In particular, we will allow for intervalley scattering, which acts to reduce the valley-polarization, and for intravalley scattering, which limits the intervalley process. In Sec. III C we will examine the effects of varying the pulse duration and decoherence time. In Sec. III D we will examine the effects of the thermal electron populations on the valley-polarization.

A. Without intraband scattering

In this section, we examine the second-order valley-polarization when there is no intraband scattering, but where there is interband decoherence. Thus, we let $\gamma_{cc}(k) = 0$ and $\gamma_{ccv}(k) = 1/\tau_0$, where $\tau_0$ is a phenomenological decoherence time. We take $\tau_0 = 30$ fs and the pulse duration $t_p = 50$ fs. We restrict ourselves to photon energies greater than 50 meV, because for lower energies, the pulse duration can be shorter than a single period. In Fig. 4 we plot $1 - \mathcal{P}^{(2)}$, that is, the deviation of the valley-polarization from perfect polarization ($\mathcal{P}^{(2)} = 1$), on a logarithmic scale as a function of the external bias $a$ and the central photon energy $h\omega$ of the exciting right-hand circularly polarized Gaussian pulse [40]. Lighter colors correspond to stronger valley-polarizations. The strongest valley-polarizations are concentrated in the low-frequency—low-bias regime, along the line $h\omega = 2a$ (indicated by a straight dashed line). The valley-polarization falls off on either side of $h\omega = 2a$. The valley-polarization degrades and broadens with increasing frequency and bias. The innermost contour corresponds to $\mathcal{P}^{(2)} > 0.97$. The two next-to-innermost contours correspond to $\mathcal{P}^{(2)} > 0.95$ and 0.90 respectively. Over the parameter space considered in Fig. 4, the valley-polarization ranges from 0.10 $\leq \mathcal{P}^{(2)} \leq 0.98$. The optimal operating frequency-bias pair occurs for $(h\omega, a) \approx (241, 127)$ meV. Note however that both the optimal operating pair and the corresponding valley-polarization depend on the pulse duration and decoherence time.

Except at very low biases, the optimal operating frequencies do not coincide with the band gap energy (indicated by the lower dashed curve). In fact, the valley-polarization appears to be somewhat suppressed for frequencies resonant with the band gap energy. The upper dashed curve in Fig. 4 gives the energy of the next-lowest transition, involving the high-energy bands that we have neglected [48]. Energies above this curve will induce significant transitions between bands other than the two low-energy bands we have considered. This curve is given explicitly by

$$\Delta E_{\text{HB}}(a) = (a^2 + t_\perp^2)^{1/2} + a,$$ (30)

which is strictly greater than $2a$. 

To help develop the main ideas, in Sec. III A we work with a simplified model in which we neglect intraband scattering, accounting only for interband decoherence. In Sec. III B we will introduce intraband scattering via optical phonons. In particular, we will allow for intervalley scattering, which acts to reduce the valley-polarization, and for intravalley scattering, which limits the intervalley process. In Sec. III C we will examine the effects of varying the pulse duration and decoherence time. In Sec. III D we will examine the effects of the thermal electron populations on the valley-polarization.

A. Without intraband scattering

In this section, we examine the second-order valley-polarization when there is no intraband scattering, but where there is interband decoherence. Thus, we let $\gamma_{cc}(k) = 0$ and $\gamma_{ccv}(k) = 1/\tau_0$, where $\tau_0$ is a phenomenological decoherence time. We take $\tau_0 = 30$ fs and the pulse duration $t_p = 50$ fs. We restrict ourselves to photon energies greater than 50 meV, because for lower energies, the pulse duration can be shorter than a single period. In Fig. 4 we plot $1 - \mathcal{P}^{(2)}$, that is, the deviation of the valley-polarization from perfect polarization ($\mathcal{P}^{(2)} = 1$), on a logarithmic scale as a function of the external bias $a$ and the central photon energy $h\omega$ of the exciting right-hand circularly polarized Gaussian pulse [40]. Lighter colors correspond to stronger valley-polarizations. The strongest valley-polarizations are concentrated in the low-frequency—low-bias regime, along the line $h\omega = 2a$ (indicated by a straight dashed line). The valley-polarization falls off on either side of $h\omega = 2a$. The valley-polarization degrades and broadens with increasing frequency and bias. The innermost contour corresponds to $\mathcal{P}^{(2)} > 0.97$. The two next-to-innermost contours correspond to $\mathcal{P}^{(2)} > 0.95$ and 0.90 respectively. Over the parameter space considered in Fig. 4, the valley-polarization ranges from 0.10 $\leq \mathcal{P}^{(2)} \leq 0.98$. The optimal operating frequency-bias pair occurs for $(h\omega, a) \approx (241, 127)$ meV. Note however that both the optimal operating pair and the corresponding valley-polarization depend on the pulse duration and decoherence time.

Except at very low biases, the optimal operating frequencies do not coincide with the band gap energy (indicated by the lower dashed curve). In fact, the valley-polarization appears to be somewhat suppressed for frequencies resonant with the band gap energy. The upper dashed curve in Fig. 4 gives the energy of the next-lowest transition, involving the high-energy bands that we have neglected [48]. Energies above this curve will induce significant transitions between bands other than the two low-energy bands we have considered. This curve is given explicitly by

$$\Delta E_{\text{HB}}(a) = (a^2 + t_\perp^2)^{1/2} + a,$$ (30)

which is strictly greater than $2a$. 

To help develop the main ideas, in Sec. III A we work with a simplified model in which we neglect intraband scattering, accounting only for interband decoherence. In Sec. III B we will introduce intraband scattering via optical phonons. In particular, we will allow for intervalley scattering, which acts to reduce the valley-polarization, and for intravalley scattering, which limits the intervalley process. In Sec. III C we will examine the effects of varying the pulse duration and decoherence time. In Sec. III D we will examine the effects of the thermal electron populations on the valley-polarization.
as a function of the external bias $P$ contours correspond to stronger valley-polarizations. From lightest to darkest, the rule is in general not exact, except for very specific states

$K$ valley-contrasting optical selection rule in favor of the $K$ while in the $A$ tional to the sum of the (positive) functions $K’$ contribution. Again, the plus and minus signs corre-

sign indicates that we are considering only the resonant introduced in Sec. II B, and where the approximation $A$ is dominated piece of the carrier-field interaction is instead proportional to $(A \mp B)$, such that when $A = B$ the selection rule favors $K’$.

The ratio $A/B$ gives a measure of the “exactness” of the optical selection rule ($A/B = 1$ when $A = B$). In Fig. [3] we plot the quantity $A/B$ as a function of $k$ for three different biases (dashed curves), along with the corresponding conduction band energies, $E_c(k)$ (solid curves). For all biases, $A/B = 1$ at $k = 0$, indicating an exact selection rule. As one moves away from $k = 0$, the selection rules softens as the $A/B$ ratio deviates from unity. The $A/B$ ratio reaches its maximum at the band-minimum. The $A/B$ ratio then decays to zero as $k \rightarrow \infty$. However, as the $A/B$ ratio decays, it passes through $A/B = 1$ at some $k = k_1$. In other words, for any given bias, the optical selection rule is exact at precisely two values of $k$: $k = 0$ and $k = k_1$. If we wish to achieve a strong valley-polarization, we should try to induce excitations at these very $k$. However, as was discussed briefly in Sec. II B, both $A$ and $B$ vanish at $k = 0$. The carrier-field interaction [Eq. (31)] therefore vanishes at $k = 0$ (for both $K$ and $K’$), and so there is no carrier injection at $k = 0$. Therefore, in what follows, we focus our discussion on the states near $k = k_1$.

In Fig. [5] we saw that the $A/B$ ratio peaks at the band-minimum. In fact, one can show that

$$A/B = \frac{a}{E_c(k)}.$$  

(32)

Thus, $A/B = 1$ when $E_c(k) = a$, which occurs precisely at $k = 0$ and $k = k_1$, where

$$k_1 = \frac{4a}{3a_0l_\parallel} = \frac{2a}{\hbar v_f}.$$  

(33)

FIG. 4. (Color online) Deviation of the second-order valley-polarization from perfect polarization $\log_{10}(1 - P(2))$ as a function of the external bias $a$ and central photon energy $\hbar \omega$ of the exciting Gaussian pulse. Lighter regions correspond to stronger valley-polarizations. From lightest to darkest, the contours correspond to $P(2) = 0.97, 0.95, 0.9, 0.8, 0.7,$ and $0.5$. The pulse duration is $\tau_p = 50$ fs, and the valley-polarization is evaluated long after the pulse has passed. Intraband scattering is neglected ($\gamma = 0$), and the interband decoherence time is taken to be $\tau_0 = 30$ fs. The straight dashed line is the line $\hbar \omega = 2a$. The lower dashed curve is the band gap energy [Eq. (1)], while the upper dashed curve is the energy of the next-lowest transition involving the higher-energy bands that we have neglected [Eq. (30)].

FIG. 5. $A/B$ ratio for several biases (dashed curves, leftmost axis), and corresponding conduction band energies $E_c(k)$ for the same biases (solid curves, rightmost axis) as a function of $k$. From lightest to darkest, $a = 100, 200,$ and $300$ meV. The dashed horizontal line indicates $A/B = 1$. We now examine in detail the origins of the main features of Fig. 4 and in particular, why the valley-polarization is generally greatest along the line $\hbar \omega = 2a$.

First, note that the first-order interband coherence $\xi_{cv}(k)$ is proportional to the carrier-field interaction $\xi_{cv}(k) \cdot E(t)$ [see Eq. (16)]. If $\xi_{cv}(k) \cdot E(t)$ can be forced to zero in one valley but not the other, a strong valley-polarization is expected. For a right-hand circularly polarized field with central frequency $\omega$, it can be easily shown using Eqs. [15] and [19] that the carrier-field interaction is given by

$$\xi_{cv}(k) \cdot E(t) \approx iE(t)e^{-i(\omega t-\theta_k)}(A(a,k) \pm B(a,k)),$$  

(31)

where $A$ and $B$ are the real, positive functions that were introduced in Sec. II B and where the approximation sign indicates that we are considering only the resonant contribution. Again, the plus and minus signs correspond to the $K$ and $K’$ valleys respectively. We see from Eq. (31) that in the $K$ valley, the interaction is proportional to the sum of the (positive) functions $A$ and $B$, while in the $K’$ valley the interaction is proportional to the difference. When $A = B$, Eq. (31) amounts to a valley-contrasting optical selection rule in favor of the $K$ valley. Because $A$ and $B$ are $k$-dependent, the selection rule is in general not exact, except for very specific states within each valley (namely, states for which $A = B$). Note that if one uses light of the opposite helicity, the dominant piece of the carrier-field interaction is instead proportional to $(A \pm B)$, such that when $A = B$ the selection rule favors $K’$.
To target these states, we must tune the frequency of the exciting field such that it is resonant with interband transitions at \( k = k_1 \). Due to electron-hole symmetry, the transition energy at \( k_1 \) is simply \( 2E_c(k_1) \), and so the optimal operating frequency for a particular bias is expected to be given by

\[
h\omega = 2E_c(k_1) = 2a,
\]

(34)

which explains why the optimal operating frequency-bias pairs in Fig. 3 lie along the line \( h\omega = 2a \). In monolayer graphene with a staggered sublattice potential, one finds an exact selection rule at the Dirac points \( (k = 0) \) [14]. In contrast, in biased bilayer graphene, because \( A \) and \( B \) both vanish at exactly \( \mathbf{K} \) and \( \mathbf{K}' \), there are no carriers injected at \( k = 0 \) and therefore there is no valley-contrast optical selection rule at \( k = 0 \). Rather, the optical selection rule is found in a ring of states with radius \( k = k_1 \) surrounding the Dirac points. The existence of such a \( k \) value seems to have gone unnoticed in the literature. However, evidence for this result can be seen in Fig. 2 of Yao et al. [14].

The valley-polarization is robust to deviations from \( h\omega = 2a \). In Fig. 3, the innermost contour corresponded to \( \mathcal{P}^{(2)} > 0.97 \). Consider operating near the optimal frequency and bias \( (h\omega, a) = (241, 127) \) meV, for which \( \mathcal{P}^{(2)} = 0.98 \). For fixed \( h\omega = 241 \) meV, external biases within the range \( 110 < a < 161 \) meV yield \( \mathcal{P}^{(2)} > 0.97 \). For fixed \( a = 127 \) meV, central photon energies within \( 198 < h\omega < 284 \) meV yield \( \mathcal{P}^{(2)} > 0.97 \). Similarly, the third-to-innermost contour corresponded to \( \mathcal{P}^{(2)} > 0.90 \). For fixed \( h\omega = 241 \) meV, \( 84 < a < 230 \) meV yields \( \mathcal{P}^{(2)} > 0.90 \). For fixed \( a = 127 \) meV, \( 146 < h\omega < 396 \) meV yields \( \mathcal{P}^{(2)} > 0.90 \). Thus, deviations of several 10s of meV in either frequency or bias from the optimal operating pair still yield valley-polarizations well-over 90%. If one cannot target the optimum precisely, it is better to err towards larger frequencies and biases.

In Fig. 3, the optimal operating frequencies coincide with the band gap for small biases. This is because the band gap energy \( \Delta E(a) \rightarrow 2a \) in the limit of small \( a \) [see Eq. (5)]. For more moderate biases, we see from Fig. 4 that the valley-polarization appears somewhat reduced at frequencies close to the band gap energy. This is because for any given bias, the \( A/B \) ratio reaches its maximum at the band edge and therefore in general leads to a poor valley-polarization [see Fig. 5 or Eq. (22)]. For frequencies less than the band gap energy, carriers are still predominantly excited at the band edge due to the finite bandwidth of the Gaussian pulse. Only once the central frequency exceeds the band gap energy do carriers away from the band edge begin to dominate the response. The valley-polarization is therefore not reduced near the band gap, rather, the valley-polarization “stalls” at the band gap energy as one sweeps upwards in frequency. This result is in agreement with Fig. 3 of Ref. [30], where a strong valley-polarization was calculated for photon energies close to resonance with the band gap energy \( \Delta E(a) \approx 2a \) for a small bias of \( 2a = 20 \) meV.

In Fig. 4 the stripe of optimal operating frequency-bias pairs broadens with increasing \( a \) and \( h\omega \). This can be understood as follows. For a given bias \( a \),

\[
\frac{d}{dE_c(k)} \left( \frac{A}{B} \right) = -\frac{a}{E^2_c(k)}. \quad (35)
\]

Consider operating under optimal conditions: \( h\omega = 2a = 2E_c(k_1) \). Eq. (35) becomes

\[
\frac{d}{d(h\omega)} \left( \frac{A}{B} \right) = -\frac{1}{2a}. \quad (36)
\]

As the bias is increased, small deviations in \( h\omega \) affect the \( A/B \) ratio to a lesser and lesser degree. Thus, the valley-polarization broadens, because the system becomes less sensitive to small deviations from the optimal configuration.

For a couple of reasons, one can never achieve a perfect valley-polarization. First, targeting states for which \( A/B = 1 \) is exclusively impossible. Even when the pulse duration is very long, the decoherence time results in linewidth broadening. This means that even when the central photon energy of the pulse is \( h\omega = 2a \), carriers will be excited not just at \( k = k_1 \), but also in the vicinity of \( k = k_1 \) where \( A \neq B \) and the valley-contrast optical selection rule is inexact. In Fig. 4, we observed that the valley-polarization degrades along \( h\omega = 2a \) with increasing frequency and bias. This degradation is only observed in the presence of finite decoherence times. Second, the valley-polarization is limited by the presence of the anti-resonant piece of the carrier field interaction. In contrast to the resonant term [Eq. (31)], the anti-resonant term is proportional to \( (A \mp B) \), meaning that the selection rule can never be truly exact. That is, when \( A = B \) the resonant term vanishes but the anti-resonant term does not. For the particular pulse duration and decoherence time chosen in Fig. 4, the difference between the valley-polarization obtained if one includes both the resonant and anti-resonant terms, and the valley-polarization obtained when one includes just the resonant terms, can be as large as \( \Delta \mathcal{P}^{(2)} = 0.05 \) along the line \( h\omega = 2a \) (over the parameter space considered). The greatest differences occur at the edges of the parameter space. Close to the optimal operating frequency-bias pair, the difference is significantly smaller, about \( \Delta \mathcal{P}^{(2)} = 0.01 \). Since the valley-polarization can reach up to \( \mathcal{P}^{(2)} = 0.98 \), a difference of 0.01 is significant. The anti-resonant term must therefore be included to obtain accurate results. Taken together, a finite decoherence time and the presence of the anti-resonant term ensures that the valley-polarization is never complete. However, as we shall discuss in Sec. [11C], the impact of these two effects can be reduced by increasing the decoherence time and pulse duration.
B. With intraband scattering

We now improve upon our model by accounting for intraband scattering processes which act to reduce the valley-polarization. Intervallen scattering (IVS) is a process in which an electron in (say) the K valley scatters to \( K' \), degrading the valley-polarization. As was discussed in the beginning of Sec. [11], the relevant relaxation pathways for photoexcited carriers in bilayer graphene are carrier-carrier and carrier-phonon scattering. Due to momentum conservation, IVS via carrier-carrier interactions is very weak [39]. We find it crucial, however, to account for carrier-phonon interactions as IVS via optical phonons can significantly reduce the valley-polarization when the frequency of the exciting pulse is large.

Carrier-phonon scattering in monolayer graphene (MLG) has been studied extensively [40–51]. Unfortunately, the literature on carrier-phonon scattering in bilayer graphene (BLG), and in particular in biased bilayer graphene (BBLG), is somewhat lacking and there is no clear consensus on the dominant scattering modes. In this work, we treat scattering to be the same as in MLG where optical phonons dominate at room temperature 50. This is certainly an approximation, for it is not clear whether scattering in BBLG is similar to scattering in MLG. For instance, a 2011 study has suggested that in contrast to MLG, low-energy acoustic phonons dominate in unbiased BLG, while optical phonon scattering is highly suppressed 52. However, it is unclear if this result holds for a biased bilayer: a 2015 study found that optical phonons are the dominant scattering mode in BBLG at biases above 150 meV [53]. If it turns out that acoustic phonons play an important role in BBLG transport, our calculations can be easily modified to account for them.

To incorporate intraband scattering into our model, we adopt the formalism of Ref. [54], where scattering rates \( \Gamma(k) \) are derived from a Fröhlich Hamiltonian up to the second order Born-Markov approximation. We shall use the formalism of Ref. [54], where scattering rates are isotropic and we may write \( \Gamma(k) = \Gamma_{cc}(k) = \Gamma_{DS}(k) \) and \( \Gamma_{IVS}(k) = \Gamma_{IVS}(k) \), noting that their precise values will have little effect on our conclusions.

Following Ref. [51], we have taken the phonons to be dispersionless near the \( K \) and \( \Gamma \) points so that \( \hbar\omega_K = 160 \text{ meV} \) and \( \hbar\omega_\Gamma = 196 \text{ meV} \). We treat the phonons as thermal baths at 300 K so that the phonon populations \( n_K \) and \( n_\Gamma \) are given by Bose-Einstein distributions at energy \( \hbar\omega_K \) and \( \hbar\omega_\Gamma \) respectively.

In Fig. [6] we plot \( \Gamma_{IVS}(k) \) and \( \Gamma_{DS}(k) \) as functions of energy for an example bias of \( a = 205 \text{ meV} \). As is demonstrated in the figure, intraband optical-phonon-scattering is forbidden unless an excited electron can afford to lose the energy of a phonon and remain in the conduction band. That is, \( \Gamma_{IVS}(k) \) and \( \Gamma_{DS}(k) \) are, respectively, strictly zero unless

\[
E_c(k) \geq \hbar\omega_K + \Delta(E(a))/2 \equiv E_{IVS/DS}(a),
\]

where \( \Delta(E(a)) = E_{valence} - E_{valence}(a) \), and \( E_{IVS/DS}(a) \) can be clearly seen. Since \( \hbar\omega_K < \hbar\omega_\Gamma \), IVS becomes energetically possible before DS. Note that due to electron-hole symmetry, the photon energy required to excite an electron from the valence band to an energy \( E_{IVS}(a) \) or \( E_{DS}(a) \) in the conduction band is given by \( \hbar\omega = 2E_{IVS}(a) \) or \( \hbar\omega = 2E_{DS}(a) \) respectively.

The scattering-out rates \( \Gamma_{IVS}(k) \) and \( \Gamma_{DS}(k) \) diverge when \( E_c(k) = E_{IVS}(a) \) and \( E_c(k) = E_{DS}(a) \), respectively. This is due to the Dirac delta functions in Eqs. [37] and [38]. Away from these points, the scattering rates are well-behaved and on the order of a few 10s of ps\(^{-1} \), which is consistent with rates found in Refs. [50–59]. These scattering rates will be incorporated into our model by setting \( \gamma_{cc}(k) = \Gamma_{IVS}(k) + \Gamma_{DS}(k) \). However, due to the approximations we made when evaluating the carrier density (Sec. [II D]), a divergent scattering rate results in numerical difficulties unless \( t_f \) is very large. In particular, large values of \( \gamma_{cc}(k) \) push the peak of the integrand of Eq. [24] towards \( t = \infty \). In requiring that

\[
\Gamma_{IVS}(k) = \frac{2\pi}{\hbar N} \sum_q \frac{1}{2} g_{n}^2 (1 - \rho_{cc}^0(k - q)) (n_K + 1)
\]

\[
\times |\delta[E_c(k - q) - E_c(k) + \hbar\omega_K]|, \quad (37)
\]

\[
\Gamma_{DS}(k) = \frac{2\pi}{\hbar N} \sum_q 2g_{c}^2 (1 - \rho_{cc}^0(k - q)) (n_T + 1)
\]

\[
\times |\delta[E_c(k - q) - E_c(k) + \hbar\omega_T]|, \quad (38)
\]
the integrand decay sufficiently to zero by the observation time $t_f$, divergent $\gamma_{cc}(k)$ forces us longer and longer $t_f$ at which to evaluate the carrier density. To address this problem, we limit $\Gamma_{DS}$ by setting $\Gamma_{IVS} = 5\text{fs}$. We find that although changing the value of $\Gamma_{max}$ affects the high-frequency results (where scattering is strong), it has very little effect in the optimal operating region.

1. Intervalley scattering

To begin, let us simplify things and neglect down-scattering by setting $\Gamma_{DS}(k) = 0$. We set the intraband scattering rate to $\gamma_{cc}(k) = \Gamma_{IVS}(k)$, and modify the interband decoherence rate according to $\gamma_{cc}(k) = 1/\tau(k)$, where

$$\frac{1}{\tau(k)} = \frac{1}{\tau_0} + \frac{\Gamma_{IVS}(k)}{2},$$

where the factor of $1/2$ arises from the usual relationship between decoherence and population decay rates. Because we treat the $K$ and $K'$ valleys as disconnected in $k$-space, and we do not account for carriers scattering-in, we cannot keep track of IVS explicitly. That is, $\gamma_{cc}(k)$ simply acts as a population decay rate for electrons in each valley independently [see Eq. (28)]. We therefore require a systematic way to keep track of the number of carriers which IVS from each valley. Once we have that, we can then add those carriers into the opposing valley before computing the valley-polarization.

To this end, we calculate the carrier density in each valley twice: once subject to decay from IVS, and a second time without population decay. If IVS is the only scattering process, then the difference between these two quantities gives an estimate of the density of carriers that have intervalley scattered. Concretely, let $n_{K,IVS}$ be the carrier density injected into the $K$ valley at time $t = t_f$, calculated using Eq. (28) with $\gamma_{cc}(k) = \Gamma_{IVS}(k)$ and $\gamma_{cc}(k) = 1/\tau(k)$. As before, $n_{K}(2)$ is the carrier density injected into $K$, calculated using Eq. (28) with $\gamma_{cc}(k) = 0$ and $\gamma_{cc}(k) = 1/\tau_0$. The quantity

$$\Delta n_{K,IVS} = n_{K}(2) - n_{K}^{IVS}$$

gives the density of carriers intervalley scattered from $K$ to $K'$. The carrier density in the $K'$ valley is then modified according to

$$n_{K'} = n_{K'}^{IVS} + \Delta n_{K,IVS}.$$  

In other words, the carrier density in $K'$ is given by the carrier density remaining in $K'$ (after allowing for population decay from IVS), plus the carriers that have intervalley scattered from $K$ to $K'$. The carrier density in the $K$ valley is obtained by interchanging $K \leftrightarrow K'$ in Eqs. (41) and (42).

The valley-polarization is modified to

$$P^{(2)} = \frac{n_{K}^{(2)} - n_{K'}^{(2)}}{n_{K}^{(2)} + n_{K'}^{(2)}},$$

which we plot as a function of the external bias and central photon energy in Fig. 7(a) under the same conditions as Fig. 4 ($t_p = 50\text{fs}$, $\tau_0 = 30\text{fs}$). As was the case in Fig. 4, the strongest valley-polarizations are concentrated in the low-frequency—low-bias regime, along $\hbar\omega = 2a$. The most obvious difference between Fig. 7(a) and Fig. 4 is the emergence of a dark blue region of poor valley-polarization in the high-frequency portion of the parameter space. This is due to the presence of intervalley scattering, the onset of which is indicated by the dashed white curve corresponding to $\hbar\omega = 2\Gamma_{IVS}(a)$. Notice that the onset of IVS intersects the line $\hbar\omega = 2a$, meaning that IVS effectively cuts off high frequencies from the parameter space of optimal operating frequency-bias pairs. Unfortunately, high frequencies may be desirable due to hot-carrier multiplication, which would result in a larger population imbalance and thus a stronger valley-polarization [60, 61]. If low-energy acoustic phonon scattering turns out to play an important role in BBLG, one would expect to see an onset curve analogous to the one in Fig. 7(a) but shifted towards lower energies. In comparison to Fig. 4, the optimal operating frequency-bias pair is shifted slightly towards lower frequencies, $(\hbar\omega, a) = (236, 126)\text{meV}$, but a very strong valley-polarization of $P^{(2)} = 0.98$ is retained. The difference is slight because the optimal operating region is largely unaffected by the presence of...
intervalley scattering, but nonzero because the pulse excites over a broad energy range. In Fig. 7(a), the valley-polarization is evaluated long after the exciting pulse has passed ($t_f = 5 t_p = 250$ fs).

At energies greater than the onset of IVS, the valley-polarization can become negative, meaning that the system is valley-polarized in favor of $K'$ electrons. In fact, the entirety of the darkest shaded region in Fig. 7(a) corresponds to $P^{(2)} < 0$. This is because the IVS rate is so strong in this region that more carriers intervalley scatter from $K$ to $K'$ than remain in the $K$ valley. This result suggests that it may be possible to achieve an inverse valley-polarization, that is, a valley-polarization in favor $K'$ electrons, even though the exciting pulse is right-hand circularly polarized.

By neglecting down-scattering thus far, the results presented in Fig. 7(a) give an underestimate of the valley-polarization. Carriers that DS contribute to the valley-polarization in the same way as if they had not scattered at all because they remain in their original valley. However, down-scattered carriers will likely have scattered into states below the energy threshold $E_{IVS}(a)$ required to IVS. In other words, DS acts to limit IVS by reducing the number of carriers that are available to IVS. We are again faced with the problem of keeping track of these carriers given that $\gamma_{cc}(k)$ simply acts as a population decay rate.

2. Down-scattering

To achieve a better estimate of $P^{(2)}$, we must adjust Eqs. (40), (42) to account for down-scattering. This essentially amounts to keeping track of another contribution to the carrier density in each valley. In the exact same way we estimated the density of carriers that intervalley scattered, let us define

$$\Delta n_{K'}^{\text{DS}} = n_{K'}^{(2)} - n_{K'}^{\text{DS}}$$

(44)

to be the density of carriers scattered within the $K'$ valley. Analogously, $n_{K'}^{\text{IVS}+\text{DS}}$ is the carrier density injected into $K'$, calculated using Eq. (28) with $\gamma_{cc}(k) = \Gamma_{\text{DS}}(k)$ and $\gamma_{cc}(k) = 1/\tau(k)$, with $\tau(k)$ given by Eq. (40) under the replacement $\Gamma_{\text{IVS}}(k) \rightarrow \Gamma_{\text{DS}}(k)$. The carrier density in the $K'$ valley becomes

$$\tilde{n}_{K'}^{(2)} = n_{K'}^{\text{IVS}+\text{DS}} + \Delta n_{K'}^{\text{IVS}} + \Delta n_{K'}^{\text{DS}}.$$  

(45)

Here, $n_{K'}^{\text{IVS}+\text{DS}}$ is the carrier density obtained by allowing for decay from both IVS and DS. That is, for $\gamma_{cc}(k) = \Gamma_{\text{IVS}}(k) + \Gamma_{\text{DS}}(k)$ and $\gamma_{cc}(k) = 1/\tau(k)$, with $\tau(k)$ given by Eq. (40) under the replacement $\Gamma_{\text{IVS}}(k) \rightarrow \Gamma_{\text{IVS}}(k) + \Gamma_{\text{DS}}(k)$. In other words, the carrier density in the $K'$ valley is given by the carrier density remaining in $K'$ (after allowing for population decay from both IVS and DS), plus the carriers that IVS from $K$ to $K'$, plus the carriers that DS within $K'$.

In Fig. 7(b) we plot Eq. (43) for $P^{(2)}$, modified to include down-scattering according to the redefini-
tions above, and under the same conditions as Fig. 7(a). The general features that were observed in Fig. 7(a) are unchanged. In comparison to Fig. 7(a), the valley-polarization recovers somewhat after the onset of down-scattering, which is indicated by the upper dashed white curve corresponding to $h\omega = 2E_{DS}(a)$. However, the valley-polarization does not recover significantly enough to make this regime attractive for valleytronics. We therefore conclude that, even when accounting for DS, one should operate in the low-frequency—low-bias regime before the onset of IVS. The optimal operating frequency-bias pair is unchanged from the previous calculation: $(h\omega, a) = (236, 126)$ meV, $P^{(2)} = 0.98$. Due to our approximate scheme for estimating the number of scattered carriers, Fig. 7(b) overestimates both IVS and DS processes. It is therefore unclear whether Fig. 7(b) represents an upper bound or a lower bound on the valley-polarization, but we believe it provides a better estimate than Fig. 7(a), which should be considered a worst-case scenario.

### C. Pulse duration and decoherence time

In this section, we study the effects of varying the pulse duration $t_p$ and decoherence time $\tau_0$ on the second-order valley-polarization. As might be anticipated, the effect of increasing $t_p$ and $\tau_0$ is to improve the valley-polarization and sharpen the response, as we explore in Fig. 8. In Fig. 8(a) we fix the bias to a constant value of $a = 205$ meV, and plot the second-order valley-polarization as a function of the central photon energy $h\omega$ of the exciting pulse. We use the scattering model developed in the previous section, accounting for both inter-valley scattering (IVS) and down-scattering (DS). Thus, Fig. 8(a) can be thought of as a vertical slice of Fig. 7(b) along constant $a$. In general, the valley-polarization increases steadily as $h\omega$ approaches $2a$, stalling briefly at the band gap energy, for reasons discussed in Sec. IIIA. The valley-polarization peaks close to $h\omega = 2a$, then decays. The decay is accelerated by the presence of IVS, and an inverse valley-polarization is observed. The presence of DS counteracts the decay due to IVS, but not significantly. These general features are observed for all biases, but are particularly prominent for $a = 205$ meV. In Fig. 8(a) the band gap energy, $h\omega = 2a$, and onsets of IVS and DS can be clearly seen (indicated with dashed vertical lines).

Five different pulse durations are considered in Fig. 8(a) and the response is evaluated at $t_f = 5t_p$ for a relatively long decoherence time of $\tau_0 = 100$ fs [62]. As can be seen, increasing the pulse duration tends to sharpen the response about $h\omega = 2a$ and improve the maximum valley-polarization. For a $t_p = 10$ fs pulse, a maximum valley-polarization of $P^{(2)} = 0.930$ is achieved, while for a $t_p = 125$ fs pulse the maximum valley-polarization achieved is $P^{(2)} = 0.989$. As was discussed in Sec. IIIA, the decoherence time leads to linewidth broadening, making targeting $E_c(k) = a$ precisely impossible even when the pulse duration is very long. In fact, we see that the maximum valley-polarization converges to a constant value as the pulse duration approaches the decoherence time. Increasing the pulse duration beyond the decoherence time will increase the density of injected carriers, however. Thus, longer pulse durations are recommended as they yield stronger valley-polarizations, but the decoherence time ultimately limits how much the valley-polarization can be improved by increasing the pulse duration alone.

In Fig. 8(a) the optimal operating frequency does not in general coincide exactly with $h\omega = 2a$, and in fact varies somewhat with the pulse duration. For a $t_p = 10$ fs pulse, the optimal operating frequency is $h\omega = 422$ meV, while for a $t_p = 100$ fs pulse the optimal operating frequency is $h\omega = 400$ meV. The value of the optimal operating frequency is a result of a complex interplay between the pulse duration, decoherence time, external bias (with corresponding $A/B$ ratio and density of states), the presence of intervalley scattering, and perhaps other factors which lead to violations of the simple rule of thumb that the optima lie along $h\omega = 2a$. For the particular bias chosen in Fig. 8(a), the presence of IVS seems to play an important role in the interplay, pushing the optima towards lower values as the pulse duration is increased. We note that in the limit of long pulse durations and decoherence times, the optimal operating frequency converges to $h\omega = 2a$. Some experimental fine-tuning may be useful for finding the optimal operating frequency-bias pair for each particular sample and laser.

In Fig. 8(b) we again set the bias to $a = 205$ meV, but this time fix the pulse duration to $t_p = 100$ fs and vary the decoherence time. The same general features of Fig. 8(a) are observed. The main difference is that the maximum valley-polarization obtained in Fig. 8(b) can be much larger than in Fig. 8(a) (note the different scales on the vertical axes). For a decoherence time of $\tau_0 = 10$ fs, a maximum valley-polarization of $P^{(2)} = 0.930$ is achieved, while for $\tau_0 = 300$ fs the maximum valley-polarization achieved is $P^{(2)} = 0.996$. The maximum valley-polarization continues to grow even as the decoherence time is increased through 500 fs. In Fig. 8(a) improving the valley-polarization by increasing the pulse duration was limited by the decoherence time. On the other hand, in Fig. 8(b) the valley-polarization can be increased indefinitely by using very pure samples with long decoherence times, regardless of the pulse duration. Comparing Figs. 8(a) and 8(b) it is clear that the decoherence time is the limiting factor with respect to the valley-polarization. While clean samples are preferable, a very strong valley-polarization can still be achieved when the decoherence time is short. As was discussed in Sec. IIIA, even in the long-pulse-duration—long-decoherence-time limit, the valley-polarization can never be complete due to the presence of the anti-resonant terms.
such that first-order perturbation theory is sufficient.

chosen to ensure that $\rho$ is independent of $E$. We ensure that the number of photons per pulse depends on the second-order valley-polarization, but will become important now as the zeroth-order response does not deplete the thermal background, we simply add the thermal carrier density to the injected carrier density [see Eq. (18)] and compute the valley-polarization. That is, we write

$$\tilde{n}_v = \tilde{n}_v^{(2)} + n_v^{(0)},$$

where $\tilde{n}_v^{(2)}$ was defined in Sec. III B to account for intraband scattering via optical phonons (both IVS and DS). Since the thermal carrier density is the same in both the $K$ and $K'$ valleys, this corresponds mathematically to adding a factor of $2n_{K'}^{(0)} = 2n_{K}^{(0)}$ to the denominator of Eq. (43). The valley-polarization is now

$$\mathcal{P} = \frac{n_{K}^{(2)} - n_{K'}^{(2)}}{n_{K}^{(2)} + n_{K'}^{(2)} + 2n_{K}^{(0)}}.$$  \hspace{1cm} \hspace{1cm} (46)

Since $n_{v}^{(2)} \propto E_0^2$ [Eq. (28)], the field strength had no effect on the zeroth-order valley-polarization, but will become important now as the zeroth-order response does not depend on $E_0$. We take $E_0 = 1.5 \times (\omega/\omega_0)^{1/2}$ kV cm$^{-1}$, with $\omega_0 = 50$ meV. By scaling the field with the pulse frequency, we ensure that the number of photons per pulse is fixed across all frequencies. The magnitude of $E_0$ is chosen to ensure that $\rho(v)_{K,K'}^{(2)}$ is at most 1% for any $k$, such that first-order perturbation theory is sufficient.

D. Thermal carriers

Up to this point we have only considered the valley-polarization arising from the second-order response, $\mathcal{P}^{(2)}$, and have so far neglected the background of thermal carriers. In this section, we examine how thermal carriers act to reduce the valley-polarization and comment on strategies for minimizing their impact. To account for the thermal background, we simply add a factor of $2$ to the denominator of Eq. (43). The valley-polarization is now

$$\mathcal{P}^{(2)} = \frac{n_{K}^{(2)} - n_{K'}^{(2)}}{n_{K}^{(2)} + n_{K'}^{(2)} + 2n_{K}^{(0)}}.$$  \hspace{1cm} \hspace{1cm} (46)

Since $n_{v}^{(2)} \propto E_0^2$ [Eq. (28)], the field strength had no effect on the zeroth-order valley-polarization, but will become important now as the zeroth-order response does not depend on $E_0$. We take $E_0 = 1.5 \times (\omega/\omega_0)^{1/2}$ kV cm$^{-1}$, with $\omega_0 = 50$ meV. By scaling the field with the pulse frequency, we ensure that the number of photons per pulse is fixed across all frequencies. The magnitude of $E_0$ is chosen to ensure that $\rho(v)_{K,K'}^{(2)}$ is at most 1% for any $k$, such that first-order perturbation theory is sufficient.

In Fig. 9(a), we plot the valley-polarization under the same conditions as Fig. 7(b) ($t_p = 50$ fs, $\tau_0 = 30$ fs, $t_f = 250$ fs, $T = 300$ K, $\mu = 0$), but this time account for the thermal background. As can be seen, the valley-polarization is reduced significantly, with a maximum valley-polarization of $\mathcal{P} = 0.70$ obtained at $(h\omega, a) \approx (368,308)$ meV. The region of optimal operating frequency-bias pairs has also moved away from $h\omega = 2a$ and now hugs the band edge. This can be understood by considering the following. Along $h\omega = 2a$, the second-order response is very strongly valley-polarized, however the actual number of carriers excited is quite small because we are operating away from the band edge where the joint density of states is small. Along the band edge, many more carriers are excited and even though the second-order response is not as pure, the difference in carrier density between the $K$ and $K'$ valleys is able to overcome the thermal background. Note also that in contrast to Fig. 7(b), the valley-polarization obtained below band gap is now essentially zero. This is because at frequencies below the band gap, very few carriers are excited and so the thermal population dominates.

In Figs. 9(a) through (d) we show the evolution of the valley-polarization as the temperature is decreased from 300 to 150 K. As the temperature decreases, the influence of the thermal population becomes less and less: the valley-polarization improves and the optimal operating region returns to $h\omega = 2a$. At 150 K, the optimal operating frequency-bias pair is $(h\omega, a) \approx (235,126)$ meV, yielding a very strong valley-polarization of $\mathcal{P}^{(2)} = 0.97$. 

FIG. 8. Deviation of the second-order valley-polarization from perfect polarization $\log_{10}(1 - \mathcal{P}^{(2)})$ as a function of the central photon energy $h\omega$ of the exciting Gaussian pulse for a fixed bias of $a = 205$ meV. (a) From lightest to darkest, $t_p = 10, 15, 25, 50$, and 100 fs for a fixed decoherence time of $\tau_0 = 100$ fs. (b) From lightest to darkest, $\tau_0 = 10, 20, 100, 200$, and 300 fs for a fixed pulse duration of $t_p = 100$ fs. In both (a) and (b), the valley-polarization is evaluated at $t_f = 5t_p$. In order of increasing energy, the vertical lines indicate the band gap energy, $h\omega = 2a$, the onset of intervalley scattering, and the onset of down scattering. The horizontal line indicates $\mathcal{P}^{(2)} = 0$. Note that the scale of the valley-polarization axis is different in (a) and (b).
FIG. 9. (Color online) Deviation of the valley-polarization (including thermal carriers) from perfect polarization $\log_{10}(1 - P)$ as a function of the external bias $a$ and central photon energy $h\omega$ of the exciting Gaussian pulse at four different temperatures and including both intervalley scattering and down-scattering. The black dashed lines are as in Fig. 4. The white dashed lines from Fig. 7 have been omitted for clarity. The field amplitude is given by $E_0 = 1.5 \times (\omega/\omega_0)^{1/2}$ kV cm$^{-1}$, with $h\omega_0 = 50$ meV. The contour values are not the same for all four plots, but the color scale is consistent throughout.

However, even at 150 K, thermal carriers significantly reduce the valley-polarization for low biases. This is because the band gap $\Delta E(a) \approx 2a$ for low biases, and so the thermal population is non-negligible. For $T = 50$ K (not shown), the thermal population is negligible, and the valley-polarization observed in Fig. 7(b) is restored. Thus, the presence of thermal carriers effectively cuts off low biases from the optimal operating region, much in the same way as intervalley scattering cut off high frequencies (Sec. III B). Small biases may be desirable due to reduced electron mobility with increasing bias [63]. In order to operate in the low-bias regime, one must work at low temperatures. If one works at a larger bias, a very strong valley-polarization can be achieved at 150 K.

It would be nice if a strong valley-polarization could be achieved at room temperature. One possibility is to increase the field strength, which has the effect of mitigating the thermal population, and follows a similar progression as Fig. 9. We must be careful with increasing the field strength however, because as alluded to earlier, we begin to push the limits of first-order perturbation theory. To restore Fig. 9(a) to Fig. 7(b) by increasing the field amplitude alone, $E_0$ would need to be increased by about a factor of 50. Since $\rho^{(2)}(k) \propto E_0^2$, this puts us well outside the realm of first-order perturbation theory.
IV. CONCLUSION

In this work, we have presented a detailed study of valley-polarization in biased bilayer graphene using circularly polarized light. At low temperatures, and in the absence of intraband scattering, the optimal operating frequency-bias pairs are found to satisfy $\hbar \omega = 2a$, where $\omega$ is the central frequency of the exciting pulse, and $2a$ is potential energy difference between the graphene layers. The valley-polarization is robust to deviations of a few 10s of meV from the optimal operating configuration. Intervaly scattering via optical phonons is found to significantly reduce the valley-polarization at high frequencies ($\hbar \omega \gtrsim 320$ meV above the band gap).

At room temperature, thermal carriers significantly reduce the valley-polarization, especially for low biases. However, their effect can be mitigated by working at low temperatures ($T < 150$ K). A strong valley-polarization can be achieved in any reasonably pure sample, but can be significantly improved by limiting defects and impurities. For typical samples, pulse durations of 10s to a few 100s of fs are recommended. For room-temperature experiments, we recommend operating at $(\hbar \omega, a) = (368, 308)$ meV, where a valley-polarization of 70% can be achieved. At room temperatures, we recommend operating at $(\hbar \omega, a) = (236, 126)$ meV, where a valley-polarization of over 97% can be achieved. We believe that these results demonstrate bilayer graphene to be a strong candidate for valleytronic applications.

ACKNOWLEDGMENTS

This work was supported by Queen’s University and the Natural Sciences and Engineering Research Council of Canada (NSERC).

Appendix A: Elliptical polarization

In this section, we examine why circularly polarized light is optimal to induce a valley-polarization in biased bilayer graphene. We also explore whether a general elliptical polarization is ever preferable to circular polarization. Consider a general electric field

$$\mathbf{E}(t) = \alpha \mathbf{E}_L + \beta \mathbf{E}_R + c.c., \quad (A1)$$

which we have written in the basis of left- and right-hand circularly polarized components

$$\mathbf{E}_L = (\mathbf{k} - i \hat{\theta}_k) e^{-i(\omega t + \theta_k)}, \quad (A2)$$

$$\mathbf{E}_R = (\mathbf{k} + i \hat{\theta}_k) e^{-i(\omega t - \theta_k)}, \quad (A3)$$

expressed in polar coordinates with origin at either $\mathbf{K}$ or $\mathbf{K}'$ (the field takes the same form in both valleys). Here, $\alpha$ and $\beta$ are complex coefficients that characterize the polarization of $\mathbf{E}(t)$, with $|\alpha|^2 + |\beta|^2 = 1$. If $\alpha = 0$ or $\beta = 0$, $\mathbf{E}(t)$ is circularly polarized. If $|\alpha| = |\beta|$, $\mathbf{E}(t)$ is linearly polarized, the polarization axis determined by the phase difference. If $\alpha \neq \beta$ and neither $\alpha$ nor $\beta$ are zero, we have some general elliptical polarization.

The first-order interband coherence $\rho_{cv}^{(1)}(\mathbf{k})$ is proportional to the carrier-field interaction $\xi_{cv}(\mathbf{k}) \cdot \mathbf{E}(t)$ [see Eq. (10)]. If $\xi_{cv}(\mathbf{k}) \cdot \mathbf{E}(t)$ can be forced to zero in one valley but not the other, a strong valley-polarization is expected. For the field of Eq. (A1), the carrier-field interaction takes the form

$$\xi_{cv}(\mathbf{k}) \cdot \mathbf{E}(t) \approx i e^{-i \omega t} \left[ (\alpha e^{i \theta_k} + \beta e^{i \theta_k}) \right]$$

$$+ B \left( -\alpha e^{-i \theta_k} + \beta e^{i \theta_k} \right), \quad (A4)$$

where the approximation sign indicates that we are considering only the resonant contributions, and where the plus and minus signs correspond to the $K$ and $K'$ valleys respectively. Here, $A$ and $B$ are the real, positive, $k$-dependent functions which were introduced in Sec. III B.

Setting Eq. (A4) to zero, we obtain the condition

$$\beta(B \pm A) = \alpha(B \mp A) e^{-2i \theta_k}. \quad (A5)$$

The problem of forcing the resonant piece of $\xi_{cv}(\mathbf{k}) \cdot \mathbf{E}(t) = 0$ has been reduced to satisfying Eq. (A5). On the surface, we are faced with serious problem: Eq. (A5) depends on $\theta_k$, suggesting that Eq. (A5) can only be satisfied along a single radial direction in $k$-space. Since $A$ and $B$ are $k$-dependent, this then implies that Eq. (A5) can only be satisfied at (at most) a couple of points in $k$-space. One way to circumvent this issue is to use circularly polarized light. If we set $\alpha = 0$ (right-hand circular polarization) we obtain

$$B = 0, \quad (A6)$$

which can be satisfied in the $K'$ valley when $A = B$, but never in K $[64]$. Similarly, if we set $\beta = 0$ (left-hand circular polarization), we obtain

$$0 = (B \mp A), \quad (A7)$$

which can be satisfied in the $K$ valley when $A = B$, but never in $K'$. Thus, for any non-circular elliptical polarization, the phase-dependence of Eq. (A5) limits the optical selection rule to just a couple of $k$-space points. By choosing circularly polarized light, the phase dependence of Eq. (A5) is removed, and a valley-contrasting optical selection rule is obtained for the specific constraint $A = B$. Since $A = B$ occurs for $k = k_1$ (see Sec. III A), the valley-contrasting optical selection rule is simultaneously satisfied around an entire ring of states in $k$-space, as opposed to just a couple of points. Circular polarization is therefore always preferable to elliptical.

Appendix B: Complex-shifted integral

In Eq. (25), $\lambda$ is in general complex, but the integral bounded by $(−\infty, \infty)$ is equivalent to a complex-shifted integral bounded by $(−\infty + i\gamma, \infty + i\gamma)$ for some real $\gamma$. 

Proof: By Cauchy’s integral theorem, an integral over the closed contour $C = [-R, R, R + iγ, -R + iγ, -R]$ is zero:

$$0 = \oint_C f(\eta) d\eta = \int_{-R}^{R} f(R + i\gamma) d\eta + \int_{R + i\gamma}^{\infty} f(R + i\gamma) d\eta + \int_{-\infty}^{R} f(-R + i\gamma) d\eta + \int_{R}^{-\infty} f(-R) d\eta,$$  \hspace{1cm} (B1)

where $R$ is real and $f(\eta) = (\text{erf}(\eta) + 1) \exp(-\eta^2)$. For $R \to \infty$, the integrals over the segments of constant $R$ are zero because $f(\eta)$ goes to zero. Thus, the integral along the real axis is equal to an arbitrarily complex-shifted integral over the same real limits

$$0 = \int_{-R}^{R} f(R + i\gamma) d\eta = \int_{-R}^{R} - f(R + i\gamma) d\eta \hspace{1cm} (B2)$$

---

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, *Science* **306**, 666 (2004).
[2] S. V. Morozov, K. S. Novoselov, M. I. Katsnelson, F. Schedin, D. C. Elias, J. A. Jaszczak, and A. K. Geim, *Phys. Rev. Lett.* **100**, 016602 (2008).
[3] A. V. Rozhkov, A. O. Sboychakov, A. L. Rakhamnov, and F. Nori, *Physics Reports* **468**, 1 (2016).
[4] C. Lee, X. Wei, J. W. Kysar, and J. Hone, *Science* **321**, 385 (2008).
[5] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009).
[6] I. Zutic, J. Fabian, and S. Das Sarma, *Rev. Mod. Phys.* **76**, 323 (2004).
[7] S. A. Vitale, D. Nozich, J. O. Varghese, P. Kim, N. Gedik, P. Jarillo-Herrero, D. Xiao, and M. Rothschild, *Small* **14**, 1801483 (2018).
[8] L. E. Golub, S. A. Tarasenko, M. V. Entin, and L. I. Magaril, *Phys. Rev. B* **84**, 195408 (2011).
[9] A. Rycerz, J. Tworzydlo, and C. W. J. Beenakker, *Nature Physics* **3**, 172 (2007).
[10] J. M. P. Jr, F. M. Peeters, R. N. C. Filho, and G. A. Farias, *Journal of Physics: Condensed Matter* **21**, 045301 (2008).
[11] T. Fujita, M. B. A. Jalil, and S. G. Tan, *Applied Physics Letters* **97**, 043508 (2010).
[12] R. V. Gorbatchev, J. C. W. Song, G. L. Yu, A. V. Kretinin, F. Withers, Y. Cao, A. Mishchenko, I. V. Grigorieva, K. S. Novoselov, L. S. Levitov, and A. K. Geim, *Science* **346**, 448 (2014).
[13] D. Xiao, W. Yao, and Q. Niu, *Phys. Rev. Lett.* **99**, 236809 (2007).
[14] W. Yao, D. Xiao, and Q. Niu, *Phys. Rev. B* **77**, 235406 (2008).
[15] G. Giovannetti, P. A. Khomyakov, G. Brock, P. J. Kelly, and J. van den Brink, *Phys. Rev. B* **76**, 073103 (2007).
[16] T. Cao, G. Wang, W. Han, H. Ye, C. Zhu, J. Shi, Q. Niu, P. Tan, E. Wang, B. Liu, and J. Feng, *Nature Communications* **3**, 887 (2012).
[17] K. F. Mak, K. He, J. Shan, and T. F. Heinz, *Nature Nanotechnology* **7**, 494 (2012).
[18] K. F. Mak, K. L. McGill, J. Park, and P. L. McEuen, *Science* **344**, 1489 (2014).
[19] F. Joucken, E. A. Quezada-López, J. Avila, C. Chen, J. L. Davenport, H. Chen, K. Watanabe, T. Taniguchi, M. C. Asensio, and J. Velasco, *Phys. Rev. B* **99**, 161406 (2019).
[20] E. McCann and V. I. Fal’ko, *Phys. Rev. Lett.* **96**, 086805 (2006).
[21] E. McCann, *Phys. Rev. B* **74**, 161403 (2006).
[22] J. B. Oostinga, H. B. Heersche, X. Liu, A. F. Morpurgo, and L. M. K. Vandersypen, *Nature Materials* **7**, 151 (2008).
[23] Y. Zhang, T.-T. Tang, C. Girit, Z. Hao, M. C. Martin, A. Zettl, M. F. Crommie, Y. R. Shen, and F. Wang, *Nature* **459**, 820 (2009).
[24] T. Ohta, A. Bostwick, T. Seyller, K. Horn, and E. Rotenberg, *Science* **313**, 951 (2006).
[25] E. V. Castro, K. S. Novoselov, S. V. Morozov, N. M. R. Peres, J. M. B. L. dos Santos, J. Nilsson, F. Guinea, A. K. Geim, and A. H. C. Neto, *Phys. Rev. Lett.* **99**, 216802 (2007).
[26] K. F. Mak, C. H. Lui, J. Shan, and T. F. Heinz, *Phys. Rev. Lett.* **102**, 256405 (2009).
[27] E. McCann, “Electronic properties of monolayer and bilayer graphene,” in *Graphene Nanoelectronics: Metrology, Synthesis, Properties and Applications*, edited by H. Raza (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012) pp. 237–275, arXiv:1205.4849 [cond-mat.mes-hall].
[28] J. Li, R.-X. Zhang, Z. Yin, J. Zhang, K. Watanabe, T. Taniguchi, C. Liu, and J. Zhu, *Science* **362**, 1149 (2018).
[29] Y. Shimazaki, M. Yamamoto, I. V. Borzenets, K. Watanabe, T. Taniguchi, and S. Tarucha, *Nature Physics* **11**, 1032 (2015).
[30] D. S. L. Abergel and T. Chakraborty, *Applied Physics Letters* **95**, 062107 (2009).
[31] D. S. L. Abergel and T. Chakraborty, *Nanotechnology* **22**, 015203 (2010).
[32] R. McGouran, I. Al-Naib, and M. M. Dignam, *Phys. Rev. B* **94**, 235402 (2016).
[33] E. Blount (Academic Press, 1962) pp. 305 – 373.
[34] Intraband connection elements, i.e. $\xi_n$, will have an extra term within the summation proportional to the gradient of the normalization factor $A_n(k)$.
[35] $B = 0$ at $k = 0$ only when $a$ is nonzero. If $a = 0$, $B \to \infty$ at $k = 0$.
[36] I. Al-Naib, J. E. Sipe, and M. M. Dignam, *Phys. Rev. B* **90**, 245423 (2014).
[37] R. McGouran and M. M. Dignam, *Phys. Rev. B* **96**, 045439 (2017).
[38] E. W. Ng and M. Geller, *Journal of Research of the National Bureau of Standards, Section B: Mathematical Sciences* **73B**, 8 (1969) Eq. (4.3.13).
[39] E. Malic, T. Winzer, E. Bobkin, and A. Knorr, *Phys. Rev. B* **84**, 205406 (2011).
[40] I. Gierz, J. C. Petersen, M. Mitrano, C. Cacho, I. C. E. Turcu, E. Springate, A. Stöhr, A. Köhler, U. Starke, and A. Cavalleri, *Nature Materials* **12**, 1119 (2013).
[41] J. C. Johansson, S. Ulstrup, F. Cilento, A. Crepaldi, M. Zacchigna, C. Cacho, I. C. E. Turcu, E. Springate,
F. Fromm, C. Raidel, T. Seyller, F. Parmigiani, M. Grioni, and P. Hofmann, Phys. Rev. Lett. 111, 027403 (2013).

[42] I. Gierz, F. Calegari, S. Aschlimann, M. Chávez Cervantes, C. Cacho, R. T. Chapman, E. Springate, S. Link, U. Starke, C. R. Ast, and A. Cavalleri, Phys. Rev. Lett. 115, 086803 (2015).

[43] S. Ulstrup, J. C. Johannsen, F. Cilento, J. A. Miwa, A. Crepaldi, M. Zacchigna, C. Cacho, R. Chapman, E. Springate, S. Mammodov, F. Fromm, C. Raidel, T. Seyller, F. Parmigiani, M. Grioni, P. D. C. King, and P. Hofmann, Phys. Rev. Lett. 112, 257401 (2014).

[44] I. Gierz, M. Mitrano, J. C. Petersen, C. Cacho, I. C. E. Turcu, E. Springate, A. Stöhr, A. Köhler, U. Starke, and A. Cavalleri, Journal of Physics: Condensed Matter 27, 164204 (2015).

[45] H. Yang, X. Feng, Q. Wang, H. Huang, W. Chen, A. T. S. Wee, and W. Ji, Nano Letters 11, 2622 (2011).

[46] The field amplitude $E_0$ does not need to be specified as it factors out in the evaluation of $P^{(2)}$. As long as the observation time $t_f$ is several times the pulse duration, $t_f$ also does not need to be specified since $n^{(2)}_V$ is time-independent for long $t_f$ when $\gamma_{cc}(k) = 0$ [see Eq. (28)].

[47] S. Cobeldick, “Colorbrewer: Attractive and distinctive colormaps,” (2020).

[48] In particular, this is the transition between either the high-energy valence band and the low-energy conduction band, or the low-energy valence band and the high-energy conduction band (see Ref. [27]).

[49] W.-K. Tse, E. H. Hwang, and S. Das Sarma, Applied Physics Letters 93, 023128 (2008).

[50] T. Fang, A. Konar, H. Xing, and D. Jena, Phys. Rev. B 84, 125450 (2011).

[51] M. Lazzeri, C. Attaccalite, L. Wirtz, and F. Mauri, Phys. Rev. B 78, 081406 (2008).

[52] K. M. Borysenko, J. T. Mullen, X. Li, Y. G. Semenov, J. M. Zavada, M. B. Nardelli, and K. W. Kim, Phys. Rev. B 83, 161402 (2011).

[53] A. K. Vallabhaneni, D. Singh, H. Bao, J. Murthy, and X. Ruan, Phys. Rev. B 93, 125432 (2016).

[54] K. M. Borysenko, J. T. Mullen, E. A. Barry, S. Paul, Y. G. Semenov, J. M. Zavada, M. B. Nardelli, and K. W. Kim, Phys. Rev. B 81, 121412 (2010).

[55] A. Fay, R. Danneau, J. K. Viljas, F. Wu, M. Y. Tomi, J. Wengler, M. Wiesner, and P. J. Hakonen, Phys. Rev. B 84, 245427 (2011).

[56] K. J. Tielrooij, J. C. W. Song, S. A. Jensen, A. Centeno, A. Pesquera, A. Zurutuza Elorza, M. Bonn, L. S. Levitov, and F. H. L. Koppens, Nature Physics 9, 248 (2013).

[57] T. Winzer, A. Knorr, and E. Malic, Nano Letters 10, 4839 (2010).

[58] We note that in Fig. 8(a) we have used pulse durations which can be shorter than a single cycle for some frequencies. While this may be unphysical or difficult to achieve experimentally, it is useful for demonstrating the effects of varying the pulse duration.

[59] X. Li, K. M. Borysenko, M. B. Nardelli, and K. W. Kim, Phys. Rev. B 84, 195453 (2011).

[60] That is, unless $A = B = 0$, but then the condition is satisfied in both valleys simultaneously so there is no valley-contrasting selection rule.