

# Charge quenching at defect states in transition metal dichalcogenide-graphene van der Waals heterobilayers

Daniel Hernangómez-Pérez,<sup>1,\*</sup> Andrea Donarini,<sup>2</sup> and Sivan Refaeli-Abramson<sup>1</sup>

<sup>1</sup>*Department of Molecular Chemistry and Materials Science,  
Weizmann Institute of Science, Rehovot 7610001, Israel*

<sup>2</sup>*Institute of Theoretical Physics, University of Regensburg, 93040 Regensburg, Germany*

We study the dynamical properties of point-like defects, represented by monoatomic chalcogen vacancies, in WS<sub>2</sub>-graphene and MoS<sub>2</sub>-graphene heterobilayers. Employing a multidisciplinary approach based on the combination of *ab initio*, model Hamiltonian and density matrix techniques, we propose a minimal interacting model that allows for the calculation of electronic transition times associated to population and depopulation of the vacancy by an additional electron. We obtain the “coarse-grained” semiclassical dynamics by means of a master equation approach and discuss the potential role of virtual charge fluctuations in the internal dynamics of impurity quasi-degenerate states. The interplay between the symmetry of the lattice and the spin degree of freedom through the spin-orbit interaction and its impact on charge quenching is studied in detail.

**Keywords:** 2D materials, transition-metal dichalcogenides, van der Waals, defects, graphene, transport, single-electron

## I. INTRODUCTION

In recent years, the study of van der Waals heterostructures [1, 2], created upon stacking atomically-thin two-dimensional layers, has unveiled a rich variety of physical phenomena in these systems [3], ranging from sequential electron transfer through defects, studied using tunneling spectroscopy [4, 5], to ultrafast interlayer charge transfer after photoexcitation with visible (yellow) light studied with time-resolved ARPES [6–8] or transient absorption spectroscopy [9]. Such layered two-dimensional heterostructures can be grown epitaxially in a controlled manner, revealing new families of materials that cannot be naturally found and whose properties extend far beyond the simple combination of the separated layers [3]. Consequently, these systems offer a vast playground where the electronic and optical properties can be modified at the atomic scale due to combinations and variations in screening, electronic confinement or spin-orbit interaction.

Of particular interest are heterostructures of monolayer transition metal dichalcogenides (TMDCs) of the type XS<sub>2</sub> adsorbed on monolayer graphene. These systems can be considered as special type I heterostructures where one of the two layers is a semimetal, and therefore, both the conduction band minimum and the valence band maximum are located in the same layer and touch at individual high symmetry points in the Brillouin zone. Moreover, these heterostructures have the peculiarity of stacking graphene, a material characterized by an extremely high carrier mobility [10, 11] and highly symmetric density of states (DoS) close to its charge neutrality point, with a strongly confined monolayer semiconductor with direct band gap and which can have a sizable spin-orbit splitting [12].

Strong confinement in two-dimensional heterostructures makes them particularly sensitive to the potential created by impurities or vacancies (defects). These vacancies are known to produce in-gap states with clear fingerprints that can be measured by local probes [13] as well as affect the heterostructure optical properties. These optical properties are also well known to be related to the large pristine [14, 15] and defect-induced [13] exciton binding energies and show valley circular dichroism due to valley-spin coupling by means of spin-orbit interaction [16–19].

From the electronic point of view, defects created by substitutions or vacancies can be considered as “electron traps”. They have been anticipated to decrease the electronic conductivity and electronic mobility when the Fermi energy is close to the localized defect levels, hybridized, for example, with the continuum from the band extrema [20, 21]. For similar reasons, defect scattering channels have also been hypothesized to be the source of longer lifetimes of electronic states compared to hole ones in WS<sub>2</sub>-graphene interfaces after ultrafast charge separation induced by photoexcitation [7, 8, 22] as they are expected to efficiently pin down the photogenerated charge carriers. However, as was recently shown by some of us for the case of oxide surfaces [23], the role of defects may be subtle and cannot be always considered as trivial charge traps simply increasing localization or decreasing photoconductivity.

Defects created by vacancies in the regular lattice of TMDCs can be seen also as quantum dots [24] with the symmetry provided by the host lattice determining their shell structure. In these “artificial atoms”, the discrete electronic charge as well as the interplay between Coulomb interaction, spin-orbit coupling and quantum statistics manifests in several well-studied physical effects (statistical crossover in the many-body interference blockade regime [25], channel blockade [26], spin blockade [27], geometrical charge frustration [28] or triplet-

\* daniel.hernangomez@weizmann.ac.il

singlet transitions induced by superexchange [29] among many others) and applications (spin qubits [30] or current rectifiers [31]). In particular, many-body quantum interference has been the subject of recent research [32–34] as well as its connection with the quantum dot symmetry and symmetry-induced degeneracies. Such type of dots can be studied both experimentally and theoretically for lateral heterostructures [26, 28, 35–37]. Vacancies in TMDCs and TMDC heterostructures thus open the possibility of studying the electronic and spin properties of these systems in the atomistic limit.

In this paper, we study the charge transfer processes occurring in  $\text{XS}_2$ –graphene (Gr) heterobilayers (here,  $X = \text{Mo}, \text{W}$ ) with chalcogen vacancies. These vacancies, understood in the dilute limit, act as isolated defects or “dots” with the symmetry provided by the hosting lattice. The vacancy states are tunnel-coupled to a graphene reservoir which provides a structured DoS. We compute the electronic transition times for charging and discharging of the empty levels of the vacancy, due to tunneling from and into the graphene layer, with a combination of *ab initio*, model Hamiltonian and density matrix-based methods. The transition rates for second order processes in the tunneling Hamiltonian are employed to understand the semiclassical many-body dynamics. We also discuss the role of degeneracies in the potentially coherent dynamics at the dot. Finally, we examine the interplay between orbital (lattice “symmetry”) and spin degrees of freedom that can occur by means of spin-orbit interaction. The population dynamics for the non-trivial charge transfer mechanism and its consequences for the semiclassical quantum dot dynamics are analyzed in detail. Overall, we show how the combination of different theoretical techniques can be employed to study quantum dot dynamics at complex interfaces, with realistic physical parameters reflecting the specific system structure. This work will also serve as a basis to understand more subtle aspects of the quantum dot dynamics and as a guidance to the interpretation of experimental observations.

The paper is organized as follows: in Sec. II, we present density functional theory (DFT) band structure calculations for the  $\text{MoS}_2$ –Gr and  $\text{WS}_2$ –Gr heterostructures upon which our model will be constructed. In Sec. III we discuss the low energy model Hamiltonian used to analyze our DFT results. A summary of the kinetic theory used in this manuscript is briefly presented in Sec. IV. The results for the dynamics and the electronic transition rates are shown in Sec. V, considering separately the case of weak and strong spin-orbit interaction. We conclude this paper in Sec. VI and devote the Supplemental Material (SM) to computational details, explicit analytical derivations and to supplementary results.

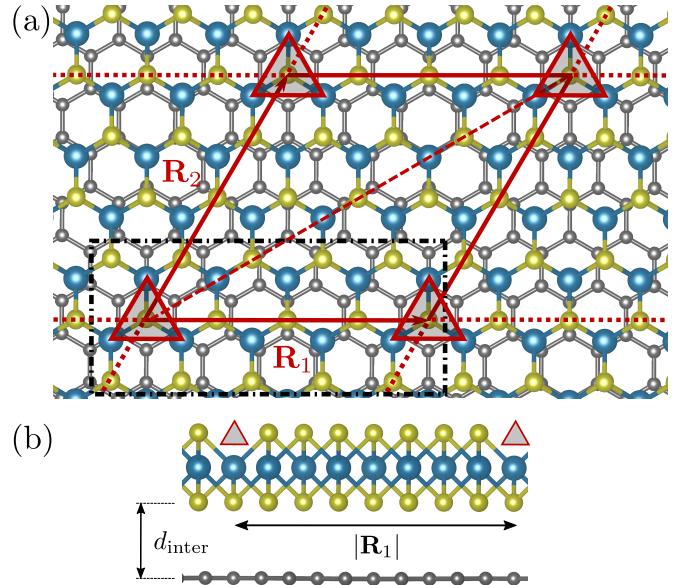


FIG. 1. (a) Top view of the  $\text{XS}_2$ –Gr supercell (here  $X = \text{W}$ ). The  $4 \times 4/5 \times 5$  supercell is indicated by the red solid parallelogram, the position of the single chalcogen vacancy (and its symmetry) is marked by the red triangle located at the origin of the unit cell, appearing four times due to the supercell periodicity in this panel. The dotted lines represent the borders of adjacent supercells. The dashed red line along the main diagonal indicates that graphene adsorption on the TMDC reduces the global symmetry of the system from  $C_{3v}$  to  $C_s$ , *i.e.* the only symmetry element is the reflection plane perpendicular to the heterostructure and bisecting the angle between the supercell basis vectors. (b) Side view of the  $\text{WS}_2$ –Gr heterostructure cut along the dashed-dotted black box marked in panel (a). The missing atom corresponds to the chalcogen vacancy in the supercell, represented here by a red triangle.

## II. DENSITY FUNCTIONAL THEORY BAND STRUCTURE ANALYSIS

### A. Geometry

To study the role of isolated defects in  $\text{XS}_2$ –Gr heterobilayers, we adopt the supercell approach. The supercell employed in this work is shown in Fig. 1 (a). This supercell is made from a heterobilayer formed by a TMDC of the type  $\text{XS}_2$  (where  $X = \text{W}, \text{Mo}$ ) and graphene. The commensuration of the corresponding lattices in each layer requires to have  $4 \times 4$  elementary cells for  $\text{XS}_2$  and  $5 \times 5$  for graphene in the supercell. The supercell was previously optimized using DFT, see details in Methods section. The in-plane lattice constant for the  $\text{XS}_2$  monolayer is equal to  $3.15 \text{ \AA}$ , while graphene has an in-plane lattice constant of  $2.52 \text{ \AA}$  (the distance between nearest neighbor carbon atoms of  $1.455 \text{ \AA}$ ). In other words, the graphene lattice is strained by  $\sim 2.4\%$  with respect to an isolated relaxed graphene monolayer while we keep TMDC layer almost not strained with respect to the experimental lattice constant [38]. In this sense, we note

the difference from other calculations performed in previous works with the same type of commensurate supercell [6], in which the strain is distributed between the two layers of the TMDC–Gr heterostructure with similar lattice mismatch. The supercell lattice vectors therefore have a length of  $|\mathbf{R}_1| = |\mathbf{R}_2| \simeq 12.6 \text{ \AA}$ . The optimized interlayer distance between the  $\text{XS}_2$  and the graphene layers in the supercell is  $3.44 \text{ \AA}$ . In order to prevent spurious interaction between the periodic replicas, an out-of-plane vacuum layer of width of  $\sim 17 \text{ \AA}$  is employed. We consider a single chalcogen vacancy per unit cell (this would correspond to a  $\sim 3\%$  vacancy concentration). The vacancy is located on the opposite side, in correspondence to a hollow position of the graphene layer and, for convenience, it is chosen to be the origin of the supercell. In our calculations, the supercell with a single vacancy has been created by removing a chalcogen atom from the previously optimized non-defected periodic supercell. We have checked that structural relaxation of the heterostructure lattice in the presence of the vacancy does not change the symmetry properties, only shrinking the defect size by roughly  $\sim 5\%$  due to a reduced metal-metal ( $\text{X}-\text{X}$ ) distance close to the vacant sulphur atom while preserving qualitatively the shape and nature of the DFT band structure.

### B. Band structure without spin-orbit interaction

We start by considering the DFT band structure of the periodic supercell in the absence of spin-orbit interaction. The computational details of our *ab initio* calculations are given in Sec. I of the SM. The band structure of the  $\text{WS}_2$ –Gr interface is shown in Fig. 2 (a). We clearly observe the superposition of the band structures of graphene (with its well-known Dirac cone located at the  $\bar{\text{K}}$  point of the mini-Brillouin zone) and  $\text{WS}_2$  monolayers, similarly to previously reported calculations [6]. For our supercell, the Dirac cone that sets the Fermi energy of the heterobilayer (and corresponds to the graphene charge neutrality point) is located in the  $\text{WS}_2$  pristine bandgap. Note that since the crystal structure of each individual layer has an hexagonal unit cell, the Brillouin zone is also hexagonal and matching at the edge for this supercell configuration sets the point  $\text{K}$  of the unfolded Brillouin zone to always map into  $\bar{\text{K}}$  of the mini-Brillouin zone by construction [39] (other high symmetry points may also map onto  $\bar{\text{K}}$ ). The non-defect conduction and valence bands of  $\text{WS}_2$  are located at  $\sim 1 \text{ eV}$  higher and lower in energy, respectively, from the graphene  $\bar{\text{K}}$  point.

The sulphur vacancies generate three defect bands stemming from the bond orbitals of the neighbouring metal atoms [40–42]. In addition, as a consequence of the  $C_{3v}$  symmetry of the sulphur vacancy, the defect bands comprise two empty and quasi-degenerate bands in the pristine  $\text{WS}_2$  gap and one occupied band located in the valence region (see Sec. II of the SM for a general discussion

on elementary symmetry properties). At the  $\bar{\Gamma}$  point, the two empty in-gap states are located at  $\sim 0.3 \text{ eV}$  above the graphene charge neutrality point. Here, the slightly broken symmetry in the supercell (see Fig. 1) manifest itself in a splitting of  $\sim 1 \text{ meV}$  between the energy levels. The vacancy states are mostly composed of  $d$ -orbitals of the W atoms closer to the vacant S atom. This situation was previously shown in other TMDCs such as  $\text{MoS}_2$  (see Refs. 40–42 and below). While the in-gap defect bands are relatively flat due to the spatial localization of the defect states in real space, they are still weakly dispersive close to the Dirac cone. This results from residual defect-defect interaction between neighbouring supercells when the TMDC layer has  $4 \times 4 \text{ XS}_2$  unit cells, see Supplemental Material in Ref. 13. We note that the in-gap defect bands hybridize with the graphene Dirac cone close to the  $\bar{\text{K}}$  point, generating anticrossings with characteristic level splittings. The different size of the two anticrossings is attributed to the symmetry properties of each of the states, being even and odd with respect to a reflection on a plane perpendicular to the heterostructure and located along the supercell main diagonal, see Fig. 1. We have checked that the electronic band structure is robust to the position of the vacancy, *i.e.* the band structure to be qualitatively invariant and the anti-crossing position and size are well-preserved, as globally breaking spatial symmetries in the supercell due to a change of the position of the vacancy does not substantially alter the local symmetry properties of the associated in-gap localized states.

In Fig. 2 (b), we show the (pseudo)-densities of the Kohn-Sham states,  $|\Psi|^2$ , marked by the red box in Fig. 2 (a) for the corresponding  $\mathbf{k}$ -point (dashed vertical line). As anticipated above, the hybridization between the two layers occurs at specific  $\mathbf{k}$ -points in the supercell Brillouin zone (forming an avoided ring around the graphene Dirac point due to the almost perfect conical shape of the graphene bands). This hybridization yields in-gap states with weight on both layers of the heterostructure at these avoided crossings [see panels (B), (C) of Fig. 2 (b) in comparison to panel (A)]. We note that hybridization can also occur far from the Fermi level, for example, in the valence band region close to the  $\bar{\text{M}}$  point, where coupling occurs between extended states on the  $\text{WS}_2$  and graphene layers as can be anticipated from the local energy level splitting.

The scenario described for the  $\text{WS}_2$ –Gr interface is similar to the one found in the  $\text{MoS}_2$ –Gr heterostructure, see Fig. 2, panels (c) and (d). In Fig. 2 (c), we display its DFT band structure. The geometry employed was obtained from the  $\text{WS}_2$ –Gr supercell by substitution of the W atoms by Mo and keeping the same lattice parameters (the experimental lattice parameters for  $\text{MoS}_2$  and  $\text{WS}_2$  are almost the same, being  $3.15 \text{ \AA}$  for the former [43] and  $3.153 \text{ \AA}$  for the latter [38]). We observe that, similar to the  $\text{WS}_2$ –Gr interface, the local  $C_{3v}$  symmetry of the vacancy yields the anticipated three defect bands, two of them being located in the pristine  $\text{MoS}_2$  bandgap

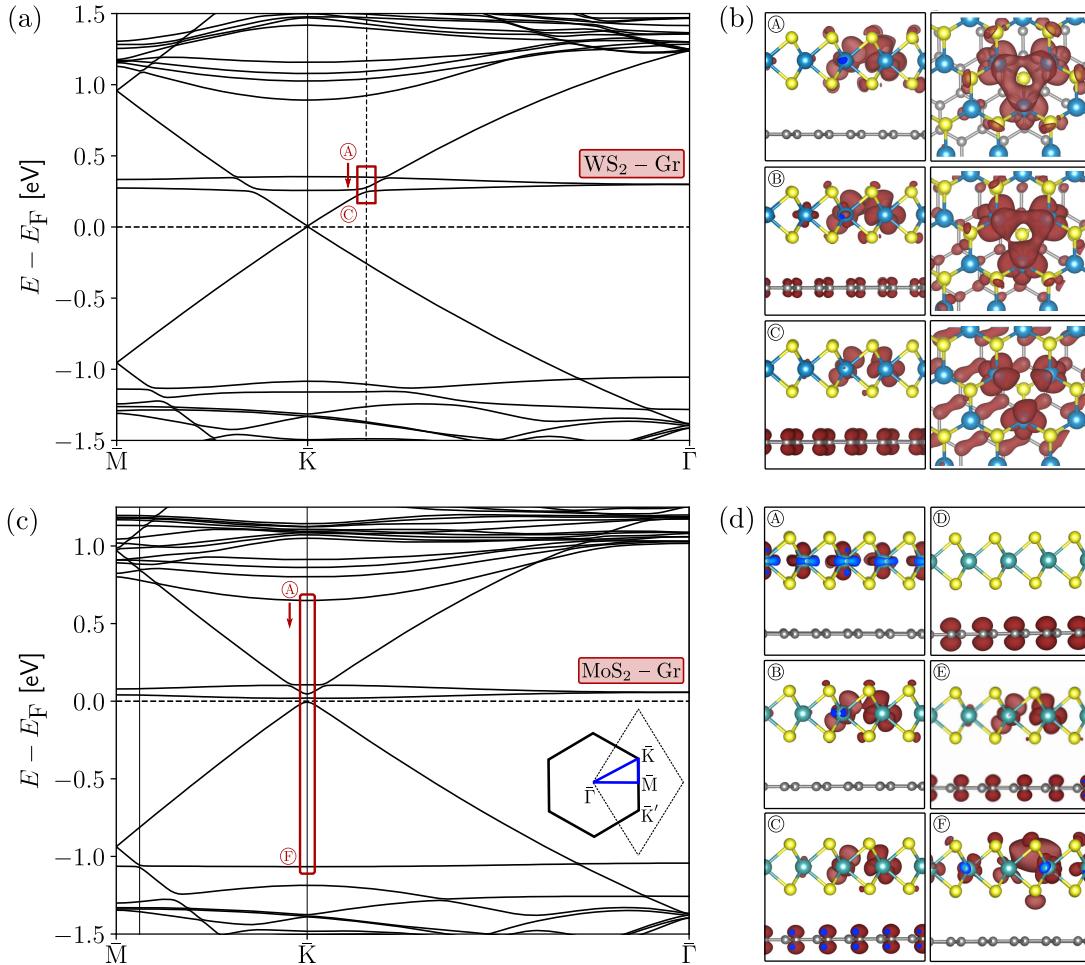


FIG. 2. (a) Band structure of the  $WS_2$ -Gr heterostructure in the absence of spin-orbit interaction computed along the  $\bar{M}-\bar{K}-\bar{\Gamma}$  path in the supercell. The chalcogen vacancy creates three defect bands, two of them appearing in the pristine  $WS_2$  gap and hybridizing with the graphene Dirac cone centered at  $\bar{K}$  and setting the charge neutrality point. (b) Kohn-Sham pseudo-charge density ( $|\Psi|^2$ ) for the empty states in the red box in panel (a), labeled by (A), (B) and (C) from highest to lowest energy. The left column shows the side view, while the right column displays the top view of the corresponding Kohn-Sham density. (c) Same as in (a) but for the  $MoS_2$ -Gr heterostructure. The inset shows the path in the hexagonal and rhombohedral mini-Brillouin zone. As in the  $WS_2$ -Gr interface, the chalcogen vacancy creates three defect bands, two of them appearing in the  $MoS_2$  pristine gap close to the Dirac point and opening a gap at the charge neutrality point of graphene located at  $\bar{K}$ . (d) Same as in (b) but for the states in the red box in panel (c), labeled from (A) to (F) from highest to lowest energy.

(about 50 meV above the charge neutrality point for  $\bar{\Gamma}$ ) and one in the valence band region [13, 40–42]. Interestingly, the graphene Dirac cone is now gapped (by about 50 meV) due to the hybridization of graphene and  $MoS_2$  defect bands occurring very close to the charge neutrality point. In Fig. 2 (d), we display the corresponding Kohn-Sham pseudodensities,  $|\Psi|^2$ , for the states located in the red box in panel (c). As expected from the presence of anti-crossings in the energy spectrum, we find again interfacial hybridized states with substantial wavefunction weight on both layers of the heterostructure, see panels (B) and (D).

### C. Band structure with spin-orbit interaction

In the presence of spin-orbit interaction, the valence and conduction bands split and the DFT band-gap from pristine  $WS_2$  is slightly reduced to  $\sim 1.75$  eV from  $\sim 1.95$  eV, see Fig. 3 (a). For comparison, the calculated  $WS_2$  bandgap (computed at the  $G_0W_0$  level for a  $5 \times 5$  supercell without adsorbed graphene) is equal to 2.8 eV while the experimental bandgap equals to 2.5 eV [13]. In general, not only screening but adsorption by graphene is expected to renormalize the  $G_0W_0$  band gap as well [44]. Therefore, in our calculations we underestimate the bandgap due to the well-known approximations in the exchange-correlation part of the DFT functional by at most  $\sim 0.75$  eV. The spin-orbit interaction also splits the

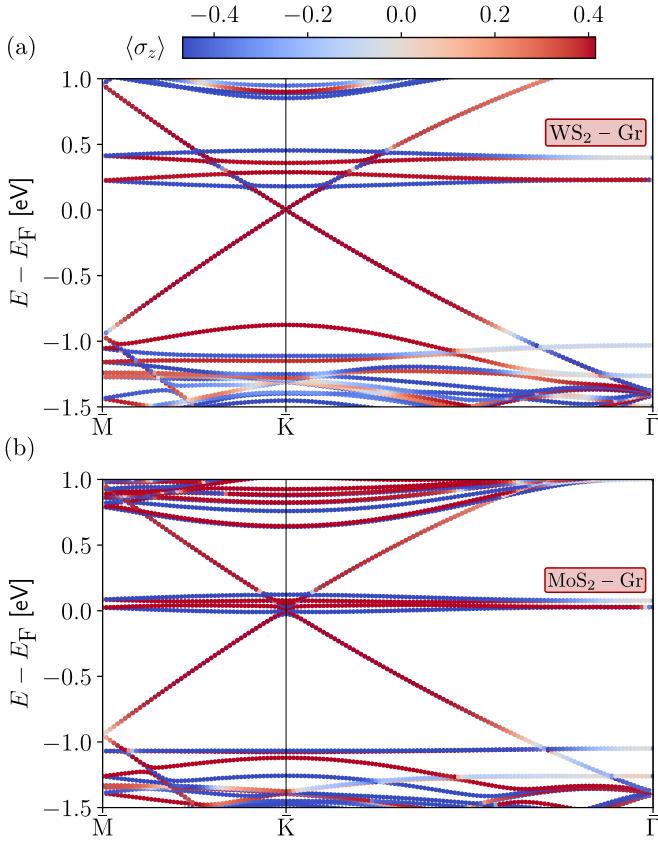


FIG. 3. Band structures of the  $\text{XS}_2\text{-Gr}$  heterostructures in the presence of spin-orbit interaction computed along the  $\bar{M}-\bar{K}-\bar{\Gamma}$  path in the supercell. Panel (a) corresponds to the case  $X = W$ , panel (b) to the case  $X = \text{Mo}$ . The spin-orbit splitting of the defect energy levels yields two degenerated levels at the  $\bar{\Gamma}$  point, which later splits into four levels clearly visible closer to the graphene Dirac cone centered at  $\bar{K}$ . The color code represents the spin magnetization projected in the direction perpendicular to the heterobilayer. The spin-orbit split defect levels have a well-defined spin direction in the vicinity of the  $\bar{K}$  point and become spin-degenerated both at high symmetry points  $\bar{M}$  and  $\bar{\Gamma}$ . The graphene Dirac cone remains spin-degenerated in the whole Brillouin zone as expected due to the small spin-orbit coupling present in the carbon atom as well as its regular structure.

pair of in-gap defect levels by 0.15 eV at the  $\bar{\Gamma}$  and  $\bar{M}$  points, suggesting that the spin-orbit coupling is strong. This spin-orbit splitting is well understood from the combination of the action of time-reversal and translational symmetry [45]. In particular, note that at the time-reversal invariant points for the honeycomb lattice  $\bar{\Gamma}$  and  $\bar{M}$ , all bands need to be doubly degenerated [45, 46], while close to the spin-degenerated Dirac cone, the vacancy in-gap states split due to their small but finite dispersion yielding a spectral structure of four spin-orbit split energy levels.

The scenario is analogous for the  $\text{MoS}_2\text{-Gr}$  interface, as displayed in Fig. 3 (b). Here, however, the spin-orbit coupling reduces the pristine band gap to  $\sim 1.7$  eV

from  $\sim 1.85$  eV as weaker spin-orbit interaction induces smaller splitting of the conduction and valence bands of  $\text{MoS}_2$  compared to  $\text{WS}_2$ . A value of  $\sim 2.8$  eV has been reported for the isolated  $\text{MoS}_2$  monolayer at the  $G_0W_0$  level [47]. In the case of a non-defected interface with a similar supercell size than the one employed here, previous calculations in the literature yield a pristine band gap upon adsorption of 1.73 eV at the DFT level and 2.43 eV at the  $G_0W_0$  level [48]. Therefore, we expect an underestimation of the band gap by  $\sim 0.7$  eV. The quasi-degenerated defect bands are split by  $\sim 0.05$  eV at the  $\bar{\Gamma}$  and  $\bar{M}$  points, roughly three time smaller than in  $\text{WS}_2$  due to weaker spin-orbit interaction in  $\text{MoS}_2$  compared to  $\text{WS}_2$ . For the same reason, the occupied defect band barely lifts its spin-degeneracy and the spin-orbit splitting is  $\sim 10$  meV at the  $\bar{K}$  point. Moreover, due to spin-orbit coupling, the lower empty defect band now appears below the graphene charge neutrality point. In other words, the system acquires partially a metallic character. However, note that this occurs only close to the  $\bar{K}$  point, where the defect bands bend due to residual defect-defect interaction between the periodic supercells, and far from the  $\bar{\Gamma}$  point, which is the point of interest for the defect states in the isolated vacancy limit. Note that this is a non-standard usage of the supercell approach, as we are employing the supercell (which contains vacancies in a  $\sim 3\%$  concentration) in order to study the dynamics at the level of the isolated vacancy. We also note that among the small effects of the geometry optimization in the presence of the vacancy is to revert the situation of partial occupancy of the vacancy energy bands by shifting the defect states again above the Dirac  $\bar{K}$  point. Because this effect is not relevant for the physics discussed in the rest of the manuscript, in our model we shall simply consider always two empty in-gap defect levels (*i.e* four states, once spin is taken into account).

To understand the impact of spin-orbit coupling beyond the energy spectrum, we also computed the spin magnetization of the defected  $\text{XS}_2\text{-Gr}$  heterostructures. The results for the spin expectation value in the direction perpendicular to the heterobilayer are shown by the color code of the electronic band structure in Fig. 3. An important observation that results from this calculation and that will be employed later for the modelling of the dynamics, is that hybridization of the vacancy levels and graphene preserves the spin orientation. This can be easily seen in Fig. 3 where the spin up defect level couples only to the spin up component of the graphene Dirac cone (here represented in red color); similarly, the spin down defect level only hybridizes with the spin down component (here shown in blue color). Therefore, we anticipate a competition between the spin and “orbital” (the angular momentum given by the lattice symmetry) degree of freedom to be manifested in the dynamical features of the population and depopulation of the vacancy once strong spin-orbit interaction is taken into consideration.

### III. MODEL

In order to microscopically describe the dynamics associated with the vacancy embedded in the TMDC–Gr interface, we consider an *ab initio* inspired model for a quantum dot coupled to a graphene “lead” (*i.e.* reservoir or bath) in a system-bath setup. Our model Hamiltonian is derived using electronic structure information, namely, the hybridization matrix elements between the TMDC and graphene monolayers are calculated by means of the combination of DFT and symmetry analysis, with the position of the vacancy levels being derived from DFT, which could potentially be corrected by a GW calculation.

We thus follow the standard partitioning of transport systems, in which the Hamiltonian is composed of three parts,

$$\hat{H} = \hat{H}_{\text{sys}} + \hat{H}_{\text{lead}} + \hat{H}_{\text{tun}}. \quad (1)$$

Here, the system Hamiltonian corresponds to our vacancy (“dot”) described as

$$\hat{H}_{\text{sys}} = \sum_{i\sigma} \epsilon_{i\sigma} \hat{d}_{i\sigma}^\dagger \hat{d}_{i\sigma} + \frac{E_c}{2} \hat{N}(\hat{N} - 1) + \hat{H}_{\text{SOC}}. \quad (2)$$

The first term describes an effective single-particle vacancy Hamiltonian without spin-orbit interaction. We represent by  $\hat{d}^\dagger$  (resp.  $\hat{d}$ ) the creation (resp. annihilation) operator of an electron in the dot, labeled by the orbital index  $i$  and spin  $\sigma$ ;  $\epsilon_{i\sigma}$  are the corresponding single-particle effective energies of the vacancy assumed to be independent of the quasi-momenta associated to the supercell [49]. The second term corresponds to the Coulomb interaction at the dot within the constant interaction picture. Here,  $E_c$  is the so-called charging energy and  $\hat{N} = \hat{d}_{i\sigma}^\dagger \hat{d}_{i\sigma}$  the total number operator. By construction, the charging energy participates in the physical processes only if more than one electron populates the vacancy. Due to the size of the defect impurity (see Fig. 2) we focus on the limit in which  $E_c$  is the largest energy scale of the system. Finally, the spin-orbit coupling term accounts for the splitting observed in the *ab initio* calculations and its form and impact on the dynamics will be discussed in Sec. V B.

The “lead” Hamiltonian corresponds to the graphene layer that acts as an fermionic reservoir/bath with

$$\hat{H}_{\text{lead}} = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}\sigma}^\dagger \hat{c}_{\mathbf{k}\sigma}. \quad (3)$$

Here,  $\hat{c}^\dagger$  (resp.  $\hat{c}$ ) corresponds to the electronic creation (resp. annihilation) operators in the graphene layer and  $\epsilon_{\mathbf{k}\sigma}$  is the low-energy graphene energy dispersion relation. Close to the  $\bar{\mathbf{K}}$  point, the graphene dispersion relation is isotropic and linear, *i.e.*  $\epsilon_{\mathbf{k}\sigma} = \hbar v_F |\mathbf{k}|$ , and with slope proportional to the Fermi velocity,  $v_F$ .

Finally, the tunneling Hamiltonian, that combines the bath and system operators, is given by

$$\hat{H}_{\text{tun}} = \sum_p \sum_{i\sigma} \sum_{\mathbf{k}} p t_{i\mathbf{k}\sigma}^p \hat{d}_{i\sigma}^p \hat{c}_{\mathbf{k}\sigma}^{\bar{p}}, \quad (4)$$

where  $p = \pm$  is employed as a short-hand notation to label the creation (+) and annihilation (-) operators, together with the conventions  $\hat{c}^+ = \hat{c}^\dagger$ ,  $\hat{c}^- = \hat{c}$ . We also adopt the notation  $\bar{p} = -p$ . The tunneling Hamiltonian describes the coupling between the vacancy at the TMDC layer and the graphene reservoir through the set of hybridization (or tunneling) matrix elements,  $t_{i\mathbf{k}\sigma}^\pm$ . These matrix elements are interpreted as the coupling between a state with momentum  $\mathbf{k}$  and spin  $\sigma$  in the graphene layer and a defect state in the  $\text{XS}_2$  layer with quantum numbers  $i$  and  $\sigma$ . By definition, the tunneling amplitudes satisfy  $t^+ = (t^-)^*$ . Note that tunneling by definition preserves the spin orientation due to the small vacuum distance that separates the two layers of the heterostructure as it was seen in Sec. II C and further discussed in Sec. V.

So far we have only considered DFT information in our model. However, a more realistic parametrization would include the correction to the position of the defect levels as obtained from a  $\text{G}_0\text{W}_0$  calculation. Under the assumption that the hybridization matrix elements are not affected at this level, as they essentially depend on the defect and graphene wavefunctions, it is possible to estimate the impact of screening in the many-body transition rates by shifting  $\epsilon_{i\sigma} \rightarrow \epsilon_{i\sigma} + \delta E$ , where  $\delta E$  would approximately be half of the change from the  $\text{G}_0\text{W}_0$  to the DFT pristine bandgap. In in Sec. IV of the SM we will discuss, by means of an illustrative case, how the model depends on the position of the defect energy levels. Reciprocally, we could also use the levels shift of a DFT band structure as a free parameter, if fit to experimental data is required, therefore, providing guidance in the interpretation of experimental results.

## IV. KINETIC EQUATION

### A. Generalities

We investigate the charge quenching of the strongly interacting impurity-bath system by means of the generalized master equation for the reduced density operator [34, 50–54]. This approach allows for the description of interference effects even for strongly interacting systems (see Ref. [54] and references therein). Moreover, its recent application to THz-STM [55] suggests a natural extension of the present study of  $\text{XS}_2$  impurities to setups with time dependent driving. Here, we briefly summarize the main steps to the derivation of a kinetic equation for the reduced density operator.

We start from the Liouville-von Neumann equation [56] for the density matrix, which can be written in a compact

form as  $\dot{\hat{\rho}} = \mathcal{L}\hat{\rho}$ , with  $\hat{\rho}$  being the density operator and the dot the time derivative. The object  $\mathcal{L}$  is the so-called Liouvillian superoperator defined as

$$\mathcal{L} := -\frac{i}{\hbar}[\hat{H}, \bullet], \quad (5)$$

with  $[\bullet, \bullet]$  the commutator and  $\hat{H}$  the total Hamiltonian combining the system, bath and hybridization terms, see Eq. (1). The equation of motion for the reduced density matrix, which contains degrees of freedom of the system only, can be obtained after tracing out the degrees of freedom of the fermionic bath,  $\hat{\rho}_{\text{red}} = \text{Tr}_{\text{lead}}\hat{\rho}$ . The bath is assumed to be at local thermodynamical equilibrium, characterized by a temperature  $T$  and chemical potential  $\mu$ , and, therefore, it is described by the equilibrium density matrix

$$\hat{\rho}_{\text{lead}} = \frac{e^{-\beta(\hat{H}_{\text{lead}} - \mu\hat{N})}}{Z_{\text{lead}}}. \quad (6)$$

Here  $\beta = (k_B T)^{-1}$  and  $Z_{\text{lead}}$  is the grand-canonical partition function which ensures the normalization of the lead density matrix,  $\text{Tr}_{\text{lead}}\hat{\rho}_{\text{lead}} = 1$ . Typically, it is assumed that the initial state for the integration of the Liouville-von Neumann equation is separable,  $\hat{\rho}(0) = \hat{\rho}_{\text{sys}}(0) \otimes \hat{\rho}_{\text{lead}}$ . Consequently, the entanglement between the system and the lead will appear as a result of the hybridization between them and manifests in the reduced density matrix dynamics at  $t > 0$ . Using the Nakajima-Zwanzig projector operator technique [57–59], an integro-differential equation for the reduced density matrix of the system,  $\dot{\hat{\rho}}_{\text{red}} = \mathcal{P}\dot{\hat{\rho}}$ , where  $\mathcal{P} := \text{Tr}\{\bullet\} \otimes \hat{\rho}_{\text{lead}}$  can be found

$$\dot{\hat{\rho}}_{\text{red}}(t) = \mathcal{L}_{\text{sys}}\hat{\rho}_{\text{red}}(t) + \int_0^t ds \mathcal{K}(t-s)\hat{\rho}_{\text{red}}(s). \quad (7)$$

The kernel superoperator,  $\mathcal{K}(t)$ , is given by

$$\mathcal{K}(t) = \mathcal{P}\mathcal{L}_{\text{tun}}\tilde{\mathcal{G}}(t)\mathcal{L}_{\text{tun}}\mathcal{P}, \quad (8)$$

and it is a combination of the Liouvillean superoperators, the projector and the propagator

$$\tilde{\mathcal{G}}(t) = \exp[\mathcal{L}_{\text{sys}} + \mathcal{L}_{\text{lead}} + (1 - \mathcal{P})\mathcal{L}_{\text{tun}}(1 - \mathcal{P})](t), \quad (9)$$

where each Liouvillian superoperator in Eqs. (7)–(9) is given by Eq. (5) after replacement of  $\hat{H}$  by  $\hat{H}_i$ .

Eq. (7) is formally exact, it contains memory effects and describes the system dynamics at all perturbative orders in the tunneling Hamiltonian. Due to its convolution form, it is natural to work in Laplace space. Employing the final value theorem, it is easy to show that the steady-state reduced density matrix satisfies the equation

$$[\mathcal{L}_{\text{sys}} + \tilde{\mathcal{K}}(0)]\hat{\rho}_{\text{red}}^{\infty} = 0, \quad (10)$$

where  $\hat{\rho}^{\infty} \equiv \hat{\rho}(t \rightarrow \infty)$  and

$$\tilde{\mathcal{K}}(0) = \lim_{\lambda \rightarrow 0^+} \int_0^{+\infty} dt e^{-\lambda t} \mathcal{K}(t). \quad (11)$$

To investigate the time evolution, the Markov approximation is applied to Eq. (7), thus leading to an equation of motion where we have discarded memory effects

$$\dot{\hat{\rho}}_{\text{red}}(t) = [\mathcal{L}_{\text{sys}} + \mathcal{K}_{\text{Markov}}]\hat{\rho}_{\text{red}}(t), \quad (12)$$

with  $\mathcal{K}_{\text{Markov}} = \tilde{\mathcal{K}}(0)$ .

## B. Sequential tunneling

From Eq. (8) we infer that the Markov approximation applies if the smallest time scale is the memory time. The latter is obtained from the correlator of the thermalized bath and reads  $\tau_{\text{corr}} = \hbar\beta$ . Therefore the condition  $\tau_{\text{corr}} \ll \tau$ , with  $\tau$  the characteristic evolution time of the system, translates into having the tunnelling rate  $\Gamma \sim 1/\tau$  smaller than the thermal energy,  $\hbar\Gamma \ll k_B T$ . The latter is in turn smaller than the charging energy,  $E_c$ . As the tunneling rate is proportional to powers of the modulus square of the tunneling matrix elements  $t_{ik\sigma}$  in Eq. (4), under the condition of weak tunnelling coupling, the Markovian kernel can be expanded perturbatively in a controlled way. Here, we only retain the first non-vanishing term in this perturbative expansion,  $\tilde{\mathcal{K}} \simeq \tilde{\mathcal{K}}^{(2)} + \mathcal{O}(\tilde{\mathcal{K}}^4)$ . From Eq. (8), this first contribution to the kernel therefore reads [59]

$$\tilde{\mathcal{K}}^{(2)}(0) = \mathcal{P}\mathcal{L}_{\text{tun}}\tilde{\mathcal{G}}_0\mathcal{L}_{\text{tun}}\mathcal{P}, \quad (13)$$

with

$$\tilde{\mathcal{G}}_0 = \lim_{\lambda \rightarrow 0^+} \frac{1}{\lambda - \mathcal{L}_{\text{sys}} - \mathcal{L}_{\text{lead}}}, \quad (14)$$

being the Laplace transform of the free propagator in the absence of tunneling from or to the bath. Note that the formal limit  $\lambda \rightarrow 0^+$  is to be taken only at the end of the calculation. The kernel in Eq. (13) corresponds to the so-called coherent sequential tunneling regime [60]. For coherent sequential tunneling, the time between each tunneling event corresponds to the largest time scale in the system and there cannot be coherence between individual tunneling events.

After integrating out the bath degrees of freedom in Eq. (13) the second order kernel reads

$$\begin{aligned} \tilde{\mathcal{K}}^{(2)}(0) = & -\lim_{\lambda \rightarrow 0^+} \frac{i}{2\pi} \sum_{\alpha, \alpha'} \sum_p \sum_{\sigma, \sigma'} \sum_{i,j} \int dE \alpha \alpha' \Gamma_{ij}^p(E, \sigma) \\ & \times \mathcal{D}_{i\sigma\alpha'}^{\bar{p}} \frac{f^{(p\alpha)}(E)}{pE - i\hbar\mathcal{L}_{\text{sys}} + i\lambda} \mathcal{D}_{j\sigma\alpha}^p, \end{aligned} \quad (15)$$

where we have used

$$\text{Tr} \left( c_{\mathbf{k}\sigma}^{p,\alpha} c_{\mathbf{k}'\sigma'}^{p',\alpha'} \hat{\rho}_{\text{lead}} \right) = \delta_{\mathbf{k}\mathbf{k}'} \delta_{\sigma\sigma'} \delta_{p\bar{p}'} f^{(p\alpha')}(E). \quad (16)$$

Here, the Liouville indices  $\alpha, \alpha' = \pm 1$ ,  $\mathcal{D}_{\alpha}$  are superoperators defined by their action from the left/right (resp.

$+/-$ ) on the density matrix

$$\mathcal{D}_{i\sigma+}^p \hat{\rho} = \hat{d}_{i\sigma}^p \hat{\rho}, \quad (17)$$

$$\mathcal{D}_{i\sigma-}^p \hat{\rho} = \hat{\rho} \hat{d}_{i\sigma}^p, \quad (18)$$

and  $f^+(E) = \{\exp[\beta(E - \mu)] + 1\}^{-1}$ ,  $f^-(E) = 1 - f^+(E)$  are Fermi distribution functions.

The geometry and the hybridization of the single-particle defect states with the fermionic reservoir are contained in the single particle (tunneling) rate matrix (also known as tunneling self energy) [34, 54]

$$\Gamma_{ij}^p(E, \sigma) = \frac{2\pi}{\hbar} \sum_{\mathbf{k}} t_{i\mathbf{k}\sigma}^p t_{j\mathbf{k}\sigma}^p \delta(\epsilon_{\mathbf{k}\sigma} - E), \quad (19)$$

where  $\Gamma_{ij}^+(E, \sigma) = \Gamma_{ji}^-(E, \sigma)$ . This matrix extends the concept of tunneling rate and its form determines the strength of the interference effects between quasi-degenerated many-body states [32, 53]. The matrix  $\Gamma(E)$  is defined in the relevant sector of the single-particle Hilbert space, spanned in this study by the two spin-full empty levels of the  $\text{XS}_2$  vacancy. The calculation of the rate matrix, combining *ab initio* results and the symmetry analysis from Sec. III in the SI, is now performed in Sec. V.

## V. RELAXATION RATES AND DYNAMICS

We now discuss the relaxation rates and kinetic equations for our minimal model. As  $E_c$  is the larger energy scale, we consider only the addition of a single electron to the vacancy. Therefore, the interaction term in Eq. (1) does not contribute explicitly to the transition rates. We denote the ground state of the system with an empty vacancy by  $|N\rangle$ ; this many-body state corresponds to a state in which two levels of the impurity are empty and the one far below the Fermi level is fully occupied, as in the equilibrium situation described by DFT. Correspondingly, the charged system in which an additional electron has tunneled into the vacancy is represented by the many-body state  $|N+1\rangle = \hat{d}_{i\sigma}^\dagger |N\rangle$  with one additional electron created at the impurity level described by the quantum numbers  $(i, \sigma)$ .

We envisage two possible pathways leading to a singly charged impurity, as shown schematically in Fig. 4 (a). In the first case, a photoexcited electron from the occupied dot levels is promoted to one of the empty levels of the vacancy. Due to hybridization of the defect states and the valence band at the  $\text{XS}_2$  layer, a hole subsequently tunnels towards the graphene layer. Since this process involves photoexcitation of the electron (photoemission for the reverse process) it naturally involves excitons. We shall not consider it here, for the sake of simplicity. The empty dot levels can be populated in a second way by increasing the chemical potential of the graphene layer by means of an external back gate or by doping. In this case, electron tunneling occurs directly from the reservoir

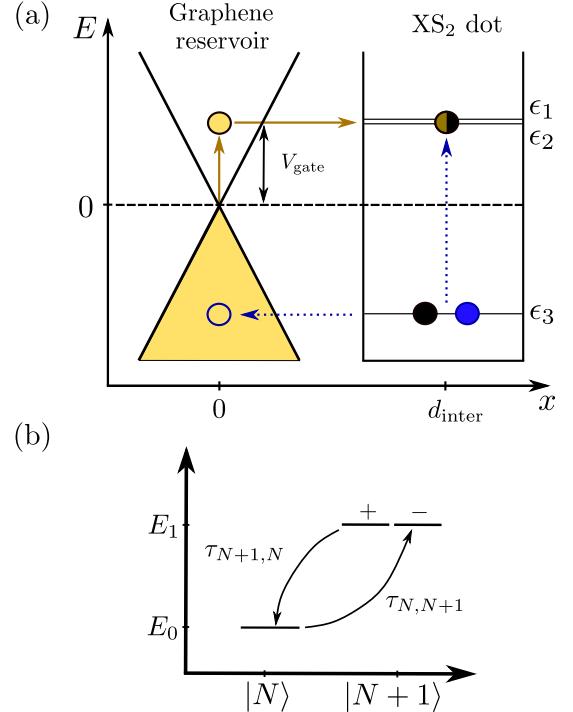


FIG. 4. (a) Sketch of the low energy landscape of the  $\text{XS}_2$  dot (here  $X = W$ ) and the graphene reservoir, providing a structured DoS. We also show the potential single-particle electronic transitions occurring within or to the vacancy. The graphene reservoir is separated by a distance  $d_{\text{inter}}$  from the dot-containing “active” layer. The position of the vacancy energy levels  $\epsilon_i$  are determined with respect to the graphene charge neutrality point ( $\mu = E_F$ , set as the zero of energies). We represent two possible pathways for a  $N \rightarrow N+1$  transition: (i) a photoexcited electron is promoted to one of the two empty quasi-degenerate levels leaving behind a hole in the originally doubly occupied state of the dot, which can be refilled with the hole tunneling towards graphene (blue dotted arrows) (ii) by means of an electric field created by gate voltage or doping the chemical potential of graphene can be changed and an electron tunnel from the graphene layer into one of the empty vacancy levels (yellow solid arrows). (b) Representation of the many-body Hilbert space and the transition times,  $\tau_{\alpha,\beta}$ , between the many-body initial state  $\alpha$  and final state  $\beta$  (characterized by the total particle number  $N, N+1$ ) in the absence or presence of very weak spin-orbit coupling. The many-body energies are considered in the grand canonical ensemble,  $E_0 = \varepsilon_0 - N\mu$  and  $E_1 = \varepsilon_1 - (N+1)\mu$ , with  $\mu \equiv E_F$  at zero gate voltage and  $\varepsilon_{0,1}$  are the total energies.

to the empty single-particle level at the vacancy. The reverse process would occur by tunneling from the previously occupied vacancy states into the graphene layer in a similar way.

### A. Weak spin-orbit interaction

**Single particle rate matrix.** If the spin-orbit interaction is weaker than the coupling to the reservoir, as it is for MoS<sub>2</sub>, we can, at first, neglect  $\hat{H}_{\text{SOC}}$  and concentrate on the single-particle Hamiltonian including only the degenerate transport levels, the tunnelling term and the reservoir, *e.g.*, the two empty defect “bands” tunnel-coupled to the graphene Dirac cone. We will apply the same analysis also to the WS<sub>2</sub>–Gr heterostructure, this time to better highlight the effects of spin-orbit interaction, presented later in Sec. V B. In both cases, the relevant Fock space has  $\dim \mathcal{F} = 5$ , with the  $N$ -sector (vacuum) being one-dimensional (as it corresponds to a closed-shell ground state) and the  $N + 1$ -sector four-dimensional (due to spin and orbital degrees of freedom). Since the  $N + 1$ -sector is a simple single-particle sector, its dimension coincides with the one of the (single-particle) Hilbert. Consequently,  $\Gamma$  is a four-dimensional matrix which, in the absence of spin-orbit interaction, simply factorizes into an orbital and a trivial spin component:  $\Gamma = \Gamma_{\text{orb}} \otimes \mathbf{1}_2$ . Moreover, we show in Sec. III in the SM that the orbital component of the single particle rate matrix is also diagonal with the form

$$\Gamma_{uu'}^{\text{orb}}(E) = \frac{\pi^2}{2\hbar} \frac{A_{\text{sc}}}{A_c} |t_u|^2 \rho_{\text{Gr}}(E) \delta_{uu'}, \quad (20)$$

where  $u$  and  $u'$  are the quantum number characterizing the symmetric/antisymmetric defect states,  $A_{\text{sc}}$  is the area of the supercell and the DoS of graphene is given by  $\rho_{\text{Gr}}(E) = 2A_c|E|/(\pi\hbar^2 v_F^2)$  [11]. Here, we considered the unit cell area of graphene given by  $A_c = 3\sqrt{3}a^2/2$  with  $a = 1.42$  Å the graphene Fermi velocity  $v_F \sim 10^6$  m/s and the ratio  $A_{\text{sc}}/A_c = 25$ . Thus,  $\Gamma^{\text{orb}}$  can also be cast into the form

$$\Gamma^{\text{orb}} = \frac{\Gamma_{uu} + \Gamma_{\bar{u}\bar{u}}}{2} \mathbf{1}_2 + \frac{\Gamma_{uu} - \Gamma_{\bar{u}\bar{u}}}{2} \boldsymbol{\tau}_z, \quad (21)$$

Note that this object has dimensions of [time]<sup>-1</sup>.

The modulus of the hybridization matrix elements can be directly extracted from the DFT band structure by looking at the anti-crossing of the energy levels close to  $\bar{K}$ , considering this problem as a local two-level problem in  $\mathbf{k}$ -space with Hamiltonian

$$\hat{H}_{\text{loc}} = \begin{pmatrix} E_1 & t \\ t^* & E_2 \end{pmatrix}. \quad (22)$$

The straightforward diagonalization of Eq. (22) gives the eigenenergies

$$E_{\pm} = \frac{E_1 + E_2}{2} \pm \frac{1}{2} \sqrt{(E_1 - E_2)^2 + 4|t|^2}. \quad (23)$$

In the literature [61, 62], this expression is sometimes known as “avoided crossing formula”. We then estimate  $|t|$  by choosing the zero of energy at the point where the unperturbed energies are degenerated,  $|t| = (E_+ - E_-)/2$ .

From the DFT band structures [Figs 2 (a) and (c)] we find  $|t_u^{\text{Mo}}| = 12$  meV,  $|t_u^{\text{W}}| = 8$  meV, together with  $E_u = 57$  meV for MoS<sub>2</sub>–Gr and  $|t_u^{\text{W}}| = 15$  meV,  $|t_u^{\text{W}}| = 5$  meV with  $E_u = 298$  meV for WS<sub>2</sub>–Gr.

**Transition rates.** Using Eq. (15), we can obtain the transition rates for the charging process between an initial uncharged and final charged vacancy,  $N \rightarrow N + 1$ , as well as the reverse - discharging - process  $N + 1 \rightarrow N$ , in the lowest non vanishing perturbative order and as shown in Fig. 4 (b). As  $\Gamma_{ij}(E, \sigma)$ , as well as the Fermi function  $f^{(p\alpha)}(E)$  are smooth functions, we can evaluate the integral by using the Sokhotski-Plemelj theorem over the real line [63]

$$\lim_{\lambda \rightarrow 0^+} \int_{-\infty}^{+\infty} d\omega \frac{h(\omega)}{\omega + i\lambda} = \text{p.v.} \int_{-\infty}^{+\infty} d\omega \frac{h(\omega)}{\omega} - i\pi h(0), \quad (24)$$

where p.v. stands for principal value and  $h(\omega)$  is a complex-valued function [64]. From the real part of  $\tilde{\mathcal{K}}^{(2)}(0)$  acting on the reduced density matrix  $\hat{\rho}_{\text{red}}$  we extract the transition rate  $W_{\alpha, \beta}$  between an initial  $\alpha$  and final  $\beta$  many-body state. The projection on the many-body Hilbert space of the vacancy yields

$$\begin{aligned} W_{N, N+1} &= \sum_u W_{0 \rightarrow u}, \\ &= \sum_{u\sigma} \int dE \langle N | \hat{d}_{u\sigma} | N + 1 \rangle \langle N + 1 | \hat{d}_{u\sigma}^\dagger | N \rangle \\ &\times \Gamma_u(E) f^+(E) \delta(E - \Delta E_{0,u}), \end{aligned} \quad (25)$$

where  $\Gamma_u := \Gamma_{uu}$  and by definition  $\Delta E_{0,u} = E_u - E_0 > 0$ . This expression simplifies to

$$W_{N, N+1} = 2 \sum_u \Gamma_u(\Delta E_{0,u}) f^+(\Delta E_{0,u}), \quad (26)$$

where the factor of two stems from the spin degeneracy. From this rate, we define the transition time for the population of the vacancy by an additional electron as  $\tau_{N, N+1} := 1/W_{N, N+1}$ . The discharging process in which the additional electron initially present at the vacancy leaves the dot by tunneling back towards the graphene layer is given by

$$W_{N+1, N} = 2 \sum_u \Gamma_u(\Delta E_{0,u}) f^-(\Delta E_{0,u}). \quad (27)$$

The corresponding transition time can be defined analogously as for the charging process,  $\tau_{N+1, N} := 1/W_{N+1, N}$ .

Evaluation of Eqs. (26)–(27) at room temperature ( $T = 300$  K) and chemical potential  $\mu = E_F$  ( $\mu \equiv 0$  at the charge neutrality point) yields  $\tau_{N, N+1}^{\text{Mo}} = 29.2$  ps and  $\tau_{N+1, N}^{\text{Mo}} = 3.1$  ps for the MoS<sub>2</sub>–Gr heterobilayer and  $\tau_{N, N+1}^{\text{W}} = 47.1$  ns and  $\tau_{N+1, N}^{\text{W}} = 0.5$  ps for the WS<sub>2</sub>–Gr heterobilayer. The charging and discharging tunnelling times are crucially influenced by the number of states in the graphene reservoir available to tunnel at the resonant

energy. Ultimately, the Dirac dispersion of graphene (linear DoS) and the Pauli exclusion principle (Fermi function) set their values.

Thus, the substantially longer charging time for the case of WS<sub>2</sub> compared to MoS<sub>2</sub> can be understood from the higher energy difference between the empty vacancy states and the chemical potential of graphene in the case of WS<sub>2</sub> as compared to the ones of MoS<sub>2</sub> [see Figs. 2 (a) and (c)]. Even at room temperature and despite the larger DoS in graphene, only few thermally excited electrons can populate the impurity. On the contrary, the discharging time is shorter for WS<sub>2</sub> compared to MoS<sub>2</sub>. This is a consequence of the larger amount of available states to tunnel to in graphene for the WS<sub>2</sub>–Gr interface, which manifests itself in the density of states, as the Pauli exclusion principle cannot hinder the population of empty states. We note that graphene being a semimetal leads to a counter intuitive behavior of the transition times compared to a metallic reservoir because the closer the state is to the Dirac point, the less available states exists to tunnel from and to.

The calculated relaxation time at zero chemical potential roughly agree to the one experimentally found in tr-ARPES experiments, in which the electron is photoexcited to a state in the conduction band of the TMDC layer [7, 8]. In the experiment, electronic lifetimes of the order of  $\sim 1$  ps are measured from the pump-probe signal for the WS<sub>2</sub>–Gr heterostructure. Though, a direct comparison between our calculations and these experimental lifetimes is *strictu senso* not easy, because in this experiment, first excitons are resonantly excited as initial state for the dynamics of hot carriers. Although thermalization and exciton dissociation are expected to occur at shorter timescales  $\lesssim 100$  fs, compatible with a binding energy of  $\sim 200$  meV found in recent calculations for the “A-like” exciton [65], the exciton dissociation and recombination to the defect states can strongly influence the dynamics. Second, the measured dynamics would involve also the pristine conduction (or valence, in the case of hole transfer) bands, a scenario different from what we are considering here, where those bands are assumed to be far in energy from the relevant defect transport channels and therefore do not participate in the dynamics. Finally, our theory considers the extremely dilute limit in which the vacancies are relatively isolated, this scenario may change in the experiment and the charge transfer rates differ if the vacancy concentration is higher. In any circumstance, the defect associated transport channel acts as a mechanism that slows down the charge transfer.

Bath temperature, electric fields or changes in the position of the single-particle defect energies due to screening can have a dramatic effect in the transition rates. We analyze here the behavior of the transition rates with the electric field created by a voltage drop in the direction perpendicular to the heterostructure (the analysis of the effect of temperature and defect level energy shifts due to screening is given in Sec. IV of the SI). In Fig. 5 (a) we display the transition rates as a function of the

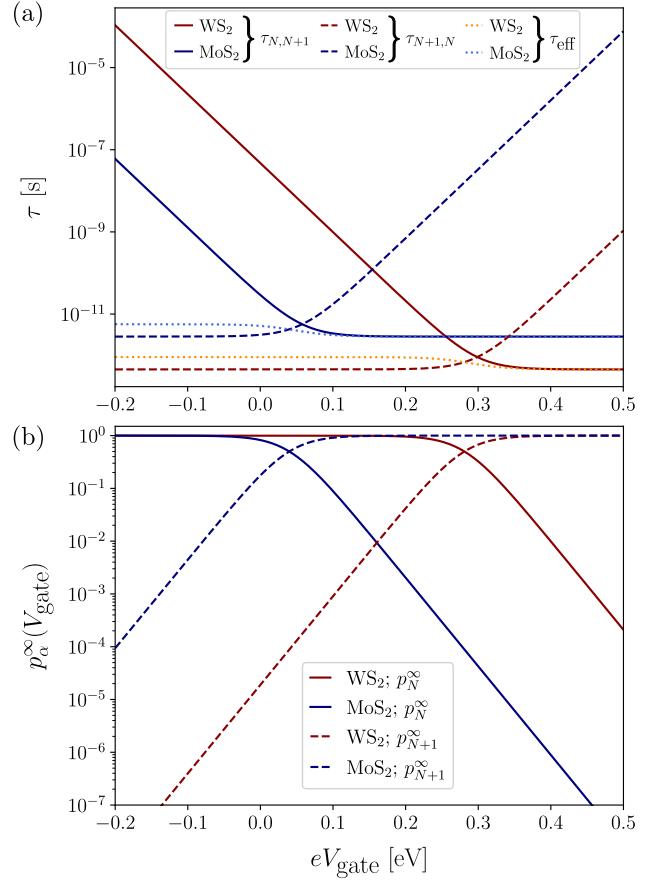


FIG. 5. (a) Electronic transition times,  $\tau$ , for MoS<sub>2</sub> and WS<sub>2</sub> vacancy states as a function of the gate voltage,  $V_{\text{gate}}$  measured in eV. The solid and dashed lines correspond to the transitions  $\alpha \rightarrow \beta$  with  $\alpha, \beta \in \{N, N + 1\}, \alpha \neq \beta$  as defined in Eqs. (26) and (27). The dotted line corresponds to the effective transition time associated to the semiclassical dynamics defined from Eq. (30) and controlling the dynamics from a local to global equilibrium. For simplicity, we set the chemical potential at zero gate voltage to be zero. (b) Corresponding stationary semiclassical probability of occupation of the many-body levels,  $p_\alpha^\infty$ , as a function of the gate voltage. As anticipated, the stationary populations follow the Fermi distribution function.

gate voltage,  $eV_{\text{gate}}$ , which may be tuned by an external gate located under the graphene layer and which creates a voltage drop between graphene and TMDC layers. The window of voltage drop between the two layers considered here corresponds to a range of electric fields between the layers of the order of  $E \lesssim 3$  V/nm. This is consistent with electric fields considered in previous theoretical works [66] and it is substantially smaller than the experimental dielectric breakdown points of these materials [67]. We first observe that for each material whether a charging ( $N \rightarrow N + 1$ ) occurs faster compared to a discharging ( $N + 1 \rightarrow N$ ) process depends on the value of the chemical potential ( $\simeq$  gate voltage) with respect to the position of the defect levels. This can be easily understood from the functional form of  $f^+(x)$  and  $f^-(x)$

in Eqs. (26)-(27): for  $\mu < \Delta E_{0,u}$ , charging is slower because the tail of the Fermi function (that has a typical width of a few  $k_B T$ ) is small and there are no occupied states that can tunnel towards the vacancy, while for  $\mu > \Delta E_{0,u}$  the reverse situation occurs and discharging becomes slower because there are very few (depending on temperature) empty states in graphene capable to receive electrons tunnelling from the vacancy. The charging and the discharging only have the same transition times and rates if the chemical potential aligns precisely with the defect energies,  $\mu = \Delta E_{0,u}$  (thus  $\tau_{N,N+1} = \tau_{N+1,N}$ ). This is a trivial consequence of  $f^+(\Delta E_{0,u}) = f^-(\Delta E_{0,u})$ , together with the fact that  $\Gamma$  is a diagonal matrix (therefore  $\Gamma^+ = \Gamma^-$ ).

We now discuss the differences between the two TMDCs in the  $\text{XS}_2\text{-Gr}$  interface. We analyze the case  $\mu \gg \Delta E_{0,u}$ , the scenario for  $\mu \ll \Delta E_{0,u}$  can be deduced with the mapping  $\tau_{N,N+1} \leftrightarrow \tau_{N+1,N}$  upon reversing the inequality. If the chemical potential is substantially larger than the defect energy,  $\mu \gg \Delta E_{0,u}$ , we observe that  $\tau_{N,N+1}$  saturates and the charging is faster for  $\text{X} = \text{W}$  compared to  $\text{X} = \text{Mo}$ . This is a consequence of the  $\text{WS}_2\text{-Gr}$  heterostructure having larger single-particle tunneling rate compared to  $\text{MoS}_2\text{-Gr}$ , mainly due to the larger number of states in graphene to tunnel from in the former compared to the latter [ $\rho_{\text{Gr}}(\Delta E_{0,u}^{\text{W}}) > \rho_{\text{Gr}}(\Delta E_{0,u}^{\text{Mo}})$ ]. The differences in the charging and descharging rates between the two heterobilayers can be tuned with the gate voltage. For example, for  $\mu > 0$ , there is a voltage threshold for which the relative ordering between the charging times of the two materials changes [in other words, the continuous lines in Fig. 5 (a) cross]. This point,  $\mu_{\text{cr}}$ , marks the change from  $\tau_{N,N+1}(\text{W}) > \tau_{N,N+1}(\text{Mo})$  if  $\mu < \mu_{\text{cr}}$  to  $\tau_{N,N+1}(\text{W}) < \tau_{N,N+1}(\text{Mo})$  if  $\mu > \mu_{\text{cr}}$ . Though their difference can be modulated, in the discharging,  $\text{WS}_2$  shows a smaller transition time  $\tau_{N+1,N}$  (faster discharging) compared to  $\text{MoS}_2$ , independently of the chemical potential. This result is attributed to the higher energy of the defect state and to its lower tunneling barrier to the graphene layer for  $\text{X} = \text{W}$  as compared to  $\text{X} = \text{Mo}$ .

**Semiclassical dynamics.** The kinetic equation Eq. (12) describes a time-local dynamics for the vacancy in presence of the graphene bath still involving quantum coherences between quasi-degenerate vacancy states. The latter are captured by the coherences (*i.e.* the off-diagonal elements of the density matrix  $\hat{\rho}_{\text{red}}$ , written in the energy eigenbasis). Starting from Eq. (12) it is easy to derive the semiclassical Pauli master equation only involving the probability of occupation of a given many-body state [68], written in terms of the charging and discharging rates, already discussed in the previous section. Thus, the probability,  $p_\alpha := p_\alpha(t)$ , of finding the heterostructure in the state  $\alpha$  with given particle number

at any time  $t$ , is obtained from the solution of

$$\dot{p}_\alpha = \sum_\beta [p_\beta W_{\beta,\alpha} - p_\alpha W_{\alpha,\beta}], \quad (28)$$

together with the conservation of probability,  $\sum_\alpha p_\alpha = 1$  [which results from  $\text{Tr}(\hat{\rho}) = 1$ ] and a given initial condition. We further assume a fast local thermalization between the tunnelling events or local thermal equilibrium, so that the population of the degenerate energy levels has to be the same,  $p_u = p_{\bar{u}}$ . As a consequence, we reformulate Eq. (28) as a two-level problem characterized by the populations  $p_N$  and  $p_{N+1} = p_u + p_{\bar{u}}$  in order to investigate the dynamics towards a final state of global equilibrium of the vacancy in contact with the thermal bath. The analytical solution of the associated system of equations gives the vector of probabilities

$$\mathbf{p}(t) = \mathbf{p}^\infty + p_0 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-W_{\text{eff}} t}. \quad (29)$$

We therefore find that the vacancy relaxes from an initial state following an expected exponential decay law typical for a tunneling process with an effective rate

$$W_{\text{eff}} := W_{N,N+1} + \tilde{W}_{N+1,N}, \quad (30)$$

where  $\tilde{W}_{N+1,N} := W_{N+1,N}/2$ . This effective rate defines an effective relaxation time  $\tau_{\text{eff}} := 1/W_{\text{eff}}$  for the population dynamics of the isolated vacancy. Substitution using Eqs. (20), (26) and (27) gives  $\tau_{\text{eff}}^{\text{W}} = 0.9$  ps and  $\tau_{\text{eff}}^{\text{Mo}} = 5.2$  ps, as the time scale at which the system reaches equilibrium. This value is of the same order of magnitude of the electronic relaxation times obtained from tr-ARPES measurements. We show  $\tau_{\text{eff}}$  in Fig. 5 (a) for  $\text{X} = \text{W}$ ,  $\text{Mo}$  as a function of  $V_{\text{gate}}$ . In comparison to  $\tau_{N,N+1}$  and  $\tau_{N+1,N}$ , we find that this relaxation rate is a smoother function of the voltage, roughly one order of magnitude larger in  $\text{MoS}_2$  compared to  $\text{WS}_2$ .

The vector  $\mathbf{p}^\infty$  corresponds to the stationary populations of the states  $|N\rangle$  and  $|N+1\rangle$

$$\mathbf{p}^\infty = \begin{pmatrix} p_N^\infty \\ p_{N+1}^\infty \end{pmatrix} = \frac{1}{W_{\text{eff}}} \begin{pmatrix} \tilde{W}_{N+1,N} \\ W_{N,N+1} \end{pmatrix}. \quad (31)$$

This object only depends on the ratio of many-body transitions rates. Note that, as expected for second order processes, the ratio of the equilibrium populations is proportional to the Boltzmann factor [59]. Indeed, it is easy to check that this ratio is  $p_{N+1}^\infty/p_N^\infty = g_{N+1} \exp(-\beta \Delta E_{0,u})$  where  $g_{N+1} = 2$  is the (Boltzmann) degeneracy of the state  $|N+1\rangle$  (by construction,  $g_N = 1$ ). The prefactor  $p_0$  in Eq. (29) depends on the initial condition for the time evolution. If the system initially had  $N$  electrons, then,  $p_0 = W_{N,N+1}/W_{\text{eff}}$ , while if the system initially was in the many-body state with particle number equal to  $N+1$  it reads  $p_0 = -W_{N+1,N}/W_{\text{eff}}$ .

The stationary many-body populations for each of the two heterostructures as a function of the gate voltage,

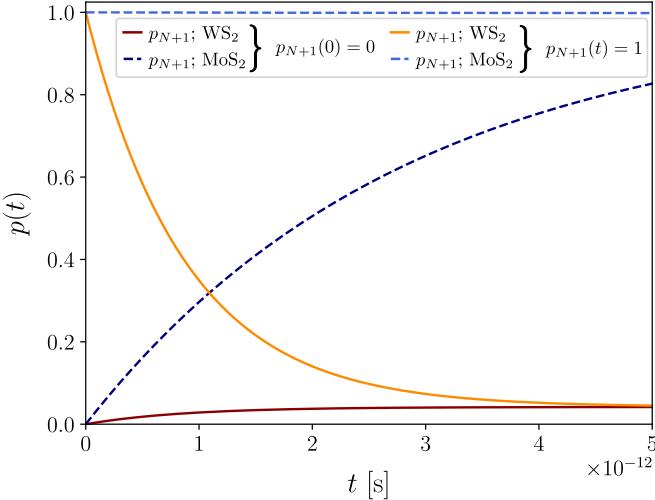


FIG. 6. Population,  $p(t)$ , of the many-body levels as a function of time,  $t$  (measured in s). The dynamics is computed for the representative gate voltage,  $V_{\text{gate}} = 0.2$  eV. We only show the population for the state  $|N+1\rangle$ , the population of the state  $|N\rangle$  being obtained as  $p_N(t) = 1 - p_{N+1}(t)$ .

$V_{\text{gate}}$ , and computed with realistic *ab initio* parametrization are shown in Fig. 5 (b). As expected, the stationary population follows a Fermi distribution function since at  $t \rightarrow +\infty$  the system has relaxed to the equilibrium distribution provided by the thermal bath. The values of the populations are controlled by the transition rates only and shifted by the difference between the defect energies for each material. In the large positive  $eV_{\text{gate}}$  limit, we have  $p_N^\infty \rightarrow 0$  and  $p_{N+1}^\infty \rightarrow 1$  as the transition time associated to the charging of the vacancy becomes smaller and saturates, as seen in Fig. 5 (a). Due to the symmetry relation  $\tau_{N,N+1} \rightarrow \tau_{N+1,N}$  with respect to the defect energy, the situation is reversed for large negative  $V_{\text{gate}}$ .

In Fig. 6 we show an example of the population dynamics for a given voltage gate,  $V_{\text{gate}} = 0.2$  V, following Eq. (29). We only display the population of the state  $|N+1\rangle$ , for two initial conditions,  $p_{N+1}(0) = 0$  and  $p_{N+1}(0) = 1$ . The population shows the expected exponential decay law, controlled by  $W_{\text{eff}}$ , as well as the asymptotic limit given in Fig. 5 (b) which corresponds to the situation of global thermodynamic equilibrium. Note that properly chosen combinations of material and  $V_{\text{gate}}$  can strongly reduce the charge relaxation, also depending on the initial conditions, effectively inhibiting charge quenching at the impurity, as it occurs here for the initially charged vacancy at the MoS<sub>2</sub>–Gr interface (see light blue dashed curve).

**Lamb shift and quantum internal dynamics.** The imaginary part of  $\tilde{\mathcal{K}}^{(2)}$  in Eq. (15) acting on the reduced density matrix  $\hat{\rho}_{\text{red}}$ , gives the so-called Lamb shift correction [32, 50]. This term induces an effective internal dynamics due to a renormalization of the system Hamiltonian generated by virtual fluctuations of

the particle number (which is conserved in average [63]). The Lamb shift correction is only relevant for quasi-degenerate states and, in presence of a single bath, it can not influence the stationary state, thus influencing, in general, only the short time dynamics of the system. For its derivation, one proceeds as in Ref. [54] to compute the principal value of the integral from Eq. (24), where a particular attention should be given to the energy dependence of the density of states in graphene. The second order kernel (15) is therefore split into the tunneling and Lamb-shift terms  $\tilde{\mathcal{K}}^{(2)} = \tilde{\mathcal{K}}_{\text{T}}^{(2)} + \tilde{\mathcal{K}}_{\text{LS}}^{(2)}$  and the equation of motion (12) can be rewritten in the form

$$\dot{\hat{\rho}}_{\text{red}} = -\frac{i}{\hbar} [\hat{H}_{\text{sys}} + \hat{H}_{\text{LS}}, \hat{\rho}_{\text{red}}] + \tilde{\mathcal{K}}_{\text{T}}^{(2)} \hat{\rho}_{\text{red}}, \quad (32)$$

where the commutator contains the Lamb-shift Hamiltonian, which, in the one particle sector, reads

$$\hat{H}_{\text{LS}} = \frac{\hbar}{2\pi} \sum_{u\sigma} \left[ \hat{d}_{u\sigma} \Upsilon_u(\hat{H}_{\text{sys}}) \hat{d}_{u\sigma}^\dagger + \hat{d}_{u\sigma}^\dagger \Upsilon_u(\hat{H}_{\text{sys}}) \hat{d}_{u\sigma} \right]. \quad (33)$$

In the previous equation we have introduced the function

$$\Upsilon_u(\hat{H}_{\text{sys}}) = \text{p.v.} \int dE \frac{\Gamma_u(E) f^+(E)}{E - \hat{H}_{\text{sys}} + E_{\text{av}}}, \quad (34)$$

where  $E_{\text{av}} = (E_u + E_{\bar{u}})/2$ .

As anticipated above, the dynamics induced by the Lamb shift Hamiltonian only influences, in the presence of a single bath, the short time scales. It produces pseudo-exchange fields leading to a “precession-like” dynamics in the quasi-degenerate subspace [54, 55]. The latter is particularly relevant in ultrafast pump-probe experiments. Here we concentrate, instead, on a semi-classical dynamics, as we always assume local thermal equilibrium, thus leaving the full analysis of the coherent dynamics of the impurity quasi-degenerate states for future work.

## B. Strong spin-orbit interaction

**Single-particle rate matrix.** In order to include spin-orbit interaction in our model, we could start from the superposition of the atomic contributions [55]

$$\hat{H}_{\text{SOC}} = \sum_{\alpha} \xi_{\alpha} \hat{\mathbf{L}}_{\alpha} \cdot \hat{\mathbf{S}}_{\alpha}, \quad (35)$$

where the summation runs over all atoms and orbitals, with  $\hat{\mathbf{L}}_{\alpha}$  being the angular momentum operator and  $\hat{\mathbf{S}}_{\alpha}$  the spin operator. Instead, we take advantage of the fact that the atomic spin-orbit interaction is already incorporated by construction into the Kohn-Sham Hamiltonian and project its spin-orbit contribution onto the vacancy Kohn-Sham empty orbitals that form the relevant Hilbert space discussed in Sec. V A. The local C<sub>3v</sub> symmetry of

the vacancy implies that the degenerate orbitals are associated to the quasi-angular momenta  $L_z = \pm 1$ . From the decomposition

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2} [L_+ S_- + L_- S_+] + L_z S_z, \quad (36)$$

with  $L_{\pm} = L_x \pm iL_y$  (similarly  $S_{\pm} = S_x \pm iS_y$ ), we thus obtain that only the diagonal term containing  $L_z$  can contribute to the projection. The resulting effective Hamiltonian is then characterized by a single spin-orbit parameter,  $\xi_{\text{eff}}$ , and reads

$$\hat{H}_{\text{SOC}} = \frac{\xi_{\text{eff}}}{2} \sum_{\ell, \sigma=\pm} \ell \sigma \hat{d}_{\ell\sigma}^{\dagger} \hat{d}_{\ell\sigma}. \quad (37)$$

Here  $\ell$  indicates the quasi angular momentum of the vacancy state, while the second index corresponds to the spin  $\sigma = \pm$ .

The value for the effective spin-orbit parameter can be estimated from the band structure looking at the spin-orbit splitting at the  $\bar{\Gamma}$  point. From Fig. 3 we find  $\xi_{\text{eff}}^{\text{Mo}} = 0.05$  eV and  $\xi_{\text{eff}}^{\text{W}} = 0.168$  eV. Evidently, due to the heavier nature of the W atoms compared to Mo, the effective spin-orbit parameter in WS<sub>2</sub>-Gr is  $\sim 3.5$  times larger compared to the MoS<sub>2</sub>-Gr heterostructure. At the  $\bar{\Gamma}$  point, where the impact of the impurity band dispersion is strongly reduced and the system displays the behavior of the isolated defect, we find that the distribution of the vacancy levels with spin-orbit interaction reads  $E_{\ell,+} = E_{\bar{\ell},-}$  and  $E_{\ell,-} = E_{\bar{\ell},+}$ . The Kramers' degeneracy is preserved and the spin-orbit coupling has opposite effect for each of the  $\ell, \bar{\ell}$  projections:  $E_{\ell\sigma} = E_{\text{av}} + \ell\sigma\xi_{\text{eff}}/2$  with  $E_{\text{av}}$  the reference mid-gap energy. For simplicity, we will use the short-hand notation  $g := \{\ell-, \bar{\ell}+\}$  and  $e := \{\ell+, \bar{\ell}-\}$ , with  $E_e > E_g$  in what follows. The ground and excited energy levels measured from the graphene charge neutrality point read  $E_g^{\text{Mo}} = 27$  meV,  $E_e^{\text{Mo}} = 77$  meV,  $E_g^{\text{W}} = 230$  meV, and  $E_e^{\text{W}} = 398$  meV.

The SOC has no influence on the tunnelling Hamiltonian. Thus, starting from the model introduced in Sec. III and parametrized in Sec. V A, we construct the single-particle rate matrix from the functional form employed there and rotate it into the eigenstate basis with SOC. The single-particle tunneling rate is still diagonal in spin, but it mixes states with different angular momentum.

A further simplification applies, though, if we additionally consider the secular or rotating wave approximation [69], which applies since  $\hbar\Gamma \ll \Delta := E_e - E_g$ . This approximation amounts to neglect the rapidly oscillating coherences, between states with different energies, as they average to zero on the typical time scales of the population dynamics. Eventually, the single-particle tunneling rate matrix can be taken completely diagonal. In the basis of the vacancy eigenstates, it reads  $\Gamma^{\text{SOC}}(E) = \Gamma^{\text{SOC}}(E) \otimes \mathbf{1}_4$  where

$$\Gamma^{\text{SOC}}(E) = \frac{\Gamma_u(E) + \Gamma_{\bar{u}}(E)}{2}, \quad (38)$$

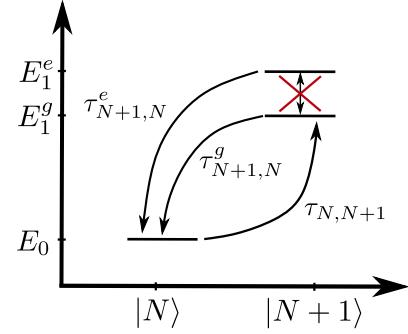


FIG. 7. Representation of the many-body Hilbert space and electron transition times,  $\tau_{\alpha,\beta}$ , between many-body states (labeled by particle number  $N, N+1$ ) in the presence of spin-orbit coupling. The energies are given in the grand canonical ensemble,  $E_0 = \varepsilon_0 - N\mu$  and  $E_1^{e,g} = \varepsilon_1^{e,g} - (N+1)\mu$ , with  $\mu \equiv E_F$  at zero gate voltage and  $\varepsilon_{0,1}$  are the total energies of the system with  $N$  and  $N+1$  electrons. Transitions that do not change the particle number are forbidden.

is the effective, still energy dependent, tunnelling rate. A direct extraction of the tunnelling amplitudes from the DFT band structures is, in this case, not possible, due to the interplay between the tunnelling coupling and the spin-orbit interaction, in determining the energy eigenstates.

**Transition rates.** Employing Eqs. (25) (and an analogous equation for the discharging rate) as well as  $\Gamma^{\text{SOC}}$ , we can compute the rates for the many-body transitions shown in Fig. 7. Charging the vacancy occurs by tunneling of the electron from the graphene layer to one of the two quasi-degenerated levels of the vacancy, with transition time  $\tau_{N,N+1}$ . The transition rate reads

$$W = \sum_{\ell\sigma} \Gamma^{\text{SOC}}(\Delta E_{0,\ell\sigma}) f^+(\Delta E_{0,\ell\sigma}), \quad (39)$$

which can be understood as  $W = W_{0g} + W_{0e}$ , as the final single-particle state in the many-body transition  $N \rightarrow N+1$  is unknown. Each transition rate composing the total charging rate is given by

$$W_{0g} = 2 \Gamma^{\text{SOC}}(\Delta E_{0,g}) f^+(\Delta E_{0,g}), \quad (40a)$$

$$W_{0e} = 2 \Gamma^{\text{SOC}}(\Delta E_{0,e}) f^+(\Delta E_{0,e}). \quad (40b)$$

Discharging occurs by tunneling of an electron originally in one of the two spin-orbit split Kramers' pairs towards the graphene bath. The corresponding transition rates are

$$W_{g0} = 2 \Gamma^{\text{SOC}}(\Delta E_{0,g}) f^-(\Delta E_{0,g}), \quad (41a)$$

$$W_{e0} = 2 \Gamma^{\text{SOC}}(\Delta E_{0,e}) f^-(\Delta E_{0,e}), \quad (41b)$$

with transition times  $\tau_{N+1,N}^g := 1/W_{g0}$ ,  $\tau_{N+1,N}^e := 1/W_{e0}$ .

We represent in Fig. 8 (a) these electron transition times as a function of the gate voltage,  $V_{\text{gate}}$ . Compared

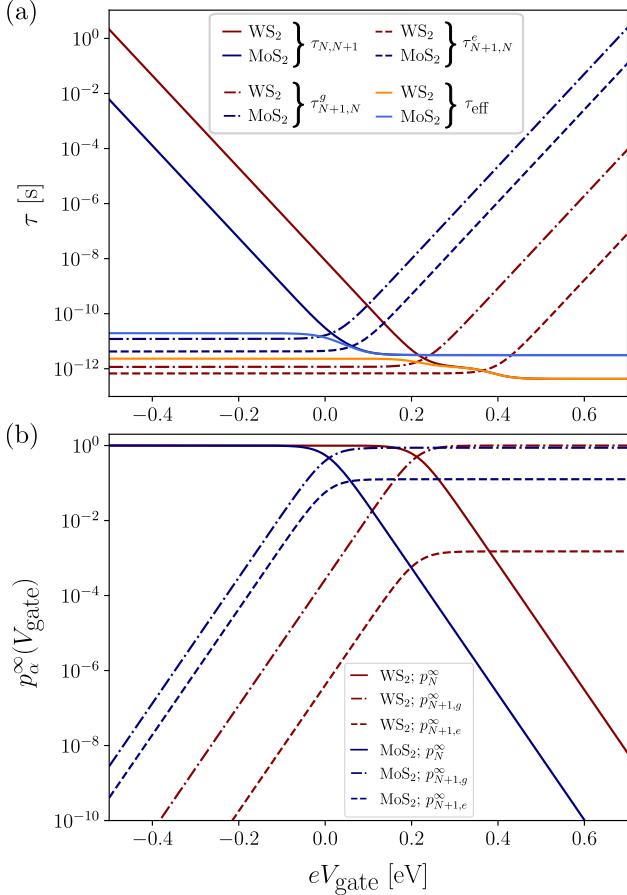


FIG. 8. (a) Electronic transition times,  $\tau$ , for MoS<sub>2</sub> and WS<sub>2</sub> vacancy states as a function of the gate voltage,  $V_{\text{gate}}$  measured in eV. The solid, dashed and dashed-dotted lines correspond to the transitions  $\alpha \rightarrow \beta$  shown in Fig. 7. The dotted line corresponds to the effective transition time associated to the semiclassical dynamics defined from Eq. (46) and controlling the dynamics from a local to global equilibrium state. For simplicity, we set the chemical potential at zero gate voltage to be zero. (b) Corresponding stationary semiclassical probability of occupation of the many-body levels,  $\mathbf{p}^\infty$ , as a function of the gate voltage.

to the case where the spin-orbit coupling is negligible, we find a similar qualitative behavior of the transition times as a function of  $V_{\text{gate}}$ . The charging rate is an effective rate were both transitions to the  $g$  and  $e$  states are allowed, thus the corresponding transition time shows a small jump when the chemical potential coincides with the higher energy level. This jump is better observed for WS<sub>2</sub> because the spin-orbit splitting is substantially larger than the characteristic thermal energy. In addition, at  $V_{\text{gate}} = 0$ , where we measure the impact of spin-orbit interaction only, we find that comparing with the case where spin-orbit coupling is absent the charging process occurs faster for the WS<sub>2</sub> by almost one order of magnitude while being almost unchanged for MoS<sub>2</sub>, thus pointing to the spin-orbit interaction as a source of change of in the many-body transition rates.

**Semiclassical dynamics.** Using the transition rates, we derive a set of semiclassical dynamical equations for the populations using Eq. (28). Differently from the case without spin-orbit interaction, the  $\bar{\Gamma}$  point degeneracy splitting of the defect states, hinders the Lamb shift induced dynamics. Technically, this happens since the Lamb shift Hamiltonian commutes with the reduced density matrix,  $[\hat{H}_{\text{LS}}, \hat{\rho}_{\text{red}}] = 0$ .

For the calculation of the semiclassical dynamics, we assume, similarly to Sec. V A, local thermal equilibrium and, therefore, an equal population of any degenerate level:  $p_g = p_{\bar{g}}$  and  $p_e = p_{\bar{e}}$ . The resulting dynamical equations can be written in matrix form as

$$\dot{\mathbf{p}} = \begin{pmatrix} -W & \tilde{W}_{g0} & \tilde{W}_{e0} \\ W_{0g} & -\tilde{W}_{g0} & 0 \\ W_{0e} & 0 & -\tilde{W}_{e0} \end{pmatrix} \mathbf{p}, \quad (42)$$

where  $\mathbf{p} = (p_N \ p_{N+1,g} \ p_{N+1,e})^T$  and we remind that  $\tilde{W}_\alpha := W_\alpha/2$ . The local thermal equilibrium further imposes a relation between the populations of the single-particle states  $|g\rangle$  and  $|e\rangle$ , namely,  $p_e = p_g \exp(-\beta\Delta)$ , where  $\Delta = E_e - E_g$ . Consequently, we can reduce Eq. (42) to a two-dimensional system of equations

$$\begin{pmatrix} \dot{p}_N \\ \dot{p}_{N+1} \end{pmatrix} = \begin{bmatrix} -W & f_0^-(\Delta)\tilde{W}_{g0} + f_0^+(\Delta)\tilde{W}_{e0} \\ W & -f_0^-(\Delta)\tilde{W}_{g0} - f_0^+(\Delta)\tilde{W}_{e0} \end{bmatrix} \begin{pmatrix} p_N \\ p_{N+1} \end{pmatrix}, \quad (43)$$

where

$$f_0^\pm(\Delta) := \frac{1}{1 + \exp(\pm\beta\Delta)}. \quad (44)$$

These equations are solved under the constraint  $\sum_\alpha p_\alpha = p_N + p_{N+1} = 1$  with  $p_{N+1} = p_g + p_e$ . The analytical solution of the system of equations yields

$$\mathbf{p}(t) = \mathbf{p}^\infty + p_0 \begin{pmatrix} 1 \\ -f_0^-(\Delta) \\ -f_0^+(\Delta) \end{pmatrix} e^{-W_{\text{eff}}t}, \quad (45)$$

where we employed the inverse relations between  $p_{N+1}$  and the populations of the  $g$  and  $e$  levels,  $p_g = f_0^-(\Delta)p_{N+1}$  and  $p_e = f_0^+(\Delta)p_{N+1}$ , as well as defined the effective decay rate

$$\begin{aligned} W_{\text{eff}} &= W + f_0^-(\Delta)\tilde{W}_{g0} + f_0^+(\Delta)\tilde{W}_{e0}, \\ &= \Gamma^{\text{SOC}}(\Delta E_{0,g})[2f^+(\Delta E_{0,g}) + f_0^-(\Delta)f^-(\Delta E_{0,g})] \\ &\quad + \Gamma^{\text{SOC}}(\Delta E_{0,e})[2f^+(\Delta E_{0,e}) + f_0^+(\Delta)f^-(\Delta E_{0,e})] \end{aligned} \quad (46)$$

These expressions reduce to Eqs. (29)-(30) for the limit  $\Delta \rightarrow 0$  in which all four states of the one particle sector are degenerated.

Eq. (46) defines the effective relaxation time  $\tau_{\text{eff}} := 1/W_{\text{eff}}$  for the population dynamics in the presence of spin-orbit coupling. We show  $\tau_{\text{eff}}$  for our heterobilayers in comparison to  $\tau_{N,N+1}$  and  $\tau_{N+1,N}$  in Fig. 8 (a). As

in the case of Fig. 5 (a), this effective time is less sensitive to the gate voltage compared to  $\tau_{N,N+1}$  and  $\tau_{N+1,N}$ . More importantly, when looking at  $\tau_{\text{eff}}$  for WS<sub>2</sub>–Gr at  $V_{\text{gate}} = 0$ , we find  $\tau_{\text{eff}}^{\text{W}} \simeq 2.3$  ps at  $\mu = 0$  (for comparison,  $\tau_{\text{eff}}^{\text{Mo}} \simeq 13.5$  ps for the MoS<sub>2</sub>–Gr interface). Shift of the vacancy levels following the G<sub>0</sub>W<sub>0</sub> for the pristine system, can reduce the relaxation time by at most  $\sim 25\%$ . In any case, this value is of the order of magnitude of the typical electronic relaxation times measured by tr-ARPES for the same type of heterostructure and same process [7, 8]. Note that within our model, we can account for the effect on the relaxation times of changes of the chemical potential due to doping at the graphene layer; our calculated relaxation times can change up to one order of magnitude within the range from  $-300$  to  $400$  meV. Our results therefore suggest that, if the vacancy concentration is low, defect charge transfer can be a potential mechanism that participates in the ultrafast relaxation times observed for photoexcited electrons in these heterobilayers.

The vector  $\mathbf{p}^\infty$  yields the stationary population at a given chemical potential and temperature of the bath for the states  $|N\rangle$ ,  $|g\rangle$  and  $|e\rangle$ ,

$$\mathbf{p}^\infty = \frac{W}{W_{\text{eff}}} \begin{pmatrix} W_{\text{eff}}/W - 1 \\ f_0^-(\Delta) \\ f_0^+(\Delta) \end{pmatrix}. \quad (47)$$

In Fig. 8 (b), we display the stationary population (47) for the two studied interfaces and evaluated for the same range of gate voltages and physical parameters as in Fig. 8 (a). The values of the stationary population are controlled by the many-body transition rates, as well as temperature. In the large negative  $V_{\text{gate}}$  limit, the chemical potential is far below the available defect levels and therefore, the vacancy remains empty,  $p_N \rightarrow 1$  and  $p_g, p_e \rightarrow 0$ . After increasing  $V_{\text{gate}}$  and once the chemical potential is larger than the single-particle energy of the empty defect states, tunneling becomes possible and the vacancy charges at equilibrium, with the relation between the  $g$  and  $e$  levels dictated by the Boltzmann factor,  $\exp(-\beta\Delta)$ . Since the splitting of the defect levels due to spin-orbit interaction is substantially smaller in MoS<sub>2</sub> compared to WS<sub>2</sub>, the population of  $|e\rangle_{\text{Mo}}$  remains several orders of magnitude larger in the large positive  $V_{\text{gate}}$  limit compared to  $|e\rangle_{\text{W}}$ . Finally, as in Sec. V A, the value for  $p_0$  depends on the initial conditions for the time evolution. It can be easily calculated, for example,  $p_0 = 1 - W/W_{\text{eff}}$  if the initial condition is fixed to  $p_N = 0$  (vacancy is occupied and therefore the system decharges over time). As an example, we show in Fig. 9 charge quenching as given by Eq. (47) for fixed gate energy  $eV_{\text{gate}} = -0.3$  eV. The dynamics shows the expected exponential decay law, controlled by  $W_{\text{eff}}$  as well as the asymptotic behavior seen in Fig. 8 (b). Also, as anticipated, if the temperature energy scale  $\beta$  is smaller than the splitting  $\Delta$ , thermal population of the  $|e\rangle$  state does not occur, as for WS<sub>2</sub> compared to MoS<sub>2</sub>.

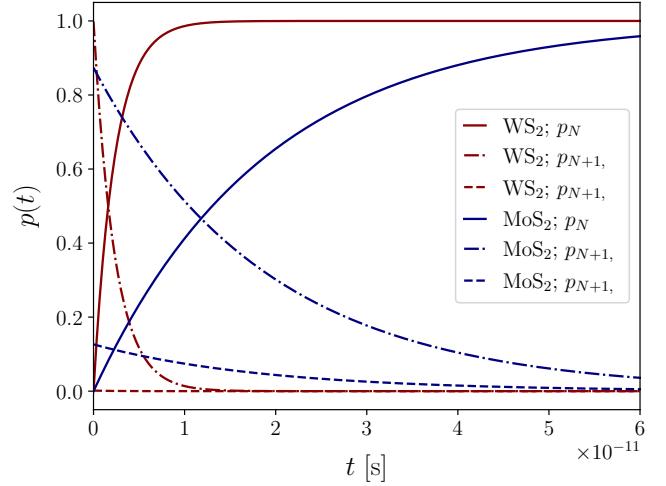


FIG. 9. Population,  $p(t)$ , of the energy levels as a function of time,  $t$  (measured in s). The dynamics is computed for the representative gate voltage,  $V_{\text{gate}} = -0.3$  eV.

## VI. CONCLUSION

In this work, we have investigated, by a combination of density functional theory, model Hamiltonian and master equation techniques, how defect states created by monoatomic chalcogen vacancies in representative MoS<sub>2</sub>–Gr and WS<sub>2</sub>–Gr interfaces are an interesting playground for electronic dynamical processes. From the *ab initio* calculations, we find that the empty vacancy levels are naturally hybridized with the Dirac cone from the graphene layer. We propose a low-energy model and derive the single-particle tunneling rate matrices employing information from the DFT band structure calculation as well as group theory arguments. This allows us to gain physical insight into the impact of symmetries into the parameter-free many-body transition rates. At the level of sequential tunneling, we perturbatively computed the transition rates for the population and depopulation of the vacancy by one electron tunneling to or from the graphene reservoir. We find that transition rates (and corresponding electronic transition times) are sensitive to external electric fields perpendicular to the heterobilayer plane, and created by a voltage gate on graphene, changing also by several orders of magnitude. A shift of the chemical potential via this gate voltage can also change the relative order of the rates for the different TMDC (MoS<sub>2</sub>, WS<sub>2</sub>) considered in this work. This is shown to influence the semiclassical population dynamics, derived here within an effective two-level model assuming local thermal equilibrium. For strong spin-orbit interaction, we proposed a modified transport model for the charge transfer using the DFT band structure information as well. We obtain a single-particle rate matrix which is still proportional to the identity, but with non-trivial energy dependence. We computed the corresponding transition rates at the level of sequential tunneling and employing

the secular approximation. A three-level model for the dynamics of the population and depopulation of the vacancy has been proposed and solved assuming local thermal equilibrium. The effect of spin-orbit interaction has been gauged in our charge transfer rates.

Our results for the effective rate suggests that defects slow down the charge transfer, effectively trapping electrons, and can be a relevant transport channel in the ultrafast charge transfer processes observed in recent experiments [7, 8] assuming the vacancy concentration considered in this work. Our research serves as a basis to study other more subtle dynamical features associated to quantum internal dynamics occurring at short time scales due to quantum degeneracy of the vacancy energy levels or electron-phonon coupling [70], and we demonstrated how simple dynamical models can be used as guidance in the understanding of experimental transport phenomena

at these heterostructures and complex interfaces.

## ACKNOWLEDGMENTS

Discussions with F. Evers, I. Gierz, M. Glazov, E. Malic, R. Perea-Causin, D. Qiu, A. Rappe and H. Steinberg are gratefully acknowledged. We are thankful to S. Latini, L. Xian and A. Rubio for providing us with the non-defected geometry of the heterostructure employed in some of the calculations in this article. The computations were performed using the resources of the Chemfarm local cluster at the Weizmann Institute of Science and the MPCDF computing center. D. H.-P. and S. R. A. acknowledge funding from a Minerva Foundation grant. D. H.-P., A. D. and S. R. A. acknowledge funding from the Collaborative Research Center SFB 1277 (Project-ID 314695032, subprojects B02 and B10).

- 
- [1] A. K. Geim and I. V. Grigorieva, *Nature* **499**, 419 (2013).  
[2] Y. Liu, N. O. Weiss, X. Duan, H.-C. Cheng, Y. Huang, and X. Duan, *Nature Reviews Materials* **1**, 16042 (2016).  
[3] C. Jin, E. Y. Ma, O. Karni, E. C. Regan, F. Wang, and T. F. Heinz, *Nature Nanotechnology* **13**, 994 (2018).  
[4] N. Papadopoulos, P. Gehring, K. Watanabe, T. Taniguchi, H. S. J. van der Zant, and G. A. Steele, *Phys. Rev. B* **101**, 165303 (2020).  
[5] T. R. Devidas, I. Keren, and H. Steinberg, *Nano Letters* **21**, 6931 (2021).  
[6] L. Yuan, T.-F. Chung, A. Kuc, Y. Wan, Y. Xu, Y. P. Chen, T. Heine, and L. Huang, *Science Advances* **4** (2018), 10.1126/sciadv.1700324.  
[7] S. Aeschlimann, A. Rossi, M. Chávez-Cervantes, R. Krause, B. Arnoldi, B. Stadtmüller, M. Aeschlimann, S. Forti, F. Fabbri, C. Coletti, and I. Gierz, *Science Advances* **6** (2020), 10.1126/sciadv.aay0761.  
[8] R. Krause, S. Aeschlimann, M. Chávez-Cervantes, R. Perea-Causin, S. Brem, E. Malic, S. Forti, F. Fabbri, C. Coletti, and I. Gierz, *Phys. Rev. Lett.* **127**, 276401 (2021).  
[9] S. Fu, I. du Fossé, X. Jia, J. Xu, X. Yu, H. Zhang, W. Zheng, S. Krasel, Z. Chen, Z. M. Wang, K.-J. Tielrooij, M. Bonn, A. J. Houtepen, and H. I. Wang, *Science Advances* **7**, eabd9061 (2021).  
[10] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, *Nature* **438**, 197 (2005).  
[11] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, *Rev. Mod. Phys.* **81**, 109 (2009).  
[12] D. W. Lutzke, W. Zhang, A. Suslu, T.-R. Chang, H. Lin, H.-T. Jeng, S. Tongay, J. Wu, A. Bansil, and A. Lanzara, *Phys. Rev. B* **91**, 235202 (2015).  
[13] S. Refaelly-Abramson, D. Y. Qiu, S. G. Louie, and J. B. Neaton, *Phys. Rev. Lett.* **121**, 167402 (2018).  
[14] G. Wang, A. Chernikov, M. M. Glazov, T. F. Heinz, X. Marie, T. Amand, and B. Urbaszek, *Rev. Mod. Phys.* **90**, 021001 (2018).  
[15] M. Goryca, J. Li, A. V. Stier, T. Taniguchi, K. Watanabe, E. Courtade, S. Shree, C. Robert, B. Urbaszek, X. Marie, and S. A. Crooker, *Nature Communications* **10**, 4172 (2019).  
[16] B. Schuler, D. Y. Qiu, S. Refaelly-Abramson, C. Kastl, C. T. Chen, S. Barja, R. J. Koch, D. F. Ogletree, S. Aloni, A. M. Schwartzberg, J. B. Neaton, S. G. Louie, and A. Weber-Bargioni, *Phys. Rev. Lett.* **123**, 076801 (2019).  
[17] E. Mitterreiter, B. Schuler, A. Micevic, D. Hernangómez-Pérez, K. Barthelmi, K. A. Cochrane, J. Kiemle, F. Sigler, J. Klein, E. Wong, E. S. Barnard, K. Watanabe, T. Taniguchi, M. Lorke, F. Jahnke, J. J. Finley, A. M. Schwartzberg, D. Y. Qiu, S. Refaelly-Abramson, A. W. Holleitner, A. Weber-Bargioni, and C. Kastl, *Nature Communications* **12**, 3822 (2021).  
[18] 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).  
[19] M. M. Glazov and E. L. Ivchenko, *JETP Letters* **113**, 7 (2021).  
[20] F. Bassani, G. Iadonisi, and B. Preziosi, *Reports on Progress in Physics* **37**, 1099 (1974).  
[21] M. Aghajanian, B. Schuler, K. A. Cochrane, J.-H. Lee, C. Kastl, J. B. Neaton, A. Weber-Bargioni, A. A. Mostofi, and J. Lischner, *Phys. Rev. B* **101**, 081201 (2020).  
[22] S. Fu, I. du Fossé, X. Jia, J. Xu, X. Yu, H. Zhang, W. Zheng, S. Krasel, Z. Chen, Z. M. Wang, K.-J. Tielrooij, M. Bonn, A. J. Houtepen, and H. I. Wang, *Science Advances* **7**, eabd9061 (2021).  
[23] R. Steinitz-Eliyahu, D. Hernangómez-Pérez, F. S. Hegner, P. Nikačević, N. López, and S. Refaelly-Abramson, *Phys. Rev. Materials* **6**, 065402 (2022).  
[24] S. M. Reimann and M. Manninen, *Rev. Mod. Phys.* **74**, 1283 (2002).  
[25] C. W. Groth, B. Michaelis, and C. W. J. Beenakker, *Phys. Rev. B* **74**, 125315 (2006).  
[26] M. Kotzian, F. Gallego-Marcos, G. Platero, and R. J. Haug, *Phys. Rev. B* **94**, 035442 (2016).  
[27] M. Busl, R. Sánchez, and G. Platero, *Phys. Rev. B* **81**, 121306 (2010).

- [28] M. Seo, H. K. Choi, S.-Y. Lee, N. Kim, Y. Chung, H.-S. Sim, V. Umansky, and D. Mahalu, Phys. Rev. Lett. **110**, 046803 (2013).
- [29] M. Korkusinski, I. P. Gimenez, P. Hawrylak, L. Gaudreau, S. A. Studenikin, and A. S. Sachrajda, Phys. Rev. B **75**, 115301 (2007).
- [30] L. Gaudreau, G. Granger, A. Kam, G. C. Aers, S. A. Studenikin, P. Zawadzki, M. Pioro-Ladrière, Z. R. Wasilewski, and A. S. Sachrajda, Nature Physics **8**, 54 (2012).
- [31] M. Stopa, Phys. Rev. Lett. **88**, 146802 (2002).
- [32] A. Donarini, G. Begemann, and M. Grifoni, Nano Letters **9**, 2897 (2009).
- [33] A. Donarini, G. Begemann, and M. Grifoni, Phys. Rev. B **82**, 125451 (2010).
- [34] A. Donarini, M. Niklas, M. Schafberger, N. Paradiso, C. Strunk, and M. Grifoni, Nature Communications **10**, 381 (2019).
- [35] L. Gaudreau, S. A. Studenikin, A. S. Sachrajda, P. Zawadzki, A. Kam, J. Lapointe, M. Korkusinski, and P. Hawrylak, Phys. Rev. Lett. **97**, 036807 (2006).
- [36] H. Pan, M. G. House, X. Hao, and H. W. Jiang, Applied Physics Letters **100**, 263109 (2012).
- [37] M. Niklas, A. Trottmann, A. Donarini, and M. Grifoni, Phys. Rev. B **95**, 115133 (2017).
- [38] W. J. Schutte, J. L. De Boer, and F. Jellinek, Journal of Solid State Chemistry **70**, 207 (1987).
- [39] T. Naimer, K. Zollner, M. Gmitra, and J. Fabian, Phys. Rev. B **104**, 195156 (2021).
- [40] J.-Y. Noh, H. Kim, and Y.-S. Kim, Phys. Rev. B **89**, 205417 (2014).
- [41] D. Liu, Y. Guo, L. Fang, and J. Robertson, Applied Physics Letters **103**, 183113 (2013).
- [42] W. Zhou, X. Zou, S. Najmaei, Z. Liu, Y. Shi, J. Kong, J. Lou, P. M. Ajayan, B. I. Yakobson, and J.-C. Idrobo, Nano Letters **13**, 2615 (2013).
- [43] N. Wakabayashi, H. G. Smith, and R. M. Nicklow, Phys. Rev. B **12**, 659 (1975).
- [44] M. H. Naik and M. Jain, Phys. Rev. Materials **2**, 084002 (2018).
- [45] Z. Y. Zhu, Y. C. Cheng, and U. Schwingenschlögl, Phys. Rev. B **84**, 153402 (2011).
- [46] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
- [47] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Phys. Rev. Lett. **111**, 216805 (2013).
- [48] C. Jin, F. A. Rasmussen, and K. S. Thygesen, The Journal of Physical Chemistry C **119**, 19928 (2015).
- [49] As discussed above and in previous work, we neglect, for the modelization, the spurious defect-defect interaction between vacancies belonging to different cells and thus we assume non-dispersive bands  $\epsilon_{i\mathbf{q}\sigma} \simeq \epsilon_{i\sigma}$  as the dispersion comes from spurious defect-defect interaction between vacancies that belong to different supercells.
- [50] M. Braun, J. König, and J. Martinek, Phys. Rev. B **70**, 195345 (2004).
- [51] I. Weymann, J. König, J. Martinek, J. Barnaś, and G. Schön, Phys. Rev. B **72**, 115334 (2005).
- [52] G. Begemann, D. Darau, A. Donarini, and M. Grifoni, Phys. Rev. B **77**, 201406 (2008).
- [53] O. Karlström, J. N. Pedersen, P. Samuelsson, and A. Wacker, Phys. Rev. B **83**, 205412 (2011).
- [54] C. Rohrmeier and A. Donarini, Phys. Rev. B **103**, 205420 (2021).
- [55] M. Frankerl and A. Donarini, Phys. Rev. B **103**, 085420 (2021).
- [56] J. J. Sakurai and J. Napolitano, *Modern quantum mechanics*, 3rd ed. (Cambridge University Press, 2021).
- [57] S. Nakajima, Progress of Theoretical Physics **20**, 948 (1958).
- [58] R. Zwanzig, The Journal of Chemical Physics **33**, 1338 (1960).
- [59] H. P. Breuer and F. Petruccione, *The theory of open quantum systems*, 1st ed. (Oxford University Press, 2002).
- [60] All the odd terms vanish due to particle-number conservation, *i.e.*  $\mathcal{P}\mathcal{L}_{\text{tun}}^n\mathcal{P} = 0$  if  $n$  is odd.
- [61] Y. Wang, Z. Wang, W. Yao, G.-B. Liu, and H. Yu, Phys. Rev. B **95**, 115429 (2017).
- [62] S. Brem, K.-Q. Lin, R. Gillen, J. M. Bauer, J. Maultzsch, J. M. Lupton, and E. Malic, Nanoscale **12**, 11088 (2020).
- [63] M. Niklas, *Current and noise properties of interacting nanojunctions*, Ph.D. thesis, Universität Regensburg (2018).
- [64] The finite band width of the graphene DoS ensures the convergence of the principal value integral.
- [65] A. Kleiner, D. Hernández-Pérez, and S. Refaelly-Abramson, “Structural and dielectric effects of single-layer graphene on the excitonic nature of single-layer WS<sub>2</sub>,” in preparation (2022).
- [66] K. Zollner, P. E. Faria Junior, and J. Fabian, Phys. Rev. B **100**, 085128 (2019).
- [67] J. Amann, T. Völk, T. Rockinger, D. Kochan, K. Watanabe, T. Taniguchi, J. Fabian, D. Weiss, and J. Eroms, Phys. Rev. B **105**, 115425 (2022).
- [68] H. Bruus and K. Flensberg, *Many-body quantum theory in condensed matter physics - an introduction* (Oxford University Press, United States, 2004).
- [69] D. Darau, G. Begemann, A. Donarini, and M. Grifoni, Phys. Rev. B **79**, 235404 (2009).
- [70] Q. Zheng, W. A. Saidi, Y. Xie, Z. Lan, O. V. Prezhdo, H. Petek, and J. Zhao, Nano Letters **17**, 6435 (2017).

## Supplemental Material

### Charge quenching at defect states in transition metal dichalcogenide-graphene van der Waals heterobilayers

Daniel Hernangómez-Pérez,<sup>1</sup> Andrea Donarini,<sup>2</sup> and Sivan Refaelly-Abramson<sup>1</sup>

<sup>1</sup>*Department of Molecular Chemistry and Materials Science,  
Weizmann Institute of Science, Rehovot 7610001, Israel*

<sup>2</sup>*Institute of Theoretical Physics, University of Regensburg, 93040 Regensburg, Germany*

Email: daniel.hernangomez@weizmann.ac.il

Keywords: 2D materials, transition-metal dichalcogenides, van der Waals, defects, graphene, transport, single-electron

## CONTENTS

I. Computational details	2
II. Elements of applied group theory	2
III. Calculation details for $\Gamma_{ij}^p(E, \sigma)$	3
IV. Temperature and defect energy dependence of the many-body transition rates	7
References	9

## I. COMPUTATIONAL DETAILS

Our *ab initio* calculations were performed using density functional theory as implemented in the QUANTUM ESPRESSO package<sup>1</sup>. For the exchange-correlation functional, we used the non-empirical generalized gradient approximation of Perdew-Burke-Ernzerhof<sup>2</sup>. We employed a plane-wave basis set and scalar relativistic (full relativistic for the spin-orbit coupling calculations) norm-conserving pseudopotentials<sup>3</sup> with a basis cut-off of 95 Ry for WS<sub>2</sub>-Gr and 90 Ry for MoS<sub>2</sub>-Gr. The self-consistent charge density was calculated within a  $6 \times 6 \times 1$  grid and the total energy considered to be converged when the difference between iterations was smaller than a threshold value of  $10^{-9}$  Ry. Preoptimizations of the supercell were performed with VASP<sup>4</sup>, using LDA as exchange correlation functional, a basis energy cut-off of 600 eV and a  $6 \times 6 \times 1$   $\mathbf{k}$ -grid. Additional relaxations in the presence of the vacancy were performed by fixing the supercell lattice vectors and optimizing the position of the atoms during the self-consistent field cycle until all the components of the ionic forces were smaller than  $10^{-3}$  Ry/ $a_0$ . For the optimization procedure we also employed the van-der Waals corrected functional vdw-df-09<sup>5,6</sup> to account for potential additional changes in the interlayer separation.

## II. ELEMENTS OF APPLIED GROUP THEORY

For the sake of completeness, we briefly summarize in section of the Supporting Information (SI) elementary point-group theory results that underlay the density functional theory findings and the calculation of the single-particle tunneling rate matrix in the main paper and Sec. III of the SI. We follow here Refs. 7–9. In the absence of the graphene layer, the XS<sub>2</sub> TMDC layer with a chalcogen vacancy possesses C<sub>3v</sub> symmetry, the point-group of the equilateral triangle. This point group is characterized by a proper rotation axis of order 3 and 3 planes of symmetry. Traditionally, E, C<sub>3</sub><sup>±</sup> and σ<sub>v</sub> ( $v = 1, 2, 3$ ) are chosen as notation for the identity, rotation and reflection elements of the group, defining three equivalence classes (see scheme in Fig. S1).

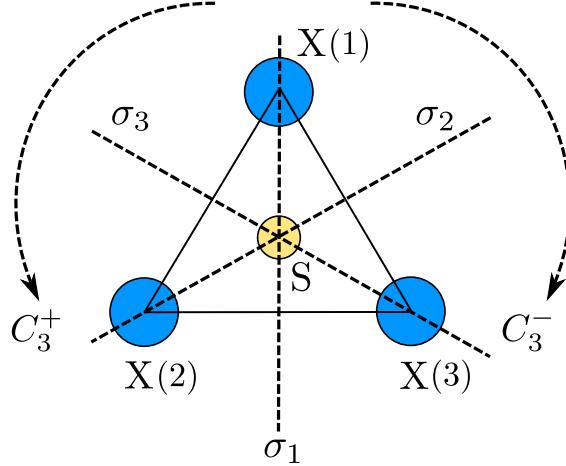
As a consequence of Schur's lemmas<sup>8</sup>, the eigenvalues and eigenstates of a Hamiltonian which commutes with all symmetry operators  $\hat{G}$  of a group can be classified according to the irreducible representations of the group itself. This is the case, if and only if, the group

is a complete group of invariance. In other words, we can use the unique reduction of the group into its irreducible representations to find a block-diagonal form of the Hamiltonian, with its invariant subspaces having the same dimension as the irreducible representations of the group. Moreover, the number of irreducible representations is equal to the number of equivalence classes in the symmetry group. In the particular case of  $C_{3v}$ , the group has three non-trivial irreducible representations of dimension 1 and 2 [typically noted as  $A_1, A_2, E$ , with  $\dim(A_i) = 1$  and  $\dim(E) = 2$ ]. Consequently, if the defect preserves the original symmetry of the lattice, the defect states are at most double degenerated. In the case of the  $XS_2$ -defected monolayers considered in this work, three defect states appear close to the TMDC: a pair of empty in-gap degenerated  $E$  states and an occupied non-degenerated  $A_1$  state close to the topmost valence band. Note that the energetics of the defect states cannot be predicted by group theory only, therefore, pointing towards the need of employing *ab initio* methods in the study of these systems.

Van der Waals adsorption of graphene to the TMDC layer lowers the global symmetry of the supercell from  $C_{3v}$  to  $C_s$ . The  $C_s$  symmetry group only has one element apart from the identity  $E$ , given by the plane of symmetry and noted  $\sigma_s$ . As a difference to  $C_{3v}$ ,  $C_s$  is Abelian (all the elements of the group commute), and all its irreducible representations are forcibly one-dimensional. Consequently, the eigenenergies and eigenstates of the system are not guaranteed to be degenerated (except for accidental degeneracies). The heterobilayer eigenstates stemming from the former vacancy  $E$  states, can now be classified according to one the reflection operators,  $\sigma_s$ . Their different nature is associated to the different size of the anti-crossings between the impurity band and the Dirac cone of graphene. We will use this observation in the calculation of the single-particle rate matrix in Sec. III of the SI.

### III. CALCULATION DETAILS FOR $\Gamma_{ij}^p(E, \sigma)$

In this section, we derive an expression for the single-particle rate matrix  $\Gamma_{ij}^p(E, \sigma)$  based on the group theory analysis from Sec. II of the SI. We start from Eq. (19), which contains as crucial component the tunneling (“hybridization”) matrix elements,  $t_{i\mathbf{k}\sigma}$ . These matrix elements are assumed to be spin-diagonal,  $t_{i\sigma, \mathbf{k}\sigma'} = t_{i\mathbf{k}\sigma} \delta_{\sigma\sigma'}$ , as the tunneling occurs over a thin vacuum layer which cannot change the spin orientation. By definition,  $t_{i\mathbf{k}\sigma}$  are given



**Fig. S1.** Schematic representation of the non-trivial operations of the group  $C_{3v}$  (rotation by  $2\pi/3$ , reflection with respect to a plane of symmetry) for the chalcogen vacancy in the  $XS_2$  monolayer.

by

$$t_{i\mathbf{k}\sigma} = \langle i\sigma | \hat{H} | \mathbf{k}\sigma \rangle := \langle i | \hat{H} | \mathbf{k} \rangle, \quad (\text{S1})$$

where  $\hat{H}$  is the Hamiltonian of the system,  $|i\rangle \equiv |i, \mathbf{0}\rangle$  is a localized impurity state, localized, for convenience, at the origin of the Bravais supercell lattice,  $|\mathbf{k}\rangle$  a graphene eigenstate and the spin degree of freedom is kept implicit as it does not intervene in the calculation.

We can express the defect band in the reciprocal space as

$$|i, \mathbf{k}\rangle = \frac{1}{\sqrt{N_{sc}}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} |i, \mathbf{R}\rangle, \quad (\text{S2})$$

where  $\{\mathbf{R}\}$  spans the supercell lattice and  $N_{sc}$  is the number of supercell lattice points. It is convenient to have the defect at the origin of coordinates (see Fig. 1 in the main text); then, the inverse transformation reads

$$|i, \mathbf{0}\rangle = \frac{1}{\sqrt{N_{sc}}} \sum_{\mathbf{k}} |i, \mathbf{k}\rangle, \quad (\text{S3})$$

where  $\{\mathbf{k}\}$  spans the quasi-momenta lattice of the supercell Brillouin zone.

Using Eq. (19), we estimate the value of the single-particle rate matrix by looking at the region of the spectrum showing level repulsion between the graphene and the vacancy states. Note that we consider only the orbital component of the system without spin-orbit interaction, spin-orbit coupling will be incorporated perturbatively into the model at a later

stage in the main paper. Substituting the matrix elements we find, for each spin sector,

$$\Gamma_{ij}^+(E) = \frac{2\pi}{\hbar} \frac{1}{N_{sc}} \sum_{\mathbf{k}} \sum_{\mathbf{k}', \mathbf{k}''} \langle i, \mathbf{k}' | \hat{H} | \mathbf{k} \rangle \langle \mathbf{k} | \hat{H} | j, \mathbf{k}'' \rangle \delta(E - \epsilon_{\mathbf{k}}), \quad (\text{S4})$$

$$= \frac{1}{N_{sc}} \sum_{\mathbf{k}} \langle i, \mathbf{k} | \hat{H} | \mathbf{k} \rangle \langle \mathbf{k} | \hat{H} | j, \mathbf{k} \rangle \delta(E - \epsilon_{\mathbf{k}}), \quad (\text{S5})$$

where we used that the graphene and the TMDC layers have the same periodicity. In other words, the translational invariance with respect to the supercell lattice vectors ensures that  $t_{i\mathbf{k}\sigma}$  only couples graphene and vacancy states with the same quasi-momentum.

The single-particle tunneling rate matrix is not diagonal in the quasi-angular momentum basis labeled by  $i, j = \pm$  because the graphene substrate does not share the same symmetry as the TMDC layer. Instead, as discussed in Sec. II of the SI, in the presence of the graphene layer the TMDC global  $C_{3v}$  symmetry is lowered to a  $C_s$  symmetry. Consequently, we can classify the states by using the action of an operator from the group  $C_s$ . For example, we can choose the major plane of symmetry on Fig. 1 whose action on Eq. (S2) is

$$\hat{\sigma}_s |+, \mathbf{k} \rangle = \frac{1}{\sqrt{N_{sc}}} \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} |-, \sigma_s \mathbf{R} \rangle = |-, \sigma_s \mathbf{k} \rangle. \quad (\text{S6})$$

Similarly, we find  $\hat{\sigma}_s |-, \mathbf{k} \rangle = |+, \sigma_s \mathbf{k} \rangle$ . We therefore construct a basis from the quasi-angular momentum states in which  $\hat{\sigma}_s$  is diagonal,  $\hat{\sigma}_s |u \rangle = u |u \rangle$ , with  $u = \pm$  from the linear combination (normalized sum/difference) of the  $i = \pm$  states. In this basis,  $\hat{\sigma}_s |u, \mathbf{k} \rangle = u |u, \sigma_s \mathbf{k} \rangle$  and  $\hat{\sigma}_s |\mathbf{k} \rangle = \exp[i\varphi(\mathbf{k})] |\sigma_s \mathbf{k} \rangle$  where  $\varphi(\mathbf{k})$  is a phase. We now show that Eq. (S5) is a diagonal matrix in this basis, which also defines a symmetry of the Hamiltonian  $\hat{H}$ . Using the observations above, it is straightforward to prove that

$$\sum_{\mathbf{k}} \langle u, \mathbf{k} | \hat{H} | \mathbf{k} \rangle \langle \mathbf{k} | \hat{H} | \bar{u} \mathbf{k} \rangle = \sum_{\mathbf{k}} \langle u, \mathbf{k} | \hat{\sigma}_s^2 \hat{H} \hat{\sigma}_s^2 | \mathbf{k} \rangle \langle \mathbf{k} | \hat{\sigma}_s^2 \hat{H} \hat{\sigma}_s^2 | \bar{u} \mathbf{k} \rangle = - \sum_{\mathbf{k}'} \langle u, \mathbf{k}' | \hat{H} | \mathbf{k}' \rangle \langle \mathbf{k}' | \hat{H} | \bar{u}, \mathbf{k}' \rangle, \quad (\text{S7})$$

with  $\mathbf{k}' = \sigma_s \mathbf{k}$  and  $\bar{u} = -u$ . Here, we used that the phase  $\varphi(\mathbf{k})$  cancels out into the projector  $|\mathbf{k}\rangle\langle\mathbf{k}|$ . As a consequence  $\Gamma_{u\bar{u}}^+ = -\Gamma_{u\bar{u}}^+ = 0$ , i.e. the single-particle rate matrix is diagonal in the orbital subspace,  $\Gamma_{uu'} = \Gamma_{uu'}^+ \delta_{uu'}$  with  $\Gamma_{uu'}^+ = \Gamma_{u'u}^-$ .

We thus consider only the diagonal elements,

$$\Gamma_{uu}^+(E) = \frac{2\pi}{\hbar} \frac{1}{N_{sc}} \sum_{\mathbf{k}} |\langle u, \mathbf{k} | \hat{H} | \mathbf{k} \rangle|^2 \delta(E - \epsilon_{\mathbf{k}}), \quad (\text{S8})$$

and, in the limit  $N_{sc} \rightarrow \infty$ , we transform the discrete sum over the quasimomenta into an integral over the supercell Brillouin zone by

$$\sum_{\mathbf{k}} \rightarrow \frac{N_{sc} A_{sc}}{(2\pi)^2} \int_{\text{SBZ}} d\mathbf{k}, \quad (\text{S9})$$

where  $A_{sc}$  is the supercell area. Thus, Eq. (S8) becomes

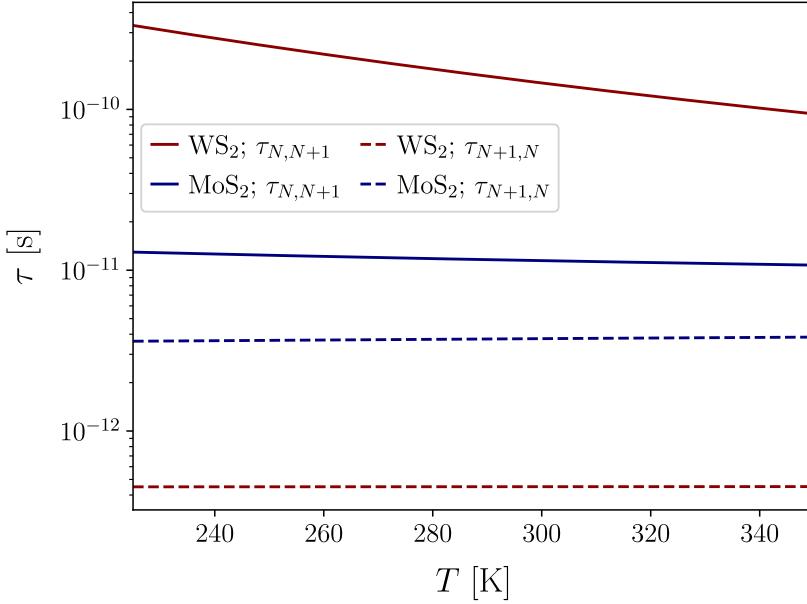
$$\Gamma_{uu}(E) = \frac{2\pi}{\hbar} \frac{A_{sc}}{(2\pi)^2} |t_u|^2 \sum_{\tau} \int_0^{2\pi} d\varphi \int_0^{\infty} dk k \delta(E - \hbar v_F k), \quad (\text{S10})$$

where  $\tau = \pm$  is the valley index,  $v_F$  the Fermi velocity and  $k$  measures the distance of the quasi-momentum from a graphene Dirac point. Moreover,  $t_u := t_{u\mathbf{k}\sigma}$  is defined as in Eq. (S1) and we dropped the  $\pm$  superindex as the matrix is diagonal. Notice that the integral has been extended to  $k \rightarrow +\infty$ , which can be justified by the  $\delta$  function, as far as  $E$  is sufficiently close to the graphene charge neutrality point.

Introducing the graphene unit cell,  $A_c$  and density of states close to its pristine chemical potential,  $\rho_{\text{Gr}}(E) = 2A_c|E|/(\pi\hbar^2v_F^2)$ , we obtain the diagonal elements of the single-particle rate matrix in their final form

$$\Gamma_{uu}(E) = \frac{\pi^2}{2\hbar} \frac{A_{sc}}{A_c} |t_u|^2 \rho_{\text{Gr}}(E). \quad (\text{S11})$$

The final step consists in estimating the absolute value of the hybridization matrix elements, which can be done by employing Eq. (23) in the main text in the context of avoided crossings in a local two-level system.

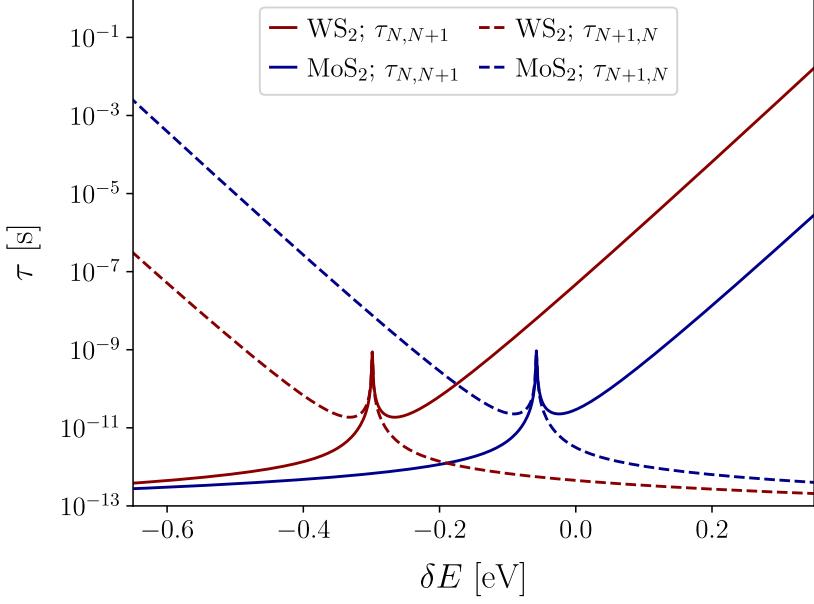


**Fig. S2.** Electronic transition time,  $\tau$ , for MoS<sub>2</sub> and WS<sub>2</sub> vacancy states as a function of bath temperature,  $T$  measured in kelvin, for the transitions  $\alpha \rightarrow \beta$  with  $\alpha, \beta \in \{N, N + 1\}, \alpha \neq \beta$ .

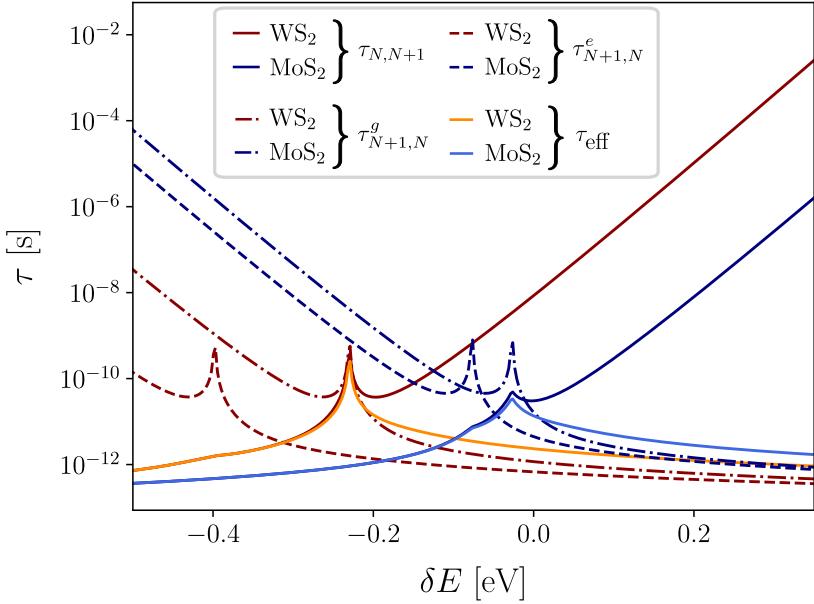
#### IV. TEMPERATURE AND DEFECT ENERGY DEPENDENCE OF THE MANY-BODY TRANSITION RATES

We display in Fig. S2 the transition times evaluated for a wide range of bath temperatures at fixed chemical potential (here  $\mu = 0$ ). For this temperature, well within the weak coupling limit  $\hbar\Gamma \ll k_B T$ , the transition rates are very smooth functions. Temperature opens (or closes) the window around the defect energies where the transition to and from the reservoir can be active, and therefore, larger sensitivity to temperature is expected in WS<sub>2</sub> compared to MoS<sub>2</sub> because its defect levels are further from the Fermi energy of undoped graphene. In addition, as we are essentially observing the effect of temperature on the Fermi function, similar behavior occurs for the case of strong spin-orbit interaction by evaluation of Eqs. (39)-(41).

In Figs. S3 and S4, we show the change in the transition time when the position of the defect levels with respect to the charge neutrality point is rigidly shift. We assume that this shift is not too large so that the hybridization matrix elements remain constant. This type of level shift can be used to understand the impact of the change of the position of the defect levels in the transition rates at lowest order, if screening would be taken into consideration, *e.g.* at the G<sub>0</sub>W<sub>0</sub> level. We consider first on the case of weak spin-orbit interaction. Overall,



**Fig. S3.** Electronic transition times,  $\tau$ , for MoS<sub>2</sub> and WS<sub>2</sub> vacancy states as function of the shift of the defect levels,  $\delta E$ , given in eV and measured from  $\mu = 0$ . This shift would result from *e.g.* including screening effects. The solid and dashed lines correspond to the transitions  $N \rightarrow N + 1$  and  $N + 1 \rightarrow N$  respectively.



**Fig. S4.** Electronic transition times with spin-orbit interaction,  $\tau$ , for vacancy states at the MoS<sub>2</sub>–Gr and WS<sub>2</sub>–Gr interface as function of the shift of the defect levels,  $\delta E$ , given in eV and measured from  $\mu = 0$ . This shift would result from *e.g.* including screening effects. The solid and dashed lines correspond to the transitions  $N \rightarrow N + 1$  and  $N + 1 \rightarrow N$  respectively.

we observe that the trends corresponding to the charging and discharging processes are reversed when comparing to the trends in Fig. 5 (a). For example, the charging of the vacancy becomes slower when the energy levels are shifted away from the charge neutrality point, while the discharging becomes faster. In other words, the positive rigid shift that would occur if screening would be included works in the opposite way as the positive gate voltage (as the pristine band gap becomes larger). In addition, a cusp over the smooth trends is present once the defect levels are shifted towards the graphene Dirac point. This cusp is originated by the alignment of the defect energy with the energy at the  $\bar{K}$  point,  $\delta E = -\Delta E_{0,u}$ , where the inverse DoS acquires a non-analytical scaling as  $\sim |E|^{-1}$ . While an enhancement of the relaxation time should be expected due to these phase space arguments, its divergence is an artefact of the perturbative approach. For example, when considering the tunnelling coupling in the dressed second order approximation<sup>10</sup>, the tunnelling rate for a given many-body transition is obtained by convolution of the DoS with a Lorentzian centered on the resonant energy, thus preventing exact vanishing of the rate. In Fig. ?? we present the electronic transition times in presence of SOC. The transition times to the single energy levels are still characterized by the diverging cusp, which is smoothed in  $\tau_{\text{eff}}$ , as the effective rate never vanishes due to the contribution of both energy levels. Evidently, at the  $G_0W_0$  level the precise value of  $\delta E$  is fixed once the screening conditions are determined, but the numerical simulations suggest both an additional way of modifying the transition rates if the dielectric environment is altered in a controlled way and a manner of interpret and understand screening from the experimentally measured transition rates.

---

<sup>1</sup> P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougaussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Maurim, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Journal of Physics: Condensed Matter **21**, 395502 (2009).

<sup>2</sup> J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

<sup>3</sup> M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze,

and G.-M. Rignanese, Computer Physics Communications **226**, 39 (2018).

<sup>4</sup> G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

<sup>5</sup> O. A. Vydrov and T. Van Voorhis, The Journal of Chemical Physics **130**, 104105 (2009).

<sup>6</sup> K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, and B. I. Lundqvist, Reports on Progress in Physics **78**, 066501 (2015).

<sup>7</sup> E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, 1st ed. (Academic Press, 1959).

<sup>8</sup> M. S. Dresselhaus, G. Dresselhaus, and A. Jorio, *Group Theory: Applications to the Physics of Condensed Matter* (Springer, 2008).

<sup>9</sup> F. Cotton, *Chemical Applications of Group Theory*, 3rd ed., Wiley-Interscience publication (Wiley, 1990).

<sup>10</sup> J. Kern and M. Grifoni, The European Physical Journal B **86**, 384 (2013).