Anomalous scaling of spin accumulation in ferromagnetic tunnel devices with silicon and germanium

S. Sharma,[1,2] A. Spiesser,[1] S. P. Dash,[3] S. Iba,[1] S. Watanabe,[1,4] B. J. van Wees,[2] H. Saito,[1] S. Yuasa,[1] and R. Jansen[1]

[1]*National Institute of Advanced Industrial Science and Technology (AIST), Spintronics Research Center, Tsukuba, Ibaraki 305-8568, Japan*

[2]*Zernike Institute for Advanced Materials, Physics of Nanodevices, University of Groningen, 9747 AG, Groningen, The Netherlands*

[3]*Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-41296, Göteborg, Sweden*

[4]*University of Tsukuba, Tsukuba, Ibaraki 305-8571, Japan*

(Received 18 November 2012; revised manuscript received 12 December 2013; published 3 February 2014)

The magnitude of spin accumulation created in semiconductors by electrical injection of spin-polarized electrons from a ferromagnetic tunnel contact is investigated, focusing on how the spin signal detected in a Hanle measurement varies with the thickness of the tunnel oxide. An extensive set of spin-transport data for Si and Ge magnetic tunnel devices reveals a scaling with the tunnel resistance that violates the core feature of available theories, namely, the linear proportionality of the spin voltage to the injected spin current density. Instead, an anomalous scaling of the spin signal with the tunnel resistance is observed, following a power law with an exponent between 0.75 and 1 over 6 decades. The scaling extends to tunnel resistance values larger than 10^9 $\Omega\,\mu\mathrm{m}^2$, far beyond the regime where the classical impedance mismatch or back flow into the ferromagnet play a role. This scaling is incompatible with existing theory for direct tunnel injection of spins into the semiconductor. It also demonstrates conclusively that the large spin signal does not originate from two-step tunneling via localized states near the oxide/semiconductor interface. Control experiments show that spin accumulation in localized states within the tunnel barrier or artifacts are also not responsible. Altogether, the scaling results suggest that, contrary to all existing descriptions, the spin signal is proportional to the applied bias voltage, rather than the (spin) current.

DOI: 10.1103/PhysRevB.89.075301 PACS number(s): 72.25.Hg, 72.25.Dc, 73.40.Gk, 85.75.−d

I. INTRODUCTION

Mainstream semiconductors such as silicon and germanium play a key role in the development of a spintronics information technology in which spin is used to represent digital data [1–4]. To create and detect spin-polarized carriers in nonmagnetic materials, the use of ferromagnetic tunnel contacts has proven to be a robust and technologically viable approach that is widely used in spin-based devices, including those with Si and Ge [5–24]. As recently reviewed [4,25], controversy has arisen because in many semiconductor spintronic devices, the magnitude of the observed spin voltage differs by several orders of magnitude from what is expected based on the available theory for spin injection and diffusion [26–29]. A common feature of all theories is that the injected spin current produces a spin accumulation $\Delta\mu$, i.e., a spin splitting in the electrochemical potential and thus a spin-dependent occupation of the electronic states in the nonmagnetic material. Conservation of spin-angular momentum requires the injected spin current to be balanced by spin relaxation, from which the steady-state nonequilibrium spin accumulation is evaluated. Consequently, the spin accumulation is predicted to be linearly proportional to the injected spin current.

A powerful way to test the predictions is to vary the thickness of the tunnel barrier, which changes the current density J exponentially. The spin accumulation is expected to exhibit a similar exponential variation, so that $\Delta\mu/J$ remains constant. Here we present an extensive set of spin-transport data on Si and Ge based magnetic tunnel devices with different tunnel oxides. The scaling of the detected spin voltage with tunnel oxide thickness violates the expected linear proportionality of spin voltage and injected spin current. The data are shown to be incompatible with any of the known theories, including those based on direct tunneling [26–29] or two-step tunneling via localized states [30,31].

II. DEVICE FABRICATION

To illustrate the generic nature of the observed scaling, we use devices with heavily doped Si (p-type and n-type) as well as p-type Ge, with an amorphous Al_2O_3 tunnel barrier and $Ni_{80}Fe_{20}$ ferromagnet, or with epitaxial, crystalline MgO/Fe contacts. Tunnel contacts on Si(001) surfaces were fabricated using n-type silicon-on-insulator wafers with a 5-μm-thick active Si layer having As doping and a resistivity of 3 mΩ cm at 300 K, or p-type silicon-on-insulator wafers with a 3-μm-thick active Si layer having B doping and a resistivity of 11 mΩ cm at 300 K. For the contacts with amorphous Al_2O_3, the Si substrate was treated by hydrofluoric acid to remove oxide, the substrate was introduced into the ultrahigh vacuum chamber, and the tunnel barrier was prepared by electron-beam deposition of Al_2O_3 from a single-crystal Al_2O_3 source, followed by plasma oxidation for 2.5 minutes, and electron-beam deposition of the ferromagnetic-metal top electrode (typically 10-nm thick) and a Au cap layer, all at room temperature. For some devices (see Appendix), the plasma oxidation step was omitted. The plasma oxidation leads to the formation of some additional silicon oxide. The actual tunnel oxide thickness is thus slightly larger than the nominal thickness of the deposited Al_2O_3, as previously confirmed and quantified [32] by transmission electron microscopy (TEM). Values of the tunnel oxide thickness quoted in this manuscript correspond to the corrected, actual oxide thickness extracted from TEM, and for convenience this is referred to as the Al_2O_3 thickness.

For epitaxial contacts [33] with MgO/Fe, after treatment with buffered hydrofluoric acid, the Si substrate was annealed in the ultrahigh vacuum deposition system to 700 °C for 10 minutes to obtain a 2×1 reconstructed Si surface. The MgO tunnel barrier and the Fe electrode (10-nm thick) were deposited at 300 °C and 100 °C, respectively, and the crystalline nature of the layers was confirmed by *in situ* reflection high-energy electron diffraction and by high-resolution TEM, as reported recently [33]. Tunnel devices on *p*-type Ge(001) were prepared using Ga-doped wafers with a resistivity of 3 mΩ cm and a carrier concentration of 8.2×10^{18} cm^{-3} at 300 K. The preparation of the epitaxial MgO/Fe tunnel contacts on Ge was as previously described [21].

To probe the spin accumulation over a large range of the tunnel barrier thickness, we employ three-terminal devices [7] in which a single ferromagnetic tunnel contact is used to inject the spin accumulation, and to detect it. This geometry, unlike four-terminal nonlocal devices [6,10], allows the contact area to be chosen arbitrarily large so as to adjust the overall device resistance and thereby ensure a sufficient signal to noise ratio. Here, the tunnel junction area is between 10×10 and 100×200 μm^2. Positive voltage corresponds to electrons tunneling from ferromagnet to semiconductor.

III. SCALING WITH TUNNEL BARRIER THICKNESS

In all devices, voltage signals corresponding to the Hanle and inverted Hanle effect [34] were detected when a magnetic field is applied perpendicular or parallel to the tunnel interface, respectively, at constant tunnel current (see Fig. 1). The Hanle (inverted Hanle) signal originates from the suppression (recovery) of the spin accumulation due to spin precession (or the reduction thereof) and is the signature of the presence of a spin accumulation [7,34]. The most striking observation is that the amplitude of the spin signal (the spin RA product, defined as $\Delta V_{\rm Hanle}/J$, the spin voltage signal per unit of J) increases by orders of magnitude when the thickness of the tunnel barrier is increased. The width of the Hanle curve and the ratio of the Hanle and inverted Hanle amplitudes are also not constant.

The tunnel resistance exhibits the expected exponential variation with thickness of the tunnel oxide [see Fig. 2(a)]. From the slope, we extract an effective tunnel barrier height $\Phi_{\rm eff}$ of 0.8 eV. Taking into account the effective electron mass in Al_2O_3 (about 0.2–0.3 times the free electron mass), this translates into a real barrier height of $\Phi = 3.2 \pm 0.8$ eV. This is a reasonable value [35] for Al_2O_3 on *p*-type Si, showing that direct tunneling from the ferromagnet into the Si is the dominant transport process (for multistep tunneling via localized states within the oxide [36], the extracted barrier height would be four times larger, which is unrealistic). More importantly, the data implies that the contact resistance is dominated by the Al_2O_3, and that the depletion region associated with the Schottky barrier in the Si contributes little to the resistance, as expected for heavily doped Si.

The spin RA product also displays an exponential variation with thickness of the tunnel oxide, and a power law is revealed when the spin RA product is plotted against tunnel resistance [see Figs. 2(b) and 2(c)]. The associated scaling exponent is about 0.75 and 0.82, respectively, for Si/Al_2O_3/$Ni_{80}Fe_{20}$

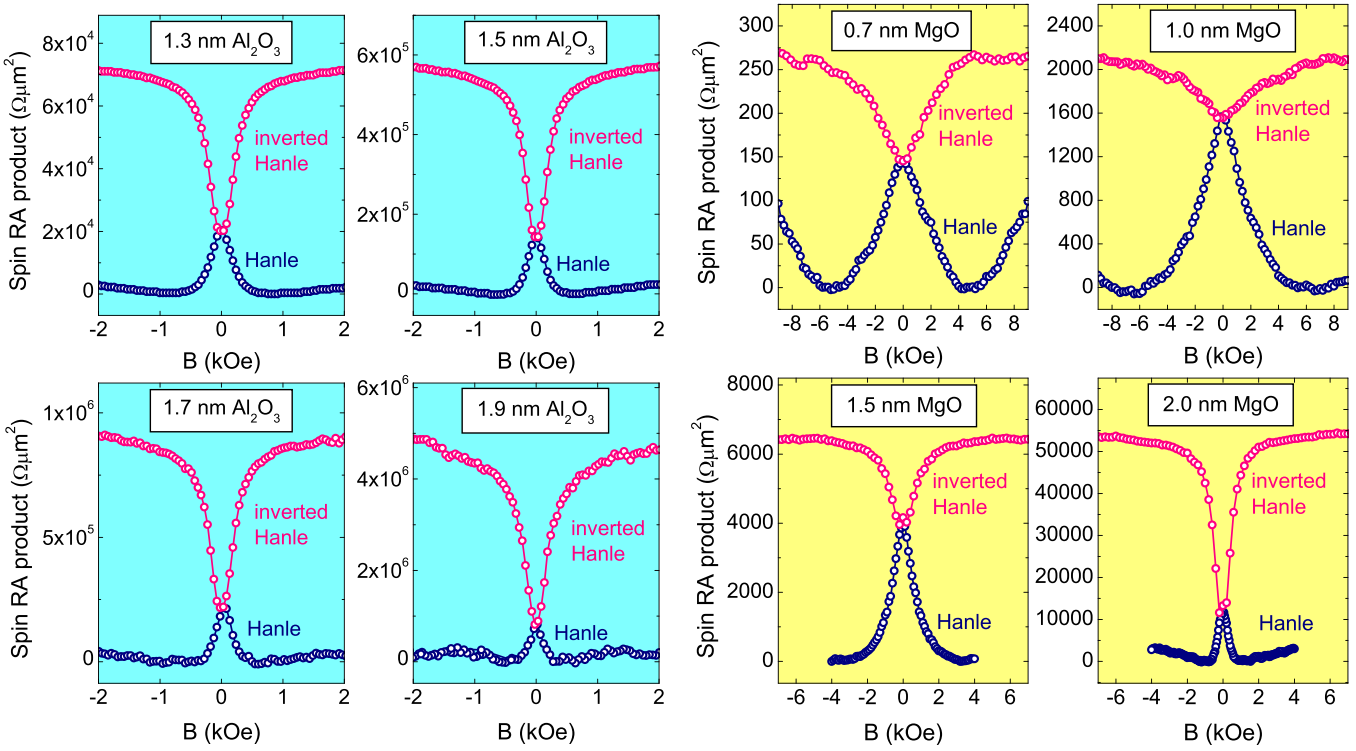

FIG. 1. (Color online) Hanle detection of spin accumulation in semiconductor/oxide/ferromagnet tunnel devices. Shown are representative Hanle and inverted Hanle curves for magnetic field B applied, respectively, perpendicular or parallel to the magnetization of the ferromagnet. Data are shown for *p*-type Si/Al_2O_3/$Ni_{80}Fe_{20}$ and *p*-type Si/MgO/Fe devices with different thickness of the tunnel barrier, all at 300 K. The vertical axis gives the spin-RA product, defined as $\Delta V_{\rm Hanle}/J$.

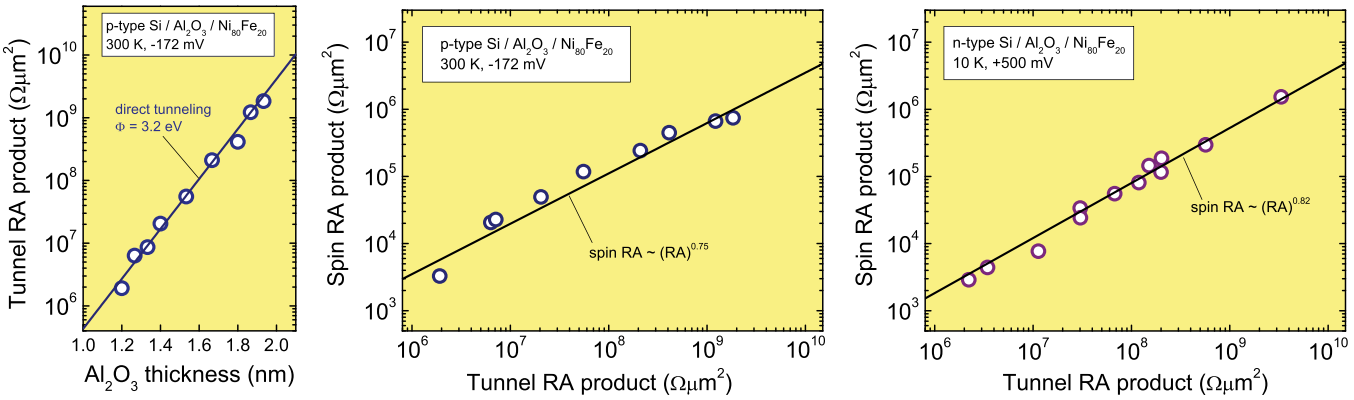

FIG. 2. (Color online) Scaling of Hanle spin signals in Si tunnel devices with amorphous Al_2O_3 barrier. (Left) tunnel resistance-area (RA) product vs Al_2O_3 thickness for p-type Si/Al_2O_3/$Ni_{80}Fe_{20}$ devices at 300 K. The extracted tunnel barrier height is 3.2 eV. (Middle) the corresponding spin RA product versus tunnel RA product. The solid line corresponds to a power law with exponent 0.75. (Right) data for similar devices but with n-type Si at 10 K. The solid line corresponds to a power law with exponent 0.82. The spin RA value is derived from the Hanle signal only, instead of the sum of the Hanle and inverted Hanle signals. See Appendix for additional data with different oxidation time (p-type) and Cs treated surfaces (n-type).

devices with p-type and n-type Si. For devices with crystalline MgO/Fe contacts, a similar exponential variation of spin RA product with MgO thickness is obtained [see Fig. 3(a)]. The contact resistance is dominated by tunneling through the MgO at larger thickness, but for small MgO thickness, a transition occurs to the regime where the contact resistance is limited by the Schottky barrier and becomes constant. Interestingly, the spin RA product displays no transition. It scales with the MgO thickness even in the low-thickness regime, suggesting that the spin signal is determined by the tunneling across the MgO. The spin RA products for p-type Si with MgO/Fe and Al_2O_3/$Ni_{80}Fe_{20}$ contacts display similar scaling as a function of resistance of the tunnel oxide [see Fig. 3(b)], with an exponent (0.75) smaller than 1. For devices on heavily doped p-type Ge with crystalline MgO/Fe contacts [20,21], we also find that R_{tun} and the spin RA product vary exponentially with MgO thickness [see Fig. 3(c)], although the data set is too limited to extract an accurate value for the scaling exponent.

Note that a similar scaling was recently reported by Uemura *et al.* for n-type Si/MgO/$Co_{50}Fe_{50}$ devices [37], although a direct comparison cannot be made because their data was taken with the same bias current for each oxide thickness, and hence with a different tunnel voltage. This, in turn, changes the tunnel spin polarization, which is known to vary with the energy of the tunnel electrons [38,39]. This additional source of variation of the spin signal with tunnel oxide thickness is not present in our data, which was obtained using the same bias voltage for each oxide thickness. Our collection of data leads to the striking and unexpected conclusion that $\Delta V_{\mathrm{Hanle}}/J$ is not a constant but scales with R_{tun}, and up to values larger than $10^9\ \Omega\,\mu m^2$. This behavior is generic, as it is observed for devices with different semiconductors, tunnel oxides, and ferromagnetic electrodes. Below we explain that this behavior is incompatible with any of the known theories for the injection, accumulation, and diffusion of spins in ferromagnetic tunnel devices.

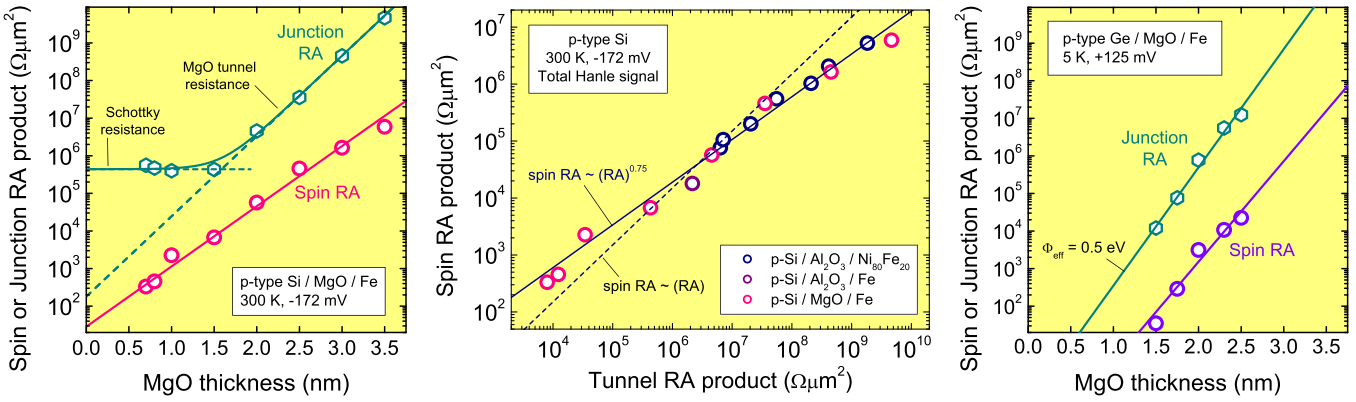

FIG. 3. (Color online) Scaling of Hanle spin signals for devices with MgO and p-type Si or Ge. (Left), spin RA product and tunnel RA product vs MgO thickness for p-type Si/MgO/Fe devices at 300 K. (Middle), corresponding spin RA product vs tunnel RA product for the same devices (pink symbols), together with data for p-type Si/Al_2O_3/$Ni_{80}Fe_{20}$ (blue symbols). For the three devices with the thinnest MgO barrier, the tunnel RA product is determined by extrapolation from the high thickness regime (dashed green line in the left panel) to remove the Schottky resistance. The solid line corresponds to a power law with exponent 0.75. (Right) Data for p-type Ge/MgO/Fe devices at 5 K. The spin RA product is the sum of the Hanle and inverted Hanle signal.

IV. COMPARISON WITH EXISTING THEORY

A. Direct tunneling

In the standard theory, the spin accumulation gives rise to a Hanle spin signal $\Delta V_{\rm Hanle} = [P_{\rm fm}^2]\, r_s\, J$, where $P_{\rm fm}$ is the tunnel spin polarization associated with the oxide/ferromagnet interface, and r_s is the spin resistance of the semiconductor [25,40] that describes the relation between spin current and spin accumulation in nonmagnetic materials. Thus $\Delta V_{\rm Hanle}/J$ is constant and independent of the resistance $R_{\rm tun}$ of the tunnel contact. This applies when $R_{\rm tun}$ is larger than r_s. If $R_{\rm tun} < r_s$, back flow of the spins into the ferromagnet limits the spin signal [25,26,40], which is then proportional to the tunnel resistance: $\Delta V_{\rm Hanle}/J = [P_{\rm fm}^2/(1 - P_{\rm fm}^2)]\, R_{\rm tun}$. Although this produces a scaling with tunnel resistance, the experimentally observed scaling extends to tunnel RA values beyond $10^9\,\Omega\,\mu{\rm m}^2$, and a value of r_s larger than this would be required for back flow to be active. This is unreasonable, since r_s is typically around 10–100 $\Omega\,\mu{\rm m}^2$ for the semiconductors used [25]. The standard description thus predicts that $\Delta V_{\rm Hanle}/J$ is independent of $R_{\rm tun}$ (and up to seven orders of magnitude smaller than observed) and cannot describe the data.

B. Inhomogeneous tunnel current density

It has previously been pointed out that an enlarged spin signal can be produced in three-terminal devices if the tunnel current density is not homogeneous across the contact area [7]. In that case, the local current density, and thereby the spin accumulation, can be significantly larger than what is expected from the applied current and the lateral dimensions of the tunnel contact. In previous work [7], the spin signal was larger than expected by 2–3 orders of magnitude and, in principle, this could be due to lateral inhomogeneity of the tunnel current. However, the new data presented here exhibit a scaling with tunnel barrier resistance that is not readily understandable with an explanation in terms of current inhomogeneity. Moreover, for devices with the thickest tunnel barrier, the observed spin signals are larger than expected by up to six orders magnitude, and this cannot be explained by inhomogeneous tunnel current. It would require that all the tunnel current goes via an area that is 10^6 times smaller than the geometric contact area of $100 \times 200\,\mu{\rm m}^2$. This translates into an effective tunnel area of only $100 \times 200\,{\rm nm}^2$ or so, which is unreasonable. We conclude that inhomogeneity of the tunnel current is not responsible for the experimental observations.

C. Two-step tunneling

Two-step tunneling via localized states near the oxide/semiconductor interface can produce an enhanced spin signal due to spin accumulation in those states [30], provided that certain conditions are satisfied [25]. To obtain a scaling of the spin signal (due to back flow) up to tunnel resistances of $10^9\ \Omega\,\mu{\rm m}^2$, the spin resistance r_s^{ls} of the localized states needs to be at least that large. Since $r_s^{ls} = \tau_s^{ls}/e\,D_{ls}$, where D_{ls} is the density of localized states, one needs localized states with a spin lifetime τ_s^{ls} of at least 10 μs for reasonable values of $D_{ls} > 10^{12}$ states/eV cm^2. Note that state-of-the-art Si/SiO_2 interfaces in commercial Si transistor devices can reach interface state densities two orders of magnitude smaller [41], but this requires a final anneal in hydrogen forming gas, which we do not apply. In fact, for the Si/Al_2O_3 interfaces in our devices, an interface state density of about 10^{13} states/eV cm^2 was previously determined [8], requiring τ_s^{ls} to be even larger (100 μs). While such large values of the spin lifetime seem unreasonable, they cannot be excluded *a priori*. Yet, there exists ample experimental evidence that shows that this mechanism is not the origin of the large spin signals observed, as recently reviewed [4,25]. The scaling data presented here provides additional and conclusive proof that two-step tunneling via interface states cannot be responsible, as it leads to a fundamental inconsistency. The scaling of the contact resistance with oxide thickness [see Fig. 2(a)] implies that the resistance is dominated by the oxide tunnel barrier, and that any resistance r_b of the Schottky barrier in the semiconductor is much smaller. According to the theory for two-step tunneling [25,30], the effective spin resistance $r_s^{\rm eff}$ of the interface states cannot be larger than the resistance r_b that couples the states to the bulk semiconductor. Taken together this would mean $r_s^{\rm eff} < r_b < R_{\rm tun}$. However, in order to obtain a scaling of the spin RA with tunnel resistance (due to back flow from the interface states into the ferromagnet), one needs the opposite, namely, $R_{\rm tun} < r_s^{\rm eff}$. These requirements cannot be satisfied simultaneously, whatever the parameters chosen. Thus Tran's model [30] for spin accumulation in interface states is inconsistent with the simultaneous exponential scaling of contact resistance and spin signal with tunnel barrier thickness. Thus Tran's explanation in terms of pure two-step tunneling via localized states near the oxide/semiconductor interface has to be discarded.

D. Two-step and direct tunneling in parallel

The model introduced by Tran *et al.* [30] for two-step tunneling via localized states assumes that *all* the current goes via the localized states and, as just noted, this cannot describe the experimental data. However, Tran's model has recently been extended [31] by including charge and spin transport by direct tunneling, in parallel with two-step tunneling. It is therefore important to examine whether this extended transport model can describe the experimental data. Depending on the details of the system, the spin accumulation created in localized states due to two-step tunneling can be much larger than that induced in the semiconductor bands by direct tunneling. If as a function of some parameter (e.g., the tunnel barrier thickness), the relative contribution of two-step and direct tunneling is changed, then a transition from a small signal (direct tunneling dominant) to a large signal (two-step tunneling dominant) or vice versa, can be produced. In this paragraph, we examine whether this can explain the observed scaling of the spin RA product with tunnel barrier thickness, and show that this is not possible.

We start by attempting to fit the experimental data for the *p*-type Si/MgO/Fe devices by setting the direct tunnel current to zero. That is, we consider transport by two-step tunneling, where the first step is by tunneling across the MgO from the ferromagnet into interface states, and the second step is by tunneling through the Schottky barrier from interface states into the bulk semiconductor. The result is shown in the left

two panels of Fig. 4, for which the spin resistance of the localized states was set to infinity, so that the spin signal is not limited by spin relaxation in the localized states. A good fit (thick solid lines) is obtained for the junction resistance. For small MgO thickness, the junction resistance is limited by the resistance of the Schottky barrier, whereas at large MgO thickness it is limited by tunneling across the MgO. According to the model, this should be accompanied by a transition in the behavior of the spin RA product, which first increases with MgO thickness, but becomes constant as soon as the junction resistance is determined by the MgO. This corroborates the statement made in the previous paragraph that the junction resistance and spin resistance cannot simultaneously exhibit a scaling with MgO thickness if transport is by two-step tunneling via localized interface states. Note that in the regime of small MgO thickness, the magnitude of the spin signal is determined by the tunnel spin polarization P_{fm} associated with the Fe/MgO interface, and a value of 20% is needed to obtain a match with the data in this regime. With a value of 75%, which is more reasonable [42], the data cannot be described, not even in the regime of small MgO thickness.

The two middle panels show the result if transport is purely by direct tunneling, setting the two-step tunnel current to zero. In principle, the data can be described, however, the required spin resistance r_s^{Si} of the silicon is of the order of $10^8\ \Omega\,\mu\mathrm{m}^2$. This is unreasonable considering that is expected to be in the range of 10–100 $\Omega\,\mu\mathrm{m}^2$ at best, for which one would obtain a spin RA product that is independent of the tunnel oxide thickness and orders of magnitude smaller than experimentally observed (horizontal black line in the bottom middle panel).

Next, we attempt to describe the data by direct and two-step tunneling in parallel, using the equations given in Ref. [31]. As already eluded to above, in order to obtain an increase of the spin RA product as a function of MgO thickness, one needs to have a transition from transport dominated by direct tunneling to transport dominated by two-step tunneling via interface states. Such a situation is depicted in the two right panels of Fig. 4. At large MgO thickness, transport is determined by tunneling across the MgO, and we have chosen the parameters such that in this regime, the resistance associated with two-step tunneling is smaller than that for direct tunneling. At small MgO thickness, the two-step tunnel current is limited by the

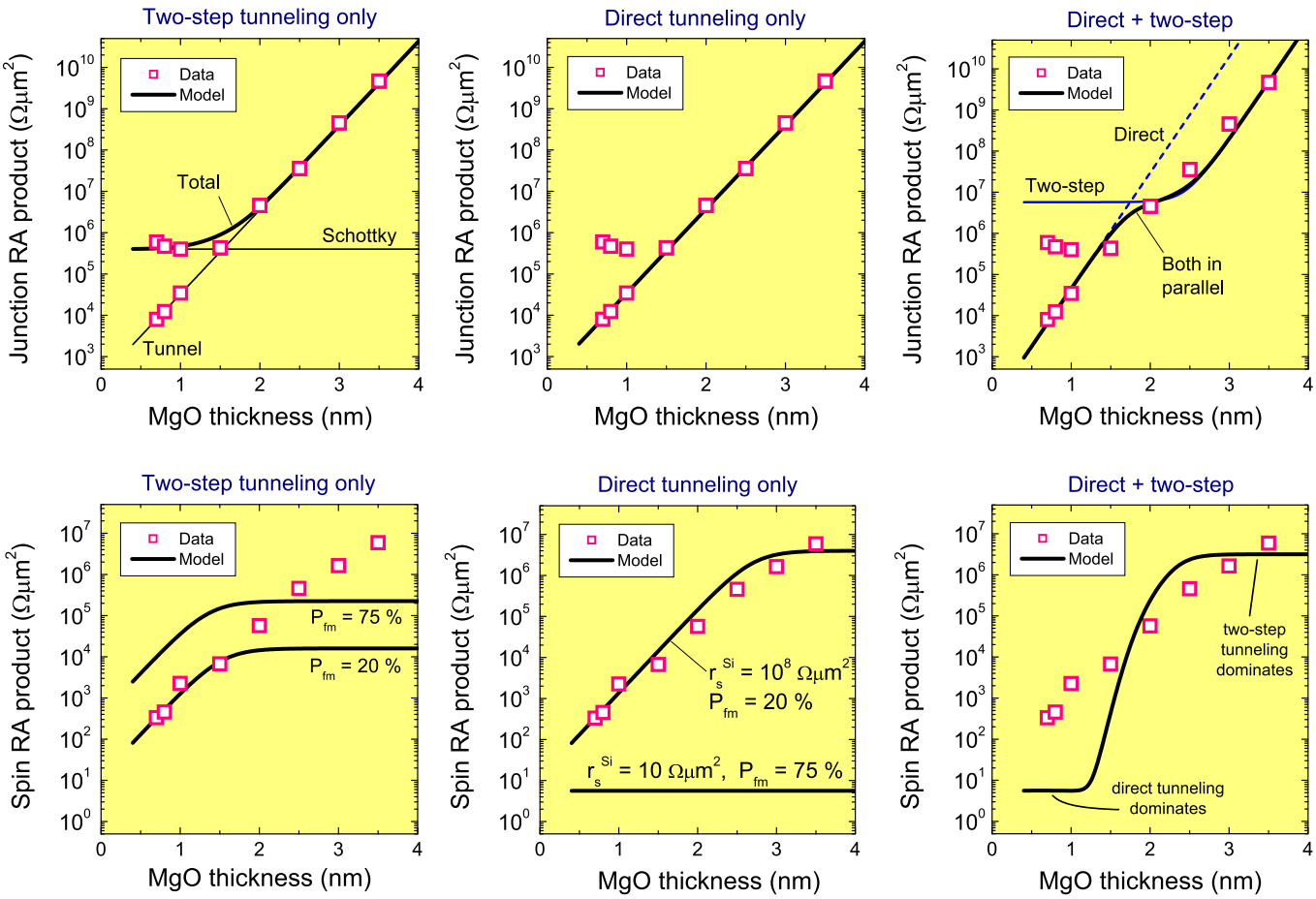

FIG. 4. (Color online) Attempts to fit the data by direct and two-step tunneling in parallel. The experimental data (pink symbols) for the junction resistance and spin RA product of *p*-type Si/MgO/Fe devices is compared with the model (solid lines) for three cases: (i) two-step tunneling only, (ii) direct tunneling only, and (iii) two-step tunneling and direct tunneling in parallel. In the top panels, two data points are given for the junction RA product of the three junctions with smallest MgO thickness; the larger value corresponds to the measured junction RA product, whereas the smaller value is obtained when the resistance of the Schottky barrier is subtracted.

resistance of the Schottky barrier. As a result, the transport at small MgO thickness is dominated by direct tunneling. This change in transport process can reasonably well describe the observed scaling of the junction resistance, but not the scaling of the spin signal. A transition from a small spin RA product, governed by direct tunneling, to an enhanced spin RA product due to two-step tunneling is indeed created, but the model does not reproduce the experimental data. It does not reproduce the observed exponential increase of the spin RA with MgO thickness, and deviates from the data in almost the entire range. We conclude that a transition in transport from direct to two-step tunneling does not describe the experimental data.

E. Two-step tunneling via localized states within the tunnel barrier

The argument used in the previous sections to rule out two-step tunneling via localized interface states is based on the original assumption [30] that the states are located at the oxide/semiconductor interface, and decoupled from the semiconductor bulk bands by a (Schottky) barrier with a characteristic coupling resistance r_b. In principle, it is possible that the relevant localized states are present *within* the oxide tunnel barrier, and that a large spin accumulation is induced in those states by two-step tunneling. Mathematically, the model derived by Tran *et al.* for two-step tunneling [30] and the modified model [31] that includes the parallel current due to direct tunneling, still apply, but for some parameters the relevant physical process is changed. Specifically, the value of r_b that couples the localized states to the semiconductor bands is no longer determined by the Schottky barrier, but by the resistance of part of the oxide tunnel barrier. It is known that two-step tunneling is more efficient for states near the center of the tunnel barrier [36]. The associated value of r_b is then determined by half of the tunnel oxide and would systematically increase with the thickness of the tunnel oxide. Depending on the parameters of the system, this could produce a spin accumulation that increases with tunnel barrier thickness, and thereby a scaling of the spin RA product with tunnel resistance.

Importantly, this scenario also has specific consequences for the spin current that is ultimately injected into the semiconductor channel and for the scaling of the effective time constant that controls the width of the Hanle curve. The behavior depends on the value of the parameters, in particular, the ratio of the spin lifetime τ_s^{ls} of the electrons in the localized states and the escape time $\tau_{\rm esc}$ that defines the leakage of the electrons out of the localized states. These parameters are related, respectively, to the spin resistance $r_s^{ls} = \tau_s^{ls}/e\,D_{ls}$ of the localized states and to $r_b = \tau_{\rm esc}/e\,D_{ls}$. We discuss the different regimes.

(1) $\tau_s^{ls} < \tau_{\rm esc}$. In this case, the spin accumulation in the localized states is limited by spin relaxation in those states and not by coupling resistance r_b. The spin RA product is then given by $[P_{\rm fm}^2]\,r_s^{ls}$ and does not scale with the width of the tunnel barrier. This is inconsistent with our data. Moreover, in this regime, the spin current into the semiconductor is negligible as spins relax in the localized states in the tunnel oxide before they escape into the semiconductor. This is inconsistent with the results of experiments with spin light emitting diodes [5,8,43–46] in which highly spin-polarized electrons were unambiguously detected in the bulk bands of the semiconductor upon electrical injection from similar oxide tunnel contacts based on Al_2O_3, MgO, or SiO_2.

(2) $\tau_s^{ls} > \tau_{\rm esc}$. In this regime, the spin accumulation in the localized states is limited by the leakage of the spins into the semiconductor. If the contribution from the parallel current due to direct tunneling [31] is small enough so that the localized states dominate the spin signal, the spin RA product would be given by $[P_{\rm fm}^2]\,r_b$ and is expected to scale with the width of the tunnel oxide. Also, as previously described [25], in this regime, the spin relaxation in the localized states is negligible because spins escape into the semiconductor fast enough, so that the spin current into the semiconductor channel is not reduced.

Because the trends predicted for the second regime are not inconsistent with the observed scaling of the spin RA product nor with the spin light emitting diode results [5,8,43–46], we cannot *a priori* rule out two-step tunneling via states in the tunnel oxide. It does require that the condition $\tau_s^{ls} > \tau_{\rm esc}$ is satisfied and that r_s^{ls} is at least 10^8–$10^9\ \Omega\,\mu{\rm m}^2$ in order to obtain a scaling of the spin RA product up to the largest values observed. Unfortunately, these conditions can neither be confirmed nor refuted experimentally since we do not know the values of τ_s^{ls} and r_s^{ls}. We can only note that such large values of r_s^{ls} require a large τ_s^{ls} and a small D_{ls}. However, a small density of localized states can only support a small current and the enhancement of the spin signal disappears if the two-step tunnel current is negligible compared to the direct tunnel current [31]. The parameter space for this scenario is thus limited.

Let us discuss another important trend that is predicted for the second regime. When $\tau_s^{ls} > \tau_{\rm esc}$, spin precession in the localized states occurs only during a time interval set by $\tau_{\rm esc}$. Therefore the effective spin precession time that controls the width of the Hanle curve is $\tau_{\rm esc}$ (and not τ_s^{ls}). Since $r_b = \tau_{\rm esc}/e\,D_{ls}$, the effective spin precession time is expected to scale with tunnel barrier thickness in the same way as r_b and the spin RA product. Thus a change of the tunnel barrier width that changes r_b and thus the spin RA product by the observed 4–5 orders of magnitude has to be accompanied by a change of the effective spin precession time and the Hanle linewidth by 4–5 orders of magnitude. We tested this latter prediction experimentally. Figure 5 shows the effective spin lifetime as extracted from the width of the Hanle curves. Although it increases as a function of the tunnel resistance, the increase is rather weak (less than a factor of 10) and much weaker than the expected 4–5 orders of magnitude. Thus the effective spin lifetime does not scale as predicted. Note that for the devices with an Al_2O_3 tunnel barrier, the line width depends slightly on the oxidation time. The spin lifetime for devices with MgO/Fe contacts is smaller than with $Al_2O_3/Ni_{80}Fe_{20}$ contacts. This can be attributed to broadening of the Hanle curve by inhomogeneous magnetostatic fields, which is more pronounced for Fe owing to its larger magnetization [34]. In the presence of such broadening, the lifetime extracted from the Hanle curve is smaller than the real spin-relaxation time. This effect may also obscure the observation of the true scaling of the spin precession time. In fact, as soon as the product of the spin precession time and the typical Larmor

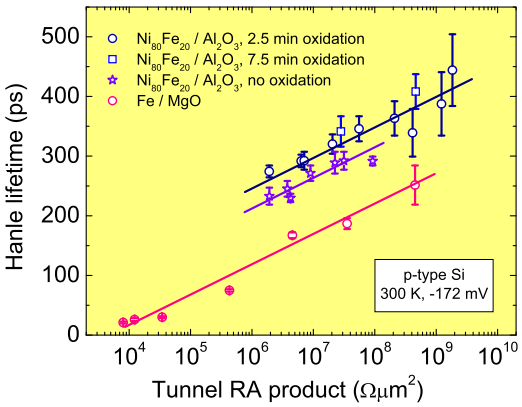

FIG. 5. (Color online) Hanle linewidth vs tunnel resistance. The Hanle linewidth is characterized by an effective lifetime obtained from a fit of the Hanle curve using a Lorentzian. This time constant is a lower bound to the spin lifetime, as previously discussed [7,25,34]. The solid lines are guides to the eye.

frequency associated with the inhomogeneous magnetostatic fields exceeds unity, the width of the Hanle curve is set by the strength of the inhomogeneous magnetostatic fields [34], and does no longer scale with the lifetime.

We conclude that based on the available experimental data, we cannot exclude (nor prove) an interpretation in terms of two-step tunneling via localized states within the tunnel oxide. However, the parameter space for this interpretation is rather limited, and it would seem unlikely that all the available experimental data obtained by many different research groups would yield such similar results. In any case, if this scenario holds, there is still a spin current injected into the semiconductor bulk bands and it is not reduced by spin relaxation in the localized states because τ_s^{ls} would be larger than τ_{esc}.

F. Localized states in the barrier near the ferromagnetic interface

It was recently suggested by Uemura *et al.* [37] that a scaling of the spin signal with tunnel resistance can be obtained if the tunnel probability depends on a magnetic field. While this is mathematically correct, no plausible physical mechanism was given to produce a tunnel probability that depends on a magnetic field in a way that would yield a Hanle or Hanle-like voltage signal. Since the tunnel time is much shorter than the spin precession time for the magnetic fields used, spin precession during the tunneling process is negligible. Uemura *et al.* therefore also invoked the existence of a spin accumulation in localized states in the tunnel oxide, but assumed that the relevant states are located near the interface with the ferromagnet. However, this is a highly unlikely scenario because such states would be efficiently coupled to the nearby ferromagnet, which acts as a spin sink. The short time for the escape of spins into the ferromagnet would prevent the built up of any spin accumulation in those localized states. Moreover, the reason why such a spin accumulation, if it would exist, affects the tunnel probability or its spin polarization remains unclear. Hence, the mechanism proposed by Uemura *et al.* has no physical basis. In addition, it is well known [36] that the probability for two-step tunneling via states in the tunnel barrier is a strongly peaked function of the location of the states, and two-step tunneling is normally dominated by states close to the middle of the barrier. One would thus not expect a significant contribution to the tunnel current from states near the interface with the ferromagnet.

V. CONTROL DEVICES

A. Control devices with metal instead of semiconductor

In order to provide further insight into the possible role of localized states within the tunnel oxide, we fabricated

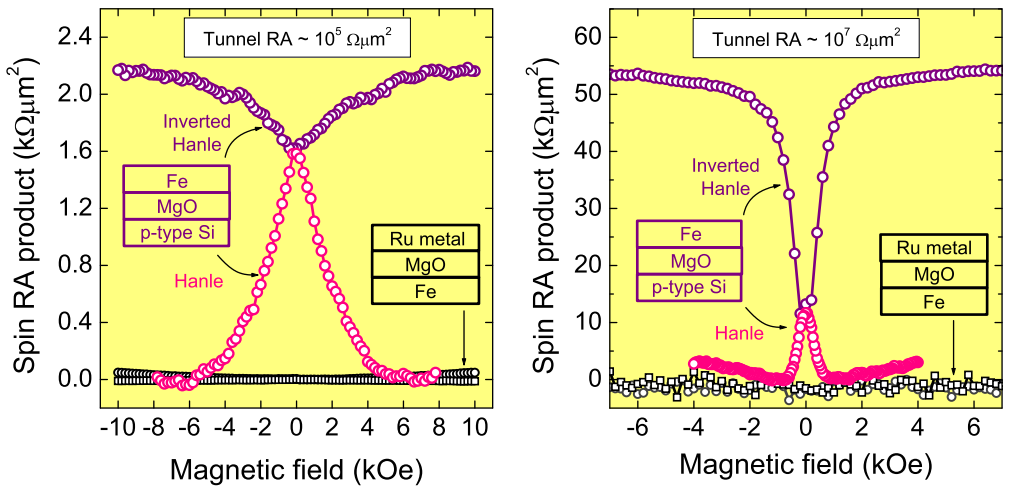

FIG. 6. (Color online) Absence of spin accumulation in control tunnel devices with the Si replaced by nonmagnetic Ru metal. Shown are the Hanle and inverted Hanle signal for *p*-type Si/MgO/Fe devices, together with similar data on control devices with Fe/MgO/Ru structure (black circles and squares). Measurements were done at 300 K, and devices with a similar tunnel RA product are compared (10^5 and 10^7 $\Omega\,\mu m^2$ for the left and right panels, respectively). The absence of any spin signal in the metallic control devices proves the absence of spin accumulation in localized states within the MgO tunnel barrier.

control devices in which the semiconductor is replaced by a nonmagnetic metal (Ru) electrode. If the large spin RA product originates from spin accumulation in states within the tunnel oxide, the spin accumulation does not depend on the spin resistance of the nonmagnetic electrode, and a similarly large spin accumulation should be observed with a Ru metal electrode. However, in control devices with the structure Fe/MgO/Ru, no spin signal could be observed, neither Hanle nor inverted Hanle (see Fig. 6). Therefore we conclude that the spin signal does not originate from spin accumulation in localized states in the tunnel oxide. This control experiment also rules out the recent proposal [37] of spin accumulation in localized states in the tunnel barrier close to the oxide/ferromagnet interface.

Our result appears to be different than that reported recently by Txoperena *et al.* [47] for metal junctions ($NiFe/Al_2O_3/Al$ and $NiFe/Al_2O_3/Au$). In those devices anomalously large Hanle and inverted Hanle signals were observed, and the width of the Hanle curve was not different for Al and Au nonmagnetic electrodes. It was therefore concluded that the signal does not originate from spin accumulation in the nonmagnetic metal electrode, but originates from the tunnel barrier (although a specific mechanism was not given). More recent work by the same authors [48] has, however, shed important new light on these results. It was found that the appearance of the Hanle signals in those metal tunnel junctions depends on the procedure used to fabricate the Al_2O_3 tunnel barrier. Whereas large spin signals are observed in all junctions prepared by natural oxidation of Al (as originally reported [47]), when the Al was oxidized in an oxygen plasma, Hanle signals were systematically absent (with one isolated exception that was included in the original report [47]). We therefore conjecture that the devices for which natural oxidation was used suffer from incomplete oxidation of the tunnel barrier. Spin injection into unoxidized or partially oxidized Al inclusions in the tunnel barrier would create a large spin accumulation by virtue of the small volume of those inclusions, whereas the spin lifetime would not depend on the nonmagnetic electrode (Al or Au). Proper oxidation of the tunnel barrier would avoid this. Since for our Al_2O_3-based devices we have prepared the barrier by deposition of Al_2O_3 from an Al_2O_3 target followed by plasma oxidation, we can be sure that such inclusions do not exist in our devices.

B. Control devices with zero tunnel spin polarization

Given that the experimental data deviate fundamentally from the theory, it is of the utmost importance to convincingly establish that the observed spin signals are genuine and originate from spin accumulation, rather than some kind of measurement artifact. Such potential artifacts can arise from (anisotropic) magnetoresistance effects related to the current through the ferromagnetic electrode itself, or from the effect of magnetic fields on charge transport in the semiconductor (Hall voltages etc.). A powerful way to exclude these artifacts is to introduce a thin nonmagnetic layer at the interface between the tunnel oxide and the ferromagnet, without removing the ferromagnet [49]. The method relies on the extreme interface sensitivity of (spin-polarized) tunneling, such that insertion of a thin nonmagnetic layer causes the tunnel spin polarization to vanish, and hence the spin accumulation. Genuine spin signals should then disappear, whereas any signals due to artifacts, if present, would still remain. This approach was previously used to rule out artifacts in the experiments by Dash *et al.* [7,50], although only the signal for out-of plane magnetic field (Hanle curve) was investigated, and only in the range of small field. In Fig. 7 , a more complete characterization is presented, showing measurements on a control device in the Hanle as well as the inverted Hanle geometry, and for fields up to 50 kOe. No spin signals are observed. This implies that the signals (Hanle and inverted Hanle) observed in the regular devices (without the nonmagnetic interlayer) are not due to an artifact but originate from spin-polarized tunneling and the spin accumulation this produces. This result corroborates previous experiments on spin injection from similar ferromagnetic tunnel contacts into a silicon light emitting diode [5,8], from which the presence of spin-polarized carriers inside the silicon was unambiguously established.

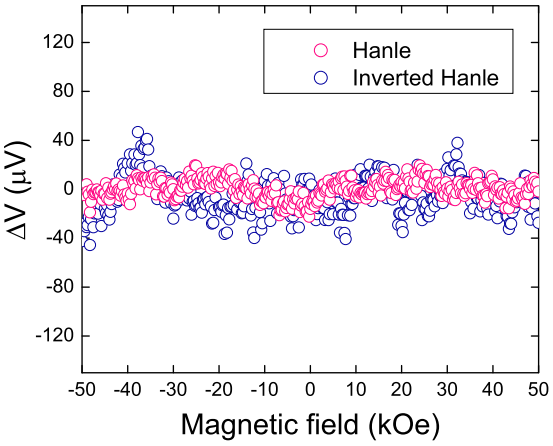

FIG. 7. (Color online) Absence of spin signals in control tunnel devices with zero tunnel spin polarization. Shown are the Hanle and inverted Hanle signal for a control device with structure *p*-type $Si/Al_2O_3/Au(10nm)/Ni_{80}Fe_{20}$, in which the nonmagnetic Au interlayer causes the tunnel spin polarization to be zero. Measurements are done at room temperature with a constant current of $-195\ \mu A$ (hole injection condition). A constant bias voltage of about -172 mV was subtracted from the data.

VI. DISCUSSION AND CONCLUSION

It has previously been noted that the magnitude of the spin accumulation signal observed in magnetic tunnel devices on semiconductors is significantly different than that predicted by the theory of spin injection, diffusion and detection, first for GaAs based devices [30], and subsequently also for Si and Ge based devices [4,7]. The results presented here provide an even larger discrepancy (of up to seven orders of magnitude), and perhaps more importantly, reveal that the scaling with tunnel barrier resistance deviates universally from theory in a fundamental way. The scaling extends over a wide range of tunnel resistance, down to the lowest tunnel RA values of about 10 $k\Omega\,\mu m^2$. It would certainly be of interest to extend the measurements to devices with even lower tunnel resistance.

Our results reveal that care has to be taken when in a particular device the observed magnitude of the spin signal is found to be in agreement with theory, because this could just be accidental. For instance, the scaling trend predicts that at small junction RA product there must be a point where experiment and theory are in agreement, but a more detailed investigation varying the tunnel barrier thickness would reveal a discrepancy. This point of "accidental agreement" will shift to larger junction RA product when the thickness of the semiconductor channel is reduced, because the theory predicts a larger spin signal when the volume of semiconductor into which spins are injected is decreased. Clearly, one needs to look beyond the magnitude of the spin signal in order to (in)validate the theory, and demonstrate that systematic trends predicted by theory are observed in a series of devices in order to support claims of agreement with theory.

While the above results are obtained with three-terminal devices and the observed signal is larger than predicted, in silicon-based nonlocal devices [10] the observed signal deviates in the opposite direction, i.e., it is about two orders of magnitude smaller than expected, as recently noted [4]. Although in the latter case there can be several other reasons, the results presented here suggest that the difference between experiment and theory in three-terminal and nonlocal devices has a common origin, namely, the incompleteness of the existing theoretical descriptions. Several explanations for the discrepancy had so far been proposed. These include two-step tunneling via localized states near the semiconductor/oxide interface [30], lateral inhomogeneity of the tunnel current density [7], two-step tunneling in parallel with direct tunneling [31], or two-step tunneling via localized states near the oxide/ferromagnet interface [37]. The scaling results presented here, together with the control experiments, show unambiguously that none of these proposals can explain the results and that localized states within the tunnel oxide are also not responsible. It is unclear whether the discrepancy arises from an incorrect description of the magnitude of the spin accumulation that is induced by spin injection, or from the description of the conversion of the induced spin accumulation into a voltage signal in a Hanle measurement. Obviously, resolving this puzzle is of crucial importance for application of magnetic tunnel contacts in semiconductor spintronic devices.

ACKNOWLEDGMENTS

The authors are grateful to T. Nozaki for his help with the growth of the Ru-based control devices. The authors acknowledge financial support from the Netherlands Foundation for Fundamental Research on Matter (FOM), and from the Japanese Funding Program for Next Generation World-Leading Researchers (No. GR099). One of the authors (A.S.) acknowledges a JSPS Postdoctoral Fellowship for Foreign Researchers.

APPENDIX: ADDITIONAL DATA

To investigate the effect of localized states produced by oxygen vacancies within the oxide tunnel barrier, we fabricated devices with p-type Si and Al_2O_3 tunnel barrier, but without the plasma oxidation step. Since the Al_2O_3 is grown by electron-beam deposition, the deposited oxide is oxygen deficient. We found that there is no effect on the spin signal, i.e., junctions with and without the plasma oxidation have the same spin RA product at the same tunnel resistance (see Fig. 8, left panel).

In order to investigate the effect of the resistance r_b of the depletion region in the Si, the Schottky barrier height was reduced (and with it r_b) using the procedure with a Cs treatment of the Si surface that was previously developed [7,8]. Here, we present similar data as in Ref. [7] for n-type $Si/Al_2O_3/Ni_{80}Fe_{20}$ devices with and without Cs, but now as a function of tunnel barrier thickness (see Fig. 8, right panel). The Cs treatment produces no change of the spin RA product, and it scales to values of $10^6\ \Omega\,\mu m^2$. This is not compatible with a description in terms of two-step tunneling via localized states at the oxide/Si interface. Owing to the small value of r_b for the devices treated with Cs, a large spin accumulation cannot built up in the interface states because spins will leak away efficiently into the silicon.

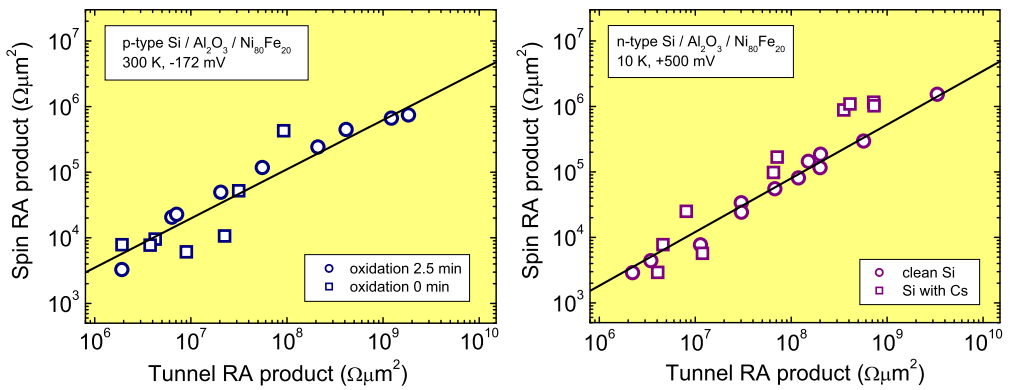

FIG. 8. (Color online) Additional data on scaling of Hanle spin signals in tunnel devices with n-type and p-type Si. The open circles are the same data as in Fig. 2, whereas squares are additional data for devices with different oxidation time (p type, left panel) and with Cs treated surfaces (n type, right panel).

[1] C. Chappert, A. Fert, and F. Nguyen van Dau, Nat. Mater. **6**, 813 (2007).
[2] D. D. Awschalom and M. E. Flatté, Nat. Phys. **3**, 153 (2007).
[3] A. Fert, Rev. Mod. Phys. **80**, 1517 (2008).
[4] R. Jansen, Nat. Mater. **11**, 400 (2012).
[5] B. T. Jonker, G. Kioseoglou, A. T. Hanbicki, C. H. Li, and P. E. Thompson, Nat. Phys. **3**, 542 (2007).
[6] O. M. J. van 't Erve *et al.*, Appl. Phys. Lett. **91**, 212109 (2007).
[7] S. P. Dash, S. Sharma, R. S. Patel, M. P. de Jong, and R. Jansen, Nature (London) **462**, 491 (2009).
[8] R. Jansen, B. C. Min, S. P. Dash, S. Sharma, G. Kioseoglou, A. T. Hanbicki, O. M. J. van 't Erve, P. E. Thompson, and B. T. Jonker, Phys. Rev. B **82**, 241305 (2010).
[9] T. Sasaki, T. Oikawa, T. Suzuki, M. Shiraishi, Y. Suzuki, and K. Noguchi, Appl. Phys. Lett. **96**, 122101 (2010).
[10] T. Suzuki, T. Sasaki, T. Oikawa, M. Shiraishi, Y. Suzuki, and K. Noguchi, Appl. Phys. Express **4**, 023003 (2011).
[11] K. R. Jeon, B. C. Min, I. J. Shin, C. Y. Park, H. S. Lee, Y. H. Jo, and S. Ch. Shin, Appl. Phys. Lett. **98**, 262102 (2011).
[12] Y. Ando, K. Kasahara, K. Yamane, Y. Baba, Y. Maeda, Y. Hoshi, K. Sawano, M. Miyao, and K. Hamaya, Appl. Phys. Lett. **99**, 012113 (2011).
[13] Y. Ando, Y. Maeda, K. Kasahara, S. Yamada, K. Masaki, Y. Hoshi, K. Sawano, K. Izunome, A. Sakai, M. Miyao, and K. Hamaya, Appl. Phys. Lett. **99**, 132511 (2011).
[14] T. Inokuchi, M. Ishikawa, H. Sugiyama, Y. Saito, and N. Tezuka, J. Appl. Phys. **111**, 07C316 (2012).
[15] M. Ishikawa, H. Sugiyama, T. Inokuchi, K. Hamaya, and Y. Saito, Appl. Phys. Lett. **100**, 252404 (2012).
[16] H. Saito, S. Watanabe, Y. Mineno, S. Sharma, R. Jansen, S. Yuasa, and K. Ando, Solid State Commun. **151**, 1159 (2011).
[17] Y. Zhou, W. Han, L. T. Chang, F. Xiu, M. Wang, M. Oehme, I. A. Fischer, J. Schulze, R. K. Kawakami, and K. L. Wang, Phys. Rev. B **84**, 125323 (2011).
[18] K.-R. Jeon, B.-C. Min, Y.-H. Jo, H.-S. Lee, I.-J. Shin, C.-Y. Park, S.-Y. Park, and S.-C. Shin, Phys. Rev. B **84**, 165315 (2011).
[19] A. Jain *et al.*, Appl. Phys. Lett. **99**, 162102 (2011).
[20] S. Iba, H. Saito, A. Spiesser, S. Watanabe, R. Jansen, S. Yuasa, and K. Ando, Appl. Phys. Express **5**, 023003 (2012).
[21] S. Iba, H. Saito, A. Spiesser, S. Watanabe, R. Jansen, S. Yuasa, and K. Ando, Appl. Phys. Express **5**, 053004 (2012).
[22] K. Kasahara, Y. Baba, K. Yamane, Y. Ando, S. Yamada, Y. Hoshi, K. Sawano, M. Miyao, and K. Hamaya, J. Appl. Phys. **111**, 07C503 (2012).
[23] O. M. J. van 't Erve, A. L. Friedman, E. Cobas, C. H. Li, J. T. Robinson, and B. T. Jonker, Nat. Nanotechnol. **7**, 737 (2012).
[24] A. Dankert, R. S. Dulal, and S. P. Dash, Sci. Rep. **3**, 3196 (2013).
[25] R. Jansen, S. P. Dash, S. Sharma, and B. C. Min, Semicond. Sci. Technol. **27**, 083001 (2012).
[26] A. Fert and H. Jaffrès, Phys. Rev. B **64**, 184420 (2001).
[27] A. Fert, J.-M. George, H. Jaffrès, and R. Mattana, IEEE Trans. Elec. Dev. **54**, 921 (2007).
[28] S. Takahashi and S. Maekawa, Phys. Rev. B **67**, 052409 (2003).
[29] Y. Song and H. Dery, Phys. Rev. B **81**, 045321 (2010).
[30] M. Tran, H. Jaffrès, C. Deranlot, J.-M. George, A. Fert, A. Miard, and A. Lemaître, Phys. Rev. Lett. **102**, 036601 (2009).
[31] R. Jansen, A. M. Deac, H. Saito, and S. Yuasa, Phys. Rev. B **85**, 134420 (2012).
[32] B. C. Min, Ph.D. thesis, Koninklijke Wöhrmann, Zutphen, The Netherlands, 2007.
[33] A. Spiesser, S. Sharma, H. Saito, R. Jansen, S. Yuasa, and K. Ando, Proc. SPIE **8461**, 84610K (2012).
[34] S. P. Dash, S. Sharma, J. C. Le Breton, J. Peiro, H. Jaffrès, J.-M. George, A. Lemaître, and R. Jansen, Phys. Rev. B **84**, 054410 (2011).
[35] J. Robertson, Rep. Prog. Phys. **69**, 327 (2006).
[36] Y. Xu, D. Ephron, and M. R. Beasley, Phys. Rev. B **52**, 2843 (1995).
[37] T. Uemura, K. Kondo, J. Fujisawa, K. Matsuda, and M. Yamamoto, Appl. Phys. Lett. **101**, 132411 (2012).
[38] S. O. Valenzuela, D. J. Monsma, C. M. Marcus, V. Narayanamurti, and M. Tinkham, Phys. Rev. Lett. **94**, 196601 (2005).
[39] B. G. Park, T. Banerjee, J. C. Lodder, and R. Jansen, Phys. Rev. Lett. **99**, 217206 (2007).
[40] J. Fabian, A. Matos-Abiague, C. Ertler, P. Stano, and I. Žutić, Acta Phys. Slov. **57**, 565 (2007).
[41] N. H. Thoan, K. Keunen, V. V. Afanas'ev, and A. Stesmans, J. Appl. Phys. **109**, 013710 (2011).
[42] S. Yuasa and D. D. Djayaprawira, J. Phys. D: Appl. Phys. **40**, R337 (2007).
[43] A. Najmaie, E. Y. Sherman, and J. E. Sipe, Phys. Rev. Lett. **95**, 056601 (2005).
[44] L. Grenet, M. Jamet, P. Noé, V. Calvo, J.-M. Hartman, L. E. Nistor, B. Rodmacq, S. Auffret, P. Warin, and Y. Samson, Appl. Phys. Lett. **94**, 032502 (2009).
[45] G. Kioseoglou, A. T. Hanbicki, R. Goswami, O. M. J. van 't Erve, C. H. Li, G. Spanos, P. E. Thompson, and B. T. Jonker, Appl. Phys. Lett. **94**, 122106 (2009).
[46] C. H. Li, G. Kioseoglou, O. M. J. van 't Erve, P. E. Thompson, and B. T. Jonker, Appl. Phys. Lett. **95**, 172102 (2009).
[47] O. Txoperena, M. Gobbi, A. Bedoya-Pinto, F. Golmar, X. Sun, L. E. Hueso, and F. Casanova, Appl. Phys. Lett. **102**, 192406 (2013).
[48] Private communication with F. Casanova and results presented at the SPIE Optics and Photonics Conference, San Diego, USA (25–29 August 2013), and at the 58^{th} MMM conference, Denver, USA (4–8 November 2013).
[49] R. S. Patel, S. P. Dash, M. P. de Jong, and R. Jansen, J. Appl. Phys. **106**, 016107 (2009).
[50] S. P. Dash, S. Sharma, J. C. Le Breton, and R. Jansen, Proc. SPIE **7760**, 77600J (2010).