Heterojunction band lineups in Si-Ge alloys using spatially resolved electron-energy-loss spectroscopy

J. F. Morar, P. E. Batson, and J. Tersoff

IBM Thomas J. Watson Research Center, Yorktown Heights, New York, New York 10598

We describe a new method for determining band lineups at semiconductor heterojunctions, applicable to Si-Ge systems. Spatially resolved transmission electron-energy-loss spectroscopy is used to detect *bulk* core-to-conduction-band transitions in alloys. The core levels then provide relatively accurate "reference levels" for aligning bands across an interface. This approach eliminates the major sources of experimental uncertainty in other methods, while permitting the study of local regions within more complex structures (e.g., real devices). Valence-band offsets in Si-Ge systems are found to vary linearly with Ge content, reaching 0.78 eV (strain-averaged) for interfaces between pure Ge and Si.

The most fundamental property of a semiconductor interface is the band lineup, i.e., the relative positions of the band edges across the interface. This lineup determines the effective potential barrier felt by an electron or hole at the interface. Yet it has proved surprisingly difficult to measure band lineups accurately.[1] Here we describe a new approach, using electron-energy-loss spectroscopy (EELS) of *bulk* Si-Ge alloys to accurately infer the band lineups at heterojunctions. This method improves upon an approach introduced by Shih and Spicer,[2] and avoids the most serious problems of previous methods.

Band offsets can be inferred from measurements of transport properties, such as interface capacitance. However, such methods are plagued by sensitivity to extraneous effects, such as charge trapping and recombination currents.[1] Photoemission spectroscopy (PES) of core levels represents an attractive alternative.[1] In that approach, one measures the energies of core levels on each side of the interface. The positions of the valence-band edges, relative to the core level, are determined in separate bulk measurements. Combining these results gives the band-edge positions on both sides of the interface, and thus the lineup.

While avoiding the complexities of transport, core-level spectroscopy still poses severe demands. One must (a) prepare samples of good interface quality, (b) ensure that band bending is negligible over the portion of the sample probed, (c) accurately measure the core-level positions in the sample, and (d) accurately measure the position of the band edge, relative to the core level, in corresponding bulk samples. For strained layers, the absence of strain in the bulk samples raises additional issues.[2]

In a potentially crucial simplification, Shih and Spicer proposed substituting *bulk* alloy samples for the actual interface.[2] They postulated that the core-level positions were approximately independent of alloy composition. The band edges could then be measured relative to a given core level in two materials, and their difference would provide the band offset at a heterojunction. This approach was also used in a theoretical analysis by Wei and Zunger.[4]

The method was applied to HgTe-CdTe,[2] giving an accurate lineup while avoiding problems (a) and (b) above, i.e., sample preparation and band bending. The principal remaining source of experimental uncertainty in this approach is (d), the determination of the position of the band edge relative to the core level. The problem here is that the valence-band edge in PES is not sharp, and its shape is not well known, being affected by wave-vector selection rules, transmission through the surface, and the angular aperture of the detector.

In addition, this approach introduces a new source of uncertainty, stemming not from the measurement itself, but from the assumption that core-level positions are constant across a heterojunction. Shih and Spicer recognized that this assumption could not be *exactly* correct, but argued that the corrections should be small for the lattice-matched common-anion systems HgTe-CdTe and AlAs-GaAs. However, subsequent calculations by Massidda, Min, and Freeman[5,6] raised questions about the accuracy of the constant-core-level approximation for AlAs-GaAs.

Here, we address the two primary sources of uncertainty in determining band lineups from core levels. First, we show that in the case of Si-Ge systems, the assumption of approximate core-level constancy can be justified more rigorously than in previous analyses. Then, we describe the use of transmission EELS to measure the energy of the transition from the core level to the conduction-band minimum (CBM). This transition is free of the wave-vector selection rules and surface effects that complicate the interpretation of valence-band PES in step (d) above. As a result, the threshold is sharp and has a simple shape, so it can be measured rather precisely.

It may seem paradoxical that one can determine the band lineup without performing a measurement on an actual interface. In fact, a common theme in discussions of band lineups is the search for a "reference level."[7] If there exists an energy level in each material, such that these levels will line up across an interface, then one need only determine the band edge *relative* to this reference level in each material separately; the difference between the two band edges will be the same whether there is an interface or not. Here we propose to use the Si $2p_{3/2}$ core

level as a reference level for Si-Ge alloys; so we must first justify the approximation of core-level constancy, and quantify the errors so introduced.

The effects of alloying can be examined by imagining a semi-infinite sample of pure Si. Atoms are then "transmuted" from Si to Ge to obtain alloy layers of various compositions. We restrict transmutations to a finite surface region, so the average lattice constant parallel to the surface remains that of the substrate. The Si $2p_{3/2}$ core level (or any convenient core level) deep in the substrate provides a convenient fixed reference energy E_c^0.

Each time a Si atom is transmuted to Ge, there are two effects that shift the energies E_c of core levels on nearby Si atoms: there is a direct electronic response to the change in the ion core; and also, the atoms move slightly in response to the larger size of Ge atoms. Because the total perturbation is weak (in a pseudopotential sense), we can use linear-response theory;[8] then to first order, the core-level shift of a given Si atom is simply the sum of the separate shifts due to the electronic and structural response to each Ge atom in the sample.

The electronic shift corresponds to the dielectric response to a local, neutral perturbation. This response is extremely short ranged,[9] so we assume that only core levels on nearest-neighbor Si atoms will be appreciably shifted, by an amount Δ_d.

We can evaluate the response to atomic displacements for pure Si, rather than for the actual composition, since the difference between these responses is of second order in the perturbation. For pure Si, it should be an excellent approximation to assume that the core-level shift of a given atom depends only on the arrangement of its four neighbors. Moreover, we are concerned only with net shifts, since splitting of a degenerate core level will not be resolved; only the distance of each neighbor has an appreciable effect on the net shift. Thus each neighbor will contribute a core-level shift $\Delta_r = (r - r_s)dE_c/dr$ associated with the atomic displacement, where r is the actual bond length to that neighbor, and r_s is the bulk Si bond length. (This formula is in fact rigorously correct for uniform strain.)

To an excellent approximation (about 0.01 Å),[10] Si-Si bonds in the alloy have an average bond length $(1+\varepsilon)r_s$, while Si-Ge bonds have bond length $(1+\varepsilon)r_{SG}$, where r_{SG} is the average of the bulk Si and Ge bondlengths, and ε is the hydrostatic component of the strain in the given layer.

Thus the total core-level shift $\Delta E_c \equiv E_c - E_c^0$ for a given Si atom can be written as the sum of a "chemical shift," which is proportional to the number n_{Ge} of Ge neighbors, and an "absolute deformation potential," which is proportional to the hydrostatic component of the strain ε:

$$\Delta E_c = n_{Ge}\Delta + \chi\varepsilon \ ,$$

$$\Delta = \Delta_d + (r_{SG} - r_s)\frac{dE_c}{dr} \ ,$$

$$\chi = 4r_s\frac{dE_c}{dr} \ .$$

(1)

Here we omit second-order terms like $\varepsilon(r_G - r_s)$.

Note that this analysis would be equally applicable to ternary systems such as GaAs-InAs (transmuting Ga into In), were it not for two factors. The calculations of Massidda, Min, and Freeman[5,6] suggest that in III-V alloys, the second-neighbor contribution to Δ_d is not negligible; and since the Born effective charge is not negligible in III-V and II-VI alloys, any lattice distortions give rise to long-ranged electrostatic fields which can shift the core levels. It is for these reasons that we restrict consideration here to Si-Ge systems.

The terms Δ_d and dE_c/dr in Eq. (1) can be calculated with standard electronic-structure methods. [Such methods could also be used to test the approximations underlying Eq. (1).] With these two parameters, ΔE_c can be calculated for any particular experimental conditions. Since such results are not yet available, we make a crude estimate. Deep core levels approximately track the local electrostatic potential near the nucleus. In the spirit of the "model solid" approach of Van de Walle and Martin,[11,12] we calculate this potential for the solid using a model charge density consisting of a superposition of spherical atomic densities. This gives $\Delta_d \sim -17$ meV and $dE_c/dr \sim 630$ meV/Å, so $\Delta \sim +13$ meV, and $\chi \sim +6$ eV. However, a more reliable calculation of these parameters would certainly be desirable. (The absolute core-level deformation potential χ could, in principle, be determined by combining the results of Schwartz et al.,[13] which implicitly contain the core-to-valence deformation potential, with other published calculations for the absolute valence-band deformation potential. However, to be useful for the present purpose, a complete and consistent calculation of both χ and Δ would be required.)

In the measurements described below, $\varepsilon \approx 0$, so the core-level shift for $Si_{1-x}Ge_x$ is $4x\Delta$. This is never more than about 50 meV by our estimate above; for typical device structures, with $x \lesssim 0.3$, the maximum shift is ~ 15 meV. Since these shifts are quite small, on the scale of the accuracy of current band lineup measurements, we henceforth treat the core level as strictly constant.

In addition to core-level shifts, we must bear in mind the conduction- and valence-band edges are also shifted and split by the tetragonal distortion in strained layers. We follow Van de Walle and Martin[11,14] in focusing on the *average* of the strain-split bands. By symmetry, only the hydrostatic component ε of the strain gives a net shift (the "absolute deformation potential"). This shift has been extensively discussed, and can be calculated, though it is typically rather small.[12,15,16] Because the strain configuration depends on details of the sample, it is best to determine what the band lineup would be in the absence of these effects, as we do here; the effects of strain in splitting and shifting the bands are then added only when discussing a particular sample.

We have applied this approach to Si-Ge systems, measuring the Si $2p_{3/2} \rightarrow$CBM energy for strain-free $Si_{1-x}Ge_x$ alloys from $x = 0.0$ to 0.95 in 5% steps. The alloys were grown epitaxially by stoichiometric coevaporation of Si and Ge from electron beam evaporators onto Si(100) substrates, in a Vacuum Generators V80 molecular-beam-epitaxy system.

A typical sample consisted of more than 100 alloy layers with thicknesses ranging from 50–4000 Å. The growth temperature started at 550 °C, and was reduced continuously as layers with higher Ge content were formed. Growth conditions were chosen such that strain resulting from the 4% lattice mismatch between Si and Ge was relieved by dislocations lying parallel to the plane of the interface, without dislocations threading to the growth surface.[17] This growth mode produced good epitaxial layers that were nominally strain-free, even in regions with high Ge content.[18]

Cross sections of the samples were prepared for transmission EELS by mechanical thinning, followed by low-temperature ion milling. Spatially resolved EELS spectra were obtained using a VG microscopes HB501 scanning transmission electron microscope, equipped with a field emission source and high-resolution electron-energy spectrometer employing photodiode array detection.[19] Spectra could be obtained with a lateral spatial resolution of less than 20 Å. Background scattering was stripped from the spectra using a power-law form.[20] Deconvolution using the *measured* instrumental response was used to obtain a Gaussian spectral resolution of 0.25 eV.[21] The $m = \frac{3}{2}$ spin component of the spectrum was extracted assuming a spin-orbit splitting of 0.608 eV and the expected statistical branching ratio of 0.5.

Spectra representative of the different alloy compositions are shown in Fig. 1. The abrupt increase in scattering, which for pure Si occurs near 99.85 eV, signals the Si $2p_{3/2} \rightarrow$ CBM threshold. The spectral shapes for the first few eV above threshold are dominated by contributions from the Δ_1 and L_1 conduction-band minima.[22] These are well separated in pure Si (99.85 and 100.65 eV in the b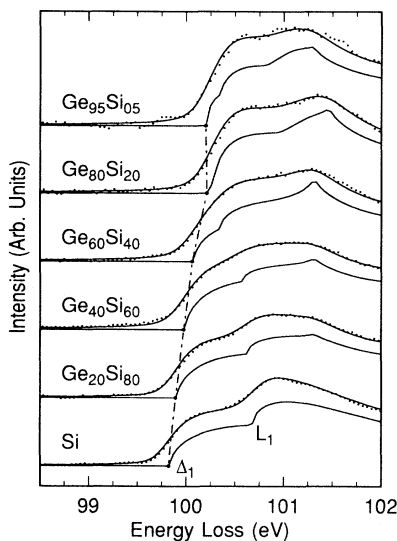ottom spectrum) but move closer in energy with increasing Ge content until they cross near 85% Ge, in agreement with photoluminescence measurements in bulk alloys.[23]

The spectral shape is well modeled using an effective-mass-like density of states (DOS). The fitted model DOS is shown as a solid curve under each spectrum. The solid line through the data points results from including the excitonic line shape and Gaussian spectral resolution,[22] and provides a good approximation to the data (dots).

The detailed evolution of the spectral shape with increasing Ge content has been described elsewhere.[22,24,25] Our concern here is to ascertain the position of the CBM. This is obtained directly as the edge of the fitted DOS. As an independent check, we note that for alloys with less than 30% Ge, where the Δ_1 and L_1 thresholds are separated by more than the 0.25 eV instrumental resolution, the threshold energy can be read directly from the point of maximum slope in the data. Agreement between these two determinations is excellent.

Figure 2 summarizes the results of the analysis presented above, for all samples. The bottom panel of the figure shows the Si $2p_{3/2}$ core level, which is assumed to remain constant, as discussed above. The circles in the upper portion of figure show the measured Si $2p_{3/2} \rightarrow$ CBM energy, and hence map out the energy of the CBM as a function of alloy content.

Since the alloy layers are virtually unstrained,[26] the energy of the valence-band maximum can be obtained by subtracting the bulk band-gap energy[23] for each alloy composition, thus obtaining the open triangles. The inferred valence-band maximum is highly linear with composition. This provides a powerful consistency check, indicating that both the conduction-band bowing and the break-in slope at 85% Ge have been accurately captured.

FIG. 1. Si $2p_{3/2}$ EELS spectra of bulk Si$_{1-x}$Ge$_x$ alloys. Dots are data. The solid line below the data is the fitted density of states for that alloy composition. The solid line through the data is obtained after including excitonic and instrumental line shapes. The dot-dashed line highlights the evolution of the threshold energy as a function of Ge content.
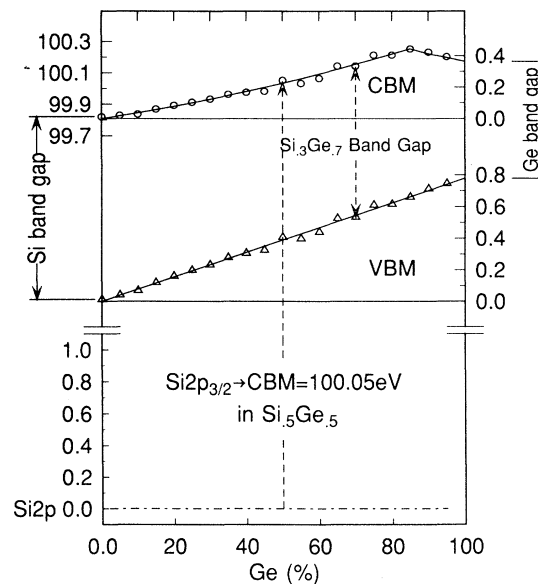

FIG. 2. Summary of threshold energies for all 20 alloy compositions. Open circles represent the Si $2p_{3/2} \rightarrow$ CBM energy. Open triangles were obtained by subtracting the band gap. The solid line through the triangles is a least-square linear fit; the curve through the circles is obtained by adding the band gap.

We estimate that the total error in each point in Fig. 2 is about 30 meV. However, much of this is not systematic, and is eliminated by the linear fit to the valence maximum. Hence we believe that the uncertainty is dominated by the small deviations from core-level constancy, as described in Eq. (1).

Results for band lineups between Si-Ge alloys of any two compositions can be read directly from Fig. 2. Splitting and shifting of the band edges due to strain must then be added, based on the specific strain configuration of the sample in question, as discussed above. An explicit prescription for this additional term has been given by Van de Walle.[12]

Our result for the valence-band discontinuity between Si and Ge (before including strain) is 0.78 eV. This is in fortuitously good agreement with "model solid" calculations, which give[12] 0.77 eV (after including spin-orbit splitting). *Ab initio* calculations (within the local-density approximation) by the same author give offsets around 0.07 eV smaller.[27] On the other hand, there is indirect evidence that such calculations systematically underestimate valence-band offsets slightly,[28] so even this small discrepancy with *ab initio* calculations is not significant. These results (after accounting for strain[14]) are consistent with band offsets measured in strained systems.[29,30]

In conclusion, we have demonstrated a new method for determining the band lineups of heterojunctions of Si, Ge, and their alloys. A convenient core level serves as a reference level, as proposed by Shih and Spicer.[2] The position of the conduction band is ascertained from a bulk core-to-conduction-band transition, using transmission EELS with high spatial resolution. This approach has numerous advantages. It is insensitive to interface quality, Fermi-level position, band bending, and surface reconstruction or contamination. The threshold has a simple shape, allowing very precise analysis. Finally, as we have shown by using multilayer samples, it is possible to study regions of small lateral extent (20 Å) within complex structures, e.g., real devices.

[1] The principal methods for measuring band lineups are discussed in *Heterojunction Band Discontinuities—Physics and Device Applications,* edited by G. Margaritondo and F. Capasso (North-Holland, Amsterdam, 1987).

[2] C. K. Shih and W. E. Spicer, Phys. Rev. Lett. **58**, 2594 (1987).

[3] J. Tersoff and C. G. Van de Walle, Phys. Rev. Lett. **59**, 946 (1987).

[4] S. Wei and A. Zunger, Phys. Rev. Lett. **59**, 144 (1987).

[5] S. Massidda, B. I. Min, and A. J. Freeman (unpublished).

[6] S. Massidda, B. I. Min, and A. J. Freeman, Phys. Rev. B **35**, 9871 (1987).

[7] J. Tersoff, *Heterojunction Band Discontinuities—Physics and Device Applications* (Ref. 1), p. 3.

[8] R. Resta, S. Baroni, and A. Baldereschi, Superlatt. Microstruct. **6**, 31 (1989).

[9] G. Srinivasan, Phys. Rev. **178**, 1244 (1969).

[10] S. de Gironcoli, P. Giannozzi, and S. Baroni, Phys. Rev. Lett. **66**, 2116 (1991).

[11] C. G. Van de Walle and R. M. Martin, Phys. Rev. B **35**, 8154 (1987).

[12] C. G. Van de Walle, Phys. Rev. B **39**, 1871 (1989).

[13] G. P. Schwartz, M. S. Hybertsen, J. Bevk, R. G. Nuzzo, J. P. Mannaerts, and G. J. Gualtieri, Phys. Rev. B **39**, 1235 (1989).

[14] C. G. Van de Walle and R. M. Martin, Phys. Rev. B **34**, 5621 (1986).

[15] M. Cardona and N. E. Christensen, Phys. Rev. B **37**, 1011 (1988).

[16] C. G. Van de Walle and R. M. Martin, Phys. Rev. Lett. **62**, 2028 (1989).

[17] F. K. LeGoues, B. S. Meyerson, and J. F. Morar, Phys. Rev. Lett. **66**, 2903 (1991).

[18] F. K. LeGoues, B. S. Meyerson, J. F. Morar, and P. D. Kirchner, J. Appl. Phys. **71**, 4230 (1992).

[19] P. E. Batson, Ultramicroscopy **28**, 32 (1989).

[20] R. F. Egerton, *Electron Energy Loss Spectroscopy in the Electron Microscope* (Plenum, New York, 1986), p. 255.

[21] P. E. Batson, D. W. Johnson, and J. C. H. Spence, Ultramicroscopy **41**, 137 (1992).

[22] P. E. Batson and J. Bruley, Phys. Rev. Lett. **67**, 350 (1991).

[23] J. Weber and M. I. Alonso, Phys. Rev. B **40**, 5683 (1989).

[24] P. E. Batson and J. F. Morar, Appl. Phys. Lett. **59**, 3285 (1991).

[25] J. F. Morar and P. E. Batson, J. Vac. Sci. Technol. B **10**, 2022 (1992).

[26] P. Mooney and J. Chu. Based on unpublished results from similar samples the 4000-Å layers (20%; 40%; 60%; 80%) are expected to be at least 90% relaxed.

[27] C. G. Van de Walle and R. M. Martin, J. Vac. Sci. Technol. B **4**, 1055 (1986).

[28] J. Tersoff, Phys. Rev. B **40**, 10 615 (1989).

[29] W. Ni and G. V. Hansson, Phys. Rev. B **42**, 3030 (1990).

[30] E. T. Yu, E. T. Cook, D. H. Chow, D. A. Collins, M. C. Phillips, T. C. McGill, J. O. McCaldin, and R. H. Miles, J. Vac. Sci. Technol B **8**, 908 (1990).