Measurement of the band gap by reflection electron energy loss spectroscopy

Maarten Vos a,*, Sean W. King b, Benjamin L. French c

a Electronic Materials Engineering Department, Research School of Physics and Engineering, The Australian National University, Canberra 0200, Australia
b Logic Technology Development, Intel Corporation, Hillsboro, OR 97124, USA
c Ocotillo Materials Laboratory, Intel Corporation, Chandler, AZ 85248, USA

ARTICLE INFO

Article history:
Received 1 May 2016
Received in revised form 7 August 2016
Accepted 11 August 2016
Available online 12 August 2016

Keywords:
Reflection electron energy loss spectroscopy
Band gap
Electron Rutherford backscattering

ABSTRACT

We investigate the possibilities of measuring the band gap of a variety of semiconductors and insulators by reflection electron energy loss spectroscopy without additional surface preparation. The band gap is a bulk property, whereas reflection energy loss spectroscopy is generally considered a surface sensitive technique. By changing the energy of the incoming electrons, the degree of surface sensitivity can be varied. Here, we present case studies to determine the optimum condition for the determination of the band gap. At very large incoming electron energies recoil effects interfere with the band gap determination, whereas at very low energies surface effects are obscuring the band gap without surface preparation. Using an incoming energy of 5 keV a reasonable estimate of the band gap is obtained in most cases.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

There is a huge range of semiconductor materials with different properties. Exploiting these properties makes it possible to produce e.g. LEDs emitting light over a large range of wavelengths or memristor devices. To obtain useful devices from the semiconductor, materials have to be part of more intricate structures. Hence, often surface layers with the desired properties are created by methods as, for example, atomic layer deposition (ALD) or ion implantation. One property that is crucial for the device performance is the band gap of the materials involved. Measurements of the band gap of a layer produced by a specific fabrication procedure is thus important. Often, when the layer is thin, and grown on a substrate, it is not straightforward to determine the band gap by optical means. Here (R)EELS ((R)eflection (E)lectron Energy Loss (S)pectroscopy) could be a useful alternative. However, this technique is surface sensitive, and in real device fabrication procedures exposure to air cannot always be avoided. The intensity at small energy losses will then be affected by the surface condition and this may obscure the onset of energy loss due to interband transitions (i.e. the band gap). Increasing the energy of the incoming beam reduces the surface sensitivity of REELS and it is then expected that the feature related to the band gap can be more easily identified.

Band gap measurements are routinely done in EELS experiments integrated in a TEM (Transmission Electron Microscopy) using energies around 100 keV [1], but complications due to the presence of Cherenkov radiation sets an upper limit to the energy one can use. This method also requires thin samples and thus extensive (and destructive) sample preparation. REELS at higher energies could be an alternative approach, avoiding the need for thin samples. One such example, that has been studied in one of our laboratories, is the development of a band gap after high-fluence O implantation in metals [2].

Photon based techniques (absorption or reflection measurements) can provide information of the optical band gap e.g. by measuring a Tauc plot [3] but probe generally much thicker layers. Photoemission experiments can also be used to obtain information about the band gap. Here the separation of a sharp core level with the onset of enhanced intensity of the background can be seen as a fingerprint of the band gap [see e.g. [4]]. Here the surface sensitivity can be decreased by increasing the photon energy. Alternatively, one can infer the band gap from the difference in position of the valence band edge, as observed in photoemission and the conduction band edge as seen by inverse photo emission (see e.g. [5]). This approach uses generally lower energies for the outgoing photoemission and incoming (inverse photoemission) electrons and well-defined surfaces are essential.

* Corresponding author.
E-mail address: maarten.vos@anu.edu.au (M. Vos).

http://dx.doi.org/10.1016/j.elspec.2016.08.001
0368-2048/© 2016 Elsevier B.V. All rights reserved.
One advantage of REELS is that it is not affected by changes in the Fermi level position near the surface ('band bending'). However, REELS depends on a large-angle deflections i.e. small impact parameter collisions with (screened) nuclei. For high energy electrons, such a deflection implies a significant momentum transfer to the scattering atom which causes, as we will see, a shift and broadening of the energy distribution of the scattered electrons. This shift increases with increasing incoming energy ($E_0$) and sets upper limits to the incoming energy one can use. Here, we investigate if it is possible to extract the band gap for a number of semiconductors that have been exposed to air, by choosing an energy that is high enough to see 'through' the surface layer but low enough that recoil effects do not interfere with the measurement. Avoiding the need for surface preparation is important from a practical point of view. Surface cleaning often requires spattering, which introduces defects [6]. Removal of these defects by annealing requires high temperatures, in many cases not compatible with the devices one wants to fabricate.

In this paper we study, using REELS, some bulk semiconductors with usually well-established band gaps to determine under which conditions one is most likely to obtain a reasonable estimate of the band gap.

2. Background

2.1. Shape of the onset of the loss spectrum

Ideally in a REELS spectrum of an insulator or wide-gap semiconductor the elastic peak is followed by a region of essentially zero intensity up to a certain energy loss after which there is additional intensity due to inelastic excitations. When we neglect defect states and excitonic excitations, the onset of the loss spectrum corresponds to the band gap. However, the onset is convoluted by the experimental resolution (and, as we will see, Doppler broadening). As statistics is limited, deconvolution of the experimental resolution is usually not an option and one has to rely on a model. The band gap value assigned to an energy loss spectrum depends then on shape of the loss function assumed. The most crude, but widely used way of interpreting such a spectrum is by approximating the leading edge by a straight line and assigning the band gap to the energy where this straight line intersects the energy axis. Drawing this line is not unique and the statistical error after this procedure is hard to trace.

For determining the band gap, one has to know the shape of the onset of the loss function. The mentioned straight-line approximation assumes that the loss function near the onset resembles $C/E - E_{gap}$, but this assumption is usually made implicitly. Similar straight-line approximations are made in the related problem of the measurement of band offset in semiconductor heterojunctions [7]. In the context of transmission EELS experiments, Rafferty and Brown [8] argued that the shape of the onset should be close to $(E - E_{gap})^{0.5}$ for direct gap semiconductors and $(E - E_{gap})^{1.5}$ for indirect gap semiconductors. This approach has been discussed in other TEM-related publications, noticeably in the context of an apparent sample thickness-dependence of the band gap and complications due to presence of Cherenkov radiation [1,9] and the effect of numerical treatment of the data on the band gap obtained [10].

Here, we will study semiconductors with well-established band gap values and compare the values extracted from the experiments when we assume that the intensity is proportional to $(E - E_{gap})$ (as is done implicitly in the straight line approximation) or either $(E - E_{gap})^{0.5}$ and $(E - E_{gap})^{1.5}$. For this approach to work, one needs a good understanding of the recoil effects that affect high-energy REELS measurements. This will be discussed next.

2.2. Recoil effects in high energy REELS measurements

In these REELS experiments the projectile electron (with momentum $k_0$) scatters over a large angle ($\theta_{scat}$) from a target atom. The transferred momentum $q = 2k_0 \sin(1/2 \theta_{scat})$ is absorbed by a single atom, and the energy transferred to this atom (mass $M$) can be calculated, assuming that this atom is free. This recoil energy $E_r$ depends then on the momentum $p$ of the target atom before the collision:

$$E_r = \frac{q^2}{2M} + \frac{q \cdot p}{M}$$

$E_r$ is thus equal to the energy transfer to a stationary atom ($q^2/2M_r$) plus a Doppler broadening term. The influence of multiple scattering on the outcome of the experiment is minor [11], and the magnitude of $q$ is determined by the incoming energy and the angle between the electron gun and the analyser $\theta_{scat}$, which is 135° in our case.

For isotropic targets (polycrystalline or powder samples) the width of the elastic peak $\sigma$ can be related to the mean kinetic energy $E_{kin}$ of atom i by:

$$\sigma = \sqrt{\frac{4E_i}{3E_{kin}}}$$

where $E_i$ is the mean recoil energy of element $i$: $q^2/2M_i$ [11]. For targets consisting of atoms of different mass the elastic peak will, for large enough $q$ values, split up into separate components.

For a scattering angle of 135°, $E_r$ corresponds to 1.58 eV for scattering from carbon at $E_0 = 10$ keV. Neglecting relativistic effects, $E_r$ is proportional to the incoming energy and inversely proportional to the atomic mass.

All detected electrons have backscattered from a target atom. Some have also created an electron–hole pair. The onset of the loss spectrum is thus strictly speaking not at $E_{gap}$ but at $E_{gap} + E_r$. If atoms of two different masses are present then the onset of the inelastic excitations depends on the atom that scattered the electron elastically and the onset for both atoms is at slightly different total energy loss values. Moreover the loss spectrum is also ‘broadened’ by the Doppler spread in the recoil energy.

In fitting these spectra one calculates first the shape of the elastic peak, based on the elastic scattering cross section and concentration of element $i$ and its mean kinetic energy. The shape of the onset of the inelastic spectrum is then proportional to the convolution of the elastic peak spectrum with $(E - E_{gap})$. Doppler broadening affects thus also the sharpness of the onset of the loss spectrum, and determines at high energies the ‘effective energy resolution’ that can be obtained.

There is an $E_0$ value for which the intensity of an electron scattered elastically from the low-Z constituent can have the same energy loss as an electron that scattered from the high-Z component and created an electron–hole pair. Then the band gap will be obscured by the elastic peak of the light element. This sets, thus, an upper limit to the incoming energy one can use.

3. Experimental details

Two electron spectrometers were used. For REELS at relatively low $E_0$ values of 500 eV, the Intel spectrometer was used, as described in Ref. [6] and for larger $E_0$ values the high energy spectrometer at the ANU as described in Ref. [12]. The low energy spectrometer has entrance and exit angles ($\theta_0$ and $\theta_1$) of 30° relative to the surface normal, corresponding to a scattering angle of 120°. In the high energy experiments the incoming beam was long.
the surface normal and the outgoing beam was 45° away from this direction, corresponding to a scattering angle of 135°.

Samples were obtained from different sources. The diamond film was a 1000 nm thick film deposited on Si (MTI corp.). The SiC (SiC-4H) and ZnO were single crystals with a <0001> surface normal (MTI corp.). The AlN and AlGaN samples were grown on 6H-SiC (0001) substrates by metal–organic chemical vapor deposition. The SiO2 layer was grown on a Si wafer by thermal oxidation and was 200 nm thick. The oxides from Ta, Ti and Nb were grown by thermal oxidation, and for these cases the experiments are described in more detail elsewhere [2,13,14].

The beam current varied from ≈1 nA at 500 eV and 5 keV to ≈10 nA at 40 keV. Spectra were typically acquired in a few hours, with the data acquisition rate smallest for the substrates consisting only of low-Z materials.

It is instructive to calculate, for $V$ - shaped trajectories, the depth that is probed. The probability $p_n(l)$ that a trajectory has $n$ - inelastic excitations depends on its length $l$ (see e.g. [15]) via:

$$p_n(l) = \left( \frac{\lambda}{l} \right)^n e^{-\lambda/l},$$

with $\lambda$ the inelastic mean free path. For the elastic peak, $n = 0$ and for the onset of the loss spectrum $n = 1$. For $V$ - shaped trajectories, there is a simple relation between $l$ and the depth of backscattering: $d\left(l = d\cos\theta_0 + d\cos\theta_1\right)$. The depth $d$ probed is sketched in Fig. 1 for some energies used in this paper and a mean free path as calculated for AlN using the TPP formalism [16]. Note that the accuracy of the $V$ - shaped trajectory approximation is better for high $E_l$ values and low Z targets. Moreover, especially at low energies, surface plasmon excitations complicate the picture. Nevertheless, Fig. 1 is a useful first-order estimate of the depth probed.

4. Results

4.1. Aluminum nitride

We discuss the case of AlN at some length to explain how the obtained spectra are interpreted. For the other materials, we summarise the results more briefly later. AlN REELS spectra for different incoming energies are displayed in Fig. 2. When plotting a REELS spectrum, where recoil effects are important, one can choose the energy scale in two different ways. If one is interested in electronic excitations only it is best to align the zero of the energy scale with the (average position of) the main elastic peak and in that case we label the energy axis as ‘Electronic Energy Loss’, as was done here. If one wants to understand the elastic peak shape it is better to align with the true zero of the energy scale and when that is done the axis is labelled ‘Total Energy Loss’.

For the 20 and 40 keV spectra, the AlN elastic peak shows a clear structure. At these energies, the recoil losses for scattering from Al and N differ enough to result in two components, separated by 1.3 eV (20 keV) and 2.6 eV (40 keV). Even the 5 keV elastic peak shows a small asymmetry as a consequence of the recoil loss. At larger energy loss values, the intensity increases again. This is attributed to electrons that have scattered elastically from a nucleus and have created an electron–hole pair. The separation of the main elastic peak and the onset of the energy loss spectrum corresponds thus to the band gap.
At 500 eV there is a strong asymmetry: the intensity a few eV away from the elastic peak is much larger at positive losses than at negative losses. Clearly the assumption that no electronic excitation occur within the band gap is not correct, probably due to the surface sensitivity of the REELS experiment at these energies, and this intensity makes it hard to identify the band gap itself. This asymmetry decreases strongly but does disappear completely for \( E_0 = 5 \) keV and higher.

Indeed at 5 keV and 20 keV the intensity in the gap region is relatively constant up to 6 eV and then increases gradually. It is thus possible to fit the elastic peak and the onset of the loss spectrum simultaneously. This is shown in more detail in Fig. 3. First the shape of the elastic peak is calculated, based on the recoil losses, intrinsic peak width (Eq. (2)) and known energy resolution (0.35 eV FWHM). The onset is modeled as the convolution of the elastic peak shape and \( c(E - E_{\text{gap}})^x \) with \( c \) a constant relating the elastic peak and inelastic loss intensity. Satisfactory fits can be obtained for \( x = 0.5 \), \( x = 1 \), or \( x = 1.5 \), but the derived values for \( E_{\text{gap}} \) are quite different (see Table 1), with the value obtained for \( x = 1 \) quite close to the established band gap of AlN.

At \( E_0 = 40 \) keV, the N part of the elastic peak starts to overlap with the onset of the inelastic losses, and it is thus not possible to obtain a band gap value for this measurement. This measurement can, however, provide valuable information about the surface composition. This is illustrated in Fig. 4. Fitting the spectra, assuming only the presence of Al and N, gives a reasonable but not perfect fit. The observed N peak is at slightly too small energy loss. There are two possible explanations for this:

- the sample is charging and the effective energy of the electrons scattering from the surface is somewhat smaller. A very good fit is obtained if one assumes \( E_0 = 38.5 \) keV.
- the AlN sample is covered with a \( \text{Al}_2\text{O}_3 \) layer at the surface. This was suggested e.g. in Ref. [17]. Indeed a good fit can be obtained, for both the \( E_0 = 40 \) keV and 20 keV measurement assuming a 2 nm thick \( \text{Al}_2\text{O}_3 \) at the surface, using the analysis method of Ref. [18]. TEM work on AlN showed slightly less than 2 nm thick \( \text{Al}_2\text{O}_3 \) layer after exposure to air for a week [17]. Thus we consider the presence of a 2 nm thick \( \text{Al}_2\text{O}_3 \) layer the most likely explanation for the small deviation in the N peak position from theory.

Such a layer is considerably thinner than the values obtained in Ref. [19] after annealing at 200 °C (5–60 nm \( \text{Al}_2\text{O}_3 \)). To see if one can still discern the band gap of AlN for thicker oxide overlayers, we annealed the sample at 200 °C for 10 min in a O₂ atmosphere. Only a very marginal increase of the \( \text{Al}_2\text{O}_3 \) thickness was found (to 3.5 nm). Note that this is a favorable case as the band gap of \( \text{Al}_2\text{O}_3 \) is \( \approx 8.8 \) eV [20], much larger than the gap of AlN and its loss spectrum will not interfere with the onset of the AlN loss intensity.

Assuming a 2 nm thickness for the \( \text{Al}_2\text{O}_3 \) overlayer, it is clear why the 500 eV REELS measurements do not give a good indication of the AlN band gap. The largest part of the intensity in this case originates then from the \( \text{Al}_2\text{O}_3 \) layer and not from AlN. Indeed the slope of the 500 eV loss spectrum becomes much steeper for losses over 8 eV.

### 4.2. Other wide-gap semiconductors

The results for some other semiconductors are summarized in Fig. 5. Both 500 eV measurements and 5 keV measurements are shown in the same plot, with the elastic peak normalized to equal
area. The 500 eV measurement showed a clear onset of the loss spectra for the oxides but there is clearly non-zero intensity inside the gap. For \( E_0 = 5 \text{ keV} \) measurement, the onset was more pronounced, except for the case of AlGaN where we could not pinpoint the band gap with any level of confidence. The band gap derived for all these semiconductors are reproduced in Table 1 for different exponents \( x \). When a band gap value could be derived from the low energy measurements, then this one is given in this table as well, based on the straight-line extrapolation method. The values of the band gap obtained for \( x = 1.5 \) is always too small. \( x = 1 \) seems to give a good description overall, in-line with the straight line extrapolation method. The onset is also at 5 keV most pronounced for the oxides and the result of the gap depends in that case not as strong on the exponent than for the other cases. For \( x = 1 \) the deviation between the obtained band gap and the literature value is 0.2 eV or less.

For ZnS the 500 eV measurement showed a minimum in intensity, rather than a plateau and an onset. Using the same spectrometer, data taken at 2.5 keV showed a more pronounced minimum (Fig. 6 upper panel). At 5 keV, data taken using the high-energy spectrometer show even less intensity in the gap region and now a plateau and onset can be distinguished. For this measurement the a band gap close to the literature value was obtained using \( x = 0.5 \), the value extracted using \( x = 1 \) was 0.5 eV too small. As all atoms in ZnS are relatively heavy, it was hoped that we could extract a band gap here for incoming energy as high as 40 keV, as the recoil energy losses for S are small. However there was a very small peak visible at both \( E_0 = 20 \text{ keV} \) and \( E_0 = 40 \) at the recoil energy of O atoms. Thus, again a surface oxide seems to be present. These impurities interfered with the determination of the onset of the loss spectrum, as is illustrated in the lower panel of Fig. 6. To help identify the O impurity contribution we plot in the central panel of Fig. 6 the elastic peak of ZnS and ZnO normalised to equal area of the Zn peak. In the lower panel, we assume that 4% of the intensity of Zn in ZnS is due to a Zn present in a surface ZnO layer and subtract this amount. Now the loss spectrum resembles the 5 keV result. A possible explanation of the low band gap value obtained at 5 keV for ZnS using \( x = 1 \) could be the presence of this thin ZnO layer at the surface.

4.3. Transition metal oxides

At 5 keV we have studied transition metal oxides in the past [2,13,14,30] and the fully oxidized states of these metals are also examples of wide-band semiconductors. Here we focus on the band gap for these cases. The 5 keV data and fits with \( x = 1 \) are given in Fig. 7. With increasing atomic number the oxygen contribution to the elastic peak decreases, and the elastic peak becomes sharper. All cases show a clear onset of the loss spectrum, and the band gap obtained for various values of \( x \) are given in Table 1 as well. For
Ta2O5 and Nb2O5 we show as well the result from the 500 eV measurement. A clear onset is seen at 500 eV as well, but the observed intensity at this incoming energy in the gap region is again much higher than at 5 keV.

5. Discussion and conclusion

It is possible to get a meaningful estimate of the band gap for many semiconductors by REELS, without surface preparation, if one use a relatively high incoming energy of several keV. Below 1 keV, the observed intensity inside the gap is far from negligible (presumably due to surface effects) and this causes a washing out of the onset, making it often impossible to derive the band gap from the experiment. An exception appears to be oxides where the onset is often quite sharp, even at lower incoming energies.

At even larger incoming energies (20–40 keV) the recoil shift of the elastic peak becomes large enough to interfere with the determination of the onset of the loss part of the spectrum. When all atoms in the semiconductor are high-Z atoms larger values of $E_0$ could be used in principle, but the elastic peak of even a small concentration of C and/or O impurities at the surface interferes with determining the onset. Measurements at these high energies are more useful to check the actual surface composition (using the recoil shifts of the elastic peak) than determining the band gap itself. Also at larger incoming energies the increased Doppler broadening reduces the sharpness of the onset.

For the case of TEM-based EELS, it was discussed if the onset of the loss spectrum should be proportional to $(E - E_{gap})^x$ with $x = 0.5$ for direct gap semiconductors and $x = 1.5$ for indirect gap semiconductors. For the reflection EELS case we find generally the best agreement with the well-established literature values of the band gap using $x = 1$, but a rigorous justification of this exponent is not obvious. Indeed for ZnO and ZnS slightly better agreement with the literature values was found using $x = 0.5$.

For narrow-band semiconductors, the REELS method, as described here, is less useful for the determination of the band gap as the tail of the very intense elastic peak then interferes with the onset of the loss spectrum. In practice band gaps of 2.5–3 eV are the lower limit for measurements as described here. A monochromator, reducing the spread in energy of the incoming beam, is required for semiconductors with a substantially smaller gap as described recently in the context of TEM-EELS [1,9], but in the REELS case improved energy resolution may dictate the use of lower incoming energies (i.e. smaller momentum transfer) to limit the Doppler broadening. Note that in the TEM-EELS case, where the measurement is done in the forward direction (at very small momentum transfer) Doppler broadening is negligible small.

For $E_0 = 5$ keV, the inelastic mean free path is of the order of 8–10 nm for most semiconductors. Thus, the minimum thickness of a semiconductor film for which the band gap can be measured is of this magnitude. Such a thickness is of interest e.g. in the area of memristors fabricated from transition metal oxides.

An overall comparison of the obtained gap values with the ‘established gap values’ is given in Fig. 8. Picking the ‘established gap values’ is often not 100% straightforward as different values are used in the literature, often without citation of the source. Part of these discrepancies are related to slightly different definition of the band gap. In optical measurements the band gap is taken to be the lowest energy where photon absorption takes place. In those cases where the created electron and hole interact to form an exciton the obtained ‘optical band gap’ differs from the band gap defined as the separation of the valence and conduction band. An example here is the extremely well-studied case of rutile (TiO2) where the value of 3.0 eV is often used, based on optical measurements (e.g. [31]) about half a volt smaller than the separation of the valence and conduction band as obtained from photoemission combined with inverse photoemission, see Ref. [32,33]. Judging from our TiO2 results, it appears that the latter definition is more applicable to the REELS case.

In summary, it is often possible to obtain meaningful estimates of the band gap without surface preparation if one uses incoming energies of $\approx 5$ keV. For lower energies the technique becomes too surface sensitive. At much higher energies the elastic peak of the lighter component and/or impurities interferes with the onset of the loss spectrum, and increased Doppler broadening limits the ‘effective energy resolution’ that can be obtained.

Acknowledgement

This work was made possible by a grant of the Australian Research Council.
References