SAND2019-14030J

Uncertainty in Linewidth Quantification of Overlapping Raman Bands

Christopher B. Saltonstall,[1,2,a)] Thomas E. Beechem,[2] Jatin Amatya,[3] Jerrold Floro,[3] Pamela M. Norris,[1] and Patrick E. Hopkins[1]

[1)]*Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, Virginia 22904, USA*

[2)]*Sandia National Laboratories, P.O. Box 5800, Albuquerque, New Mexico 87185, USA*

[3)]*Materials Science Department, University of Virginia, Charlottesville, Virginia 22904, USA*

(Dated: 19 September 2018)

Spectral linewidths are used to assess a variety of physical properties, even as spectral overlap makes quantitative extraction difficult owing to uncertainty. Uncertainty, in turn, can be minimized with the choice of appropriate experimental conditions used in spectral collection. In response, we assess the experimental factors dictating uncertainty in the quantification of linewidth from a Raman experiment highlighting the comparative influence of: 1) spectral resolution, 2) signal to noise and 3) relative peak intensity of the overlapping peaks. Practically, Raman spectra of SiGe thin films were obtained experimentally and simulated virtually under a variety of conditions. Relative peak intensity is found to be the most impactful parameter in specifying linewidth followed by the spectral resolution and signal to noise. While developed for Raman experiments, the results are generally applicable to spectroscopic linewidth studies illuminating the experimental trade-offs inherent in quantification.

Keywords: Raman, linewidth, lifetime, uncertainty, scattering rate, phonon, resolution, noise, spectroscopy

I. INTRODUCTION

Spectroscopic techniques (Raman,[1–3] XRD,[4–6] EELS,[7] FTIR,[8] XPS,[9] NMR[10]) utilize the linewidth of spectral peaks as the metric for a wide variety of physical properties: phonon lifetimes,[2] crystallinity,[4–6] electron spin lifetimes,[10,11] plasmon lifetimes[7]. In many cases the peak of interest may overlap with other spectroscopic features. Such a convolution makes linewidth extraction difficult even when a model function (i.e., the sum of two peaks) fits the total response quite well. Simply stated, fitting additional peaks adds significant uncertainty to the quantification of the linewidth.[12,13] Experimental conditions can be chosen to minimize the emergence of this uncertainty, however. To that end, this work examines the impact of major experimental parameters (spectral resolution, noise and spatial resolution) on linewidth quantification in a Raman experiment.

Spatial resolution defines the ability to selectively collect signal from the region of interest while rejecting signal from other sources. Due to the finite penetration depth and spot size of the focused laser, a Raman signal will be generated and collected from a volume of the sample. If the material of interest is smaller than this collection volume, an overlapping signal may be generated from adjacent materials. Depending on the sample geometry–e.g., quantum wells[14] or nanoscale thin films[15]–the spatial resolution may be limited by either lateral or depth resolution. For uniform thin films where lateral resolution is immaterial, depth resolution dictates the ability to differentiate between film and substrate. This differentiation can be quantified by the relative peak intensity (RPI) between the substrate and film, which is dependent upon the Raman cross section, thickness and optical properties of the materials. Practically, the RPI is controllable through selection of the laser wavelength.[16] Impacting the intensity ratio of the overlapping peaks, wavelength selection will therefore impact linewidth uncertainty of the Raman measurement.

Spectral resolution quantifies the ability to discern two peaks of similar energy. When the spectral resolution is poor, there are few data points mapping out a peak leading to artificial broadening and blurring with nearby spectral features. Blurring directly limits accurate extraction of the linewidth. Additionally, poor resolution reduces the number of points mapping out the peak. Less data leads to uncertainties in the least squares algorithms quantifying the linewidth.

Spectral resolution in Raman spectroscopy is controlled through selection of both the laser wavelength and diffraction grating condition. The diffraction grating groove density and order of operation determines the wavenumber range that falls on each pixel of the detecting CCD camera (cm^{-1}/px) and thus the spectral resolution. However, changing the laser wavelength also influences spectral resolution since diffraction gratings separate light at roughly a constant wavelength per pixel, while a Raman signal is shifted in energy (i.e., wavenumber). This complicates optimizing experimental conditions since reducing the laser wavelength to improve spatial resolution will lead to a diminished spectral resolution.

In an ideal experiment possessing no noise, it would be possible to fit two peak spectra with nearly any realistic spatial and spectral resolution. However, noise distorts the peak shapes and in turn increases uncertainties of the extracted linewidths. Signal to noise is controlled by both the laser power and the integration time (similar to adjusting the f/stop and exposure time on a photographic camera). The effects of noise are exacerbated in poor spatial and spectral resolution conditions. One must therefore understand how spatial and spectral resolution influence linewidth uncertainties in conjunction with noise.

Together, these experimental parameters–spatial and spec-

[a)]Electronic mail: cbsalto@sandia.gov

tral resolution along with noise to signal ratio (N/S)–dictate uncertainty in a linewidth measurement with any degree of spectral overlap. Their impact is convoluted, thus obscuring their comparative influence and any heuristic by which to arrive at optimized experimental parameters. In response, uncertainty in the linewidth of SiGe films possessing spectral overlap with the underlying silicon was measured while varying spatial and spectral resolution. Interpretation of the trends and deconvolution of the relative impact of each parameter was then accomplished via the use of a computational "virtual Raman" experiment in which individual parameters were adjusted independently. We find that uncertainty is primarily dictated by spectral overlap defined by spatial resolution and only secondarily by the spectral resolution and lastly by noise. Leveraging these results, we have developed a means to link experimental conditions with uncertainty in the quantification of linewidth and any parameters derived from it.

II. EXPERIMENT

In this study, two "real" experiments and three "virtual" experiments (where the Raman signal is numerically simulated) were performed under a variety of experimental conditions. The first experiment measured the single-peak silicon spectrum under varying noise levels and spectral resolution. These results were then used as inputs to develop a noise model of the experiment, that was used to construct the virtual Raman experiment. Subsequently, silicon germanium alloy (SiGe) thin films of varying thicknesses on a Si substrate were measured under the same variable experimental conditions used in the Si investigation. Importantly, the SiGe Raman spectrum convoluted with the spectrum from the underlying Si substrate results in two overlapping spectral features. The virtual Raman experiment was then tested against the SiGe results from the real Raman experiment to ensure that linewidth and linewidth uncertainties were accurately modeled. Finally, the virtual Raman experiment was used to investigate the comparative influence of experimental conditions, e.g., spectral resolution, spatial resolution and noise. Ultimately, these results were used to identify a heuristic by which to select the ideal experimental conditions for minimizing uncertainty in linewidth quantification when investigating two overlapping peaks.

The real spectra of Si and SiGe thin films were collected using a commercially built Renishaw InVia Raman spectrometer. The system is equipped with three laser wavelengths (405, 488 and 514 nm) and two diffraction gratings (1800 and 3000 g/mm). By varying the combination of laser wavelength, grating condition and sample thickness, we investigated the influence of spatial resolution, spectral resolution and noise.

Spectral resolution is quantified by the energy range collected by each pixel in units of wavenumber per pixel (cm^{-1}/px). High spectral resolution values therefore respresent poor resolution and vice versa. The spectral resolution at the center of the CCD camera chip for each laser and grating combination is shown in Table I, which was quantified using the spectra of Ne and Ar calibration lamps. In order to extend the spectral resolutions available, each grating was operated at two diffraction orders indicated by the number in parentheses. Laser wavelengths were varied at each grating condition to investigate the effects of spatial resolution which was quantified by the RPI. To further extend the range of examined RPI's, SiGe films of three thicknesses were assessed.

TABLE I. Spectral resolution (cm^{-1}/px) for various laser wavelength and grating combinations. The number in parenthesis represents the grating order. The second column presents the laser penetration depths, δ_p, in $Si_{75}Ge_{25}$.

Wavelength (nm)	δ_p (nm)	Grating (g/mm) 1800 (+1)	3000 (+1)	3000 (-1)	1800 (-2)
405	59	–	1.84	1.16	0.71
488	297	2.24	1.12	0.56	–
514	376	1.99	0.95	0.40	–

$Si_{75}Ge_{25}$ films with various thicknesses (d = 40, 90 and 210 nm) were grown epitaxially on 5 cm diameter, 250 μm thick, undoped Si(001) wafers. The epitaxial alloy films were synthesized in an ultra-high vacuum, hyperthermal molecular beam epitaxy (MBE) system. The MBE system has a base pressure of 1x10^{-10} Torr and is equipped with variable-distance magnetron sputter guns with Ge and Si as target sources. Prior to insertion into the MBE system, Si(001) substrates were chemically cleaned using a modified IMEC/Shiracki process,[17] creating a passive SiO_x layer in the final step. The passive SiO_x layer was desorbed *in situ* at 800 °C after an overnight temperature ramp plus 5 hour prebake at 600 °C. The substrate was then cooled to 720 °C for the deposition of a 50 nm Si buffer layer. $Si_{75}Ge_{25}$ films were grown at 400 °C with the chamber pressure maintained at 5 mTorr of getter- and LN_2-purified Ar throughout the deposition. All samples were 0.5 cm square sections taken from the center of the growth wafer. Silicon wafers were commercially grown unintentionally doped in the (111) orientation.

The thickness of the alloy film, d, was characterized *ex situ* by X-ray reflectivity (XRR). The X-ray diffraction (XRD) instrument used was a Rigaku SmartLab X-ray diffractometer[18] with a 3kW copper sealed tube generator. All the measurements were made at a wavelength of λ = 1.5406Å obtained using a parallel beam Ge 220×2-bounce monochromator. Germanium composition, x, and the strain relaxation, r, in the films were determined by measuring the 004 and 224 reflections from the sample via conventional "$\theta - 2\theta$" and rocking curve (ω) scans in XRD.[19,20] Relaxation was found to be negligible and alloy concentrations were nominal.

Twenty five Raman spectra were collected for Si and each SiGe sample using the laser and grating combination shown in Table I. The laser power for each combination was selected to prevent heating as monitored by the peak position. The integration time was chosen in each instance to maintain a constant noise to signal ratio (N/S) of 0.0025. See Ref. 15 for more details on the procedure for selecting the laser power, integration time and quantifying N/S. The natural/Lorentzian linewidth of the SiGe peak was extracted from the experimental spectra from a fit to the summation of two Voigt functions. Here we define the Lorentzian and Gaussian linewidths as the

full width at half maximum (FWHM). The linewidth uncertainties for each sample, grating and laser combination were calculated through the linewidth standard deviation of the 25 spectra. Similarly, the virtual experiment uses a Monte Carlo approach in which 50 numerically generated spectra are fit and uncertainty deduced from the standard deviation of fit results.

Alternatively, uncertainty can also be quantified using an analytical approach employing the Jacobian of the model function and the residuals of the fit. Being analytical, this method is much less computationally expensive than Monte Carlo approaches and is capable of accounting for the uncertainty implicit with poor fitting. However, it assumes that the model function is linear between sampling points, which is not the case for Raman spectra collected with poor resolution. Therefore, while more computationally expensive, the Monte Carlo approach is utilized here as it is more accurate when the model accurately represents the data. The two approaches can be directly compared with the code supplied within the Supplementary Information.

III. RESULTS AND DISCUSSION

A Raman spectrum is determined by the experimental conditions under which it was collected. Experimental conditions influence how pixelated a spectrum appears, the relative intensity of the spectral peaks, the noise, among a variety of other effects. When quantifying spectral linewidths, these characteristics directly impact uncertainty. This concept is demonstrated for the measured spectra of SiGe thin films on a silicon substrate in Fig. 1, where the obvious changes in the spectral shape are apparent for different experimental conditions.

Figure 1 a) plots the spectra of a 210 nm film acquired with the 488 nm laser under three different grating conditions to investigate the effects of spectral resolution. As the spectral resolution worsens, the peaks become pixelated–fewer pixels map out the peak–and blur together. This decreased spectral resolution causes the uncertainty of the extracted linewidth to increase linearly, as shown in Fig. 1 b). Poor spectral resolution can also lead to errors in the measurement beyond statistical uncertainties due to reduced accuracy of the fitting algorithm (see Section III D). Additionally, as the spectral resolution increases, the N/S increases for a constant integration time (not shown) since the Raman signal is spread over more pixels. The N/S is further increased when operating gratings in higher diffraction orders since they are designed to be most efficient in the +1-order. This change in N/S can lead to a compounded change in linewidth uncertainty if not compensated for by an increased integration time.

Figure 1 c) plots the spectra of the 40 nm film using the 3000 g/mm (-1) grating employing various laser wavelengths. As the wavelength decreases, the Si substrate signal ($\sim$520 cm^{-1}) diminishes relative to the SiGe film signal ($\sim$512 cm^{-1}) signal, i.e., the RPI decreases as the wavelength decreases. This increase in film signal results in a corresponding decrease in linewidth uncertainty as shown in Fig. 1 d). Decreased uncertainty with reduced wavelength occurs in spite of the worsening spectral resolution that is implicit with decreasing wavelength.

The results of Fig. 1 indicate that the uncertainty in the linewidth emerges from a convolution of N/S, spatial, and spectral resolution. Thus, it is difficult to optimize a linewidth study *a priori*. Linewidth uncertainties can only be tested under the available conditions of one's particular experiment. However, these conditions are often very limited without the purchase of a many lasers and gratings. In order to optimize uncertainties, a more continuous range of parameters needs to be tested. Here, a virtual experiment is employed to these ends to simulate the effects of spectral resolution, spatial resolution and noise on a Raman experiment.

A. The Virtual Experiment

A virtual experiment is performed by numerically creating spectra analogous to those measured, which can then be fit to quantify error and uncertainty. Unlike an actual experiment, each factor influencing the response is represented mathematically allowing separate effects to be independently controlled and thus assessed. Here, the single-peak spectrum of silicon is first examined to both quantify the underlying noise sources of the measurement while at the same time establishing the approach's applicability. Subsequently, the two-peak spectrum of SiGe is investigated to link how experimental conditions affect the uncertainty of linewidth quantification for overlapping features.

Practically, the measured Raman signal is composed of a discrete series of numeric pairs corresponding to the energy of the Raman shift and the intensity of the signal at this energy. Creation of a virtual Raman spectrum is therefore an exercise in creating these energy-intensity pairs. The underlying signal is represented by a Lorentzian energy distribution representing the natural linewidth of a crystalline solid. With measurement, the spectrometer itself artificially broadens the signal. The Gaussian nature of this so-called "spectrometer response function" is therefore convoluted with the underlying Lorentzian signal resulting in the measured response having a Voigt profile. Approaches for measuring the spectrometer response function and the Voigt profile have been previously established.[21]

The Voigt profile, $V(\nu)$, is the mathematical convolution of two continuous functions and so is itself continuous. The Raman measurement is not, as it derives from the output of a discrete set of pixels that each bin over a finite range of wavenumbers. Pixelation of the continuous signal must therefore occur in the virtual experiment. Mathematically, this is accomplished by assigning a central wavenumber (ν_i) and spectral width ($\delta\nu_i$) for each pixel and then integrating the continuous Voigt profile to arrive at a pixel-specific intensity using,

$$V_{Pixel}(\nu_i) = \int_{\nu_i-\delta\nu_i/2}^{\nu_i+\delta\nu_i/2} V(\nu)d\nu, \qquad (1)$$

where both ν_i and $\delta\nu_i$ are determined by the spectral resolution of the system. Note that Eq. 1 is not simply a discrete sample of $V(\nu_i)$ but instead an integration of V over

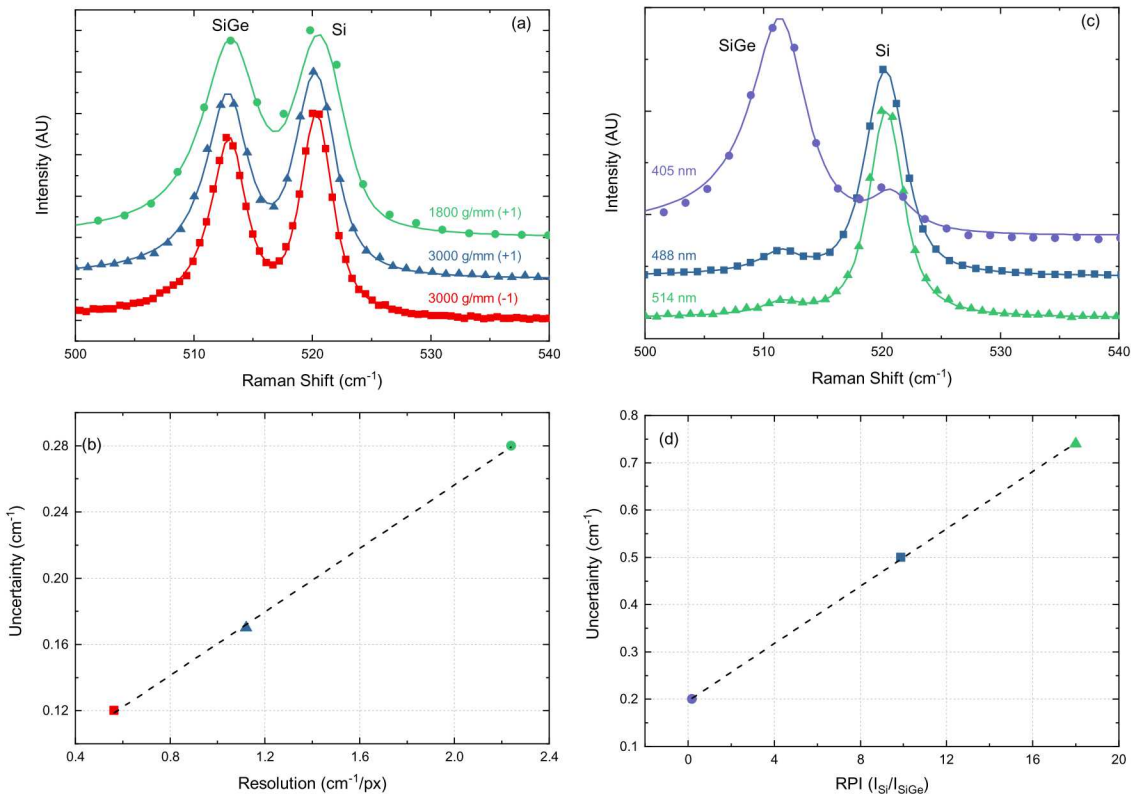

FIG. 1. (a) Experimental Raman spectra (dots) of 210 nm SiGe on Si using the 488 nm laser with various grating conditions with fits (solid lines) showing picelation of the spectra with worsening resolution. (b) The corresponding linewidth uncertainty caused by the worsening spectral resolution. (c) Experimental Raman spectra of 40 nm SiGe on Si using various laser wavelengths and the 3000 g/mm (-1) grating showing how spectra and spatial resolution change with wavelength. (d) Corresponding linewidth uncertainty as a function of inverse RPI showing how uncertainty grows as the Si substrate peak grows in relative intensity.

a finite range. This impacts the lineshape of the resulting spectral curve and thus quantifications of the linewidth. Additionally, the location of the discrete pixels (ν_i) relative to the maximum of V also influences the fitted lineshape and hence linewidth measurements. The resulting wavenumber-intensity pairs from Eq. 1 represent the virtual Raman spectrum presuming no noise.

Some degree of noise is present for all measurements. To model noise in the virtual experiment, all sources are lumped into two categories, namely that arising from the background and shot noise. Background noise originates from light leakage into the spectrometer and any thermal noise within the detector itself. It is presumed uniform across the analyzed spectral range and was therefore modeled as a random sampling from a Gaussian distribution with a magnitude that is proportional to the signal based on a stipulated N/S. Shot noise, on the other hand, originates from the probabilistic nature of the light-matter interaction as well as variations in the detector's quantum efficiency. Since shot noise arises from signal collection, it is proportional to the signal intensity. Mathematically, noise from these sources is quantified via

$$V_{Sim}(\nu_i) = V_{Pixel}(\nu_i) + V_{Pixel}(\nu_i)N_S P_1 + N_B P_2, \quad (2)$$

where N_S and N_B are the the magnitude of the shot and background noise, respectively, and P_j is a random number sampled from a normal distribution possessing a standard deviation of one. Controlled by the integration time of the experiment, N_B is the N/S ratio of 0.0025. The shot noise was determined to be 0.018 by examining the fit residuals of single-peak silicon spectra. Due to the random nature of noise, multiple virtual spectra were generated under each experimental condition by sampling from the random distribution defining P_j. In this manner, an ensemble of virtual experiment spectra were collected and then fit using a Voigt function such that linewidth uncertainties could be compared to those directly obtained from experiment.

The Voigt function is the mathematical convolution of a Gaussian and Lorentzian function and so the total linewidth, Γ, is an indirect convolution of the Gaussian, Γ_G, and

Lorentzian, Γ_L, linewidths.[21] Therefore, in order to specify the Voigt function, both linewidths must be known. While the latter is material dependent, the former is tied to the state of the experimental equipment. For this reason, the Gaussian linewidth will vary with the experimental resolution while the Lorentzian component remains constant. To demonstrate, the Raman response of silicon was acquired at several spectral resolutions and then fit to a Voigt curve to deduce the Gaussian and Lorentzian widths. Figure 2 presents the resulting linewidths in which the the Lorentzian component remains constant, as expected, whereas the Gaussian width increases with resolution. The change in Gaussian linewidth is proportional to the image of the beam spot on the CCD camera and so can be modeled by,

$$\Gamma_G = R\Gamma_S + \Gamma_0, \tag{3}$$

where R is the spectral resolution in wavenumbers per pixel (cm^{-1}/px), Γ_S=1.25 px is the laser spot size on the detecting CCD camera and Γ_0 is the y-intercept.

The entirety of this analysis rests on the natural lineshape being Lorentzian. However, in some cases, this presumption is not valid such as is the case for highly doped semiconductors,[22] heavily disordered[23–25] or nanoscaled materials.[26–28] In such a cases, a Voigt function cannot be employed but rather a convolution of the appropriate natural lineshape with the spectrometer induced Gaussian must be used.

B. One Peak - Silicon

Using the input parameters for the noise and spectral broadening of the Raman spectrum described in the previous section, 50 silicon spectra were generated and fit for a range of spectral resolutions and noise levels. The resolution was varied from 0.01 to 2.5 cm^{-1}/px and the noise ranged from 0.01 to 2 times that of N_S and N_B. Note, a noise fraction of one is equal to the experimental values. Uncertainty was calculated from the standard deviation of measured linewidth. The results are presented in Fig. 3. Uncertainties derived for the virtual experiment are comparable to that measured experimentally as seen in Fig. 3 c). For this reason, we conclude that the virtual approach can therefore be employed to gain insight into the individual sources of uncertainty.

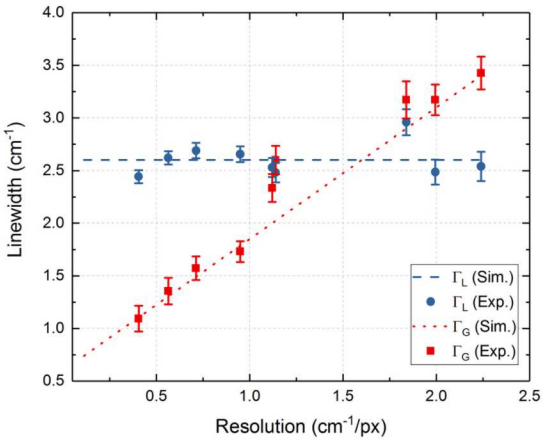

FIG. 2. Experimental (markers) and simulated (lines) corresponding to the natural, Γ_L, and instrument, Γ_G, linewidths of the Si Raman spectrum for various spectral resolution. The spectral resolution was varied by changing the grating and laser used, see Table I. Both the natural and instrument linewidth were extracted from Si spectra with a 4-parameter (peak intensity, peak position, Γ_L and Γ_G) Voigt fit.

As the noise and resolution worsen, the uncertainty in the linewidth increases. The contours have a roughly circular shape indicating the resolution and noise fraction are roughly of equal importance for single peak linewidth measurements. Consequently, doubling the noise has a similar effect as doubling the resolution as seen more directly in the cross section of the surface plot, Figs. 3 b) and c). However, spectral resolution can cause errors beyond uncertainties as will be addressed in Section III D.

While the trends of these results are general, the magnitudes will depend on the specifics of the spectral parameters. Explicitly, these results are quantitatively specific to a natural linewidth of 2.6 cm^{-1}, as the magnitude of the uncertainty is proportional to the natural linewidth (not shown). Broader natural linewidths result in lower uncertainties. This is because as the natural linewidth increases, it becomes greater in proportion to the instrument broadening. Additionally, a broader peak extends over more pixels providing more data to map out the curve. Thus, resolution is more consequential in measuring the linewidths of narrow highly crystalline systems than for the broad linewidths characteristic of disordered systems.

C. Two Peak - SiGe on Si

Using the same methodology, linewidth uncertainty is examined for the case of two overlapping peaks. In addition to including noise and spectral resolution, spatial resolution must be examined in this case to assess the substrate's impact on linewidth extraction. To this end, a range of RPIs (ratio of the substrate and film peak intensities, I_{Si}/I_{SiGe}) consistent with experimental observations of the SiGe film samples was investigated along with the same spectral resolution range studied previously. Noise and instrument broadening were modeled the same way described above. The signal from the two peak system is modeled as the sum of two Voigt functions, with natural widths of 4 cm^{-1} and 2.6 cm^{-1} for the film and substrate, respectively. The difference in the widths arises from alloy disorder broadening in the SiGe film.

Figure 4 plots the experimental (text boxes) and simulated uncertainties of the SiGe/Si system as a function of both resolution and RPI at a noise fraction of one, i.e., equal to experimental noise. First note that the simulated parameters closely match the experimental results, as shown in the supplementary information (Table S2), indicating the model accurately predicts the combined effects of resolution and RPI on linewidth measurements. The contours are roughly circular in shape indicating that resolution and RPI are of equal consequence over the examined region. However, as previously mentioned,

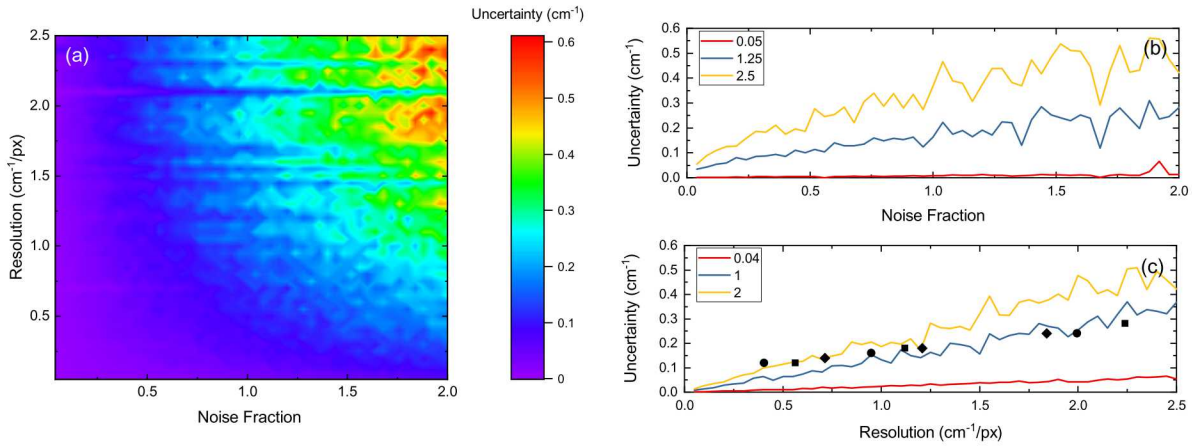

FIG. 3. (a) Contour plot of the uncertainty of the Si Lorentzian linewidth vs resolution and noise fraction (experimental noise, N_S, and N_B, multiplier) calculated by the virtual Raman experiment. (b) Cross section of the contour plot in the noise fraction direction at various values of resolution. (c) Cross section of the contour plot in the resolution direction at various values of noise fraction. Markers are experimental measurements (see Table S1 in supplemental for details) using the (circles) 514 nm, (squares) 488 nm and (diamonds) 405 nm lasers at various grating conditions outlined in Table I. Noise fraction of one is equal to experimental noise.

poor resolution can lead to errors in the fit. The cause of this error and the associated increase in uncertainty, observed in Figs. 3 and 4 as "streaks" of erroneous low uncertainty, are discussed more fully in Section III D. Since RPI and resolution are nearly equally important, we must identify regions on this surface plot which correspond to achievable experimental conditions with acceptable uncertainties.

Generally, the linewidth is used as a metric for a physical parameter of interest, for example temperature.[3,29] The uncertainty of the linewidth will therefore translate into uncertainty in temperature. The contour plots in Fig. 4 can then be used to identify the uncertainty of a deduced physical parameter. Conversely, knowing the precision needed in the measurement of a physical parameter, Fig. 4 can be used to select capable experimental conditions. For example, the tolerable uncertainty of a deduced physical parameter can be transformed into an uncertainty in linewidth using physical or emperical relationships. An uncertainty of 0.2 cm^{-1} in SiGe, for instance, will correspond to a 20 K temperature uncertainty at room temperature.[30] The maximum linewidth uncertainty can then be used to identify an elliptical exclusion zone of experimental conditions, bounded by the solid line in Fig. 4, defining the 0.2 cm^{-1} contour. Thermometry measurements to within 20 K for our hypothetical SiGe device would require experimental conditions outside of this bounded ellipse. Several conditions, most especially those for thinner SiGe films, do not meet this requirement because of the large substrate signal, i.e., large RPI. Note that this exclusion zone will shift depending on the noise in the system (dashed lines represent $\pm 50\%$ noise shift), but since the uncertainty is linearly related to noise, any increase in noise will simply yield a proportional increase in the uncertainties on this plot without changing the trends.

Taken together, experimental conditions which yield required uncertainties can therefore be identified using the virtual experiment developed here (see supplemental for code). This is extendable to a general Raman system and implemented via an accurate modeling of noise and spectral resolution on a particular system combined with the methods outlined above. With this analysis, experimental conditions can be optimized to minimize uncertainty.

D. Pixelation Error

Uncertainty has been shown to increase as spectral resolution worsens for both one and two-peak systems. While intuitively obvious, the underlying causes for this observation are more subtle. Specifically, additional broadening occurs in the signal owing to the pixelation of the continuous "signal" to that of the discrete intensity wavenumber pairs. This broadening occurs because each CCD camera pixel does not sample only from a given wavelength but rather over a range as seen in Eq. 1. The broadening impacts both uncertainty and error due to a change in spectral shape and a reduced number of fitting points.

The uncertainty of a fit is related to both the number of data points (i.e., samples or observations) and the residuals of the "fitted" model to the data.[31,32] Worsening resolution adversely affects each. Reduced data points, for instance, do not definitively map out the curve. The resulting fit will be less constrained and thus of greater uncertainty. Residuals, meanwhile, increase as worsening resolution distorts the curve from the model function utilized in the fit, see Fig 1 a). It is here that the implications of pixel induced broadening are most consequential.

The Voigt function is utilized to analyze Raman spectra as it is presupposed that the signal is accurately represented by the convolution of the spectrometer response function (Gaussian) and the natural linewidth of the analyte (Lorentzian). This

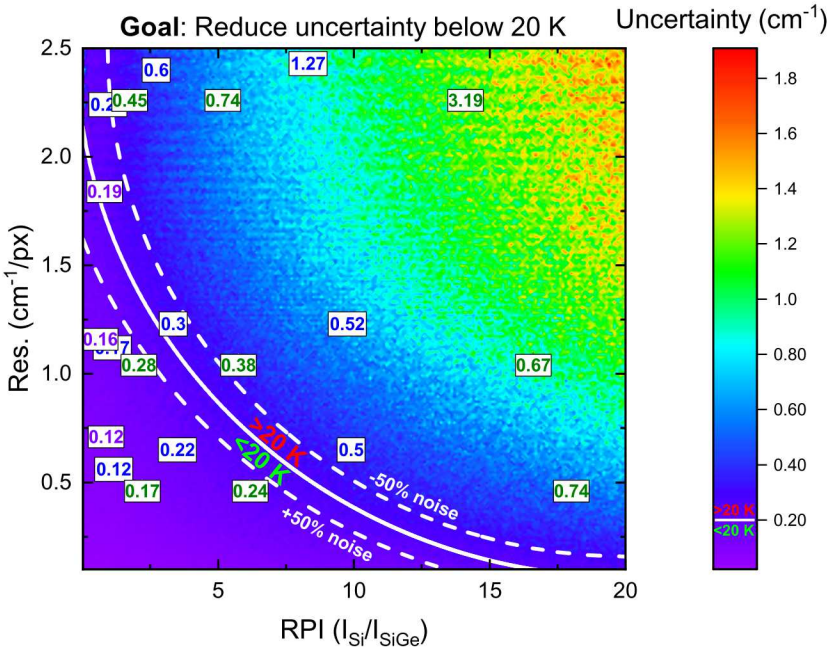

FIG. 4. Contour plot of the uncertainty of the SiGe Lorentzian linewidth vs resolution and RPI. Text indicates experimental uncertainty measurements, using the 514 nm (green), 488 nm (blue) and 405 nm laser (violet) using various grating conditions on the various SiGe thicknesses. RPI and resolution were varied using grating, laser wavelength and SiGe film thickness. See Table S2 in supplemental for details. A boundary (solid line) can be identified which separates experimental conditions which meet uncertainty criteria (<20 K and < 0.2 cm^{-1} in our hypothetical case) from conditions which do not meet the uncertainty criteria. This boundary shifts linearly with noise, N/S (dashed lines).

presupposition loses its veracity with worsening resolution as each pixel will sample over a larger bandwidth. The additional non-Gaussian broadening results in data pairs that are no longer a pure Voigt function. Residuals will therefore increase as exemplified by the poor fit of the 1800 (+1) grating condition shown in Fig. 1 a) and thus so too uncertainty as quantitatively illustrated along the ordinates of Figs. 3 and 4.

Pixel broadening can also induce errors in the extraction of the natural linewidth where the "true" linewidth falls outside the uncertainty bounds for the fit. For example, with poor spectral resolution, only a few pixels capture the spectral feature, and thus it is then likely that no pixel will be centered at the apex of the spectrum. The resulting fit can then have an arbitrary choice of peak height. Arbitrary peak height, in turn, causes the total linewidth (i.e., the full width at half maximum) to be ill defined leading to erroneous under predictions. On the other hand, if a pixel is centered at the apex, pixel broadening will be weighted towards the peak maximum distorting the Voigt to be more "Gaussian-like" leading to an under prediction of the Lorentzian component. These cases results in streaking of the uncertainty surface plots at poor resolution in Figs. 3 a) and 4.

To demonstrate, the influence of pixel broadening was examined through further investigation of the Raman response of Si. First, the Si-spectrum was defined as a Voigt function possessing a constant 2 cm^{-1} wide Gaussian component with a natural linewidth of 2.6 cm^{-1}. Virtual spectra were then generated at varying spectral resolutions using Eq. 1 and then fit to a Voigt function. Figure 5 a) plots the ratio of the extracted linewidth over the actual (input) linewidth versus the ratio of spectral resolution to linewidth. Non-dimensionalization generalizes these findings to arbitrary experimental conditions. Representative virtual spectra and fits at "good" (low) and "bad" (high) resolution are shown in Figs. 5 b) and c), respectively.

When the spectral resolution is much narrower than the linewidth, the natural linewidth is preserved and error negligible. However, the Voigt function in total is affected even in this case by pixel broadening. However, as seen in Fig. 5 a), the Gaussian component of the Voigt function increases and "absorbs" the pixel broadening resulting in an accurately deduced natural linewidth. With worsening resolution, the Gaussian component continues to increase but is no longer able to fully compensate for the pixel broadening as the discrete data diverges from a Voigt shape. This leads to significant errors in the extracted Lorentzian linewidth as exemplified for values above 0.38 px^{-1} in Fig. 5 a) where uncertainty not only increases but the "true" value lies outside of these bounds.

Figure 5 b) and c) provide qualitative insight into the quantitative results of Fig. 5 a) where identical spectra are shifted from one another to highlight the influence of pixel location relative to peak maximum. For "good" spectral resolution

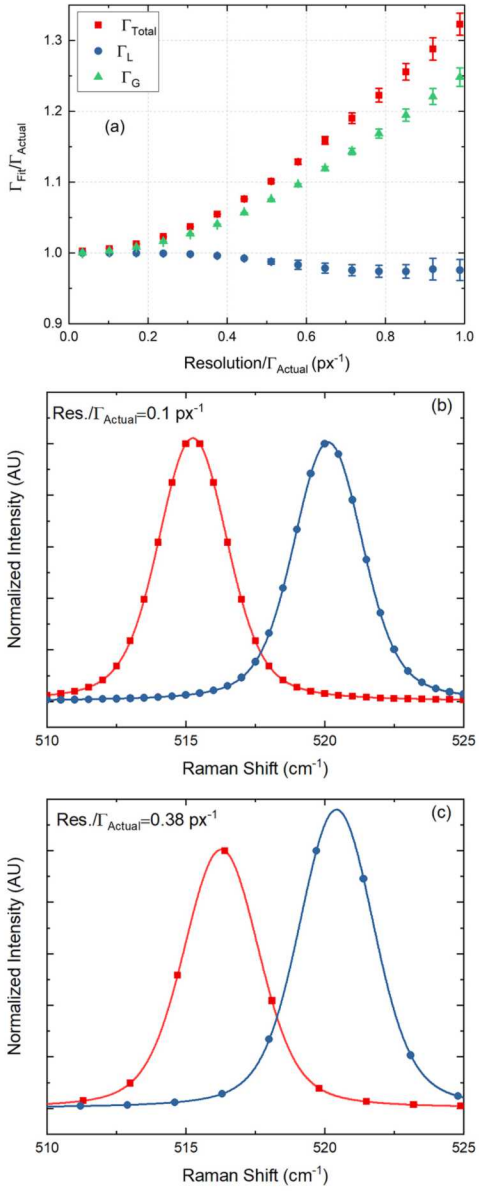

FIG. 5. (a) Normalized fit linewidth versus normalized resolution showing how resolution biases the fit of the Lorentzian linewidth. The fit linewidth and resolution were normalized by the actual linewidth of the simulation input peak. Error bars represent fit uncertainty. (b) Two spectra (markers) at high resolution with pixels located at different point on the peaks and their corresponding fit (lines) (c) Two spectra (markers) with pixels located at different locations on the peak at low resolution and their corresponding fits (lines).

(i.e., Res/linewidth<0.38 px^{-1}), several data points map out the lineshape. Thus, each data point samples over a comparatively small bandwidth thus mitigating the influence of pixel broadening. Therefore, pixel location in this instance is immaterial. In contrast, few data points compose the peak when using poor spectral resolution. In this case, the impact of pixel broadening is acute and hence the locations of the pixels themselves are of consequence which leads to the "streaks" in Fig. 3 a). This is demonstrated by the very different fit results observed for the same "true" spectrum seen in Fig. 5 c) which results in different fit parameters and uncertainties. Practically, these errors manifest when examining extremely narrow lines of a neon calibration lamp oftentimes used to quantify the spectrometer response function (i.e., instrument broadening).

Ultimately, pixel broadening manifests owing to the curvature of the spectral response. Its influence is therefore most apparent in regions of high curvature. For a single peak spectrum, this occurs only near the apex. Spectra possessing overlapping features have additional regions of high curvature at the saddle point between peaks. Thus, the influence of pixel broadening will be exacerbated and errors will increase when examining two peak spectra with insufficient spectral resolution. This is seen in Fig. 1 a) where the lineshapes are appreciably modified with changes in grating condition. These changes are most apparent by the increase in minima of the saddle point with worsening resolution. This is accompanied by poorer fits–compare fitted lines with data points–that result with departure from a Voigt response caused by pixel broadening. Similar to the case of a single peak, the fit will depend on the location of the pixels relative to peaks and saddle points in poor resolution conditions which leads to the "streaks" in Fig. 4.

The consequence of pixel broadening induced errors is that the range of experimental conditions which are appropriate for a given uncertainty limit, discussed in the previous section and Fig. 4, are reduced. The delimiting solid line now has the caveat that resolutions with few pixels mapping out a spectral feature must be avoided even if they fall in the region with uncertainties below the uncertainty limit (bottom left of the solid line in Fig. 4). Therefore, selection of experimental conditions should be weighted towards lower resolutions to minimize pixel induced biases.

IV. CONCLUSIONS

Spectral linewidths are used to measure numerous physical parameters in many spectroscopic experiments. When spectral features overlap, extraction of linewidth becomes difficult leading to large uncertainty. The magnitude of uncertainty is linked to several different experimental parameters that are oftentimes interdependent. To explore the relative influence of these experimental parameters, a virtual experiment methodology was employed to quantify uncertainty. Uncertainty is linearly related to relative peak strength, noise and spectral resolution with relative influences that are nearly equivalent. A holistic approach considering each of theses parameters must therefore be employed to identify the experimental conditions required to achieve a given linewidth precision. Experimental conditions not only impact uncertainty but also error as pixel induced broadening distorts spectral shape. Taken together, these results highlight the relative finesse required when quantifying linewidths compared to peak position in not only Raman spectroscopy but spectroscopic techniques in general.

V. ACKNOWLEDGEMENTS

Thanks to Elbara Ziade for critical review of this manuscript. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA0003525. This material is based upon work supported in part by the Air Force Office of Scientific Research under award number FA9550-18-1-0352.

[1] T. Beechem, S. Graham, S. P. Kearney, L. M. Phinney, and J. R. Serrano, Review of Scientific Instruments **78**, 061301 (2007).

[2] T. E. Beechem, *Metrology of GaN Electronics Using Micro-Raman Spectroscopy*, Ph.D. thesis, Georgia Institute of Technology (2008).

[3] T. E. Beechem and J. R. Serrano, Spectroscopy **33**, 26 (2011).

[4] U. Schwertmann, R. W. Fitzpatrick, R. M. Taylor, and D. G. Lewis, Clays and Clay Minerals **27**, 105 (1979), http://ccm.geoscienceworld.org/content.

[5] O. Lemine, Superlattices and Microstructures **45**, 576 (2009).

[6] K. Pantleon and M. A. Somers, Scripta Materialia **55**, 283 (2006).

[7] E. T. Jensen, R. E. Palmer, W. Allison, and J. F. Annett, Phys. Rev. Lett. **66**, 492 (1991).

[8] W. R. Kuhn and J. London, Journal of the Atmospheric Sciences **26**, 189 (1969), http://dx.doi.org/10.1175/1520-0469(1969)026¡0189:IRCITM¿2.0.CO;2.

[9] G. M. Bancroft, H. W. Nesbitt, R. Ho, D. M. Shaw, J. S. Tse, and M. C. Biesinger, Phys. Rev. B **80**, 075405 (2009).

[10] O. O. Bernal, D. E. MacLaughlin, H. G. Lukefahr, and B. Andraka, Phys. Rev. Lett. **75**, 2023 (1995).

[11] B. kyu Shin, Journal of Magnetic Resonance **249**, 1 (2014).

[12] W. F. Maddams, Applied Spectroscopy **34**, 245 (1980), http://asp.sagepub.com/content/34/3/245.full.pdf+html.

[13] J. Pitha and R. N. Jones, Canadian Journal of Chemistry **44**, 3031 (1966), http://dx.doi.org/10.1139/v66-445.

[14] K. T. Tsen, O. F. Sankey, and H. Morko, Applied Physics Letters **57**, 1666 (1990).

[15] C. B. Saltonstall, *Raman Measurments of Optical Phonon Scattering in Sub-Micron $Si_{1-x}Ge_x$ Thin Films*, Ph.D. thesis, University of Virginia (2016).

[16] The collection volume can also be controlled through confocality of the Raman spectrometer. However, the penetration depth of our samples and wavelength is much smaller (300 nm) than the confocal collection volume (1μm.).

[17] J. Kassim, C. Nolph, M. Jamet, P. Reinke, and J. Floro, Journal of Applied Physics **113**, 073910 (2013).

[18] T. R. Watkins, O. B. Cavin, C. R. Hubbard, B. Matlock, and R. D. England, Rigaku Journal **23**, 52 (2006).

[19] J. M. Hartmann, B. Gallas, R. Ferguson, J. Fernández, J. Zhang, and J. J. Harris, Semiconductor Science and Technology **15**, 362 (2000).

[20] P. Zaumseil, physica status solidi (a) **141**, 155 (1994).

[21] D. W. Posener, Australian Journal of Physics **12**, 184 (1959).

[22] F. Cerdeira, T. A. Fjeldly, and M. Cardona, Solid State Communications **13**, 325 (1973).

[23] K. Sinha, A. Mascarenhas, G. S. Horner, K. A. Bertness, S. R. Kurtz, and J. M. Olson, Phys. Rev. B **50**, 7509 (1994).

[24] A. C. Ferrari and J. Robertson, Phys. Rev. B **64**, 075414 (2001).

[25] M. Grimsditch, A. Polian, and R. Vogelgesang, Journal of Physics: Condensed Matter **15**, S2335 (2003).

[26] P. Zhang, Y. Feng, R. Anthony, U. Kortshagen, G. Conibeer, and S. Huang, Journal of Raman Spectroscopy **46**, 1110 (2015).

[27] P. M. Fauchet and I. H. Campbell, Critical Reviews in Solid State & Materials Sciences **14**, 79 (1988).

[28] I. Campbell and P. Fauchet, Solid State Communications **58**, 739 (1986).

[29] T. Beechem and S. Graham, Journal of Applied Physics **103**, 093507 (2008).

[30] H. H. Burke and I. P. Herman, Phys. Rev. B **48**, 15016 (1993).

[31] N. Börlin and P. Grussenmeyer, The Photogrammetric Record **28**, 396 (2013).

[32] J. Yang, E. Ziade, and A. J. Schmidt, Review of Scientific Instruments **87**, 014901 (2016), https://doi.org/10.1063/1.4939671.