Published in final edited form as:

Surf Interface Anal. 2017 December ; 49(12): 1187–1205. doi:10.1002/sia.6285.

Quantitative analysis of trace levels of surface contamination by X-ray photoelectron spectroscopy Part I: statistical uncertainty near the detection limit

Shannon B. Hill, Nadir S. Faradzhev*, and Cedric J. Powell

National Institute of Standards and Technology, Gaithersburg, MD 20899

Abstract

We discuss the problem of quantifying common sources of statistical uncertainties for analyses of trace levels of surface contamination using X-ray photoelectron spectroscopy. We examine the propagation of error for peak-area measurements using common forms of linear and polynomial background subtraction including the correlation of points used to determine both background and peak areas. This correlation has been neglected in previous analyses, but we show that it contributes significantly to the peak-area uncertainty near the detection limit. We introduce the concept of relative background subtraction variance (RBSV) which quantifies the uncertainty introduced by the method of background determination relative to the uncertainty of the background area itself. The uncertainties of the peak area and atomic concentration and of the detection limit are expressed using the RBSV, which separates the contributions from the acquisition parameters, the background-determination method, and the properties of the measured spectrum. These results are then combined to find acquisition strategies that minimize the total measurement time needed to achieve a desired detection limit or atomic-percentage uncertainty for a particular trace element. Minimization of data-acquisition time is important for samples that are sensitive to x-ray dose and also for laboratories that need to optimize throughput.

Introduction

Detection and quantification of trace levels of surface contamination is a challenge for many modern analytical techniques. X-ray photoelectron spectroscopy (XPS) is widely used to obtain quantitative information about surface composition and the chemical states of elements in the near-surface region from which the signals originate. Core-level electrons excited by X-ray photons deliver quantitative information about the species in the surface region defined by the information depth for the detected photoelectrons and the XPS measurement configuration.¹ The measured intensity for a particular peak is determined by the number of electrons in the relevant subshell, the photoionization cross section, the depth distribution of analyzed atoms, elastic- and inelastic-scattering processes in each layer, possible structural and chemical effects, the measurement configuration, the analyzer acceptance angle, and the efficiency of the detector.¹ The surface composition can be inferred from a measurement of the intensities of the peaks associated with transitions for

* current affiliation: KLA Tencor, 3 Technology Drive, Milpitas, CA 95035

each species present on the surface. We discuss here the accuracy and detection limits of XPS quantitative analyses of species present at trace levels. In Part 1 we describe the statistical uncertainties associated with the measurement and analysis of XPS intensities. In Part 2, we provide a method for estimating the absolute amount and systematic uncertainties of a particular trace element.²

The simplifying assumptions that are commonly used to estimate surface composition from XPS measurements are described in greater detail in Part 2 of this work. Briefly, the distribution of species is assumed to be homogenous throughout the information depth of the measurement, and the variation of scattering parameters, atomic densities, chemical states and possible structural ordering of the different species is neglected. The concentration of a particular species x relative to all species (indexed by q) measured in the surface region is then expressed as

$$C_x = \frac{I_x/R_x}{\sum_q I_q/R_q} \quad (1)$$

Here I_q and R_q are the background-subtracted peak intensity and the elemental relative sensitivity factor (RSF), respectively, for the relevant energy level and mode of operation for species q . (Note that instrumental software often automatically adjusts RSFs to account for variations in the transmission function of the XPS instrument due to changes in the operating mode.) As is common practice, we define the intensity of the peak associated with species q as the number of counts in the background-subtracted area of the spectral feature times the energy step size, dE_q , normalized to the acquisition dwell time, dt_q (taken to be the product of the dwell time per spectrum point and the number of spectral sweeps). I_q therefore, has units of counts per second times electron volts and is given by

$$I_q = A_{BS,q} dE_q / dt_q \quad (2)$$

Here $A_{BS,q}$ is the peak area given by the direct sum of the counts at each measurement point spanning the peak for species q after subtraction of a suitable background. (The species subscript will be suppressed unless needed to distinguish between peaks from multiple species.)

Industrial laboratories often utilize XPS to determine whether a certain type of surface contamination is present at levels in excess of some critical limit. In these cases, it is desirable to minimize the measurement time to improve throughput. Even in situations where measurement time is not a critical factor, it is often important to minimize the exposure of the sample to x-rays which may modify the sample and affect the relevant measurement. The trade-off between sensitivity and measurement time is a concern for XPS as for other analytical techniques; a similar trade-off may have to be made with spatial resolution. Currie provided a statistical basis for the often loosely defined concepts of

detection and quantification limits first in radiochemistry³ and then in a broader historical context⁴. The definition of detection limit by Technical Committee 201 on Surface Chemical Analysis of the International Organization for Standardization (ISO) is based on the common convention that the detection limit is the amount of material for which the background-subtracted signal is three times its standard deviation.⁵

Powell and Seah⁶ presented a detailed outline of the factors influencing both systematic and random uncertainties encountered in surface-analysis measurements using XPS and Auger-electron spectroscopy. Harrison and Hazell⁷ and also Evans⁸ derived relationships between the statistical uncertainties of atomic concentrations and the acquisition parameters, peak areas and number of points used for the most common method of linear-background subtraction. The former work proposed strategies for minimizing acquisition time and uncertainty, and the latter examined the impact on least-squares peak-fitting methods. Cumpson and Seah⁹ proposed a procedure for estimating the statistical uncertainties of peak position, peak width and peak area using common peak-fitting methods for overlapping or multiple peaks with complex backgrounds. This effort was extended by applying a Bayesian statistical model to determine the uncertainties of the resulting atomic concentrations calculated in two different ways.¹⁰ There have also been a number of empirical studies that describe the statistical variation of peak intensities reported by different analysts when they were presented with the same spectra.¹¹ Specific to the case of trace elements near the detection limit, Suzuki et al. reported variation among analysts of peak detection by eye¹² and with commercial software.¹³ Recently, Shard presented a method for systematically estimating detection limits for elements present in more than 6000 binary systems.¹⁴ The Shard approach was found useful in a recent experimental test.¹⁵

Many of the commercially available XPS-analysis algorithms use a subset of spectrum points to determine both the background and the peak area. The impact of the resulting correlation has not been explored in previous studies of the sensitivity and uncertainty of XPS measurements. This is understandable for pronounced peaks that are easily distinguished from the background. In this “large-peak-area regime,” the contribution from the background to the total uncertainty of the peak area is typically negligible compared to the intrinsic uncertainty of the peak area itself; however, we will show that the statistical uncertainty of peaks near the detection limit can be significantly overestimated by the commonly employed “uncorrelated-area approximation” that neglect the correlation of the background and peak areas. We begin by including the correlation of the background and peak areas in the derivation of general expressions to estimate the statistical uncertainty of peak areas for linear backgrounds. This process naturally leads to the concept of the relative background-subtraction variance (RBSV) which quantifies the contribution to the statistical uncertainty of the peak area from the method of background determination relative to the contribution from the background area itself. The RBSV allows the statistical uncertainty of a small peak area to be expressed in a form that separates the dependence on acquisition parameters, properties of the spectrum, and the analysis methods. We demonstrate that this approach is valid for commonly used linear backgrounds determined by endpoint averaging as well as for polynomial backgrounds determined by least-squares fitting. We then review the relationships between the uncertainties of atomic concentrations and peak areas. We combine these results to provide optimal strategies for both analysis and acquisition that will

minimize the measurement time required to achieve a target detection limit or uncertainty for an element of interest. This analysis reveals that the RBSV is useful as a metric to determine the conditions under which the large-peak-area approximation is valid and to establishing the limits of fruitful background averaging.

Statistical uncertainty of peak-area measurements

In the absence of interference from features of other species, the background in the vicinity of an XPS peak of a trace contaminant element is usually well approximated by a straight line. Due to the inherently small area of the peak, different choices of background functions will have small effects on the derived peak areas. We will show, however, that the uncertainty of the peak area is dramatically affected by the method of background determination. We first derive the statistical uncertainty of the peak area determined by subtracting a linear background defined by averaging symmetric numbers of points on either side of the peak. We then discuss the more general cases of asymmetric endpoints and least-squares fitting of polynomial backgrounds. In both cases the derived relationships are compared to results of Monte-Carlo simulations. Although polynomial backgrounds are not commonly employed for prominent XPS peaks, they are useful when estimating detection limits for small or absent peaks on a smoothly varying background.

The simulated XPS peak shown in Fig. 1 illustrates the various parameters used to determine the background-subtracted peak area and its uncertainty for an arbitrary background-subtraction method. (These and other relevant symbols are summarized in Table 1. Tables 2 and 3 describe modifiers that indicate specific applications of some of the symbols in Table 1.) N_P is the minimum number of spectrum points comprising the entire peak with no background points, and $E_P = N_P dE$ is the corresponding measured peak energy range. N_P is often taken to be the largest integer less than $3 E_{FWHM}/dE$, where E_{FWHM} is the expected full width at half maximum for the peak. The user must select N_L and N_R points on the left and right sides of the peak, respectively, that will be used to determine the background. The background-subtracted peak area will be determined by the difference between the measured signal and the calculated background value over the analysis energy range, E_A , which usually includes some background points in addition to the peak points. N_A is the total number of analysis points in E_A ; N_{LA} and N_{RA} are the number of background points within E_A (correlated with peak area); and N_{LB} and N_{RB} are the number background points outside E_A (uncorrelated with peak area). With these definitions,

$$\begin{aligned} N_L &= N_{LA} + N_{LB}, \\ N_R &= N_{RA} + N_{RB}, \\ N_A &= N_P + N_{LA} + N_{RA}, \end{aligned} \quad (3)$$

and the meanings of the subscripts are summarized in Table 2.

Most XPS-analysis algorithms use symmetric endpoints so that $N_S = N_L = N_R$ and $N_{SA} = N_{LA} = N_{RA}$. It is also common practice to select the analysis region so that equal numbers of

background points are inside and outside E_A so that $N_{LA} = N_L/2$ and $N_{RA} = N_R/2$. Since half of the background points are correlated with the peak area, we describe the background points in this scenario as “semi-correlated” with peak area. Frequently, these conditions are combined using symmetric background points that are semi-correlated with peak area so that $N_A = N_P + N_S$. To illustrate the two most common correlation cases, Fig. 1 shows the unlikely combination of asymmetric endpoints that are semi-correlated with the peak area on the left ($N_{LA} = N_L/2$) and uncorrelated with peak area on the right ($N_{RA}=0$). Specific examples are provided in Part 2 of this work.²

Linear background with symmetric endpoint averaging

The most common method of determining a linear background is to use the line connecting the two average endpoints defined by the mean energy and mean signal of each set of background points on the left and right of the peak. The standard uncertainty of the background-subtracted peak area, $\sigma_{A_{BS}}$, for an arbitrary selection of background points is derived in Appendix A. For the case of symmetric endpoints, the uncertainty of the background-subtracted peak area is given by

$$\left(\sigma_{A_{BS}}\right)^2 = \left(\sigma_{A_P}\right)^2 + \left(\sigma_{\langle y_B \rangle}\right)^2 N_A \left(1 + \frac{N_A - 4N_{SA}}{2N_S}\right) \quad (4)$$

where $\sigma_{\langle y_B \rangle}^2$ is the variance of the mean background signal, and $\sigma_{A_P}^2$ is the intrinsic variance of the true peak area, A_P . Although A_P and its measured estimate A_{BS} are presumed to be numerically similar, Eq. (4) shows that their uncertainties are not. The uncertainty σ_{A_P} is determined solely by the statistics of the measured signal comprising the peak area, while $\sigma_{A_{BS}}$ includes the contribution from the uncertainty of the background determination and its subtraction from the measured signal.

Equation (4) predicts the variance for any statistical distribution of the measured signal. Poisson statistics are appropriate if the signal is reported in counts from a single detector (not the average of multiple detectors) and has not been normalized by a sensitivity factor or transmission function. In this case, $\sigma_{A_P}^2 = \hat{A}_P$ and $\sigma_{\hat{y}_B}^2$ is well approximated by the average background signal, $\langle \hat{y}_B \rangle$. (Caret is used to indicate quantities based on unscaled counts.) With these substitutions in Eq. (4), the relative background-subtracted peak-area uncertainty for Poisson-distributed statistics is given by

$$\left(\frac{\sigma_{\hat{A}_{BS}}}{\hat{A}_P}\right)^2 = \frac{1}{\hat{A}_P} \left[1 + \frac{\langle \hat{y}_B \rangle N_A}{\hat{A}_P} \left(1 + \frac{N_A - 4N_{SA}}{2N_S} \right) \right] \quad (5)$$

For species present at trace levels, the average background area, $\langle \hat{y}_B \rangle N_A$, is typically at least two orders of magnitude greater than the peak area, $\hat{A}_P$. Hence, for measurements governed by Poisson statistics, the primary contribution to the peak-area uncertainty is from

the second term in Eq. (5). Even if the signal has been averaged or scaled in some way that leads to deviations from Poisson statistics, the small peak area of a trace element will typically have an absolute intrinsic variance that is significantly smaller than the absolute variance of the much larger background area. The total uncertainty will therefore likely be dominated by the second term in Eq. (4), and $\sigma_{\langle y_B \rangle}$ can be determined from the goodness-of-fit parameter resulting from a polynomial fit to a nearby section of contiguous background data as described by Shard.^{7,14}

The unity term in the parentheses of the second term of Eqs. (4) and (5) is the uncertainty contribution from the variation in the intrinsic number of counts contained in the background area. The second term in parentheses accounts for the uncertainty introduced by the methods used to determine and subtract the background. As expected, this contribution decreases inversely with number of points used to define the background on each side, N_S . It also decreases linearly with the number of background points that are inside the analysis region, N_{SA} . This result is due to the correlation introduced by using the same N_{SA} points to determine both the total peak area and the background. The existence of this correlation has been mentioned in previous work,^{7,16} but it was neglected because the measurements occurred near the large-peak-area regime. N_S and N_{SA} in these earlier works were also typically much smaller than N_A which decreased the impact of the correlation. Near the detection limit, however, the correlation term is usually not negligible because many background points ($N_S \gtrsim N_A$) are often needed to reduce the contribution from the background-subtraction procedure, which is dominated by the peak-area uncertainty.

Figure 2 demonstrates excellent agreement between the relative variance of the background-subtracted area predicted by Eq. (5) and that obtained from Monte-Carlo simulations of 3,000 spectra. The results shown are for Gaussian peaks with the same amplitude (100 counts) and full-width at half maximum ($E_{FWHM} = 2$ eV), but the average background signal was chosen to be $\langle \hat{y}_B \rangle = (300, 1000 \text{ or } 3000)$ counts for panels (a), (b) and (c), respectively, of Fig. 2. Each of these plots shows the relative variance as a function of N_S for symmetric endpoints that are both semi-correlated ($N_{SA} = N_S/2$) and uncorrelated ($N_{SA}=0$) with the peak area. To demonstrate the dependence on the energy step size of the acquired spectrum, results are shown for $dE = (0.2, 0.5 \text{ and } 1)$ eV in each plot. In all cases the expected peak energy range is taken to be three times the FWHM so that N_P is the largest integer less than $3 E_{FWHM}/dE$. Examples of the three different simulated spectra using the three different dE values are shown in Fig. 3.

The common shapes of the curves in Figs. 2(a–c) suggest that the variance of the background-subtracted peak area can be expressed in a form that separates the dependences on quantities inherent to the acquired spectrum ($\sigma_{AP} \sigma_{\langle y_B \rangle}$ and N_P) from parameters associated with the background-subtraction method (N_S and N_{SA}). To decouple the background points from the peak points we substitute $N_A = N_P + 2N_{SA}$ [Eq. (3) for symmetric endpoints] and factor N_P from all terms in Eq. (4) to obtain

$$\left(\sigma_{ABS}\right)^2 = \left(\sigma_{AP}\right)^2 + \left(\sigma_{\langle y_B \rangle}\right)^2 N_P \left[(1 + 2n_{SA}) \left(1 + \frac{1 - 2n_{SA}}{2n_S} \right) \right]. \quad (6)$$

A lower-case n indicates normalization to the number of peak points with $n_{SA} = N_{SA}/N_P$ and $n_S = N_S/N_P$. The term in the square bracket of Eq. (6) depends only on these ratios of background to peak points; therefore, it is independent of the acquisition parameters (e.g., dE) and depends only on the method of background determination (e.g., the fraction of the data used to define the background relative to that comprising the peak). This observation motivates the definition of the relative background-subtraction variance (RBSV) as the contribution to σ_{ABS}^2 remaining after subtracting the intrinsic variance of the peak area, σ_{AP}^2 , and normalizing to the intrinsic variance of the background area under the peak, $(\sigma_{\langle y_B \rangle})^2 N_P$.

$$RBSV(n_\Lambda) = \frac{\left(\sigma_{ABS}\right)^2 - \left(\sigma_{AP}\right)^2}{\left(\sigma_{\langle y_B \rangle}\right)^2 N_P}. \quad (7)$$

The general form of the RBSV in Eq. (7) is expressed as a function of n_Λ to emphasize that the RBSV depends only on the ratios of the numbers of points in the relevant background sets (indicated by Λ values from Table 2) to the number of points comprising the peak. These background-to-peak point-number ratios are equivalent to the ratios of the energy ranges used to define the background and peak regions of the spectrum, so the RBSV depends not on the actual number of measurement points, but on the fraction of the spectrum data that is used to define the background relative to the amount of data in the peak.

The RBSV describes the uncertainty contribution from the method used to determine the background, independent of the signal strength, spectrum characteristics and acquisition parameters. For the specific case of a linear background defined by symmetric endpoint averaging, substitution of Eq. (6) into Eq. (7) gives

$$RBSV_{EA}(n_S, n_{SA}) = (1 + 2n_{SA}) \left(1 + \frac{1 - 2n_{SA}}{2n_S} \right). \quad (8)$$

For comparison, it will be helpful to use Eq. (7) to define the effective RBSV for symmetric endpoint averaging resulting from the commonly employed uncorrelated-area approximation. Using the σ_{ABS} defined by Equations (1) and (2) in Shard¹⁴ and with the appropriate change of variables[†] this gives

[†] $\sigma_1^{(Shard)} \rightarrow \sigma_{AP}, \sigma_A^{(Shard)} \rightarrow \sigma_{ABS}, \sigma_\beta^{(Shard)} \rightarrow \sigma_{\langle y_B \rangle}, N^{(Shard)} \rightarrow N_A$ and $w^{(Shard)} \rightarrow N_S$

$$RBSV_{EA, Approx}(n_S) = \frac{(1 + n_S)(1 + 3 n_S)}{2 n_S}. \quad (9)$$

We will see that $RBSV_{EA, Approx}(n_S) > RBSV_{EA}(n_S, n_{SA})$ for all n_S and n_{SA} and that this difference increases with n_S .

From the general RBSV definition in Eq. (7), the uncertainty of the background-subtracted peak area can be written as

$$\left(\sigma_{A_{BS}}\right)^2 = \left(\sigma_{A_P}\right)^2 + \left(\sigma_{\langle y_B \rangle}\right)^2 N_P RBSV(n_\Lambda). \quad (10)$$

Equation (10) has the same form as Eq. (6) which was derived for a linear background determined by endpoint averaging. We will show that Eq. (10) also applies to both linear and polynomial backgrounds determined by least-squares fitting. (This formulation is not appropriate for larger peaks that require Shirley or other peak-dependent background types.) According to Eq. (10) the uncertainty of the background-subtracted peak area is determined by three components:

1. The intrinsic variance of the true peak area, $(\sigma_{A_P})^2$
2. The intrinsic variance of the background area under the peak, $N_P(\sigma_{\langle y_B \rangle})^2$
3. The RBSV which depends solely on the type of background and on the amount of spectrum data used to define the background relative to that which defines the peak.

The first two components are properties of the measured spectrum and depend on the strength of the peak and background as well as on the acquisition parameter dE which determines the number of points comprising the peak, N_P . The uncertainty introduced by the method of background determination and subtraction is entirely described by the RBSV, which is independent of acquisition parameters. Figure 2(d) demonstrates this result by comparing the RBSV predicted by Eq. (8) (curves) to the values determined by the statistical definition of the RBSV in Eq. (7) (points). For the latter calculation, the quantities $\sigma_{A_{BS}}$, σ_{A_P} and $\sigma_{\langle y_B \rangle}$ were determined from the Monte-Carlo simulations that produced the results plotted in Figs. 2(a–c). The excellent agreement between the points and curves in Fig. 2(d) indicates that the RBSV for linear endpoint averaging given by Eq. (8) is accurate. This plot also demonstrates that the RBSV depends only on the background-to-peak point-number ratio, n_S , since all points lie on the same curve for simulated spectra with three different background rates acquired using three different energy step sizes. Appendix A shows that the uncertainty of the background-subtracted peak area can be expressed in the same form as Eq. (10) for a linear background determined by asymmetric endpoint averaging. As demonstrated in Appendix B and discussed in the next section, this formulation is also

consistent with the uncertainty for polynomial backgrounds determined by least-squares fitting.

With the appropriate RBSV, Eq. (10) can be used to determine optimal analysis strategies for minimizing the peak-area uncertainty independent of the properties of the spectrum or the acquisition parameters. For endpoint averaging of symmetric endpoints, the RBSV given by Eq. (8) shows that an increase in n_S has two competing effects. Including more points lowers the uncertainty of the average background signal at each endpoint, but it also pushes the endpoints to energies further from the peak, which increases the uncertainty of the background function since it must be interpolated over a larger range. The lowest thin black curve in Fig. 2(d) shows that if the endpoints are uncorrelated with the peak area ($n_{SA} = 0$), these competing effects balance as n_S increases, and the RBSV decreases monotonically to an asymptotic value of unity. For the semi-correlated case ($n_{SA} = n_S/2$), however, there is an additional contribution to the background uncertainty because the number of points at which the background function must be evaluated increases with n_S . [I.e., N_A increases with $N_{LA} = N_{RA} = N_S$ in Eq. (3)]. This case is shown by the thick black curve in Fig. 2(d). For $n_S < 1$ (i.e., $N_S < N_P$), increasing n_S reduces the endpoint uncertainty more than it increases the background area uncertainty, but for $n_S > 1$ the latter effect dominates, and the total uncertainty increases with n_S . Hence, for the common case of symmetric, semi-correlated endpoint averaging, the uncertainty is minimized when the background is defined using the same number of points (on each side of the peak) as comprise the peak, regardless of the peak width or the energy step size used to acquire the spectrum:

$$\begin{aligned} n_S^{Opt} = 1 & \quad \text{To minimize } \sigma_{ABS} \text{ for linear} \\ \text{(i.e., } N_S = N_P) & \quad \text{background defined by symmetric,} \\ & \quad \text{semi-correlated endpoint averaging} \end{aligned} \quad (11)$$

Even for this optimal condition, Fig. 2(d) shows that the RBSV for the semi-correlated case is 33 % larger than the uncorrelated case at $n_S = 1$. In fact, for symmetric endpoints, the uncertainty as a function of the relative number of peak-area-correlated endpoints, n_{SA} , is *maximized* for the *commonly* used semi-correlated condition ($n_{SA} = n_S/2$) and minimized for the less commonly encountered uncorrelated case ($n_{SA} = 0$). It is thus always better to use uncorrelated endpoint averaging when available.

Comparison of the black and brown curves in Fig. 2(d) shows the error introduced by using the uncorrelated-area approximation [Eq. (9)]. For the optimal condition of Eq. (11), the approximation overestimates the peak-area variance by a factor of 2 for semi-correlated endpoint averaging and by a factor of 2.7 for the uncorrelated case. The uncorrelated-area approximation would also underestimate by about a factor of 2 the number of background points needed to minimize the uncertainty since the minimum of Eq. (9) occurs near $n_S \approx 0.5$ rather than 1.0 [Eq. (11)]. Shard's method for estimating the detection limit of binary systems¹⁴ is based on the uncorrelated-area approximation. His results may be corrected for the effects of correlation by using $4.9\sqrt{eV}$ (rather than $9\sqrt{eV}$) for Shard's symbol of κ defined as the proportionality factor between the peak area at the detection limit and the

statistical uncertainty of the background signal divided by the square root of the energy step size.

Linear and polynomial backgrounds using asymmetric endpoints

The data available for defining a linear background may be restricted to an energy range much less than the extent of the peak, E_P , on one side due to a neighboring spectral feature. If the range of points consistent with a linear background has a background-to-peak point-number ratio that is smaller on one side of the peak, n_{Small} , than the other, n_{Large} , symmetric endpoint averaging is constrained to using n_{Small} on both sides with an RBSV given by Eq. (8) with $n_S = n_{Small}$. In this section we compare the RBSV of a linear background defined by constrained symmetric endpoint averaging to three common methods that utilize more of the available background data by using different numbers of points on each side of the peak. We examine linear backgrounds defined by endpoint averaging, and linear and polynomial backgrounds determined by least-squares fitting.

Linear endpoint averaging—The RBSV for asymmetric endpoint averaging is given by Eq. A.8 with $n_L = n_{Small} < n_R = n_{Large}$. Figure 4(a) illustrates the advantages of using asymmetric rather than symmetric endpoint averaging when the background on one side is linear over just 10 % of the energy range spanned by the peak ($n_{Small}=0.1$). The asymmetric RBSV normalized to the symmetric, constrained RBSV ($n_L = n_R = 0.1$) is plotted as a function of n_{Large} for both semi-correlated and uncorrelated endpoints in Fig. 4(a). For both correlation cases, the asymmetric RBSVs initially decrease, pass through a minimum and then increase with n_{Large} because of the tradeoff between the various uncertainty contributions discussed in the previous section. The optimal value of n_{Large} for endpoints that are semi-correlated with the peak area is

$$n_{Large}^{Opt} \equiv \frac{n_{Small}}{4} \left(\sqrt{9 + \frac{16}{n_{Small}}} - 1 \right). \quad \begin{array}{l} \text{To minimize } \sigma_{A_{BS}} \text{ for linear} \\ \text{background defined by} \\ \text{asymmetric endpoint averaging} \end{array} \quad (12)$$

A closed-form expression for the uncorrelated case does not exist; however, the value of the uncorrelated RBSV at $n_{Large} = n_{Large}^{Opt}$ given by Eq. (12) is within 6 % of the true minimum value for all n_{Small} , as indicated in Fig 4(a) for $n_{Small}=0.1$. Thus Eq. (12) is a good approximation of the optimal number of unconstrained endpoints for linear backgrounds determined by asymmetric endpoint averaging regardless of correlation.

Figure 4(b) plots the minimum ratio of the RBSV for asymmetric endpoints with the optimal n_{Large} of Eq. (12) to the RBSV for the symmetric, constrained case ($n_L = n_R = n_{Small}$) as a function of n_{Small} . Very little reduction over the symmetric, constrained case is obtained for $n_{Small} > 0.15$, but in cases where only one or two linear background points are available on one side ($n_{Small} < 0.15$), the asymmetric case using Eq. (12) can significantly reduce the uncertainty over the symmetric, constrained method. Comparison of the thick and thin curves in Fig. 4(a) and 4(b) demonstrate that the uncertainty for uncorrelated endpoint

averaging is always lower than the semi-correlated case, just as was observed for symmetric endpoints. Indeed, the dashed line in Fig. 4(b) shows that the peak-area variance resulting from the common default choice of constrained symmetric semi-correlated endpoints can always be reduced by 15 % to 25 % by using asymmetric uncorrelated endpoints instead.

A somewhat counterintuitive result of the above analyses for symmetric and asymmetric endpoint averaging is that the peak-area uncertainty may be *increased* if more than an optimal number of background points are used for most commonly employed methods. *Only* for the special case of symmetric endpoints that are uncorrelated with the peak area will the uncertainty *always* be reduced by including more background points. To ensure the lowest possible uncertainty for the symmetric semi-correlated case, Eq. (11) indicates that no more than N_P points should be used on each side of the peak ($n_S^{Opt} = 1$). For the asymmetric case (regardless of endpoint-peak-area correlation), Eq. (12) indicates that fewer than N_P background points should be used on the unconstrained side if fewer than N_P background points are available on the constrained side ($n_{Large}^{Opt} < 1$ if $n_{Small} < 1$), and more than N_P unconstrained background points may be used if there are more than N_P points on the constrained side ($n_{Large}^{Opt} > 1$ if $n_{Small} > 1$). For example, when the number of points consistent with a linear background is only 10 % of N_P on one side of the peak ($n_{Small} = 0.1$), Fig. 4(a) shows that the RBSV will be 20 % lower if $n_{Large} = n_{Large}^{Opt} = 0.3$ [Eq. (12)] is used on the other side of the peak compared to using more background points as in the $n_{Large} = 1$ condition for the symmetric, unconstrained case of Eq. (11).

Linear least-squares fitting—Appendix B describes the general procedure for finding the RBSV for a background determined by least-squares fitting (LSF) of a polynomial of arbitrary order to endpoints that are not correlated with the peak area ($n_{LA} = n_{RA} = 0$). The first row of Table B.1 gives the RBSV for LSF of asymmetric endpoints to a linear background. Note that for symmetric endpoints, the RBSVs for a linear background determined by LSF and by endpoint averaging [Eq. (8) with $n_{SA} = 0$] are equal. The thin black and gray curves in Fig. 4(a) compare these RBSVs for asymmetric endpoints as a function of n_{Large} when $n_{Small} = 0.1$. When the optimal condition of Eq. (12) is used for both methods, the RBSVs are approximately equal for all n_{Small} . Figure 5 shows that while the RBSV is always lower for LSF, it only offers significant reduction in the uncertainty over the endpoint-averaging method when $n_{Large} \gg n_{Small}$. For example, the linear-LSF RBSV is at least 20 % smaller than the endpoint-average RBSV only when $n_{Large} \gtrsim 10 n_{Small}$. Since the peak-area uncertainty is proportional to the square root of the RBSV [Eq. (10), neglecting σ_{Ap}], linear LSF would offer a 20 % reduction in uncertainty over endpoint averaging only when $n_{Large} \gtrsim 20 n_{Small}$. This suggests that LSF may offer significant benefits when only one or two background points are available on *one* side of the peak and a linear background extends for many multiples of the peak width on the other side. Of course, LSF is the only option if there are no background points consistent with a linear background on one side of the peak.

Polynomial least-squares fitting—If the background can only be considered linear for one or two points on *both* sides of the peak, the uncertainty may be reduced by LSF of a

polynomial over a larger range of background points. Although a polynomial background is rarely appropriate for large peaks, it can be useful to determine the minimum detectable peak area in a region of the spectrum where the background is not linear on one or both sides of the expected peak location. However, the additional degrees of freedom introduced by even a low-order polynomial fit dramatically increase the uncertainty of the peak area. Here we determine the conditions under which a polynomial background will produce a lower peak-area uncertainty (and detection limit) than a truncated linear background. We consider only the case of endpoints that are not correlated with the peak area ($n_{LA} = n_{RA} = 0$), since the uncertainty is always larger for the semi-correlated case.

The RBSV for a polynomial-background of order M fit with endpoints spanning energy ranges n_L and n_R is denoted $RBSV_M(n_L, n_R)$ and derived in Appendix B. The $RBSV_M(n_L, n_R)$ for even and odd polynomial orders are plotted in Fig. 6 as a function of n_L for three different n_R values. The RBSVs for all polynomial orders decrease monotonically with increasing total number of background points, but the oscillations show that the rate of decrease can vary significantly depending on the balance of n_L and n_R . (The dependence on background-point asymmetry is discussed further at the end of Appendix B.) Monte Carlo simulations for LSF of polynomial backgrounds were performed over the entire range of parameters shown in Fig. 6 and compared to the predictions of Appendix B. The agreement was found to be just as good as that demonstrated in Fig. 2 for linear backgrounds; however, this degree of accuracy holds only if the number of background points on each side of the peak is at least large as the order of the polynomial (N_L and N_R both $\geq M$). In the extreme condition when fewer than M endpoints are used, the RBSV can no longer be considered independent of N_P .

The $RBSV_M(n_S)$ for symmetric endpoints $n_S \equiv n_L = n_R$ are plotted as thick black curves in Fig. 6. There is only one such curve in each panel because the RBSV for symmetric endpoints fit to an even-order polynomial, $RBSV_{M_{Even}}(n_S)$, is identically equal to that for the subsequent odd-order polynomial, $RBSV_{M_{Even}+1}(n_S)$. As discussed in Appendix B, this equality is due to the symmetry of applying LSF to two sets of points of equal number and spacing separated by a region with no points (i.e., the peak region). This result, however, does not necessarily imply that the total uncertainty of A_{BS} will be the same for subsequent even- and odd-ordered polynomials. For example, if the background data are not fit well by a polynomial of order M_{Even} , the total uncertainty of A_{BS} will be dominated by the systematic error of the poor fit, not by the statistical uncertainty resulting from the RBSVs plotted in Fig. 6. Hence, the total uncertainty would likely be improved by increasing the polynomial order to $M_{Even} + 1$. If, on the other hand, the background is well described by a polynomial of order M_{Even} , increasing the order to $M_{Even} + 1$ will not improve the statistical uncertainty and may risk increasing the systematic uncertainty.

Table B.3 provides approximate expressions for $RBSV_M(n_S)$. Using these formulae, we find the empirical relationship $RBSV_{M_{Even}}(n_S) \cong RBSV_2(n_S)^{(M_{Even}/2)/1.06^{(M_{Even}-2)}}$ is accurate to within $\pm 15\%$ for $M_{Even} = 4, 6, \text{ or } 8$ and $0.5 < n_S < 5$. Although the accuracy of this approximation is not as good as that for the formulae in Table B.3, it demonstrates that the statistical uncertainty of the peak area near the detection limit increases exponentially with the highest even order of the polynomial used to fit the background. It is, therefore, not

advisable to increase the order of the polynomial fit to the background data beyond the lowest value that provides a reasonable fit. To do so would only increase the statistical uncertainty of the peak area and may also risk increasing the systematic uncertainty as discussed above.

The horizontal dashed lines in Fig. 6 indicate the RBSV for a linear background determined by symmetric, uncorrelated endpoint averaging with $n_L = n_R = 0.2$ (i.e., the marginal conditions under which a linear background could be used for a peak consisting of 10 points on a background that is linear over just two endpoints on each side). The logarithmic plots of $RBSV_M(n_L, n_R)$ in Fig. 6 demonstrate that the statistical uncertainty resulting from a LSF-polynomial background will be many times larger than this marginal linear example unless the background is well approximated by a low-order polynomial over an energy range larger than E_P on both sides of the peak (i.e., $n_L > 1$ and $n_R > 1$). For example, σ_{ABS} for subtraction of a polynomial background from a peak spanning $N_P = 10$ points is only lower than the marginal linear two-endpoint case if the background can be well fit by a polynomial of the following orders, M , and numbers of background points, N_S , on *each side* of the peak: $M=2$ or 3 over $N_S = 9$ points ($n_S = 0.9$), $M=4$ or 5 over $N_S = 16$ points ($n_S = 1.6$) or $M=6$ or 7 over $N_S = 22$ points ($n_S = 2.2$). To generalize this point, Fig. 7 plots the ratio of the number of symmetric background points to peak points, $n_S = n_{Poly}$, that must be included in a LSF-polynomial background of order M_{Even} to produce an RBSV equal to that for an endpoint-averaged linear background with $n_S = n_{Lin}$. The filled circles are the result of direct calculations using the techniques described in Appendix B, and the curves are given by the empirical expression determined by parametric fitting of the points to the form

$$n_{Poly} = (a_1 + b_1 M_{Even}) + (a_2 + b_2 M_{Even}) n_{Lin}^p \quad (13)$$

$$\text{with } p = 0.851, \text{ and } \begin{matrix} a_1 = -0.137, & a_2 = 1.07 \\ b_1 = 0.175, & b_2 = 0.745 \end{matrix}$$

The value of n_{Poly} for an odd order, $M_{Even} + 1$, is the same as the value for M_{Even} .

Figure 7 shows that a prohibitive number of points would be needed in most cases for a LSF-polynomial background to have a lower statistical uncertainty than a linear background using just a few points. A polynomial, however, would be appropriate for an isolated peak on a smoothly varying background that extends several multiples of the peak width on both sides of the peak. If the relative number of suitable background points for polynomial LSF is limited to n_{Poly} , it may be more useful to solve Eq. (13) for M_{Even} to determine the maximum polynomial order that would yield an RBSV no larger than the linear-endpoint-averaging case using n_{Lin} .

Finally, it is important to distinguish between the polynomial used for background subtraction, and the polynomial that may be used to determine the variance of the average background signal, σ_{yB} , used in Eqs. (4) and (10). If the measured signal is not described by Poisson statistics (e.g., due to the application of a transmission function or averaging by

multiple detectors), $\sigma_{\langle y_B \rangle}$ may be approximated by the goodness of fit parameter, χ , for a polynomial fit to a contiguous set of smoothly varying background points near the peak.¹⁴ The choice of polynomial order and number of points to include in this fit should be determined by trial and error to find the lowest χ for a representative set of background data. The relationships discussed above between the RBSV and the polynomials used for background subtraction have no bearing on the determination of $\sigma_{\langle y_B \rangle}$.

Optimal measurement strategies for small peaks

The preceding section describes optimal analysis methods for minimizing the peak-area uncertainty from a previously measured XPS spectrum based on the available distribution of background points which may be limited by features inherent to the spectrum or by the acquired energy range. In this section we determine the acquisition parameters that will minimize the total measurement time needed to achieve a desired level of sensitivity for a specific method of background subtraction. In subsection (A) the sensitivity is defined in terms of the smallest measurable atomic concentration that is statistically different from zero (i.e., the detection limit), and in subsection (B) it is generalized to the uncertainty of a low-level ($\lesssim 10\%$) atomic-concentration.

As in Eq. (1), we use the subscript q to index the background-subtracted areas, A_q , of all measured XPS peaks. The specific peak corresponding to a trace element of interest is indicated by $q = x$. The relationship between the standard uncertainties of the background-subtracted area, $\sigma_{A_{BS,x}}$, and the resulting atomic concentration, σ_{C_x} , is obtained by propagation of error through Eq. (1).^{7,8} The atomic concentrations of the measured species are correlated due to the sum of RSF-normalized peak areas in the denominator of Eq. (1); however, for samples consisting of a substrate with only trace amounts of other elements totaling $\lesssim 10$ atomic percent, this sum will be dominated by the areas of the substrate peaks, and the dependence on the peak areas of the trace elements can be neglected. In this case, the relative standard uncertainties of the atomic concentration and of the background-subtracted peak area for the trace element x [as defined by Eq. (2)] are approximately equal:

$$\left(\frac{\sigma_{C_x}}{C_x} \right)^2 \cong \left(\frac{\sigma_{A_{BS,x}}}{A_{BS,x}} \right)^2. \quad (14)$$

This approximation holds for the condition

$$\left(\frac{\sigma_{A_{BS,x}}}{A_{BS,x}} \right)^2 \gg \sum_{q \neq x} \left(C_q \frac{\sigma_{A_{BS,q}}}{A_{BS,q}} \right)^2, \quad (15)$$

where it has been assumed that $C_x \ll 1$, and the sum is over all measured species except x . For a measurement consisting of a few strong substrate peaks and a single trace element, this condition is easily satisfied.

Equations (14) and (15) are based on the assumption that the peak areas for different species are not correlated. For samples with a trace element on a substrate that is then covered with an overlayer, the peak areas of the substrate, trace element and overlayer become strongly correlated as do their uncertainties. This situation is quite common for air-exposed samples with an overlayer of adventitious carbon. If the overlayer attenuates each elemental peak by a similar proportion, the correlation can be avoided by excluding the overlayer species from the atomic-percentage calculation.

(A) Species present at or near the detection limit

The uncertainty of the atomic concentration of a species present at trace levels can be estimated using Eqs. (10) and (14) and the RBSV appropriate to the background-subtraction method. This approach provides a way to relate the acquisition parameters to the detection limit, which has been defined in many ways.^{4,1214} Generally speaking, the detection limit for a measured quantity such as peak area, A_{DL} , is described as some multiple of the standard deviation of the background-subtracted area, σ_{ABS} . The ISO Technical Committee on Surface Chemical Analysis cites the common practice of $A_{DL}=3\sigma_{ABS}$.⁵ While this is adequate for many situations, it does not contain any information on the confidence level of the measurement. Here we use the more complete definition $A_{DL}=2k\sigma_{ABS}$ proposed by Currie^{3,4} for normally distributed errors with constant σ_{ABS} and coverage factor k . The factor of two is the result of requiring that the probabilities for *both* false positives *and* false negatives are equal and given by one minus the integral of the standard normal distribution from 0 to $k+1$. With this definition and Eq. (10) in the limit of zero peak area (and hence zero intrinsic peak-area variance, $\sigma_{AP,x}$), the background-subtracted peak area at the limit of detection for species x is given by

$$A_{DL,x} = 2k \lim_{\sigma_{AP,x} \rightarrow 0} (\sigma_{ABS,x}) = 2k \sigma_{yB,x} \sqrt{N_{P,x} RBSV_x}. \quad (16)$$

The appropriate value for the coverage factor, k , will depend upon the fidelity required for a given measurement. For example, if the location of the peak is known, the probability that a peak will be incorrectly detected when it is not present is $p_0=5\%$ for $k=1.65$.³ If the peak location is uncertain such that the peak is equally likely to appear in one of $\mathcal{N}$ resolvable locations (and not in the $\mathcal{N}-1$ others), the probability of erroneous detection increases to $\mathcal{N} p_0 (1-p_0)^{\mathcal{N}-1}$. To retain the same overall false-detection probability of 5% when $\mathcal{N}=5$, p_0 must be reduced to 1% by increasing k to 2.33. Note that a more detailed analysis is required if the region of potential peak locations extends far into the energy range used for background determination. Under these conditions the peak-area uncertainty may vary significantly with peak location due to the dependence of the appropriate RBSV on endpoint asymmetry. This variation is usually much larger for polynomial backgrounds than for linear backgrounds. (See Fig. B.1 and discussion in Appendix B.)

For the common case of Poisson-distributed statistics, the variance of the background signal is $\sigma_{yB,x}^2 = dt_x \hat{\beta}_x$, where $\hat{\beta}_x$ is the average background count rate (counts per second), and

the caret indicates that the counts are unscaled. In this case, the detection limit of atomic concentration can be expressed using Eqs. (1), (2), (16) and $E_{P,x} = N_{P,x} dE_x$ as

$$C_{DL,x} = \frac{A_{DL,x}(dE_x/dt_x)}{Z} = \frac{2k}{R_x Z} \sqrt{RBSV_x \hat{\beta}_x \Delta E_{P,x} dE_x/dt_x} \quad (17)$$

$$Z = \frac{I_{sub}}{C_{sub} R_{sub}} \quad (18)$$

Here Z represents the denominator of Eq. (1) which is the sum of the RSF-normalized intensities of all the species included in the atomic-percentage calculation. By inverting Eq. (1), Eq. (18) gives Z in terms of the measurable peak intensity, I_{sub} (counts per second times electron volts), and atomic concentration of the substrate, C_{sub} .

Equations (17) and (18) give the detection limit for Poisson-distributed uncertainties as a function of: the acquisition parameters (dE_x , dt_x); characteristics of the spectra and sample that are measured (I_{sub} , C_{sub} , $\hat{\beta}_x$), known (R_x , R_{sub}), and assumed ($E_{P,x}$); the desired fidelity of the detection limit (k); and the variance due the method of background determination normalized to the intrinsic variance of the background area ($RBSV_x$). In this context, we define the *achievable* detection limit as the smallest atomic concentration that can be considered statistically different from zero when a specific background-subtraction method is applied to a measured spectrum (e.g., automated measurements of atomic concentration.) This quantity may differ from the *ultimate* detection limit of an acquired spectrum corresponding to the smallest atomic concentration capable of producing a peak that is detectable above the background by eye¹² or by a peak-detection algorithm.^{13,17} Derivation of the RBSV for such peak-detection methods is beyond the scope of this work.

In addition to estimating the detection limit for a previously acquired set of data, Eqs. (17) and (18) can be inverted to determine the dwell time, $dt_{DL,x}$, needed to achieve a desired detection limit, $C_{DL,x}$. From this result we also determine the total spectrum acquisition time, $t_{DL,x}$, needed to acquire the total number of peak and symmetric background points.

$$dt_{DL,x} = \frac{RBSV_x}{N_{P,x}} dt_{(A_{DL} = \sigma_B),x} \quad (19)$$

$$t_{DL,x} = N_{P,x} (1 + 2n_{S,x}) dt_{DL,x} = RBSV_x (1 + 2n_{S,x}) dt_{(A_{DL} = \sigma_B),x} \quad (20)$$

$$\text{with } dt_{(A_{DL}=\sigma_B),x} = \hat{\beta}_x \left(\frac{2k \Delta E_{P,x}}{C_{DL,x} R_x Z} \right)^2 \quad (21)$$

Here $dt_{(A_{DL}=\sigma_B),x}$ is the dwell time at which the peak area corresponding to the detection limit equals $2k$ times the uncertainty of the total background area under the peak. This characteristic dwell time is determined only by the desired detection limit and the characteristics of the spectrum, such as the background count rate, $\hat{\beta}_x$, and the energy range spanned by the peak, $E_{P,x}$, but it is independent of acquisition and analysis parameters. The prefactor in Eq. (19) indicates that the dwell time $dt_{DL,x}$ depends on both the RBSV and the energy step size of the acquisition (through $N_{P,x} = E_{P,x}/dE_x$); however, the total acquisition time in Eq. (20) depends only on the background-determination method through the prefactor $RBSV_x(1+2n_{S,x})$ and is independent of the acquisition parameters. The latter dependence occurs because the detection limit is determined by the total number of counts collected under the peak during the entire acquisition, regardless of how they are accumulated (e.g., many energy steps with short dwell times or vice versa). Hence, the strategy to minimize total acquisition time is determined solely by the background-subtraction method that will minimize the prefactor of $dt_{(A_{DL}=\sigma_B),x}$ in Eq. (20).

To illustrate this result, Fig. 8(a) shows the dependence of the total acquisition time on the relative number of symmetric endpoints used to determine a linear background by endpoint averaging. The curve for the uncorrelated case indicates that a desired detection limit can be achieved 15 % faster by acquiring fewer background points ($n_{S,x} = 0.5$, i.e., background containing half the number of peak points on each side) for a longer time compared to the conditions that minimize peak-area uncertainty without regard to acquisition time [$n_{S,x} = 1$ from Eq. (11), i.e., background with the same number of points as the peak on each side]. For the semi-correlated case, the minimum acquisition time at $n_{S,x} = 0.39$ is 28 % smaller than the $n_{S,x} = 1$ case.

Equation (20) is also plotted in Fig. 8 (b) for LSF of symmetric endpoints (uncorrelated with peak area) to quadratic and quartic backgrounds using the approximate RBSVs from Table B.3. This plot emphasizes the point made previously that the uncertainty of the peak-area measurement is greatly increased by the additional degrees of freedom introduced by polynomial fits. Even for the optimal case of $n_{S,x} \cong 1.5$ for quadratic ($n_{S,x} \cong 2.5$ for quartic) backgrounds, the minimum acquisition time is more than twice (more than three times for quartic) as long as that for a linear background with $n_{S,x} = 0.5$ to achieve the same detection limit.

(B) Detectable species with atomic concentrations below $\approx 10\%$

In this section we consider a peak with atomic concentration greater than the detection limit but low enough that its relative uncertainty is approximately independent of the concentrations of the substrate and any other species so that Eq. (14) is valid (typical for atomic concentrations $\leq 10\%$). This approach is consistent with the assumption that the corresponding peak area is small enough that a linear or polynomial background is

appropriate, rather than a peak-dependent form like the Shirley or Tougaard backgrounds. In this case, both terms of Eq. (10) may contribute significantly to the total uncertainty of the background-subtracted peak area. It would be counterproductive to use more background points than are necessary to decrease the RBSV below the value at which the variance due to the background determination [second term in Eq. (10)] is equal to the intrinsic peak area variance [first term in Eq. (10)]. Expanding the background region to include more points than this would provide little reduction in the total statistical uncertainty of the peak area, and it may increase the systematic uncertainty if the additional points are not consistent with the assumed background form. Equating the two terms of Eq. (10), we see that this condition of diminishing returns will be avoided if the RBSV remains greater than the ratio of the variances of the intrinsic areas for the peak, $\sigma_{A_{P,x}}^2$, and the background under the peak, $N_{P,x}\sigma_{(Y_{B,x})}^2$. For Poisson counting statistics, $\sigma_{A_{P,x}}^2 = \hat{A}_{P,x}$ and $N_{P,x}\sigma_{(Y_{B,x})}^2 = N_{P,x}\hat{\beta}_x dt_x$, which is the area under the peak, $\hat{B}_x$. The critical value for the RBSV 2 for species x is then

$$RBSV_{crit,x} \equiv \hat{A}_x / \hat{B}_x. \quad (22)$$

This result indicates that the statistical uncertainties of *both* the background and intrinsic peak areas contribute significantly to $\sigma_{ABS,x}$ when the associated RBSV is comparable to the ratio of the peak and background areas, which is independent of the acquisition parameters. Equation (22), therefore, determines whether a peak is in the large-peak-area regime or not: if $RBSV_x \gg \hat{A}_x / \hat{B}_x$, then $\sigma_{A_{P,x}}$ may be neglected and the methods of the previous section are applicable; if $RBSV_x \ll \hat{A}_x / \hat{B}_x$, then the peak-area uncertainty is dominated by $\sigma_{A_{P,x}}$ and large-peak-area approximation is valid.

The practical consequences of Eq. (22) are evident when applied to the exponential growth of the RBSV with the order of a LSF-polynomial background. $\sigma_{ABS,x}$ can be dominated by the uncertainty of the background-subtraction method for much larger peaks on a polynomial background than on a linear background. For example, for symmetric endpoints equal to the number of peak points ($n_{S,x} = 1$), $\sigma_{ABS,x}$ will be limited by the uncertainty of the LSF-polynomial background of order M when $\hat{A}_x / \hat{B}_x$ is < 1.5 for $M=1$, < 3 for $M=2$ or 3 , < 7 for $M=4$ or 5 , or < 20 for $M=6$ or 7 . In other words, for a fourth-order polynomial background, $\sigma_{ABS,x}$ can be significantly reduced by increasing the number of background points beyond N_P even for peak areas that are up to seven times the background area under the peak. While for linear-LSF backgrounds, using more than N_P background points will only significantly reduce $\sigma_{ABS,x}$ for peaks with areas up to 1.5 times the background area.

The dwell and total acquisition times needed to achieve a desired uncertainty of atomic concentration for species x , are found by solving Eqs. (10) and (14) with the variances for Poisson counting statistics $\sigma_{A_{P,x}}^2 = \hat{A}_{P,x} = C_x R_x Z dt_x / dE_x$ and $\sigma_{(Y_{B,x})}^2 = \hat{\beta}_x dt_x$:

$$dt_{\sigma_{C_x}} = \frac{RBSV_{crit,x} + RBSV_x}{N_{P,x}} dt_{(A_{\sigma_C} = \sigma_B),x} \quad (23)$$

$$t_{\sigma_{C_x}} = N_{P,x} (1 + 2n_{S,x}) dt_{\sigma_{C_x}} = \quad (24)$$

$$(RBSV_{crit,x} + RBSV_x) (1 + 2n_{S,x}) dt_{(A_{\sigma_C} = \sigma_B),x}$$

$$\text{with } dt_{(A_{\sigma_C} = \sigma_B),x} = \hat{\beta}_x \left(\frac{\Delta E_{P,x}}{\sigma_{C_x} R_x Z} \right)^2 \quad (25)$$

Here $dt_{(A_{\sigma_C} = \sigma_B),x}$ is the dwell time for which the area (counts) corresponding to the desired atomic-concentration uncertainty, σ_{C_x} , equals the uncertainty of the total background area (counts) under the peak. Comparison of Eqs. (20) and (21) with Eqs. (24) and (25) reveals that the total acquisition time needed to achieve a specific σ_{C_x} will be a factor of $(1 + RBSV_{crit,x}/RBSV_x)$ longer than the time to achieve a detection limit equal to $2k\sigma_{C_x}$. Together with Eq. (22), this result indicates that for measurements of peaks with $\hat{A}_x/\hat{B}_x \ll RBSV_x$, Eqs. (19) through (21) with the substitution $C_{DL,x} \rightarrow 2k\sigma_{C_x}$ can be used to estimate the acquisition times to produce a desired σ_{C_x} . The more complicated Eqs. (23) through (25) are required only when the ratio of peak to background areas is comparable to the RBSV.

The total acquisition time, $t_{\sigma_{C_x}}$, given by Eq. (24) will depend on the selection of background points in a similar way to the $t_{DL,x}$ predicted by Eq. (20) and displayed in Fig. 8 for LSF of linear and polynomial backgrounds; however, the $RBSV_{crit,x}$ term in Eq. (24) will shift the minima for $t_{\sigma_{C_x}}$ to slightly lower values of $n_{S,x}$ than those for $t_{DL,x}$. For cases where the ratio of peak to background areas is less than half the RBSV ($RBSV_{crit,x}/RBSV_x < 0.5$), the $t_{\sigma_{C_x}}$ predicted by Eq. (24) by using the optimal $n_{S,x}$ from Eqs. (20) and (21) with $C_{DL,x} = 2k\sigma_{C_x}$ will be within 10 % of the true acquisition-time minimum.

SUMMARY

We analyze the statistical uncertainty of common procedures for obtaining background-subtracted peak areas from XPS spectra. These results prescribe optimal analysis and acquisition strategies for measurements of species present near the limit of detection. Unlike previous studies, our analysis includes the correlation of points used to determine both the peak and the background areas. For peak areas with variances comparable to that of the background area (i.e., near the detection limit), the uncertainties can be significantly overestimated if this correlation is neglected. The default algorithms of many commercial instruments and analysis-software packages introduce this correlation.

1. We express the uncertainties, detection limits and optimal measurement conditions in a form that separates the dependence on acquisition parameters, spectrum properties and the method used to determine the background. The latter factor is quantified by the relative background-subtraction variance (RBSV) defined as the ratio of the variance introduced by the background-determination

method to the intrinsic variance of the background area. The RBSV has the following properties:

- a.** It depends only on the ratios of the numbers of points used to define the background to the number of points comprising the peak. These ratios are equivalent to the ratios of the energy ranges spanned by the background regions to the energy range of the peak.
 - b.** It is independent of the acquisition parameters (e.g., energy step size and dwell time) and the spectrum attributes (e.g., peak width and intensities of the peak and background).
 - c.** The total peak-area uncertainty is dominated by the uncertainty of the background when the RBSV is much greater than ratio of the background-subtracted peak area to the background area under the peak. When the RBSV is much smaller than the peak-to-background-area ratio, the uncertainty is dominated by the intrinsic variance of the peak area (i.e., the large-peak-area approximation is valid).
 - d.** The background-subtracted peak-area uncertainty and the detection limit are proportional to the square root of the RBSV.
 - e.** The total acquisition time required to achieve a specific detection limit is proportional to the RBSV; therefore, the method of background subtraction that will be used should be considered when planning a measurement.
- 2.** We derive the RBSVs and determined optimal analysis strategies to minimize the uncertainty for several common background-subtraction methods:
 - a.** Linear backgrounds determined by averaging *symmetric* endpoints (i.e., equal number of background points on each side of peak)
 - i.** For the common case of endpoints that are semi-correlated with the peak area (i.e., equal numbers of background points inside and outside the peak analysis region), the background-subtracted peak-area uncertainty is minimized when the number of points used for background determination is equal to the number of points comprising the peak. Using more background points than this will increase the total peak-area uncertainty.
 - ii.** For the less common case of endpoints with no peak-area correlation (i.e., the peak-analysis region includes only peak points and no background points), the uncertainty is always lower than the semi-correlated case and decreases monotonically with increasing background points to a value 40 % lower than the minimum for the semi-correlated case.
 - b.** Linear backgrounds determined by endpoint averaging of *asymmetric* endpoints.

- i. Asymmetric endpoints may be useful when the ratio of background to peak points is limited to n_{Small} on one side of the peak but may extend to a larger value, n_{Large} , on the other.
 - ii. We determine the optimal n_{Large} for a given n_{Small} to minimize background-subtracted peak-area uncertainty. Using more background points than this will increase the total peak-area uncertainty.
 - iii. If the number of available background points on one side is <10 % of the number of peak points ($n_{Small} < 0.1$), using asymmetric endpoints can significantly reduce the RBSV compared to using n_{Small} on both sides. This improvement decreases dramatically for $n_{Small} > 0.2$.
 - c. Linear and polynomial backgrounds determined by least-squares fitting (LSF) of endpoints that are uncorrelated with the peak area
 - i. We provide expressions to approximate the RBSVs for LSF of asymmetric endpoints. For odd-ordered polynomial backgrounds, the RBSV is always lowest for symmetric endpoints. For even orders the RBSV for asymmetric endpoints can be lower than the symmetric case over a certain range of endpoint asymmetry.
 - ii. For $n_{Large} > 20 n_{Small}$ (i.e., 20 times more background points on one side of the peak than the other), we show that the uncertainty for linear LSF is significantly lower than that for endpoint averaging. For $n_{Large} < 20 n_{Small}$ the difference between the two methods is <20 %.
 - iii. We demonstrate that the RBSV increases approximately exponentially with the polynomial order of the background.
 - iv. We provide methods to determine the conditions under which the peak-area uncertainty for a polynomial background of arbitrary order will be comparable to or less than that for a linear background.
- 3. We present a general method for an arbitrary RBSV that determines the acquisition parameters needed to minimize total measurement time and attain a specific detection limit or atomic-percentage uncertainty for a species present at trace levels.
 - a. The optimal number of background points to acquire on each side of the peak is significantly less than that needed to minimize peak-area uncertainty of a previously acquired spectrum with arbitrary dwell time.
 - b. The total measurement time needed to attain a specific sensitivity increases exponentially with the order of the polynomial background for a fixed number of background points. For example, if the numbers

of background points on each side are equal to the number of points comprising a peak near the detection limit, the measurement times needed to achieve the same level of sensitivity for polynomial backgrounds of orders 2, 4, and 6 must be increased by factors of 2, 5, and 14, respectively, compared to that for a linear background.

Acknowledgments

The authors would like to thank Dr. Alex Shard for helpful discussions on the practical application of detection limits and their corresponding confidence intervals.

References

1. Powell CJ, Jablonski A. *J Electron Spectrosc Relat Phenom*. 2010; 178–179:331–46.
2. Faradzhev NS, Hill SB, Powell CJ. *Surf Interface Anal*. 2017; 49:1214–24.
3. Currie LA. *Anal Chem*. 1968; 40:586–593.
4. Currie LA. *Anal Chim Acta*. 1999; 391:127–134.
5. ISO 18115-1, Surface Chemical Analysis – Vocabulary, Part 1 – General Terms and Terms Used in Spectroscopy. International Organization for Standardization; Geneva: 2013.
6. Powell CJ, Seah MP. *J Vac Sci Technol Vac Surf Films*. 1990; 8:735.
7. Harrison K, Hazell LB. *Surf Interface Anal*. 1992; 18:368–76.
8. Evans S. *Surf Interface Anal*. 1992; 18:323–32.
9. Cumpson PJ, Seah MP. *Surf Interface Anal*. 1992; 18:345–60.
10. Cumpson PJ, Seah MP. *Surf Interface Anal*. 1992; 18:361–7.
11. Powell CJ, Conny JM. *Surf Interface Anal*. 2009; 41:269–94.
12. Suzuki M, Fukushima S, Tanuma S. *Surf Interface Anal*. 2008; 40:1337–43.
13. Suzuki M, Fukushima S, Tanuma S. *Surf Interface Anal*. 2013; 45:1174–81.
14. Shard AG. *Surf Interface Anal*. 2014; 46:175–85.
15. Atzei D, Elsener B, Fantauzzi M, Brundu F, Rossi A. *Spectrochim Acta Part B At Spectrosc*. 2016; 121:38–46.
16. Seah MP. *Surf Interface Anal*. 2005; 37:300–9.
17. Furukawa Y, Otsuki M, Ikeo N, Nagatsuka Y, Yoshitake M, Tanaka A. *J Surf Anal*. 2016; 23:73–97.
18. Végh J. *J Electron Spectrosc Relat Phenom*. 2006; 151:24–30.
19. Richter PH. *Telecommun Data Acquis Rep*. 1995; 42:107–37.
20. Mathews, J., Walker, RL. *Mathematical methods of physics*. 2. WA Benjamin; New York: 1970.

Appendix A

We derive the uncertainty, σ_{ABS} of the background-subtracted peak area, A_{BS} , for a linear background defined by the averaging of asymmetric numbers of endpoints on either side of a peak from a single species as illustrated in Fig. 1. Since the subscripts x and q are not needed to distinguish different species, we use subscripts i and j in both Appendices to index the individual points in the measured spectrum. With this notation A_{BS} is defined as the sum of the counts, y_i , measured at energies, E_i , within the analysis range, E_B , minus the value of the background function, $f_B(E_i)$. This choice provides a simple definition of peak area that does not depend on a particular numerical integration technique. It also avoids the confusion cited by Vegh¹⁸ when spectra are misinterpreted as energy distributions. To keep track of which points are used to determine only the background, only the peak or both, we use the

five contiguous, non-overlapping sets of spectrum points indexed with the label Λ which takes on values LB, LA, P, RA, RB as described in Table 2 and shown in Fig. 1. Each set contains N_Λ spectrum points with coordinates denoted as $(E_{\Lambda,i}, y_{\Lambda,i})$. Using this notation

$$A_{BS} = \sum_{i=1}^{N_{LA}} [y_{LA,i} - f_B(E_{LA,i})] + \sum_{i=1}^{N_P} [y_{P,i} - f_B(E_{P,i})] + \sum_{i=1}^{N_{RA}} [y_{RA,i} - f_B(E_{RA,i})]. \quad (\text{A.1})$$

The commonly assumed definition of a linear background determined by endpoint averaging is

$$f_B(E) = aE + b, \quad (\text{A.2})$$

where

$$a = \frac{\langle y \rangle_L - \langle y \rangle_R}{\langle E \rangle_L - \langle E \rangle_R} \text{ and } b = \frac{\langle E \rangle_L \langle y \rangle_R - \langle E \rangle_R \langle y \rangle_L}{\langle E \rangle_L - \langle E \rangle_R}$$

Here $\langle E \rangle$ and $\langle y \rangle$ are the average energy and counts, respectively, of all the background points on the left or right side of the peak indicated by the subscript L or R respectively:

$$\begin{aligned} \langle E \rangle_L &= \frac{1}{N_{LB} + N_{LA}} \left(\sum_{j=1}^{N_{LB}} E_{LB,j} + \sum_{j=1}^{N_{LA}} E_{LA,j} \right) \\ \langle E \rangle_R &= \frac{1}{N_{RB} + N_{RA}} \left(\sum_{j=1}^{N_{RB}} E_{RB,j} + \sum_{j=1}^{N_{RA}} E_{RA,j} \right) \end{aligned} \quad (\text{A.3})$$

Corresponding expressions for $\langle y \rangle_L$ and $\langle y \rangle_R$ are obtained by replacing $E_{\Lambda,j}$ with $y_{\Lambda,j}$ in Eq. (A.3).

To determine the statistical uncertainty of A_{BS} , we assume that each $y_{\Lambda,i}$ represents an independent sampling of an arbitrary statistical distribution with standard deviation $\sigma_{y_{\Lambda,i}}$ at an exact energy $E_{\Lambda,i}$. In this case, the standard deviation of A_{BS} is calculated by first-order error propagation

$$\begin{aligned} (\sigma_{A_{BS}})^2 = & \sum_{i=1}^{N_{LB}} \left(\frac{\partial A_{BS}}{\partial y_{LB,i}} \right)^2 (\sigma_{y_{LB,i}})^2 + \sum_{i=1}^{N_{LA}} \left(\frac{\partial A_{BS}}{\partial y_{LA,i}} \right)^2 (\sigma_{y_{LA,i}})^2 + \\ & \sum_{i=1}^{N_P} \left(\frac{\partial A_{BS}}{\partial y_{P,i}} \right)^2 (\sigma_{y_{P,i}})^2 + \sum_{i=1}^{N_{RA}} \left(\frac{\partial A_{BS}}{\partial y_{RA,i}} \right)^2 (\sigma_{y_{RA,i}})^2 + \sum_{i=1}^{N_{RB}} \left(\frac{\partial A_{BS}}{\partial y_{RB,i}} \right)^2 (\sigma_{y_{RB,i}})^2 \end{aligned} \quad (\text{A.4})$$

The partial derivatives can be expressed using Eqs. (A.1) and (A.2) as

$$\begin{aligned} \frac{\partial A_{BS}}{\partial y_{\Lambda,i}} = & \frac{\partial}{\partial y_{\Lambda,i}} \left(\sum_{j=1}^{N_{LA}} y_{LA,j} + \sum_{j=1}^{N_P} y_{P,j} + \sum_{j=1}^{N_{RA}} y_{RA,j} \right) - \\ & \left[\frac{\partial b}{\partial y_{\Lambda,i}} (N_{LA} + N_P + N_{RA}) + \frac{\partial a}{\partial y_{\Lambda,i}} \left(\sum_{j=1}^{N_{LA}} E_{LA,j} + \sum_{j=1}^{N_P} E_{P,j} + \sum_{j=1}^{N_{RA}} E_{RA,j} \right) \right] \end{aligned} \quad (\text{A.5})$$

with

$$\begin{aligned} \frac{\partial a}{\partial y_{\Lambda,i}} = & \frac{1}{\langle E \rangle_L - \langle E \rangle_R} \begin{cases} (N_{LB} + N_{LA})^{-1}, & \text{for } \Lambda = LA \text{ or } LB \\ 0, & \text{for } \Lambda = P \\ (N_{RB} + N_{RA})^{-1}, & \text{for } \Lambda = RA \text{ or } RB \end{cases} \quad (\text{A.6}) \\ \frac{\partial b}{\partial y_{\Lambda,i}} = & - \frac{\partial a}{\partial y_{\Lambda,i}} \begin{cases} \langle E \rangle_R, & \text{for } \Lambda = LA \text{ or } LB \\ 0, & \text{for } \Lambda = P \\ \langle E \rangle_L, & \text{for } \Lambda = RA \text{ or } RB \end{cases} \end{aligned}$$

The partial derivative of the three sums in the first line of Eq. (A.5) is trivial since $y_{\Lambda_1,i}$ $y_{\Lambda_2,j}$ equals 1 only if both $\Lambda_1 = \Lambda_2$ and $i = j$. It equals 0 otherwise because the $y_{\Lambda,i}$ are independent. Since the energies of each set, $E_{\Lambda,j}$, are equally spaced by the same dE , the sums over the energies in Eqs. (A.5) and (A.3) reduce to simple arithmetic series multiplied by dE . Upon substitution into Eqs. (A.5) and (A.6), the dE factors cancel, and $\sigma_{A_{BS}}$ is independent of dE .

Using Eqs. (A.5) and (A.6), the sums in Eq. (A.4) all reduce to sums over the variances of the signals, $\sigma_{y_{\Lambda,i}}^2$, in each set. For a slowly varying background, it is reasonable to approximate the standard deviations of the signals by σ_{y_B} , the mean standard deviation of all background points: $\sigma_{y_{\Lambda,i}} = \sigma_{y_B}$ for $\Lambda = LB, LA, RA, RL$ and all i . With this approximation, the sums over the four background sets become trivial. The sum over the variances of the peak points is the variance of the total area spanning E_P (prior to background subtraction) which can be divided into the separate sums of the background and peak areas. For small peaks, it is reasonable to assume that these two components of the total area are not correlated. (This assumption would not be appropriate for larger peaks requiring

a peak-dependent background such as the Shirley form.) The variance of their sum can therefore be expressed as the sum of the intrinsic variances of the background-subtracted peak area, σ_{AP}^2 , and of the background area spanning the peak range, $N_P \sigma_{YB}^2$. This approach gives $\sum_{i=1}^{N_P} \sigma_{y_{P,i}}^2 = \sigma_{AP}^2 + N_P \sigma_{YB}^2$. With this simplification, the total variance of A_{BS} reduces to

$$\sigma_{A_{BS}}^2 = \sigma_{AP}^2 + \frac{N_T(N_P + N_{LA} + N_{RA})(N_P + N_{LB} + N_{RA})(N_P + N_{LA} + N_{RB})(N_P + N_{LB} + N_{RB})}{(N_{LA} + N_{LB})(N_{RA} + N_{RB})(2N_P + N_T)^2} \sigma_{YB}^2. \quad (\text{A.7})$$

Here $N_T = N_{LA} + N_{LB} + N_{RA} + N_{RB}$ is the total number of all background points. Equation (A.7) reduces Eq. (4) for the special case of symmetric endpoints with $N_{LA} = N_{RA} = N_{SA}$, $N_{LB} = N_{RB} = N_S - N_{SA}$, $N_T = 2N_S$ and $N_P = N_A - 2N_{SA}$.

To determine the RBSV for the general case, we express Eq. (A.7) in terms of the ratio n_Λ of the number of background to peak points for the different sets by substituting $N_\Lambda = n_\Lambda N_P$. The RBSV for endpoint averaging is then given by Eq. (7) as

$$RBSV_{EA}(n_{LB}, n_{RB}, n_{LA}, n_{RA}) = \frac{n_T(1 + n_A)(1 + n_T - n_A)(1 + n_{RB} + n_{LA})(1 + n_{LB} + n_{RA})}{(n_{LB} + n_{LA})(n_{RB} + n_{RA})(2 + n_T)^2} \quad (\text{A.8})$$

with $n_T = n_{LA} + n_{LB} + n_{RA} + n_{RB}$ and $n_A = n_{LA} + n_{RA}$. For spectra with Poisson-distributed statistics, Figs. 2 (a–c) demonstrate that the variances predicted by Eq. (A.7) agree with those resulting from Monte Carlo simulations for symmetric endpoints that are semi-correlated and uncorrelated with the peak area. Although not shown here, Eq. (A.7) was also found to agree with simulated results for asymmetric endpoints with arbitrary peak-area correlation.

Appendix B

We derive the statistical uncertainty, $\sigma_{A_{BS}}$, of the background-subtracted peak area, A_{BS} , for a background determined by least-squares fitting (LSF) of a polynomial of order M to asymmetric endpoints. Since $\sigma_{A_{BS}}$ is always larger for endpoints with any degree of peak-area correlation, we avoid the increased complexity and consider only endpoints with no

peak-area correlation. In this case $N_{LA} = N_{RA} = 0$, and the set-index subscripts LB and RB are replaced by L and R , respectively, as described in Table 2. A_{BS} is then given by the middle sum over N_P in Eq. (A.1), and the background function is defined by

$$f_B(E) = \sum_{k=1}^{M+1} g_k E^{k-1}. \quad (\text{B.1})$$

Using the same notation described in Appendix A, the variance of A_{BS} is defined in terms of the variance of the measured points, $\sigma_{y_{\Lambda,i}}^2$, by first-order error propagation [as in Eq. (A.4)]

$$\begin{aligned} (\sigma_{A_{BS}})^2 = & \sum_{i=1}^{N_L} \left(\frac{\partial A_{BS}}{\partial y_{L,i}} \right)^2 (\sigma_{y_{L,i}})^2 + \sum_{i=1}^{N_P} \left(\frac{\partial A_{BS}}{\partial y_{P,i}} \right)^2 (\sigma_{y_{P,i}})^2 + \sum_{i=1}^{N_R} \left(\frac{\partial A_{BS}}{\partial y_{R,i}} \right)^2 (\sigma_{y_{R,i}})^2 = \\ & \sum_{i=1}^{N_L} \left(\sum_{j=1}^{N_P} \sum_{k=1}^{M+1} \frac{\partial g_k}{\partial y_{L,i}} E_{P,j}^{k-1} \sigma_{y_{L,i}} \right)^2 + \sum_{i=1}^{N_P} \sigma_{y_{P,i}}^2 + \sum_{i=1}^{N_R} \left(\sum_{j=1}^{N_P} \sum_{k=1}^{M+1} \frac{\partial g_k}{\partial y_{R,i}} E_{P,j}^{k-1} \sigma_{y_{R,i}} \right)^2 \end{aligned} \quad (\text{B.2})$$

The coefficients, g_k , are found by minimizing the mean-squared error between $f_B(E)$ and the background data sets. The solution described in many references^{19,20} is expressed here in terms of the two sets of background points $(E_{L,i}, y_{L,i})$ and $(E_{R,i}, y_{R,i})$ as

$$g_k = \sum_{m=1}^{M+1} \left[V_{k,m} \left(\sum_{i=1}^{N_L} y_{L,i} E_{L,i}^{m-1} + \sum_{i=1}^{N_R} y_{R,i} E_{R,i}^{m-1} \right) \right] \quad (\text{B.3})$$

The $V_{k,m}$ are the elements of the $(M+1)$ by $(M+1)$ covariance matrix defined as the inverse of the Hessian matrix, with elements $H_{k,m} = \sum_{i=1}^{N_L} E_{L,i}^{k+m-2} + \sum_{i=1}^{N_R} E_{R,i}^{k+m-2}$. Since the $V_{k,m}$ are independent of $y_{L,i}$ and $y_{R,i}$, the derivatives of the g_k in Eq. (B.2) reduce to

$$\frac{\partial g_k}{\partial y_{\Lambda,i}} = \sum_{m=1}^{M+1} V_{k,m} E_{\Lambda,i}^{m-1}, \text{ for } \Lambda = L \text{ or } R$$

As in Appendix A, we assume that the variances of all background points are equal and well approximated by the mean variance of the background signal, $\sigma_{\langle y_B \rangle}^2$. We express the total variance of the peak area [$\sum_{i=1}^{N_P} \sigma_{y_{P,i}}^2$ in Eq. (B.2)] as the sum of the intrinsic variances of the background-subtracted peak area, $\sigma_{A_P}^2$, and of the background area spanning the peak range, $N_P \sigma_{\langle y_B \rangle}^2$. Equation (B.2) can then be written as

$$\sigma_{A_{BS}}^2 = \sigma_{A_P}^2 + N_P \sigma_{\langle y_B \rangle}^2 \quad (\text{B.4})$$

$$\left\{ 1 + \frac{1}{N_P} \left[\sum_{i=1}^{N_L} \left(\sum_{j=1}^{N_P} \sum_{k=1}^{M+1} \sum_{m=1}^{M+1} V_{k,m} E_{L,i}^{m-1} E_{P,j}^{k-1} \right)^2 \right. \right. \\ \left. \left. + \sum_{i=1}^{N_R} \left(\sum_{j=1}^{N_P} \sum_{k=1}^{M+1} \sum_{m=1}^{M+1} V_{k,m} E_{R,i}^{m-1} E_{P,j}^{k-1} \right)^2 \right] \right\}$$

Since the $E_{L,i}$ and the $E_{R,i}$ are equally spaced by the same increment, dE , all sums in Eq. (B.4) reduce to geometric series, and the dependence on dE drops out. The $V_{k,m}$ can be determined by inverting the Hessian matrix with elements defined above. This process yields a closed-form expression for arbitrary N_L , N_R and N_P for a particular polynomial order, M .

As defined in Eq. (7), the RBSV is given by the terms in the curly brackets of Eq. (B.4). Once the summations are performed, we substitute $N_L = n_L N_P$ and $N_R = n_R N_P$ which produces a high-order rational polynomial in n_L , n_R and N_P . In most cases n_L and n_R are both $\ll N_P$, so the polynomials in the numerator and denominator will be dominated by the terms with the highest power of N_P . By keeping only these terms, the dependence on N_P drops out, and the RBSV for polynomial order M reduces to a function of only n_L and n_R denoted as $RBSV_M(n_L, n_R)$. The results of these calculation are presented in Table B.1 for polynomials up to order 3.

With the use of symbolic computational software, we also found closed-form expressions for orders 4 and 5 which are plotted in panel (b) of Fig. 6. Since these expressions were rational polynomials of very high orders, they were not included in Table B.1. Due to the increasing complexity of the closed-form expressions with polynomial order, for $M > 5$ the $V_{k,m}$ of Eq. (B.3) and the sums in Eq. (B.4) were calculated using numerical techniques for specific values of N_L , N_R and N_P . The results of these numerical calculations are plotted in panel (c) of Fig. 6.

For ease of use, the RBSVs for polynomial orders 3 through 5 can be approximated by the following rational polynomial in n_{Small} and n_{Large} where the former is the smaller of n_L and n_R , and the latter is the larger:

$$RBSV_{M, Approx}(n_{Small}, n_{Large}) = \frac{\sum_{k=0}^3 \sum_{m=0}^3 \mathcal{P}_{k,m} n_{Small}^k n_{Large}^m}{\sum_{k=0}^3 \sum_{m=0}^3 \mathcal{Q}_{k,m} n_{Small}^k n_{Large}^m}. \quad (B.5)$$

The coefficients $\mathcal{P}_{k,m}$ and $\mathcal{Q}_{k,m}$ were determined by least-squares fitting to the closed-form expressions and are provided in Table B.2. The approximate RBSVs given by Eq. (B.5) were found to be accurate to within $\pm 5\%$ over the ranges of $n_{Large} = n_L$ plotted for each polynomial order in panels (a) and (b) of Fig. 6. The $RBSV_{M, Approx}$ for $M=6$ and 7 were not included because they were not adequately approximated by third-order rational polynomials in n_{Small} and n_{Large} .

The exact RBSVs in Table B.1 and the approximations given by Eq. (B.5) and Table B.2 all simplify significantly for the commonly encountered case of symmetric endpoints ($n_S = n_L = n_R$). Table B.3 summarizes the expressions that may be more easily employed to determine the RBSVs for LSF of symmetric endpoints (uncorrelated with peak area) to polynomials up to order 7. The approximate RBSVs for orders 4 through 7 are within 3 % of the closed-form and numerical calculation for $0.5 < n_S < 10$. Note that the symmetric-endpoint RBSV for an even-order polynomial is exactly equal to that for the subsequent odd-order polynomial. This result is due to the symmetry of the $V_{k,m}$ and of Eq. (B.3) when $N_L = N_R$. When the degree of the polynomial is increased from an even to an odd integer, the additional degree of freedom is entirely accommodated by the addition of an inflection point at the center of the gap (the peak region) between the two sets of background points. Outside the peak region, the M_{Even} and $M_{Even} + 1$ polynomials are identical. Inside the peak region the difference between the even- and odd-ordered polynomials is completely symmetric about the inflection point. When the order is increased from M_{Even} to $M_{Even} + 1$, the statistical uncertainty will remain unchanged for the $M_{Even} + 1$ polynomial since the deviations from the M_{Even} polynomial on each side of the inflection point are completely correlated. There is, however, a risk of increasing the systematic uncertainty if the inflection point of the $M_{Even} + 1$ polynomial causes the background to exceed the signal for very small peaks near the detection limit. For these reasons, it is never advisable to increase the polynomial order beyond the lowest M_{Even} that adequately describes the background over equal numbers of points on each side of the peak.

Figure 6 suggests that the RBSV can vary significantly for a given total number of background points depending on how they are distributed on the two sides. The dependence of the RBSV on endpoint asymmetry is shown more directly in Fig. B.1 which plots the ratio of the asymmetric $RBSV_M(n_L, n_R)$ to the symmetric $RBSV_M(n_T/2)$ for the same total number of background points $n_T = n_L + n_R$. This ratio is plotted for three different n_T values as a function of the asymmetry of the endpoint range defined as $\alpha = n_{Large} / n_{Small}$. Comparing the plots for pairs of polynomial orders M_{Even} and $M_{Even} + 1$ in each panel, it is not surprising that the RBSV is always larger for the odd-ordered $M_{Even} + 1$ polynomial because of the extra degree of freedom. For a fixed n_T , the sign of the ratio's rate of change with α is reversed for even- and odd-ordered polynomials. This inverse behavior is most relevant as the asymmetry initially increase away from the symmetric condition $\alpha = 1$. For

even-ordered polynomials, the RBSV initially *decreases* with increasing endpoint asymmetry to a local *minimum* near the empirically determined value $\alpha_c = 4 n_T^{-1/5} (M_{Even} - 1.5)^{-1/3}$, and the asymmetric RBSV becomes larger than the symmetric RBSV for $\alpha \gtrsim 2\alpha_C$. For odd-ordered polynomials, the RBSV initially *increases* with α to a local *maximum* near α_C , and the asymmetric RBSV is always larger than the symmetric RBSV. The different values of n_T plotted in Fig. B.1 indicate that the magnitude of these variations with asymmetry decreases as the total number of background points increases. These observations indicate that for LSF of even-ordered polynomials the lowest uncertainty is obtained by using different numbers of background points on each side of the peak as long as $n_{Large} / n_{Small} < 2\alpha_C$ with a minimum near $n_{Large} / n_{Small} = \alpha_C$. For odd-ordered polynomials the uncertainty is always lowest for symmetric endpoints.

Figure B.1.

Relative background-subtraction variance, $RBSV_M(n_L, n_R)$, for least-squares fitting of a polynomial background of even (solid) and odd (long dashed) orders, M , for asymmetric endpoints (uncorrelated with the peak area) with background-to-peak point-number ratios n_L and n_R . The ratio of the asymmetric $RBSV_M(n_L, n_R)$ to the symmetric $RBSV_M(n_T/2)$ for the same total number of background points, $n_T \equiv n_L + n_R$, are shown on a semi-log plot as a function of the endpoint-number asymmetry $\alpha = n_L / n_R$ (assuming $n_L \geq n_R$) for three different values of n_T indicated by red, blue and green, respectively, for (a) $M=2$ and 3 with $n_T=1, 2, 4$; (b) $M=4$ and 5 with $n_T=2, 4, 8$; (c) $M=6$ and 7 with $n_T=3, 6, 12$.

Table B.1

Relative background-subtraction variances, $RBSV_M(n_L, n_R)$, for least-squares fitting of polynomials of order M to asymmetric endpoints uncorrelated with peak area. n_L and n_R are the background-to-peak point-number ratios for the sets of points used in the fit on the left and right sides of the peak, respectively. Here $n_{LR} = n_L n_R$ and $n_T = n_L + n_R$.

M	$RBSV_M(n_L, n_R)$
1	$\frac{n_T(1+n_T)(3+3n_T+n_T^2)}{n_T^4+12n_{LR}(1+n_T)}$
2	$\frac{(1+n_T)[60n_{LR}(1+n_T)^3+n_T^4(20+40n_T+28n_T^2+8n_T^3+n_T^4)]}{n_T^9+(1+n_T)n_Tn_{LR}[36n_T^2(5+5n_T+2n_T^2)-180n_{LR}(3+3n_T+n_T^2)]}$
3	$\frac{n_T(1+n_T)[-2100n_{LR}^2(1+n_T)^3(5+10n_T+10n_T^2+5n_T^3+n_T^4)+100n_{LR}n_T^2(1+n_T)^3(35+70n_T+77n_T^2+42n_T^3+9n_T^4)+n_T^8(175+42n_T+42n_T^2+14n_T^3+n_T^4)]}{n_T^{16}+(1+n_T)n_{LR}[42000n_{LR}^3(1+n_T)^3+80n_T^8(35+70n_T+53n_T^2+18n_T^3+3n_T^4)+2800n_{LR}^2n_T^4(20+40n_T+28n_T^2+8n_T^3+n_T^4)+40n_{LR}n_T^5(10+10n_T+5n_T^2+n_T^3)]}$

Table B.2

Coefficients for terms in numerator, $\mathcal{P}_{k,m}$, and denominator, $\mathcal{Q}_{k,m}$, of rational polynomials defined in Eq. (B.5) that approximate the RBSVs for LSF of asymmetric endpoints to polynomial backgrounds of order M .

M	Numerator coefficients, $\mathcal{P}_{k,m}$					Denominator coefficients, $\mathcal{Q}_{k,m}$				
3		1	n_{Large}	n_{Large}^2	n_{Large}^3		1	n_{Large}	n_{Large}^2	n_{Large}^3
	1	1190	10300	-10100	4760	1	5.25	9.03	12.1	249
	n_{Small}	-7920	27200	16800	-2120	n_{Small}	-193	-374	-3900	7810
	n_{Small}^2	204	-39900	-15400	14500	n_{Small}^2	2960	21300	-46000	7250
	n_{Small}^3	89300	34100	-14700	4720	n_{Small}^3	-29200	95100	-22800	5830
4		1	n_{Large}	n_{Large}^2	n_{Large}^3		1	n_{Large}	n_{Large}^2	n_{Large}^3
	1	18.5	-154	1560	-200	1	-1.14	5.55	-59.1	31.7
	n_{Small}	746	-5520	562	219	n_{Small}	3.41	72.3	299	25.1
	n_{Small}^2	5090	-31.7	-945	28.4	n_{Small}^2	-13.5	-621	-390	27.9
	n_{Small}^3	139	1230	-197	43.3	n_{Small}^3	89.2	1110	-258	41.7
5		1	n_{Large}	n_{Large}^2	n_{Large}^3		1	n_{Large}	n_{Large}^2	n_{Large}^3

M	Numerator coefficients, $\phi_{k,m}$					Denominator coefficients, $\mathcal{D}_{k,m}$				
	1	5560	3480	4500	3000	1	30.9	7.38	55.6	59
	n_{Small}	-15700	-17600	-13300	3140	n_{Small}	-402	-306	-1500	3140
	n_{Small}^2	35400	20300	-13100	-222	n_{Small}^2	2110	4960	-11500	-281
	n_{Small}^3	-16200	17400	394	-30.3	n_{Small}^3	-6990	12300	783	-37.4

Table B.3

RBSVs for least-squares fitting of a polynomial of order M to symmetric background points (uncorrelated with peak area) with peak-to-background point-number ratio n_S . The RBSVs are given for $M = 1$ through 3 as exact solutions to Eq. (B.4), while the RBSVs for $M = 4$ through 7 are approximated by $RBSV_M(n_S) = 1 + e^{F_M(n_S)}$ using the given $F_M(n_S)$.

M	RBSV or $F_M(n_S)$
1	$1 + \frac{1}{2n_S}$
2 and 3	$\frac{(1 + 2n_S)(5 + 20n_S + 40n_S^2 + 40n_S^3 + 16n_S^4)}{2n_S^3(15 + 30n_S + 16n_S^2)}$
4 and 5	$F_4(n_S) = \frac{11.7 + 3.59n_S - 2.84n_S^2}{0.735 + 5.10n_S + n_S^2}$
6 and 7	$F_6(n_S) = \frac{33.5 + 2.96n_S - 2.93n_S^2}{1.58 + 8.86n_S + n_S^2}$

Figure 1.

Simulated XPS spectrum of an isolated peak on a linear background acquired with energy step size dE . Bold blue and red capital letters indicate the sets of points described in Table 2. Vertical dashed lines are spaced at intervals equal to the peak energy range $E_P = N_P dE$ with N_P equal to the number of points comprising the peak (red dots in set indicated by P). The background-subtracted peak area, A_{BS} , shaded in light red is obtained by summing the measured counts at each of the points within the analysis energy range E_A (red and blue points inside black circles in set indicated by blue A) after subtraction of the background function, $f(E)$ (thick blue line). For the method of endpoint averaging shown here, $f(E)$ is the line connecting the endpoints (blue squares) defined by averaging the background points (blue dots) on the left (set of N_L points indicated by blue L) and right (set of N_R points indicated by blue R) sides of the peak. The endpoints may be divided into two sets depending on their correlation with the peak area. LA indicates the set of N_{LA} background points that are inside E_A and used to determine both the background and the peak area. LB indicates the set of N_{LB} background points that are outside E_A and are not correlated with the peak area. The scenario shown is for asymmetric endpoints ($N_L > N_R$) that are semi-correlated with peak area on the left ($N_{LA} = N_{LB} = N_L/2$) and uncorrelated with the peak area on the right ($N_{RA} = 0$, $N_{RB} = N_R$).

Figure 2.

(a–c) Plots of the relative standard deviation of the background-subtracted peak areas as a function of the number of symmetric points, N_S , used on each side of the peak to determine the linear background by endpoint averaging. Lines are plots of Eq. (5), and data points are the results of Monte Carlo simulations of 3000 spectra (See Fig. 3) assuming Poisson-distributed statistics. Results are shown for spectra with fixed Gaussian peak amplitude of 100 counts and full-width at half maximum of $E_{FWHM} = 2$ eV on a linear background with average value, $\langle \hat{y}_B \rangle$, at peak center (a) (circles) 300 counts, (b) (squares) 1000 counts and (c) (diamonds) 3000 counts. The relative uncertainty for each spectrum is calculated using energy step sizes, dE , of 0.2 eV (red), 0.5 eV (green) and 1.0 eV (blue) and for endpoints that are semi-correlated (thick lines and open symbols) and uncorrelated (thin lines and filled symbols) with the peak area. For each case, the number of points comprising the peak, N_P , is taken to be the largest integer less than $3 E_{FWHM}/dE$. (d) The relative background-subtraction variance (RBSV) plotted as function of the background-to-peak point-number ratio, $n_S = N_S/N_P$, for all three background amplitudes and energy step sizes shown in (a–c). The black thick and thin lines are plots of Eq. (8) for the semi- and un-correlated cases, respectively. The points are calculated from Eq. (7) using the results of the simulations with colors and symbols corresponding to the different dE and $\langle \hat{y}_B \rangle$ values, respectively, as in (a–c). The upper brown line is the RBSV resulting from the uncorrelated-area approximation [Eq. (9)] which neglects the correlation of the N_{SA} points used to determine both the background and peak areas.

Figure 3.

Examples of the 3000 spectra used for each Monte-Carlo simulation of the background-subtracted peak-area uncertainties plotted in Fig. 2 assuming Poisson-distributed statistics. All spectra have Gaussian peaks with the same amplitude (100 counts) and full-width at half maximum ($E_{FWHM} = 2$ eV) on a linear background. Each column shows spectra with different average background signals, $\langle \hat{y}_B \rangle = (300, 1000 \text{ or } 3000)$ counts, and each row shows the simulated spectra sampled with a different energy step size, $dE = (0.2, 0.5 \text{ and } 1)$ eV. The vertical dashed lines are spaced at intervals of $3 E_{FWHM}$.

Figure 4.

Ratio of RBSV for asymmetric endpoints to that for constrained, symmetric endpoints used to determine a linear background by endpoint averaging (black) and linear least-squares fitting (gray) of endpoints that are semi-correlated (thick curve) and un-correlated (thin curves) with the peak area. Only the uncorrelated case is shown for linear LSF. (a) Ratios plotted as a function of n_{Large} for fixed $n_{Small} = 0.1$ where n_{Large} and n_{Small} are the ratios of the larger and smaller, respectively, numbers of background points on each side of the peak to the number of peak points. The vertical dashed line indicates $n_{Large} = n_{Large}^{Opt}$ from Eq. (12) where the RBSV is a minimum for the endpoint-averaged, semi-correlated case. (b) Minimum RBSV ratios as a function of n_{Small} for $n_{Large} = n_{Large}^{Opt}$ for endpoint averaging. The vertical dashed gray line indicates the $n_{Small} = 0.1$ condition of (a). The dashed curve is the ratio of the minimum RBSV for asymmetric uncorrelated endpoints to the RBSV for the commonly employed constrained symmetric semi-correlated endpoints.

Figure 5.

Ratio of RBSVs for linear backgrounds determined by least-squares fitting, $RBSV_1$, and by uncorrelated endpoint averaging, $RBSV_{EA}$, of asymmetric endpoints. n_{Large} and n_{Small} are the ratios of the larger and smaller, respectively, numbers of background points on each side of the peak to the number of peak points. The solid, short-dashed and long-dashed curves show the ratio when $n_{Large} = 1, 1.5$ and 2 , respectively, for LSF and $n_{Large} = n_{Opt}$ is used for endpoint averaging.

Figure 6.

Relative background-subtraction variances, $RBSV_M(n_L, n_R)$, for least-squares fitting of a polynomial background of order M for asymmetric endpoints (uncorrelated with the peak area) with background-to-peak point-number ratios n_L and n_R . Log-log plots of $RBSV_M(n_L, n_R)$ are shown for successive even (solid) and odd (long dashed) polynomial orders as functions of n_L for three different n_R values indicated by the colors red, green and blue, respectively with (a) $M = 2$ and 3 , $n_R = 0.3, 0.6, 1.2$; (b) $M = 4$ and 5 , $n_R = 0.5, 1, 2$; (c) $M = 6$ and 7 , $n_R = 0.7, 1.4, 2.8$. For comparison, the horizontal dashed lines indicate $RBSV_1$ (0.2,0.2) (i.e., least-squares fitting of a linear background to symmetric endpoints spanning 20 % of the peak width).

Figure 7.

Background-to-peak point-number ratio, n_{poly} , of symmetric endpoints needed for least-squares fitting of a polynomial background of order $M=2$ through $M=7$ that will result in an RBSV equal to that for a linear background determined by symmetric endpoint averaging with background-to-peak point-number ratio, n_{Lin} . Points are calculated using the methods of Appendix B. Curves are parameterized fits given by Eq. (13).

Figure 8.

Total acquisition time to achieve a specific detection limit [Eq. (20)] as a function of the background-to-peak point-number ratio, n_S , for symmetric endpoints using: (a) linear endpoint averaging assuming endpoints are uncorrelated (thin curve) and semi-correlated (thick curve) with peak-area points; (b) least-squares fitting of a polynomial background of order $M=2$ and 3, and $M=4$ and 5 for the uncorrelated case. In (a) and (b), the plotted acquisition times are normalized to the minimum at $n_S = 0.5$ shown in (a) for the linear uncorrelated case.

Table 1

Partial list of symbols.

Symbol	Refers to
A_{BS}	Background-subtracted area used to determine peak intensity; direct sum of measured counts at each spectrum point in the energy range E_A after subtraction of background
β	Average background count rate (counts per second)
C_x	Atomic concentration of species x [Eq. (1)]
dt	Acquisition dwell time for each point in spectrum
dE	Energy step size of spectrum
E_A	Analysis energy range spanning all points used to determine peak area A_{BS}
E_P	Minimum energy range that spans all points comprising the peak and no background points
I_x	Peak intensity of relevant energy level for species x in units of counts per second times electron volts [Eq. (2)]
k	Coverage factor for detection limit [Eq. (16)]
M, M_{Even}	Order of polynomial used for least-squares fitting of background. “Even” subscript indicates restriction to even orders.
N_P	Minimum number of points that fully comprise <i>only</i> the peak including no background points
N_Λ	Number of measured spectrum points in set corresponding to label Λ as described in Table 2.
n_Λ	Ratio of number of background points, N_Λ , to number of peak points, N_P $n_\Lambda = N_\Lambda/N_P$
R_x	Relative sensitivity factor for the relevant XPS peak of species x .
$\sigma_{A_{BS}}$	Uncertainty of the measured background-subtracted peak area, A_{BS} .
σ_{A_P}	Standard deviation of the intrinsic peak area, independent of the background-subtraction process.
$\sigma_{\langle y_B \rangle}$	Standard deviation of the average measured background signal
y	Signal measured in XPS spectrum in units of counts
$\langle y_B \rangle$	Mean background signal in units of counts

Table 2

List of subscript labels used to indicate specific sets of spectrum points.

Label index, Λ	Refers to
LB, RB	Background points on left and right, respectively, of peak used <i>only</i> to determine background
SB	Symmetric background points on either side of peak used <i>only</i> to determine background; $N_{SB} = N_{LB} = N_{RB}$
LA, RA	Background points on left or right, respectively, of peak used to determine <i>both</i> background and peak area
SA	Symmetric background points on each side of peak used to determine <i>both</i> background and peak area; $N_{SA} = N_{LA} = N_{RA}$
L, R	All background points on left or right, respectively, side of peak regardless of peak-background correlation; e.g., $N_L = N_{LB} + N_{LA}$
S	All symmetric background points on each side of peak regardless of peak-background correlation; $N_S = N_L = N_R$
$Small, Large$	Smaller or larger, respectively, set of asymmetric background points regardless of which side of the peak they are on
P	Points comprising the entire peak <i>only</i> , including no background points
A	All points used to determine the peak area; $N_A = N_P + N_{LA} + N_{RA}$

Table 3

List of modifiers used to indicate specific instances of the symbols in Table 1.

Symbol modifier	Refers to
$\wedge$	Caret indicates that quantity is derived from a measured signal of unscaled counts that is governed by Poisson statistics.
Λ	Subscript index for sets of spectrum points described in Table 2
DL	Subscript indicating quantities corresponding to detection limit
Opt	Superscript indicating the optimal value to minimize the RBSV
q	Subscript index for all measured species
sub	Subscript for quantities associated with the substrate
x	Subscript referring to a particular species
EA	Subscript indicating background determined by endpoint averaging