Probability Density Function Transformation Using Seeded Localized Averaging

Nedialko B. Dimitrov and Valentin T. Jordanov, *Senior Member, IEEE*

***Abstract*—Seeded Localized Averaging (SLA) is a spectrum acquisition method that averages pulse-heights in dynamic windows. SLA sharpens peaks in the acquired spectra. This work investigates the transformation of the original probability density function (PDF) in the process of applying the SLA procedure. We derive an analytical expression for the resulting probability density function after an application of SLA. In addition, we prove the following properties: 1) for symmetric distributions, SLA preserves both the mean and symmetry. 2) for unimodal symmetric distributions, SLA reduces variance, sharpening the distribution's peak. Our results are the first to prove these properties, reinforcing past experimental observations. Specifically, our results imply that in the typical case of a spectral peak with Gaussian PDF the full width at half maximum (FWHM) of the transformed peak becomes narrower even with averaging of only two pulse-heights. While the Gaussian shape is no longer preserved, our results include an analytical expression for the resulting distribution. Examples of the transformation of other PDFs are presented.**

***Index Terms*—Energy resolution, energy spectrum, seeded localized averaging, spectrum acquisition.**

I. Introduction

UNIQUE radioactive elements and their associated decay chains produce unique energy spectra. This property has motivated extensive study in the ability to identify radioactive material based on its energy spectrum [1]–[3]. The main objective of the radiation spectroscopy is to record a spectrum from a radiation detector, and a standard method of recording a spectrum is a multi-channel pulse height analyzer (MCA). The spectrum provides information about both the incident radiation and the response of the radiation detector. However, to accurately identify a radioactive material, we would like accurate information solely on the incident radiation.

An example of radiation spectrum is shown in Fig. 1. The horizontal axis is the channel number. The vertical axis represents the channel content—number of counts. There are different features in the spectra that are associated not only with the energy of the detector events but also with the detector itself and the signal processor properties. The most important characteristic of the spectroscopy system is the resolution (energy, time, etc.). The full-width at half-maximum (FWHM) of the spectral peaks is a measure of the spectroscopy system capability to resolve incident radiation. The FWHM depends on various factors such as statistical fluctuations of the detector signal, noise contribution of the signal processor, external interference, temperature and long term drifts, etc [4].

Statistical fluctuations determine the theoretical limit of the energy resolution for a given detector. The effect of the other sources can be reduced by using appropriate noise filtering and other electronic techniques. One way to reduce the detector statistical fluctuations is to use averaging of more than one pulse height in the process of spectrum acquisition. In this paper, we analyze a spectrum acquisition technique called Seeded Localized Averaging (SLA), that allows for a real time reduction of the FWHM in acquired spectra [5], [6]. In addition to demonstrating the benefits of SLA through both experimental and simulated examples (see Fig. 1), we prove basic theoretical properties of the transformation.

There are methods to post-process the acquired energy spectra using spectral-deconvolution that reduce peak FWHM [7]. Those methods are based on modeling the observed energy spectrum as a function of two random variables, the input energy spectrum and the detector response function. The detector response function can be modeled as a joint probability distribution on input pulse heights to output pulse heights. At least three such methods exist, using regularisation [8], maximum likelihood [9], and maximum entropy [10]. However, modeling the detector's energy-response function for spectral-deconvolution "can be a very computing-intensive task [7]."

The main difference between SLA and spectral-deconvolution methods is that SLA is a real-time spectrum acquisition technique which requires no knowledge of the detector response function. In other words, SLA can be carried out in hardware, in real time, as pulse heights are measured by the detector. In addition, because SLA uses averaging to reduce instrument noise, SLA does not require an explicitly formulated noise model for the detector. This is practically advantageous, because SLA can be readily used with varying detector equipment. On the other hand, having a computationally-intensive, data-driven model of the noise from a specific instrument may result in cleaner post-processing spectra.

The remainder of this paper is structured as follows. In Section II, we describe Seeded Localized Averaging, its variants, and its basic theoretical properties. In Section III, we demonstrate some specific examples of the SLA transformation to different distributions and its sensitivity to its input parameters. In Section V, we draw some conclusions. The detailed proofs of the properties of SLA are left to Appendix A.

Manuscript received June 18, 2011; revised September 08, 2011; accepted November 03, 2011. Date of publication January 04, 2012; date of current version August 14, 2012

N. B. Dimitrov is with the Operations Research Department, Naval Postgraduate School, Monterey, CA 93943 USA (e-mail: ned@nps.edu).

V. T. Jordanov is with the Yantel, LLC, Santa Fe, NM 87508 USA (e-mail: jordanov@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2011.2177861

0018-9499/$31.00 © 2012 IEEE

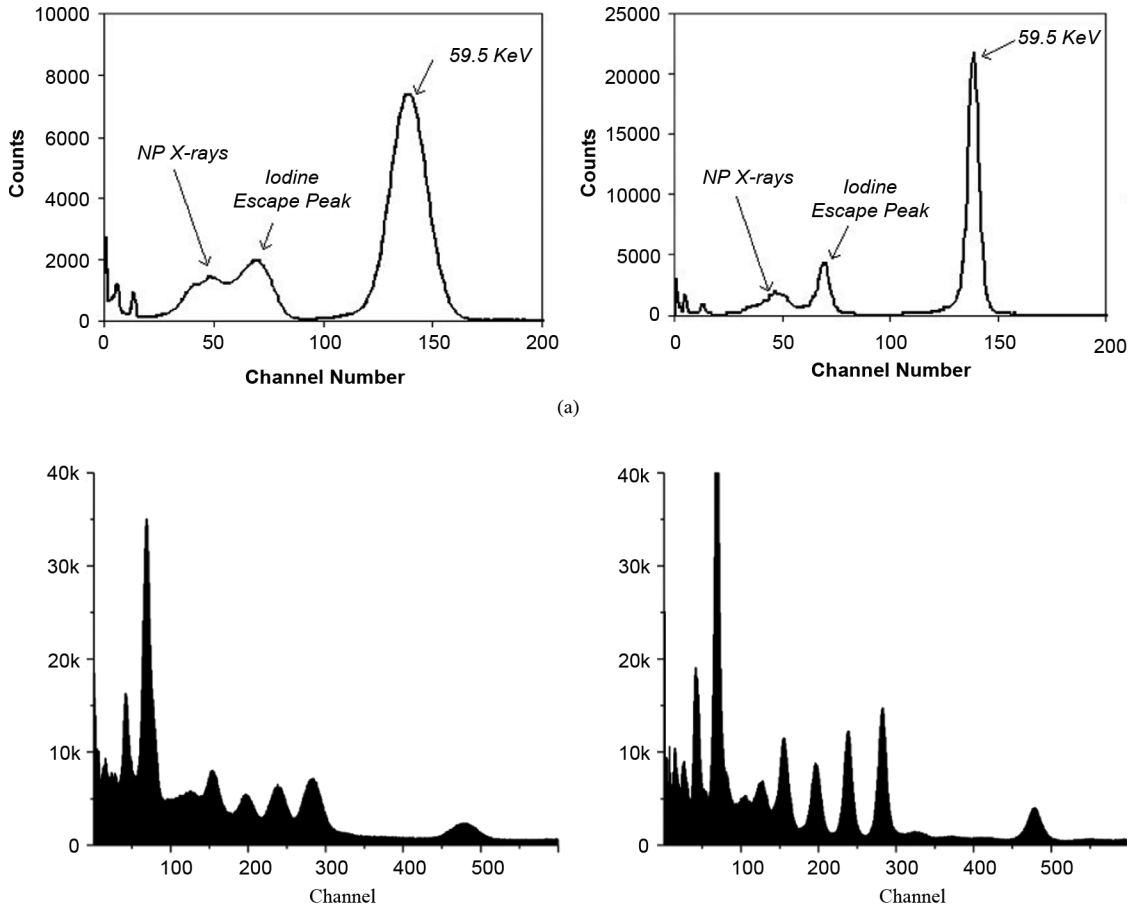

Fig. 1. Comparison of conventional spectrum acquisition to a spectrum acquired with seeded localized averaging (SLA). Fig. 1(a) shows a low energy spectrum example from NaI(Tl) scintillation detector and Am241 radiation source. Both spectra have the same total number of counts. The spectrum on the left was acquired with conventional pulse height analysis, the spectrum on the right was acquired using SLA. The full-width half-maximum of the peak is significantly reduced. Fig. 1(b) also shows conventional pulse height analysis (left) and SLA (right) using an old fluorescent watch dial source, Ra226, and a NaI(Tl) scintillation detector. Both spectra have the same total number of counts. SLA again reduces the FWHM.

II. Description of Seeded Localized Averaging

In conventional spectroscopy, the increment of the spectroscopy channels is based on a single pulse height measurement. That is, every pulse height measurement causes increment of the channel content.

In order to improve energy resolution, Seeded Localized Averaging uses the average of a predetermined number of pulse heights to determine a channel to increment. The essential feature of SLA and the fundamental difference from the conventional pulse height analysis method is the use of more than one pulse height measurement to increment the spectroscopy channels. That is, the channel number is derived and the channel content increments only after the average of two or more pulse heights is obtained.

Using simple pulse height averaging of all measured pulse heights destroys the differential between distinct energy peaks in the spectrum. Therefore, SLA implements a selective approach to carry out the averaging, averaging only over a narrow window of pulse heights. The window size naturally alters the results, and has to be carefully chosen in order to preserve the spectral information. The window size selection should be made taking into account the FWHM of the spectrum obtained by the conventional pulse height analyzer.

Selection of an appropriately narrow averaging window is not sufficient by itself; it is also necessary to select the position of the averaging window. If the averaging window has a fixed position, the post-averaging spectrum would simply be a step-function of the original spectrum, not providing any decrease in the FWHM. Thus, the key to SLA is that it alters the position of the averaging window as the averaging is performed.

The SLA process has two parameters, the size of the *averaging window*, and the predetermined *number of pulse heights to average*. Simply, in SLA, when the first pulse height is measured, it creates an *averaging range* around it based on the averaging window. For example, if the averaging window is 3 and a pulse height of 100 is measured, the averaging range is from 97 to 103. When a second pulse height is measured that falls within that averaging range, the two pulse heights are averaged and the center of the averaging range is moved to the computed average. For example, suppose the second pulse height measured is 98. The second pulse height is within the averaging range created by the first pulse height. The averaging range is re-centered at the average of the two pulse heights, 99, and now spans 96 to 102. The process is repeated until a fixed number of pulse heights have been averaged, as specified by the SLA parameter. For example, if the number of pulse heights to average is 2, the running example has averaged two pulse heights, and we output 99 as a channel increment.

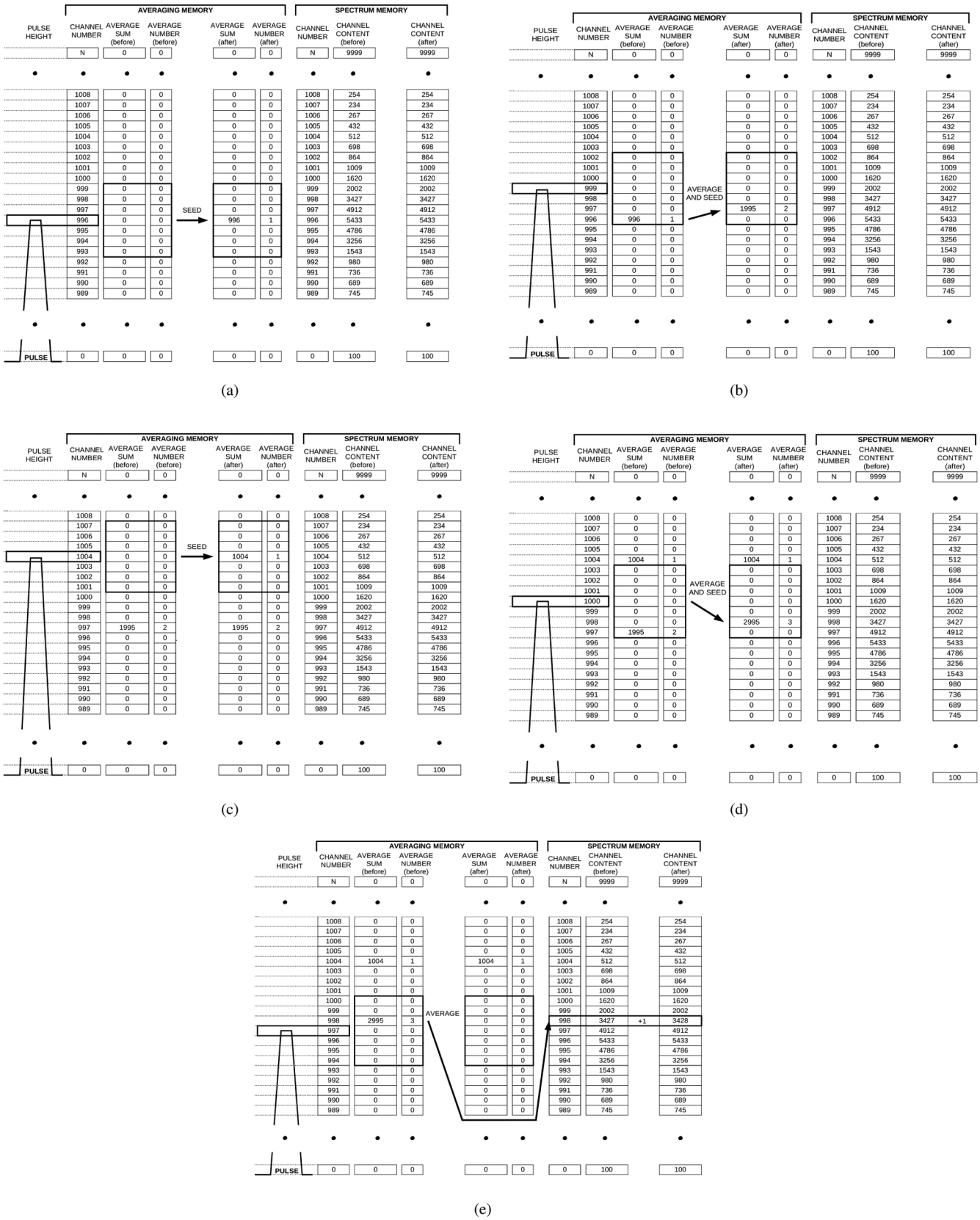

Fig. 2. Example of SLA implementation. The averaging window in this example is set to 3, making the search range encompass a total width of seven channels—three from each side of the pulse height channel and the pulse height channel itself. Figs. 2(a) and 2(c) show a seeding operation, when there is no average sum within the search range. The rest of the figures illustrate the average sum build-up process. The number of pulse heights to average in this example is set to four. This number is reached in Fig. 2(e) which illustrates the process of incrementing the spectrum memory channels.

Fig. 2 graphically depicts an implementation of SLA. To implement SLA, the averaging of the pulse heights is done using an *averaging memory*. The averaging memory has similar structure as spectral memory; it has channels that typically have the same

width as the spectral channels. Each of the averaging memory channels holds two values: average sum and average number.

Since there may be multiple averaging ranges active at the same time, it is easier to think of SLA's averaging window parameter as specifying a *search range* around the recently measured pulse height.

Initially, the entire content of the averaging memory is set to zero. When a pulse arrives its pulse height is used to center the search range. The average number is checked within the search range. If all channels within the search range hold zero than the pulse height becomes the seed of new average sum. The pulse height seeds the average sum, and the average number is incremented by one Fig. 2(a).

After some average sums are seeded, when a pulse height is measured, not all channels in the search range will hold zero. If there is only one channel with non-zero average sum in the search range, then the measured pulse height is added to the average sum of that channel. If more than one channel has non-zero average sum in the search range, the measured pulse height is added to the average sum of the channel closest to the measured pulse height. If the measured pulse height is exactly between two channels containing average sums, a fair coin is flipped and the pulse height is added to one of the two channels. Every time a pulse height is added to the average sum, the average number is incremented by one. Then, an average pulse height is computed by dividing the average sum by the average number. If the average pulse height points to a new channel in the averaging memory, the average sum and the average number are moved to this new channel. The old average sum and average number locations are set to zero Fig. 2(b).

Eventually, the average number of an averaging channel reaches the predetermined number of pulse heights to average. At that point, the average sum is divided by the average number. The result of this division is an average pulse height that is used to point to the channel of the spectral memory. The channel of the spectral memory increments. The locations of the average sum and the average number are set to zero Fig. 2(e).

A. Theoretical Properties of SLA

Theoretically, SLA can be viewed as a probability density function (pdf) transformation. The energy spectrum produced by the directly measured pulse heights describes a probability density over energy levels. SLA takes as input such a probability density, and transforms it to a new probability density over the same energy levels.

Let the energy spectrum before SLA describe pdf $f(x)$ and associated cumulative distribution function (cdf) $F(x)$. Let r be the averaging window parameter. Let n be the number of pulse heights to average parameter, and for simplicity let us fix it to 2. Let X_0 be a random variable (r.v.) describing the initial pulse height measured; with a pdf $f(x)$ and cdf $F(x)$. The pulse height X_0 specifies an averaging range, $[X_0 - r, X_0 + r]$. Let X_1 be a r.v. describing the second pulse height to fall within the averaging range. The distribution of X_1 has the same pdf as X_0 but restricted to the interval $[X_0 - r, X_0 + r]$ and normalized so that it integrates to 1. The energy channel incremented at the end of SLA is the average of the two measured pulse heights, corresponding to the random variable $P_0 = X_0 + X_1/2$. An input probability density function, $f(x)$, completely specifies the distributions of X_0 and X_1, and the SLA transformation produces the probability density function of P_0.

For the remainder of this section, we describe the theoretical properties of SLA, and what they mean practically. The proofs of these properties are relegated to the appendix. We discuss the properties with the parameter n fixed to 2, for simplicity. For a larger value of n, we would have $P_0 = \sum_{i=0}^{n-1} X_i/n$, where the distribution of X_j for $j = 1 \ldots n-1$ is the same as that of X_0 but restricted to the interval $[\sum_{i=0}^{j} X_i/j - 1 - r, \sum_{i=0}^{j-1} X_i/j + r]$ and normalized so that it integrates to 1. Properties for those versions of the SLA transformation can be derived in a similar manner as the ones stated here.

It is possible to explicitly characterize the pdf and cdf resulting from the SLA transformation, in terms of the parameter r and the initial pdf, $f(x)$ and cdf, $F(x)$, regardless of their specific shapes.

Lemma II.1: The random variable P_0 has cdf

$$\Pr[P_0 \le y] = \int_{-\infty}^{y-r/2} f(x)\,dx + \int_{y-r/2}^{y+r/2} \times f(x_0) \int_{x_0-r}^{2y-x_0} \frac{f(x_1)}{F(x_0+r) - F(x_0-r)}\,dx_1 dx_0$$

and pdf

$$f_p(y) = \int_{y-r/2}^{y+r/2} \frac{2f(x)f(2y-x)}{F(x+r) - F(x-r)}\,dx.$$

The most basic, desirable property of the SLA transformation is not to shift the locations of peaks in the energy spectrum. This is clearly not true for an arbitrary setting for the parameters. For example, setting r to infinity and n to a very large constant makes to SLA transformation take a simple average of n measured pulse heights, regardless of their location. For those parameter settings, the SLA transformation would produce a delta function at the mean of the original distribution, $f(x)$. Using the characterization of Lemma II.1, it is possible to show that SLA does not shift or alter the symmetry of the original distribution.

Lemma II.2: If $f(x)$ is symmetric about μ, then $f_p(y)$ is also symmetric about μ. As a corollary, if $f(x)$ is symmetric with mean μ, then $f_p(y)$ is also symmetric with the same mean, μ.

Intuitively, the above lemma states that regardless of the parameter values for SLA, the SLA transformation does not shift the distribution, and maintains symmetry. In particular, if the original distribution were a single Gaussian peak, the resulting distribution would have the same mean, maintaining the peak location, and would also be symmetric.

The second basic, desirable property of the SLA transformation is that it reduces FWHM of the original distribution, sharpening the peaks. For a single Gaussian peak, the FWHM is related to the variance of the distribution. We can characterize SLA's impact the variance of $f(x)$ with the following lemma.

Lemma III.3: If $f(x)$ is symmetric and increasing towards its mean, μ, then P_0 has a smaller variance than X_0.

Intuitively, the above lemma states that for any distribution that looks like a single peak, in other words is symmetric and increasing towards the mean, the SLA transformation reduces the variance, and thus the FWHM. Intuitively, for distributions with multiple peaks, if r is small, at a local level of radius r, the distribution behaves like a distribution with a single peak. In Section III, we show explicit examples of these theoretical results that help in their interpretation.

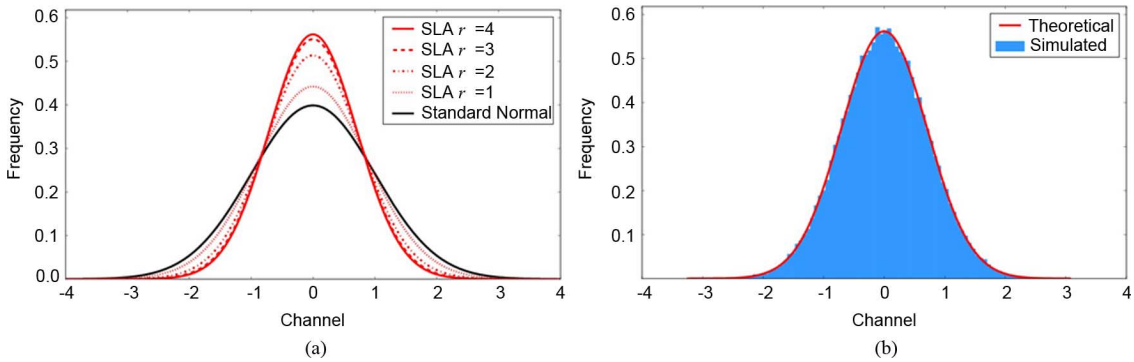

Fig. 3. The Seeded Localized Averaging transformation applied to a standard normal distribution. Fig. 3(a) depicts the original Gaussian pdf, and the resulting pdfs from several variants of SLA, all with $n = 2$ but with varying averaging window, r. The figure demonstrates the decrease in the FWHM. Fig. 3(b) demonstrates agreement between a simulated histogram of SLA with $n = 2, r = 4$ applied to a standard normal against the theoretical distribution as specified by Lemma II.1.

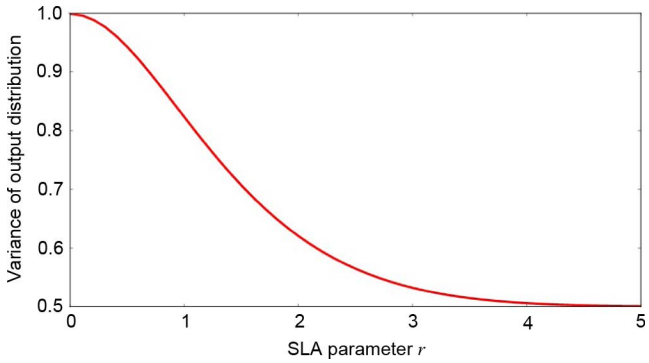

Fig. 4. SLA variance reduction when applied to a standard normal distribution. The horizontal axis shows different values of the averaging window parameter, r. The parameter n is fixed at 2. The vertical axis shows the variance of the resulting distribution. The results agree with theoretical arguments showing that the variance of the output distribution must be between 1 and 0.5.

III. Examples of Seeded Localized Averaging

To get a better understanding of the SLA transformation, Fig. 3 depicts the transformation as applied to a standard normal distribution. Fig. 3 depicts the original distribution, and the resulting transformed distributions from SLA with n fixed to 2 and different values of r. The decrease in the FWHM of the peak is evident in the figure. Fig. 3(b) depicts the theoretical distribution resulting from SLA with $n = 2$ and $r = 4$ versus a histogram derived through simulation.

Different values of r provide different levels of variance reduction, as can be seen in Fig. 3(a). We can understand what levels of variance reduction are possible, both theoretically and practically. Theoretically, as r goes to zero, X_1 is approximately equal in value to X_0, and thus $P_0 = X_0$, providing no variance reduction. As r goes to infinity, X_1 has the same distribution as X_0, and thus $P_0 = X_0 + X_1/2$ is the sum of two independent, identically distributed Gaussians divided by two. Because of standard results on the sums of Gaussians, if X_0 is standard normal, the variance of P_0 is $1/2$, providing a factor of 2 variance reduction. As n increases, this variance reduction would be greater. Fig. 4 depicts the variance reduction of SLA when changing the parameter r and holding n fixed to 2.

For a symmetric distribution without a peak, such as the uniform distribution, SLA maintains both the mean and symmetry, as guaranteed by Lemma II.2. The result of the SLA transformation on the uniform distribution is depicted in Fig. 5. When the averaging window, r, is equal to 1, the SLA transformation outputs the average of two independent identically distributed uniform variables, giving the familiar triangular shaped distribution in Fig. 5(a). Fig. 5(b) demonstrates agreement between a simulated histogram of SLA with $n = 2, r = 0.25$ applied to the uniform against the theoretical distribution as specified by Lemma II.1.

Finally, consider a distribution with two peaks, such as a mixture of two Gaussians. Fig. 6 depicts the SLA transformation acting on a mixture of two standard normals, one centered at -2 and the other at 2. For small values of r, the SLA transform reduces the FWHM, as shown for the value 1. For larger values of r, the original distribution no longer locally (an interval with radius r) acts as a distribution with one peak. Thus, for larger values of r, 4 in the example, we see a new peak introduced at 0, the mean of the original distribution. For any distribution, there is a value of r small enough so as not to introduce such artifacts. And, even for large values of r, the SLA transform maintains the mean of the original pdf and symmetry as guaranteed by Lemma II.2.

IV. SLA Simulations and Experimental Tests

Analytical solutions of the PDF transformation due to SLA become complicated when a larger number of pulse heights are averaged. Numerical simulations allow investigation of the properties of SLA by synthesizing artificial peaks that model real spectroscopy peaks. These simulations allow for controlling parameters such as FWHM and the total number of counts under the peak. To illustrate this, an example of two overlapping peaks is shown in Fig. 7(a). A random number generator produces a sequence of pulse heights that can be processed either conventionally (MCA) or using SLA. The two Gaussian peaks P1 and P2 in Fig. 7 are obtained using an MCA technique. In this example P1 represents exactly 1 million counts while P2 has counts that are fraction of the P1 counts. This fraction of counts is indicated in percentage in Fig. 7(a). The FWHM MCA of both peaks is 50 channels. The space between the centroids of the peaks is exactly 1.5 FWHM MCA. The same sequence of pulse heights is processed by SLA routine with averaging window of 0.5 FWHM MCA. The number of averaging pulse

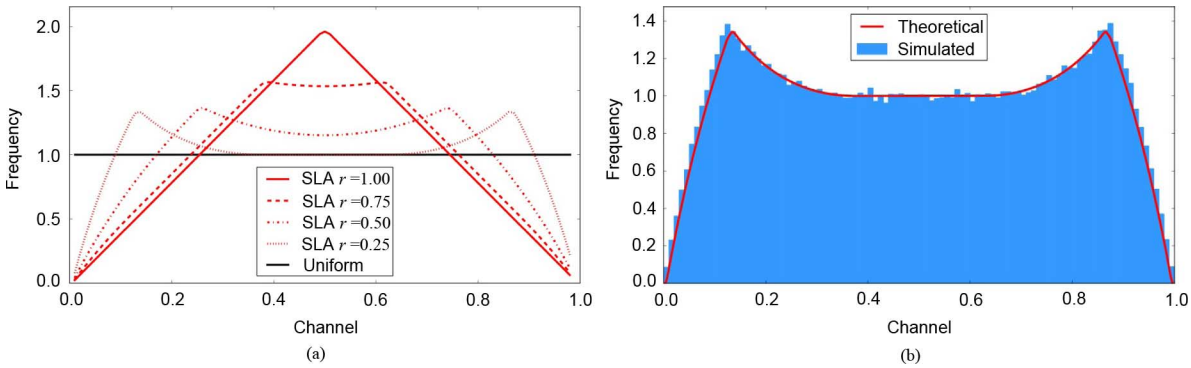

Fig. 5. The Seeded Localized Averaging transformation applied to a uniform distribution. Fig. 5(a) depicts the original uniform pdf, and the resulting pdfs from several variants of SLA, all with $n = 2$ but with varying averaging window, r. Fig. 5(b) demonstrates agreement between a simulated histogram of SLA with $n = 2, r = 0.25$ applied to the uniform against the theoretical distribution as specified by Lemma II.1.

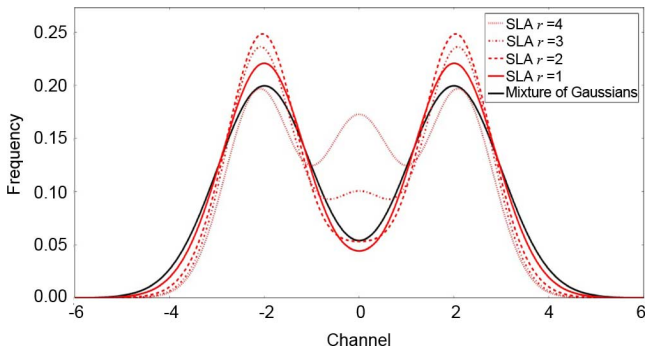

Fig. 6. SLA applied to a mixture of two Gaussian distributions. The original pdf is of two standard normals, centered at 2 and -2. When r is sufficiently small, about 1 in this example, the FWHM of both peaks is reduced, and the dip between the peaks is exaggerated. When r becomes larger, the original pdf no longer behaves like a single peak locally. With $r = 4$, the output density's pdf has a peak at 0, the mean of the original density. Because the original distribution is symmetric, the output distribution is also symmetric, following the theoretical result of Lemma II.2.

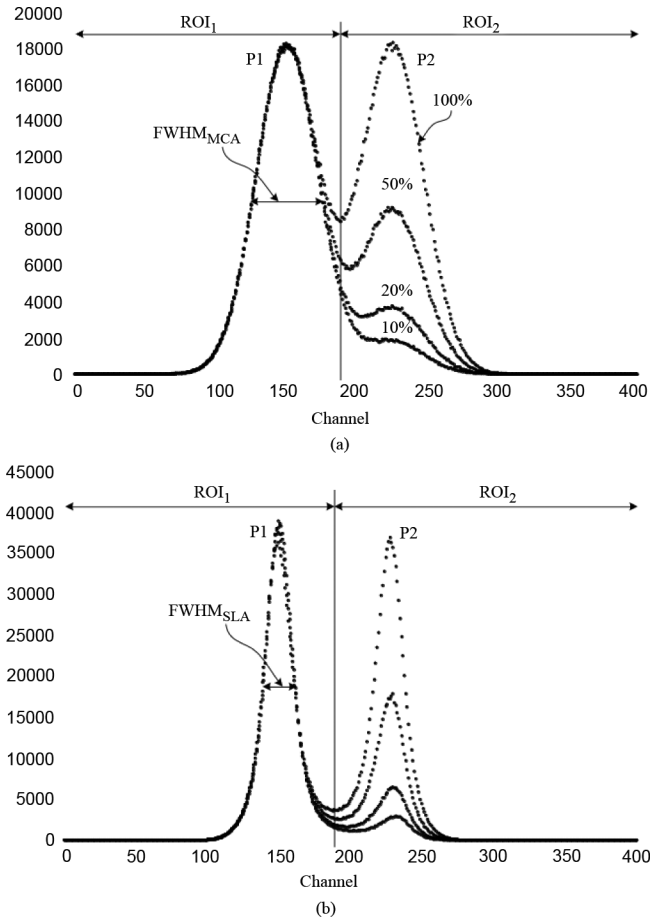

Fig. 7. A simulation of SLA for two Gaussian peaks of differing sizes. In this example P1 represents exactly 1 million counts while P2 has counts that are fraction of the P1 counts, as indicated in Fig. 7(a). Both peaks have a FWHMA MCA of 50 channels, and the space between their centroids is 1.5 FWHM MCA. Fig. 7(a) depicts spectra acquired using MCA. Fig. 7(b) depicts spectra acquired with SLA, an averaging window of 0.5 FWHM MCA, and averaging 16 pulse heights. To preserve the total number of counts SLA with pipeline pulse height recycling is used [11].

heights is variable in general (Table I), however, Fig. 7(b) shows an example of an SLA spectrum averaging 16 pulse heights. To preserve the total number of counts SLA with pipeline pulse height recycling is used [11].

One important question about SLA is its ability to preserve the relative area of the peaks in the process of PDF transformation. While a theoretical or in-depth experimental analysis of the relative peak area is out of the scope of this paper, we provide some quick check simulation analysis for this question. We consider two separate simulations. In the first, we consider peaks with centroid spacing less than 3 FWHM MCA, and in the second we consider peaks with centroid spacing more than 3 FWHM MCA.

For close peaks, with centroid spacing less than 3 FWHM MCA, the simulation shown in Fig. 7 allows for performing a quick check on the peak area ratio (P2/P1) using simple count integration in the regions of interest (ROI). A set of SLA spectra using different numbers of pulse heights to average are recorded and the P2/P1 ratio is calculated (see Table I). Table I also shows the dependence of FWHM SLA on the number of pulse heights to average. For this particular example it is evident that SLA improves not only the FWHM but the peak area ratio which can be viewed as a direct sequence of the resolution improvement. Specifically, consider the example where P2 has 10% of the counts of P1. Because the two peaks are close and overlap, the ratio of the ROI for P2/P1 in an MCA acquired spectrum is not 10% but about 14%. On the other hand, when 16 pulse

TABLE I
THE BEHAVIOR OF SLA AS WE VARY THE NUMBER OF PULSE HEIGHTS TO AVERAGE. THE FIRST ROW OF THE SLA WHEN ONLY 1 PULSE HEIGHT IS AVERAGED, WHICH IS EQUIVALENT TO STANDARD MCA SPECTRUM ACQUISITION. THE COLUMNS ON THE RIGHT HAND SIDE OF THE STUDIES FOR DIFFERENT SPECTRA DEPICTED IN FIG. 7(A). FOR EXAMPLE, A COUNT RATIO OF P2/P1 OF 10% MEANS THAT P2 HAS 10% OF THE COUNTS OF P1. BECAUSE THE TWO PEAKS ARE VERY CLOSE AND OVERLAP, WHEN INTEGRATING THE REGIONS OF INTERESTS (ROI) UNDER EACH PEAK, THE RATIO NEED NOT BE 10%. FOR EXAMPLE, AN MCA ACQUIRED SPECTRUM WHEN THE COUNT RATIO IS 10% RESULTS IN AN ROI AREA RATIO OF 13.97%. SLA IMPROVES THE FWHM AND ACHIEVES AN ROI AREA RATIO OF P2/P1 THAT IS CLOSER TO THE ORIGINAL COUNT RATIO WHEN COMPARED TO MCA

Number of pulse heights to average	$\frac{\text{FWHM SLA}}{\text{FWHM MCA}}$	SLA $\frac{P_2}{P_1}$ when count ratio is 10%	SLA $\frac{P_2}{P_1}$ when count ratio is 20%	SLA $\frac{P_2}{P_1}$ when count ratio is 50%	SLA $\frac{P_2}{P_1}$ when count ratio is 100%
1 (MCA)	1	13.97 $\pm$ 0.04 %	23.53 $\pm$ 0.05 %	52.18 $\pm$ 0.07 %	100.02 $\pm$ 0.1 %
2	0.87	13.04 $\pm$ 0.04 %	22.67 $\pm$ 0.05 %	51.60 $\pm$ 0.07 %	99.98 $\pm$ 0.1%
4	0.75	12.03 $\pm$ 0.03 %	21.82 $\pm$ 0.05 %	51.01 $\pm$ 0.07 %	100.00 $\pm$ 0.1%
8	0.58	11.34 $\pm$ 0.03 %	21.22 $\pm$ 0.05 %	50.68 $\pm$ 0.07 %	100.04 $\pm$ 0.1%
16	0.42	10.80 $\pm$ 0.03 %	20.72 $\pm$ 0.05 %	50.44 $\pm$ 0.07 %	100.01 $\pm$ 0.1%

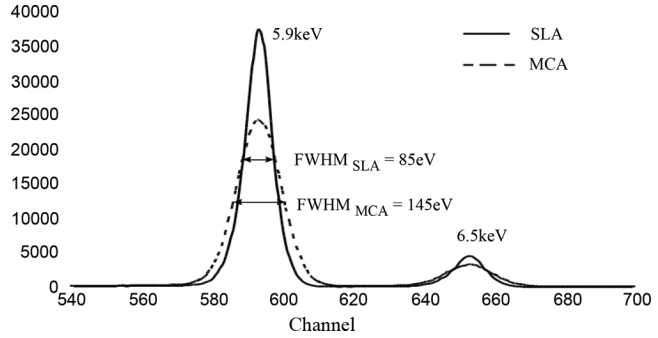

Fig. 8. X-ray fluorescent spectrum of the Mn K-lines obtained with a silicon drift detector. SLA reduces the FWHM from 145 eV (MCA) to 85 eV (SLA). The peaks are more than 3 FWHM MCA, and the ratio of the peak areas under MCA and SLA are identical.

heights are averaged for an SLA acquired spectrum, the ratio of the ROI for P2/P1 is 10.8%. More detailed results are available in Table I.

The study in Fig. 8 considers peaks with centroids spaced at distance larger than 3 FWHM MCA. The figure shows an x-ray fluorescent spectrum of the Mn K-lines obtained with a silicon drift detector. The resolution obtained with MCA is 145 eV. The resolution of the SLA spectrum is 85 eV and exceeds the statistical limit of the conventional spectroscopy using silicon detectors. The peak areas and the ratio between K-alpha and K-beta peaks obtained with MCA and SLA are identical.

V. CONCLUSIONS

Fig. 2 compares spectra obtained with conventional spectroscopy system and spectroscopy system using SLA. The figure shows the improvement in FWHM when the SLA technique is used.

This paper demonstrates both a practical and theoretical description of the Seeded Localized Averaging transformation. The theoretical description leads to proofs of the fundamental properties of SLA, and investigation of the SLA transformation and its sensitivity on basic examples such as a Gaussian peak (Figs. 3 and 4), the uniform distribution (Fig. 5), and a mixture of two Gaussians (Fig. 6). For a symmetric distribution with a single peak, the SLA transformation always maintains the mean, maintains symmetry, and reduces the FWHM. For distributions with multiple peaks, if the averaging window is small enough, SLA reduces the FWHM. Simulations and experimental results have shown advantages of using SLA to improve the FWHM of the spectral peaks.

One drawback of SLA is the reduced number of counts in the channels compared to the conventional method. The total number of the counts in the SLA spectra is reduced by a factor equal to the predetermined maximum number of pulse heights in the average sum (the parameter n). Some techniques are available to mitigate this effect [11]. For example, a large number of pulse height measurements may be collected, and then repeatedly, randomly sub-sampled to apply the SLA transformation. However, the theoretical analysis of such techniques is beyond the scope of this presentation.

SLA provides a real-time processing method for reducing detector noise, with provable theoretical guarantees, and contributes to a related body of work on spectral-deconvolution methods.

APPENDIX
RANDOM VARIABLE DEFINITIONS

Let r be SLA averaging window. Let X_0 have a pdf $f(x)$ and cdf $F(x)$. Let X_1, which depends on X_0, have the same pdf as X_0 but restricted to the interval $[X_0-r, X_0+r]$ and normalized so that it integrates to 1.

We define the seeded localized average as the random variable $P_0 = X_0 + X_1/2$.

PDF and CDF of Seeded Localized Averaging:

Lemma A.1: The random variable P_0 has cdf:

$$\Pr[P_0 \le y] = \int_{-\infty}^{y-r/2} f(x)\,dx + \int_{y-r/2}^{y+r/2} f(x_0) \int_{x_0-r}^{2y-x_0} \frac{f(x_1)}{F(x_0+r)-F(x_0-r)}\,dx_1 dx_0.$$

Proof: We can split the event $\{P_0 \le y\}$ into two mutually exclusive events. The first event is $\{X_0 \le y - r/2\}$, in which any value of X_1 makes $\{P_0 \le y\}$. The second event is that X_0 is bigger, but X_1 is small enough that P_0 is less than y anyway: this happens when $(\{y-r/2 \le X_0 \le y+r/2\})$ and $(\{X_0-r \le X_1 \le 2y - X_0\})$. The two terms in the cdf expression in the lemma reflect the probabilities of these two mutually exclusive events, where $f(x_1)/F(x_0+r) - F(x_0-r)$ is the pdf of X_1 in the interval $(X_0 - r, X_0 + r)$.

Lemma A.2: The random variable P_0 has pdf:

$$f_p(y) = \int_{y-r/2}^{y+r/2} \frac{2f(x)f(2y-x)}{F(x+r)-F(x-r)}\,dx.$$

Proof: We begin with the definition of pdf and the result of Lemma A.1:

$$\begin{aligned}
f_p(y) &= \tfrac{d}{dy}\Pr[P_0 \le y] \\
&= \tfrac{d}{dy}\left(\int_{-\infty}^{y-r/2} f(x)\,dx \right. \\
&\qquad \left. + \int_{y-r/2}^{y+r/2} f(x_0) \int_{x_0-r}^{2y-x_0} \tfrac{f(x_1)}{F(x_0+r)-F(x_0-r)}\,dx_1 dx_0\right) \\
&= f\big(y-\tfrac{r}{2}\big) \\
&\quad + \tfrac{d}{dy}\left(\int_{y-r/2}^{y+r/2} f(x_0) \int_{x_0-r}^{2y-x_0} \tfrac{f(x_1)}{F(x_0+r)-F(x_0-r)}\,dx_1 dx_0\right) \\
&= f\big(y-\tfrac{r}{2}\big) + \tfrac{d}{dy}\left(\int_{y-r/2}^{y+r/2} f(x_0) \tfrac{F(2y-x_0)-F(x_0-r)}{F(x_0+r)-F(x_0-r)}\,dx_0\right) \\
&= f\big(y-\tfrac{r}{2}\big) - f\big(y-\tfrac{r}{2}\big) \tfrac{F(2y-\left(y-\frac{r}{2}\right))-F((y-\frac{r}{2})-r)}{F((y-\frac{r}{2})+r)-F((y-\frac{r}{2})-r)} \\
&\quad + \int_{y-r/2}^{y+r/2} \tfrac{f(x_0)f(2y-x_0)\cdot 2}{F(x_0+r)-F(x_0-r)}\,dx_0 \\
&\quad + f\big(y+\tfrac{r}{2}\big) \tfrac{F(2y-(y+\frac{r}{2}))-F((y+\frac{r}{2})-r)}{F((y+\frac{r}{2})+r)-F((y+\frac{r}{2})-r)} \\
&= f\big(y-\tfrac{r}{2}\big) - f\big(y-\tfrac{r}{2}\big) \tfrac{F(y+\frac{r}{2})-F(y-\frac{3d}{2})}{F(y+\frac{r}{2})-F(y-\frac{3d}{2})} \\
&\quad + \int_{y-r/2}^{y+r/2} \tfrac{f(x_0)f(2y-x_0)\cdot 2}{F(x_0+r)-F(x_0-r)}\,dx_0 \\
&\quad + f\big(y+\tfrac{r}{2}\big) \tfrac{F\left(y-\frac{r}{2}\right)-F\left(y-\frac{r}{2}\right)}{F(y+\frac{3d}{2})-F(y-\frac{r}{2})} \\
&= \int_{y-r/2}^{y+r/2} \tfrac{f(x_0)f(2y-x_0)\cdot 2}{F(x_0+r)-F(x_0-r)}\,dx_0,
\end{aligned}$$

where we have used the fundamental theorem of calculus to take derivatives, and a change of dummy variables completes the theorem statement.

Properties of Seeded Localized Averaging:

Lemma A.3: If $f(x)$ is symmetric about μ, then $f_p(y)$ is also symmetric about μ.

Proof: We show that $f_p(\mu-\delta) = f_p(\mu+\delta)$. We start with the left hand side and the definition of $f_p(x)$ from Lemma A.2:

$$f_p(\mu-\delta) = \int_{\mu-\delta-r/2}^{\mu-\delta+r/2} \frac{2f(x)f(2(\mu-\delta)-x)}{F(x+r)-F(x-r)}\,dx.$$

Now, we substitute $x = \mu-\delta-r/2+z$ to get

$$\begin{aligned}
f_p(\mu-\delta) &= \int_0^r \frac{2f\left(\mu-\delta-\frac{r}{2}+z\right) f\left(\mu-\delta+\frac{r}{2}-z\right)}{F\left(\mu-\delta+\frac{r}{2}+z\right)-F\left(\mu-\delta-\frac{3d}{2}+z\right)}\,dz \\
&= \int_0^r \frac{2f\left(\mu-\left(\delta+\frac{r}{2}-z\right)\right) f\left(\mu+\left(-\delta+\frac{r}{2}-z\right)\right)}{F(\mu+(-\delta+\frac{r}{2}+z))-F\left(\mu-\left(\delta+\frac{3d}{2}-z\right)\right)}\,dz.
\end{aligned} \tag{1}$$

We continue with the right hand side

$$f_p(\mu+\delta) = \int_{\mu+\delta-r/2}^{\mu+\delta+r/2} \frac{2f(x)f(2(\mu+\delta)-x)}{F(x+r)-F(x-r)}\,dx,$$

and we substitute $x = \mu+\delta+r/2-z$ to get

$$\begin{aligned}
f_p(\mu+\delta) &= \int_r^0 -\frac{2f\left(\mu+\delta+\frac{r}{2}-z\right) f\left(\mu+\delta-\frac{r}{2}+z\right)}{F\left(\mu+\delta+\frac{3d}{2}-z\right)-F\left(\mu+\delta-\frac{r}{2}-z\right)}\,dz \\
&= \int_0^r \frac{2f\left(\mu+\left(\delta+\frac{r}{2}-z\right)\right) f\left(\mu-\left(-\delta+\frac{r}{2}-z\right)\right)}{F\left(\mu+\left(\delta+\frac{3d}{2}-z\right)\right)-F\left(\mu-\left(-\delta+\frac{r}{2}+z\right)\right)}\,dz.
\end{aligned} \tag{2}$$

Because $f(x)$ is symmetric about μ, for any value b we have $f(\mu-b) = f(\mu+b)$, and thus the numerators of (1) and (2) are equal. Also because $f(x)$ is symmetric about μ, for any values a and b we have $F(\mu+a)-F(\mu-b) = F(\mu+b)-F(\mu-a)$, and thus the denominators of (1) and (2) are equal. Intuitively, the property $F(\mu+a)-F(\mu-b) = F(\mu+b)-F(\mu-a)$ simply says that the area under the curve $f(x)$ in the interval $(\mu-b, \mu+a)$ is the same as the area in the interval $(\mu-a, \mu+b)$.

Because both the numerators and denominators of (1) and (2) are equal, we have that $f_p(\mu-\delta) = f_p(\mu+\delta)$.

Corollary A.4: If $f(x)$ is symmetric with mean μ, then $f_p(y)$ is also symmetric with the same mean, μ.

Proof: Follows from Lemma A.3.

Lemma A.5: If $f(x)$ is symmetric and increasing towards its mean, μ, then P_0 has a smaller variance than X_0.

Proof: Without loss of generality, we assume $\mu = 0$. The lemma statement then reduces to showing $E[(X_0+X_1/2)^2] < E[X_0^2]$. We begin by using linearity of expectation

$$\begin{aligned}
E\left[\left(\frac{X_0+X_1}{2}\right)^2\right] &= E\left[\frac{1}{4}\left(X_0^2+X_1^2+2X_0X_1\right)\right] \\
&= \frac{1}{4}\left(E\left[X_0^2\right]+E\left[X_1^2\right]+2E\left[X_0X_1\right]\right) \\
&\le \frac{1}{4}\left(2E\left[X_0^2\right]+2E\left[X_0X_1\right]\right) \\
&= \frac{1}{2}\left(E\left[X_0^2\right]+E\left[X_0X_1\right]\right)
\end{aligned}$$

where the inequality is because X_1 is a restricted version of X_0.

To show the lemma statement, the only thing that remains is to show $E\left[X_0X_1\right] < E[X_0^2]$. We have

$$\begin{aligned}
E\left[X_0X_1\right] &= \int_{-\infty}^{\infty} f(x_0)E[x_0X_1|X_0=x_0]\,dx_0 \\
&= \int_{-\infty}^{\infty} f(x_0)x_0E[X_1|X_0=x_0]\,dx_0.
\end{aligned} \tag{3}$$

Because $f(x)$ is symmetric and increasing towards the mean (0, since we assume $\mu = 0$) we have: 1) if $x_0 = \mu = 0$ then $E[X_1|X_0 = x_0] = x_0$ and 2) if $x_0 \ne \mu$ then $|E[X_1|X_0 = x_0]| < |x_0|$. Thus, continuing from (3),

$$\begin{aligned}
E\left[X_0X_1\right] &= \int_{-\infty}^{\infty} f(x_0)x_0E[X_1|X_0=x_0]\,dx_0 \\
&< \int_{-\infty}^{\infty} f(x_0)x_0^2\,dx_0 \\
&= E[X_0^2]
\end{aligned}$$

completing the proof.

ACKNOWLEDGMENT

N.B.D. thanks the Naval Postgraduate School's Research Initiation Program that partially supported this research.

REFERENCES

[1] T. B. Gosnell, J. M. Hall, C. L. Jam, D. A. Knapp, Z. M. Koenig, S. J. Luke, B. A. Pohl, A. S. von Wittenau, and J. K. Wolford, Gamma-Ray Identification of Nuclear Weapon Materials Lawrence Livermore National Lab, Tech. Rep. UCRL-ID-127436, 1997.

[2] J. V. Candy, E. Breitfeller, B. L. Guidry, D. Manatt, K. Sale, D. H. Chambers, M. A. Axelrod, and A. M. Meyer, "Physics-based detection of radioactive contraband: A sequential bayesian approach," *IEEE Transactions on Nuclear Science*, vol. 56, no. 6, pp. 3694–3711, 2009.

[3] J. V. Candy, D. H. Chambers, E. F. Breitfeller, B. L. Guidry, J. M. Verbeke, M. A. Axelrod, K. E. Sale, and A. M. Meyer, "Threat detection of radioactive contraband incorporating compton scattering physics: A model-based processing approach," *IEEE Transactions on Nuclear Science*, vol. 58, no. 1, pp. 214–230, 2011.

[4] G. F. Knoll, *Radiation Detection and Measurement*, 4th ed. : Wiley, 2010.

[5] V. T. Jordanov, "Radiation spectroscopy using seeded localized averaging (SLA)," in *IEEE Nuclear Science Symposium Conference Record*, 2005, vol. 1. IEEE, pp. 216–220.

[6] V. T. Jordanov, "Method and Apparatus for Radiation Spectroscopy Using Seeded Localized Averaging (SLA)," US Patent 7,269,523, Sep. 2007.

[7] L. J. Meng and D. Ramsden, "An inter-comparison of three spectral-deconvolution algorithms for gamma-ray spectroscopy," *IEEE Transactions on Nuclear Science*, vol. 47, no. 4, pp. 1329–1336, 2000.

[8] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *Journal of the ACM*, vol. 9, no. 1, pp. 84–97, 1962.

[9] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 113–122, 1982.

[10] E. T. Jaynes, *How Does the Brain do Plausible Reasoning?*. : Kluwer Academic Publishers, 1988, vol. 1, pp. 1–24.

[11] V. T. Jordanov, "Method for Radiation Spectroscopy Using Seeded Localized Averaging (SLA) and Channel-Address Recycling," US Patent 7,302,353, Nov. 2007.