Ibrahim Al-Nahhal, Octavia A. Dobre, Ertugrul Basar,
Cecilia Moloney, and Salama Ikki

A Fast, Accurate, and Separable Method for Fitting a Gaussian Function

The Gaussian function (GF) is widely used to explain the behavior or statistical distribution of many natural phenomena as well as industrial processes in different disciplines of engineering and applied science. For example, the GF can be used to model an approximation of the Airy disk in image processing, a laser heat source in laser transmission welding [1], practical microscopic applications [2], and fluorescence dispersion in flow cytometric deoxyribonucleic acid histograms [3]. In applied sciences, the noise that corrupts the signal can be modeled by the Gaussian distribution according to the central limit theorem. Thus, by fitting the GF, researchers can develop a sound interpretation of the corresponding process or phenomenon behavior.

This article introduces a novel fast, accurate, and separable (FAS) algorithm for estimating the GF parameters to fit observed data points. A simple mathematical trick can be used to calculate the area under the GF in two ways. Then, by equating these two areas, the GF parameters can be easily obtained from the observed data.

GF-fitting approaches

A GF has a symmetrical bell shape around its center, with a width that smoothly decreases as it moves away from its center on the x-axis. The mathematical form of the GF is

$$y(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{1}$$

with three shape-controlling parameters, A, μ, and σ, where A is the maximum height (amplitude) that can be achieved on the y-axis, μ is the curve center (mean) on the x-axis, and σ is the standard deviation (SD), which controls the width of the curve along the x-axis. This article presents a new method for the accurate estimation of these three parameters. The difficulty of this lies in estimating the three shape-controlling parameters (A, μ, and σ) from observations, which are generally noisy, by solving an overdetermined nonlinear system of equations.

The standard solutions for fitting the GF parameters from noisy observed data are obtained by one of the following two approaches:

1) *Solving the problem as a nonlinear system of equations using one of the least-squares optimization algorithms*: This solution employs an iterative procedure, such as the Newton–Raphson algorithm [4]. The drawbacks of this approach are the iterative procedure, which may not converge to the true solution, and its high cost from the computational complexity perspective.
2) *Solving the problem as a linear system of equations based on the fact that the GF is an exponential of a quadratic function*: By taking the natural logarithm of the observed data, the problem can be solved in polynomial time as a 3×3 linear system of equations. Two traditional algorithms have been proposed in this context: Caruana's algorithm [5] and Guo's algorithm [6]. Furthermore, instead of taking the natural logarithm, the partial derivative is used in Roonizi's algorithm [7].

In this article, we consider only the second approach, which is more suitable for most scientific applications, due to its simplicity and because it avoids the drawbacks of the first approach. Let us start with a brief introduction of the existing three algorithms for the second approach.

Caruana's algorithm

Caruana's algorithm exploits the fact that the GF is an exponential of a quadratic function and transforms it into a linear form by taking the natural logarithm of (1) to obtain

$$\begin{aligned} \ln(y) &= \ln(A) + \frac{-(x-\mu)^2}{2\sigma^2} \\ &= \ln(A) - \frac{\mu^2}{2\sigma^2} + \frac{2\mu x}{2\sigma^2} - \frac{x^2}{2\sigma^2} \\ &= a + bx + cx^2, \end{aligned} \tag{2}$$

where $a = \ln(A) - \mu^2/(2\sigma^2)$, $b = \mu/\sigma^2$, and $c = -1/(2\sigma^2)$. Accordingly, the

Digital Object Identifier 10.1109/MSP.2019.2927685
Date of current version: 29 October 2019

unknowns become a, b, and c in the linear equation (2) instead of A, μ, and σ in the nonlinear equation (1). Next, if the observations y are noisy, then they can be modeled as $\hat{y} = y + w$. Each contains the ideal data point, y, that is corrupted by the noise, w, with SD of σ_w. Note that in (2), we consider only the observations that have values above zero.

Once we have an overdetermined linear system, the unknowns can be estimated using the least-squares method. Caruana's algorithm estimates the three unknowns (a, b, and c) in (2) using the least-squares method by forming the error function, ε, for (2) as

$$\begin{aligned}\varepsilon &= \ln(\hat{y}) - \ln(y)\\ &= \ln(\hat{y}) - (a + bx + cx^2).\end{aligned} \tag{3}$$

Then, by differentiating the sum of ε^2 with respect to a, b, and c and equating the results to zero, we obtain three equations, which represent the following linear system:

$$\begin{bmatrix} N & \sum x_n & \sum x_n^2 \\ \sum x_n & \sum x_n^2 & \sum x_n^3 \\ \sum x_n^2 & \sum x_n^3 & \sum x_n^4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum \ln(\hat{y}_n) \\ \sum x_n \ln(\hat{y}_n) \\ \sum x_n^2 \ln(\hat{y}_n) \end{bmatrix}, \tag{4}$$

where N is the number of observed data points and $\sum$ denotes $\sum_{n=1}^{N}$. In this case, the parameters a, b, and c can be determined simply by solving (4) as a determined linear system of equations. Subsequently, the original parameters of the GF are determined as

$$A = e^{a - \frac{b^2}{4c}}, \quad \mu = \frac{-b}{2c}, \quad \sigma = \sqrt{\frac{-1}{2c}}. \tag{5}$$

The weighted least-squares method is the second candidate method to estimate the unknowns, and it is expected to have a better estimation accuracy than the least-squares method.

Guo's algorithm

Guo's algorithm, a modified version of the Caruana algorithm, finds the unknowns a, b, and c in (2) using the weighted least-squares method. It uses the noisy observed data, $\hat{y}$, to weight the error function in (3). Therefore, the error equation in (3) becomes $\delta = \hat{y}\varepsilon = \hat{y}[\ln(\hat{y}) - (a + bx + cx^2)]$, and the linear system of equations in (4) becomes

$$\begin{bmatrix} \sum \hat{y}_n^2 & \sum x_n \hat{y}_n^2 & \sum x_n^2 \hat{y}_n^2 \\ \sum x_n \hat{y}_n^2 & \sum x_n^2 \hat{y}_n^2 & \sum x_n^3 \hat{y}_n^2 \\ \sum x_n^2 \hat{y}_n^2 & \sum x_n^3 \hat{y}_n^2 & \sum x_n^4 \hat{y}_n^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum \hat{y}_n^2 \ln(\hat{y}_n) \\ \sum x_n \hat{y}_n^2 \ln(\hat{y}_n) \\ \sum x_n^2 \hat{y}_n^2 \ln(\hat{y}_n) \end{bmatrix}. \tag{6}$$

Moreover, the values of A, μ, and σ can be computed from (5).

One of the problems that affects the estimation accuracy is the long-tail GF. This occurs when the number of small values in the observed data is large compared to the observed data length, N, which means that a large amount of noise exists in those observations. Thus, an iterative procedure is required to improve the estimation accuracy.

Guo's algorithm with iterative procedure

The estimation accuracy of the Guo's algorithm deteriorates for a long-tail GF. To increase the accuracy of fitting the long-tail Gaussian parameters, an iterative procedure for (6) is given as

$$\begin{bmatrix} \sum \hat{y}_{n,(k-1)}^2 & \sum x_n \hat{y}_{n,(k-1)}^2 & \sum x_n^2 \hat{y}_{n,(k-1)}^2 \\ \sum x_n \hat{y}_{n,(k-1)}^2 & \sum x_n^2 \hat{y}_{n,(k-1)}^2 & \sum x_n^3 \hat{y}_{n,(k-1)}^2 \\ \sum x_n^2 \hat{y}_{n,(k-1)}^2 & \sum x_n^3 \hat{y}_{n,(k-1)}^2 & \sum x_n^4 \hat{y}_{n,(k-1)}^2 \end{bmatrix} \times \begin{bmatrix} a_{(k)} \\ b_{(k)} \\ c_{(k)} \end{bmatrix} = \begin{bmatrix} \sum \hat{y}_{n,(k-1)}^2 \ln(\hat{y}_n) \\ \sum x_n \hat{y}_{n,(k-1)}^2 \ln(\hat{y}_n) \\ \sum x_n^2 \hat{y}_{n,(k-1)}^2 \ln(\hat{y}_n) \end{bmatrix}, \tag{7}$$

where $\hat{y}_{n,(k)} = \hat{y}_n$ for $k = 0$ and $\hat{y}_{n,(k)} = e^{a_{(k)} + b_{(k)} x_n + c_{(k)} x_n^2}$ for $k > 0$, with the parenthesized subscripts denoting the indices of iteration.

Roonizi's algorithm

Roonizi's algorithm is designed to fit the GF riding on a polynomial background. It can be used to fit a GF by taking the partial derivative of (1), and then taking the integral of the result to obtain

$$y(x) = \beta_1 \phi_1(x) + \beta_2 \phi_2(x), \tag{8}$$

where $\beta_1 = -1/\sigma^2$, $\beta_2 = \mu/\sigma^2$, and

$$\phi_1(x) = \int_{-\infty}^{x} u y(u)\, du, \quad \phi_2(x) = \int_{-\infty}^{x} y(u)\, du. \tag{9}$$

In a manner similar to the steps in the Caruana and Guo algorithms, the error of (8) becomes $\zeta = \hat{y} - (\beta_1 \phi_1(x) + \beta_2 \phi_2(x))$. A linear system of equations results as follows:

$$\begin{bmatrix} \sum |\phi_1(x_n)|^2 & \sum \phi_1(x_n)\phi_2(x_n) \\ \sum \phi_1(x_n)\phi_2(x_n) & \sum |\phi_2(x_n)|^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum \phi_1(x_n)\hat{y}_n \\ \sum \phi_2(x_n)\hat{y}_n \end{bmatrix}. \tag{10}$$

By solving (10) in terms of β_1 and β_2, the estimated $\hat{\mu}$ and $\hat{\sigma}$ of the GF can be calculated as

$$\hat{\sigma} = \sqrt{\frac{-1}{\beta_1}}, \quad \hat{\mu} = \frac{-\beta_2}{\beta_1}. \tag{11}$$

Finally, using $\hat{\mu}$ and $\hat{\sigma}$ from (11), the estimated $\hat{A}$ of the GF can be calculated as

$$\hat{A} = \frac{\sum \left(\hat{y}_n \exp\left(\frac{-(x_n - \hat{\mu})^2}{2\hat{\sigma}^2} \right) \right)}{\sum \exp\left(\frac{-(x_n - \hat{\mu})^2}{2\hat{\sigma}^2} \right)}. \tag{12}$$

Note that the Roonizi's algorithm has no iterative procedure to increase the accuracy of fitting long-tail GF parameters.

Motivation

Guo's and Roonizi's algorithms have better estimation accuracy than Caruana's algorithm, while their computational complexity burden is comparable. Moreover, the three algorithms dependently estimate the GF parameters (A, μ, and σ). This means that, in some applications that require the estimation of only one parameter, the fitting algorithm may require unnecessary parameters to be estimated as well. Therefore, there is a need for a new method that provides better estimation accuracy with an efficient computational complexity as well as

the capability for a separable parameter estimation.

Proposed algorithm

In this article, we propose a novel FAS algorithm for a GF that accurately fits the observed data. The basic idea of the proposed FAS algorithm is to find a direct formula for the SD (i.e., σ) parameter from the noisy observed data. Then the amplitude A and mean μ can be determined using the weighted least-squares method for only two unknowns.

Derivation of the SD formula

To derive an approximation formula for the SD, a simple mathematical trick is applied. For N observations that represent the GF, as shown in Figure 1, the area under the GF can be divided into thin vertical rectangles with a width of Δx_n, where Δx_n is the nth step size of two successive observation points on the x-axis. Therefore, the total area under the GF, Λ, is numerically calculated as the summation of the areas of the vertical rectangles:

$$\Lambda \approx \sum_{n=1}^{N} \Delta x_n \hat{y}_n. \tag{13}$$

Note that (13) reflects at least 99.7% of the GF area in case of an available observation width greater than $\mu \pm 3\sigma$. Now, let us calculate the area under the GF using a different method. From the GF and Q-function properties, the total area under the GF is given as

$$\Lambda = \int_{-\infty}^{\infty} A e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = A\sigma\sqrt{2\pi}. \tag{14}$$

Equating (13) and (14), and replacing the amplitude A by the maximum value of the observed data, $\hat{y}_{\max}$, the estimated σ is obtained as

$$\hat{\sigma} = \frac{\sum_{n=1}^{N} \Delta x_n \hat{y}_n}{\sqrt{2\pi}\,\hat{y}_{\max}}. \tag{15}$$

Thus, in certain applications that require the estimate of the SD of the GF, the FAS algorithm directly outputs this estimate, without estimating the other two parameters. This is referred to as the *separable property* of the FAS algorithm.

Error analysis

To study the error of (15), first let us discuss the systematic error resulting from equating (13) and (14). This error becomes notable when a small portion of the GF is sampled, and the GF curve is approximated by rectangles (as in Figure 1). Based on extensive testing of the algorithm with varying parameters, as discussed further later in the article, the systematic error can be considered negligible when $W > 6$ and the observation samples are dense enough [e.g., $(N/W) > 10$], where W is the ratio of the SD to the observation width on the x-axis (i.e., the observation width equals $W\sigma$, or equivalently, it varies from $\mu - (W/2)\sigma$ to $\mu + (W/2)\sigma$).

To calculate the relative error in the numerator in (15), let the numerator equal $\sqrt{2\pi}A\sigma + \Delta x \sum_{n=1}^{N} w_n$, where $\sqrt{2\pi}A\sigma$ represents the actual area of the GF and $\Delta x \sum_{n=1}^{N} w_n$ is normally distributed with its SD being $\sqrt{N}\,\sigma_w \Delta x = \sqrt{N}\,\sigma_w (W\sigma/N)$. For simplicity of analysis, Δx is considered fixed for all observations. The relative error of the numerator, α_N, can be written as

$$\alpha_N \approx k_1 \frac{\sigma_w W}{\sqrt{2\pi}A\sqrt{N}} = k_1 \frac{W}{\mathrm{SNR}\sqrt{2\pi N}}, \tag{16}$$

where k_1 is a constant value, which can be considered 2 for the 95.5% confidence interval, and $\mathrm{SNR} = A/\sigma_w$ is the signal-to-noise ratio (SNR).

For the denominator, let us assume that it equals $\sqrt{2\pi}(A \pm \Delta A)$, where ΔA is the maximum of the normally distributed noise samples with SD of σ_w. The relative error of the denominator in (15), α_D, can be written as

$$\alpha_D \approx \frac{k_2 \sigma_w}{A} = \frac{k_2}{\mathrm{SNR}}, \tag{17}$$

where k_2 is a constant whose value can be assumed to be 3. (Based on comprehensive simulations, $k_2 = 3$ is the worst-case scenario for the error. Also, the probability of such a scenario is very low.) Hence, the total relative error in (15), α, can be approximated using a Taylor series as

$$\alpha \approx \alpha_N + \alpha_D = \frac{1}{\mathrm{SNR}}\left(k_1 \frac{W}{\sqrt{2\pi N}} + k_2\right). \tag{18}$$

If the samples are dense enough (i.e., large enough N/W), a reduced relative error can be attained for a high SNR.

Estimates of the remaining two parameters

To estimate the remaining two parameters A and μ using $\hat{\sigma}$ estimated from (15), we can differentiate the sum of δ^2 with respect to a and b and then equate the results to zero (i.e., using the same steps as in Guo's algorithm). The resulting linear system of equations becomes

$$\begin{bmatrix} \sum \hat{y}_n^2 & \sum x_n \hat{y}_n^2 \\ \sum x_n \hat{y}_n^2 & \sum x_n^2 \hat{y}_n^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum \hat{y}_n^2 \ln(\hat{y}_n) - c \sum x_n^2 \hat{y}_n^2 \\ \sum x_n \hat{y}_n^2 \ln(\hat{y}_n) - c \sum x_n^3 \hat{y}_n^2 \end{bmatrix}, \tag{19}$$

where $c = -1/(2\hat{\sigma}^2)$ and $\hat{\sigma}$ is the estimated SD, which is calculated from (15). Therefore, the values of a and b are obtained by solving the 2×2 linear system in (19); then, the original parameters A and μ can be calculated from (5).

Figure 2 shows the superiority of the proposed FAS algorithm over the traditional algorithms in the presence of a noise with SD $\sigma_w = 0.1$ for different values of N; the proposed algorithm

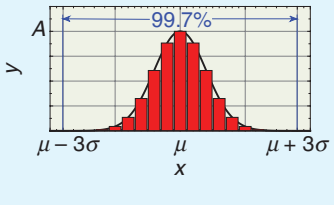

FIGURE 1. A graph illustrating a Gaussian function.

provides the best fit to the observed data points compared to the other fitting algorithms for all values of N. Figure 2 shows that $\hat{y}_{\max}$ is obviously different from the actual amplitude A. However, $\hat{\sigma}$ from (15) provides reasonable results using $\hat{y}_{\max}$ even if a small number of observation points are available, as in Figure 2(c).

Since the FAS algorithm provides poorer accuracy in fitting long-tail GF parameters, an iterative procedure is required to improve the fitting accuracy.

FAS algorithm with iterative procedure

For the long-tail GF, we propose an iterative algorithm that improves the fitting accuracy of the FAS algorithm. The recursive version of (19) is given as

$$\begin{bmatrix} \sum \hat{y}_{n,(k-1)}^2 & \sum x_n \hat{y}_{n,(k-1)}^2 \\ \sum x_n \hat{y}_{n,(k-1)}^2 & \sum x_n^2 \hat{y}_{n,(k-1)}^2 \end{bmatrix} \begin{bmatrix} a_{(k)} \\ b_{(k)} \end{bmatrix} = \begin{bmatrix} \sum \hat{y}_{n,(k-1)}^2 \ln(\hat{y}_n) - c \sum x_n^2 \hat{y}_{n,(k-1)}^2 \\ \sum x_n \hat{y}_{n,(k-1)}^2 \ln(\hat{y}_n) - c \sum x_n^3 \hat{y}_{n,(k-1)}^2 \end{bmatrix}, \tag{20}$$

where $\hat{y}_{n,(k)} = \hat{y}_n$ for $k = 0$, $\hat{y}_{n,(k)} = e^{a_{(k)} + b_{(k)} x_n + c x_n^2}$ for $k > 0$, and $\hat{\sigma}$ is estimated from (15) only once. This means that (15) can provide accurate results in fitting the long-tail GF without iteration, while the other two parameters still need to be estimated through iterations. However, after a few iterations, $\hat{\sigma}$ can be further improved by including an updated SD from (15) in the iterations, using A obtained by (20).

Figure 3 shows results of the iterative Guo's and proposed FAS algorithms for fitting a long-tail GF with $N = 200$, $A = 1$, $\sigma = 2$, and $\sigma_w = 0.1$ for $\mu = 18$ and 19, respectively. As we can see from the figure, the number of iterations required for the FAS algorithm to fit the long-tail GF is lower than that for Guo's algorithm. For example, in Figure 3(a), the FAS algorithm needs only three iterations to fit the observation; however, Guo's algorithm provides poor fitting for the same number of iterations. Note that, from Figure 3(b), as the tail of the

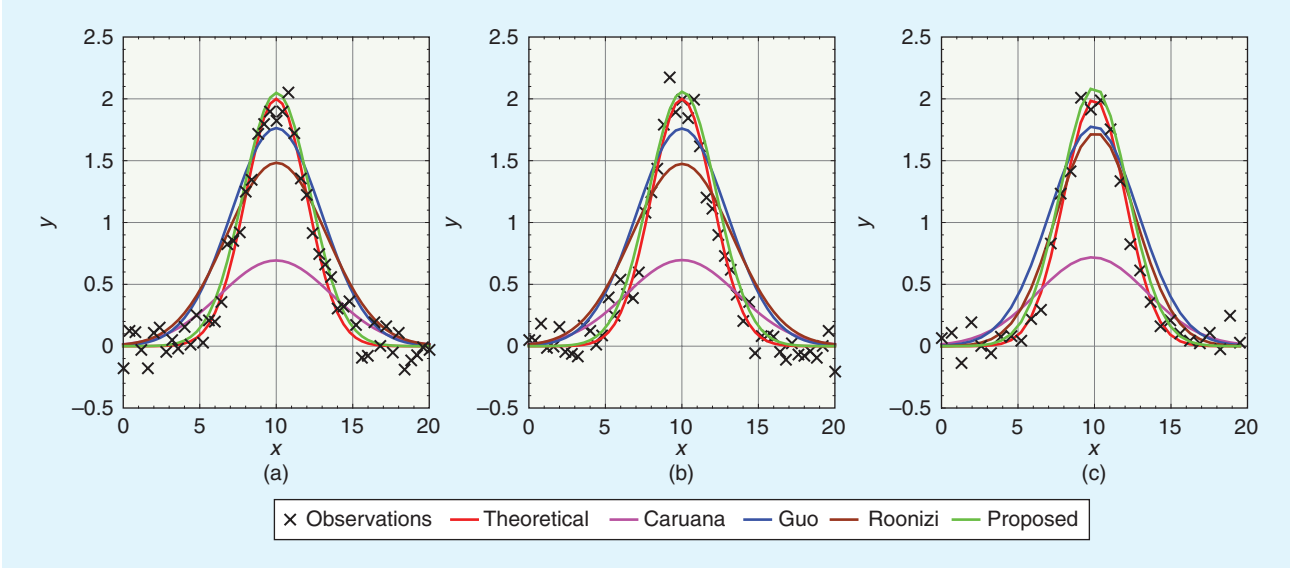

FIGURE 2. Graphs showing the results of different algorithms for fitting the GF with $A = 2$, $\sigma = 2$, and $\mu = 10$ in the presence of observation noise with $\sigma_w = 0.1$ (i.e., SNR = 10). (a) $N = 50$. (b) $N = 40$. (c) $N = 30$.

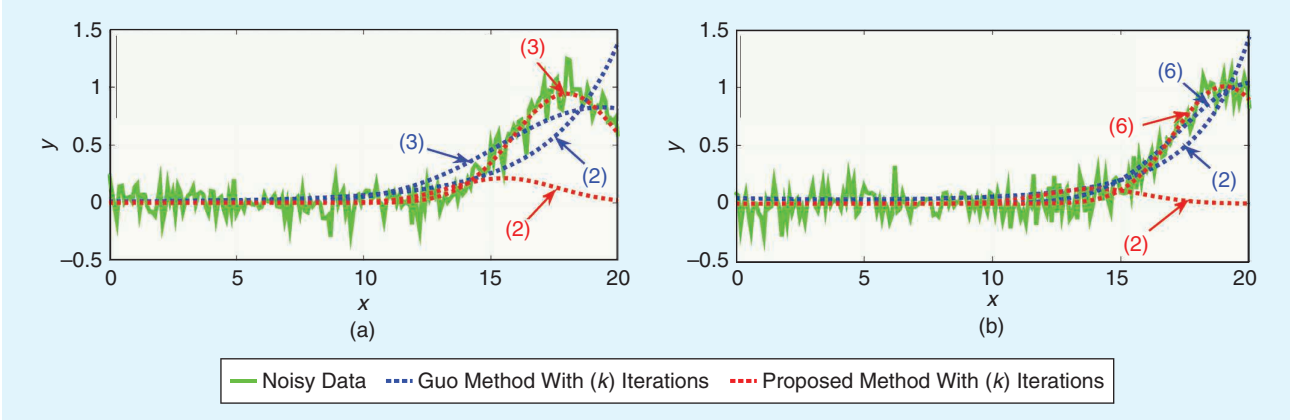

FIGURE 3. Graphs showing the results of the proposed FAS iterative algorithm in comparison with Guo's algorithm for fitting the GF of $N = 200$, $A = 1$, $\sigma = 2$, and $\sigma_w = 0.1$ (i.e., SNR = 10). (a) $\mu = 18$. (b) $\mu = 19$.

GF lengthens, more iterations are needed (i.e., six iterations are needed instead of three to provide a good fitting to the longer-tail GF). Even in the presence of considerable noise and with only a small portion of the GF, the iterative procedure of the proposed algorithm can nicely fit the GF after only a few iterations.

Accuracy comparison

In this section, Monte Carlo simulation results for at least 10^4 simulated trials are considered for comparing the average absolute relative error (ARE) of the fitting accuracy for the SD estimated using (15) and by traditional algorithms. The ARE percentile of the SD is given as $\text{ARE}\%(\sigma) = (|\hat{\sigma} - \sigma|/\sigma) \times 100\%$, where $|\cdot|$ denotes the absolute value and σ is the true SD. The GF parameters used for this simulation are $A = 1$, $\mu = 10$, and $\sigma = 2$. As demonstrated by the total relative error estimated in (18), three parameters can be used for assessing the accuracy of estimation (i.e., SNR, W, and N).

For the evaluation of the estimation accuracy, we calculate the average ARE%(σ), where one of the three parameters varies while the other two parameters are fixed. Figure 4 shows such results, where the SD is estimated using the proposed FAS algorithm in comparison with the three previously presented traditional algorithms. In Figure 4(a) and (b), $W = 12$ and the SNR varies from 1 to 100 for $N = 30$ and 200, respectively. Figure 4(c) and (d) depicts the effect of W, which varies from 2 to 24, for $N = 30$ and 200, respectively, in the case of SNR = 25 (i.e., $\sigma_w = 0.04$). Figure 4(e) and (f) shows the effect of N, which varies from 20 to 100 and from 200 to 1,000, respectively, with $W = 12$ and SNR = 25. It is obvious from these figures that the SD estimated from (15) has the lowest ARE% in all cases, except for $W < 6$ when Guo's algorithm is the best. This is called the *accurate property* of the FAS algorithm. In many practical applications, an adequate portion of the GF (i.e., $W \geq 6$) is sampled with more than 200 observation points (i.e., $N \geq 200$). Roonizi's algorithm is more general than the other techniques since it can also fit a Gaussian riding on a polynomial background. This might explain its poorer performance in comparison to the other algorithms that fit a sole GF as described by (1).

The plots in Figure 4 also depict the worst-case ARE% of the proposed algorithm. The simulated worst-case ARE% represents the maximum ARE% that occurs during the 10^4 simulated trials, which is compared to (18) with $k_1 = 2$ and $k_2 = 3$ to show the accuracy of our derived error estimated in (18). Note that the probability of such a worst-case error is very low. Notably, the worst-case theoretical and simulated ARE% match, except when $W < 6$ due to the

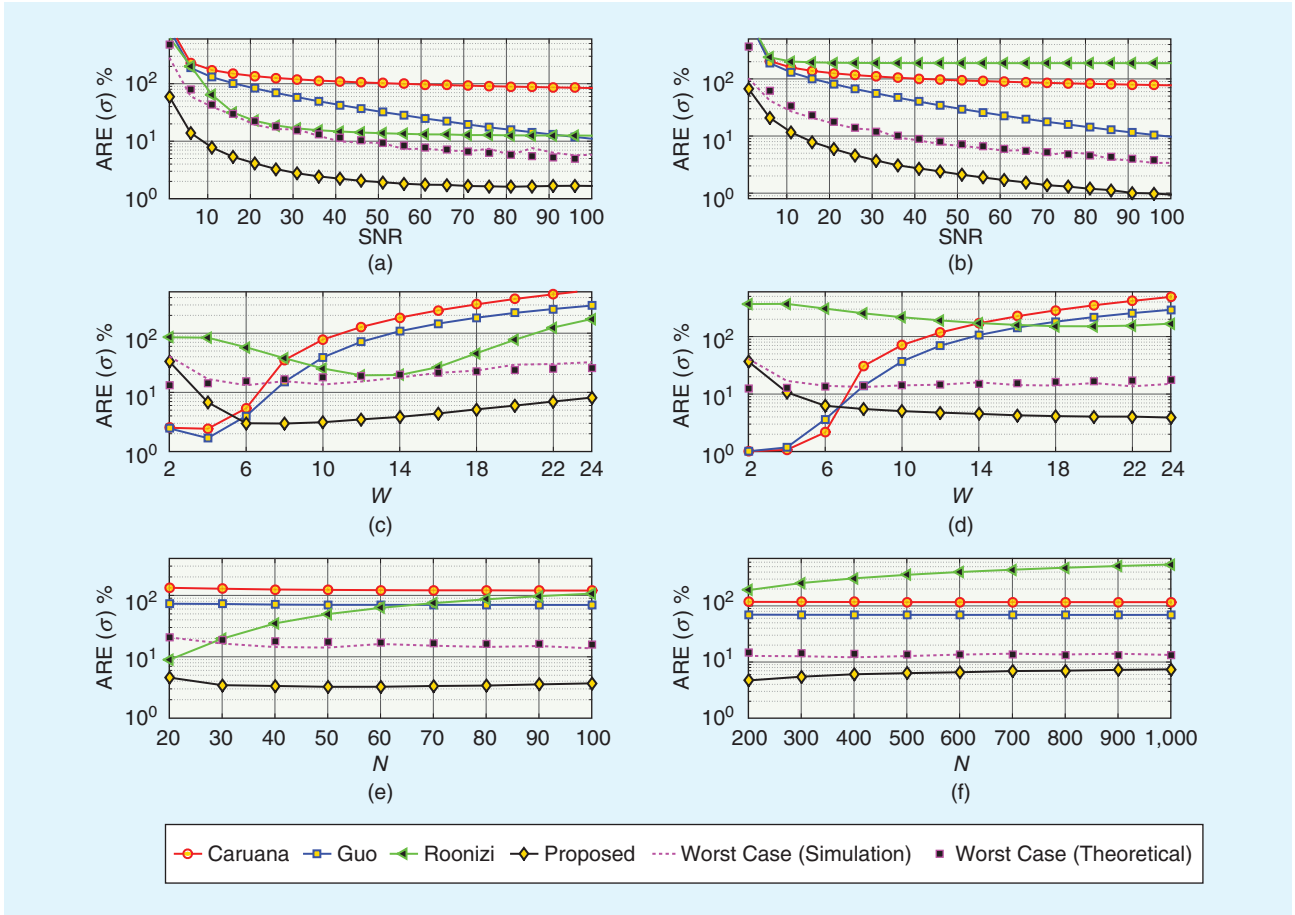

FIGURE 4. The ARE% of σ estimated from different fitting algorithms. (a) $W = 12$ and $N = 30$. (b) $W = 12$ and $N = 200$. (c) SNR = 25 and $N = 30$. (d) SNR = 25 and $N = 200$. (e) $W = 12$ and SNR = 25. (f) $W = 12$ and SNR = 25.

considerable systematic error. The superiority of the proposed algorithm versus the traditional ones holds for the worst-case ARE% as well; however, for the clarity of the plots in Figure 4, curves corresponding to the latter algorithms were not included. As shown in Figure 4(f), after a particular value of N, the error of the denominator in (15) becomes dominant. As N increases, there will be many samples around the peak of the GF, and the ARE% of the proposed algorithm slightly increases when N increases, finally approaching the worst-case scenario.

Complexity comparison

We address the computational complexity comparison of Guo's, Roonizi's, and the proposed FAS algorithms in terms of the number of additions and multiplications required to complete the fitting procedure. We assume that subtraction and division operations are respectively equivalent to addition and multiplication operations in complexity. It should be noted that solving an $n \times n$ linear system of equations using Gauss elimination requires $(2n^3 + 3n^2 - 5n)/6$ additions and $(n^3 + 3n^2 - n)/3$ multiplications [8]. Therefore, the total number of additions (Add) and multiplications (Mul) for the Guo, Roonizi, and FAS algorithms are given as follows:

$$\begin{aligned} \text{Add}^{(\text{Guo})} &= N(A_{\ln} + 8) + 3, \\ \text{Mul}^{(\text{Guo})} &= N(M_{\ln} + 11) + 17, \end{aligned} \tag{21}$$

$$\begin{aligned} \text{Add}^{(\text{Roonizi})} &= N^2 + 8N + NA_{\exp} - 5, \\ \text{Mul}^{(\text{Roonizi})} &= 0.5N^2 + 9.5N + NM_{\exp} + 9, \end{aligned} \tag{22}$$

$$\begin{aligned} \text{Add}^{(\text{FAS})} &= N(A_{\ln} + 8) - 3, \\ \text{Mul}^{(\text{FAS})} &= N(M_{\ln} + 10) + 12, \end{aligned} \tag{23}$$

where $A_{\ln}$ and $M_{\ln}$ represent the number of additions and multiplications required to calculate the natural logarithm, respectively, while $A_{\exp}$ and $M_{\exp}$ represent the number of additions and multiplications, respectively, required to calculate the natural exponential in (12). Note that the term of N^2 in (22) comes from the calculation of $\phi_1(x)$ in (9), which requires an accumulated numerical integration of $(uy(u))$ from the first observation point to the current value of x for all N observations.

It can be seen from (21) to (23) that the proposed algorithm requires fewer additions and multiplications when compared with Guo's and Roonizi's algorithms. Assuming $A_{\ln} = A_{\exp}$ and $M_{\ln} = M_{\exp}$, the proposed algorithm saves six additions and $\mathrm{O}(N)$ multiplications over Guo's algorithm, while it saves $\mathrm{O}(N^2)$ additions and multiplications over Roonizi's algorithm. This is referred to as the *fast property* of the proposed FAS algorithm.

Conclusions

This article proposed a simple approximation expression for the SD of a GF to fit a set of noisy observed data points. This expression results from a simple mathematical trick, which is based on the equality between the area under the GF calculated numerically and based on the Q-function properties. Then, the amplitude and mean of the GF can be calculated using the weighted least-squares method. Through comprehensive simulations and mathematical analysis, it has been shown that the proposed algorithm is not only faster than Guo's and Roonizi's algorithms, but also provides better estimation accuracy when an adequate interval of the GF is sampled. Additionally, an iterative procedure is proposed, which is suitable to fit the GF when the observed data points are contaminated with substantial noise, as in the case of a long-tail GF. It has been shown by extensive computer simulations that the proposed iterative algorithm fits the GF faster than the iterative Guo's algorithm. The proposed algorithm could be useful for several applications, such as Airy disk approximation, laser transmission welding, fluorescence dispersion, and many others involving digital signal processing.

Acknowledgments

We thank Prof. Roberto Togneri, *IEEE Signal Processing Magazine*'s area editor, columns and forum, and the anonymous reviewers for providing valuable suggestions to improve the manuscript. We are grateful to Prof. Balazs Bank for his assistance and insightful feedback during the revision of this article. The support of the Natural Sciences and Engineering Research Council of Canada through its Discovery program is also gratefully acknowledged. The work of Ertugrul Basar is supported in part by the Turkish Academy of Sciences Young Scientists Award Program.

Authors

Ibrahim Al-Nahhal (ioalnahhal@mun.ca) received his B.Sc. and M.Sc. degrees in electronics and communications engineering from Al-Azhar University and Egypt–Japan University for Science and Technology, Egypt, in 2007 and 2014, respectively. He is a Ph.D. student at Memorial University, Canada. Between 2008 and 2012, he was an engineer in industry and a teaching assistant in the Faculty of Engineering, Al-Azhar University in Cairo, Egypt. From 2014 to 2015, he was a physical-layer expert at Nokia, Belgium. He holds three patents. His research interests include designs for low-complexity receivers for emerging technologies; spatial modulation; multiple-input, multiple-output systems; and sparse code multiple access.

Octavia A. Dobre (odobre@mun.ca) received her Dipl. Ing. and Ph.D. degrees from the Polytechnic Institute of Bucharest, Romania, in 1991 and 2000, respectively. She is a professor and research chair at Memorial University, Canada. She was a visiting professor at the Massachusetts Institute of Technology and a Royal Society and Fulbright scholar. Her research interests include technologies for 5G and beyond, as well as optical and underwater communications. She has published more than 250 refereed papers in these areas. She was the editor-in-chief of *IEEE Communications Letters*. She has been a senior editor and an editor with prestigious journals as well as general chair and technical cochair of flagship conferences in her area of expertise. She is a Distinguished Lecturer of the IEEE Communications Society and a fellow of the Engineering Institute of Canada.

Ertugrul Basar (ebasar@ku.edu.tr) received his B.Sc. degree from Istanbul

University, Turkey, in 2007, and his M.S. and Ph.D. degrees from Istanbul Technical University, Turkey, in 2009 and 2013, respectively. He is currently an associate professor with the Department of Electrical and Electronics Engineering, Koç University, Istanbul, Turkey, and the director of Communications Research and Innovation Laboratory. His primary research interests include multiple-input, multiple-output systems; index modulation; waveform design; visible light communications; and signal processing for communications.

Cecilia Moloney (cmoloney@mun.ca) received her B.Sc. degree in mathematics from Memorial University, Canada, and her M.A.Sc. and Ph.D. degrees in systems design engineering from the University of Waterloo, Ontario, Canada. Since 1990, she has been a faculty member at Memorial University, where she is now a professor of electrical and computer engineering. From 2004 to 2009, she held the Natural Sciences and Engineering Research Council of Canada/Petro-Canada chair for Women in Science and Engineering, Atlantic Region. Her research interests include nonlinear signal and image processing methods, signal representations, radar signal processing, and methods for ethics in engineering and engineering education.

Salama Ikki (sikki@lakeheadu.ca) received his B.Sc. degree from Al-Isra University, Amman, Jordan, in 1996 and his Ph.D. degree in electrical engineering from Memorial University, Canada, in 2009. He is an associate professor in the Department of Electrical Engineering, Lakehead University, Ontario, Canada. From 2009 to 2010, he was a postdoctoral researcher at the University of Waterloo, Ontario, Canada. From 2010 to 2012, he was a research assistant with the Institut national de la recherche scientifique, University of Québec, Canada. He is the author of 100 journal and conference papers and has more than 4,000 citations and an H-index of 30. His research interests include cooperative networks; multiple-input, multiple-output systems; spatial modulation, and wireless sensor networks.

References

[1] H. Liu, W. Liu, X. Zhong, B. Liu, D. Guo, and X. Wang, "Modeling of laser heat source considering light scattering during laser transmission welding," *Mater. Des.*, vol. 99, pp. 83–92, June 2016.

[2] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Appl. Opt.*, vol. 46, no. 10, pp. 1819–1829, Apr. 2007.

[3] F. Lampariello, G. Sebastiani, E. Cordelli, and M. Spano, "Comparison of Gaussian and t-distribution densities for modeling fluorescence dispersion in flow cytometric DNA histograms," *Cytometry*, vol. 12, no. 4, pp. 343–349, Jan. 1991.

[4] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York: Cambridge Univ. Press, 2007.

[5] R. Caruana, R. Searle, T. Heller, and S. Shupack, "Fast algorithm for the resolution of spectra," *Anal. Chem.*, vol. 58, no. 6, pp. 1162–1167, May 1986.

[6] H. Guo, "A simple algorithm for fitting a Gaussian function," *IEEE Signal Process. Mag.*, vol. 28, no. 5, pp. 134–137, Sept. 2011.

[7] E. K. Roonizi, "A new algorithm for fitting a Gaussian function riding on the polynomial background," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1062–1065, Sept. 2013.

[8] G. Strang, *Introduction to Linear Algebra*, 3rd ed. Cambridge, MA: Wellesley-Cambridge Press, 1993.

SP

APPLICATIONS CORNER *(continued from page 132)*

seismic events from Llaima volcano (Chile)," in *Proc. 2018 Int. Joint Conf. Neural Networks,* pp. 1–8. doi: 10.1109/IJCNN.2018.8489285.

[45] R. Soto, F. Huenupan, P. Meza, M. Curilem, and L. Franco, "Spectro-temporal features applied to the automatic classification of volcanic seismic events," *J. Volcanology Geothermal Res.*, vol. 358, pp. 194–206, June 2018. doi: 10.1016/j.jvolgeores.2018.04.025.

[46] M. Curilem, F. Huenupan, D. Beltrán, C. San Martin, G. Fuentealba, L. Franco, C. Cardona, G. Acuña et al., "Pattern recognition applied to seismic signals of Llaima volcano (Chile): An evaluation of station-dependent classifiers," *J. Volcanology Geothermal Res.*, vol. 315, pp. 15–27, Apr. 2016. doi: 10.1016/j.jvolgeores.2016.02.006.

[47] S. M. Bhatti, M. S. Khan, J. Wuth, F. Huenupan, M. Curilem, L. Franco, and N. B. Yoma, "Automatic detection of volcano-seismic events by modeling state and event duration in hidden Markov models," *J. Volcanology Geothermal Res.*, vol. 324, pp. 134–143, Sept. 2016. doi: 10.1016/j.jvolgeores.2016.05.015.

[48] D. A. Firoozabadi, F. Seguel, I. Soto, D. Guevara, F. Huenupan, M. Curilem, and L. Franco, "Evaluation of Llaima volcano activities for localization and classification of LP, VT and TR events," *J. Electr. Eng.*, vol. 68, no. 5, pp. 325–338, 2017. doi: 10.1515/jee-2017-0064.

[49] J. C. Lahr, B. A. Chouet, C. D. Stephens, J. A. Power, and R. A. Page, "Earthquake classification, location, and error analysis in a volcanic environment: Implications for the magmatic system of the 1989–1990 eruptions at redoubt volcano, Alaska," *J. Volcanology Geothermal Res.*, vol. 62, no. 1–4, pp. 137–151, 1994. doi: 10.1016/0377-0273(94)90031-0.

[50] C. Bouvet de Maisonneuve, M. A. Dungan, O. Bachmann, and A. Burgisser, "Insights into shallow magma storage and crystallization at Volcán Llaima (Andean Southern Volcanic Zone, Chile)," *J. Volcanology Geothermal Res.*, vol. 211–212, pp. 76–91, Jan. 2012. doi: 10.1016/j.jvolgeores.2011.09.010.

[51] L. E. Franco, J. L. Palma, F. Gil-Cruz, and J. J. San Martín, "Descripción de la actividad sísmica relacionada a erupciones estrombolianas violentas: Vn. Llaima, Chile (2007–2010)," *Earth Sci. Res. J.*, vol. 18, pp. 338–339, July 2014.

[52] M. Curilem, R. F. de Mello, F. Huenupan, C. San Martin, L. Franco, E. Hernández, and R. A. Rios, "Discriminating seismic events of the Llaima volcano (Chile) based on spectrogram cross-correlations," *J. Volcanology Geotherm. Res*, vol. 367, pp. 63–78, Nov. 2018. doi: 10.1016/j.jvolgeores.2018.10.023.

[53] M. Titos, A. Bueno, L. García, M. C. Benítez, and J. Ibañez, "Detection and classification of continuous volcano-seismic signals with recurrent neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1936–1948, 2019. doi: 10.1109/TGRS.2018.2870202.

[54] M. Malfante, M. Dalla Mura, J. I. Mars, J.-P. Métaxian, O. Macedo, and A. Inza, "Automatic classification of volcano seismic signatures," *J. Geophys. Res.: Solid Earth*, vol. 123, no. 12, pp. 10,645–10,658, 2018. doi: 10.1029/2018JB015470.

[55] C. Hibert, F. Provost, J.-P. Malet, A. Maggi, A. Stumpf, and V. Ferrazzini, "Automatic identification of rockfalls and volcano-tectonic earthquakes at the Piton de la Fournaise volcano using a random forest algorithm," *J. Volcanology Geothermal Res.*, vol. 340, pp. 130–142, June 2017. doi: 10.1016/j.jvolgeores.2017.04.015.

[56] L. García, I. Álvarez, M. Titos, A. Díaz-Moreno, M. C. Benítez, and Á. de la Torre, "Automatic detection of long period events based on subband-envelope processing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 11, pp. 5134–5142, 2017. doi: 10.1109/JSTARS.2017.2739690.

[57] U. Stańczyk, B. Zielosko, and L. C. Jain, Eds., *Advances in Feature Selection for Data and Pattern Recognition*. Cham, Switzerland: Springer, 2018.

[58] L. Yu and H. Liu, "Efficient feature selection via analysis of relevance and redundancy," *J. Mach. Learning Res.*, vol. 5, pp. 1205–1224, Dec. 2004.

SP