IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 15 September 2022, accepted 2 October 2022, date of publication 5 October 2022, date of current version 12 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3212388

METHODS

A Fast Method for Fitting a Multidimensional Gaussian Function

ANITA GRIBL KOŠČEVIĆ, (Graduate Student Member, IEEE), AND DAVOR PETRINOVIĆ, (Member, IEEE)
Faculty of Electrical Engineering and Computing, University of Zagreb, 10 000 Zagreb, Croatia
Corresponding author: Anita Gribl Koščević (anita.gribl@fer.hr)

This work was supported in part by the Croatian Science Foundation under Project IP-2019-04-6703 and Project DOK-01-2018, and in part by the European Regional Development Fund (DATACROSS) under Grant KK.01.1.1.01.0009.

ABSTRACT This paper estimates the multidimensional Gaussian profile parameters from the noisy measurements in the exponential function's argument domain. The proposed method minimizes the weighted squared error between the natural logarithm of the model and the logarithm of the normalized input data with the weights set to the input data values or model values. The proposed method is an iterative method where the parameters of the covariance matrix and the profile's peak position are alternatively estimated. The main advantage of the proposed method is a one-step analytical solution for the parameters of the covariance matrix and the linear profile scale for the given initial centroid position for arbitrary dimensions. The profile's peak position is then updated given the estimated parameters by solving a system of nonlinear coupled equations using an iterative optimization procedure. Finally, the proposed method in the log domain is compared with the LS method in the domain of Gaussian profile values, where all profile parameters are simultaneously estimated using an iterative procedure for solving a system of nonlinear equations using numerical optimization. The proposed log domain estimation method yields similar results as the numerical LS method in the value domain for sufficiently high signal-to-noise ratios (SNRs) and narrow regions-of-interest (ROIs) concerning their precision. However, it converges much faster due to the analytic solution.

INDEX TERMS Multidimensional Gaussian profile fitting, weighted least-squares method, estimation in the log domain.

I. INTRODUCTION

The Gaussian profiles of various dimensions are widely applied in many engineering fields. For example, the 1D Gaussian profile is used in spectroscopy to fit emission or absorption spectral lines [1], [2], [3]. The 2D Gaussian profile is used in image processing to approximate an Airy disk in the case of a diffraction-limited imaging system or to approximate the image blur of the point source caused by different degradations during the imaging process. Also, it has an application in astrometry for the identification and tracking of stellar objects [2], [4], [5]. The 2D and 3D Gaussian profiles are also applied in PET/CT imaging for image reconstruction of a volume of interest [6] or microscopy for single-particle tracking [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Gaggero.

In this paper, we propose the iterative method for the parameter estimation of the Gaussian profile of arbitrary dimension in the log domain from the data corrupted with additive noise. Several papers have already dealt with the estimation of the 1D Gaussian profile parameters in the log domain from the noisy measurements using the least-squares (LS) method, where the estimation problem is simplified to parabola fitting [8], [9], [10]. However, for the multidimensional Gaussian profile, the number of parameters quadratically increases, and consequently, the estimation complexity. Moreover, the analytic LS solution for all model parameters exists only for the 1D case since optimal parameters are non-linearly coupled for higher dimensionality. The inverse covariance matrix in the quadratic form and the determinant of the covariance matrix in the normalization term in front of exponential additionally complicates the estimation of the proper linear profile's scale for higher profile dimensions.

The original approach proposed in this paper exploits the fact that minimizing the weighted sum of squared residuals in the log domain concerning the parameters of the inverse covariance matrix and one additional unknown, which is the residual vertical shift of the log target, yields a system of linear equations with the one-step analytical solution given the initial profile's peak position. However, minimizing the same objective function concerning the remaining parameters, i.e., the profile's peak position, yields a system of nonlinear equations that require an iterative optimization procedure given the previously analytically estimated parameters. Therefore, to reduce the computational costs and speed up the estimation of the multidimensional Gaussian profile parameters, we proposed a two-stage weighted least-squares (WLS) method in the log domain that combines both the analytical and iterative approaches described above. These two stages are iteratively exchanged depending on the required accuracy, but the method usually converges in only a few iterations.

The proposed method's accuracy and complexity are compared with the numerical LS method in the domain of values, where all profile parameters are estimated simultaneously. In addition, the 2D Gaussian profile parameters are estimated from data corrupted with additive noise, and the results are directly compared with those previously published in [11] for conventional value domain estimation.

The estimation weights were introduced to ensure that the reduction of weighted squared error in the log domain simultaneously reduces the sum of squared errors in the domain of values and two types are considered: model-driven or input data-driven estimation weights. In addition, we have also analyzed the influence of negative input samples and the difference in the informativeness of input data depending on the chosen size of the estimation domain, also called region-of-interest (ROI), on the accuracy of the model estimation in the log domain as a function of the input noise level.

II. RELATED WORK

There are a few approaches to estimating the Gaussian profile's parameters from the noisy data depending on the required precision, implementation complexity, and prior knowledge of some parameters or noise statistics [7]. Two approaches considering the estimation domain are the estimation in the domain of values or the domain of exponential function's argument.

The simplest method in the domain of values is the fast and non-iterative method of moments. The first moment calculates the expected peak location while the second central moment yields the expected profile's variances [12], [13]. Disadvantages of the moment-based methods are the sensitivity to noise and selected ROI and a bias with underestimated profile variances. However, this method can be applied in real-time applications such as particle tracking [14].

The second group of estimation methods consists of iterative methods that search for optimal profile parameters either in the least-squares (LS) or maximum likelihood (ML) sense. Both approaches require solving an overdetermined system of nonlinear equations using some of the iterative optimization techniques such as the quasi-Newton method [15], downhill simplex method [14], [16], and Levenberg–Marquardt algorithm [17], [18].

The commonly used LS method in the value domain minimizes the sum of squared residuals between the observations and the fitted model values [19]. However, it has high computational costs, does not guarantee the optimal solution, and highly depends on the initial guess of model parameters. The modified and more robust version of the least squares method for fitting a multidimensional Gaussian function was proposed in [20] where the regularization term that measures the closeness of the observations and the Gaussian function using the Kullback-Leibler divergence [21] was added.

On the other hand, the optimal solution can be searched in the ML sense wondering which parameters are most likely to yield the observed data. Moreover, unlikely the LS method, the ML estimation requires prior knowledge of noise distribution. Hagen et al. [22], [23] proposed the ML estimation of 1D and 2D Gaussian profile parameters from the data corrupted with additive Gaussian noise. Also, they provided analytical gradients and Hessian matrix of log-likelihood function and used a Newton method for parameter update. Additionally, they provided analytical expressions for estimated parameter variances by using the Cramer-Rao bound and inverting the Fisher information matrix. Namely, the ML estimate is unbiased and achieves the Cramer-Rao lower bound but requires many input samples.

The estimation can be transformed to the argument domain by taking the logarithm of the observed data and exponential function of the model (profile), thus avoiding the nonlinear optimization and simplifying the estimation problem to polynomial fitting. Caruana et al. [8] proposed a fast algorithm for the 1D Gaussian fitting using the LS method in the log domain by solving the overdetermined system of linear equations. However, despite the high speed of the proposed non-iterative method, the method's accuracy is strongly reduced in the presence of additive noise. Namely, the ratio of the noise variance and the squared profile value occurs in the expectation of the quadratic error in the log domain; consequently, the small profile values can significantly increase the total error. The loss of precision was especially emphasized at wide ROIs far from the profile's peak position, where additive noise and small profile values dominate. Guo [9] proposed the WLS estimation in the log domain with weights equal to Gaussian profile values to account for the influence of additive noise. Such weighting yields the expectation of the quadratic error independent of the Gaussian profile values and reduces the method's sensitivity to additive noise. Since the actual profile values were unknown, the input data values were used to estimate ideal weights. With the thresholding of small values, the method becomes even more precise. Additionally, in the case of long-tailed contamination, the method degrades significantly since the approximation of

error function with only the first term of the Taylor series is insufficient on such wide ROIs where the difference between the actual profile value and observation can be considerable. Guo [9] introduced the iterative procedure using the estimated model values from the last iteration as a weight for the next iteration to solve the problem of long-tailed noise contamination. Al-Nahal et al. [10] proposed a fast, accurate, and separable method for 1D Gaussian fitting in the log domain where the direct formula for the profile's standard deviation (STD) is derived by equating the total area under the Gaussian function obtained numerically and Q-function properties. At the same time, the remaining two parameters were calculated the same as in Guo's method. An iterative procedure is further introduced to solve the problem of long-tailed noise.

Another approach that avoids iterative procedure and transforms the nonlinear least-squares fitting into a standard linear least-squares fitting was proposed by Roonizi [24], which uses differentiation and integration and assumes that the Gaussian function is placed on the polynomial background. The method suffers from accumulated noise error from the numerical integration process which was solved in [25]. However, both methods are proposed only for the case of 1D Gaussian.

Anthony et al. [16] estimated the parameters of the 2D Gaussian profile in the log domain, but only the uncorrelated 2D Gaussian profile was considered. Furthermore, the influence of noise was neglected, and it was assumed that the background was removed.

This paper proposes the iterative method for the generalized multidimensional Gaussian profile fitting using the WLS method in the log domain. The method analytically estimates all remaining profile parameters, including the covariance matrix and the linear scale for the Gaussian profile of arbitrary dimension for a priorly known profile's peak position. The profile's peak position is then updated, given the previously estimated parameters. Moreover, the iterative procedure was introduced to increase the method's accuracy. To the best of our knowledge, this is the first case of approaching the estimation of Gaussian profile parameters in this way, and in addition, a more generalized multidimensional profile estimation is considered.

III. MULTIDIMENSIONAL GAUSSIAN PROFILE

The multi-dimensional Gaussian profile is a function of the form

$$f(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, A) = \frac{A}{\sqrt{(2\pi)^n |\boldsymbol{\Sigma}|}} \exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big), \quad (1)$$

where $\mathbf{x}^{(i)} = [x_1^{(i)}, \ldots, x_n^{(i)}]^T$ corresponds to the column vector of the ith input sample positions in n-dimensional space $\forall i \in (1, m)$ where m is the total number of estimation input samples, A is the unknown linear scale, $\boldsymbol{\mu} \in \mathbb{R}^n$ is the vector of centroid position, i.e., profile's peak position and $\boldsymbol{\Sigma} \in \mathbf{S}_{++}^n$ is the $n \times n$ covariance matrix that is symmetric, positive definite, invertible and of full rank. The estimation process involves optimization over all samples enclosed within the estimation input domain, the so-called, region-of-interest (ROI).

$(\mathbf{x^1}, \ldots, \mathbf{x^m}, \mathbf{z})$ are the given empirical input data, where the set of vectors $\mathbf{X} = (\mathbf{x^1}, \ldots, \mathbf{x^m})$ defines the domain of input samples, i.e., ROI, which does not necessarily need to be defined over a rectangular grid in n-dimensional space, but it is only assumed that it is bounded and symmetrically distributed around the expected peak position. The vector of input sample values $\mathbf{z} = [z^{(1)}, \ldots, z^{(m)}]^T$ represents the discrete sample values of n-dimensional probability density function (PDF) multiplied by an unknown linear scale A. Since the input domain is bounded, the PDF should fit the truncated Gaussian PDF, as described in subsection IV-F.

IV. AN ITERATIVE METHOD FOR GAUSSIAN PARAMETERS ESTIMATION IN THE ARGUMENT DOMAIN

This section describes the proposed iterative method for fitting a multidimensional Gaussian profile in the log domain. The illustration of the proposed method is given in Fig. 1. The main steps involve:

1) Initialization
   - Calculation of the spatial coverage factor k that determines the width of enclosed convex region-of-interest (ROI) around the profile's peak position
   - Calculation of the initial centroid position $\boldsymbol{\mu}_{init}$
   - Normalization of input data $\mathbf{z}$
2) Taking the logarithm of the normalized input data values $\ln(\mathbf{z}_n)$
3) Selection of the normalized estimation weights $\mathbf{w}_n$ as either initial model or input data-driven weights
4) Minimizing the weighted sum of squared residuals in the log domain by solving the system of linear equations to obtain the parameters of the inverse covariance matrix $\hat{\boldsymbol{\Sigma}^{-1}}$ and the residual vertical shift of the log target $\hat{z_0}$ given the centroid $\hat{\boldsymbol{\mu}}$
5) Update centroid position $\hat{\boldsymbol{\mu}}$ given the estimated parameters $\hat{\boldsymbol{\Sigma}}$ and $\hat{z_0}$
6) Update linear scale $\hat{A}$ by fitting the obtained model to input data as follows

$$\hat{A} = \frac{f(\mathbf{X}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 1)^T \mathbf{z}}{f(\mathbf{X}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 1)^T f(\mathbf{X}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 1)} \quad (2)$$

7) Repeat steps 4-6 using the profile's parameters estimated in the previous iteration until achieving the given stop criterion, such as convergence or a maximal number of iterations

The method finds an optimal solution that minimizes the weighted L_2 norm in the log domain concerning noisy input data. In the noiseless case, $m_{min} = \frac{n(n+1)}{2} + 1$ input samples are sufficient to find a unique and universal analytical solution to the linear system of equations in the first stage of the proposed method, regardless of chosen estimation weights, as long as those samples meet certain conditions described

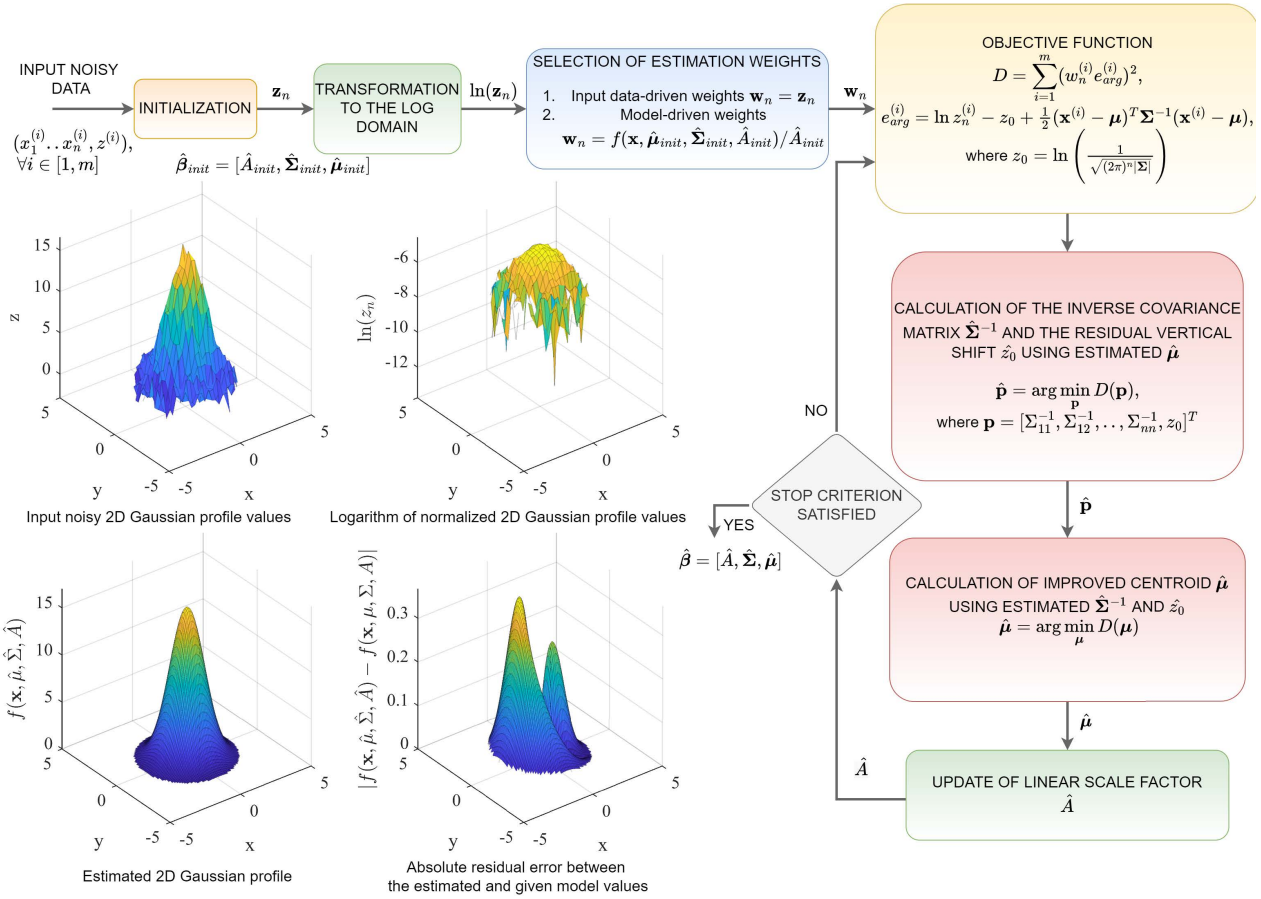


FIGURE 1. Illustration of the proposed method.

in subsection IV-C. Namely, m_{min} represents the number of unknowns including the unique elements of the covariance matrix $\mathbf{\Sigma}$, i.e., its inverse $\mathbf{\Sigma}^{-1}$, and the residual vertical shift of the log target z_0. The minimal number of input samples m_{min} is sufficient to define the model uniquely, but in that case, noise in input data can strongly affect the accuracy of the estimated parameters and their variances. Therefore, a much larger number of input samples is recommended to reduce the noise variance by averaging and to achieve a higher estimation accuracy. The proposed method attempts to iteratively correct the covariance matrix $\mathbf{\Sigma}$ and centroid $\boldsymbol{\mu}$ to obtain a model in each iteration closer to empirical noisy input data.

Due to the nonlinearity of the exponential function, the sum of squared errors in the log domain does not have the same meaning as the sum of squared errors in the domain of values. Thus, the sum of weighted squared errors is used in the log domain to obtain results that mimic those obtained by minimizing the sum of squared errors in the domain of values. Thereby, the estimation weights can be model-driven or input data-driven weights.

The advantage of optimization in the domain of argument is in the fact that the optimal inverse covariance matrix can be found as an analytical solution to the system of linear equations with $\frac{n(n+1)}{2}$ unknowns for a given centroid vector $\boldsymbol{\mu}$. Moreover, it is even better to extend this linear problem to the $\frac{n(n+1)}{2} + 1$ variables where an additional unknown is the residual vertical shift of the target in the log domain z_0, thus automatically solving the problem of the unknown residual scale factor of normalized input data that are not perfectly normalized. Namely, the initial scale factor A_{init} used for input data normalization is not optimal, and there is also an unknown determinant of the covariance matrix in the normalization term in front of the exponential. Ignoring these two quantities, the scale factor and determinant of the covariance matrix prevents the transition of the input data to the log domain, because both quantities cause the inseparable vertical shift of the log target.

After finding the optimal covariance matrix, $\mathbf{\Sigma}$, and the residual vertical shift in log domain z_0, an additional centroid correction can be made to find a better solution for the fixed covariance matrix that will further reduce the weighted squared error of the argument. Unfortunately, this correction has no analytical expression because the optimal solution is defined with n complete the third order polynomials in variables $[\mu_1 \ldots \mu_n]$. However, it is possible to find analytical

gradients of the weighted squared error of the argument for $[\mu_1 \dots \mu_n]$, as well as the corresponding Hessian matrix, and find the optimum using a numerical optimization procedure that typically converges in two or three iterations. By repeating the whole procedure with reasonable assumption on the convergence of such a two-step optimization procedure (alternately $\mathbf{\Sigma}$, then centroid $\boldsymbol{\mu}$), the new model will always be better than the previous one.

A. DETERMINATION OF ESTIMATION WEIGHTS

Estimating the Gaussian profile parameters in the log domain by minimizing the sum of weighted squared residuals requires a properly selected weight function to ensure that the sum of squared residuals in the domain of values is minimized as well. It is easily shown that the optimal sample weight is precisely equal to the noiseless profile value for the same sample position. The proof of this statement is given in Appendix A. Such a choice of estimation weights for weighting the error of the exponential function's argument is the best prediction of the error in the exponential function's value domain.

Since the ideal Gaussian profile values are usually unknown, the empirical input sample values can be used as estimation weights instead of the ideal profile values, but only for small argument error that consequently yields a small error in the domain of values. Therefore, this paper chooses the estimation weights either as input sample values or initial model values for the same sample positions. We also analyzed the effects of both types of estimation weights on estimation accuracy. To regularize the dynamic range of estimated profile parameters, we used normalized input data and, consequently, the normalized estimation weights, which are

$$\mathbf{w}_n = \mathbf{z}_n, \tag{3}$$

where $\mathbf{z}_n$ is the normalized input data column vector. In the case of model-driven weights, a normalized initial moment-based model of the form $f(\mathbf{X}, \hat{\boldsymbol{\mu}}_{init}, \hat{\mathbf{\Sigma}}_{init}, 1)$ is used. Such chosen weights are calculated only once, and the same weights are used in all iterations of the proposed method. Normalization of the input data is described in Subsection IV-F.

B. THE OBJECTIVE FUNCTION

The vector of parameters of the n-dimensional Gaussian profile is

$$\boldsymbol{\beta} = [A,\ \Sigma_{11}, \Sigma_{12}, \dots, \Sigma_{nn},\ \mu_{x_1}, \dots, \mu_{x_n}], \tag{4}$$

which includes the linear scale A, the unique members of the symmetric covariance matrix $\Sigma_{11}, \Sigma_{12}, \dots, \Sigma_{nn}$, and the profile's peak position $\mu_{x_1}, \dots, \mu_{x_n}$. The unique members of the symmetric covariance matrix, $\Sigma_{11}, \Sigma_{12}, \dots, \Sigma_{nn}$, are unambiguously defined by the unique members of its symmetric inverse, $\Sigma_{11}^{-1}, \Sigma_{12}^{-1}, \dots, \Sigma_{nn}^{-1}$, which occur directly in the quadratic form of the exponential function's argument. For each pixel position $\mathbf{x}^{(i)}$ within the selected ROI and its corresponding normalized value $z_n^{(i)}$, the sample's error in the log domain can be calculated as

$$e_{arg}^{(i)} = \ln z_n^{(i)} - \ln\left(f(\mathbf{x}^{(i)}, \boldsymbol{\mu}, \mathbf{\Sigma}, A)/A\right) \tag{5}$$

$$= \ln z_n^{(i)} - \ln\left(\frac{1}{\sqrt{(2\pi)^n |\mathbf{\Sigma}|}}\right) \tag{6}$$

$$+ \frac{1}{2}(\mathbf{x}^{(i)} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}^{(i)} - \boldsymbol{\mu}), \quad \forall i \in [1, m]. \tag{7}$$

If the residual vertical shift of the log target is expressed as an auxiliary variable z_0

$$z_0 = \ln\left(\frac{1}{\sqrt{(2\pi)^n |\mathbf{\Sigma}|}}\right), \tag{8}$$

the sample's error in the domain of argument becomes

$$e_{arg}^{(i)} = \ln z_n^{(i)} - z_0 + \frac{1}{2}(\mathbf{x}^{(i)} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}^{(i)} - \boldsymbol{\mu}). \tag{9}$$

The vector of samples' errors is then the column vector of the form $\mathbf{e}_{arg} = [e_{arg}^{(1)}, \dots, e_{arg}^{(m)}]^T$. The objective function is the weighted squared error as follows

$$D = \sum_{i=1}^{m} (w_n^{(i)} e_{arg}^{(i)})^2. \tag{10}$$

C. DETERMINATION OF THE INVERSE COVARIANCE MATRIX IN THE LOG DOMAIN

The first stage of the proposed method involves the estimation of the inverse covariance matrix and the auxiliary variable z_0 so the vector of unknown parameters $\mathbf{p}$ has the following form $\mathbf{p} = [\Sigma_{11}^{-1}, \Sigma_{12}^{-1}, \dots, \Sigma_{nn}^{-1}, z_0]^T$. The vector of unknowns $\mathbf{p}$ is calculated by minimizing the objective function (10) as

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} D(\mathbf{p}), \tag{11}$$

where

$$e_{arg}^{(i)} = \ln z_n^{(i)} - z_0 + \frac{1}{2}(\mathbf{x}^{(i)} - \hat{\boldsymbol{\mu}})^T \mathbf{\Sigma}^{-1}(\mathbf{x}^{(i)} - \hat{\boldsymbol{\mu}}). \tag{12}$$

The partial derivatives of the objective function (10) considering vector $\mathbf{p}$ given an initial centroid $\hat{\boldsymbol{\mu}}$ yield a system of linear equations with a one-step solution. The gradients of the objective function for the vector of unknowns $\mathbf{p}$ in the matrix form are

$$\nabla D(\mathbf{p}) = \mathbf{A}\mathbf{p} - \mathbf{b}, \tag{13}$$

where $\mathbf{A}$ is the matrix of coefficients and $\mathbf{b}$ is the vector of constant terms. The solution can be found as

$$\hat{\mathbf{p}} = \mathbf{A}^{-1}\mathbf{b}. \tag{14}$$

The matrix of coefficients $\mathbf{A}$ is calculated as

$$\mathbf{A} = \sum_{i=1}^{m} 2\, w_n^{(i)} \mathbf{d}^{(i)} \mathbf{d}^{\mathbf{T}(i)}, \tag{15}$$

while vector of constant terms $\mathbf{b}$ is

$$\mathbf{b} = \sum_{i=1}^{m} -2 w_n^{(i)} \ln(z_n^{(i)}) \mathbf{d}^{(i)}. \tag{16}$$

The auxiliary vector $\mathbf{d}^{(i)}$ is the vector of the form

$$\mathbf{d}^{(i)} = \Big[\quad dx_{1,1}^{(i)}/2, dx_{1,2}^{(i)}, dx_{1,3}^{(i)}, \cdots, dx_{1,n}^{(i)}, \tag{17}$$

$$dx_{2,2}^{(i)}/2, dx_{2,3}^{(i)}, \cdots, dx_{2,n}^{(i)}, \cdots, \tag{18}$$

$$dx_{n-1,n-1}^{(i)}/2, dx_{n-1,n}^{(i)}, dx_{n,n}^{(i)}/2, -1\Big]^T, \tag{19}$$

where $dx_j^{(i)} = x_j^{(i)} - \mu_{x_j}, dx_{r,s}^{(i)} = dx_r^{(i)} \cdot dx_s^{(i)}$ and the index pairs (r, s), with $1 \le r \le n$ and $r \le s \le n$, denote the unique row and column indices of the inverse covariance matrix elements, taking into account the symmetry property.

The Hessian matrix of the objective function (10) for the vector of unknowns $\mathbf{p}$ is equal to the matrix of coefficients $\mathbf{A}$. According to (15), the matrix $\mathbf{A}$ depends only on the input samples' positions and corresponding estimation weights. We use the positive estimation weights according to (3) and the matrix $\mathbf{A}$ is formed as the weighted sum of sample autocorrelation matrices which are positive semidefinite by definition. Consequently, the matrix $\mathbf{A}$, which is also the Hessian matrix, is positive semidefinite too. Hence, the solution of the linear system $\hat{\mathbf{p}}$ is the unique and optimal solution for the global minimum of the quadratic objective function (10) for the given centroid $\hat{\boldsymbol{\mu}}$ and for the chosen and fixed estimation weights, provided that $\mathbf{A}$ is nonsingular, i.e., that the Hessian is strictly positive definite. The matrix $\mathbf{A}$ can become singular and the system of linear equations yields infinitely many solutions only for the exceptional cases which are: an insufficient number of input samples or their collinearity relative to the centroid position, resulting in a non-full rank system. It occurs if the number of input samples is minimal (m_{min}) and if their absolute distances from the profile's peak position are equal concerning any one of the n axes ($\text{rank}(\mathbf{A}) < n$ if $|dx_j^{(1)}| = |dx_j^{(2)}| = \ldots = |dx_j^{(m_{min})}| = \text{const}$, $1 \le j \le n$). In that situation, the positions of all of the m_{min} input samples are on a flat hyperplane perpendicular to one of the n axes, either left (−const) or right (+const) of the peak position. Therefore, in the case of a regular grid, the geometric centre of those m_{min} input samples must not coincide with the profile's peak position since, in that case, all input samples are equidistant from the peak position for all axes. In all other cases, the matrix $\mathbf{A}$ and the Hessian matrix are positive definite, ensuring that the obtained solution for the inverse covariance matrix and the residual vertical shift of the log-target is unique and represents the global minimum of the objective function for the given centroid position and the chosen fixed estimation weights.

Finally, the optimal solution for the covariance matrix $\hat{\boldsymbol{\Sigma}}$ is found by inverting $\hat{\boldsymbol{\Sigma}}^{-1}$ whose unique elements are contained in $\hat{\mathbf{p}}$ by imposing the symmetry of the inverse. Thus, the solution $\hat{\boldsymbol{\Sigma}}$ will also be a symmetric matrix, but to be a valid covariance matrix, it also has to be positive semidefinite. Theoretically, the estimated covariance matrix $\hat{\boldsymbol{\Sigma}}$ can have some of the axes (eigenvalues) of zero length, thus indicating that the given Gaussian is actually of a lower dimension than the dimension of the input vectors, n. The condition for positive semidefiniteness of $\hat{\boldsymbol{\Sigma}}$ also requires positive definiteness of $\hat{\boldsymbol{\Sigma}}^{-1}$, because if any eigenvalue of $\hat{\boldsymbol{\Sigma}}^{-1}$ is negative, then its reciprocal value, which is the eigenvalue of $\hat{\boldsymbol{\Sigma}}$, will also be negative. Also, none of the eigenvalues of $\hat{\boldsymbol{\Sigma}}^{-1}$ must be zero since, in that case, $\hat{\boldsymbol{\Sigma}}^{-1}$ would be singular, and would not have an inverse. Since the solution for $\hat{\mathbf{p}}$ in (14) depends on the input data (sample positions, their values, and weights), the estimated matrices $\hat{\boldsymbol{\Sigma}}^{-1}$ and $\hat{\boldsymbol{\Sigma}}$ might become invalid for estimation from noisy data. Experiments have shown that such exceptions occur only for very low SNRs and for a small number of input samples k, which is comparable with the number of unknowns.

D. DETERMINATION OF IMPROVED CENTROID

The second step is the calculation of the improved centroid $\hat{\boldsymbol{\mu}}$ using the estimated inverse covariance matrix $\hat{\boldsymbol{\Sigma}}^{-1}$ which will further reduce the summary weighted squared error of the argument. The centroid was calculated by minimizing the same objective function (10) as

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} D(\boldsymbol{\mu}), \tag{20}$$

where

$$e_{arg}^{(i)} = \ln z_n^{(i)} - \hat{z_0} + \frac{1}{2}(\mathbf{x}^{(i)} - \boldsymbol{\mu})^T \hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x}^{(i)} - \boldsymbol{\mu}). \tag{21}$$

If $\mathbf{e}_{arg}$ is the column vector of errors of the exponential function's argument and $\mathbf{w}_n$ is the column vector of normalized estimation weights with m elements that correspond to the number of input samples, then the expressions for gradients and Hessian matrix of the objective function concerning the centroid position have the following matrix forms,

$$\nabla D(\boldsymbol{\mu}) = -2\boldsymbol{\Sigma}^{-1^T}\mathbf{L}^T\mathbf{W}_n\mathbf{e}_{arg}, \tag{22}$$

$$\mathbf{H}(D(\boldsymbol{\mu})) = 2(\boldsymbol{\Sigma}^{-1^T}\mathbf{L}^T\mathbf{W}_n\mathbf{L}\boldsymbol{\Sigma}^{-1} + \mathbf{w}_n^T\mathbf{e}_{arg}\boldsymbol{\Sigma}^{-1}), \tag{23}$$

where $\mathbf{W}_n = diag(\mathbf{w}_n)$, and $\mathbf{dx}_j = (dx_j^{(1)}, \ldots, dx_j^{(m)})^T$ denotes the column vector of distances of all m input samples from the given initial centroid for the jth dimension, and $\mathbf{L} = (\mathbf{dx}_1, \ldots, \mathbf{dx}_n)$ is the matrix of such column vectors for all of n dimensions.

E. UPDATE LINEAR SCALE

The previously estimated linear scale $\hat{A}$ is updated either from the solution of the system of linear equations (14) by using the estimated $\hat{\boldsymbol{\Sigma}}$ and $\hat{z}_0$ or by fitting the obtained normalized model to input data in the value domain according to (2) using the estimated $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\mu}}$.

In the first case, the previously estimated linear scale is multiplied by the residual scale calculated from (8) to obtain the updated scale of the form

$$\hat{A} = \hat{A}\exp(\hat{z_0})\sqrt{(2\pi)^n|\hat{\boldsymbol{\Sigma}}|}. \tag{24}$$

However, it is easier and more accurate to update a linear scale in the value domain according to (2), but adding the residual vertical shift of the log target z_0 to the set of unknowns is necessary to decouple the scale and shape of the Gaussian profile.

F. INITIALIZATION

Applying the proposed method requires an appropriate normalization of the input data values and the initial estimate of centroid position.

The ROI used for model estimation can be defined as a region enclosed within an arbitrarily rotated hyper-ellipsoid in n-dimensional space centered around the profile's peak position, where the chosen (or available) domain of input samples determines the ROI width. To simplify the domain selection for a multidimensional case and to describe it using a single scalar width parameter, it is convenient to bound it using the chosen maximum Mahalanobis distance of samples' positions to the centroid. Regrettably, the profile's parameters must be known (centroid and covariance) to find these distances. However, since the values of the samples are known, and these are assumed to belong to the underlying Gaussian profile, simple thresholding of sample values can be used to extract those expected to be within the desired ROI. For example, if the maximum value of input samples is $\max(\mathbf{z})$, then the chosen marginal threshold $\overline{\mathbf{z}}_{edge}$ will extract only those samples whose Mahalanobis distance to the profile center is k at most, where k equals

$$k = \sqrt{-2\log\left(\overline{\mathbf{z}}_{edge}/\max(\mathbf{z})\right)}, \tag{25}$$

Such selection is invariant concerning the unknown linear scale A since it is canceled in the quotient in (25). Moreover, since input sample values may be contaminated with noise, the binary domain matrix obtained through thresholding can be further refined using morphological smoothing to ensure it is convex and homogeneous in all dimensions.

After extracting only the samples within the chosen bounded hyper-ellipsoidal ROI, these samples approximate the truncated multidimensional Gaussian profile, which is truncated at Mahalanobis distance of k. The total probability of all samples within such a truncated region is known to be $erf(k/\sqrt{2})$, which can be used to normalize the input data (i.e., to remove the unknown scale A_{init}).

In the case of uniformly sampled ROI, the initial scale factor A_{init} can be calculated as a quotient of the total probability under the n-dimensional histogram of input data and the total probability of the truncated Gaussian profile as

$$A_{init} = \frac{\sum \mathbf{z} \cdot \delta x_1 \cdot \ldots \cdot \delta x_n}{erf\left(\frac{k}{\sqrt{2}}\right)}, \tag{26}$$

where δx_j denotes the distance between neighboring samples for the jth axis in n-dimensional space in the case of regularly and uniformly sampled ROI.

Therefore, the input data values $\mathbf{z}$ are normalized according to

$$\mathbf{z}_n = \frac{\mathbf{z} \cdot \delta x_1 \cdot \ldots \cdot \delta x_n}{A_{init}} = \frac{\mathbf{z} \cdot erf\left(\frac{k}{\sqrt{2}}\right)}{\sum \mathbf{z}}, \tag{27}$$

where $\mathbf{z}_n$ represents the normalized n-dimensional histogram, approximating the regularly sampled Gaussian profile truncated at maximal Mahalanobis distance k.

The initial centroid can be determined by any method known in the literature, but the simplest case of centroid determination using the method of moments in the domain of values is presented below. The first moment yields the expected centroid, i.e., the profile's peak position, along the x_j-axis as

$$\mu_{x_j} = \frac{\sum_{i=1}^{m} x_j^{(i)} z_n^{(i)}}{\sum_{i=1}^{m} z_n^{(i)}}, \tag{28}$$

where the denominator represents the total probability within the enclosed input region of truncated Gaussian, m is the number of input samples within the ROI, and j is the notation of the axis in n-dimensional space ($j \in [1, n]$).

In the case of using the normalized moment-based model of the form $\mathbf{f_1} = f(\mathbf{X}, \hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\Sigma}}_1, \hat{A})/\hat{A} = f(\mathbf{X}, \hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\Sigma}}_1, 1)$ as estimation weights, the corresponding covariance matrix has to be calculated as well. The second moment yields the variances that form the covariance matrix as follows:

$$\Sigma_{rr} = \sigma_r^2 = \sum_{i=1}^{m} z_n^{(i)} (x_r^{(i)} - \mu_{x_r})^2, \tag{29}$$

$$\Sigma_{rs} = \sigma_r \sigma_s \rho_{rs} = \sum_{i=1}^{m} z_n^{(i)} (x_r^{(i)} - \mu_{x_r})(x_s^{(i)} - \mu_{x_s}), \tag{30}$$

where σ_r and σ_s denotes the profile's STDs for the r and s axes and $\rho_{rs} \in [-1, 1]$ is the Pearson product-moment correlation coefficient of r and s. Since the Gaussian profile is truncated at the maximal Mahalanobis distance k, the calculated profile's STDs have to be additionally scaled using the following expression

$$\sigma_{trunc} = \sigma / \sqrt{1 - \frac{k \exp(-k^2/2)}{\sqrt{\frac{\pi}{2}} erf\left(\frac{k}{\sqrt{2}}\right)}} \tag{31}$$

Additionally, the optimal scale factor can be calculated by fitting the obtained model to empirical data as the ordinary LS estimate in a simple linear regression model as

$$\hat{A}_1 = \frac{\mathbf{f}_1^T \mathbf{z}}{\mathbf{f}_1^T \mathbf{f}_1}, \tag{32}$$

where the numerator in (32) represents the sample covariance between the initial model and input data values, while the denominator represents the sample variance of the estimated model.

V. EXPERIMENTS AND RESULTS

This section consists of two main parts. In the first part, the parameters of the 3D Gaussian profile were estimated analytically for a given actual centroid position, comparing the results with other methods. In the second part, the 2D Gaussian profile was fitted to the noisy data by using the proposed iterative method, and an extensive analysis of the iterative method's accuracy, complexity, and convergence was made.

A. 3D GAUSSIAN PROFILE FITTING FOR A GIVEN CENTROID

To demonstrate the accuracy and the speed of the proposed method, we estimated the parameters of the 3D Gaussian profile from the synthesized noiseless data and data contaminated with additive Gaussian noise by using the proposed method and compared the results with the results of the commonly used method of moments [13], LS method [18], [26], [27] and ML method [23]. Two examples of 3D Gaussian fitting to the noisy data using the proposed method are shown in Fig. 2.

To emphasize the method's speed and to easily compare it with the mentioned methods, we fixed the centroid position and measured the accuracy and the execution time for estimation of all other profile parameters: the covariance matrix and the linear scale. The experiments were executed on a system with Intel(R) Core(TM) i5-7200U CPU @ 2.5 GHz and 8 GB of RAM, with MATLAB implementation of all methods. The solution of the method of moments was provided as an initial guess for the LS and ML methods which minimize the objective function using the iterative optimization procedure. The optimization was performed using the native MATLAB *fminunc* solver without specified analytical gradients. The maximum number of iterations was set to 50, while the optimality tolerance and the current point tolerance were set to 10^{-7}.

As a measure of accuracy, the total modeling error was calculated according to the following formula,

$$e_{total} = 10\log_{10}\left(\frac{\sum_{\forall i\in eval}\left(f(\mathbf{x}^{(i)},\hat{\boldsymbol{\beta}}) - f(\mathbf{x}^{(i)},\boldsymbol{\beta})\right)^2}{\sum_{\forall i\in eval} f(\mathbf{x}^{(i)},\boldsymbol{\beta})^2}\right), \tag{33}$$

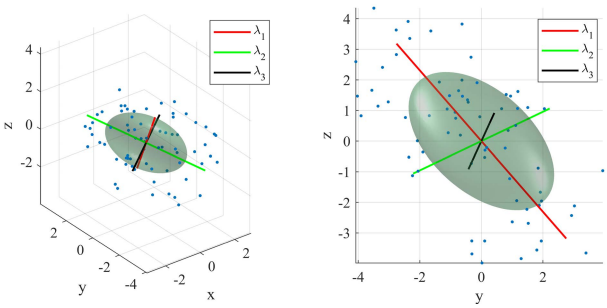
where $\mathbf{x}^{(i)}$ represents the position of the ith sample on the arbitrary evaluation grid $eval$, $f(\mathbf{x}^{(i)},\hat{\boldsymbol{\beta}})$ represents the estimated model value and $f(\mathbf{x}^{(i)},\boldsymbol{\beta})$ represents the given actual model value for the same sample position. The total modeling error represents the sum of the squared residuals between the estimated and given model over the same evaluation grid and therefore, indirectly aggregates the accuracy of all model parameters.

For experimental purposes, to obtain the synthetic noisy measurements, the profile with the given parameters was firstly synthesized and then the additive noise with the appropriate STD σ_n was added to the profile to ensure the desired SNR within the ROI. In this case, the SNR was defined as PSNR according to the following formula

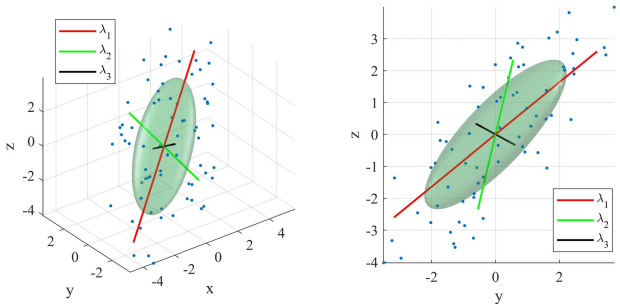
$$\mathrm{SNR_{dB}} = 20\log_{10}\frac{A/\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}}{\sigma_n}. \tag{34}$$

Fig. 2 illustrates the estimation of the given 3D Gaussian profile using the proposed method. Its actual parameters in the uncorrelated form are

$$[A, \lambda_1, \lambda_2, \lambda_3, \theta_1, \theta_2, \theta_3] = [1000, 3, 2, 1, \pi/3, \pi/4, \pi/6],$$



(a) $A = 1000, \lambda_1 = 3, \lambda_2 = 2, \lambda_3 = 1, \theta_1 = \pi/3, \theta_2 = \pi/4, \theta_3 = \pi/6, \hat{\lambda_1} = 3.01, \hat{\lambda_2} = 2.03, \hat{\lambda_3} = 1.02, \hat{\theta_1} = \pi/3.00, \hat{\theta_2} = \pi/4.09, \hat{\theta_3} = \pi/6.21, \hat{A} = 1019.02$



(b) $A = 1000, \lambda_1 = 4, \lambda_2 = 2, \lambda_3 = 0.5, \theta_1 = \pi/5, \theta_2 = -\pi/7, \theta_3 = \pi/3, \hat{\lambda_1} = 4.13, \hat{\lambda_2} = 2.00, \hat{\lambda_3} = 0.50, \hat{\theta_1} = \pi/5.03, \hat{\theta_2} = -\pi/7.09, \hat{\theta_3} = \pi/3.00, \hat{A} = 1020.49$

FIGURE 2. Examples of 3D Gaussian fitting results in 3D and 2D views using the proposed method for a given centroid. Estimation was done from $m = 70$ input samples with noise level of $\mathrm{SNR_{dB}} = 30$. Estimated profile parameters are: the linear scale $\hat{A}$, the semiaxes widths $(\hat{\lambda_1}, \hat{\lambda_2}, \hat{\lambda_3})$, and the rotation angles $(\hat{\theta_1}, \hat{\theta_2}, \hat{\theta_3})$. The fitted 3-D surfaces represent the points where the estimated model is equal to the given model value at $k = 1$, which is $f = A\exp(-1/2)$. The blue dots represent only the input samples' positions but not their values (due to the limitation of the 3D example visualization).

where λ_1, λ_2 and λ_3 denote the semiaxes widths, θ_1, θ_2, θ_3 denote the Euler rotation angles and A is the linear scale factor. The centroid position was fixed to $\boldsymbol{\mu} = [0, 0, 0]$ and passed as input to all estimation methods. The actual profile's parameters yield the following rotation matrix $\mathbf{R}$

$$\mathbf{R} = \mathbf{R}_z(\theta_1)\mathbf{R}_y(\theta_2)\mathbf{R}_x(\theta_3) = \begin{bmatrix} 0.3536 & -0.5732 & 0.7392 \\ 0.6124 & 0.7392 & 0.2803 \\ -0.7071 & 0.3536 & 0.6124 \end{bmatrix},$$

and the corresponding covariance matrix $\boldsymbol{\Sigma}$

$$\boldsymbol{\Sigma} = \mathbf{R}\begin{bmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{bmatrix}\mathbf{R}' = \begin{bmatrix} 2.9858 & 0.4609 & -2.6080 \\ 0.4609 & 5.6392 & -2.6801 \\ -2.6080 & -2.6801 & 5.3750 \end{bmatrix}.$$

Finally, the vector of the actual profile's parameters can be expressed as

$$\boldsymbol{\beta} = [A, \Sigma_{11}, \Sigma_{12}, \Sigma_{13}, \Sigma_{22}, \Sigma_{23}, \Sigma_{33}] = [1000, 2.9858, 0.4609, -2.6080, 5.6392, -2.6801, 5.3750].$$

The proposed method estimated the unique terms of the covariance matrix by solving the system of linear equations (14) and inverting the obtained solution, while the linear

scale factor was calculated from the estimated additional unknown z_0 as $\hat{A} = \exp(\hat{z_0})\sqrt{(2\pi)^3|\hat{\mathbf{\Sigma}}|}$, thus avoiding the need for initialization. In the noiseless case, the unit estimation weights were used, while for the estimation from the noisy data, the estimation weights were fixed to input sample values.

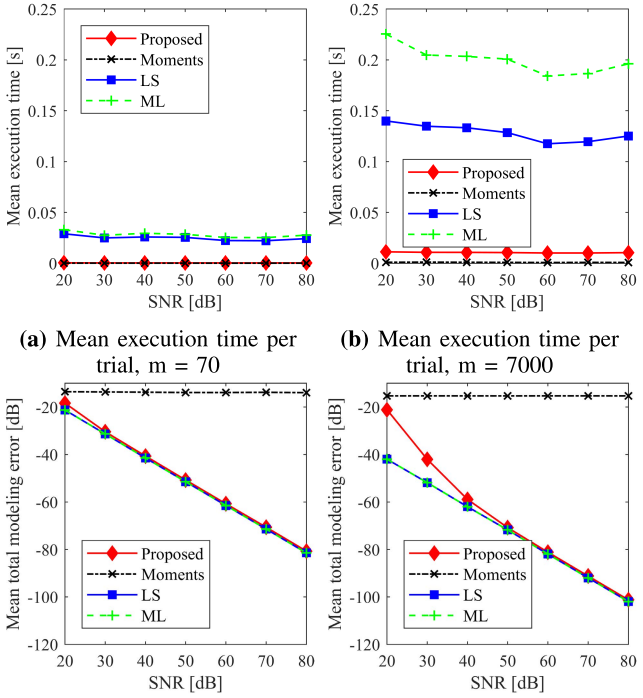
To compare the estimation results, the parameters of the given profile were estimated for different numbers of input samples ($m = \{7, 70, 7000\}$) and different SNRs ($\text{SNR}_{\text{dB}} = \{20, 30, 40, 50, 60, 70, 80\}$). The Monte Carlo simulation with 1000 trials was performed for each combination of those parameters, including the noiseless case. In each trial, the input samples were randomly picked within convex ROI determined with maximal Mahalanobis distance $k = 2$ from the actual centroid position. The evaluation region was selected as uniformly sampled ROI enclosed within Mahalanobis distance $k = 3$ with the step size of 0.1 in each direction and is defined by the given model's parameters. For the same 3D example, for $\text{SNR}_{\text{dB}} = 40$, the estimated parameters from 70 random input samples using the proposed method in one trial were

$$\hat{\boldsymbol{\beta}} = [\hat{A}, \hat{\Sigma_{11}}, \hat{\Sigma_{12}}, \hat{\Sigma_{13}}, \hat{\Sigma_{22}}, \hat{\Sigma_{23}}, \hat{\Sigma_{33}}] = [1004.5, 3.0018, 0.4664, -2.6003, 5.6972, -2.6807, 5.3330]$$

with the total modeling error $e_{total} = -42.8741$dB.

The solution of the method of moments was provided as an initial guess for LS and ML methods. Since the objective of the ML method assumes the presence of stochastic additive Gaussian noise, in the noiseless case, $\sigma_n = 10^{-4}$ was passed as an input to the optimization procedure to mimic the almost ideal measurements ($\text{SNR}_{\text{dB}} = 100$). In the case of noisy measurements, the considered numbers of input samples were only 70 and 7000 since, in that case, the minimum number of input samples which corresponds to the number of unknowns was insufficient for a valid solution (for the 3D Gaussian with the given centroid position, the number of unknowns is 7). The mean execution times and total modeling errors for different numbers of input samples for the noiseless case and for $\text{SNR}_{\text{dB}} = 40$ are given in Table 1. The trends of mean total modeling errors and execution times for different SNRs and numbers of input samples are shown in Fig. 3.

In the noiseless case, the proposed method finds the optimal (ideal) solution of all seven unknowns from only seven randomly positioned input samples with average modeling error of −255dB for 1000 trials ($e_{total} = -255.58$dB, $m = 7$)). For such a small number of input samples, the LS and ML estimates are inaccurate due to the bad initial guess obtained by the method of moments. The LS and ML methods also converge to the optimal solution in the noiseless case for a larger number of input samples ($m = 70$ and $m = 7000$), but their speed and accuracy highly depend on the initial guess. The method of moments has poor accuracy ($e_{total} = -15.31$dB, $m = 7000$) since it underestimates all parameters. In the estimation from noisy measurements, the LS and the ML methods behave similarly considering



(a) Mean execution time per trial, m = 70

(b) Mean execution time per trial, m = 7000

(c) Mean total modeling error, m = 70

(d) Mean total modeling error, m = 7000

FIGURE 3. Comparison of the speed and the accuracy of the following methods: proposed method, method of moments, LS method, and ML method.

the accuracy since they yield the same optimal solution for the case of additive Gaussian noise contamination, as can be seen in Fig. 3c and Fig. 3d. The 100-fold increase in the number of input samples from 70 to 7000 caused the 20dB increase in the accuracy of the LS and ML methods due to averaging, as expected. The proposed method follows the trends of LS and ML methods concerning accuracy, except for very low SNRs (20dB) where the approximation of the exponential function with the first term of the Taylor series for estimation weights determination in the argument domain is insufficient. However, the proposed method requires at least 10 times less time for calculation of the covariance matrix and linear amplitude scale than LS and ML methods yielding comparable estimates without prior knowledge. Fig.3a and Fig.3b show that the mean execution time of the proposed method is comparable with the mean execution time of the method of moments since both methods do not use the iterative procedure and find solution analytically. Both methods are almost up to 100 times faster than the LS method for the smaller number of input samples (70) and up to 10 times for the larger number of input samples according to Table 1. The LS and ML methods have a similar mean execution time for the small number of input samples, but with the increasing number of samples, the ML method becomes slower than the LS method, as shown in Fig. 3b.

To sum up, the proposed method is much faster than the commonly used LS and ML methods and has a mean

TABLE 1. The mean total modeling errors in dB and the mean execution times in seconds for the estimation of the 3D Gaussian profile given the actual centroid position.

	Mean total modeling error [dB]					Mean execution time per trial [s]				
	Ideal noiseless case			$SNR_{dB} = 40$		Ideal noiseless case			$SNR_{dB} = 40$	
Number of input samples	7	70	7000	70	7000	7	70	7000	70	7000
Proposed method	**-255.5804**	**-293.3372**	**-293.5277**	**-40.6434**	**-58.8981**	**0.0004**	**0.0004**	**0.0101**	**0.0004**	**0.0105**
Method of moments	-7.1845	-13.9014	-15.3146	-13.7786	-15.3044	0.0001	0.0001	0.0007	0.0001	0.0009
LS method	-67.0825	-134.5511	-135.3360	-41.4402	-61.9902	0.0475	0.0349	0.1239	0.0255	0.1286
ML method	-89.9811	-134.5511	-135.3360	-41.4402	-61.9902	0.0638	0.0391	0.1922	0.0285	0.2008

execution time comparable with the fastest method of moments. Thereby, it yields the estimate of almost the same accuracy as the LS and ML methods except for very low SNRs.

B. ITERATIVE METHOD ACCURACY

In the second experiment, the parameters of the rotationally symmetric 2D Gaussian profile were estimated from noisy data using the proposed method to verify its accuracy. In addition, the results are compared with those previously published in [11], where the same 2D Gaussian profile was estimated in the value domain using the numerical LS method. Although we fitted the 2D Gaussian profile, the proposed method can be applied to estimate the Gaussian profile of arbitrary dimensions and arbitrary shapes, as was demonstrated in the 3D experiment.

In [11], it was shown that the accuracy of the LS method in the estimation of the 2D Gaussian parameters from data corrupted with additive noise varies for different ROI widths as the consequence of the difference in the informativeness of the input data, i.e., the difference in the differential entropy of the input data. The maximal differential entropy is obtained when the Gaussian profile is truncated at Mahalanobis distance $k = 2$. For other non-optimal ROI widths, the reduction of differential entropy compared to the optimal case for $k = 2$ was compensated by the proportional reduction of the given noise level. The difference in differential entropy of the input data precisely predicts a decrease in the LS method accuracy for different ROI widths in the value domain. We performed the same experiments in this paper and re-applied the data entropy compensation.

In addition to differential entropy, we also analyzed the influence of negative samples on the method's accuracy since it is necessary to handle negative input sample values before the log transformation. Therefore, the probability of negative samples when a Gaussian profile is contaminated with additive noise was derived and given in Appendix B. The total number of input samples was increased in proportion to the predicted percentage of negative samples for the given ROI width and SNR to eliminate the loss of method accuracy due to the occurrence of negative samples that alone cannot participate in the estimation in the log domain. Negative samples were either set to $eps = 2^{-52}$ and logarithmized or simply removed from the input data and the estimation process. However, the compensation of negative samples did not significantly improve the accuracy of the proposed method for the ROI widths considered in this paper.

Furthermore, two types of estimation weights were considered and compared. In the first case, the weights in the objective function (10) were the values of the input samples, while, in the second case, the weights were selected as the initial moment-based model values for the same sample positions.

Two experiments were made considering circular ROI sampling: random sampling with a fixed number of input samples or uniform sampling with a fixed density of input samples. In each experiment, the total number of performed test cases was 16, depending on the input parameters given in Table 2. In each test case, the Monte Carlo simulation with 50 trials was performed for all Mahalanobis distances k from the set $k = \{0.5, 1, 1.5, 2, 2.5, 3\}$ and for all given SNRs from the set $SNR_{dB} = \{20, 40, 60, 80\}$. As an evaluation grid, we used a uniformly sampled circular region of fixed width $r = 3\sigma$, where $\sigma = \sigma_1 = \sigma_2$ is the profile's STD with equal x and y spacing of $\sigma/10$.

In both experiments, the same six parameters of the specified 2D Gaussian profile were estimated from noisy measurements: profile's STDs σ_1, σ_2, the correlation coefficient ρ_{12}, the profile's peak position $\boldsymbol{\mu} = [\mu_{x_1}, \mu_{x_2}]$, and the linear profile's scale A. The given profile parameters in the correlated form were $[A, \sigma_1, \sigma_2, \rho_{12}, \mu_{x_1}, \mu_{x_2}] = [100, 1, 1, 0, 0, 0]$, i.e., $\boldsymbol{\beta} = [A, \Sigma_{11}, \Sigma_{12}, \Sigma_{22}, \mu_{x_1}, \mu_{x_2}] = [100, 1, 0, 1, 0, 0]$. In the case of compensation of differential entropy reduction compared to the nominal case for $k = 2$, the noise STD calculated from (34) was reduced according to the formula [11]

$$\sigma_{n_{reduced}} = \sigma_n \cdot 10^{(\Delta h_{z_{dB}}/20)}, \tag{35}$$

where $\Delta h_{z_{dB}}$ denotes the reduction of differential entropy in dB for the chosen factor k.

TABLE 2. Input parameters combined in experiments.

Weights $\mathbf{w_n}$	$\mathbf{z_n}$	$\mathbf{f_1}$
Number of input samples m (only for random sampling)	100	10000
Compensation of predicted number of negative samples	yes	no
Negative input samples	set to eps	removed
Compensation of differential entropy reduction compared to the case for $k = 2$	yes	no

1) THE FIXED NUMBER OF INPUT SAMPLES

In the first experiment, the estimation input samples were randomly picked within circular ROIs, and their number was fixed for all ROI widths. The number of input samples was chosen as $m_1 = 10000$ or $m_2 = 100$. Thus, the total number of test cases in this experiment was 32, considering all combinations of input parameters in Table 2. The comparison of the mean total modeling errors averaged over 50 trials of the LS method in the domain of values and the proposed method in the argument domain are shown in Fig. 4. As shown in [11], by compensating the differential entropy reduction, compared to the nominal optimal case for $k = 2$, the efficiency of the LS method in the domain of values became invariable of ROI widths for each given SNR (flat dashed lines). Additionally, as expected, the 100-fold increase in the number of random input samples increases the accuracy of the LS method by 20 dB ($10\log_{10}(m_1/m_2)$) due to averaging, which effectively reduces the noise level and, therefore, improves model accuracy.

The accuracy of the proposed argument domain method depends on the selected combination of input parameters used in each test case. The results of a few representative test cases are shown in Fig. 4 where differential entropy reduction was compensated, and the compensation of the predicted number of negative samples was not applied.

As shown in Fig. 4, the accuracy of the proposed method is smaller than the LS method in the value domain only for very low SNR ratios and extensive ROI widths. For sufficiently high SNR ratios and narrower ROIs, the efficiency of the proposed method is practically identical to the efficiency of the LS method in the value domain. In Figs. 4a and 4b, the input sample values are used as estimation weights in the objective function (10), while in Figs. 4c- 4f the initial model values are used as estimation weights. When input sample values are utilized as estimation weights, both methods' mean total modeling errors are almost the same whether negative samples are removed before estimation or set to *eps*. However, for the case of using initial model values as weights, the results of the proposed method are closer to the results of the LS method in the value domain, but only if negative samples are removed from the estimation process (Figs. 4c and 4d). The largest deviation in the accuracy of the proposed method compared to the LS method in the value domain is observed in the case of replacing negative samples with *eps* when the initial model values were used as weights, especially for low SNR and wide ROI (SNR = 20dB, $k = 3$) as Figs. 4e and 4f show. Additionally, in Figs. 4a, 4c and 4e is shown that a 100-fold increase in the number of input samples does not increase the estimation accuracy by 20 dB on wide ROIs and low SNRs. The causes of such behavior are inappropriate handling of negative samples and inappropriate weight selection since the assumption that the exponential function can be approximated with the first term of the Taylor series holds only for small errors.

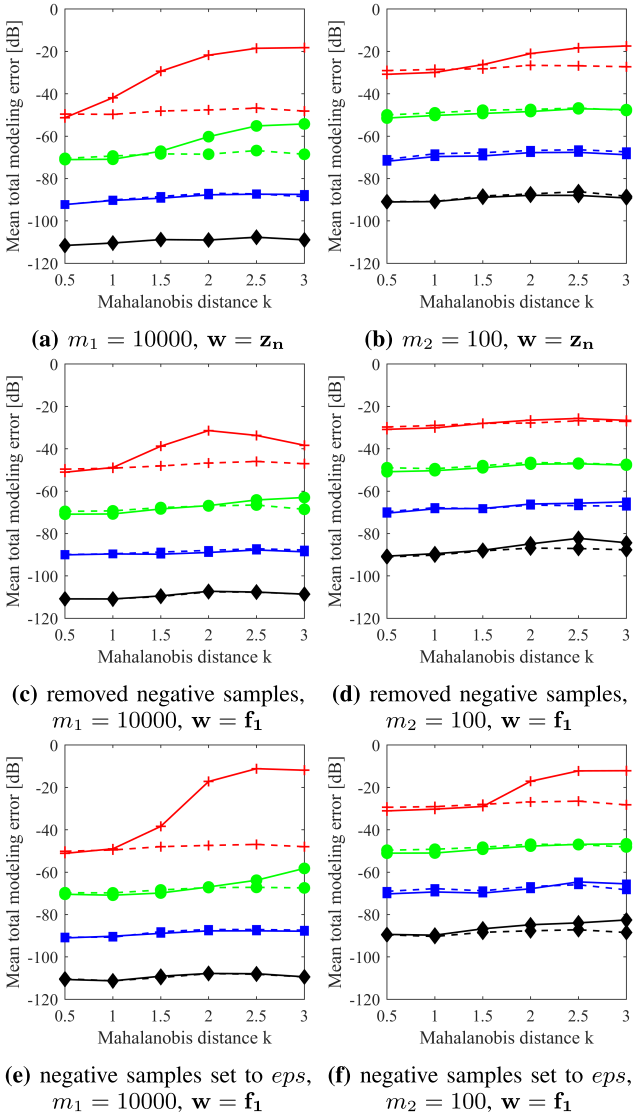
In Fig. 4 can be seen that the relative degradation for low SNRs (20 dB) and wide ROIs is not the same for $m_1 = 10000$ as for $m_2 = 100$ but increases more for the larger m and less for the small m. However, it is simply a consequence of direct bias in the input data which are falsified, so no averaging can help. The averaging can only attenuate a stochastic error and not a consistent one, so the error is almost the same in absolute terms for both cases ($m_1 = 10000$ and $m_2 = 100$ when SNR = 20 dB and ROI = 3). It can be concluded that increasing the number of input samples reduces the model error only for very narrow ROIs when the percentage of such negative samples is tiny, so in fact, the contribution of stochastic noise is effectively attenuated. With very low SNRs, the initial moment-based model might also be inaccurate. Using such a model for weights and even throwing out negative samples results in a model denoted with the red line that is not parallel to green, blue, and black but instead increases incrementally (SNR = 20 dB).

2) THE FIXED DENSITY OF INPUT SAMPLES

The circular ROI was uniformly sampled in the second experiment, so the number of input samples was variable for different ROI widths depending on the ratio between the profile's STD and the pixel size. Since the STD of the given 2D Gaussian profile was $\sigma_1 = \sigma_2 = 1$, the selected pixel width was chosen as $\delta x_1 = \delta x_2 = 0.25$ to ensure the sufficient number of input samples for estimation of all given profile's parameters even for the narrowest ROI ($k = 0.5$), as it was described in [11]. From the ratio of the ROI size and individual pixel size, it can be concluded that the number of input samples increases quadratically with the factor of Mahalanobis distance k ($m = P_{\mathrm{ROI}}/P_{\mathrm{pix}} = k^2\sigma_1\sigma_2\pi/(\delta x_1\delta x_2)$ so $m_1/m_2 = k_1^2/k_2^2$) so it is expected that the total modeling error decreases with the increase of ROI width as $10\log_{10}(m_1/m_2) = 10\log_{10}(k_1^2/k_2^2)$. Since the minimal ROI width considered in this experiment is for $k = 0.5$ while the maximal for $k = 3$, the number of input samples increases 36-fold, and consequently, the expected reduction of the total modeling error as a result of ROI size increase should be 15.56 dB. The mean total modeling errors are shown in Fig. 5. Again, only some of the 16 test cases are shown considering all input parameters in Table 2, while the number of input samples is directly determined by k. The results show that the proposed method's accuracy follows the LS method's accuracy trend in the value domain when using initial model values as weights and by removing negative samples. The results are worse for the case when input sample values are used as weights for low SNRs (20 dB) and wide ROIs. However, both methods achieve the same accuracy in cases of high SNRs.

C. DISCUSSION

This subsection analyses the influence of each input parameter given in Table 2 on the estimation accuracy of the proposed method. In the case of using input sample values as

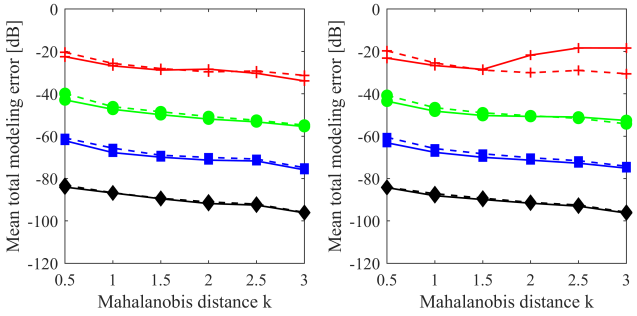


(a) $m_1 = 10000$, $\mathbf{w} = \mathbf{z_n}$ **(b)** $m_2 = 100$, $\mathbf{w} = \mathbf{z_n}$

(c) removed negative samples, $m_1 = 10000$, $\mathbf{w} = \mathbf{f_1}$ **(d)** removed negative samples, $m_2 = 100$, $\mathbf{w} = \mathbf{f_1}$

(e) negative samples set to *eps*, $m_1 = 10000$, $\mathbf{w} = \mathbf{f_1}$ **(f)** negative samples set to *eps*, $m_2 = 100$, $\mathbf{w} = \mathbf{f_1}$

FIGURE 4. Comparison of the mean total modeling errors in dB of the LS method in the value domain (dashed lines) and the proposed method in the argument domain (solid lines) in the case of random sampling with fixed number of input samples for different ROI widths determined by the factor *k* and different SNRs: 20 dB (red), 40 dB (green), 60 dB (blue), and 80 dB (black). $\mathbf{f_1}$ denotes a vector of initial model values, while $\mathbf{z_n}$ denotes input sample values.



(a) negative samples are removed, $\mathbf{w_n} = \mathbf{f_1}$, where $\mathbf{f_1}$ is a vector of initial model values

(b) negative samples are removed, $\mathbf{w_n} = \mathbf{z_n}$, where $\mathbf{z_n}$ is a vector of input sample values

FIGURE 5. Comparison of the mean total modeling errors in dB of the LS method in the value domain (dashed lines) and the proposed method in the argument domain (solid lines) in the case of uniform sampling and a variable number of input samples for different ROI widths determined by the factor *k* and different SNRs: 20 dB (red), 40 dB (green), 60 dB (blue), and 80 dB (black).

estimation weights, the weighted squared residual error is the same whether the negative samples are removed or set to *eps*. Since the squared residual error between the log of *eps* and the log of model value is further weighted with eps^2, it does not contribute to the total error sum, thus achieving the same effect as the negative sample removal. However, when the model values are used as estimation weights in the log domain, setting negative samples to *eps* or removing them yields different estimation results. Namely, the log transform of such a small value (*eps*) yields a significantly negative value. The initial moment-based model is sensitive to noise and imprecise at wide ROIs where negative samples typically occur and yields falsely and potentially more enormous model values at those positions than *eps*. When such a model value is used as a weight at a place of negative input sample set to *eps*, it can significantly degrade the estimation since such weight is false and arbitrary. This can be observed only for the lowest SNR ratio of 20dB when such negative samples predominantly occur and, of course, only for the widest ROIs. For narrower ROIs, the probability of negative samples is lower, even for such low SNRs, so it is much better to remove negative samples from the estimation process and avoid the problems of weight selection. Therefore, using the initial model values as weights yields better results only if negative samples are removed, and SNR is sufficiently high.

In addition, the experimental results show that the contribution of differential entropy compensation is not ideal. For example, in the case of a fixed number of input samples, an alignment of the error curves similar to the estimation in the value domain is achieved, but these curves are still not wholly flat across all ROI widths.

Noise level compensation performed by predicting the expected number of negative samples for a given width and SNR also does not significantly improve the estimation accuracy in the log domain since these negative samples occur mainly for very low SNRs and wide ROIs. In these cases, the problem is not a lack of valid positive samples but rather a poor approximation of quadratic error in the value domain by the weighted quadratic error in the argument domain by using the weights that describe only the first term of Taylor's series of that approximation. Such simplified weights are insufficient for considerable errors when measurements significantly deviate from the given model.

D. ITERATIVE METHOD COMPLEXITY

The complexity of the proposed two-step method in the argument domain is comparable to the complexity of only a single iteration of the LS method in the domain of values that simultaneously estimates all profile parameters using

Newton's optimization technique that requires the calculation of analytical derivatives and Hessian matrix concerning all profile parameters.

The complexity of both methods is parameterized with the dimension of the Gaussian profile n and the number of estimation input samples m. The number of input samples m must be at least equal to the number of unknowns, which is of order n^2. When the number of input samples is of order n^2, the complexity of both methods in a big O notation is $O(n^6)$. For the number of input samples of order n^3 or higher, the complexity of both methods in a big O notation equals $O(mn^4)$.

Although the complexity of a single iteration is comparable, we have experimentally determined that the proposed method converges in at most four such iterations since most of the optimal model parameters are found analytically, which explains the significant advantage of our method.

E. CONVERGENCE OF THE ITERATIVE METHOD

This subsection explains the iterative method convergence where the centroid and the covariance matrix are updated alternately. If the given initial centroid is close enough to the actual solution and if the estimation weights are fixed, the proposed method's first step finds a unique solution that minimizes the objective function by solving a system of linear equations. The proof of this statement with the description of exceptions is given in subsection IV-C. In the second step of the method, we search for a better centroid that gives a smaller weighted squared error than the centroid from the previous iteration for the calculated covariance matrix by solving the system of coupled nonlinear equations. Due to the minimization criterion itself and the fixed estimation weights, an iterative optimization procedure such as Newton's method returns, in the worst case, the current centroid, and the iterative procedure terminates. In all other situations, the new centroid will yield a better fit (smaller value of the objective function), and a new estimation of the covariance matrix and scale can be performed for the new centroid position that will further improve the model fit. The method can get stuck in a local minimum if the initial centroid is very far from the real solution or if SNR is low but the method's first step with the analytical solution and fixed estimation weights ensure convergence, at least according to the local minimum.

VI. FUTURE WORK

In the case of low SNRs, the initial model can significantly deviate from the specified model, and using the initial model weights yields even worse parameter estimates than using the input data-driven weights. This issue can be approached by iterating the estimation procedure where the weights in each iteration are selected as estimated model values from the previous step. Under the assumption of getting a better model in each iteration and consequently more appropriate weights, the updated estimates in each iteration will be closer to the given model. Such a procedure with adaptive weights should converge for high SNRs. As noise gradually grows and SNR reduces, the estimation should still converge and get a usable solution due to robustness; if the noise does not become so high, the estimation diverges and becomes unfeasible. However, the guarantee of convergence is much more difficult to prove in the case of an iterative procedure with adaptive weights since in each iteration, in addition to the model, the optimization criterion also changes. However, this approach is undoubtedly worth future work.

VII. CONCLUSION

The experimental results show that in the case of high SNRs, both the WLS method in the log domain and the LS method in the domain of Gaussian profile values achieve the same accuracy. Still, the proposed method in the log domain converges faster, especially for the Gaussian profiles of higher dimensions.

The main advantage of the proposed method is a one-step solution for the covariance matrix and linear scale for a given centroid position, thus avoiding the nonlinear optimization for the estimation of those parameters. For many practical applications where the initial centroid is already sufficiently accurately determined, only one single iteration of the proposed algorithm is sufficient to determine all remaining model parameters using the analytical procedure without the need for further iteration. Experiments showed that in the case of the 3D Gaussian profile estimation from data contaminated with additive Gaussian noise, the covariance matrix and the linear scale were estimated one or even two orders of magnitude faster using the analytical solution of the proposed method than using the iterative LS method in the domain of values. At the same time, for high SNRs ($\geq$40dB) the proposed method achieved almost the same total modeling error (with suboptimality of less than 3dB) without prior knowledge except for the assumed position of the peek's center in comparison to the iterative LS method in the domain of values, which yields an optimal solution for the case of Gaussian noise contamination. Moreover, experiments showed that such a model could be further improved by the described correction of the initial centroid through a maximum of four iterations of this two-step method. For comparison, numerical optimization procedures in the value domain require a significantly larger number of iterations than the proposed method to find all model parameters, which explains the significant advantage of our method. Due to its rapid convergence, the method can be used in real-time applications.

The LS method in the domain of values is more accurate than the proposed WLS method in the case of low SNRs ($\leq$20 dB) and wide ROIs. The loss of precision of the proposed method is dominantly caused by small profile values on such wide ROIs that can introduce significant errors in the log domain when the approximation of the exponential function with the first term of the Taylor series does not hold. The proposed method achieves higher accuracy in the case of model-driven estimation weights, which are selected as values of the initial, moment-based model. Such selected

estimation weights are fixed for all iterations of the proposed method, thus guaranteeing the method's convergence.

APPENDIX A ESTIMATION WEIGHTS

In this section, we derive the estimation weights that relate the error of the Gaussian profile argument with the error of the Gaussian profile value. Let us start with the simplified relation of the 1D Gaussian profile value y and the corresponding argument x of the exponential function

$$y = K\exp(x), \tag{36}$$

where K equals the product of the linear scale A and the normalization term. Let us assume that the argument x equals the sum of the ideal argument value x_0 and added argument error dx. The question is how this argument error is transmitted to the domain of values.

The argument error adds the perturbation dy to the ideal profile value y_0, which equals $y_0 = K\exp(x_0)$, so the total sample value can be expressed as

$$y = y_0 + dy = K\exp(x_0 + dx) = y_0\exp(dx). \tag{37}$$

The perturbation in the domain of values dy from the above expression then equals

$$dy = y_0(\exp(dx) - 1). \tag{38}$$

For minor errors dx, the exponential function can be approximated by the first term of Taylor's series as

$$\exp(dx) \simeq 1 + dx, \tag{39}$$

and the final expression for the relation of the exponential function's argument error dx and the error of the exponential function's value dy equals

$$dy = y_0\, dx. \tag{40}$$

This proves that the optimal sample weight equals the ideal sample value y_0.

APPENDIX B PROBABILITY OF NEGATIVE SAMPLES

Estimating the Gaussian profile parameters in the log domain assumes that the input estimation samples are positive. This is because the parameters are estimated from the profile's values, usually contaminated with noise. In the case of additive Gaussian noise of the zero mean and the variance of σ_n^2, the noise samples can have positive and negative values, and consequently, the sum of signal and noise values can be negative thus preventing the estimation of the Gaussian parameters in the log domain. This is the motivation for calculating the probability that the sum of these random variables, Gaussian profile value, and additive noise is less than zero. The total input signal can be written as

$$Z = Y + X, \tag{41}$$

where Y is Gaussian profile value as the function of uniform random variables and X is additive noise. Y is the random variable that follows a log-normal distribution, as shown in [11], while X is the random noise of normal distribution. We are interested in the probability that the sum of these variables is less than zero, i.e., $P(Z < 0)$. For Z to be less than z, X must be less than $z - Y$. The cumulative distribution function (CDF) of Z can be calculated as

$$F_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_{XY}(x, y)dxdy. \tag{42}$$

If X and Y are independent, the CDF of Z can be calculated as

$$F_Z(z) = \int_{-\infty}^{\infty} f_Y(y)\Big(\int_{-\infty}^{z-y} f_x(x)dx\Big)dy. \tag{43}$$

In our case, $z = 0$ and the PDF of variables X and Y are

$$f_Y(y) = \frac{2}{k^2 y}, A\exp(-k^2/2) \le y \le A, \tag{44}$$

$$f_X(x) = \frac{1}{\sigma_n\sqrt{2\pi}}\exp\frac{-1}{2}\frac{x^2}{\sigma_n^2}, \mu = 0 \text{ (zero-mean noise)}, \tag{45}$$

while the CDF of normal noise and of total signal Z are

$$F_X(x) = \frac{1}{2}\left(1 + erf\left(\frac{x}{\sigma_n\sqrt{2}}\right)\right), \tag{46}$$

$$F_Z(z) = \int_{A\exp(-k^2/2)}^{A} f_y(y)\Big(\int_{-\infty}^{-y} f_X(x)dx\Big)dy \tag{47}$$

$$= \int_{A\exp(-k^2/2)}^{A} f_y(y)F_x(-y)dy \tag{48}$$

$$= \int_{A\exp(-k^2/2)}^{A} f_y(y)\left(\frac{1}{2}\left(1 + erf\left(\frac{-y}{\sigma_n\sqrt{2}}\right)\right)\right)dy \tag{49}$$

$$= \frac{1}{2} - \frac{1}{(2k^2\sigma_n\sqrt{\pi})} \times \left(2\sqrt{2}A\,{}_2F_2\left(\left[\frac{1}{2}, \frac{1}{2}\right], \left[\frac{3}{2}, \frac{3}{2}\right], \frac{-A^2}{2\sigma_n^2}\right)\right. \tag{50}$$

$$\left. - 2\sqrt{2}A\exp\left(\frac{-k^2}{2}\right){}_2F_2 \times \left(\left[\frac{1}{2}, \frac{1}{2}\right], \left[\frac{3}{2}, \frac{3}{2}\right], \frac{-A^2\exp(-k^2)}{2\sigma_n^2}\right)\right) \tag{51}$$

where ${}_2F_2$ is hypergeometric function. The derived expression for the CDF of total signal $F_Z(z)$ can be used for predicting the percentage of negative samples depending on the linear scale A and the maximal Mahalanobis distance k.

ACKNOWLEDGMENT

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Croatian Science Foundation.

REFERENCES

[1] D. D. Lenz and T. R. Ayres, "Errors associated with fitting Gaussian profiles to noisy emission-line spectra," *Publications Astronomical Soc. Pacific*, vol. 104, p. 1104, Nov. 1992.

[2] J. J. Condon, "Errors in elliptical Gaussian FITS," *Publications Astronomical Soc. Pacific*, vol. 109, p. 166, Feb. 1997.
[3] J. Ireland, "Precision limits to emission-line profile measuring experiments," *Astrophys. J.*, vol. 620, no. 2, p. 1132, 2005.
[4] T. Delabie, J. De Schutter, and B. Vandenbussche, "An accurate and efficient Gaussian fit centroiding algorithm for star trackers," *J. Astron. Sci.*, vol. 61, no. 1, pp. 60–84, Mar. 2014.
[5] Z. Lei, X. Chen, and T. Lei, "Sub-pixel location of motion blurred weak celestial objects in optical sensor image based on elliptical 2D Gaussian surface fitting," in *Proc. Int. Conf. Ind. Informat.-Comput. Technol., Intell. Technol., Ind. Inf. Integr. (ICIICII)*, Dec. 2016, pp. 100–104.
[6] K. Okada, D. Comaniciu, and A. Krishnan, "Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT," *IEEE Trans. Med. Imag.*, vol. 24, no. 3, pp. 409–423, Mar. 2005.
[7] H. Shen, L. J. Tauzin, R. Baiyasi, W. Wang, N. Moringo, B. Shuang, and C. F. Landes, "Single particle tracking: From theory to biophysical applications," *Chem. Rev.*, vol. 117, no. 11, pp. 7331–7376, Jun. 2017.
[8] R. A. Caruana, R. B. Searle, T. Heller, and S. I. Shupack, "Fast algorithm for the resolution of spectra," *Anal. Chem.*, vol. 58, no. 6, pp. 1162–1167, 1986.
[9] H. Guo, "A simple algorithm for fitting a Gaussian function [DSP tips and tricks]," *IEEE Signal Process. Mag.*, vol. 28, no. 5, pp. 134–137, Sep. 2011.
[10] I. Al-Nahhal, O. A. Dobre, E. Basar, C. Moloney, and S. Ikki, "A fast, accurate, and separable method for fitting a Gaussian function [tips & tricks]," *IEEE Signal Process. Mag.*, vol. 36, no. 6, pp. 157–163, Nov. 2019.
[11] A. G. Koscevic and D. Petrinovic, "Maximizing accuracy of 2D Gaussian profile estimation using differential entropy," in *Proc. 45th Jubilee Int. Conv. Inf., Commun. Electron. Technol. (MIPRO)*, May 2022, pp. 888–893.
[12] Y. Feng, J. Goree, and B. Liu, "Accurate particle position measurement from images," *Rev. Sci. Instrum.*, vol. 78, no. 5, 2007, Art. no. 053704.
[13] S. Campbell, "Comparison of the generalized centroid with Gaussian and quadratic peak localization methods," 2018, *arXiv:1807.08355*.
[14] M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophys. J.*, vol. 81, no. 4, pp. 2378–2388, 2001.
[15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[16] S. M. Anthony and S. Granick, "Image analysis with rapid and accurate two-dimensional Gaussian fitting," *Langmuir*, vol. 25, no. 14, pp. 8152–8160, Jul. 2009.
[17] C. S. Smith, N. Joseph, B. Rieger, and K. A. Lidke, "Fast, single-molecule localization that achieves theoretically minimum uncertainty," *Nature Methods*, vol. 7, no. 5, pp. 373–375, May 2010.
[18] X. Zhu and D. Zhang, "Efficient parallel Levenberg–Marquardt model fitting towards real-time automated parametric imaging microscopy," *PLoS ONE*, vol. 8, no. 10, Oct. 2013, Art. no. e76665.
[19] M. Leutenegger and M. Weber, "Least-squares fitting of Gaussian spots on graphics processing units," 2021, *arXiv:2106.02045*.
[20] E. Chouzenoux, T. T.-K. Lau, C. Lefort, and J.-C. Pesquet, "Optimal multivariate Gaussian fitting with applications to PSF modeling in two-photon microscopy imaging," *J. Math. Imag. Vis.*, vol. 61, no. 7, pp. 1037–1050, Sep. 2019.
[21] M. Basseville and J.-F. Cardoso, "On entropies, divergences, and mean values," in *Proc. IEEE Int. Symp. Inf. Theory*, Sep. 1995, p. 330.
[22] N. Hagen, M. Kupinski, and E. L. Dereniak, "Gaussian profile estimation in one dimension," *Appl. Opt.*, vol. 46, no. 22, pp. 5374–5383, 2007.
[23] N. Hagen and E. L. Dereniak, "Gaussian profile estimation in two dimensions," *Appl. Opt.*, vol. 47, no. 36, pp. 6842–6851, 2008.
[24] 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, Nov. 2013.
[25] M. Kwak, B. Lkhagvasuren, and X. Sun, "An efficient method for fitting Gaussian functions," *Iranian J. Sci. Technol., Trans. A, Sci.*, vol. 45, no. 3, pp. 1043–1056, Jun. 2021.
[26] M. Merriman, *A List of Writings Relating to the Method of Least Squares: With Historical and Critical Notes*, vol. 4. New Haven, CT, USA: Transactions of the Connecticut Academy of Arts and Sciences, 1877.
[27] M. L. Johnson and L. M. Faunt, "[1] parameter estimation by least-squares methods," in *Methods in Enzymology*, vol. 210. Amsterdam, The Netherlands: Elsevier, 1992, pp. 1–37.

ANITA GRIBL KOŠČEVIĆ (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering and information technology and the M.Sc. degree in control engineering and automatization, in 2016 and 2018, respectively. She is currently pursuing the Ph.D. degree in technical sciences in the scientific field of signal processing with the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. Her research interests include image processing and applied statistics.

DAVOR PETRINOVIĆ (Member, IEEE) was born in Zagreb, Croatia, in 1965. He received the Dipl. (Ing.), M.Sc., and Dr.Sc. degrees in electrical engineering from the Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, in 1988, 1996, and 1999, respectively.

In 2010, he was appointed as a Full Professor with the Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb. He was a Fulbright Postdoctoral Scholar with the Signal Compression Laboratory, University of California at Santa Barbara, Santa Barbara, CA, USA, from 2000 to 2001, and a Visiting Researcher with the Sound and Image Processing Laboratory, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden, from 2005 to 2006. His current research interests include signal and image processing, speech and audio modeling, processing, and coding.