Noise parameter mismatch in variance stabilization, with an application to Poisson-Gaussian noise estimation

Markku Mäkitalo and Alessandro Foi

***Abstract*—In digital imaging, there is often a need to produce estimates of the parameters that define the chosen noise model. We investigate how the mismatch between the estimated and true parameter values affects the stabilization of variance of signal-dependent noise. As a practical application of the general theoretical considerations, we devise a novel approach for estimating Poisson-Gaussian noise parameters from a single image, combining variance-stabilization and noise estimation for additive Gaussian noise. Further, we construct a simple algorithm implementing the devised approach. We observe that when combined with optimized rational variance-stabilizing transformations (VST), the algorithm produces results that are competitive with those of a state-of-the-art Poisson-Gaussian estimator.**

***Index Terms*—parameter estimation, photon-limited imaging, Poisson-Gaussian noise, variance stabilization.**

I. Introduction

The random noise sources in a typical digital imaging chain are often collectively modelled as Poisson-Gaussian noise, where the Poisson noise component accounts for the inherent randomness associated with photon sensing, and the additive Gaussian noise component comprises the other random noise sources, such as thermal noise. In particular, this noise model can be characterized with two main parameters: the scaling factor of the Poisson component and the standard deviation of the Gaussian component.

Even though a variety of techniques exist for estimating the standard deviation of additive white Gaussian noise (see, e.g., [1], [2], [3], [4]), parameter estimation for the Poisson-Gaussian noise model is a less studied problem. A notable approach is to segment the image into approximately uniform regions, locally estimate the expected value (mean) and standard deviation of each region, and finally obtain global estimates of the implicit functional relationship between the image intensity and the noise standard deviation through regression of these local estimates [5], [6], [7], [8]. This can be thought as fitting a curve to a scatter plot in the Cartesian plane where the abscissa and ordinate are, respectively, the mean and standard deviation (or variance) of the image samples. In case of a time-variant model (as due to decay of markers in fluorescence imaging), the problem has also been addressed through an expectation-maximization algorithm [9].

This work was supported by the Academy of Finland (project no. 252547, Academy Research Fellow 2011–2016) and in part by Tampere Doctoral Programme in Information Science and Engineering (TISE).

The authors are with the Department of Signal Processing, Tampere University of Technology, P.O. Box 553, 33101 Tampere, Finland (e-mail: firstname.lastname@tut.fi).

Regardless of the particular noise model, the problem of estimating the parameter(s) associated with it can be approached either by taking advantage of multiple images (see, e.g., [10]), or by considering only a single image. We will concentrate on parameter estimation from a single image. While [5] and [9] mainly target parameter estimation from multiple images, [6] is able to provide very good results also for the single image case, with [8] reporting performance comparable to [6].

More than parameter estimation as such, we focus on the effect of inaccurate parameter estimation in stabilizing the variance of signal-dependent noise. In particular, we consider a general parametric noise model and investigate the relationship between the parameter mismatch and the standard deviation of stabilized noisy data. We present theoretical results concerning this mismatch, and as a proof of concept, we construct a novel algorithm for estimating Poisson-Gaussian noise parameters from a single image using an iterative variance-stabilization scheme. The main purpose of the proposed algorithm is to expose the potential of the novel viewpoint in contrast with the more traditional methods, and to confirm the practical significance of the presented theoretical considerations.

The paper is organized as follows. In Section II we introduce the Poisson-Gaussian noise model. As a leading example, Section III presents the generalized Anscombe transformation (GAT) that is typically used for stabilizing the variance of Poisson-Gaussian data. Section IV, which is the core contribution of the paper, considers the mismatch between the estimated and true parameter values for a general noise distribution. We show that under certain simplifying assumptions, the unitary contours of a function returning the standard deviation of stabilized data are smooth in a neighbourhood of the true parameter values and intersect at these true values; we prove this result explicitly for the Poisson-Gaussian noise model.[1] In Section V we apply these results by devising a proof-of-concept algorithm for estimating the parameters of the Poisson-Gaussian noise model. Section VI consists of experimental results, including comparisons to the regression-based parameter estimation algorithm [6], followed by the conclusions in Section VII. Finally, the appendices elaborate on two topics: in Appendix A, we discuss the implementation details of the proposed estimation algorithm; and in Appendix B, we present our construction of a family of

[1]The theoretical results of Section IV were first introduced in the first author's doctoral thesis [11]. However, the proof-of-concept parameter estimation algorithm presented therein is different and less refined than the algorithm presented in this paper, so the former is no longer considered here.

optimized rational variance-stabilizing transformations (VST) used within this algorithm.

II. Poisson-Gaussian noise model

Let z_i, $i = 1, \ldots, N$ be the observed pixel intensity values obtained through an imaging device. In the Poisson-Gaussian noise model, each z_i is composed of a scaled Poisson component and an additive Gaussian component. Specifically, we can write

$$z_i = \alpha p_i + n_i, \tag{1}$$

where $\alpha > 0$ is a constant scaling term, p_i is an independent random Poisson variable with an underlying expected value $y_i \geq 0$ (the unknown noise-free pixel value), and n_i is a random Gaussian variable with zero mean and standard deviation $\sigma \geq 0$. In other words, $p_i \sim \mathcal{P}(y_i)$ and $n_i \sim \mathcal{N}(0, \sigma^2)$. Hence, Poisson-Gaussian noise is defined as

$$\eta_i = z_i - \alpha y_i. \tag{2}$$

While the Poisson-Gaussian noise model can also be defined to have a Gaussian component with possibly non-zero mean (as in, e.g., [9], [14]), for our purposes this constant shift can be considered a part of the noise-free signal y, as it does not affect the noise variance estimation.

III. Generalized Anscombe transformation

For the observed Poisson-Gaussian data z (1), the mean and the standard deviation of the data are linked by the following relations:

$$E\{z \mid y\} = \alpha y, \quad \operatorname{std}\{z \mid y\} = \sqrt{\alpha^2 y + \sigma^2}.$$

Processing z by a VST aims at rendering the distribution of z into approximately Gaussian with unitary variance (i.e., $\operatorname{var}\{f_{\alpha,\sigma}(z) \mid y\} \approx 1$). A common choice of VST is the generalized Anscombe transformation (GAT) [12]

$$f_{\alpha,\sigma}(z) = \begin{cases} \frac{2}{\alpha}\sqrt{\alpha z + \frac{3}{8}\alpha^2 + \sigma^2}, & z > -\frac{3}{8}\alpha - \frac{\sigma^2}{\alpha} \\ 0, & z \leq -\frac{3}{8}\alpha - \frac{\sigma^2}{\alpha} \end{cases}. \tag{3}$$

Since the GAT (3) is derived by zeroing the first-order term in the Taylor expansion of the stabilized variance, it satisfies $\operatorname{var}\{f_{\alpha,\sigma}(z) \mid y\} = 1 + \mathcal{O}(y^{-2})$ for $y \to \infty$. For the pure Poisson case (i.e., $\alpha = 1$, $\sigma = 0$), this transformation coincides with the traditional Anscombe transformation [13]. However, even though the GAT is constructed in order to provide accurate asymptotic stabilization, it is unable to provide very accurate stabilization for low mean intensities y. This deficiency may translate into mediocre performance in practical applications dealing with low-intensity data, as shown for instance in [14], [15], [16]. Even though exact stabilization of Poisson-Gaussian data is not possible [17], [18], more accurate stabilizers than the GAT can be obtained by optimization. Hence, in this paper we also construct and employ a family of optimized rational VSTs to mitigate this inherent limitation of the GAT. Our construction is presented in Appendix B, and for a more general and detailed discussion on the optimization of VSTs we refer to [19].

IV. Standard deviation contours

In this section, we investigate the parameter mismatch in variance stabilization. More specifically, we consider a scenario where the stabilization is done through possibly inaccurate estimated noise model parameters, instead of using the true parameter values. Our goal is then to describe how the standard deviation of stabilized data depends on the mismatch between the estimated parameters and the true parameters.

Let z_θ be a noisy image, where the noise model is parametrized by the vector $\theta = (\theta_1, \ldots, \theta_n)$, and $y \geq 0$ is the expected value of z_θ. Further, let $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_n)$ be the estimated values for the noise parameters θ, and $f_{\hat{\theta}}$ a variance-stabilizing transformation constructed using the estimated parameters. For instance, in the Poisson-Gaussian case, $z_{\alpha,\sigma}$ is an image corrupted by (1). Then, we can set $f_{\hat{\alpha},\hat{\sigma}}$ to be, e.g., an optimized VST or the GAT (3) with estimated parameters $\hat{\alpha}$ and $\hat{\sigma}$.

The following propositions are based on ideal hypotheses, which typically hold only asymptotically (as, e.g., in the Poisson-Gaussian case with GAT, as was commented in Section III).

Proposition 1. *Assume that*

1) *The variance-stabilizing transformation constructed using the correct noise parameters achieves exact stabilization:*
$$\operatorname{std}\{f_\theta(z_\theta) \mid y\} = 1 \ \ \forall y \geq 0 \ \ \forall \theta. \tag{4}$$
2) *For any combination of parameters $\hat{\theta}$ and θ, the conventional approximation (which follows from a first-order Taylor expansion of $f_{\hat{\theta}}(z_\theta)$)*
$$\operatorname{std}\{f_{\hat{\theta}}(z_\theta) \mid y\} \approx \operatorname{std}\{z_\theta \mid y\}\, f'_{\hat{\theta}}(E\{z_\theta \mid y\}) \ \ \forall y \geq 0 \tag{5}$$
holds exactly (i.e., that the above symbol $\approx$ can be replaced by $=$).

Further, we denote the mean standard deviation over a block B as

$$\mathfrak{E}_B\{z_\theta\} := \int \operatorname{std}\{z_\theta \mid y\}\, p_B(y)\, dy, \tag{6}$$

where $p_B(y)$ is the probability density (i.e., the normalized histogram) of y over B.

Under the two assumptions, the mean standard deviation of the stabilized image block $f_{\hat{\theta}}(z_\theta)$ can be written as

$$\mathfrak{E}_B\{f_{\hat{\theta}}(z_\theta)\} = \int \frac{\operatorname{std}\{z_\theta \mid y\}}{\operatorname{std}\{z_{\hat{\theta}} \mid y\}} p_B(y)\, dy. \tag{7}$$

Proof. Applying (5) to (4) and noting that $E\{z_\theta \mid y\} = y$, we can write

$$\operatorname{std}\{f_\theta(z_\theta) \mid y\} = \operatorname{std}\{z_\theta \mid y\}\, f'_\theta(y) = 1,$$

and hence

$$f'_\theta(y) = \frac{1}{\operatorname{std}\{z_\theta \mid y\}}. \tag{8}$$

Thus, for the VST $f_{\hat{\theta}}$ with the estimated parameters $\hat{\theta}$,

combining (4) and (8) leads to

$$\begin{aligned}\operatorname{std}\left\{f_{\hat{\theta}}\left(z_\theta\right)|y\right\} &= \operatorname{std}\left\{z_\theta|y\right\} f'_{\hat{\theta}}\left(y\right)\\ &= \operatorname{std}\left\{z_\theta|y\right\} f'_{\theta}\left(y\right)\frac{f'_{\hat{\theta}}\left(y\right)}{f'_{\theta}\left(y\right)}\\ &= \operatorname{std}\left\{f_{\theta}\left(z_\theta\right)|y\right\}\frac{f'_{\hat{\theta}}\left(y\right)}{f'_{\theta}\left(y\right)}\\ &= \frac{f'_{\hat{\theta}}\left(y\right)}{f'_{\theta}\left(y\right)} = \frac{\operatorname{std}\left\{z_\theta|y\right\}}{\operatorname{std}\left\{z_{\hat{\theta}}|y\right\}}. \qquad (9)\end{aligned}$$

In particular, in the last equality we used (8) and its equivalent form $f'_{\hat{\theta}}\left(y\right) = 1/\operatorname{std}\left\{z_{\hat{\theta}}|y\right\}$, which is valid since the assumptions hold for arbitrary combinations of $\hat{\theta}$ and θ.

Let us now consider an image block B, with $p_B\left(y\right)$ being the probability density of y over this block. It then follows from (9) that

$$\begin{aligned}\mathfrak{E}_B\left\{f_{\hat{\theta}}\left(z_\theta\right)\right\} &= \int \operatorname{std}\left\{f_{\hat{\theta}}\left(z_\theta\right)|y\right\} p_B\left(y\right) dy\\ &= \int \frac{\operatorname{std}\left\{z_\theta|y\right\}}{\operatorname{std}\left\{z_{\hat{\theta}}|y\right\}} p_B\left(y\right) dy. \qquad (10)\end{aligned}$$

■

In what follows, we concentrate on the Poisson-Gaussian noise model, with true parameters α, σ and their respective estimates $\hat{\alpha}$, $\hat{\sigma}$. Let us denote

$$F_B\left(\hat{\alpha},\hat{\sigma}\right) := \mathfrak{E}_B\left\{f_{\hat{\alpha},\hat{\sigma}}\left(z_{\alpha,\sigma}\right)\right\}. \qquad (11)$$

Our aim is to show that, given the assumptions in Proposition 1, the function $F_B\left(\hat{\alpha},\hat{\sigma}\right)$ returning the standard deviation of a stabilized block has a well-behaving unitary contour near the true parameter values α, σ. More precisely, we show that such a contour is locally a simple smooth curve in a neighbourhood of (α,σ) in the $(\hat{\alpha},\hat{\sigma})$ plane.

Proposition 2. *Let F_B be defined as in (11), and let the assumptions of Proposition 1 hold. Then the solution of $F_B\left(\hat{\alpha},\hat{\sigma}\right) = 1$ is locally a simple smooth curve in a neighbourhood of (α,σ) in the $(\hat{\alpha},\hat{\sigma})$ plane.*

Proof. For Poisson-Gaussian noise, we have $\operatorname{std}\left\{z_{\alpha,\sigma}|y\right\} = \sqrt{\alpha^2 y + \sigma^2}$. Thus, according to Proposition 1, we can write (11) as

$$F_B\left(\hat{\alpha},\hat{\sigma}\right) = \int \sqrt{\frac{\alpha^2 y + \sigma^2}{\hat{\alpha}^2 y + \hat{\sigma}^2}} p_B\left(y\right) dy.$$

Clearly, $F_B\left(\alpha,\sigma\right) = 1$. Let us consider the implicit function theorem for F_B at (α,σ). For the theorem to hold, we need to show that either $\frac{\partial F_B}{\partial\hat{\alpha}}\left(\alpha,\sigma\right) \neq 0$ or $\frac{\partial F_B}{\partial\hat{\sigma}}\left(\alpha,\sigma\right) \neq 0$. With straightforward calculations, we see that

$$\begin{aligned}\frac{\partial F_B}{\partial\hat{\alpha}} &= \int \frac{\partial}{\partial\hat{\alpha}}\sqrt{\frac{\alpha^2 y + \sigma^2}{\hat{\alpha}^2 y + \hat{\sigma}^2}} p_B\left(y\right) dy\\ &= -\int \hat{\alpha} y \frac{\sqrt{\left(\alpha^2 y + \sigma^2\right)\left(\hat{\alpha}^2 y + \hat{\sigma}^2\right)}}{\left(\hat{\alpha}^2 y + \hat{\sigma}^2\right)^2} p_B\left(y\right) dy,\end{aligned}$$

and similarly

$$\frac{\partial F_B}{\partial\hat{\sigma}} = -\int \hat{\sigma} \frac{\sqrt{\left(\alpha^2 y + \sigma^2\right)\left(\hat{\alpha}^2 y + \hat{\sigma}^2\right)}}{\left(\hat{\alpha}^2 y + \hat{\sigma}^2\right)^2} p_B\left(y\right) dy.$$

Thus,

$$\begin{aligned}\frac{\partial F_B}{\partial\hat{\alpha}}\left(\alpha,\sigma\right) &= -\int \frac{\alpha y}{\alpha^2 y + \sigma^2} p_B\left(y\right) dy, \qquad (12)\\ \frac{\partial F_B}{\partial\hat{\sigma}}\left(\alpha,\sigma\right) &= -\int \frac{\sigma}{\alpha^2 y + \sigma^2} p_B\left(y\right) dy. \qquad (13)\end{aligned}$$

The partial derivative (13) equals 0 if and only if $\sigma = 0$. However, if $\sigma = 0$, then $\frac{\partial F_B}{\partial\hat{\alpha}}\left(\alpha,\sigma\right) = -\frac{1}{\alpha} < 0$. Therefore we have that either $\frac{\partial F_B}{\partial\hat{\alpha}}\left(\alpha,\sigma\right) < 0$ or $\frac{\partial F_B}{\partial\hat{\sigma}}\left(\alpha,\sigma\right) < 0$, and hence $F_B\left(\hat{\alpha},\hat{\sigma}\right) = 1$ is locally a simple smooth curve in a neighbourhood of (α,σ). ■

Let us further discuss the properties of the slope. In particular, the implicit function theorem also states that the slope of the curve at (α,σ) is given by the ratio of the above partial derivatives (12) and (13):

$$s := -\frac{\frac{\partial F_B}{\partial\hat{\alpha}}\left(\alpha,\sigma\right)}{\frac{\partial F_B}{\partial\hat{\sigma}}\left(\alpha,\sigma\right)} \qquad (14)$$

Hence, we can write

$$\hat{\sigma} = s\left(\hat{\alpha} - \alpha\right) + \sigma. \qquad (15)$$

For a uniform block B with value y, the probability density p_B is a Dirac delta centered at y. In that case,

$$\begin{aligned}\frac{\partial F_B}{\partial\hat{\alpha}}\left(\alpha,\sigma\right) &= -\frac{\alpha y}{\alpha^2 y + \sigma^2},\\ \frac{\partial F_B}{\partial\hat{\sigma}}\left(\alpha,\sigma\right) &= -\frac{\sigma}{\alpha^2 y + \sigma^2},\end{aligned}$$

and the slope equals

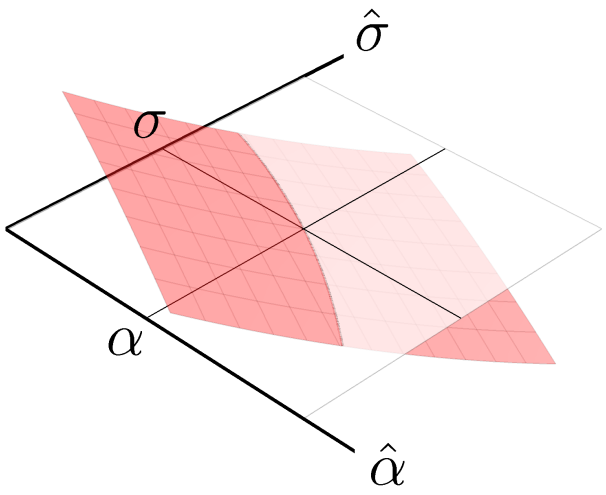
$$s = -\frac{\alpha y}{\sigma}. \qquad (16)$$

Thus, the slope can attain values in the range $[-\alpha/\sigma \cdot \text{peak}, 0]$. In particular, the minimum (negative) value of the slope is attained with a uniform white block, and the maximum value 0 with a uniform black block.

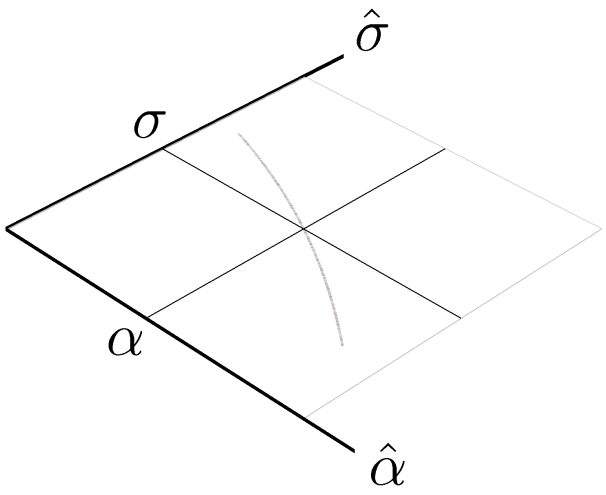
While it is easy to construct examples of blocks having distinct normalized histograms yet identical slopes (14), typically different blocks yield different slopes, as we show in the following examples. We also point out that since all the curves are locally smooth and pass through (α,σ), different slopes imply a unique intersection located at (α,σ); this observation is also visualized in Figure 1, which depicts (idealized) stabilized standard deviation surfaces and their unitary contours for blocks with different slopes.

In order to illustrate the results of the above propositions, Figure 2(b) shows ten contours $F_{B_i}\left(\hat{\alpha},\hat{\sigma}\right) = 1$ computed with the GAT from ten randomly selected 32×32 blocks B_i of a Poisson-Gaussian corrupted test image "Armchair" (peak= 120, $\alpha = 1$, $\sigma = 12$), which is shown in Figure 2(a). The stabilized standard deviation of each block is estimated by the sample standard deviation of the wavelet detail coefficients of the block. In this high-intensity case, the GAT is able to provide very good stabilization. As the assumptions of Proposition 1 hold for these ranges of α and σ, the contours intersect near the true parameter values $\alpha = 1$, $\sigma = 12$.

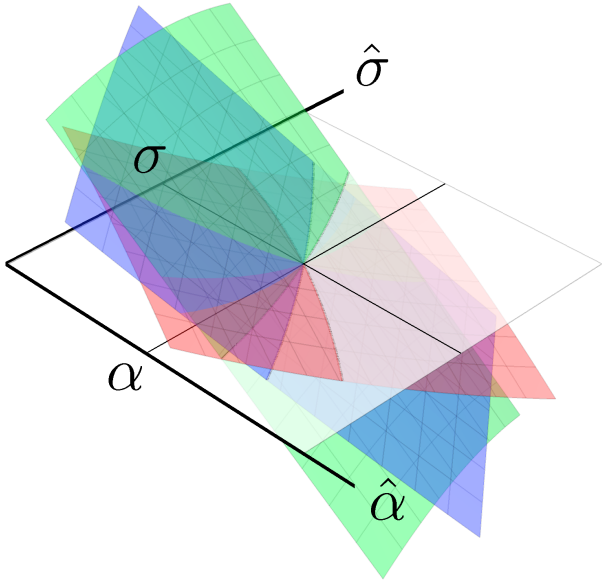
Figure 3 shows a similar visualization for a low-intensity case (Piano, peak= 5, $\alpha = 1$, $\sigma = 0.5$); in this case it is clear that when using the GAT (Figure 3(b)), our initial assumptions do not all hold, and the contours fail to intersect at a point. In contrast, the stabilization obtained with the optimized VST



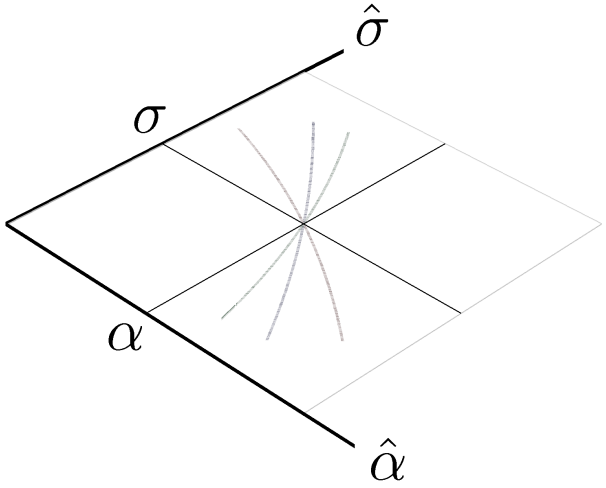
(a) $F_{B_i}(\hat{\alpha}, \hat{\sigma}) - 1$ for one block ($i = 1$).



(b) Unitary contour $F_{B_i}(\hat{\alpha}, \hat{\sigma}) = 1$ of the surface shown in (a).



(c) $F_{B_i}(\hat{\alpha}, \hat{\sigma}) - 1$ for three blocks ($i = 1, 2, 3$).



(d) Unitary contours $F_{B_i}(\hat{\alpha}, \hat{\sigma}) = 1$ of the surfaces shown in (c).

Fig. 1. Idealized illustration of the stabilized standard deviation surfaces and their corresponding unitary contours $F_{B_i}(\hat{\alpha}, \hat{\sigma}) = 1$ in the $(\hat{\alpha}, \hat{\sigma})$ plane, for different blocks B_i; all contours intersect at the true parameters (α, σ).

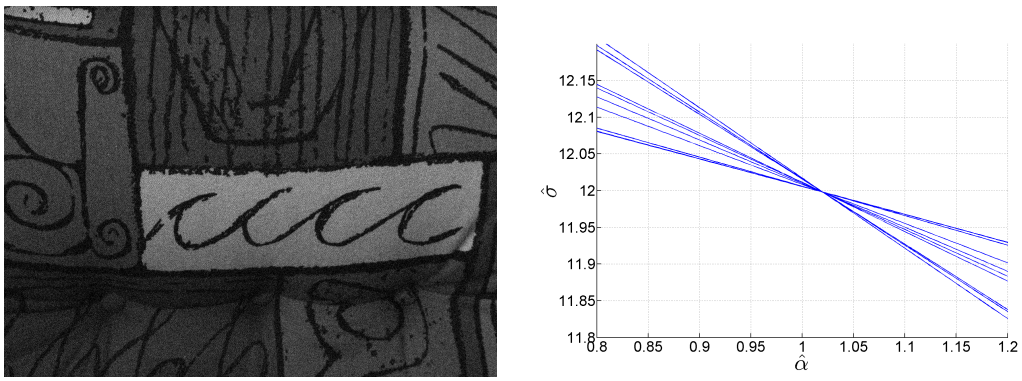
(as described in Appendix B) is highly accurate, in which case the contours intersect very close to the true parameter values $\alpha = 1$, $\sigma = 0.5$, as shown in Figure 3(c).

Since the slopes range in the interval $[-\alpha/\sigma \cdot \text{peak}, 0]$, the intersection is likely to be less stable for smaller values of α and peak. Note that such values also bring the problem far from the asymptotical regime on which the propositions are built. Nevertheless, the above examples demonstrate good agreement between the practical case and the idealized model; the use of an optimized VST helps in maintaining this agreement over a wider range of the parameter values.

V. The Poisson-Gaussian parameter estimation approach

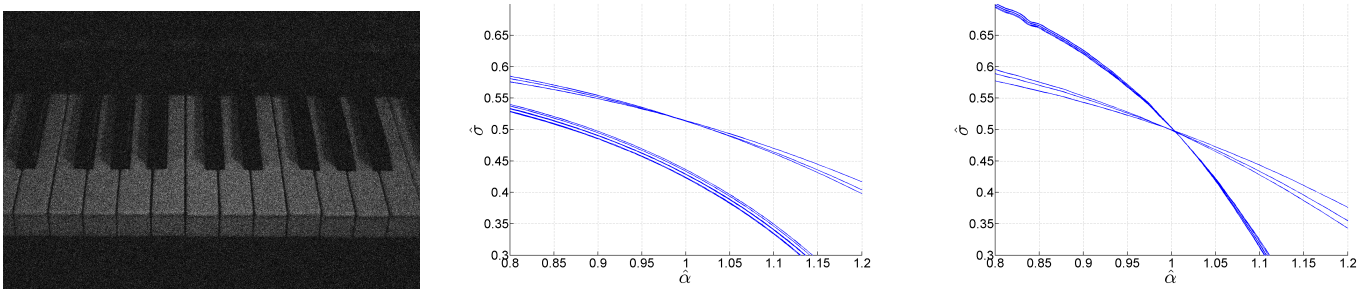
As shown in Section IV, the unitary variance contours corresponding to different stabilized image blocks are locally smooth in the $(\hat{\alpha}, \hat{\sigma})$ plane. Moreover, when the probability densities of these image blocks are different, the contours are typically differently oriented and intersect each other. Hence, these results suggest that we could estimate the true parameter values α and σ of Poisson-Gaussian noise by constructing an algorithm that searches for this intersection of the unitary variance contours.

Specifically, we suggest that the estimation is performed by



(a) (b)

Fig. 2. (a) Poisson-Gaussian corrupted test image "Armchair" (1224×922), with peak intensity value 120 and noise parameters $\alpha = 1$, $\sigma = 12$. (b) Ten standard deviation contours $F_{B_i}(\hat{\alpha}, \hat{\sigma}) = 1$ computed from ten randomly selected 32×32 blocks B_i of the test image (a), using the GAT. On this range, the stabilization achieved with the GAT is very good.



(a) (b) (c)

Fig. 3. (a) Poisson-Gaussian corrupted test image "Piano" (1193×795), with peak intensity value 5 and noise parameters $\alpha = 1$, $\sigma = 0.5$. (b) Ten standard deviation contours $F_{B_i}(\hat{\alpha}, \hat{\sigma}) = 1$ computed from ten randomly selected 32×32 blocks B_i of the test image (a), using the GAT. On this range, the GAT is unable to provide accurate stabilization. (c) Ten standard deviation contours $F_{B_i}(\hat{\alpha}, \hat{\sigma}) = 1$ computed from the same ten blocks as in (b), using the optimized VST.

stabilizing the variance of a collection of sub-blocks of an image, and then searching for the intersection by minimizing a suitable cost functional. The main steps of the general approach are as follows:

1) Initialize the estimates $\hat{\alpha}$ and $\hat{\sigma}$.
2) Choose K random blocks B_i, $i = 1, \ldots, K$ from the noisy image $z_{\alpha,\sigma}$.
3) Apply a variance-stabilizing transformation $f_{\hat{\alpha},\hat{\sigma}}(z_{\alpha,\sigma})$ to each block.
4) Compute an estimate $F_{B_i}(\hat{\alpha}, \hat{\sigma}) = \mathfrak{E}_{B_i}\{f_{\hat{\alpha},\hat{\sigma}}(z_{\alpha,\sigma})\}$ for the standard deviation of each stabilized block (using any AWGN standard deviation estimator $\mathfrak{E}$).
5) Optimize the estimates $\hat{\alpha}$ and $\hat{\sigma}$ in such a way that they minimize the difference between $F_B(\hat{\alpha}, \hat{\sigma})^2$ and 1 (target variance) over the K blocks, according to any cost functional of choice.

In order to confirm the validity of this approach, we also construct an algorithm implementing it in practice. For more details on our implementation, we refer the reader to Appendix A.

VI. Experiments

Our experiments are twofold. In the first part, we evaluate the accuracy and effectiveness of the proposed parameter estimation algorithm by artificially corrupting images with Poisson-Gaussian noise, allowing comparisons to the original ground truth images. In the second part, we estimate the parameters directly from raw images captured with various cameras.

A. Estimation with ground truth comparisons

For these experiments, we first create two virtually noise-free ground truth images by averaging and downsampling a set of raw images obtained with high-quality digital cameras at a low ISO setting. The resulting 1-megapixel images "Piano" and "Armchair" are shown in Figure 4.

Each test image is scaled to a specific peak intensity value, yielding an image y. A set of noisy images z are then produced from y according to the model (1), using various combinations of the Poisson-Gaussian noise parameters α and σ. The parameters α and σ are estimated from each noisy image with the proposed algorithm, using either the optimized VST described in Appendix B, or the GAT (3). Moreover, we compare their estimation accuracy with that of the algorithm presented in [6]. Specifically, the accuracy is measured with a normalized root mean squared error metric Φ:

$$\Phi^2 = \int_{\mathbb{R}^+} p(\xi) \frac{\left(\sqrt{\alpha^2\xi + \sigma^2} - \sqrt{\hat{\alpha}^2\xi + \hat{\sigma}^2}\right)^2}{\alpha^2\xi + \sigma^2} d\xi, \quad (17)$$

where $p(\xi)$ is the normalized histogram of y; α, σ are the

(a) (b)

Fig. 4. The ground truth test images used in the experiments of Section VI-A. (a) Piano (1193×795), (b) Armchair (1224×922).

true noise parameter values, and $\hat{\alpha}$, $\hat{\sigma}$ are their corresponding estimates.

Table I presents the parameter estimation results for peak intensity values $2, 10, 30$ and 120. When combined with the optimized VST, the proposed algorithm yields results that are competitive with the results obtained with [6]. We also observe that both algorithms have a slightly better overall performance on Piano than on Armchair, which is likely explained by the highly textured nature of Armchair. Moreover, it is evident that the optimized VST plays an important role in the estimation performance for the low-intensity cases; since it provides highly accurate stabilization also for low intensities, our algorithm performs acceptably also in these cases. In contrast, the GAT is inherently unable to perform accurate variance stabilization in regions with low mean intensity values; this lack of accurate stabilization is in clear violation with the condition (4) assumed in Proposition 1, on the grounds of which our algorithm was devised.

Figure 5 illustrates, through three example cases, how the estimated parameter values typically evolve as the proposed algorithm performs the iterative updating. In Figure 5(a), the parameters are estimated from Piano with peak $= 2$, $\alpha = 1$, $\sigma = 0.2$. Despite the low-intensity setting, the optimized VST used within the estimation algorithm helps in producing accurate estimates, converging to $\hat{\alpha} = 1.008$, $\hat{\sigma} = 0.186$ after 37 iterations. Figures 5(b) and 5(c) examine two different noise realizations of Armchair with peak $= 10$, $\alpha = 2.5$, $\sigma = 1$. In Figure 5(b), the algorithm converges to $\hat{\alpha} = 2.495$, $\hat{\sigma} = 1.028$ after 19 iterations. However, in Figure 5(c), we come across a case where the true parameter values $\alpha = 2.5$, $\sigma = 1$ do not correspond to the global minimum of the cost functional used in our implementation of Step 5 of Section V (our cost functional is defined in (18) in Appendix A). Hence, our algorithm eventually converges to $\hat{\alpha} = 2.708$, $\hat{\sigma} = 0.036$, yielding a slightly (less than 0.5 %) lower cost functional value than what would be obtained with the true parameter values. This suggests that there is still some room for improvement in the definition of our cost functional, even though in the large majority of the test cases the correspondence between the global minimum and the true parameter values is more adequate.

Figures 6–8 provide an illustration for the obtained estimation accuracy in terms of the error metric (17). Specifically, they visualize the estimated and true standard deviation curves ($\mathrm{std}\{z_{\hat{\alpha},\hat{\sigma}}|y\} = \sqrt{\hat{\alpha}^2 y + \hat{\sigma}^2}$ and $\mathrm{std}\{z_{\alpha,\sigma}|y\} = \sqrt{\alpha^2 y + \sigma^2}$, respectively) for the three example cases discussed above. In addition, the histograms of the corresponding noisy and noise-free images are overlaid to each comparison. This aids in identifying the intensity ranges, where the accuracy or inaccuracy of the estimates will have practical significance; we remind that (17) is evaluated over the range of the noise-free image. In particular, despite the discrepancy between the true and estimated parameter values in our third example, Figure 8 shows that the estimated standard deviation curve still looks reasonable over the range of the noise-free image.

B. Estimation from raw data

In the second part of our experiments, we consider noisy raw images (Figure 9) captured by three different cameras. Specifically, we use each camera to image one target with two different ISO settings, after which we extract the blue colour channel of each image; the raw data is scaled to range $[0, 1]$. First, we have two images labeled "Cube" (2385×1589), captured by a Canon EOS 500D with ISO 200 and ISO 3200. The second pair is "Abacus" (1224×922), taken with a Fujifilm FinePix S9600 with ISO 200 and ISO 800. Finally, a Canon PowerShot S90 is used to obtain "Flute" (1842×1380), with ISO 400 and ISO 1600.

We estimate the parameters α and σ from each of these images, using both the proposed algorithm and [6]. As earlier, the proposed algorithm is combined with either the optimized VST or the GAT. The results are summarized in Table II. Overall, we see that in all six test cases the estimates produced by each algorithm are in reasonably good agreement with each other. Further, all of the cases fall into a range where the GAT is also able to stabilize the variance adequately, so it produces very similar estimates as the optimized VST.

VII. Discussion

We investigated the effect of inaccurate parameter estimation on variance stabilization for a general noise distribution.

TABLE I

A COMPARISON OF THE ESTIMATION ERROR VALUES Φ (17) FOR THE PROPOSED ALGORITHM AND THE ALGORITHM PRESENTED IN [6], FOR VARIOUS PEAK INTENSITY VALUES WITH $\sigma = \text{PEAK}/10$ AND VARIOUS VALUES OF α. EACH ERROR VALUE IS REPORTED AS AN AVERAGE OVER 10 NOISE REALIZATIONS, FOLLOWED BY THE STANDARD DEVIATION OF THESE 10 ERROR VALUES.

Image	Peak	α	σ	Noisy PSNR	Proposed + opt. VST	Proposed + GAT	[6]
Piano	2	0.5	0.2	7.06	0.042 ± 0.002	0.286 ± 0.008	**0.024 ± 0.009**
	2	1.0	0.2	7.77	**0.006 ± 0.002**	0.448 ± 0.005	0.028 ± 0.011
	2	2.5	0.2	8.00	**0.007 ± 0.005**	0.676 ± 0.007	0.056 ± 0.016
	10	0.5	1.0	11.46	**0.006 ± 0.003**	0.021 ± 0.002	0.011 ± 0.007
	10	1.0	1.0	13.83	**0.006 ± 0.002**	0.023 ± 0.004	0.014 ± 0.006
	10	2.5	1.0	14.82	**0.005 ± 0.004**	0.013 ± 0.005	0.016 ± 0.008
	30	0.5	3.0	12.97	**0.006 ± 0.003**	**0.006 ± 0.003**	0.016 ± 0.007
	30	1.0	3.0	16.89	**0.005 ± 0.002**	0.009 ± 0.002	0.016 ± 0.007
	30	2.5	3.0	19.18	**0.005 ± 0.003**	0.008 ± 0.002	0.014 ± 0.006
	120	0.5	12.0	13.70	**0.004 ± 0.002**	**0.004 ± 0.002**	0.015 ± 0.007
	120	1.0	12.0	18.99	**0.004 ± 0.002**	**0.004 ± 0.002**	0.013 ± 0.006
	120	2.5	12.0	23.75	**0.005 ± 0.003**	**0.005 ± 0.003**	0.014 ± 0.007
Armchair	2	0.5	0.2	7.30	0.073 ± 0.002	0.300 ± 0.024	**0.025 ± 0.006**
	2	1.0	0.2	8.06	**0.010 ± 0.004**	0.467 ± 0.006	0.041 ± 0.010
	2	2.5	0.2	8.30	**0.008 ± 0.007**	0.659 ± 0.010	0.077 ± 0.008
	10	0.5	1.0	11.60	**0.010 ± 0.004**	0.025 ± 0.001	0.012 ± 0.004
	10	1.0	1.0	14.06	**0.005 ± 0.002**	0.031 ± 0.004	0.015 ± 0.004
	10	2.5	1.0	15.11	**0.013 ± 0.018**	0.033 ± 0.002	0.015 ± 0.006
	30	0.5	3.0	13.03	**0.007 ± 0.002**	**0.007 ± 0.002**	0.015 ± 0.003
	30	1.0	3.0	17.04	**0.007 ± 0.003**	0.014 ± 0.002	0.023 ± 0.007
	30	2.5	3.0	19.44	**0.014 ± 0.006**	0.019 ± 0.003	0.030 ± 0.007
	120	0.5	12.0	13.72	**0.005 ± 0.003**	**0.005 ± 0.003**	0.015 ± 0.004
	120	1.0	12.0	19.05	**0.012 ± 0.004**	**0.012 ± 0.004**	0.023 ± 0.005
	120	2.5	12.0	23.94	**0.038 ± 0.003**	**0.038 ± 0.003**	0.058 ± 0.010

TABLE II

A COMPARISON OF THE ESTIMATED PARAMETER VALUES $\hat{\alpha}$ AND $\hat{\sigma}$ FOR THE PROPOSED ALGORITHM AND THE ALGORITHM PRESENTED IN [6], FOR TEST IMAGES (FIGURE 9) CAPTURED WITH VARIOUS CAMERAS.

Image	Proposed + opt. VST		Proposed + GAT		[6]	
	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\sigma}$
Cube (ISO 200)	$6.16 \cdot 10^{-5}$	$5.53 \cdot 10^{-4}$	$6.16 \cdot 10^{-5}$	$5.53 \cdot 10^{-4}$	$6.21 \cdot 10^{-5}$	$5.31 \cdot 10^{-4}$
Cube (ISO 3200)	$8.90 \cdot 10^{-4}$	$2.85 \cdot 10^{-3}$	$9.13 \cdot 10^{-4}$	$2.84 \cdot 10^{-3}$	$9.20 \cdot 10^{-4}$	$2.47 \cdot 10^{-3}$
Abacus (ISO 200)	$5.05 \cdot 10^{-4}$	$2.74 \cdot 10^{-3}$	$5.05 \cdot 10^{-4}$	$2.74 \cdot 10^{-3}$	$5.17 \cdot 10^{-4}$	$2.71 \cdot 10^{-3}$
Abacus (ISO 800)	$2.94 \cdot 10^{-3}$	$7.84 \cdot 10^{-3}$	$2.93 \cdot 10^{-3}$	$8.01 \cdot 10^{-3}$	$2.21 \cdot 10^{-3}$	$1.02 \cdot 10^{-2}$
Flute (ISO 400)	$5.86 \cdot 10^{-4}$	$3.12 \cdot 10^{-3}$	$5.86 \cdot 10^{-4}$	$3.12 \cdot 10^{-3}$	$5.97 \cdot 10^{-4}$	$3.12 \cdot 10^{-3}$
Flute (ISO 1600)	$2.32 \cdot 10^{-3}$	$1.27 \cdot 10^{-2}$	$2.32 \cdot 10^{-3}$	$1.27 \cdot 10^{-2}$	$2.38 \cdot 10^{-3}$	$1.27 \cdot 10^{-2}$

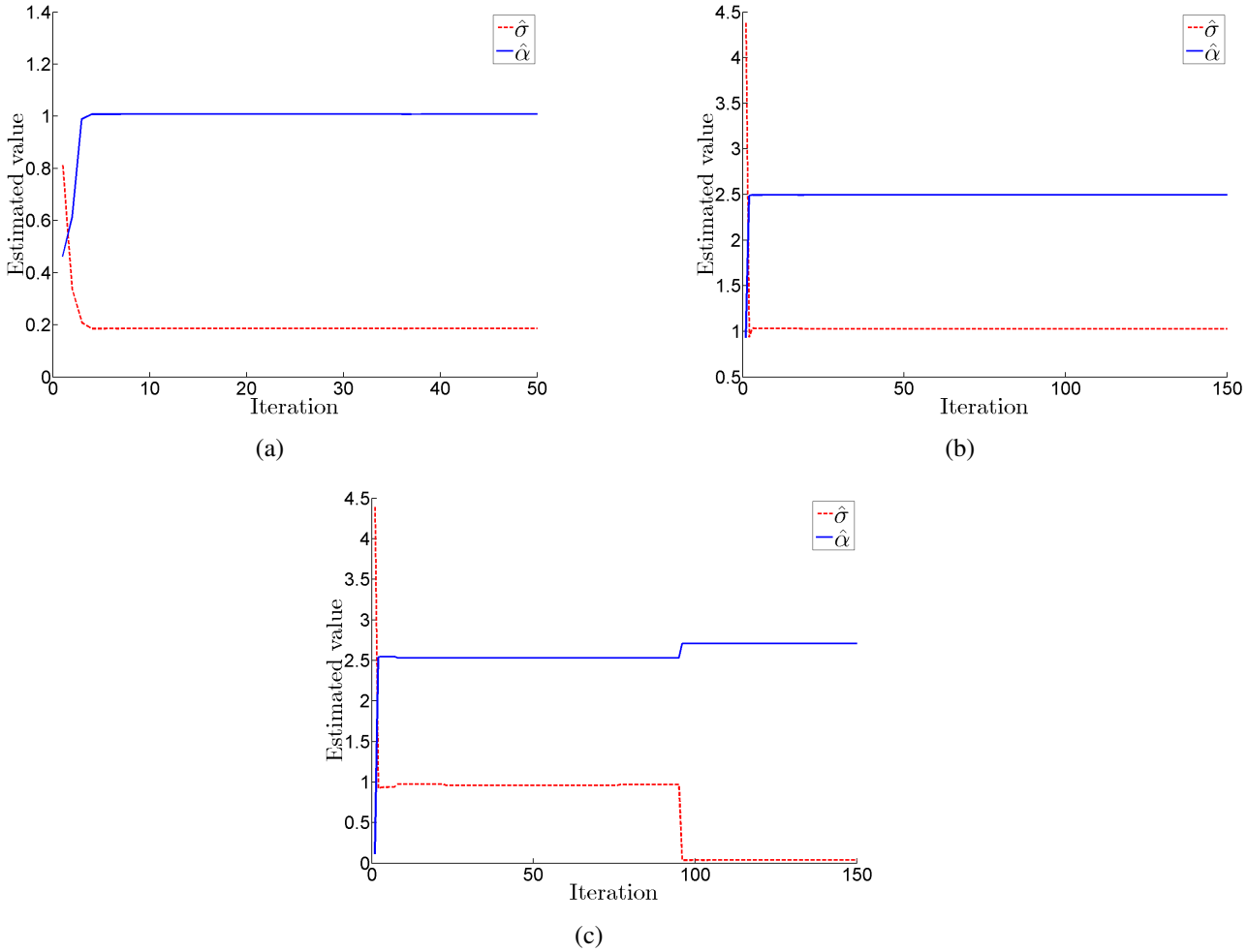


Fig. 5. Evolution of the parameter estimates $\hat{\alpha}$ and $\hat{\sigma}$. (a) Piano (peak $= 2$, $\alpha = 1$, $\sigma = 0.2$); estimates converge to $\hat{\alpha} = 1.008$, $\hat{\sigma} = 0.186$. (b) Armchair (peak $= 10$, $\alpha = 2.5$, $\sigma = 1$), noise realization 1; estimates converge to $\hat{\alpha} = 2.495$, $\hat{\sigma} = 1.028$. (c) Armchair (peak $= 10$, $\alpha = 2.5$, $\sigma = 1$), noise realization 2; estimates converge to $\hat{\alpha} = 2.708$, $\hat{\sigma} = 0.036$.

In particular, we showed that under some simplifying assumptions, the unitary contours of a function returning the standard deviation of stabilized data are smooth in a neighbourhood of the true parameter values. We derived a general expression for this function, and proved the smoothness explicitly for the Poisson-Gaussian noise model (Proposition 2). However, we expect the proposition to hold also for other common distribution families, with the proof proceeding analogously (showing that at least one partial derivative of $F_B(\hat{\theta})$ does not vanish in a neighbourhood of θ).

Based on the results, we proposed a variance-stabilization based iterative algorithm for estimating the parameters of Poisson-Gaussian noise, and compared its estimation performance against an algorithm based on segmentation, local estimates and regression [6]. Moreover, for the purposes of the proposed algorithm, we constructed a family of optimized rational VSTs for stabilizing Poisson-Gaussian data.

When combined with the optimized rational VSTs proposed in Appendix B, the estimation performance of our algorithm is generally comparable to that of the estimator presented in [6] (and consequently comparable to [8]). More importantly, the practical results corroborate the validity and usefulness of our theoretical analysis.

Compared to [6], our method has the advantage that it does not require a segmentation of the image. On the other hand, [6] is clearly faster than our current implementation of the proposed method, but the latter is not optimized for speed in any way. For instance, for the Armchair test image, [6] completes the estimation in about 3 seconds, whereas our estimator takes about 57 seconds per 10 iterations. Moreover, even though our estimator typically reaches a minimum within relatively few iterations, it is not always clear whether it is a local or global minimum; sometimes a slightly more favourable solution (i.e. lower cost functional value) is found after a considerable amount of iterations, as Figure 5(c) showed. Nevertheless, the main point of our implemented algorithm is to serve as a proof of concept, not to provide a fully polished estimator; drawbacks in our implementation should not be viewed as inherent limitations of the proposed general approach.

We wish to emphasize that the use of variance stabilization for the estimation of noise parameters is an untraditional use of VSTs, which has not been explored until recently. Pioneering work was presented in [20], introducing an estimation method

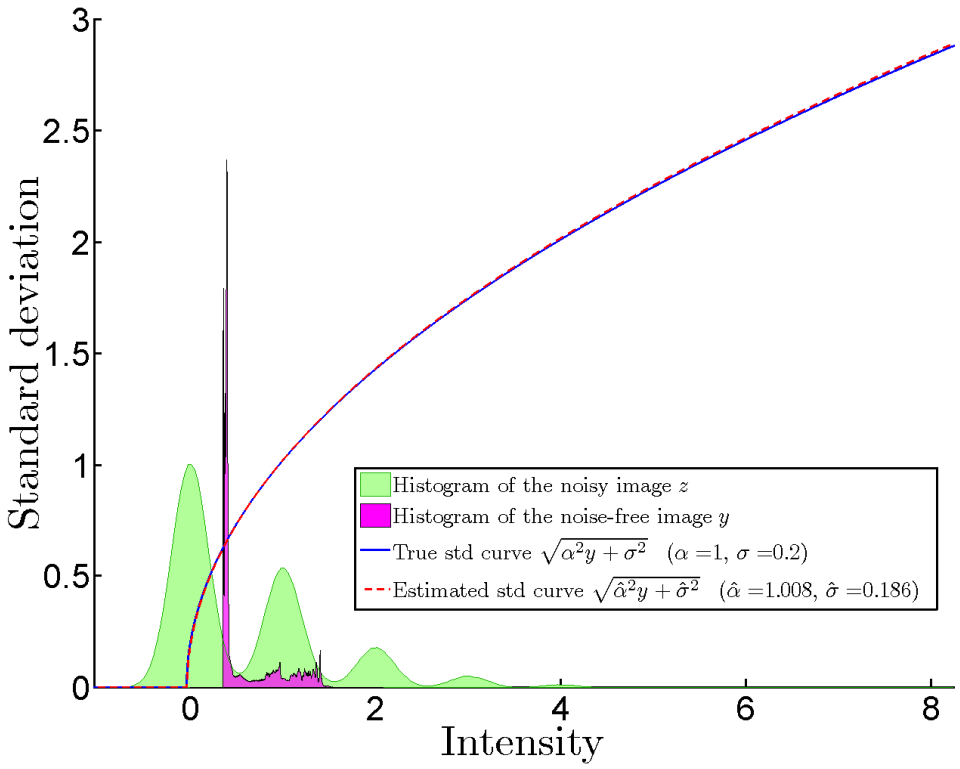


Fig. 6. Estimated and true standard deviation curves for the Piano image (peak = 2, $\alpha = 1$, $\sigma = 0.2$), overlaid with the histograms of the noisy and noise-free images.

that is much different than ours, yet still based on variance stabilization. However, that method and its theoretical development are not applicable to our case; whereas we consider the two-parameter Poisson-Gaussian case, a single parameter is sufficient for characterizing Rician noise, making its estimation relatively straightforward. Another early application of variance stabilization to noise estimation can be found in [21], which considers scaled Poisson noise including an additive constant. However, their method relies on additional data obtained from a blank image taken with the same camera under the same acquisition settings; if the additive constant is neglected, their method reduces to straightforward non-iterative estimation of a single Poisson scaling parameter, which can be treated as a particular case of Proposition 1. The results originally appearing in [21] were recently more widely published in [22]. Finally, we note that an equation corresponding to (9), which describes the parameter mismatch, was very recently independently[2] suggested for the Poisson-Gaussian case in [23] within a similar variance-stabilization framework. The equation is based on a first-order Taylor expansion that is equivalent to our assumption (5). However, it is not developed further towards the analysis of their sophisticated estimation algorithm, which employs PCA as well as sample kurtosis.

[2]Our equations (4)–(13) originally appeared in the first author's doctoral thesis [11].

APPENDIX A
IMPLEMENTATION OF THE ESTIMATION APPROACH

This appendix addresses our implementation of the estimation approach proposed in Section V, following the chronological order of the steps.

A. Initialization of the estimates

We initialize $\hat{\sigma}$ by estimating the standard deviation of the non-stabilized whole image $z_{\alpha,\sigma}$:

$$\hat{\sigma} = \mathfrak{E}\{z_{\alpha,\sigma}\} \approx \operatorname{std}\{z_{\alpha,\sigma}|y\}.$$

In particular, the estimate is obtained via the sample standard deviation of the wavelet detail coefficients of $z_{\alpha,\sigma}$. On the other hand, according to our observations, the choice of the initial values does not have a noticeable impact on the convergence of the algorithm. Hence, we initialize $\hat{\alpha}$ to a random positive value.

B. Choosing the random blocks

As shown in Section IV, the unitary variance contours corresponding to different stabilized image blocks are locally smooth in the $(\hat{\alpha}, \hat{\sigma})$ plane. Moreover, they are differently oriented and intersect each other, provided that the probability densities of these image blocks are different. Thus, after the initialization step, we process only a random collection of K (possibly overlapping) sub-blocks B_i of the noisy image $z_{\alpha,\sigma}$. In our approach, the block size is fixed to 32×32 pixels

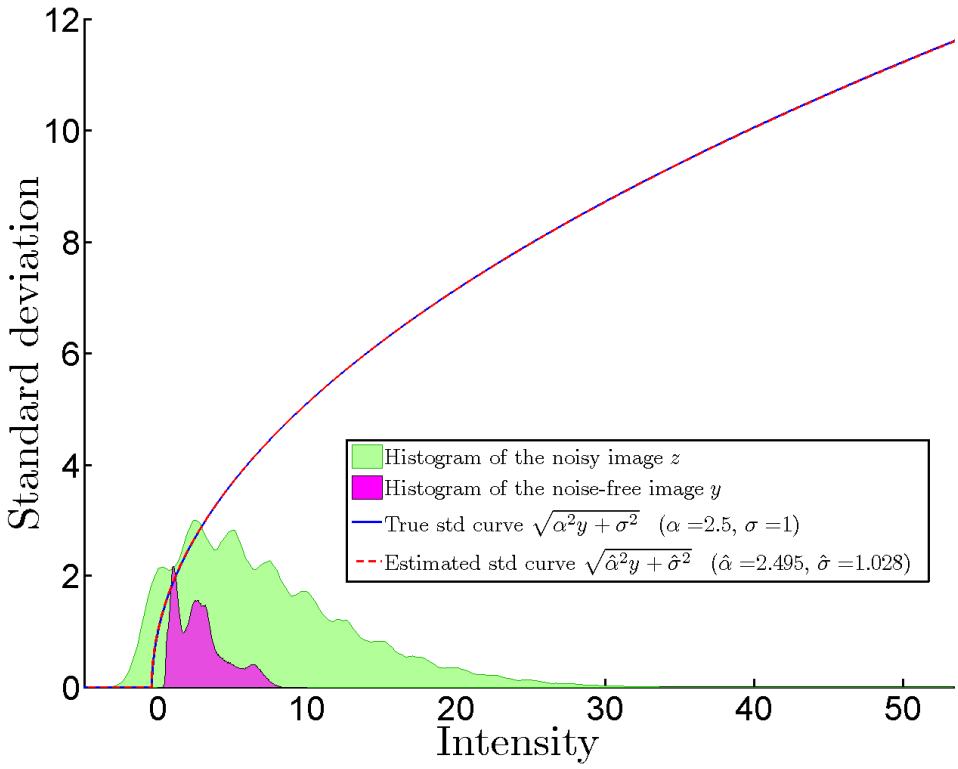


Fig. 7. Estimated and true standard deviation curves for the Armchair image (peak $= 10$, $\alpha = 2.5$, $\sigma = 1$), overlaid with the histograms of the noisy and noise-free images. For this noise realization, our algorithm successfully converges near the true values $\alpha = 2.5$, $\sigma = 1$.

and the number of blocks set to $K = 2000$, but adapting these parameters based on suitable heuristics may also be considered.

C. Stabilizing the variance of a random block

After extracting the random blocks, each of them is processed with a VST in order to stabilize its variance to (approximately) unity. Specifically, we use the optimized rational VST $f_{\hat{\alpha},\hat{\sigma}}$ presented in Appendix B, with $\hat{\alpha}$ and $\hat{\sigma}$ being the current parameter estimates. For a comparison, we perform the stabilization also with the GAT (3), where α and σ are similarly replaced by their current estimates $\hat{\alpha}$ and $\hat{\sigma}$.

D. Estimating the standard deviation

After applying the VST on each block B_i, $i = 1, \ldots, K$, the noise is treated as homoskedastic Gaussian. Thus, we can compute the standard deviation estimate $F_B(\hat{\alpha}, \hat{\sigma})$ for the stabilized image block $f_{\hat{\alpha},\hat{\sigma}}(z_{\alpha,\sigma})$ as

$$F_B(\hat{\alpha}, \hat{\sigma}) = \mathfrak{E}_B\{f_{\hat{\alpha},\hat{\sigma}}(z_{\alpha,\sigma})\},$$

where $\mathfrak{E}$ is any sample estimator of the standard deviation designed for the additive white Gaussian noise case. In particular, our estimate is obtained as the sample standard deviation of the wavelet detail coefficients of the block; this is the same estimator that was used in computing the initial estimate of σ.

E. Optimization of the estimates

The parameter estimates $\hat{\alpha}, \hat{\sigma}$, which determine the optimized VST $f_{\hat{\alpha},\hat{\sigma}}$ to be used, are optimized by minimizing the cost functional

$$C(\hat{\alpha}, \hat{\sigma}) = \underset{i=1,\ldots,K}{\text{mean}} \left| \Gamma\left(F_{B_i}(\hat{\alpha}, \hat{\sigma})^2 - 1 \right) \right|. \quad (18)$$

In practice, errors in the stabilized variance are regulated so that very large errors will have only a limited impact to the cost functional, and on the other hand very small errors will still have at least some impact. This is done by passing the error through a nonlinear response function Γ, which is illustrated in Figure 10.

The minimization is done by applying the Nelder-Mead algorithm iteratively, each time restarting the process with the values $\hat{\alpha}$ and $\hat{\sigma}$ from the previous round as initial values, as long as the value of the cost functional keeps decreasing. However, if a restart does not yield a lower value of the cost functional than in the preceding restart, we perturb $\hat{\alpha}$ and $\hat{\sigma}$ in order to avoid getting stuck at local minima. This is a simple practitioner's approach to such a non-convex optimization problem [24].

APPENDIX B
OPTIMIZATION OF VSTS USING RATIONAL FUNCTIONS

In this appendix, we construct a family of optimized VSTs for stabilizing Poisson-Gaussian data z. Specifically, this family is parametrized by σ, and assumes $\alpha = 1$. The general case

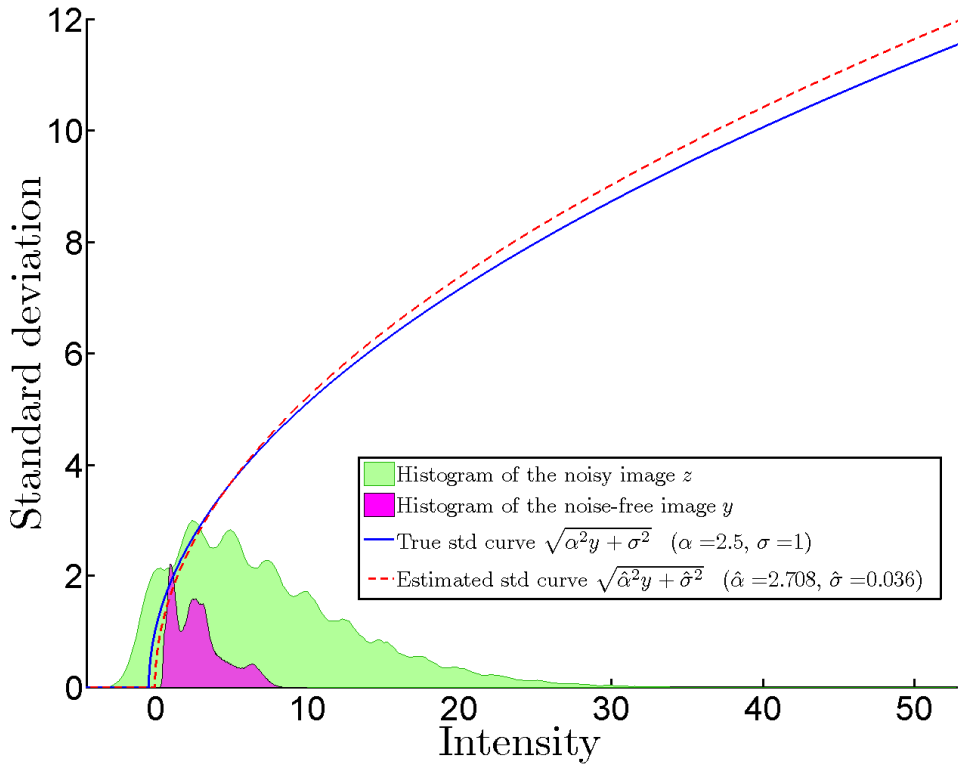


Fig. 8. Estimated and true standard deviation curves for the Armchair image (peak = 10, $\alpha = 2.5$, $\sigma = 1$), overlaid with the histograms of the noisy and noise-free images. For this noise realization, the global minimum of the cost functional (18) is not in the vicinity of the true values $\alpha = 2.5$, $\sigma = 1$; instead, our algorithm converges to $\hat{\alpha} = 2.708$, $\hat{\sigma} = 0.036$, corresponding to a slightly lower cost functional value. Despite this discrepancy, the estimated standard deviation curve remains reasonably accurate over the range of the noise-free image.

$\alpha \neq 1$ can then be handled by simple affine scaling $z \to z/\sigma$, $\sigma \to \sigma/\alpha$ as in [14].

In particular, we consider the ratio of two polynomials inside the typical square root transformation:

$$f_{1,\sigma}(z) = 2\sqrt{\frac{\sum_{i=0}^{N} p_i z^i}{\sum_{i=0}^{M} q_i z^i}} = 2\sqrt{\frac{P(z)}{Q(z)}}, \tag{19}$$

where the polynomials $P(z)$ and $Q(z)$ are to be optimized for each fixed value of σ. If $P(z)/Q(z) < 0$ for some $z \in \mathbb{R}$, we define $f_{1,\sigma}(z) = 0$.

For our purposes, we want to construct the polynomials in such a way that the optimized VST approaches the GAT asymptotically. In this way, the optimized VST always attains good asymptotic stabilization, as can be shown through simple calculus following the results in [25]. In practice, we ensure that

$$\frac{P(z)}{Q(z)} - z - \frac{3}{8} - \sigma^2 \to 0 \text{ as } z \to +\infty \tag{20}$$

at a rate of $\mathcal{O}(z^{-1})$. The constraint (20) immediately requires that at least $M = N-1$ and $p_N = q_M$. Now, let us set $N = 3$ and $M = 2$, so that our rational polynomial $P(z)/Q(z)$ has a cubic numerator and a quadratic denominator. Then, we see that (20) holds if we additionally set $q_1 = p_2 - p_3(3/8 + \sigma^2)$. Hence, we can define

$$f_{1,\sigma}(z) = 2\sqrt{\frac{p_3 z^3 + p_2 z^2 + p_1 z + p_0}{p_3 z^2 + [p_2 - p_3\,(3/8 + \sigma^2)]\, z + 1}}, \tag{21}$$

which depends solely on $\{p_i\}_{i=0}^{3}$.

From now on, we assume the VST $f_{1,\sigma}$ to have the form (21). We optimize $f_{1,\sigma}$ for each fixed σ by minimizing an integral stabilization cost functional:

$$\underset{\{p_i\}_{i=0}^{3}}{\arg\min} \int_{\mathbb{R}^+} \Gamma\left(\text{var}\left\{f_{1,\sigma}(z) \mid y\right\} - 1\right) dy, \tag{22}$$

where the response function Γ is the same as in Appendix A.

Specifically, the integral in (22) penalizes errors in the stabilized variance with respect to the target value 1. Furthermore, during the optimization, we impose a number of additional constraints on $f_{1,\sigma}$ in order to avoid degenerate solutions. First, we ensure that the mapping $h : E\{z \mid y\} \to E\{f_{1,\sigma}(z) \mid y\}$ is strictly increasing, because this in turn guarantees the existence of the exact unbiased inverse VST $E\{f_{1,\sigma}(z) \mid y\} \to E\{z \mid y\}$. Note that the monotonicity of h does not restrict $f_{1,\sigma}$ to be monotone. Besides the monotonicity of h, we ensure that the cubic numerator $P(z)$ has only one root and the quadratic denominator $Q(z)$ is strictly positive.

In practice, the cost functional is minimized through an iterative application of the Nelder-Mead algorithm; the process is analogous to the one presented in Section E of Appendix A. We compute optimized VSTs for 200 values of σ in the range $[0, 4]$, with exponential spacing (i.e., a more dense spacing for low values of σ); for larger values of σ, the GAT is already accurate, so we use it instead of the optimized VST.

Finally, Figure 11 shows an example of the obtained results

for $\sigma = 0.357$, comparing the optimized VST $f_{1,\sigma}$ with the GAT. In Figure 11(a), we see that the optimization procedure enhanced the stabilization accuracy by introducing a large oscillation in $f_{1,\sigma}$. This is further illustrated in Figure 11(b), showing that the stabilized standard deviation achieved with the optimized VST is superior to that achieved with the GAT, with a drastic improvement in the low-intensity region.

REFERENCES

[1] D. Donoho and J. Johnstone, "Ideal spatial adaptation via wavelet shrinkage", *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[2] A. Danielyan and A. Foi, "Noise variance estimation in nonlocal transform domain", *Proc. Int. Workshop on Local and Non-Local Approx. in Image Process., LNLA 2009*, Tuusula, Finland, pp. 41–45, Aug. 2009.

[3] A. Buades, Y. Lou, J. M. Morel, and Z. Tang, "Multi image noise estimation and denoising", preprint, Aug. 2010.

[4] S. I. Olsen, "Noise Variance Estimation in Images", *8th Scandinavian Conference on Image Analysis*, Tromsø, Norway, May 1993.

[5] S. Delpretti, F. Luisier, S. Ramani, T. Blu, and M. Unser, "Multiframe SURE-LET denoising of timelapse fluorescence microscopy images", *Proc. IEEE Int. Symp. on Biomedical Imaging, ISBI 2008*, Paris, France, pp. 149–152, May 2008.

[6] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical Poissonian-Gaussian noise modeling and fitting for single image raw-data", *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1737–1754, Oct. 2008.

[7] P. Gravel, G. Beaudoin, and J. A. De Guise, "A Method for Modeling Noise in Medical Images", *IEEE Trans. Medical Imaging*, vol. 23, no. 10, pp. 1221–1232, Oct. 2004.

[8] B. G. Jeong, B. C. Kim, Y. H. Moon, and I. K. Eom, "Simplified noise model parameter estimation for signal-dependent noise", *Signal Processing*, vol. 96, pp. 266–273, 2014.

[9] A. Jezierska, C. Chaux, J-C. Pesquet, H. Talbot, and G. Engler, "An EM Approach for Time-Variant Poisson-Gaussian Model Parameter Estimation", *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 17–30, Jan. 2014.

[10] G. E. Healey and R. Kondepudy, "Radiometric CCD Camera Calibration and Noise Estimation", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, pp. 267–276, March 1994.

[11] M. Mäkitalo, *Exact Unbiased Inverse of the Anscombe Transformation and its Poisson-Gaussian Generalization*, Ph.D. Thesis, Tampere University of Technology, Finland, March 2013. http://urn.fi/URN:ISBN:978-952-15-3039-5

[12] J. L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis*, Cambridge University Press, Cambridge, 1998.

[13] F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data", *Biometrika*, vol. 35, no. 3/4, pp. 246–254, Dec. 1948. http://dx.doi.org/10.1093/biomet/35.3-4.246

[14] M. Mäkitalo and A. Foi, "Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise", *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 91–103, Jan. 2013. http://dx.doi.org/10.1109/TIP.2012.2202675

[15] J. Salmon, Z. Harmany, C-A. Deledalle, and R. Willett, "Poisson Noise Reduction with Non-local PCA", *J. Math. Imaging Vis.*, vol. 48, no. 2, pp. 279–294, Feb. 2014.

[16] R. Giryes and M. Elad, "Sparsity Based Poisson Denoising with Dictionary Learning", 2014, preprint available at http://arxiv.org/pdf/1309.4306v2.pdf.

[17] J. H. Curtiss, "On transformations used in the analysis of variance", *The Annals of Mathematical Statistics*, vol. 14, no. 2, pp. 107–122, June 1943.

[18] B. Efron, "Transformation theory: How normal is a family of distributions?", *The Annals of Statistics*, vol. 10, no. 2, pp. 323–339, 1982.

[19] A. Foi, "Optimization of variance-stabilizing transformations", 2009, preprint available at http://www.cs.tut.fi/~foi/.

[20] A. Foi, "Noise Estimation and Removal in MR Imaging: the Variance-Stabilization Approach", *Proc. 2011 IEEE International Symposium on Biomedical Imaging, ISBI 2011*, Chicago, USA, pp. 1809–1814, Apr. 2011.

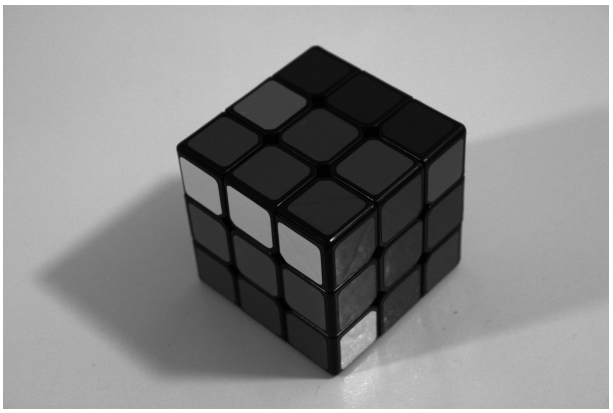
[21] X. Jin, *Poisson Approximation to Image Sensor Noise*, M.Sc. Thesis, University of Dayton, Ohio, USA, Dec. 2010.

[22] X. Jin, Z. Xu, and K. Hirakawa, "Noise Parameter Estimation for Poisson Corrupted Images Using Variance Stabilization Transforms", *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1329–1339, March 2014.

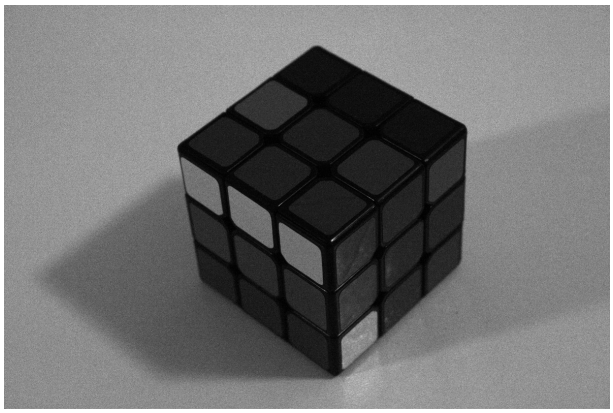
[23] S. Pyatykh and J. Hesser, "Image Sensor Noise Parameter Estimation by Variance Stabilization and Normality Assessment", *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3990–3998, Sept. 2014.

[24] A. R. Conn, K. Scheinberg, and L. N. Vicente, *Introduction to derivative-free optimization*, MPS-SIAM Series on Optimization, vol. 8, 2009.

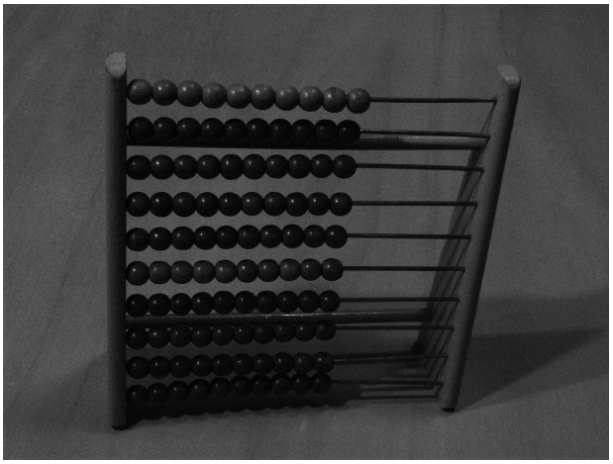
[25] S. K. Bar-Lev and P. Enis, "On the construction of classes of variance stabilizing transformations", *Statistics & Probability Letters*, vol. 10, pp. 95–100, July 1990.



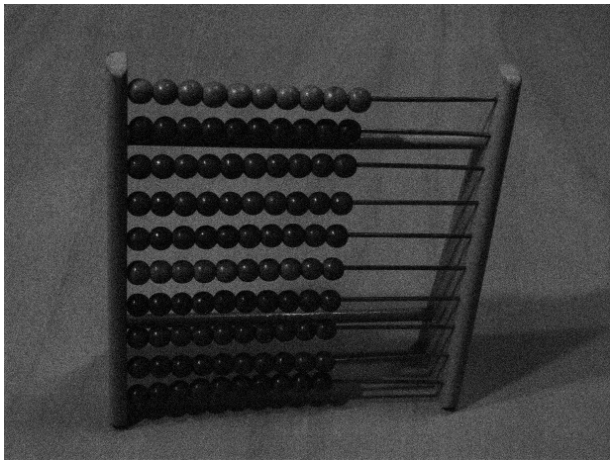
(a) Cube, ISO 200



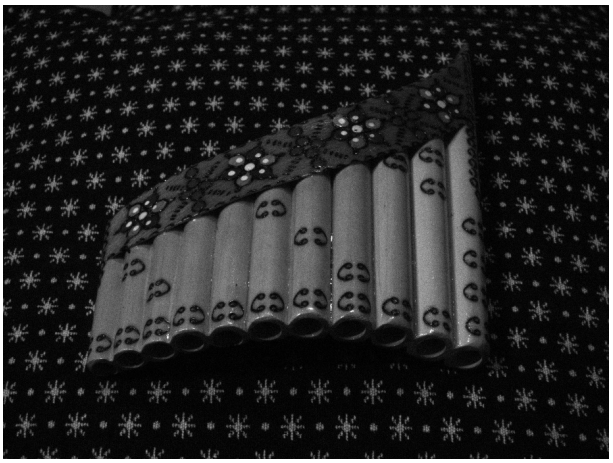
(b) Cube, ISO 3200



(c) Abacus, ISO 200



(d) Abacus, ISO 800



(e) Flute, ISO 400

(f) Flute, ISO 1600

Fig. 9. The test images "Cube" (2385 × 1589, Canon EOS 500D), "Abacus" (1224 × 922, Fujifilm FinePix S9600) and "Flute" (1842 × 1380, Canon PowerShot S90) used in the experiments of Section VI-B. Each image is the blue colour channel extracted from raw Bayer data.

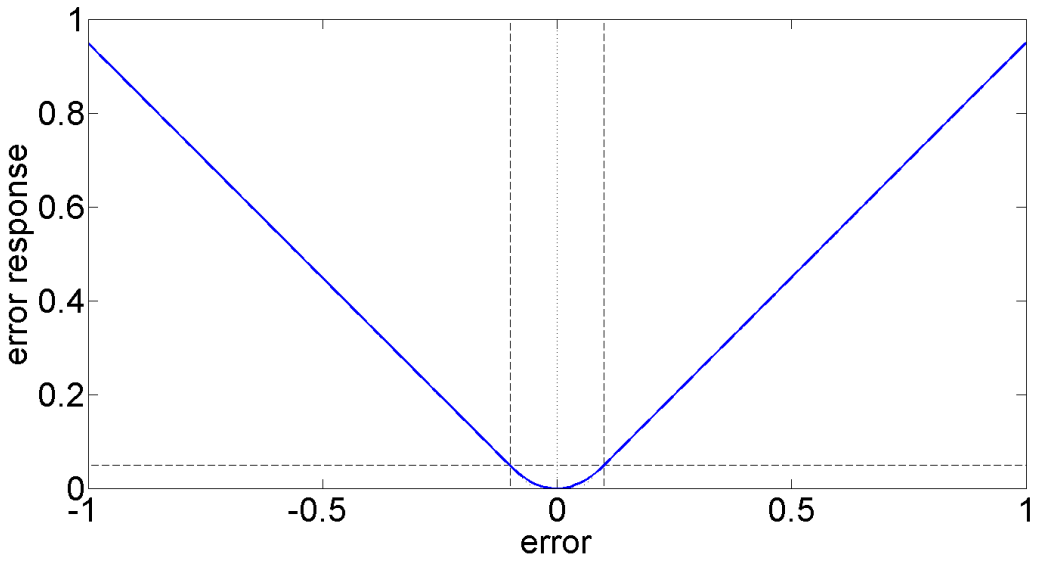
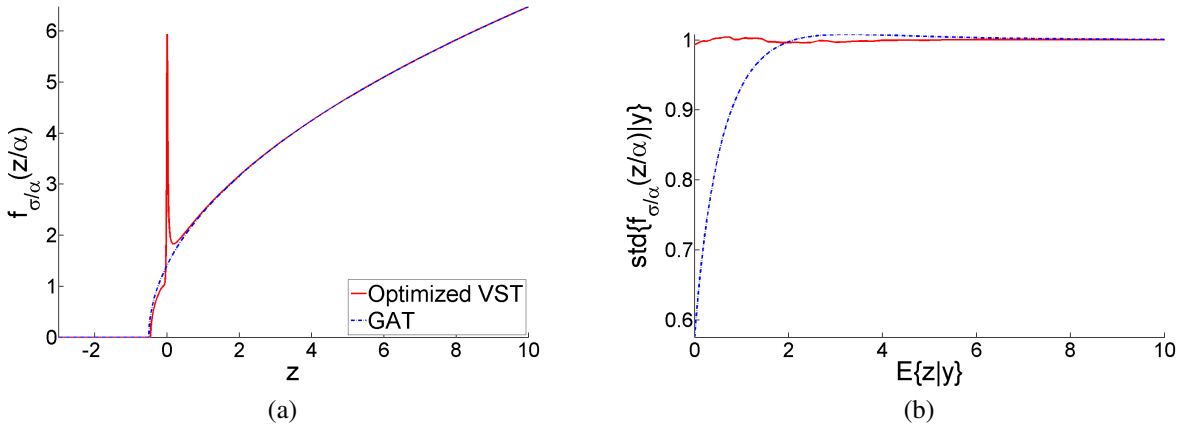


Fig. 10. Nonlinear response function Γ used in modulating errors in the stabilized variance, when optimizing the cost functionals (18) and (22).



(a) (b)

Fig. 11. (a) Optimized rational VST $f_{1,\sigma}(z)$ and the GAT, for $\sigma = 0.357$ ($\alpha = 1$). (b) The stabilized standard deviation obtained with the VSTs in (a).