Clipped noisy images:
Heteroskedastic modeling and practical denoising

Alessandro Foi

Department of Signal Processing, Tampere University of Technology

P.O. Box 553, 33101, Tampere, FINLAND

Abstract

We study the denoising of signals from clipped noisy observations, such as digital images of an under- or over-exposed scene. From a precise mathematical formulation and analysis of the problem, we derive a set of homomorphic transformations that enable the use of existing denoising algorithms for non-clipped data (including arbitrary denoising filters for additive independent and identically distributed (i.i.d.) Gaussian noise). Our results have general applicability and can be "plugged" into current filtering implementations, to enable a more accurate and better processing of clipped data. Experiments with synthetic images and with real raw data from charge-coupled device (CCD) sensor show the feasibility and accuracy of the approach.

Key words: denoising; noise modeling; signal-dependent noise; heteroskedasticity; raw data; overexposure; underexposure; clipping; censoring; homomorphic transformations; variance stabilization.

1. Introduction

We consider the problem of recovering a signal from its noisy observations that have been clipped, i.e. observations whose range is limited by a lower and by an upper bound. A prominent example are images acquired with under- or over-exposed areas and particularly the raw-data images acquired by a digital imaging sensor (e.g., a charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) imager). As a continuation of our previous work [13] on modeling and automatic noise parameter estimation for clipped images corrupted by Poissonian-Gaussian noise, in this paper we provide a complete methodology for the accurate denoising of such images. The goal of this paper is general and pragmatic: rather than developing a novel denoising algorithm specially designed for clipped noisy images, we derive a set of homomorphic transformations that enable the use of existing denoising algorithms for non-clipped data (including arbitrary denoising filters for additive independent and identically distributed (i.i.d.) Gaussian noise). Thus, our results can be "plugged" into current filtering implementations, allowing a more accurate and better processing of clipped data.

An interesting and important feature of clipped noisy signals is that they carry usable information about the signals' values outside the acquisition range. By properly processing and denoising these signals, it is possible to extract this information

and then produce images with a dynamic range that exceeds that of the clipped signal. Thus, we can partly overcome the problem of saturation of the imaging device when a scene is captured with overexposed regions. Furthermore, our procedure can be utilized for the accurate linearization of the sensor's output with respect to the expectation of its input. Hence, we provide a practical solution which has direct application for a number of image processing problems, including deblurring/deconvolution, single-image and multiple-image high dynamic range imaging [23],[6], sensor characterization, and digital forensics [3].

This paper extends our preliminary results presented in [8] with a complete and detailed mathematical analysis of the properties of clipped signal-dependent noise, establishing a rigorous framework for processing data subject to this kind of degradation. Clipped heteroskedastic normal variables are used as the principal tool for carrying out the mathematical study.

The paper is organized as follows. In the next section we introduce the general observation models for images corrupted by signal-dependent noise; the models are given for the observations both before and after the clipping. Next, we give the relations and transformations which exist between the expectations and standard deviations of the clipped and the non-clipped random variables. The core of our contributions is given in Section 3, where we consider the denoising of a clipped signal using generic filters. In particular, we discuss the use of conventional filters designed for additive i.i.d. Gaussian noise and derive specific homomorphic transformations to stabilize the variance of the clipped observations, to compensate the bias due to the clipped distribution in the variance-stabilized domain, and finally to compensate the estimation bias between the denoised clipped variables and the non-clipped true variables. Experi-

ments with synthetic as well as with real raw data from a commercial CCD camera are presented in Section 4, showing the feasibility and accuracy of the developed approach. Relevant conclusions are given in the last section while mathematical details and proofs are included in the Appendix.

2. Preliminaries

2.1. Signal-dependent noise model

Let $z(x)$, $x \in X \subset \mathbb{Z}^2$, be noisy observations with the expectations $E\{z(x)\} = y(x) \in Y \subseteq \mathbb{R}$, where the errors (noise) $\eta(x) = z(x) - y(x)$ are random independent and the standard deviation of these observations is modeled by std$\{\eta(x)\} = \sigma(y(x))$, $\sigma : Y \to \mathbb{R}^+$ being a deterministic function of y. Equivalently, we consider the generic signal-dependent noise observation model of the form

$$z(x) = y(x) + \sigma(y(x))\,\xi(x), \quad x \in X, \qquad (1)$$

where $\xi : X \to \mathbb{R}$ is zero-mean independent random noise with standard deviation equal to 1. Clearly, $\eta(x) = \sigma(y(x))\,\xi(x)$. Although std$\{\xi(x)\} = 1$ for all $x \in X$, the probability distribution of ξ can be different at different samples (i.e. $\xi(x_1) \nsim \xi(x_2)$ if $x_1 \neq x_2$). However, to allow a simple mathematical formulation, we approximate ξ as a standard normal random variable, $\xi \sim \mathcal{N}(0, 1)$, which is formally equivalent to considering a heteroskedastic Gaussian noise $\eta(x) \sim \mathcal{N}(0, \sigma^2(y(x)))$ having signal-dependent variance. As discussed in [13], this can be a suitable approximation when dealing with the noise in the raw data from digital imaging sensors, for which the typical form of the function σ is

$$\sigma^2(y(x)) = ay(x) + b, \qquad y \geq -\frac{b}{a}, \qquad (2)$$

with the constants $a \in \mathbb{R}^+$ and $b \in \mathbb{R}$ depending on the sensor's specific characteristics and on the particular acquisition settings (e.g., analog gain or ISO value, temperature, pedestal, etc.). Obviously, the trivial i.i.d. (signal-independent and homoskedastic) additive Gaussian noise case, where std$\{\eta(x)\} \equiv \sqrt{b}$, is obtained when $a = 0$, for which the lower bound $-\frac{b}{a}$ in (2) becomes $-\infty$ and, thus, $y \in \mathbb{R}$.

2.2. Clipping

In practice, the range of the acquisition system is always limited. Without loss of generality, we consider data given with respect to the range $[0, 1]$, where the extremes correspond to the maximum and minimum allowed pixel values for the considered noisy image (e.g., raw data) format. Values above or below these bounds are replaced by the bounds themselves. This corresponds to the behavior of digital imaging sensors in the case of over- or under-exposure. Thus, we define the *clipped* observations $\tilde{z}$ as

$$\tilde{z}(x) = \max\{0, \min\{z(x), 1\}\}, \qquad x \in X, \qquad (3)$$

where z is given by the (non clipped) signal-dependent noise model (1). In what follows, we use the tilde to denote variables directly related to clipped observations. The corresponding noise model for the clipped observations (3) is then

$$\tilde{z}(x) = \tilde{y}(x) + \tilde{\sigma}(\tilde{y}(x))\,\tilde{\xi}(x), \qquad (4)$$

where $\tilde{y}(x) = E\{\tilde{z}(x)\} \in \tilde{Y}$, $\tilde{\sigma} : \tilde{Y} \to \mathbb{R}^+$ gives the standard-deviation of the clipped noisy data as a function of their expectation, i.e. $\tilde{\sigma}(\tilde{y}(x)) = \text{std}\{\tilde{z}(x)\}$, and $E\{\tilde{\xi}(x)\} = 0$, var$\{\tilde{\xi}(x)\} = 1$. Because of clipping, in general, we have that

$$\tilde{y}(x) = E\{\tilde{z}(x)\} \neq E\{z(x)\} = y(x), \qquad (5)$$
$$\tilde{\sigma}(\tilde{y}(x)) = \text{std}\{\tilde{z}(x)\} \neq \text{std}\{z(x)\} = \sigma(y(x)),$$

and $Y \neq \tilde{Y} \subseteq [0, 1]$. Rewriting (4) as

$$\tilde{z}(x) = y(x) + \left[\tilde{y}(x) - y(x) + \tilde{\sigma}(\tilde{y}(x))\,\tilde{\xi}(x)\right],$$

we can see that, with respect to the true signal y, the clipped observations $\tilde{z}$ are corrupted by an error (the term in square brackets) which has non-zero mean. Observe also that, even though var$\{\tilde{\xi}(x)\} = $ var$\{\xi(x)\} = 1$, the distributions of ξ and $\tilde{\xi}$ are different. In particular, assuming $\xi(x) \sim \mathcal{N}(0, 1)$, we have that $\tilde{\xi}(x)$ follows a doubly censored normal distribution [4] supported on $\left[\frac{-\tilde{y}}{\tilde{\sigma}(\tilde{y})}, \frac{1-\tilde{y}}{\tilde{\sigma}(\tilde{y})}\right]$.

2.3. Conventions

To be able to define $\tilde{\sigma}$ as a function of $\tilde{y}$ (and not as a multi-valued mapping) some mild restrictions need to be imposed on σ or Y. For instance, as proved in the Appendix, it suffices that σ is smooth and concave on $Y \ni \frac{1}{2}$, as from these hypotheses follows that $y \longmapsto \tilde{y}$ is injective. All standard-deviation functions which appear in imaging applications satisfy such basic conditions and thus, in what follows, we will always assume that $\tilde{\sigma}$ is well-defined as a function of $\tilde{y}$[1]. Not to overload the notation, unless there is risk of confusion we will omit the argument x and similarly we will not explicitly write the conditioning on y of the various expectations, standard-deviations, and variances (thus, for example, we simply write $E\{z\}$ instead of $E\{z|y\}$ or $E\{z(x)\}$).

Figure 1 gives an example of the curves $(y, \sigma(y))$ and $(\tilde{y}, \tilde{\sigma}(\tilde{y}))$, for $\sigma(y) = \sqrt{0.01y + 0.04^2}$. We emphasize that each curve is drawn in the corresponding expectation/standard-deviation Cartesian plane (i.e. we plot the "non-clipped" $\sigma(y)$ against the y,σ axes and the "clipped" $\tilde{\sigma}(\tilde{y})$ against the $\tilde{y},\tilde{\sigma}$ axes). The figure illustrates the correspondence between points on the two curves given by the equations (3)-(5).

---

[1] Although in this paper we do not investigate denoising in the case of a multivalued $\tilde{\sigma}$, we wish to note that it is technically possible to deal with such cases by decomposing X and Y into sub-domains X_k and Y_k, $k = 1, \ldots, K$, such that to each pair of these corresponds a single-valued $\tilde{\sigma}_k$. Examples of multivalued $\tilde{\sigma}$ are illustrated in the Appendix.
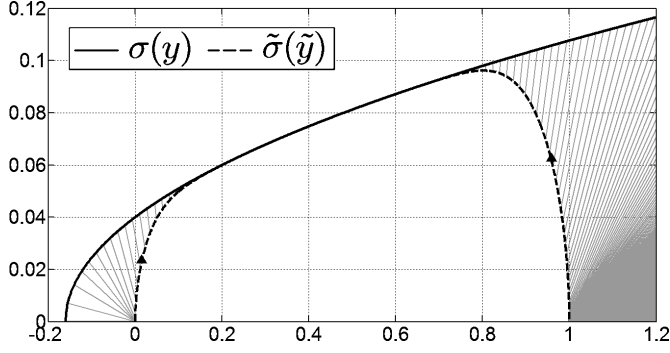
Figure 1: Standard-deviation function $\sigma(y) = \sqrt{0.01y + 0.04^2}$ (solid line) and the corresponding standard-deviation curve $\tilde{\sigma}(\tilde{y})$ (dashed line). The gray segments illustrate the mapping $\sigma(y) \longmapsto \tilde{\sigma}(\tilde{y})$. The small black triangles ▲ indicate points $(\tilde{y}, \tilde{\sigma}(\tilde{y}))$ which correspond to $y = 0$ and $y = 1$.

2.4. Expectations and standard deviations of clipped variables and their transformations

A crucial point when working with clipped noisy signals is to understand how the variables and functions of the observation model (1) relate to those of the clipped observations' model (4). In particular, it is important to compute the functions $\tilde{y}$ and $\tilde{\sigma}$ given σ and y, and vice versa. The probability density function (p.d.f.) of the unobserved non-clipped noisy data $z \sim \mathcal{N}(y, \sigma^2(y))$ is simply $\frac{1}{\sigma(y)}\phi\left(\frac{\zeta - y}{\sigma(y)}\right)$, whereas the clipped $\tilde{z} = \max\{0, \min\{z, 1\}\}$ is distributed according to a doubly censored Gaussian distribution having a generalized p.d.f. $\wp_{\tilde{z}}$ of the form

$$\wp_{\tilde{z}}(\zeta) = \Phi\left(\tfrac{-y}{\sigma(y)}\right)\delta_0(\zeta) + \tfrac{1}{\sigma(y)}\phi\left(\tfrac{\zeta - y}{\sigma(y)}\right)\chi_{[0,1]} + \Phi\left(\tfrac{y-1}{\sigma(y)}\right)\delta_0(1 - \zeta),$$

(6)

where $\chi_{[0,1]}$ denotes the characteristic function of the interval $[0, 1]$ and δ_0 is the Dirac delta impulse at 0. Here ϕ and Φ are the p.d.f. and cumulative distribution function (c.d.f.) of the standard normal $\mathcal{N}(0, 1)$, respectively. The first and last addends in (6) correspond to the probabilities of clipping from below and from above (under- or over-exposure). Very tedious calculations provide the following exact expressions of the expectation and variance of $\tilde{z}$:

$$E\{\tilde{z}\} = \tilde{y} = \Phi\left(\tfrac{y}{\sigma(y)}\right)y - \Phi\left(\tfrac{y-1}{\sigma(y)}\right)(y - 1) +$$
$$+ \sigma(y)\phi\left(\tfrac{y}{\sigma(y)}\right) - \sigma(y)\phi\left(\tfrac{y-1}{\sigma(y)}\right),$$

(7)

$$\operatorname{var}\{\tilde{z}\} = \tilde{\sigma}^2(\tilde{y}) = \Phi\left(\tfrac{y}{\sigma(y)}\right)\left(y^2 - 2\tilde{y}y + \sigma^2(y)\right) +$$
$$+ \tilde{y}^2 - \Phi\left(\tfrac{y-1}{\sigma(y)}\right)\left(y^2 - 2\tilde{y}y + 2\tilde{y} + \sigma^2(y) - 1\right) +$$
$$+ \sigma(y)\phi\left(\tfrac{y-1}{\sigma(y)}\right)(2\tilde{y} - y - 1) - \sigma(y)\phi\left(\tfrac{y}{\sigma(y)}\right)(2\tilde{y} - y).$$

(8)

For a given function σ, these expressions explicitly define two mappings $\mathcal{A}_\sigma : Y \to \tilde{Y}$ and $\mathcal{B}_\sigma : \mathbb{R}^+ \to \mathbb{R}^+$ as follows:

$$\mathcal{A}_\sigma : y \longmapsto \tilde{y} = \mathcal{A}_\sigma(y),$$

(9)

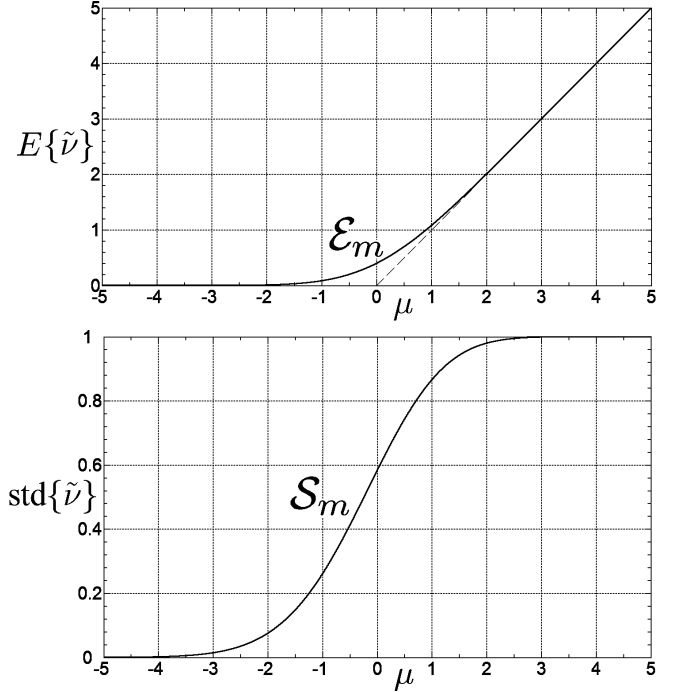$$\mathcal{B}_\sigma : \sigma(y) \longmapsto \tilde{\sigma}(\tilde{y}) = \mathcal{B}_\sigma(y).$$

(10)

Figure 2: Expectation $E\{\tilde{\nu}\}$ and standard deviation $\operatorname{std}\{\tilde{\nu}\}$ of the clipped $\tilde{\nu} = \max\{0, \nu\}$ as functions $\mathcal{E}_m$ and $\mathcal{S}_m$ of μ, where $\mu = E\{\nu\}$ and $\nu \sim \mathcal{N}(\mu, 1)$.

In particular, (9) and (10) yield a transformation that brings the standard deviation curve $(y, \sigma(y))$ to its clipped counterpart $(\tilde{y}, \tilde{\sigma}(\tilde{y}))$.

The inverse mapping of $\mathcal{A}_\sigma$ will be denoted as $\mathcal{C}_\sigma : \tilde{Y} \to Y$,

$$\mathcal{C}_\sigma : \tilde{y} \longmapsto y = \mathcal{C}_\sigma(\tilde{y}).$$

(11)

Although the expressions (7) and (8) can be eventually useful for a numerical implementation, they are cumbersome and cannot be easily manipulated for further analysis.

2.5. Approximation by singly-clipped variables

As in [13], we can simplify the analysis by assuming that the two clippings, the one from below ($z < 0$, $\tilde{z} = 0$) and the one from above ($z > 1$, $\tilde{z} = 1$), are not mixed by the randomness of the noise. It means that, for a fixed y, at most one of the impulses in the p.d.f. (6) has mass appreciably larger than 0. In practice, this assumption is always verified, except for those extreme situations where the noise is dramatically strong for relatively small signal values (e.g., $\sigma(y) \gg 0.2$ for $y \in [0, 1]$).

Let $\nu \sim \mathcal{N}(\mu, 1)$ be a normally distributed random variable with mean $E\{\nu\} = \mu$ and unitary variance, and let $\tilde{\nu} = \max\{0, \nu\}$. It can be easily shown (see, e.g., [14] Chapter 20 or [15]) that the expectation $E\{\tilde{\nu}\}$ and the variance $\operatorname{var}\{\tilde{\nu}\}$ of the clipped (from below) $\tilde{\nu}$ are

$$E\{\tilde{\nu}\} = \mathcal{E}_m(\mu) = \Phi(\mu)\mu + \phi(\mu),$$

(12)

$$\operatorname{var}\{\tilde{\nu}\} = \mathcal{S}_m^2(\mu) = \Phi(\mu) + \mathcal{E}_m(\mu)\mu - \mathcal{E}_m^2(\mu) =$$

(13)

$$= \Phi(\mu) + \phi(\mu)\mu - \phi^2(\mu) +$$
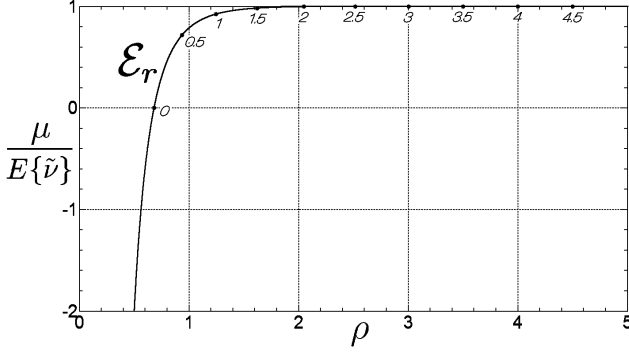$$+ \Phi(\mu)\mu(\mu - \Phi(\mu)\mu - 2\phi(\mu)).$$

3

Figure 3: The function $\mathcal{E}_r = \frac{\mu}{E\{\tilde{\nu}\}}$ of $\rho = \frac{E\{\tilde{\nu}\}}{\text{std}\{\tilde{\nu}\}}$. The small italic numbers indicate the corresponding value of μ.

The plots of the expectation $E\{\tilde{\nu}\} = \mathcal{E}_m(\mu)$ and of the standard deviation $\text{std}\{\tilde{\nu}\} = \mathcal{S}_m(\mu)$ are shown, as functions of μ, in Figure 2.

Exploiting these functions, the direct and inverse transformations which link σ and y to $\tilde{y}$ and $\tilde{\sigma}$ can be expressed in the following compact forms [13].

2.5.1. Direct transformation (obtain $\tilde{y}$ and $\tilde{\sigma}$ from y and σ)

Provided that $y = E\{z\}$ and $\sigma(y) = \text{std}\{z\}$ from the basic model (1) are known, the expectation $\tilde{y} = E\{\tilde{z}\}$ and the standard deviation $\tilde{\sigma}(\tilde{y}) = \text{std}\{\tilde{z}\}$ from the observation model (4) are obtained as

$$\tilde{y} = \mathcal{A}_\sigma(y) \approx \mathcal{A}(y, \sigma(y)) =$$
$$= \sigma(y)\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right) + 1 - y - \sigma(y)\mathcal{E}_m\left(\frac{1-y}{\sigma(y)}\right), \quad (14)$$

$$\tilde{\sigma}(\tilde{y}) = \mathcal{B}_\sigma(y) \approx \mathcal{B}(y, \sigma(y)) = \sigma(y)\mathcal{S}_m\left(\frac{y}{\sigma(y)}\right)\mathcal{S}_m\left(\frac{1-y}{\sigma(y)}\right). \quad (15)$$

Compared to the exact (9) and (10), the approximate equations (14) and (15) provide a more intuitive description of the transformations that bring the standard-deviation curve $(y, \sigma(y))$ to its clipped counterpart $(\tilde{y}, \tilde{\sigma}(\tilde{y}))$. For instance, provided y is sufficiently smaller than 1, by observing Figure 2 it is easy to realize that $\mathcal{E}_m\left(\frac{1-y}{\sigma(y)}\right)$ and $\mathcal{S}_m\left(\frac{1-y}{\sigma(y)}\right)$ can be substituted by $\frac{1-y}{\sigma(y)}$ and 1, respectively (the substitution is asymptotically exact). Thus, for describing the clipping from below, (14) and (15) can be reduced to, respectively, $\sigma(y)\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)$ and $\sigma(y)\mathcal{S}_m\left(\frac{y}{\sigma(y)}\right)$, which allows to construct the graph of $(\tilde{y}, \tilde{\sigma}(\tilde{y}))$ in the vicinity of $(0,0)$ by simple manipulations of the graphs of $\mathcal{E}_m$ and $\mathcal{S}_m$.

2.5.2. Inverse transformation (obtain y from $\tilde{\sigma}$ and $\tilde{y}$)

The approximation of (11), for calculating the non-clipped y (1) from the clipped $\tilde{y}$ and $\tilde{\sigma}(\tilde{y})$, can be given as

$$y = \mathcal{C}_\sigma(\tilde{y}) \approx \mathcal{C}(\tilde{y}, \tilde{\sigma}(\tilde{y})) =$$
$$= \tilde{y}\mathcal{E}_r\left(\frac{\tilde{y}}{\tilde{\sigma}(\tilde{y})}\right) - \tilde{y} + 1 - (1-\tilde{y})\mathcal{E}_r\left(\frac{1-\tilde{y}}{\tilde{\sigma}(\tilde{y})}\right), \quad (16)$$

where $\mathcal{E}_r$ is defined implicitly as function of $\rho = \frac{\mathcal{E}_m(\mu)}{\mathcal{S}_m(\mu)} = \frac{E\{\tilde{\nu}\}}{\text{std}\{\tilde{\nu}\}}$ by $\mathcal{E}_r(\rho) = \frac{\mu}{\mathcal{E}_m(\mu)}$. Figure 3 shows the plot of $\mathcal{E}_r$ as function of ρ.

We refer the reader to [13] for a detailed study of these approximate transformations, including their indirect polynomial interpolation.

3. Denoising clipped signals

A generic denoising algorithm can be modeled as an operator whose output is an estimate of the expectation of the noisy input. Formally, let $\mathbf{D}: \mathbb{R}^{|X|} \to \mathbb{R}^{|X|}$ be the denoising operator, then

$$\mathbf{D}(\tilde{z}) = \widehat{E\{\tilde{z}\}} \approx E\{\tilde{z}\} = \tilde{y}. \quad (17)$$

It means that when we denoise $\tilde{z}$, as output we do not get an estimate of y, but rather an estimate of $\tilde{y}$. Analogously, we may say that $\mathbf{D}(\tilde{z})$ is a biased estimator of y, in the sense that $E\{\mathbf{D}(\tilde{z})\} \approx \tilde{y} \neq y$; in such a case, the bias error can be expressed as $\tilde{y} - y$.

For the sake of simplification, we may even assume that $\mathbf{D}$ is an ideal operator that can always accurately recover the expectation of a given (non clipped) input image corrupted by additive i.i.d. Gaussian noise. This is an appealing assumption, but there remain a number of important issues that need to be considered when we apply $\mathbf{D}$ for the denoising of clipped data.

Mainly, the clipping noise is signal-dependent and as such it requires special care: first, when the unknown noise parameters are estimated from the image; and second, when we aim at suppressing the noise. Furthermore, the noise samples follow essentially non-Gaussian and asymmetrical distributions. Finally, the estimation bias $y - \tilde{y}$ caused by the clipping needs to be compensated.

3.1. Noise estimation

In [13], we proposed an algorithm for the automatic estimation of the parameters of clipped signal-dependent noise models from a single noisy image. The algorithm utilizes a one-level wavelet decomposition of the noisy image and the transformations of Section 2.4 for the maximum-likelihood estimation of the noise parameters and hence of the curves $\tilde{\sigma}$ and σ. If the parameters of the noise model are not known in advance, this algorithm can be used as the first step in the processing of clipped noisy images[2].

3.2. Noise removal

In general, when dealing with signal-dependent noise, we can either use filters specifically designed for such heteroskedastic noise (e.g., [12], [10]), or we can exploit a variance-stabilizing homomorphic transformation (e.g., [1], [21]) and then apply practically any filter for homoskedastic noise on the transformed noisy signal. Here, we discuss both alternative approaches and concentrate our attention on *transform-domain* filters designed for either heteroskedastic or homoskedastic *Gaussian* noise, as these are in practice the most powerful and most widely used ones (see, e.g., [19] and references therein). In what follows, we refer to the two approaches

---

[2]To simplify notation, here we use the symbols σ and $\tilde{\sigma}$ for the true as well as for the estimated curves. In [13] the latter are denoted by $\hat{\sigma}_{\text{fit}}$ and $\hat{\tilde{\sigma}}_{\text{fit}}$.
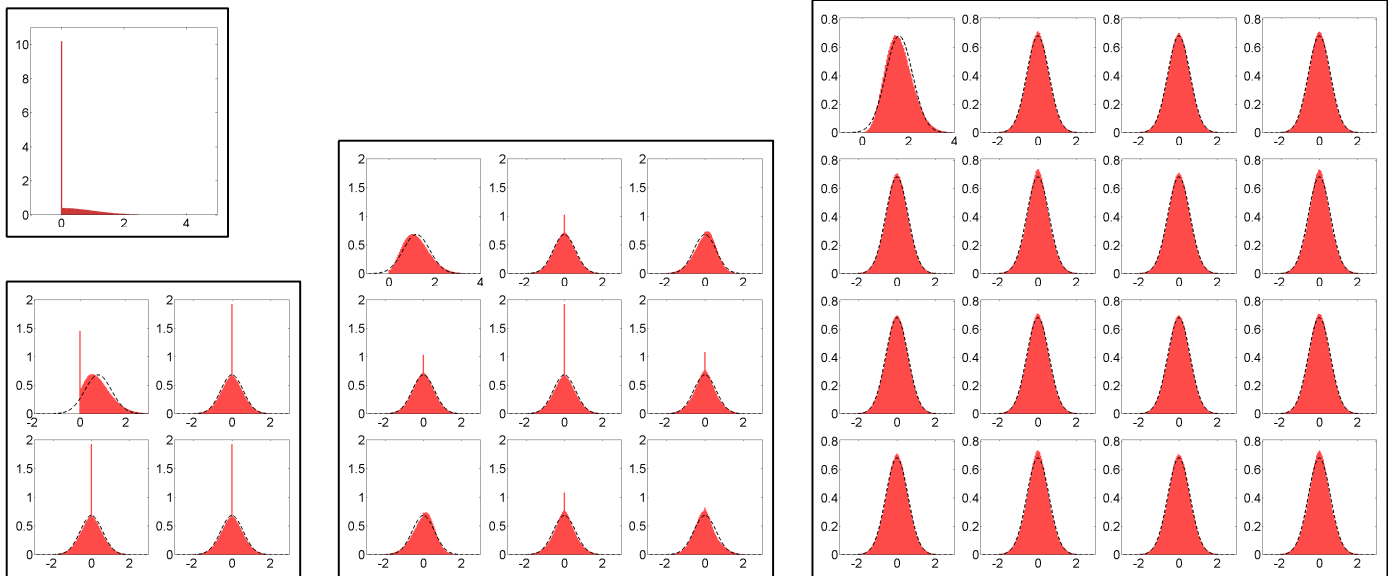
Figure 4: Normalized histograms of an i.i.d. clipped noise field $\tilde{\nu}$, $\tilde{\nu}(\cdot) = \max\{0, \nu(\cdot)\}$ (where $\nu(\cdot) \sim \mathcal{N}(0, 1)$) and of each coefficient of its 2-D DCT transforms of size $n \times n$, n=2,3,4. The dashed lines show the probability distribution functions of $\mathcal{N}(n\mathcal{E}_m(0), \mathcal{S}_m(0))$ (DC terms) and $\mathcal{N}(0, \mathcal{S}_m(0))$ (AC terms). As the size of the transform increases, the distribution of the coefficients converges very quickly to these Gaussian bells. The histograms have equispaced bins of width 0.05 and are computed from $2 \cdot 10^7$ independent realizations of $\tilde{\nu}(\cdot)$.

as *heteroskedastic filtering* and *homoskedastic filtering*, respectively, and denote the corresponding denoising filters as $\mathbf{D}_{\mathsf{he}}$ and $\mathbf{D}_{\mathsf{ho}}$. Due to the essential non-Gaussianity of the clipped noise, the latter approach requires additional care.

3.2.1. Heteroskedastic filtering

Here, we assume that the denoising filter $\mathbf{D}_{\mathsf{he}}$ can handle the signal-dependent standard deviation $\tilde{\sigma}$. However, the noise distributions of different samples are all different and non-Gaussian. Fortunately, this does not constitute a serious problem for transform-based algorithms because, in practice, the transform coefficients of clipped data have distributions which are nearly Gaussians. In particular, let us consider a clipped noise field $\tilde{\nu} = \max\{0, \nu\}$, where $\nu(\cdot) \sim \mathcal{N}(\mu, 1)$. For any basis element ψ of an orthonormal transform, the distribution of the transform coefficient $\langle \psi, \tilde{\nu} \rangle$ can be approximated by $\mathcal{N}\big(\sum_j \psi(j)\, \mathcal{E}_m(\mu), \mathcal{S}_m(\mu)\big)$. Figure 4 illustrates this approximation when $\mu = 0$ in the case of the 2-D discrete cosine transform (DCT) transform of size $n \times n$, n=2,3,4. As can be seen in the figure, the Gaussianization of the AC coefficients is faster. This is due to the positive and negative signs in the samples of the corresponding basis elements. The magnitude of the impulse at 0 (visible in some the histograms of smaller size DCT) decays exponentially with the number K of nonzero samples of the basis element, $\Phi(-\mu)^K$ [13]. For example, the nine basis elements $\{\psi_{i,j}\}_{i,j=1,2,3}$ of the 3×3 DCT have $K(i, j)$=[3 2 3]^T[3 2 3] non-zero samples, respectively, and for μ=0 we have $\Phi(-\mu)=\frac{1}{2}$; it means that the magnitude of the impulse at 0 in the histogram for $i = j = 2$ (K=4) is approximately 2^{6-4}=4 times larger than those for $i + j = 3, 5$ (K=6), 2^{9-4}=32 times larger than those for $i, j \in \{1, 3\}$ (K=9), and 2^{4-1}=8 times smaller than the impulse in the histogram of $\tilde{\nu}$ (K=1). Each basis element of the 4×4 DCT has K=16 non-

zero samples, hence the magnitude of the impulse in their histograms is 2^{16-1}=32768 times smaller than the impulse in the histogram of $\tilde{\nu}$ and thus practically negligible.

Of course, because of the central-limit theorem, one always gets some sort of "Gaussianization", regardless of the distribution of the original samples. In the case of clipped normal samples, the convergence is remarkably fast, as can be seen in the figure.

We should also remark that, in the transform domain, the clipped noise is no longer independent and that instead some correlation exists in the noise of different coefficients, even when the transform is orthogonal. Nevertheless, the near totality of modern transform-based filters already employs over-complete and highly redundant decompositions (for which the noise in spectral domain is necessarily correlated) and the correlation due to heteroskedasticity is in practice a secondary issue when compared to the correlation due to the redundancy of the transform.

Therefore, we conclude that a transform-domain filter $\mathbf{D}_{\mathsf{he}}$ can indeed be applied successfully on clipped data and that the approximation (17) holds.

3.2.2. Homoskedastic filtering

To improve the readability, we assume here that $\tilde{Y} \supseteq (0, 1)$, confining the considerations and technicalities about the case $\tilde{Y} \supsetneq (0, 1)$ to the Appendix.

A. Variance-stabilizing transformation. A specifically designed homomorphic transformation $f : \tilde{Y} \to \mathbb{R}$ can be utilized in order to stabilize the signal-dependent standard deviation of the clipped observations $\tilde{z}$ to a desired constant $c \in \mathbb{R}^+$ and thus apply a denoising algorithm $\mathbf{D}_{\mathsf{ho}}$ for homoskedastic noise on $f(\tilde{z})$. Following the simple approach which appears

in many works (e.g., [21]), we use a first-order Taylor expansion for a monotonically increasing f, from which follows $\mathrm{std}\{f(\tilde{z})\} \approx f'(E\{\tilde{z}\})\,\mathrm{std}\{\tilde{z}\} = f'(\tilde{y})\,\tilde{\sigma}(\tilde{y})$, and then solve for $\mathrm{std}\{f(\tilde{z})\} \equiv c$. Up to an arbitrary additive constant, this yields

$$f(t) = \int_{t_0}^{t} \frac{c}{\tilde{\sigma}(\tilde{y})}\,d\tilde{y}, \quad t, t_0 \in [0, 1], \tag{18}$$

i.e. the c-stabilizing homomorphic transformation is an indefinite integral of $\frac{c}{\tilde{\sigma}(\tilde{y})}$, with the integration with respect to the argument $\tilde{y}$. As shown in the Appendix, under mild and rather generic assumptions on $\sigma(y)$, the integral (18) is convergent when $\tilde{\sigma}(\tilde{y}) \to 0$ for $\tilde{y} \to 0^+$ or $\tilde{y} \to 1^-$, which implies the boundedness of f. Therefore, in what follows, we will always assume that f is a bounded, strictly increasing (hence invertible) function, and that $t_0 = 0$, $f(0) = 0$.

B. Transform-domain filtering. In the light of the previous first-order approximation, $f(\tilde{z})$ can be treated as a clipped normal random variable, with clipping from below at $f(0) = 0$ and from above at $f(1)$, and variance (after clipping) equal to c^2. Thus, qualitatively, the rapid Gaussianization discussed in Section 3.2.1 holds also after the variance stabilization, enabling the effectiveness of the homoskedastic denoising filter. Hence, in the case of homoskedastic filtering, the approximation (17) has the form

$$\mathbf{D}_{\mathsf{ho}}(f(\tilde{z})) \approx E\{f(\tilde{z})\}. \tag{19}$$

C. Debiasing and inverse transformation. Because of the non-linearity of f, the first-order approximation is never exact. While minor approximation errors on the variance are typically acceptable, the errors on the expectations are not, because they result in a systematic estimation bias. More precisely, we have that

$$E\{f(\tilde{z})\} \neq f(E\{\tilde{z}\}); \tag{20}$$

this discrepancy must be compensated before inverting f.

The two terms in (20) can be clearly computed from the generalized p.d.f. (6) as

$$f(E\{\tilde{z}\}) = f\left(\int_0^1 \zeta \wp_{\tilde{z}}(\zeta)\,d\zeta\right),$$

$$E\{f(\tilde{z})\} = \int_0^1 f(\zeta)\,\wp_{\tilde{z}}(\zeta)\,d\zeta =$$

$$= \int_0^1 f(\zeta)\frac{1}{\sigma(y)}\phi\left(\frac{\zeta-y}{\sigma(y)}\right)d\zeta + f(1)\left(1 - \Phi\left(\frac{1-y}{\sigma(y)}\right)\right).$$

We note that $f(0) = 0$ and therefore the mass at 0 from (6) does not show up in the last equation above. Let now h be the function defined (implicitly varying y in Y) by

$$f(E\{\tilde{z}\}) \overset{h}{\longmapsto} E\{f(\tilde{z})\} = h(f(E\{\tilde{z}\})). \tag{21}$$

The invertibility of h follows essentially from the monotonicity of f (a proof is given in the Appendix). It means that

$$h^{-1}(E\{f(\tilde{z})\,|y\}) = f(E\{\tilde{z}|y\}) \quad \forall y \in Y.$$

Thus, we have that

$$h^{-1}\big(\mathbf{D}_{\mathsf{ho}}(f(\tilde{z}))\big) \approx f(E\{\tilde{z}\}) \tag{22}$$
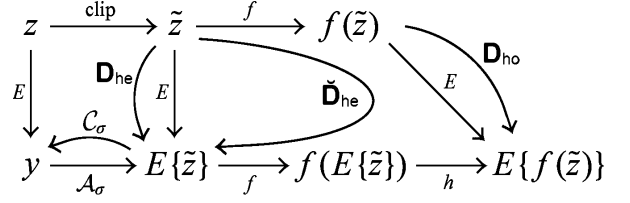

Figure 5: Diagram of operations linking the random and the deterministic variables used in our modeling. The curved arrows show the actions of the denoising filters and of the declipping transformation.

and hence that

$$f^{-1}\big(h^{-1}\big(\mathbf{D}_{\mathsf{ho}}(f(\tilde{z}))\big)\big) \approx E\{\tilde{z}\}, \tag{23}$$

which is the final form of the approximation (17) for the case of homoskedastic filtering. Due to (23), we can define a denoising filter $\check{\mathbf{D}}_{\mathsf{he}}$ for heteroskedastic noise as $\check{\mathbf{D}}_{\mathsf{he}} = f \circ \mathbf{D}_{\mathsf{ho}} \circ h^{-1} \circ f^{-1}$.

In the implementations, before applying h^{-1} and the subsequent steps of the inversion procedure, one needs to ensure (e.g., by a projection operator) that the output of the denoising filter $\mathbf{D}_{\mathsf{ho}}(f(\tilde{z}))$ (19) does not exceed the range of $E\{f(\tilde{z})\}$, which coincides with the codomain of h.

Let us also observe that, in most image processing algorithms exploiting variance-stabilization in the integral form (18), the role of h is neglected and the two terms in (20) are mistakenly assumed as equal (see e.g., [21], [2], [16], [18]). However, the compensation (22) turns out to be crucial, particularly when dealing with asymmetric distributions such as (6), and the difference between $E\{f(\tilde{z})\}$ and $f(E\{\tilde{z}\})$ can be significant, as illustrated in Section 4.

3.3. Output declipping

Figure 5 provides a diagram of the various operations which link the random and the deterministic variables used in our modeling. As shown in Section 3.2, using either $\mathbf{D}_{\mathsf{he}}$ or $\check{\mathbf{D}}_{\mathsf{he}}$ (from now on denoted collectively as $\mathbf{D}$), we can get an estimate of $E\{\tilde{z}\} = \tilde{y}$. However, our goal is to estimate the non-clipped y. Its estimate can be obtained exploiting the inverse transformation (11) as

$$\hat{y} = \mathcal{C}_\sigma(\mathbf{D}(\tilde{z})), \tag{24}$$

or, via the approximate inverse (16), as $\hat{y} = \mathcal{C}(\mathbf{D}(\tilde{z}), \tilde{\sigma}(\mathbf{D}(\tilde{z})))$. As can be intuited from Figure 3 and by the very definition of $\mathcal{E}_r$, the range of $\hat{y}$ can in principle be the whole $(-\infty, \infty)$. Nevertheless, because $\mathcal{E}_r$ has unbounded derivative and because of finite precision, the actual range that can be obtained is rather limited. Firstly, with double-precision floats (as, e.g., in Matlab) the limit is set at about 8σ beyond the 0 and 1 bounds, which roughly means that the whole range cannot be extended more than $(16\sigma + 1)$ times (note that $\Phi(\mu)$ approaches the relative spacing between adjacent double-precision floats at $\mu = -8.1$). However, in practice, the achievable range can be much smaller, mainly because of estimation errors. In this sense, it is important to emphasize that the inverse $\mathcal{C}$ (16) treats $\mathbf{D}(\tilde{z})$ as an exact $E\{\tilde{z}\}$ and thus it does not consider the estimation errors

6

in the estimation of y or $\sigma(y)$ and it is not an optimal inverse (such as a MMSE or ML inverse). In particular, if the quality of estimation for $\mathbf{D}(\tilde{z})$ is not good, applying $\mathcal{C}$ may lead to dramatic consequences, with $\hat{y}$ being a worse estimate of y than $\mathbf{D}(\tilde{z})$ itself.

The fact that the extended range increases with σ follows directly from the form of (14)-(16). It means that by denoising we can "take advantage of the noise" to obtain an image with a wider range, since the range of $\tilde{z}$ and $\tilde{y}$ is always smaller than that of y. Qualitatively, the estimation accuracy of $\hat{y}$ depends on the smoothness (e.g., bandwidth) of y: smooth clipped (e.g., overexposed) areas can be recovered quite accurately, whereas fine details or singularities are usually compromised, because denoising is not able to provide a good enough estimate for reliable declipping.

4. Experimental results

We devote most of the experiments to the homoskedastic filtering, firstly because this is the approach that has more general applicability and secondly because there is lack of algorithms for heteroskedastic filtering implemented for a special signal-dependent noise model such as (4). As denoising homoskedastic filter $\mathbf{D}_{\text{ho}}$, we consider the BM3D [5], the BLS-GSM [13], the TLS [18], the K-SVD [7], and the SA-DCT [11] algorithms for additive white Gaussian noise removal. For all these algorithms we use the publicly available codes with default parameters[3]. A modified SA-DCT algorithm for signal-dependent noise [12] is used as heteroskedastic filter $\mathbf{D}_{\text{he}}$. Similar to [13], for all experiments we follow the signal-dependent noise models (1)-(4), with the noise term in (1) composed of two mutually independent parts, a *Poissonian* signal-dependent component η_{p} and a *Gaussian* signal-independent component η_{g}: $\sigma(y(x))\xi(x) = \eta_{\text{p}}(y(x)) + \eta_{\text{g}}(x)$. In particular, these components are defined as $\big(y(x) + \eta_{\text{p}}(y(x))\big)\chi \sim \mathcal{P}(\chi y(x))$, where $\chi > 0$ and $\mathcal{P}$ denotes the Poisson distribution, and $\eta_{\text{g}}(x) \sim \mathcal{N}(0, b)$. From the elementary properties of the Poisson distribution, we obtain the following equation for mean and variance:

$$E\big\{\big(y(x) + \eta_{\text{p}}(y(x))\big)\chi\big\} = \text{var}\big\{\big(y(x) + \eta_{\text{p}}(y(x))\big)\chi\big\} = \chi y(x).$$

Since $E\big\{\big(y(x) + \eta_{\text{p}}(y(x))\big)\chi\big\} = \chi y(x) + \chi E\big\{\eta_{\text{p}}(y(x))\big\}$ and $\chi^2 \text{var}\big\{\eta_{\text{p}}(y(x))\big\} = \chi y(x)$, it follows that

$$E\big\{\eta_{\text{p}}(y(x))\big\} = 0 \quad \text{and} \quad \text{var}\big\{\eta_{\text{p}}(y(x))\big\} = y(x)/\chi.$$

Thus, as discussed in Section 2.1, $\sigma^2(y(x)) = ay(x) + b$, with $a = \chi^{-1}$.

4.1. Experiments with synthetic noise

We begin from simulations for two 1024×1024 test images [24], *Testpat* and *Man*. For the simulations, the images
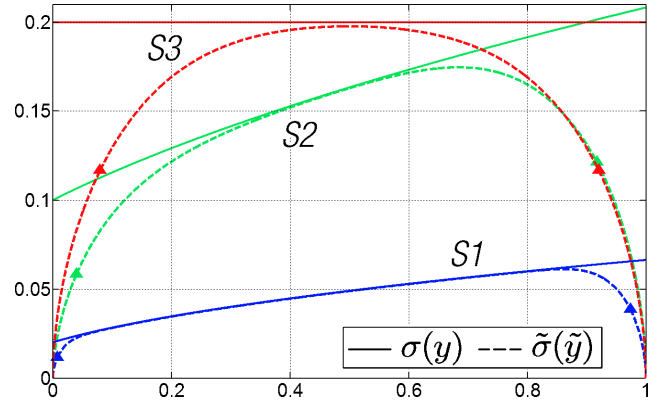
Figure 6: Plots of the standard-deviation functions σ and $\tilde{\sigma}$ for $\sigma(y) = \sqrt{ay+b}$ with three different sets of parameters: "*S1*" a=0.004, b=0.02^2; "*S2*" a=1/30, b=0.1^2; and "*S3*" a=0, b=0.2^2.

are first corrupted by synthetic noise[4] with three different sets of parameters for the Poissonian and Gaussian components: "*S1*" a=0.004, b=0.02^2, "*S2*" a=1/30, b=0.1^2, and "*S3*" a=0, b=0.2^2 (Figure 6). Clipping produces the observed images $\tilde{z}$. Two of the clipped noisy images are shown in Figures 7 and 8. The standard-deviation functions σ and $\tilde{\sigma}$ corresponding to the three sets of parameters are shown in Figure 6. Let us emphasize that even though for "*S3*" we have σ constant, the $\tilde{\sigma}$ is not and that the clipped noise is always signal-dependent.

We first assume that the noise parameters are exactly known a priori, in order to avoid the potential influence of noise mis-estimation on the final result.

Figure 9 presents the computed variance-stabilizing transformations (18)[5]. The inverse mappings h^{-1} are shown in Figure 10; observe that in all three cases the mapping deviates from the identity and that this difference is significant in correspondence with the markedly non-affine portions of the variance-stabilizing transformation. Figure 11 illustrates both the direct transformations $y \mapsto \tilde{y} = \mathcal{A}_\sigma(y)$ and, by transposition of the plot, the inverse transformations $\tilde{y} \mapsto y = \mathcal{C}_\sigma(\tilde{y})$. The latter are applied on $\mathbf{D}(\tilde{z})$ for the declipping. The peak signal-to-noise ratio (PSNR) values (dB) of $\tilde{z}$, $\mathbf{D}(\tilde{z})$ (23), and $\hat{y}$ are given in Table 1. In the table we also provide PSNR results obtained exploiting the additional information about the actual range of the original images y being the unit interval $[0, 1]$, by defining a constrained estimate $\hat{y}_0^1 = \max\big\{0, \min\{\hat{y}, 1\}\big\}$.

We can also observe that while for some weaker estimators (e.g., BLS-GSM, SA-DCT) the PSNR gain from $\hat{y}$ to $\hat{y}_0^1$ can be dramatic, for more powerful and stable algorithms (e.g.,

---

[4]The observations are generated according to the following Matlab code:
```
randn('state',0);   % initializes pseudo-random generator
rand('state',0);    % initializes pseudo-random generator
if a==0             % only Gaussian component
  z=y+sqrt(b)*randn(size(y));
else                % Poissonian and Gaussian components
  chi=1/a;
  z=poissrnd(chi*y)/chi+sqrt(b)*randn(size(y));
end
ztilde=max(0,min(1,z));
```

[5]The integral (18) converges and f is bounded for any $a \geq 0$ and $b > 0$, as proved in Proposition A.3.6 of the Appendix. Thus, we can set $t_0 = 0$ and have $f(0) = 0$.

---

[3]For the BLS-GSM we use the full steerable pyramid implementation and for the K-SVD we use four times redundancy with 256 atoms on blocks of size 8×8.
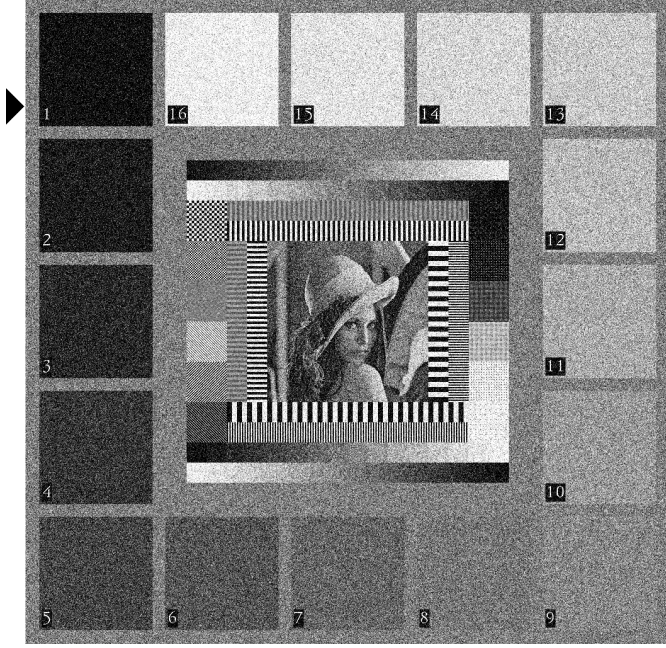
Figure 7: Noisy $\tilde{z}$ (4) *Testpat* test image of size 1024×1024. Noise parameters are "*S3*" : $a=0$, $b=0.2^2$. The triangle indicates the row for the cross-sections plotted in Figure 14.

Figure 8: Noisy $\tilde{z}$ (4) *Man* test image of size 1024×1024. Noise parameters are "*S2*" : $a=1/30$, $b=0.1^2$

BM3D) this gain exists but is much less than the PSNR difference between $\hat{y}$ and $\mathbf{D}(\tilde{z})$. Although the numbers in Table 1 highlight the obvious instability of the declipping transformation $\mathcal{C}_\sigma$, at the same time they demonstrate that, provided successful denoising in $\mathbf{D}(\tilde{z})$, substantial improvement is further obtained by compensating the bias due to the clipped noise. The final declipped and denoised estimates $\hat{y}$ (24) obtained using the BM3D algorithm as homoskedastic filter $\mathbf{D}_{ho}$ are shown
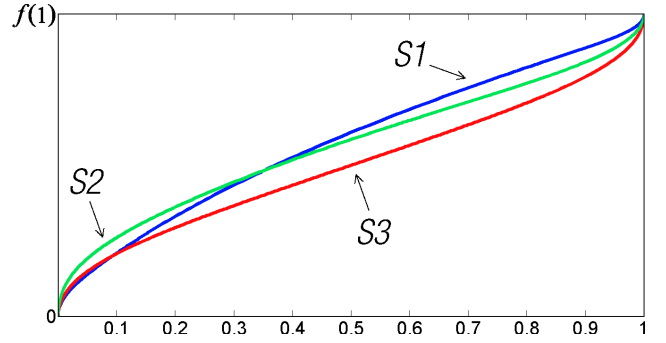

Figure 9: Variance-stabilizing homomorphic transformation f (18) for the three standard-deviation functions shown in Figure 6. The curves are plotted with respect to a normalized vertical scale $\left[0, f(1)\right]$.
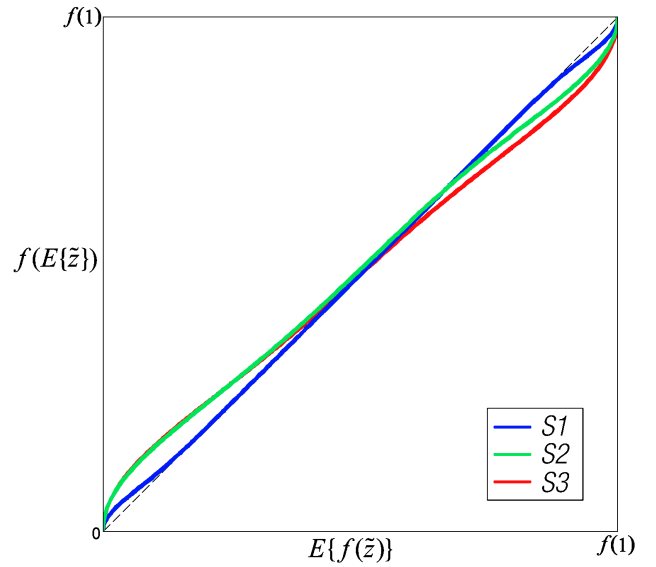

Figure 10: Plots of the inverse mapping h^{-1}, $h^{-1}\left(E\{f(\tilde{z})\,|y\}\right) = f(E\{\tilde{z}|y\})$, $y \in Y$. The curves are plotted with respect to normalized horizontal and vertical scales $\left[0, f(1)\right]$.
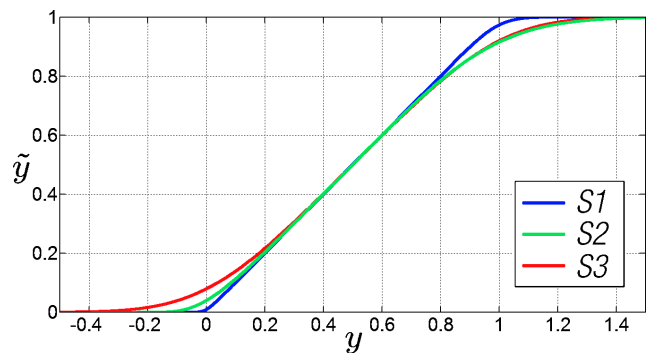

Figure 11: The direct transformations $y \mapsto \tilde{y} = \mathcal{A}_\sigma(y)$. The inverse transformations (declipping) $\tilde{y} \mapsto y = \mathcal{C}_\sigma(\tilde{y})$ can be visualized by transposition of this plot.

in Figures 12 and 13.

It is important to emphasize that the BM3D filter leads to consistent improvement throughout all experiments and that its estimates $\hat{y}$ are numerically better than both the estimates $\hat{y}$ and

Table 1: PSNR (dB) values for the clipped noisy images $\tilde{z}$ (4), and for the denoised $\mathbf{D}(\tilde{z})$ (23), denoised and declipped $\hat{y}$ (24), and range-constrained estimates $\hat{y}_0^1 = \max\{0, \min\{\hat{y}, 1\}\}$ obtained using different denoising algorithms. Best results are boldfaced.

| Noise parameters	S1 a=0.004, b=0.02^2						S2 a=1/30, b=0.1^2						S3 a=0, b=0.2^2					
	Testpat			*Man*			*Testpat*			*Man*			*Testpat*			*Man*		
Noisy $\tilde{z}$	26.83			27.51			16.70			17.07			14.99			14.86		
	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$
BM3D [5]	**39.03**	**41.20**	**42.29**	**33.75**	**33.75**	**33.76**	**28.49**	**31.53**	**32.33**	**27.72**	27.99	28.03	**26.45**	**29.36**	**30.15**	**25.26**	**26.32**	26.44
TLS [18]	32.50	31.32	32.95	33.48	33.46	33.49	25.54	25.69	26.78	27.28	27.36	27.50	23.90	24.26	25.23	24.72	25.24	25.54
K-SVD [7]	35.49	34.29	36.43	33.05	33.01	33.05	26.03	26.66	27.48	26.08	26.06	26.15	23.99	24.77	25.36	23.47	23.80	23.96
BLS-GSM [20]	33.26	28.78	33.72	33.58	33.57	33.59	24.64	22.52	25.63	26.85	26.96	27.04	22.89	21.45	23.99	24.16	24.79	24.96
SA-DCT hom.[11]	37.86	37.66	39.96	33.57	33.56	33.57	26.83	27.20	28.87	27.37	27.61	27.66	24.49	24.65	26.26	24.80	25.81	25.91
SA-DCT het. [12]	38.57	35.85	41.40	33.52	33.53	33.54	27.91	22.53	30.95	27.69	**28.07**	**28.13**	25.97	19.94	28.94	25.25	25.92	**26.82**
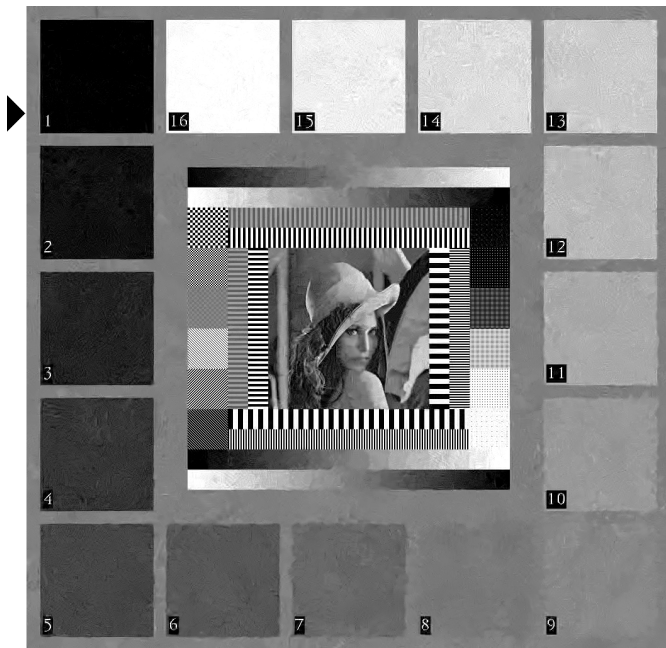
Figure 12: Denoised and debiased $\hat{y}$ (24) *Testpat* image using the BM3D algorithm as homoskedastic filter (input image $\tilde{z}$ shown in Figure 7). The triangle indicates the row for the cross-sections plotted in Figure 14.

Figure 13: Denoised and debiased $\hat{y}$ (24) *Man* image using the BM3D algorithm as homoskedastic filter (input image $\tilde{z}$ shown in Figure 8).

$\hat{y}_0^1$ produced by any of the other homoskedastic filters.

These features are well illustrated by the cross-sections of *Testpat* shown in Figure 14 and 15. We can observe that the inversion of the clipping operated by $\mathcal{C}$ is able to properly correct the misestimation of the BM3D $\mathbf{D}(\tilde{z})$ estimate (leading to an almost 3 dB gain in PSNR), whereas it results in dramatic amplification of the overshooting artifacts of the BLS-GSM estimate (with the consequent drop of about 1.5 dB).

The results obtained by the heteroskedastic SA-DCT on the *Testpat* image require a separate explanation. The SA-DCT algorithms exploit filtering on adaptive neighborhoods and for small neighborhoods the Gaussianization considered in Section 3.2.1 is marginal or even absent (in case of singletons). It turns out that the adaptive-scale selection in the heteroskedastic filter is much more sensitive than that in the homoskedastic implementation. As a result, the heteroskedastic SA-DCT operates on smaller adaptive neighborhoods, which leads to sharper estimates: indeed, in Table 1, the PSNR values of $\mathbf{D}(\tilde{z})$ for the heteroskedastic SA-DCT is usually among the highest and the differences between the two SA-DCT implementations are more conspicuous for the edge-rich *Testpat* image. However, this sharpness leads also to a higher number of pixels close to the range boundaries (which are then declipped to extreme values well outside [0, 1]) and to the consequent PSNR drop for $\hat{y}$.

We note also that the range of the noise-free *Man* image is concentrated well inside the interval [0, 1]. Therefore, for relatively low levels of noise such as in "S1", the impact of clipping is negligible and thus the resulting PSNR numbers for the $\mathbf{D}(\tilde{z})$, $\hat{y}$, and $\hat{y}_0^1$ estimates are nearly identical.

Often the noise parameters are not known in advance and need to be estimated from the given image. The standard-deviation functions σ and $\tilde{\sigma}$ estimated with the algorithm [13] for the above observations $\tilde{z}$ are shown in Figures 16 and 17.

Table 2: Noise parameter values for the model (2) estimated using the algorithm [13] from the individual clipped noisy images $\tilde{z}$ (4).

Noise parameters	S1 a=0.004, b=0.02^2		S2 a=1/30, b=0.1^2		S3 a=0, b=0.2^2	
	Testpat	*Man*	*Testpat*	*Man*	*Testpat*	*Man*
Estimated a	0.00448	0.00445	0.0334	0.0349	0.00103	0.00612
Estimated b	0.000463	0.000504	0.0102	0.00954	0.0393	0.0373

Table 3: PSNR (dB) values for the denoised $\mathbf{D}(\tilde{z})$ (23), denoised and declipped $\hat{y}$ (24), and range-constrained estimates $\hat{y}_0^1 = \max\{0, \min\{\hat{y}, 1\}\}$ obtained using different denoising algorithms. These results differ from those in Table 1 since here the noise paremeters were estimated from each image $\tilde{z}$ using the algorithm [13], as reported in Table 2 and shown in Figures 16 and 17. Best results are boldfaced.

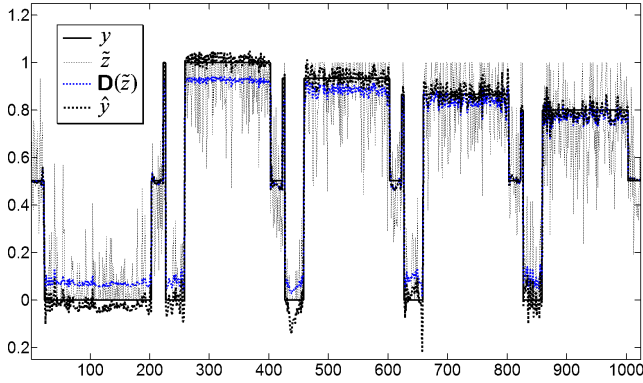|	S1						S2						S3					
	Testpat			*Man*			*Testpat*			*Man*			*Testpat*			*Man*		
	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$	$\mathbf{D}(\tilde{z})$	$\hat{y}$	$\hat{y}_0^1$
BM3D [5]	**38.95**	**41.18**	**42.50**	**33.67**	**33.68**	**33.68**	**28.49**	31.55	32.39	**27.71**	27.96	28.01	**26.45**	**29.33**	**30.12**	**25.25**	**26.24**	26.35
SA-DCT het. [12]	38.60	35.47	41.59	33.44	33.46	33.46	27.91	22.53	30.97	27.69	**28.07**	**28.13**	25.98	19.96	28.95	**25.25**	25.91	**26.81**

Figure 14: Cross-sections of the original y (1), noisy $\tilde{z}$ (4), denoised $\mathbf{D}(\tilde{z})$ (23), and of the denoised and declipped $\hat{y}$ (24) (shown in Figure 7), using the BM3D denoising algorithm [5], for the 171th row of the *Testpat* image (the row is indicated by the black triangle next to the images in Figures 7 and 12).
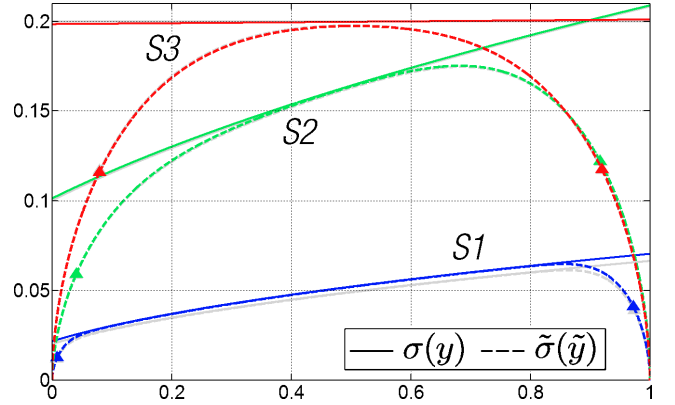

Figure 16: Standard-deviation functions σ and $\tilde{\sigma}$ estimated from the clipped noisy $\tilde{z}$ *Testpat* image. The true curves are drawn in light gray beneath the respective estimated lines.
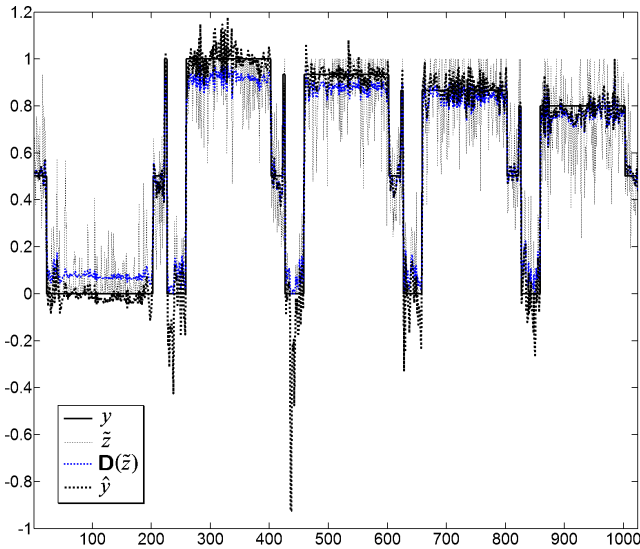

Figure 15: Cross-sections of the original y (1), noisy $\tilde{z}$ (4) (shown in Figure 7), denoised $\mathbf{D}(\tilde{z})$ (23), and of the denoised and declipped $\hat{y}$ (24), using the BLS-GSM denoising algorithm [20], for the 171th row of the *Testpat* image (the row is indicated by the black triangle next to the images in Figures 7 and 12).
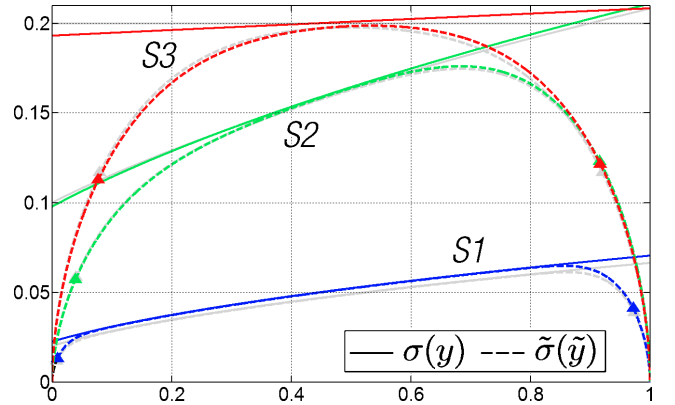

Figure 17: Standard-deviation functions σ and $\tilde{\sigma}$ estimated from the clipped noisy $\tilde{z}$ *Man* image. The true curves are drawn in light gray beneath the respective estimated lines.

The estimated curves closely match with the respective true curves. The values of the estimated parameters are reported in Table 2, while Table 3 gives the PSNR results obtained by the BM3D homoskedastic filter and by the SA-DCT heteroskedas-

Figure 18: Top: noisy raw-data image *Bottle* $\tilde{z}$ (4) (Fujifilm FinePix S9600 camera). Middle: denoised $\mathbf{D}(\tilde{z})$ (23). Bottom: denoised and declipped $\hat{y}$ (24).

Figure 19: Top: noisy raw-data image *Staircase* $\tilde{z}$ (4) (Fujifilm FinePix S9600 camera). Middle: denoised $\mathbf{D}(\tilde{z})$ (23). Bottom: denoised and declipped $\hat{y}$ (24).

tic filter. These results are essentially the same as those obtained if the noise parameters were known exactly in advance.

4.2. Experiments with raw-data images

We also show results of processing real raw-data 1224×922 images from the CCD sensor (green subcomponent) of a Fujifilm FinePix S9600 camera at ISO 1600. For these experiments

we use BM3D algorithm as homoskedastic filter. Figures 18 and 19 (top) show the raw-data images $\tilde{z}$ *Bottle* and *Staircase*. The standard-deviation functions σ and $\tilde{\sigma}$ had been estimated from each image using the algorithm [13]; in particular, the esti-
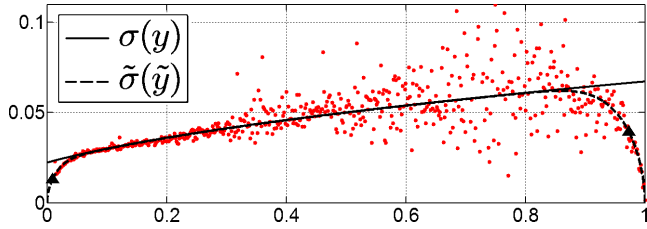
Figure 20: Standard-deviation functions σ and $\tilde{\sigma}$ estimated from the raw-data image *Staircase* acquired by the CCD sensor of a Fujifilm FinePix S9600 camera, shown in Figure 19 (top). The red dots are conditioned expectation/standard-deviation pairs measured on the image by the algorithm [13].
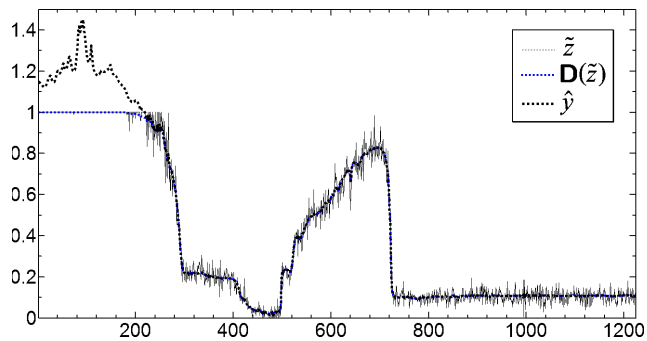

Figure 21: Cross-sections for the 200th line of the raw image *Bottle* $\tilde{z}$ (Figure 18) and of its denoised $\mathbf{D}(\tilde{z})$ (23) and denoised and declipped $\hat{y}$ (24) images (the line is indicated by the black triangle next to the images).
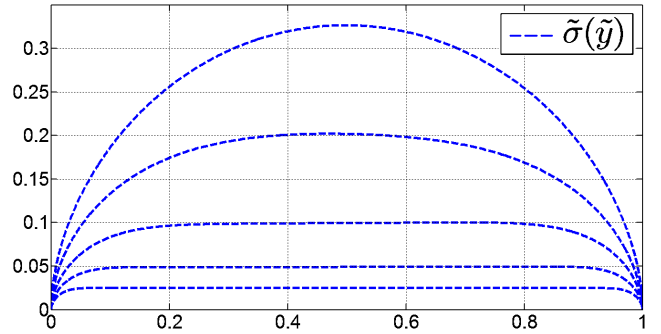

Figure 22: Estimated standard-deviation functions $\tilde{\sigma}(\tilde{y})$ for the clipped sine images with $\sigma(y) = 0.4, 0.2, 0.1, 0.05, 0.025$.

these curves are not constant, as opposed to $\sigma(y)$. The cross-sections of the of denoised and declipped images (using BM3D as homoskedastic filter) are given in Figure 23, where we can observe that for $\sigma(y) \leq 0.05$ we are no longer able to effectively declip the denoised $\mathbf{D}(\tilde{z})$. A close inspection reveals that our implementation is able to extend the range of about 5σ both above 1 and below 0 (thus, roughly comparable to the theoretical limit of about 8σ discussed in Section 3.3), however the quality of the reconstruction is much affected by the accuracy of the estimation of $\tilde{\sigma}(\tilde{y})$ and by that of the denoising of $\tilde{z}$.

5. Discussion and conclusions

We identified conditions and derived transformations for the rigorous processing of clipped noisy signals using generic filters. Experiments demonstrate the success of the proposed approach. The shown techniques can also be used for the accurate linearization of the sensor response, including the correct pedestal or black level measurement. They are thus relevant for high dynamic range imaging applications and for all fields where accuracy in the estimation of the sensor response is crucial, such as inverse imaging (allowing to use of linear deconvolution for deblurring over-exposed raw-data), digital forensics (sensor identification), etc.

A delicate issue concerns the intrinsic instability of the declipping transformation $\mathcal{C}_\sigma$; future work shall investigate optimal inverse transformations which can account for the variance of the denoised estimate $\mathbf{D}(\tilde{z})$, thus providing a stable way to declip this estimate. Unfortunately, such approach would not be directly applicable as "plug-in" element, because denoising filters do not usually return the pointwise variance of the output denoised image. In this paper, our goal had been to provide a solution "pluggable" into existing implementations.

The mathematical analysis included in the Appendix demonstrates an intimate interplay between the qualitative characteristics of the standard-deviation function $\sigma(y) = \sqrt{ay + b}$ and the main features of clipped data, mainly due to its concavity and sublinear growth. As shown in our previous work [13], complex nonlinearities observed in the sensor output can be accurately explained by this conceptually simple clipped signal-dependent noise model. We expect that the new general results

mation for the *Staircase* image is illustrated in Figure 20[6]. The denoised images $\mathbf{D}(\tilde{z})$, and the declipped images $\hat{y}$ are shown in Figures 18 and 19 (middle and bottom). The declipped images have an extended range of about $[-0.03, 1.45]$ and are thus shown after scaling to $[0, 1]$. The extension of the range is clearly visible in the overexposed highlights of the two scenes and, in particular, in the cross-sections plotted in Figure 21.

4.3. Declipping vs. noise standard-deviation

To illustrate the interaction between the range of the declipped signal and the standard-deviation of the original signal prior to clipping, we consider the following experiment. Define the true image y of size 512×512 pixels as

$$y(x_1, x_2) = 0.7 \sin\left(\frac{2\pi}{511}x_1\right) + 0.5, \quad x_1, x_2 = 0, \dots, 511.$$

The range of y is $Y = [-0.2, 1.2]$. We generate five different noisy images $z(x) = y(x) + \eta_g(x)$ where $\eta_g(x) \sim \mathcal{N}(0, b)$, having constant standard-deviations $\sigma(y) = \sqrt{b} = 0.4, 0.2, 0.1, 0.05, 0.025$. The standard-deviation curves $\tilde{\sigma}(\tilde{y})$ for the clipped images $\tilde{z}$ are shown in Figure 22; observe that

---

[6]The parameter set "*S1*", used for the synthetic experiments, had been purposely chosen to roughly match the noise of these raw-data images. Thus, the transformations shown in Figures 9, 10, 11 for the "*S1*" case qualitatively correspond to those actually used for processing the raw data. Likewise, the PSNR values shown in Tables 1 and 3 give an objective indication of the improvement achievable when processing real images.
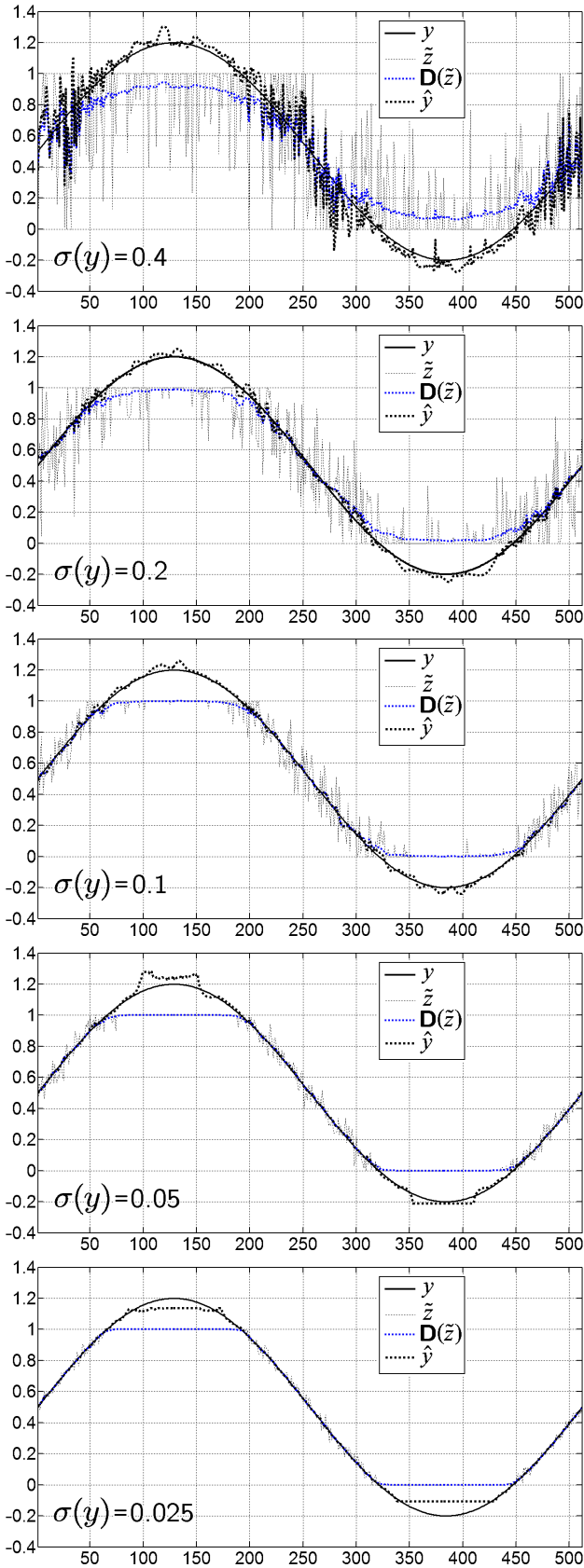
Figure 23: Cross-sections of the original y (1), noisy $\tilde{z}$ (4), denoised $\mathbf{D}(\tilde{z})$ (23), and of the denoised and declipped $\hat{y}$ (24), for the clipped sine images with $\sigma(y) = 0.4, 0.2, 0.1, 0.05, 0.025$.

proved in the Appendix can find a useful role in the modeling of sensing devices other than the conventional digital imagers.

We concentrated on the variance-stabilizer (18) mainly because of its intuitive simplicity. Other transformations may have been used as well, including optimized ones such as those proposed in [9].

Finally, we highlight that the proposed declipping approach is applicable also to data interpolated from clipped noisy samples, including color raw Bayer data: operatively there is no difference whether $\mathbf{D}(\tilde{z}(x))$, $x \in X$, is obtained from denoising $\tilde{z} : X \to [0, 1]$ (as illustrated in this paper) or from denoising/interpolating a smaller-size $\tilde{z} : X_{\mathrm{sub}} \to [0, 1]$, where $X_{\mathrm{sub}} \subsetneq X$ is a subset or sublattice of X, as long as $\mathbf{D}(\tilde{z}(x))$ can be treated as an approximation of $\tilde{y}(x)$.

Related Matlab software is provided online at www.cs.tut.fi/~foi/sensornoise.html

A. Appendix

This mathematical appendix is organized in four parts. First, we investigate the monotonicity with respect to y of the conditional expectations of a clipped variable $\tilde{z}$ or of a transformed clipped variable $f(\tilde{z})$. These results are mostly needed to guarantee that the main elements used in the presented procedure are indeed well defined. Second, we study the ranges of the expectation and standard deviation of the clipped variables. Third, we consider the boundedness of the variance stabilizing transformations defined by (18). In these three parts, the main emphasis is placed on the standard-deviation function σ. Fourth and last, we analyze the special case of a $\sigma(y) = \sqrt{ay + b}$ with $b < 0 < a$.

In the sequel we assume $z \sim \mathcal{N}(y, \sigma^2(y))$, and, unless explicitly noted otherwise, $\tilde{z} = \max\{0, \min\{1, z\}\}$, with $\tilde{y} = E\{\tilde{z}|y\}$.

A.1. Monotonicity of $E\{\tilde{z}|y\}$, $E\{f(\tilde{z})|y\}$ and $f(E\{\tilde{z}|y\})$

Throughout the proofs we utilize Lebesgue integration. Absolute continuity is used to enable integration by parts and to provide differentiability almost everywhere (in Lebesgue sense). The reader can refer to mathematical analysis textbooks such as Rudin [22] for the basic definitions.

Proposition A.1.1. If $f : \mathbb{R} \to \mathbb{R}$ is a monotone nondecreasing absolutely continuous function such that $f(\zeta) e^{-|\zeta|^\alpha} \underset{|\zeta| \to \infty}{\longrightarrow} 0$ for some $\alpha < 2$, $\sigma(y)$ is absolutely continuous and concave, and $f'(\zeta) = 0$ if ζ does not belong to the domain Y of σ, then $E\{f(z)|y\}$ is a monotone nondecreasing function of $y = E\{z|y\}$. If f is also not identically constant, then $E\{f(z)|y\}$ is strictly increasing.

Proof. We show that the derivative of $E\{f(z)|y\}$ with respect to y,

$$\frac{\partial E\{f(z)|y\}}{\partial y} = \int_{-\infty}^{+\infty} f(\zeta) \frac{\partial}{\partial y} \left(\frac{1}{\sigma(y)} \phi\left(\frac{\zeta-y}{\sigma(y)}\right) \right) d\zeta,$$

is nonnegative almost everywhere. Let us proceed by integrating by parts, defining

$$g'(\zeta) = \frac{\partial}{\partial y} \left(\frac{1}{\sigma(y)} \phi\left(\frac{\zeta-y}{\sigma(y)}\right) \right)$$
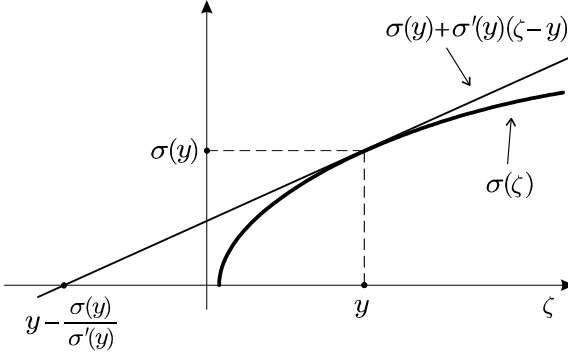
13

Figure 24: Illustration of the role of concavity in ensuring that $E\{f(z)|y\}$ is a monotone function of y (see proof of Proposition A.1.1).
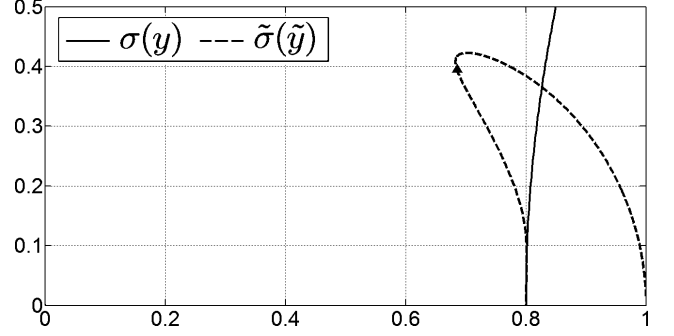

Figure 25: Standard-deviation function $\sigma(y) = \sqrt{5y - 4}$ (solid line) and the corresponding multi-valued mapping $\tilde{y} \longmapsto \tilde{\sigma}$ (dashed line, formally denoted as $\tilde{\sigma}(\tilde{y})$).

and thus

$$g(\zeta) = \frac{\partial}{\partial y}\left(\Phi\left(\frac{\zeta - y}{\sigma(y)}\right)\right) = -\phi\left(\frac{\zeta - y}{\sigma(y)}\right)\frac{\sigma(y) + \sigma'(y)(\zeta - y)}{\sigma^2(y)}.$$
(25)

We have

$$\frac{\partial E\{f(z)|y\}}{\partial y} = \int_{-\infty}^{+\infty} f(\zeta) g'(\zeta)\, d\zeta =$$

$$= [f(\zeta) g(\zeta)]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f'(\zeta) g(\zeta)\, d\zeta =$$

$$= -\int_{-\infty}^{+\infty} f'(\zeta) g(\zeta)\, d\zeta,$$
(26)

where $[f(\zeta) g(\zeta)]_{-\infty}^{+\infty} = 0$ because the hypotheses on f ensure that $f(\zeta) g(\zeta) \underset{|\zeta|\to\infty}{\longrightarrow} 0$. Let us examine (25) and (26) and consider separate cases depending on the sign of $\sigma'(y)$ (we remark that $\sigma'(y)$ is defined almost everywhere on Y). If $\sigma'(y) = 0$, then $g(\zeta) < 0\ \forall \zeta$ and thus $\frac{\partial E\{f(z)|y\}}{\partial y} \geq 0$. Here, since $g(\zeta) < 0$ and $f'(\zeta) \geq 0$, we deduce that $\frac{\partial E\{f(z)|y\}}{\partial y} = 0$ if and only if $f'(\zeta) g(\zeta) = 0$ almost everywhere. Thus, provided f is not identically constant, we have $\frac{\partial E\{f(z)|y\}}{\partial y} > 0$. If instead $\sigma'(y) > 0$, we have that $g(\zeta) < 0$ for all $\zeta > y - \frac{\sigma(y)}{\sigma'(y)}$. The concavity of σ on its domain implies that $\sigma(\zeta) \leq \sigma'(y)(\zeta - y) + \sigma(y)$, where the last expression gives the line tangent to σ at y, as illustrated in Figure 24. However, since $\sigma(\zeta)$ cannot be negative, we necessarily have that σ cannot be defined for ζ beyond the solution of $\sigma'(y)(\zeta - y) + \sigma(y) = 0$, namely $\zeta = y - \frac{\sigma(y)}{\sigma'(y)}$. Figure 24 provides a sketch of this situation. Then, from the hypotheses, it follows that $f'(\zeta) = 0$ for all $\zeta < y - \frac{\sigma(y)}{\sigma'(y)}$. The situation for the case where $\sigma'(y) < 0$ is analogue, with $g(\zeta) < 0$ for all $\zeta < y - \frac{\sigma(y)}{\sigma'(y)}$ and $f'(\zeta) = 0$ for all $\zeta > y - \frac{\sigma(y)}{\sigma'(y)}$.

We now split the integral in two terms:

$$\int_{-\infty}^{+\infty} f'(\zeta) g(\zeta)\, d\zeta =$$

$$= \int_{-\infty}^{y - \frac{\sigma(y)}{\sigma'(y)}} f'(\zeta) g(\zeta)\, d\zeta + \int_{y - \frac{\sigma(y)}{\sigma'(y)}}^{+\infty} f'(\zeta) g(\zeta)\, d\zeta.$$
(27)

From the above considerations, we see immediately that one of the two terms is zero (because $f'(\zeta) = 0$) while the other term cannot be larger than zero (because $g(\zeta) < 0$ and $f'(\zeta) \geq 0$). Therefore, we obtain $\frac{\partial E\{f(z)|y\}}{\partial y} \geq 0$. This inequality becomes strict (i.e., $E\{f(z)|y\}$ becomes strictly increasing) provided f is not identically constant, as this implies that one term in (27) is strictly negative. □

Remark A.1.2. The hypothesis that $f(\zeta) e^{-|\zeta|^\alpha} \underset{|\zeta|\to\infty}{\longrightarrow} 0$ for some $\alpha < 2$ corresponds to an exponential bound on the possible growth of f, which forces $f(\zeta) g(\zeta) \underset{|\zeta|\to\infty}{\longrightarrow} 0$. This hypothesis is trivially satisfied by all bounded functions. As it is clearly visible in the proof, the condition can be relaxed, because for the proposition to hold it is not necessary that $[f(\zeta) g(\zeta)]_{-\infty}^{+\infty} = 0$ but it is enough that this term is well-defined and non-negative.

Corollary A.1.3. Let $\tilde{z} = \max\{0, \min\{1, z\}\}$. If $f : [0, 1] \to \mathbb{R}$ is any monotone nondecreasing absolutely continuous function, σ is absolutely continuous and concave and $f'(\zeta) = 0$ if ζ does not belong to the domain of σ, then $E\{f(\tilde{z})|y\}$ is a monotone nondecreasing function of y.

Proof. Let us define a function $\tilde{f} : \mathbb{R} \to \mathbb{R}$ by $\tilde{f}(\zeta) = f(\max\{0, \min\{1, \zeta\}\})$. Clearly, $E\{f(\tilde{z})|y\}$ and $E\{\tilde{f}(z)|y\}$ coincide and, as from the above proposition, the latter expectation is a monotone nondecreasing function of y. □

Corollary A.1.4. If σ is absolutely continuous and concave and its domain includes the interval $[0, 1]$, then $\tilde{y} = E\{\tilde{z}|y\}$, where $\tilde{z} = \max\{0, \min\{1, z\}\}$, is a monotone nondecreasing function of y.

Proof. It suffices to take $f(\zeta) = \zeta$, $\zeta \in [0, 1]$, in Corollary A.1.3. □

Remark A.1.5. Note that the concavity of σ alone is not sufficient for the last corollary to hold. An example is given by taking $\sigma(y) = \sqrt{5y - 4}$, as shown in Figure 25. In the figure we can observe that $\tilde{y} = E\{\tilde{z}|y\}$ is not monotonic and thus $\tilde{\sigma}(\tilde{y})$ becoming multi-valued. In this example, $f' > 0$ outside of the domain of the concave σ. Nevertheless, the hypotheses of Proposition A.1.1 are somewhat redundant and part of these requirements can be relaxed provided more is known about σ or f, as shown by the next Proposition A.1.7. The proposition requires the following lemma.

14

Lemma A.1.6. Let σ be absolutely continuous on Y and differentiable at y. If any of the following three alternative cases holds:

    (i) $\sigma'(y) = 0$,
    (ii) $\sigma'(y) > 0$ and $y \le \frac{1}{2}$, or
    (iii) $\sigma'(y) < 0$ and $y \ge \frac{1}{2}$;

we have $\frac{\partial E\{\tilde{z}|y\}}{\partial y} > 0$.

Proof. As in Corollary A.1.4, we define $f(\zeta) = \max\{0, \min\{1, \zeta\}\}$. Using the same notation as in the proof of Proposition A.1.1, we observe that (25) can be rewritten as

$$g(\zeta) = -\frac{1}{\sigma(y)}\phi\left(\frac{\zeta-y}{\sigma(y)}\right) + \sigma'(y)\frac{\partial}{\partial\zeta}\phi\left(\frac{\zeta-y}{\sigma(y)}\right).$$

This form allows to easily exploit the symmetry of the p.d.f. ϕ. From (26), we have

$$\frac{\partial E\{\tilde{z}|y\}}{\partial y} = -\int_{-\infty}^{+\infty} f'(\zeta)\,g(\zeta)\,d\zeta = -\int_0^1 g(\zeta)\,d\zeta = $$
$$= \frac{1}{\sigma(y)}\int_0^1 \phi\left(\frac{\zeta-y}{\sigma(y)}\right)d\zeta - \sigma'(y)\int_0^1 \frac{\partial}{\partial\zeta}\phi\left(\frac{\zeta-y}{\sigma(y)}\right)d\zeta.$$

These two addends are both nonnegative. In particular, the former is strictly positive while for the latter we note that

$$\int_0^1 \frac{\partial}{\partial\zeta}\phi\left(\frac{\zeta-y}{\sigma(y)}\right)d\zeta = \phi\left(\frac{1-y}{\sigma(y)}\right) - \phi\left(\frac{y}{\sigma(y)}\right)$$

is negative if and only if $y < \frac{1}{2}$. Thus, provided either $\sigma'(y) = 0$, $\sigma'(y) > 0$ and $y \le \frac{1}{2}$, or $\sigma'(y) < 0$ and $y \ge \frac{1}{2}$, we obtain $\frac{\partial E\{\tilde{z}|y\}}{\partial y} > 0$. $\square$

Proposition A.1.7. Let σ be an absolutely continuous concave function on its domain Y, with $\frac{1}{2} \in Y \subseteq \mathbb{R}$. Then $\tilde{y} = E\{\tilde{z}|y\}$ is a monotone nondecreasing function of $y = E\{z|y\}$.

Proof. It suffices to prove that $\frac{\partial E\{\tilde{z}|y\}}{\partial y} > 0$ for all combinations of y and $\sigma'(y)$ not considered by Lemma A.1.6. We begin by proving the assertion for the case $y \ge \frac{1}{2}$ and $\sigma'(y) > 0$. Again, let $f(\zeta) = \max\{0, \min\{1, \zeta\}\}$. Because of the concavity of σ and since $\frac{1}{2} \in Y$, we deduce that the value p of ζ which solves $\sigma'(y)(\zeta - y) + \sigma(y) = 0$ is beyond $\frac{1}{2}$, i.e. $p = y - \frac{\sigma(y)}{\sigma'(y)} \le \frac{1}{2}$. Thus, $g(\zeta) < 0$ for all $\zeta \ge \frac{1}{2}$. We can now rewrite (26) as

$$\frac{\partial E\{\tilde{z}|y\}}{\partial y} = -\int_{-\infty}^{+\infty} f'(\zeta)\,g(\zeta)\,d\zeta = -\int_0^1 g(\zeta)\,d\zeta = $$
$$= \int_0^1 \left[\frac{1}{\sigma^2(y)}\phi\left(\frac{\zeta-y}{\sigma(y)}\right)\right]\left[\sigma(y) + \sigma'(y)(\zeta-y)\right]d\zeta = $$
$$= \int_0^1 A(\zeta, y)\,B(\zeta, y)\,d\zeta.$$

The two factors A and B inside the integral have specific characteristics: the one on the left, $A(\zeta, y)$, is even-symmetric with respect to $\zeta = y \ge \frac{1}{2}$, is positive, and for $\zeta > y$ it is strictly decreasing; the one on the right, $B(\zeta, y)$, is odd-symmetric with respect to $\zeta = p \le \frac{1}{2}$ and nondecreasing. From these considerations follows that for any pair of points symmetric with

respect to p, say $\zeta_1 = p - \xi$ and $\zeta_2 = p + \xi$ with $\xi > 0$, we have $-B(\zeta_1, y) = B(\zeta_2, y) > 0$ and $A(\zeta_1, y) \le A(\zeta_2, y)$. Consequently, defining $\tilde{p} = \max\{0, p\}$, we obtain

$$\frac{\partial E\{\tilde{z}|y\}}{\partial y} = $$
$$= \int_0^{2\tilde{p}} A(\zeta, y)\,B(\zeta, y)\,d\zeta + \int_{2\tilde{p}}^1 A(\zeta, y)\,B(\zeta, y)\,d\zeta > 0, \quad (28)$$

because both addends are nonnegative, with the first one strictly positive if $p > 0$ and the second one is strictly positive if $p < \frac{1}{2}$. The proof for the remaining case, where $y \le \frac{1}{2}$ and $\sigma'(y) < 0$, is simply specular to the one above, with now $p \ge \frac{1}{2}$. Thus, $\frac{\partial E\{\tilde{z}|y\}}{\partial y} > 0$ for almost every $y \in Y$ and we obtain the statement. $\square$

Corollary A.1.8. Let $\sigma(y) = \sqrt{ay + b}$, with $a > 0$ and $-\frac{b}{a} \le \frac{1}{2}$. Then $\tilde{y} = E\{\tilde{z}|y\}$ is a monotone nondecreasing function of $y = E\{z|y\}$.

Proof. σ is concave and its domain $Y = \left[-\frac{b}{a}, +\infty\right) \ni \frac{1}{2}$. $\square$

Remark A.1.9. The last proposition and, particularly, the last corollary guarantee that the transformations presented in this paper are indeed well-defined and applicable to signal-dependent noise models having $\sigma(y) = \sqrt{ay + b}$, with $a > 0$ and $-\frac{b}{a} \le \frac{1}{2}$. We also understand now that, in the example of Remark A.1.5, illustrated in Figure 25, the non-monotonicity of $\tilde{y}$ (leading to a multi-valued $\tilde{\sigma}(\tilde{y})$) is possible because there $-\frac{b}{a} = 0.8 > \frac{1}{2}$.

Remark A.1.10. Of course, if $E\{\tilde{z}|y\}$ is monotone (resp. strictly monotone) with respect to y and f is monotone (resp. strictly monotone) on $\tilde{Y}$, then also $f(E\{\tilde{z}|y\})$ is monotone (resp. strictly monotone) with respect to y. However, we emphasize that, even though f needs to be defined for all values of $\tilde{z}$, i.e. on $[0, 1]$, it is not required that f is monotone (resp. strictly monotone) on the whole $[0, 1]$, unless $\tilde{Y}$ itself coincides with $[0, 1]$. This issue is discussed in the third part of this appendix.

Remark A.1.11. The function h (21) is clearly invertible, provided both $f(E\{\tilde{z}|y\})$ and $E\{f(\tilde{z})|y\}$ are strictly monotone functions of y.

Remark A.1.12. We note that, to obtain the declipped estimate from homoskedastic filtering, it is not really necessary to go first through the separate inversions of f and h, followed by composition with $\mathcal{C}_\sigma = \mathcal{A}_\sigma^{-1}$ (see Figure 5). Instead, one can simply invert their composition, which is defined as the mapping $\varrho = \mathcal{A}_\sigma \circ f \circ h$ by

$$y = E\{z|y\} \longmapsto E\{f(\tilde{z})|y\} = \varrho(y),$$

and then apply ϱ^{-1} to $\mathbf{D}_{\text{ho}}(f(\tilde{z}))$ (19).

A.2. Ranges of $\tilde{y}$ and $\tilde{\sigma}$

Here, we investigate the most basic properties of the set $\tilde{Y}$. We remark that $\tilde{Y}$ can be equivalently considered to be the range of $\tilde{y}$, as well as the domain of the mapping $\tilde{y} \mapsto \tilde{\sigma}(\tilde{y})$. In particular, we are interested in providing sufficient conditions to ensure $\tilde{Y} = [0, 1]$.

We begin from three easy lemmas.

Lemma A.2.1. Let $y > \frac{1}{2}$ ($y < \frac{1}{2}$, or $y = \frac{1}{2}$), then $\tilde{y} > \frac{1}{2}$ (resp., $\tilde{y} < \frac{1}{2}$, or $\tilde{y} = \frac{1}{2}$).

Proof. Exploiting the symmetries of ϕ and Φ, it suffices to observe that if $y > \frac{1}{2}$ (resp., $y < \frac{1}{2}$, or $y = \frac{1}{2}$) then:

   (i) $\phi\left(\frac{\zeta - y}{\sigma(y)}\right) > \phi\left(\frac{1 - \zeta - y}{\sigma(y)}\right)$ (resp., $<$, or $=$) $\forall \zeta \in [0,1]$,

   (ii) $\Phi\left(\frac{y-1}{\sigma(y)}\right) > \Phi\left(\frac{-y}{\sigma(y)}\right)$ (resp., $<$, or $=$).

From (i) follows that the mean of the absolutely continuous part of the distribution of $\tilde{z}$ [i.e. the middle term in (6)] is larger than (resp., smaller than, or equal to) $\frac{1}{2}$; while (ii) implies that the mean of the two masses at 0 and 1 in (6) is larger than (resp., smaller than, or equal to) $\frac{1}{2}$. Hence, the mean of (6) is larger than (resp., smaller than, or equal to) $\frac{1}{2}$. $\square$

Lemma A.2.2. If $y \in \mathbb{R}$ then $\tilde{\sigma}(\tilde{y}) < \frac{1}{2}$.

Proof. The inequality $\tilde{\sigma}(\tilde{y}) \leq \frac{1}{2}$ holds regardless of the particular distribution of z. It suffices to take the definition of the variance and observe that $0 \leq \tilde{z}^2 \leq \tilde{z} \leq 1$. Hence, $0 \leq E\{\tilde{z}^2 | y\} \leq E\{\tilde{z}|y\} \leq 1$ and

$$\tilde{\sigma}^2(\tilde{y}) = E\{\tilde{z}^2|y\} - E^2\{\tilde{z}|y\} < E\{\tilde{z}^2|y\} - E^2\{\tilde{z}^2|y\} \leq \frac{1}{4}.$$

However, for a normally distributed z, unless $\sigma(y) = 0$, we always have $E\{\tilde{z}^2|y\} < E\{\tilde{z}|y\}$, from which follows the statement. $\square$

Remark A.2.3. In the proof of the above lemma, to have $E\{\tilde{z}^2|y\} = E\{\tilde{z}|y\}$ it is necessary that $\tilde{z} \in \{0, 1\}$ with probability 1. This happens if $\sigma(y) = 0$ and $y \notin (0, 1)$. Otherwise, for a normally distributed z, the two expectations $E\{\tilde{z}^2|y\}$ and $E\{\tilde{z}|y\}$ can be equal only asymptotically, as described by the next lemma.

Lemma A.2.4. Let $\bar{y} \in Y \cup \partial Y$ be a point in Y or in the boundary of Y, and assume $\frac{y}{\sigma(y)} \xrightarrow[y \to \bar{y}]{} \lambda \in \mathbb{R}$. Then, either of the three situations below holds:

   (a) $\sigma(y) \xrightarrow[y \to \bar{y}]{} +\infty$; then $\tilde{y} \xrightarrow[y \to \bar{y}]{} \Phi(\lambda)$, and $\tilde{\sigma}(\tilde{y}) \xrightarrow[y \to \bar{y}]{} \sqrt{\Phi(\lambda)\,\Phi(-\lambda)}$.

   (b) $\sigma(y) \xrightarrow[y \to \bar{y}]{} \varsigma \in [0, +\infty)$; then $\bar{y} \in \mathbb{R}$, and $\tilde{y} \xrightarrow[y \to \bar{y}]{} \mathcal{A}_\sigma(\bar{y})$, $\tilde{\sigma}(\tilde{y}) \xrightarrow[y \to \bar{y}]{} \mathcal{B}_\sigma(\bar{y})$, where $\sigma(\bar{y})$ in (7),(8) is defined as ς (continuous extension).

   (c) $\lim_{y \to \bar{y}} \sigma(y)$ does not exist; then $\lambda = \bar{y} = 0$ and (i) or (ii), respectively, can be applied for sequences y_k such that $\lim_{y_k \to \bar{y}} \sigma(y_k)$ equals $+\infty$ or $\varsigma \in [0, +\infty)$.

Proof. We first note that these three situations are exhaustive and we remark that, if $\lim_{y \to \bar{y}} \sigma(y)$ exists, then $\frac{y-1}{\sigma(y)} \xrightarrow[y \to \bar{y}]{} \lambda - \lim_{y \to \bar{y}} \frac{1}{\sigma(y)}$. To prove (a), we can exploit the following first-order expansions for a small $\frac{1}{\sigma(y)}$,

$$\Phi\left(\tfrac{y-1}{\sigma(y)}\right) = \Phi\left(\tfrac{y}{\sigma(y)}\right) - \phi\left(\tfrac{y}{\sigma(y)}\right)\tfrac{1}{\sigma(y)} + o\left(\tfrac{1}{\sigma(y)}\right),$$
$$\phi\left(\tfrac{y-1}{\sigma(y)}\right) = \phi\left(\tfrac{y}{\sigma(y)}\right) + \tfrac{y}{\sigma(y)}\phi\left(\tfrac{y}{\sigma(y)}\right)\tfrac{1}{\sigma(y)} + o\left(\tfrac{1}{\sigma(y)}\right).$$

Substituting these into (7) yields

$$\tilde{y} = \phi\left(\tfrac{y}{\sigma(y)}\right)\tfrac{y}{\sigma(y)} + o\left(\tfrac{1}{\sigma(y)}\right)y + \Phi\left(\tfrac{y-1}{\sigma(y)}\right) +$$
$$- \tfrac{y}{\sigma(y)}\phi\left(\tfrac{y}{\sigma(y)}\right) + o\left(\tfrac{1}{\sigma(y)}\right)\sigma(y) =$$
$$= \Phi\left(\tfrac{y}{\sigma(y)}\right) + o(1) + o\left(\tfrac{1}{\sigma(y)}\right)y. \quad (29)$$

By letting $\sigma(y) \xrightarrow[y \to \bar{y}]{} \infty$, we obtain $\tilde{y} \xrightarrow[y \to \bar{y}]{} \Phi(\lambda)$. This result can be obtained also by replacing $\frac{y}{\sigma(y)}$ and $\frac{1}{\sigma(y)}$ in (6) with their limiting values, which gives the asymptotic p.d.f. as $\Phi(-\lambda)\,\delta_0(\zeta) + \Phi(\lambda)\,\delta_0(1 - \zeta)$ and, thus (because of bounded convergence), the limiting value of $\tilde{y}$ as $\Phi(\lambda)$. Similar substitutions into (8) give the limit variance as $\tilde{\sigma}^2(\tilde{y}) \xrightarrow[y \to \bar{y}]{} \Phi^2(\lambda)\,(1 - \Phi(\lambda)) + (1 - \Phi(\lambda))^2\,\Phi(\lambda) = \Phi(\lambda)\,\Phi(-\lambda)$. Hence, we see that the asymptotic envelope of the curve $(\tilde{y}, \tilde{\sigma}(\tilde{y}))$ is the semicircle of radius $\frac{1}{2}$ centered at $\left(\frac{1}{2}, 0\right)$ given by

$$\left(\Phi(\lambda),\, \sqrt{\Phi(\lambda)\,\Phi(-\lambda)}\right), \quad \lambda \in \mathbb{R}. \quad (30)$$

Under any circumstances, it is not possible for the graph of $(\tilde{y}, \tilde{\sigma}(\tilde{y}))$ to be outside of this envelope, simply because this already corresponds to the case of an arbitrary two-impulse distribution supported on $\{0, 1\}$ (thus giving the highest variance among all distributions supported on $[0, 1]$ with the same mean). Proving (b) is equivalent to proving the convergence of (7),(8), which indeed follows easily from the continuity of Φ and ϕ and from the existence of the finite limits of y, $\sigma(y)$, $\frac{y}{\sigma(y)}$, and $\frac{y-1}{\sigma(y)}$. We can now turn to (c). Since we assume $\frac{y}{\sigma(y)} \xrightarrow[y \to \bar{y}]{} \lambda$, $\lim_{y \to \bar{y}} \sigma(y)$ may not exist only if $\lambda = 0$; otherwise, if $\lambda \neq 0$, either $|\bar{y}| = \infty$ and hence $\lim_{y \to \bar{y}} \sigma(y) = +\infty$, or $\bar{y} \in \mathbb{R}$ and $\lim_{y \to \bar{y}} \sigma(y) = \bar{y}/\lambda$. Likewise, if $\lambda = 0$ and $\bar{y} \neq 0$, we deduce that $\lim_{y \to \bar{y}} \sigma(y) = +\infty$, which completes the proof. $\square$

Lemma A.2.5. Let $\bar{y} \in Y \cup \partial Y$ be a point in Y or in the boundary of Y, and assume $\frac{y}{\sigma(y)} \xrightarrow[y \to \bar{y}]{} +\infty$. Then, either of the two situations below holds:

   (d) $\bar{y} \in [0, +\infty)$; then $\tilde{y} \xrightarrow[y \to \bar{y}]{} \min\{1, \bar{y}\}$, and $\tilde{\sigma}(\tilde{y}) \xrightarrow[y \to \bar{y}]{} 0$.

   (e) $\bar{y} = +\infty$; then $\tilde{y} \xrightarrow[y \to \bar{y}]{} 1$ and $\tilde{\sigma}(\tilde{y}) \xrightarrow[y \to \bar{y}]{} 0$.

Specular situations are obtained if we assume $\frac{y}{\sigma(y)} \xrightarrow[y \to \bar{y}]{} -\infty$.

Proof. In case (d), we have necessarily $\sigma(y) \xrightarrow[y \to \bar{y}]{} 0$ and hence the two limits follow easily from the asymptotic p.d.f. (6) given as $\delta_0(\min\{1, \bar{y}\} - \zeta)$. To prove (e) we first observe that if $y \to +\infty$ then also $\frac{y-1}{\sigma(y)} \xrightarrow[y \to \bar{y}]{} +\infty$. Thus, the asymptotic p.d.f. (6) becomes $\delta_0(1 - \zeta)$, which yields the statement. The case $\frac{y}{\sigma(y)} \xrightarrow[y \to \bar{y}]{} -\infty$ is handled analogously. $\square$

Proposition A.2.6. Let $\sigma(y) = \sqrt{ay + b}$. Then, depending on the values of a and b, the sets Y and $\tilde{Y}$ are as follows:

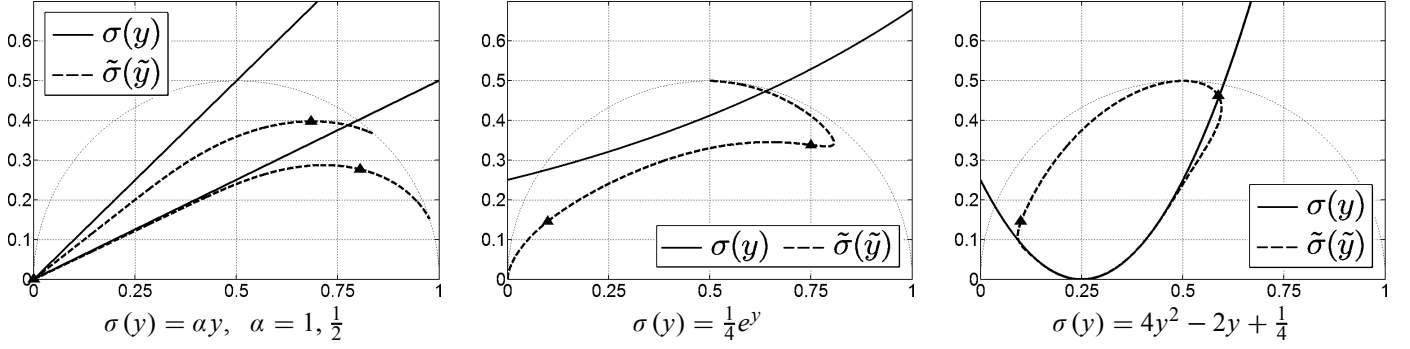   (i) $a = 0$, $b \geq 0$, $Y = \mathbb{R}$, $\tilde{Y} = (0, 1)$;

Figure 26: Four examples of standard-deviation functions σ for which $\sup \tilde{Y} < 1$ despite $\sup Y = +\infty$. In each plot, the asymptotic envelope semicircle (30) is also drawn as a thin dotted line.

(ii) $a > 0, -\frac{b}{a} \leq 0, \ Y = \left[-\frac{b}{a}, +\infty\right), \ \tilde{Y} = [0, 1)$;

(iii) $a > 0, 0 < -\frac{b}{a} \leq \frac{1}{2}, \ Y = \left[-\frac{b}{a}, +\infty\right), \ \tilde{Y} = \left[-\frac{b}{a}, 1\right)$;

(iv) $a > 0, \frac{1}{2} < -\frac{b}{a} < 1, \ Y = \left[-\frac{b}{a}, +\infty\right),$
$\tilde{Y} = [\gamma, 1), \frac{1}{2} < \gamma \leq -\frac{b}{a}$;

(v) $a > 0, 1 \leq -\frac{b}{a}, \ Y = \left[-\frac{b}{a}, +\infty\right),$
$\tilde{Y} = [\gamma, 1), \frac{1}{2} < \gamma < 1$.

Further, $\tilde{\sigma}$ can be extended with continuity to the closure of $\tilde{Y}$, by defining $\tilde{\sigma}\left(\inf \tilde{Y}\right) = \tilde{\sigma}\left(\sup \tilde{Y}\right) = 0$. Conditions specular to (ii)-(v) hold for $a < 0$.

Proof. These conditions follow directly from some of the results proved so far in this appendix. More precisely, Lemma A.2.5(e) is used for the supremum in all five cases, while for the lower bound we use: (i) Lemma A.2.5(e) (specular), (ii) Lemma A.2.5(e) (specular) (if $b > 0$) or Lemma A.2.4(b) (if $b = 0$), (iii) Lemma A.2.5(e) (specular) and Corollary A.1.8, (iv)-(v) Lemma A.2.5(e) (specular) and Lemma A.2.1. $\square$

The reader might refer to Figure 25 in order to better appreciate the meaning of γ and the role of Corollary A.1.8 and Lemma A.2.1 in the above proof.

Corollary A.2.7. Let $\sigma(y) = \sqrt{ay + b}$ and $Y \ni \frac{1}{2}$. If f satisfies the hypotheses of Corollary A.1.3, then the function h (21) maps the set $f(\tilde{Y})$ onto itself.

Proof. The hypotheses imply that both $f(E\{\tilde{z}|y\})$ and $E\{f(\tilde{z})|y\}$ are monotone functions of y. It remains to prove that $\sup f(\tilde{Y})$ and $\inf f(\tilde{Y})$ coincide with $f(\sup \tilde{Y})$ and $f(\inf \tilde{Y})$, respectively. Since $\tilde{\sigma} \to 0$ at both $\sup Y$ and $\inf Y$, we have that the asymptotic p.d.f. (6) approaches an impulse. Therefore, $f(E\{\tilde{z}|y\}) - E\{f(\tilde{z})|y\} \to 0$ as y approaches either $\sup Y$ or $\inf Y$. Due to the monotonicity of $f(E\{\tilde{z}|y\})$ and $E\{f(\tilde{z})|y\}$, $\tilde{y}$ approaches $\sup \tilde{Y}$ and $\inf \tilde{Y}$ as y approaches $\sup Y$ and $\inf Y$. $\square$

Figure 26 gives few examples where $\sup \tilde{Y} < 1$ despite $\sup Y = +\infty$. It is maybe surprising that even simple standard-deviation functions σ, such as those shown in the figure, can lead to rather exotic standard-deviation curves for the clipped variables. The linear $\sigma(y) = \alpha y$ is concave on its domain $Y = [0, \infty)$, therefore, from Corollary A.1.4, we deduce that $\tilde{y} = E\{\tilde{z}|y\}$ is a monotone nondecreasing function of y. It fol-

lows that

$$\tilde{Y} = \left[E\{\tilde{z}|y = 0\}, \lim_{y \to +\infty} E\{\tilde{z}|y\}\right] = \left[0, \Phi\left(\frac{1}{\alpha}\right)\right].$$

Thus, for the two examples in Figure 26(left), the upper bounds of $\tilde{Y}$ are $\Phi(1) = 0.841$ and $\Phi(2) = 0.977$, as visible in the plots. The remaining two examples use non-concave functions, so it is not possible to guarantee monotonicity of $E\{\tilde{z}|y\}$, and indeed they both yield multi-valued mappings $\tilde{y} \longmapsto \tilde{\sigma}$, as can be seen in the figure. Nevertheless, because of continuity, we necessarily have that $\inf \tilde{Y} \neq 0$ or $\sup \tilde{Y} \neq 1$, if any of the two limits $\lim_{y \to \inf Y} E\{\tilde{z}|y\}$ or $\lim_{y \to \sup Y} E\{\tilde{z}|y\}$ belongs to the open interval $(0, 1)$. In particular, for $\sigma(y) = \frac{1}{4}e^y$ we have $\lim_{y \to -\infty} E\{\tilde{z}|y\} = 0$ (because $\sigma(y) \underset{y \to -\infty}{\longrightarrow} 0$) and $\lim_{y \to +\infty} E\{\tilde{z}|y\} = \Phi(0) = \frac{1}{2}$, and hence $\left(0, \frac{1}{2}\right] \subsetneq \tilde{Y} \subsetneq [0, 1)$. Similar, for $\sigma(y) = 4y^2 - 2y + \frac{1}{4}$ the limit of $\frac{y}{\sigma(y)}$ is zero both for $y \to -\infty$ and for $y \to +\infty$, therefore $\frac{1}{2} = \Phi(0) \in \tilde{Y} \subsetneq (0, 1)$.

A.3. Boundedness of the variance-stabilizing transformation

Let us now discuss the convergence of the integral (18) (and thus the boundedness of f) when $\tilde{\sigma}(\tilde{y}) \to 0$.

A.3.1. Clipping from below

In the light of the approximations of Section 2.5, we first restrict to the case of clipping from below (e.g., we may assume $y + \sigma(y) \ll 1$), i.e. $\tilde{z} = \max\{0, z\}$, for which (14) and (15) reduce to

$$\tilde{y} = \sigma(y)\,\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right), \quad \tilde{\sigma}(\tilde{y}) = \sigma(y)\,\mathcal{S}_m\left(\frac{y}{\sigma(y)}\right). \quad (31)$$

Let $\mu = \frac{y}{\sigma(y)}$; we have

$$\tilde{\sigma}(\tilde{y}) = \sigma(y)\sqrt{\Phi(\mu) + \mathcal{E}_m(\mu)\,\mu - \mathcal{E}_m^2(\mu)}. \quad (32)$$

We give proofs of divergence and convergence for the four main cases which are of practical interest. For the proofs, we consider the asymptotic behavior of $\tilde{\sigma}(\tilde{y})$ (denoted by the symbol "$\asymp$") within neighborhoods of its points of zero, with the leading distinctions based on the limiting value of μ. Figure 27
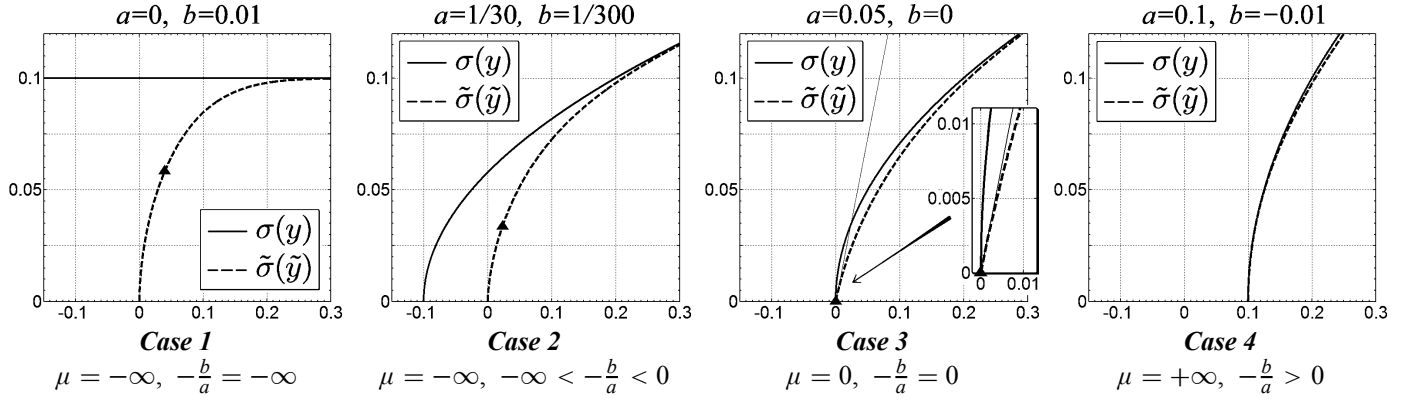
17

Figure 27: Illustration of the four cases considered by Propositions A.3.1 - A.3.5 for standard-deviation functions of the form $\sigma(y) = \sqrt{ay+b}$. In the plot corresponding to *Case 3*, observe the line tangent to $\tilde{\sigma}(\tilde{y})$ at $\tilde{y} = 0$; the equation of this line is given by the right-hand side of (33) with $\lambda = 0$.

illustrates these four cases with examples where the standard-deviation function has the form $\sigma(y) = \sqrt{ay+b}$ with $a \geq 0$.

Proposition A.3.1 (*Case 1*). Let $\tilde{\sigma}(\tilde{y}) \xrightarrow[\tilde{y}\to 0^+]{} 0$ and assume that exists some $\varepsilon > 0$ such that $\sigma(y) \geq \varepsilon$ as $y \to -\infty$. Then, the integral (18) converges at 0 and hence f is bounded.

Proof. Because of (31), the hypotheses imply that $\mu \xrightarrow[\tilde{y}\to 0^+]{} -\infty$ and hence that $y \xrightarrow[\tilde{y}\to 0^+]{} -\infty$. From (32), we obtain

$$\tilde{\sigma}(\tilde{y}) = \sigma(y)\sqrt{\mathcal{E}_m(\mu)\left(\frac{\Phi(\mu)}{\mathcal{E}_m(\mu)} + \mu - \mathcal{E}_m(\mu)\right)} =$$

$$= \sigma(y)\sqrt{\mathcal{E}_m(\mu)\left(\left(\frac{\Phi(\mu)\mu+\phi(\mu)}{\Phi(\mu)}\right)^{-1} + \mu - \mathcal{E}_m(\mu)\right)} =$$

$$= \sigma(y)\sqrt{\mathcal{E}_m(\mu)\left(\left[(\mu + R^{-1}(-\mu))^{-1} + \mu\right] - \mathcal{E}_m(\mu)\right)},$$

where R is the Mill's ratio, $R(\mu) = \frac{1-\Phi(\mu)}{\phi(\mu)}$. Exploiting a classical asymptotic expansion [17],

$$R(\mu) = \frac{1}{\mu} - \frac{1}{\mu^3} + \frac{1\cdot 3}{\mu^5} - \frac{1\cdot 3\cdot 5}{\mu^7} + \dots,$$

we can approximate the term in the square brackets above as $-\frac{2}{\mu} + O(\frac{1}{\mu^3})$. Moreover, $\mathcal{E}_m(\mu) \xrightarrow[\mu\to-\infty]{} 0$ at exponential rate and, in particular, much faster than $\left(-\frac{2}{\mu}\right)^2$. It means that $-\frac{2}{\mu} + O(\frac{1}{\mu^3})$ is much larger than $\sqrt{\mathcal{E}_m(\mu)}$. Hence, for $-\mu$ large enough, we have

$$\tilde{\sigma}(\tilde{y}) = \sigma(y)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)\left(-\frac{2}{\mu} + O(\mu^{-3})\right)} >$$

$$> \sigma(y)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)}} = \sigma^{\frac{1}{4}}(y)\,\tilde{y}^{\frac{3}{4}} \geq \varepsilon^{\frac{1}{4}}\tilde{y}^{\frac{3}{4}}.$$

This proves our result, because in a neighborhood of 0 the integrand in (18) is positive and bounded from above by $c\varepsilon^{-\frac{1}{4}}\tilde{y}^{-\frac{3}{4}}$, which has itself a convergent indefinite integral at 0. □

Proposition A.3.2 (*Case 2*). Let $\tilde{\sigma}(\tilde{y}) \xrightarrow[\tilde{y}\to 0^+]{} 0$ and $\sigma(y) \xrightarrow[y\to q^+]{} 0$ for some $q < 0$. Then, (18) converges at 0.

Proof. This is straightforward adaptation of Case 1. As in the proof above, we have $\mu \xrightarrow[\tilde{y}\to 0^+]{} -\infty$. Now, let us observe that $\sigma(y) = \frac{y}{\mu} \asymp \frac{q}{\mu}$ and thus that $\sigma(y) = \sigma^{\frac{3}{4}}(y)\sigma^{\frac{1}{4}}(y) \asymp \sigma^{\frac{3}{4}}(y)\left(\frac{q}{\mu}\right)^{\frac{1}{4}}$. Then, by the same arguments in the proof of Proposition A.3.1,

$$\tilde{\sigma}(\tilde{y}) = \sigma(y)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)\left(-\frac{2}{\mu} + O(\mu^{-3})\right)} =$$

$$= \sigma^{\frac{3}{4}}(y)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)O\left(\left(\frac{q}{\mu}\right)^{-1/2}\right)\left(-\frac{2}{\mu} + O(\mu^{-3})\right)} =$$

$$= \sigma^{\frac{3}{4}}(y)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)\frac{2}{\sqrt{-q}}O((-\mu)^{-3/2})} >$$

$$> \sigma^{\frac{3}{4}}(y)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)\sqrt{\mathcal{E}_m\left(\frac{y}{\sigma(y)}\right)}} = \tilde{y}^{\frac{3}{4}},$$

from which we deduce the convergence at 0 of the indefinite integral (18). □

Proposition A.3.3 (*Case 3*). Let both $\tilde{\sigma}(\tilde{y}) \xrightarrow[\tilde{y}\to 0^+]{} 0$ and $\sigma(y) \xrightarrow[y\to 0^+]{} 0$. If $\mu = \frac{y}{\sigma(y)} \xrightarrow[y\to 0^+]{} \lambda \in \mathbb{R} \cup \{+\infty\}$, then (18) diverges and thus f is unbounded. In particular, the divergence is logarithmical if $\lambda \in \mathbb{R}$ whereas it is logarithmical or faster if $\lambda = +\infty$.

Proof. From (31)-(32) and (12)-(13), we obtain

$$\tilde{\sigma}(\tilde{y}) = \tilde{y}\frac{\mathcal{S}_m(\mu)}{\mathcal{E}_m(\mu)} = \tilde{y}\sqrt{\frac{\Phi(\mu)}{\mathcal{E}_m^2(\mu)} + \frac{\mu}{\mathcal{E}_m(\mu)} - 1}.$$

Let first $\lambda \in \mathbb{R}$. Because of continuity and because both $\mathcal{E}_m(\lambda)$ and $\mathcal{S}_m(\lambda)$ are finite and strictly positive numbers,

$$\tilde{\sigma}(\tilde{y}) \asymp \tilde{y}\frac{\mathcal{S}_m(\lambda)}{\mathcal{E}_m(\lambda)} = \tilde{y}\sqrt{\frac{\Phi(\lambda)}{\mathcal{E}_m^2(\lambda)} + \frac{\lambda}{\mathcal{E}_m(\lambda)} - 1}, \quad (33)$$

which implies that (18) diverges logarithmically at 0. If $\lambda = +\infty$, it suffices to observe that $\mathcal{E}_m(\mu) > \mu$ and $\Phi(\mu) < 1$, from which we immediately obtain

$$\frac{\Phi(\mu)}{\mathcal{E}_m^2(\mu)} + \frac{\mu}{\mathcal{E}_m(\mu)} - 1 < \frac{1}{\mu^2}$$

18

and hence $\tilde{\sigma}(\tilde{y}) < \frac{\tilde{y}}{\mu}$. The statement follows because $\mu \to +\infty$ and $\tilde{\sigma}(\tilde{y}) \geq 0$. $\square$

Remark A.3.4. There is a synergy between the vanishing of $\sigma(y)$ and the clipping, which makes $\tilde{\sigma}(\tilde{y})$ to vanish even faster. An example is given by $\sigma(y) = \sqrt{ay}$, $a > 0$, for which $\mu \xrightarrow[y \to 0^+]{} 0$ and hence $\tilde{\sigma}(\tilde{y}) \asymp \tilde{y}\sqrt{\Phi(0) - \phi^2(0)}/\phi(0)$. This example is illustrated in Figure 27, for $a = 0.05$. It is interesting to note that while (18) diverges, the integral $\int_{t_0}^t \frac{1}{\sqrt{ay}} dy$ is convergent for $t \to 0^+$.

Whereas the above proposition and example demonstrate that the effect of clipping can be strong enough to change the character of convergence, the following propositions prove that clipping alone is not sufficient to cause divergence when $\sigma(0) > 0$ and that, roughly speaking, the effect of clipping vanishes as $\sigma(y) \to 0$ for y far from 0.

Proposition A.3.5 (Case 4). If $\sigma(y) \xrightarrow[y \to q]{} 0$ for some $q \in (0, 1)$, the boundedness of $\int_q^t \frac{1}{\tilde{\sigma}(\tilde{y})} d\tilde{y}$ is equivalent to the boundedness of $\int_q^t \frac{1}{\sigma(y)} dy$.

Proof. We have $\mu \to +\infty$ and simple majorization and minorizations give $\tilde{\sigma}(\tilde{y}) \asymp \sigma(y)$ and $\tilde{y} = y + o(|\tilde{y} - q|)$. It then follows that

$$\int_q^t \frac{1}{\tilde{\sigma}(\tilde{y})} d\tilde{y} = (1 + o(|t - q|)) \int_q^t \frac{1}{\sigma(y)} dy,$$

which completes the proof. $\square$

A.3.2. Clipping from above and below

Of course, all the above propositions, which are derived assuming $\tilde{z} = \max\{0, z\}$, can be immediately reformulated for the specular case of a variable singly clipped from above $\tilde{z} = \min\{1, z\}$. Let us now consider the complete case of doubly clipped variables $\tilde{z} = \max\{0, \min\{1, z\}\}$ and question the validity of the statements of Propositions A.3.1 - A.3.5 in this more complex scenario. Using a similar strategy as that used in the proof of Proposition A.3.5, we consider the differences Δ_E and Δ_σ between the expectations and the standard-deviations of doubly clipped and singly clipped variables:

$$E\{\max\{0, \min\{1, z\}\}|y\} - E\{\max\{0, z\}|y\} = \Delta_E(y) =$$
$$= \int_1^{+\infty} 1 \frac{1}{\sigma(y)} \phi\left(\frac{\zeta - y}{\sigma(y)}\right) d\zeta - \int_1^{+\infty} \zeta \frac{1}{\sigma(y)} \phi\left(\frac{\zeta - y}{\sigma(y)}\right) d\zeta =$$
$$= \int_1^{+\infty} \frac{1 - \zeta}{\sigma(y)} \phi\left(\frac{\zeta - y}{\sigma(y)}\right) d\zeta,$$

$$\text{std}\{\max\{0, \min\{1, z\}\}|y\} - \text{std}\{\max\{0, z\}|y\} = \Delta_\sigma(y) =$$
$$= \sqrt{\int_1^{+\infty} (1 - \tilde{y})^2 \frac{1}{\sigma(y)} \phi\left(\frac{\zeta - y}{\sigma(y)}\right) d\zeta +}$$
$$- \sqrt{\int_1^{+\infty} (\zeta - \tilde{y} + \Delta_E(y))^2 \frac{1}{\sigma(y)} \phi\left(\frac{\zeta - y}{\sigma(y)}\right) d\zeta}.$$

It is easy to realize that both $\Delta_E(y)$ and $\Delta_\sigma(y)$ vanish at exponential rate compared to $\tilde{y}$ and $\tilde{\sigma}(\tilde{y})$, provided either of the following two condition holds:

(i) $y \to q < 1$ and $\sigma(y) \to 0$;
(ii) $y \to -\infty$ and $\frac{y}{\sigma(y)} \to -\infty$.

Moreover, the conditions (i) and (ii) are alternatively satisfied by the cases considered in the above Propositions A.3.1 - A.3.6. Thus, the change of variables realized by introducing the compensating terms Δ_E and Δ_σ cannot influence the convergence or divergence results proved above.

Proposition A.3.6. Let $\sigma(y) = \sqrt{ay + b}$ with either $a \geq 0$ and $b > 0$, or $a < 0$ and $b > -a$. The integral (18) converges and f is bounded on $[0, 1]$.

Proof. If $a = 0$ and $b > 0$, the statement is a direct consequence of Proposition A.2.1, applied first to prove convergence at 0, and then (in specular form) to prove convergence at 1. If $a > 0$ and $b > 0$, the convergence at 0 is ensured by Proposition A.2.2, while Proposition A.2.1 is again used (in specular form) for the convergence at 1. The case $0 > a > -b$ is exactly specular to $a > 0$ and $b > 0$. $\square$

Remark A.3.7. For obvious reasons of numerical stability, it is preferable and recommended always to deal with bounded functions; therefore, in practice, for those cases (e.g., Case 3) where (18) may diverge, we can use $f(t) = \int_{t_0=0}^t \frac{c}{\max\{\tilde{\sigma}(\tilde{y}), \varepsilon\}} d\tilde{y}$, where $\varepsilon > 0$ is a small regularization constant. This also ensures that f is Lipschitz.

A.4. The case $b < 0 < a$

We conclude this appendix by considering the special case of $\sigma(y) = \sqrt{ay + b}$ with $a > 0$ and $-\frac{b}{a} \in \left(0, \frac{1}{2}\right]$. Such case is particularly relevant in practical applications with sensors whose output raw-data includes a manifest pedestal value (refer to [13] for details on how the sensor characteristics affect the two parameters of the Poissonian-Gaussian modeling).

First of all, due to Corollary A.1.8 and Proposition A.2.6 (iii), $\tilde{\sigma}$ can be well defined as a continuous function $\tilde{\sigma} : \left[-\frac{b}{a}, 1\right] \to [0, +\infty)$, with $\tilde{\sigma}(\tilde{y}) = 0$ if and only if $\tilde{y} \in \left\{-\frac{b}{a}, 1\right\}$. Proposition A.3.5 and Proposition A.3.1 (in specular form) ensure the boundedness of the integral

$$\int_{t_0}^t \frac{c}{\tilde{\sigma}(\tilde{y})} d\tilde{y}, \quad t, t_0 \in \left[-\frac{b}{a}, 1\right].$$

Note that this integral is defined only on $\left[-\frac{b}{a}, 1\right] \subsetneq [0, 1]$, while the range of $\tilde{z}$ for $y > -\frac{b}{a}$ is the whole closed interval $[0, 1]$. In order to construct a variance stabilizing transformation over $[0, 1]$, we follow the statement of Proposition A.1.1 and define f as

$$f(t) = \begin{cases} 0, & 0 \leq t < -\frac{b}{a}, \\ \int_{-\frac{b}{a}}^t \frac{c}{\tilde{\sigma}(\tilde{y})} d\tilde{y}, & -\frac{b}{a} \leq t \leq 1, \end{cases} \tag{34}$$

which, since f' is null outside the domain of σ, leads to strictly monotone $E\{f(\tilde{z})|y\}$, for $y \in \left[-\frac{b}{a}, +\infty\right)$. As observed in Remark A.1.10, while we certainly need f to be defined over the whole range $[0, 1]$ of $\tilde{z}$, for $E\{f(\tilde{z})|y\}$ to be strictly monotone it suffices that f is strictly monotone over $\tilde{Y}$, which is a subset of $[0, 1]$. Moreover, we are practically interested in the invertibility of f only as a function defined over $\tilde{Y}$. This is because the inversion is applied after denoising, after which we

assume to work with expectations, as declared by (19). Indeed, restricted on $\left[-\frac{b}{a}, 1\right]$, f is strictly monotone. It is then easy to verify that $h : f(\tilde{Y}) \to f(\tilde{Y})$ is invertible, since both $\tilde{y} = E\{\tilde{z}|y\}$ and $E\{f(\tilde{z})|y\}$, and thus also $f(E\{\tilde{z}|y\})$, are strictly monotone functions of y. Summarizing, all functions used for the stabilization, debiasing, inversion, and declipping can be properly defined and used also when $\tilde{Y} = \left[-\frac{b}{a}, 1\right]$ and $0 < -\frac{b}{a} \le \frac{1}{2}$, provided f is made constant equal to $f\left(-\frac{b}{a}\right) = 0$ on $\left[0, -\frac{b}{a}\right]$ as described by (34).

Of course, analogous considerations apply to the specular case where $a < 0$ and $-\frac{b}{a} \in \left[\frac{1}{2}, 1\right)$. We do not investigate further the cases where $a > 0$ and $-\frac{b}{a} > \frac{1}{2}$ or $a < 0$ and $-\frac{b}{a} < \frac{1}{2}$, since for these we are not able to guarantee monotonicity of $\tilde{y}$ with respect to y, as shown in Remarks A.1.5 and A.1.7.

References

[1] Anscombe, F.J., "The Transformation of Poisson, Binomial and Negative-Binomial Data", *Biometrika*, vol. 35, no. 3/4, pp. 246-254, 1948.

[2] Arsenault, H.H., and M. Denis, "Integral expression for transforming signal-dependent noise into signal-independent noise," *Opt. Lett.* vol. 6, no. 5, pp. 210-212, May 1981.

[3] Chen, M., J. Fridrich, M. Goljan, and J. Lukáš, "Determining Image Origin and Integrity Using Sensor Noise", *IEEE Trans. Infor. Secur. Forensics*, vol. 3, no. 1, pp. 74-90, March 2008.

[4] Cohen, A.C., *Truncated and Censored Samples: Theory and Applications*, CRC Press, 1991.

[5] Dabov, K., A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering", *IEEE Trans. Image Process.*, vol. 16, no. 8, Aug. 2007. MATLAB software available online at http://www.cs.tut.fi/~foi/GCF-BM3D

[6] Debevec, P.E., and J. Malik, "Recovering High Dynamic Range Radiance Maps from Photographs", *Proc. SIGGRAPH 97*, August 1997.

[7] Elad, M., and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries", *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736-3745, December 2006. MATLAB software available online at http://www.cs.technion.ac.il/~elad/Various/KSVD_Matlab_ToolBox.zip

[8] Foi, A., "Practical denoising of clipped or overexposed noisy images", *Proc. 16th Eur. Signal Process. Conf., EUSIPCO 2008*, Lausanne, August 2008.

[9] Foi, A., "Direct optimization of nonparametric variance-stabilizing transformations", presented at 8èmes Rencontres de Statistiques Mathématiques, CIRM, Luminy, December 2008.

[10] Foi, A., R. Bilcu, V. Katkovnik, and K. Egiazarian, "Adaptive-Size Block Transforms for Signal-Dependent Noise Removal", *Proc. 7th Nordic Signal Processing Symposium, NORSIG 2006*, Reykjavik, Iceland, June 2006.

[11] Foi, A., V. Katkovnik, and K. Egiazarian, "Pointwise Shape-Adaptive DCT for High-Quality Denoising and Deblocking of Grayscale and Color Images", *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1395-1411, May 2007. MATLAB software available online at http://www.cs.tut.fi/~foi/SA-DCT

[12] Foi, A., V. Katkovnik, and K. Egiazarian, "Signal-dependent noise removal in Pointwise Shape-Adaptive DCT domain with locally adaptive variance", *Proc. 15th Eur. Signal Process. Conf., EUSIPCO 2007*, Poznań, September 2007.

[13] Foi, A., 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, October 2008.

[14] Greene, W., *Econometric Analysis*, 4th ed., Prentice Hall, 2000.

[15] Johnson, N.L., S. Kotz, and N. Balakrishnan, *Continuous univariate distributions*, 2nd ed., Wiley, 1994.

[16] Kasturi, R., J.F. Walkup, and T.F. Krile, "Image restoration by transformation of signal-dependent noise to signal-independent noise," *Applied Optics*, vol. 22, no. 22, pp. 3537-3542, November 1983.

[17] Kendall, M., and A. Stuart, *The advanced theory of statistics*, *Volume 1*, 2nd ed., Charles Griffin and co., 1963.

[18] Hirakawa, K., and T.W. Parks, "Image denoising using total least squares", *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2730-2742, September 2006. MATLAB software available online at http://www.accidentalmark.com/research/packages/TLSdenoise.zip

[19] Lansel, S., D. Donoho, and T. Weissman, "DenoiseLab: A Standard Test Set and Evaluation Method to Compare Denoising Algorithms", benchmarks and MATLAB software online at http://www.stanford.edu/~slansel/DenoiseLab/.

[20] Portilla, J., V. Strela, M. Wainwright, and E.P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain", *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338-1351, November 2003. MATLAB software available online at http://decsai.ugr.es/~javier/denoise/software/BLS-GSM_Denoising.zip

[21] Prucnal, P.R., and B.E.A. Saleh, "Transformation of image-signal-dependent noise into image signal-independent noise", *Optics Letters*, vol. 6, no. 7, July 1981.

[22] Rudin, W., *Real and complex analysis*, third edition, McGraw-Hill, New York, 1987.

[23] Wang, L., L. Wei, K. Zhou, B. Guo and H. Shum, "High Dynamic Range Image Hallucination", *Proc. Eurographics Symposium on Rendering*, Grenoble, France, June, 2007.

[24] The USC-SIPI Image Database, University of Southern California, http://sipi.usc.edu/database/.