# Collaborative Filtering of Correlated Noise: Exact Transform-Domain Variance for Improved Shrinkage and Patch Matching

Ymir Mäkinen, Lucio Azzari, and Alessandro Foi

*Abstract*—Collaborative filters perform denoising through transform-domain shrinkage of a group of similar patches extracted from an image. Existing collaborative filters of stationary correlated noise have all used simple approximations of the transform noise power spectrum adopted from methods which do not employ patch grouping and instead operate on a single patch. We note the inaccuracies of these approximations and introduce a method for the exact computation of the noise power spectrum. Unlike earlier methods, the calculated noise variances are exact even when noise in one patch is correlated with noise in any of the other patches. We discuss the adoption of the exact noise power spectrum within shrinkage, in similarity testing (patch matching), and in aggregation. We also introduce effective approximations of the spectrum for faster computation. Extensive experiments support the proposed method over earlier crude approximations used by image denoising filters such as Block-Matching and 3D-filtering (BM3D), demonstrating dramatic improvement in many challenging conditions.

*Index Terms*—Image denoising, colored noise, correlated noise, collaborative filtering, BM3D

## I. INTRODUCTION

**T**RANSFORM-based denoising algorithms [1], [2] perform noise removal in a chosen domain where the signal to recover is sparse, i.e. it can be represented with a small amount of coefficients significantly different from zero. The shrinkage depends crucially on the transform-domain noise variance, which, in the case of stationary correlated noise, can be computed from the noise power spectral density (PSD) and the transform basis functions [3], [4].

Transform-based algorithms can be combined with the principles of nonlocal denoising (e.g., [5], [6]) to exploit the mutual similarity between patches at different locations in the image. BM3D [7] is one of the leading methods in this hybrid class known as *collaborative filters*. Mutually similar patches are jointly processed by applying first a 2-D transform $\mathcal{T}^{2D}$ to each patch and then a 1-D transform $\mathcal{T}^{NL}$ across the obtained $\mathcal{T}^{2D}$-spectra. This results in a 3-D transform $\mathcal{T}^{3D}$ that decorrelates both local and nonlocal image regularity. The advantage of collaborative filtering lies in the enhanced sparsity in this $\mathcal{T}^{3D}$ domain where shrinkage is performed. However, the effectiveness of denoising hinges on a correctly set shrinkage threshold, which in turn requires knowledge of the noise variance in this transform domain.

To model the $\mathcal{T}^{3D}$ noise power spectrum, [7] and subsequent works, such as the BM3D filter for correlated noise [8], have adopted a simplified modelling borrowed from local filters like [4], [9], so that the PSD in $\mathcal{T}^{3D}$ domain is calculated by merely replicating the $\mathcal{T}^{2D}$ PSDs. However, this model presumes $\mathcal{T}^{NL}$ is orthonormal and, most importantly, that noise in one patch is always independent from that in any other patch. The latter requirement is often not fulfilled. As noted in [7] for i.i.d. noise, noise correlation between patches may occur due to their overlap. Furthermore, with stationary correlated noise, noise may be correlated across different patches even if they do not overlap, potentially creating large inaccuracies in the simplified approximations.

Correlated noise has broad relevance in many realistic scenarios. In x-ray medical imaging, for example, noise affecting the projections is commonly assumed stationary correlated [10]–[16]. Thermal cameras are also affected by spatially correlated noise and a significant component of fixed-pattern noise which can be considered as stationary correlated in time [17]–[19]. Furthermore, satellite and remote-sensing imagery, particularly those captured using push-broom sensors and radar, are affected by spatially correlated noise [20]–[22].

Although neural networks have recently achieved state-of-the-art performance in white noise denoising, neural network based methods for correlated noise denoising have not yet caught up. Their adaptivity is presently limited to varying levels of uncorrelated noise [23]–[25] and expensive retraining of the network is required especially in the case of noise with visible long-range correlation. The proposed method can instead adapt to varying correlation without any prior training and can thus be utilized within filters and iterative recovery schemes that require online adjustment of the noise model, such as [19], [26]. Furthermore, in terms of computational complexity, collaborative filters are still an order of magnitude less expensive than deep neural networks [27]–[29], making them preferable for various real-time and low-power applications (e.g., [30], [31]).

Our contributions are summarized as follows[1]:
- We introduce a method for the exact computation of the noise variance in transform domain in any number of

[1]This work is an extended version of the authors' paper [32]. Here, we extend the calculation of the transform-domain variances to any dimensionality. We elaborate further their use in shrinkage and aggregation, provide an interpretation for the previous results on patch-matching, and analyze the selection of algorithm parameters. We extend the set of experiments and consider the case of non-periodic noise, which better reflects realistic imaging. We also evaluate the refiltering procedure with other denoisers.

dimensions (Section III-C) as well as effective approximations for faster computation (Section IV-D).

- We embed the new variance calculation into the BM3D algorithm, where it can be used to improve shrinkage accuracy (Section IV-B), adjust patch matching (Sections III-D and IV-A), and to weigh patches in the aggregation of the denoised patches (Section IV-C).

- We discuss the inherent limitations of patch-based collaborative filtering of correlated noise, and how denoising results can be improved by applying a simple global Fourier thresholding on the denoising residual and refiltering (Section V).

- We demonstrate significant, sometimes dramatic improvements both in visual quality and PSNR by endowing BM3D with the proposed algorithmic contributions, and note that refiltering can restore further details in BM3D but also several other tested denoising algorithms (Section VII).

## II. PROBLEM FORMULATION

Let us consider a noisy observation $z : X \to \mathbb{R}$ of an unknown deterministic noise-free image $y$ corrupted by stationary additive correlated noise $\eta$

$$z(x) = y(x) + \eta(x), \quad x \in X, \tag{1}$$

where $x \in X \subset \mathbb{Z}^d$ is the coordinate in the finite $d$-dimensional regular image domain $X$, and

$$\eta = \nu \circledast g, \qquad \nu(\cdot) \sim \mathcal{N}(0, 1), \tag{2}$$

$\nu$ being zero-mean i.i.d. Gaussian noise with unit variance, and $g$ being a convolution kernel representing the spatial correlation of the noise. Since $\text{var}\{\nu\} = 1$, $\text{var}\{\eta\} = \|g\|_2^2$. An equivalent way of representing correlated noise is by its PSD $\Psi$:

$$\Psi = \text{E}\left\{|\mathscr{F}[\eta]|^2\right\} = \text{var}\left\{\mathscr{F}[\eta]\right\} = |X| \, |\mathscr{F}[g]|^2, \tag{3}$$

$\mathscr{F}$ being the $d$-dimensional Fourier transform. The goal of denoising is to estimate $y$ from $z$; we consider the case of non-blind denoising, assuming that $g$, or equivalently $\Psi$, is known. When $g$ is a scaled Dirac delta, (1)-(2) reduces to the additive white Gaussian noise (AWGN) model and $\Psi$ is constant. When $\Psi$ is markedly non-constant, the noise is often described as *colored*, as opposed to white, in analogy with optics.

Four examples of noise correlation as well as white noise are shown in Figure 1, illustrated through the kernel $g$, a noise realization $\eta$, and the PSD $\Psi$. The first kernel $g_w$ is a Dirac delta, resulting in white, uncorrelated noise. The horizontal kernel $g_1$ is representative of sensor crosstalk in digital image acquisition, $g_2$ is representative of band-pass noise, and $g_3$ is a line pattern noise common to analog videos [33]. The kernel $g_4$ realizes an example of pink noise, which is found in analog electronic devices due to resistor voltage fluctuations [34]. Definitions of the kernels $g_1$, $g_2$, $g_3$, and $g_4$ are given in Table I. We further define four kernels $g_5$, $g_6$, $g_7$, and $g_8$
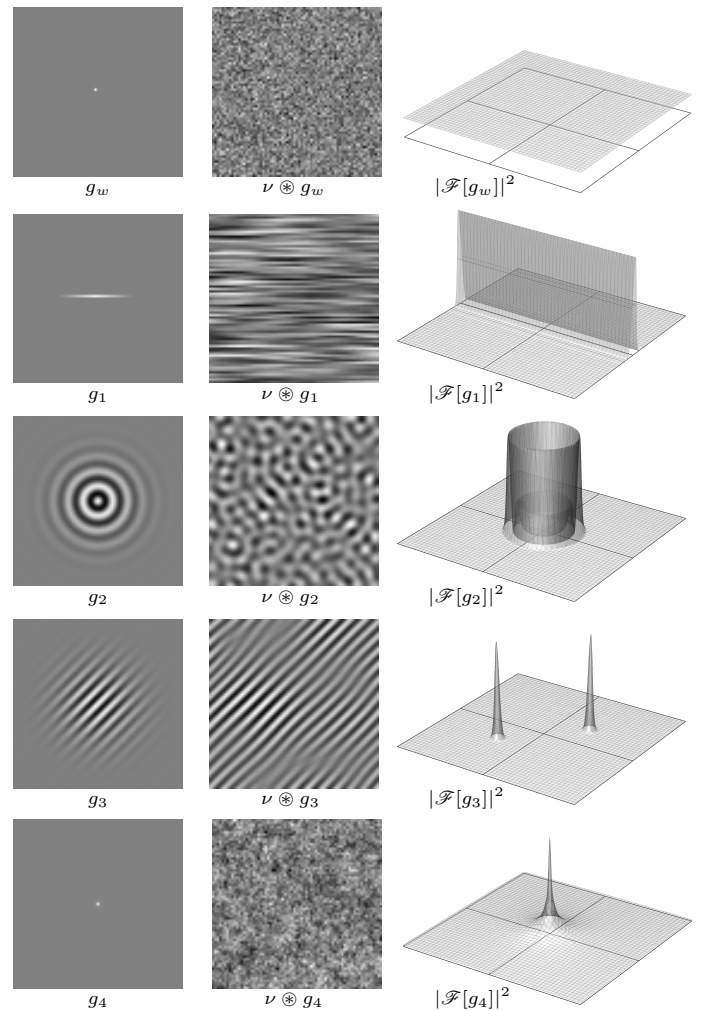


Figure 1. Left: correlation kernels (displayed on a $71 \times 71$-pixel canvas). Center: correlated noise fields obtained by convolving the kernels with white noise. Right: corresponding power spectral densities $\Psi$. As $g_w$ is a Dirac delta, $\nu \circledast g_w$ is white noise and the corresponding PSD is flat.

combining the flat white-noise PSD of $g_w$ with that of $g_n$, $n = 1, 2, 3, 4$, through

$$g_{n+4} = \mathscr{F}^{-1}\left[\sqrt{(1-\beta) + \beta |\mathscr{F}[g_n/\|g_n\|_2]|^2}\right], \tag{4}$$

with $\beta = 0.8$ determining the proportion of noise energy represented by $g_n$. Throughout the manuscript we refer to these examples (upon suitable rescaling to a desired $\text{var}\{\eta\} = \|g\|_2^2$) to discuss the properties of collaborative filtering of correlated noise and to assess the proposed model under challenging conditions.

## III. NOISE PSD OF NONLOCAL COLLABORATIVE TRANSFORMS

Collaborative filters operate on groups of similar patches extracted from the image. Let $\{z_{x_1}, \ldots, z_{x_M}\}$ be a group of $M$ $d$-dimensional patches extracted from $z$ at coordinates[2] $x_1, \ldots, x_M$, respectively, where each patch is composed of $N$ elements (e.g., pixels, voxels) and all patches in the group have

---

[2]As coordinate of a patch we intend the coordinate of its first element, which for a rectangular block is the pixel in the top-left corner. The patch coordinates are treated as deterministic.

Table I

REPRODUCIBLE MODELS FOR THE NOISE CORRELATION EXAMPLES IN
FIG. 1, WHERE $x^{(1)}$ AND $x^{(2)}$ DENOTE THE HORIZONTAL AND VERTICAL
INTEGER COORDINATES, AND $G_\varsigma$ DENOTES A GAUSSIAN FUNCTION WITH
STANDARD DEVIATION $\varsigma$ CENTERED AT THE ORIGIN. THE FIRST THREE
ARE DEFINED BY THEIR KERNEL, UPON SUITABLE RENORMALIZATION TO
A DESIRED VARIANCE LEVEL $\mathrm{var}\{\eta\} = \|g\|_2^2$. THE LAST IS DEFINED
THROUGH ITS PSD $|\mathscr{F}[g_4]|^2$ OF EQUAL SIZE TO THE IMAGE.

| | |
|---|---|
| $g_1$ | $\max\{0, 1 - |x^{(2)}|\}\max\{0, 16 - |x^{(1)}|\}$ |
| $g_2$ | $\cos\left(((x^{(1)})^2 + (x^{(2)})^2)^{1/2}\right)G_{10}(x^{(1)}, x^{(2)})$ |
| $g_3$ | $\cos(x^{(1)} + x^{(2)})G_{10}(x^{(1)}, x^{(2)})$ |
| $|\mathscr{F}[g_4]|^2$ | $\left(((x^{(1)})^2 + (x^{(2)})^2)^{1/2} + 10^{-2}|X|^{1/2}\right)^{-1}$ |

the same shape. Let $\mathcal{T}^{d\mathrm{D}}$ be a $d$-dimensional patch transform, and denote by $s_i^{x_t} = \langle z_{x_t}, b_i^{d\mathrm{D}} \rangle$ a generic $\mathcal{T}^{d\mathrm{D}}$-spectrum coefficient of $z_{x_t}$, where $b_i^{d\mathrm{D}}$ is the $i$-th basis function of $\mathcal{T}^{d\mathrm{D}}$. Further, we denote by $\{s_{i,j}^{x_1,\ldots,x_M}, i=1,\ldots,N, j=1,\ldots,M\}$ the $\mathcal{T}^{(d+1)\mathrm{D}}$ spectrum of the group $\{z_{x_1},\ldots,z_{x_M}\}$, computed by applying a 1-dimensional transform $\mathcal{T}^{\mathrm{NL}}$ of length $M$ to $[s_i^{x_1},\ldots,s_i^{x_M}]$, $i=1,\ldots,N$. Here, and throughout the rest of the manuscript, $i$ and $j$ index spectral components within the local and nonlocal dimensions of a group, respectively.

The core of this work is about the calculation and use of the variances $\mathrm{var}\{s_{i,j}^{x_1,\ldots,x_M}\}$ of the $\mathcal{T}^{(d+1)\mathrm{D}}$ spectrum $s_{i,j}^{x_1,\ldots,x_M}$, which we denote by $v_{i,j}^{x_1,\ldots,x_M}$. The indexing and notation of the transform spectrum coefficients are illustrated in Figure 2. Note the superscript making explicit the coordinates of the grouped patches, which play an important role in characterizing the variance, as demonstrated in what follows.

### A. Preliminaries

To calculate the noise variance of the $\mathcal{T}^{d\mathrm{D}}$ spectrum coefficients, we note that $s_i^{x_t} = \langle z_{x_t}, b_i^{d\mathrm{D}} \rangle = (z \circledast \overleftrightarrow{b}_i^{d\mathrm{D}})(x_t)$, where the $\overleftrightarrow{\phantom{x}}$ decoration denotes the reflection about the origin of $\mathbb{Z}^d$. Thus

$$\mathrm{var}\{s_i^{x_t}\} = \mathrm{var}\left\{\left(\nu \circledast g \circledast \overleftrightarrow{b}_i^{d\mathrm{D}}\right)(x_t)\right\} = \mathrm{var}\{\nu\}\left\|g \circledast \overleftrightarrow{b}_i^{d\mathrm{D}}\right\|_2^2,$$

which shows that this variance does not depend on the coordinate $x_t$ of the patch. Hence we can adopt the simple notation $v_i = \mathrm{var}\{s_i^{x_t}\}$ and since $\mathrm{var}\{\nu\} = 1$,

$$v_i = \left\|g \circledast \overleftrightarrow{b}_i^{d\mathrm{D}}\right\|_2^2 = \left\||X|^{-2}\Psi\left|\mathscr{F}\left[\overleftrightarrow{b}_i^{d\mathrm{D}}\right]\right|^2\right\|_1, \quad (5)$$

where the last equality follows from Parseval's isometry and (3).

As $\mathcal{T}^{(d+1)\mathrm{D}}$ is a separable transform, the $\mathcal{T}^{(d+1)\mathrm{D}}$-spectrum coefficients are calculated through the direct tensor product of the $\mathcal{T}^{d\mathrm{D}}$ and $\mathcal{T}^{\mathrm{NL}}$ transforms, as

$$s_{i,j}^{x_1,\ldots,x_M} = \left\langle [z_{x_1}; \cdots; z_{x_M}], b_i^{d\mathrm{D}} \otimes b_j^{\mathrm{NL}} \right\rangle =$$
$$= \left\langle [s_i^{x_1}, \cdots, s_i^{x_M}], b_j^{\mathrm{NL}} \right\rangle = \sum_{t=1}^M b_j^{\mathrm{NL}}(t)s_i^{x_t}, \quad (6)$$

where $b_j^{\mathrm{NL}}(t)$ is the $t$-th element of the $j$-th basis function $b_j^{\mathrm{NL}}$ of $\mathcal{T}^{\mathrm{NL}}$, and $[\cdot\,;\cdots;\cdot]$ denotes the stacking along the $(d+1)$-th dimension.
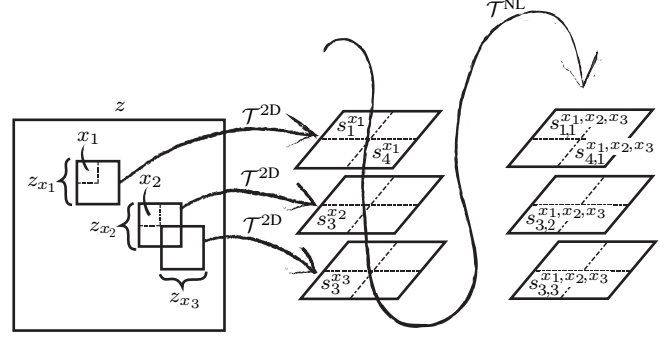


Figure 2. Notation and indexing of patch coordinates $x_k$, patches $z_{x_k}$, and coefficients $s_i^{x_k}$ and $s_{i,j}^{x_1,\ldots,x_M}$ in the corresponding $\mathcal{T}^{d\mathrm{D}}$ and $\mathcal{T}^{(d+1)\mathrm{D}}$ spectra. The illustration is for a group of three blocks of size $2\times2$ at coordinates $x_1 = (4,3)$, $x_2 = (7,5)$, $x_3 = (8,6)$ within a $10\times10$-pixel image ($d=2$). The variances of $s_i^{x_k}$, for any $x_k$, are denoted by $v_i$, whereas the variances of $s_{i,j}^{x_1,\ldots,x_M}$ are denoted by $v_{i,j}^{x_1,\ldots,x_M}$.

The exact calculation of the variance of the spectrum coefficients may not be immediate from (6). The variance of the spectrum coefficients is

$$v_{i,j}^{x_1,\ldots,x_M} = \mathrm{var}\left\{\sum_{t=1}^M b_j^{\mathrm{NL}}(t)s_i^{x_t}\right\} =$$
$$= \sum_{t=1}^M \left(b_j^{\mathrm{NL}}(t)\right)^2 \mathrm{var}\{s_i^{x_t}\} + \sum_{k\neq t} b_j^{\mathrm{NL}}(k)b_j^{\mathrm{NL}}(t)\,\mathrm{cov}\{s_i^{x_t}, s_i^{x_k}\} =$$
$$= v_i \sum_{t=1}^M \left(b_j^{\mathrm{NL}}(t)\right)^2 + \sum_{k\neq t} b_j^{\mathrm{NL}}(k)b_j^{\mathrm{NL}}(t)\,\mathrm{cov}\{s_i^{x_t}, s_i^{x_k}\}, \quad (7)$$

where $\mathrm{cov}\{s_i^{x_t}, s_i^{x_k}\}$ is the covariance between same $\mathcal{T}^{d\mathrm{D}}$ spectrum coefficients for different patches.

### B. Conventional methods for approximating $v_{i,j}^{x_1,\ldots,x_M}$

The common simplifying assumption (e.g., [7], [8], [35]) is the independence of the noise between stacked patches; in other words, noise may be correlated within each patch but not across distinct patches. Thus, the covariance terms vanish from (7). Under the further assumption that $\mathcal{T}^{\mathrm{NL}}$ is orthonormal, (7) simplifies to (5):

$$v_{i,j}^{x_1,\ldots,x_M} \approx v_i \sum_{t=1}^M \left(b_j^{\mathrm{NL}}(t)\right)^2 = v_i, \quad (8)$$

meaning that $\mathcal{T}^{(d+1)\mathrm{D}}$ spectrum variances are identical to those of $\mathcal{T}^{d\mathrm{D}}$ and become independent of the patch coordinates.

The approximation (8) is used by BM3D and the cited derivative works[3]. However, the requirement of independence between patches is not always fulfilled: it obviously fails if the patches are overlapping, but when noise is correlated it may fail even if they do not overlap. The failure of this requirement results in potentially significant imprecision in the calculation of the variances.

---

[3]Notably RF3D video denoiser [19] interpolates the spectrum variances under fixed-pattern noise from the exact variances computed for two extreme cases of idealized alignment of the blocks, as further discussed in Section VIII.
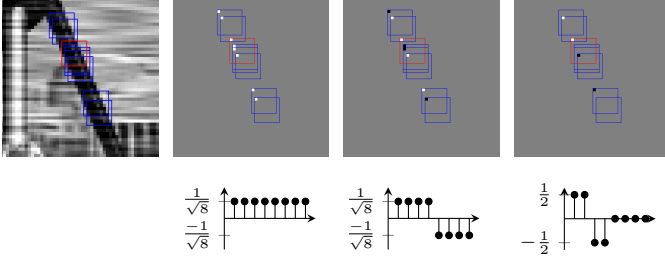
Figure 3. A noisy image and a group of 8 blocks (reference block in red), and the corresponding $\tilde{b}_j^{\mathrm{NL}}$ for $j = 1, 2, 3$, as well as the respective $b_j^{\mathrm{NL}}$ where $\mathcal{T}^{\mathrm{NL}}$ is the Haar transform. Gray in the $\tilde{b}_j^{\mathrm{NL}}$ images indicates the zero level.

## C. Exact calculation of $v_{i,j}^{x_1,\ldots,x_M}$

We observe that $s_i^{x_t}$ can equivalently be written as[4]

$$s_i^{x_t} = \left\langle z, b_i^{d\mathrm{D}} \circledast \delta_{x_t} \right\rangle, \tag{9}$$

where $\delta_{x_t}$ is a Dirac delta at the coordinate $x_t$. The coefficient $s_{i,j}^{x_1,\ldots,x_M}$ is thus computed as

$$s_{i,j}^{x_1,\ldots,x_M} = \sum_{t=1}^{M} b_j^{\mathrm{NL}}(t) \left\langle z, b_i^{d\mathrm{D}} \circledast \delta_{x_t} \right\rangle = \tag{10}$$

$$= \sum_{t=1}^{M} \left\langle z, b_i^{d\mathrm{D}} \circledast \left( b_j^{\mathrm{NL}}(t) \, \delta_{x_t} \right) \right\rangle. \tag{11}$$

The sum in (11) can be finally seen as the convolution [36]

$$s_{i,j}^{x_1,\ldots,x_M} = \left( \overleftrightarrow{z} \circledast b_i^{d\mathrm{D}} \circledast \tilde{b}_j^{\mathrm{NL}} \right)(0), \tag{12}$$

where $\tilde{b}_j^{\mathrm{NL}} = \sum_{t=1}^{M} b_j^{\mathrm{NL}}(t) \, \delta_{x_t}$ is an array of the same size of $z$ that is zero everywhere except at the coordinates $x_t$ where it assumes the corresponding values $b_j^{\mathrm{NL}}(t)$ (see Figure 3). Even though (12) is arithmetically identical to (6), it physically embeds the spatial locations $x_1, \ldots, x_M$ that (6) had lost through the stacking.

If we assume $z$ as in (1), the variance of the generic spectrum coefficient $s_{i,j}^{x_1,\ldots,x_M}$ computed from a group of blocks extracted from $z$ is

$$v_{i,j}^{x_1,\ldots,x_M} = \mathrm{var}\left\{ s_{i,j}^{x_1,\ldots,x_M} \right\} = \mathrm{var}\left\{ (\overleftrightarrow{z} \circledast b_i^{d\mathrm{D}} \circledast \tilde{b}_j^{\mathrm{NL}})(0) \right\} =$$

$$= \mathrm{var}\left\{ (\nu \circledast \overleftrightarrow{g} \circledast b_i^{d\mathrm{D}} \circledast \tilde{b}_j^{\mathrm{NL}})(0) \right\} =$$

$$= \left\| \overleftrightarrow{g} \circledast b_i^{d\mathrm{D}} \circledast \tilde{b}_j^{\mathrm{NL}} \right\|_2^2 = \tag{13}$$

$$= \left\| |X|^{-2} \, \Psi \left| \mathscr{F}\left[ b_i^{d\mathrm{D}} \right] \right|^2 \left| \mathscr{F}\left[ \tilde{b}_j^{\mathrm{NL}} \right] \right|^2 \right\|_1. \tag{14}$$

Note how (13) incorporates a convolution against $\tilde{b}_j^{\mathrm{NL}}$ which is instead completely missing from (5); as $\tilde{b}_j^{\mathrm{NL}}$ varies with the relative displacement of the patches, so do (13)-(14), as opposed to (5) and (8) which are the same for all groups.

The above procedure is directly applicable to nonlocal collaborative transforms used by filters for arbitrary $d$-dimensional data, for instance by BM3D for filtering images ($d=2$) and by BM4D [37] for filtering volumetric data ($d=3$). Note that, unlike (8), (14) does not presume anything about orthogonality or normalization of the $\mathcal{T}^{d\mathrm{D}}$ and $\mathcal{T}^{\mathrm{NL}}$ transforms.

[4]The first element of $b_i^{d\mathrm{D}}$ is used as its origin in convolution.
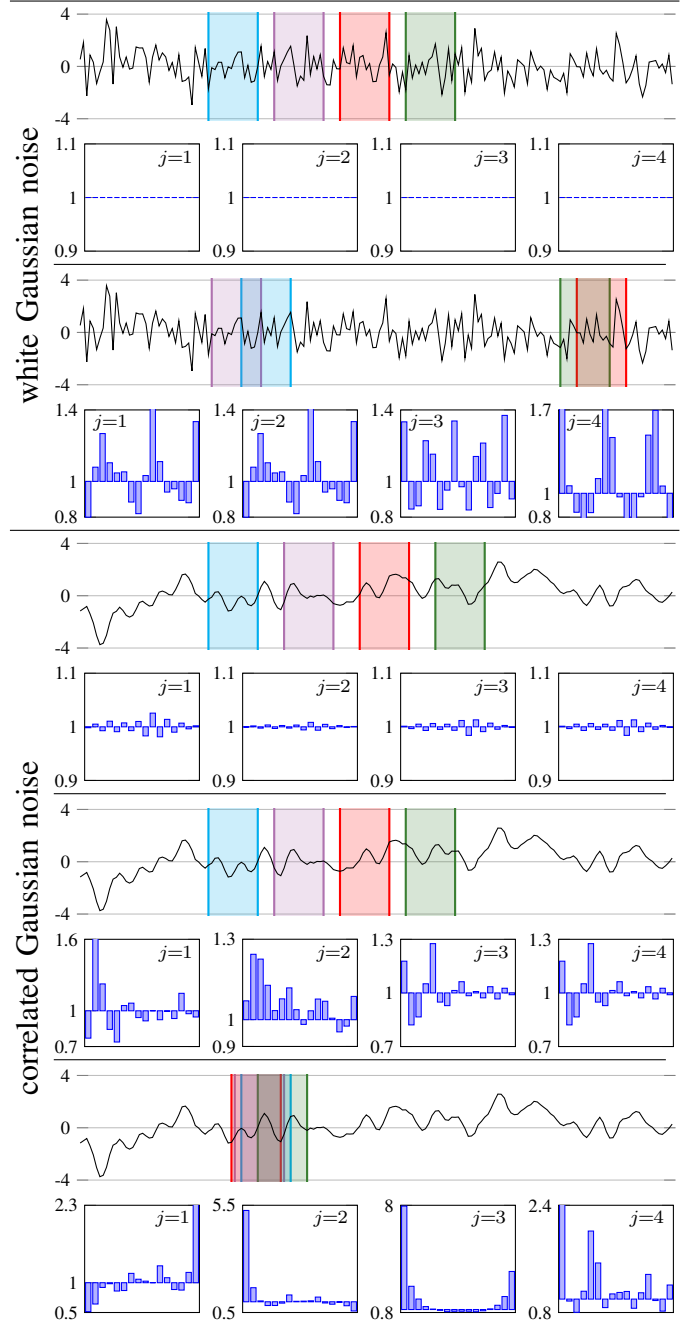


Figure 4. Comparison between $v_i$ and $v_{i,j}^{x_1,\ldots,x_M}$. Each group of plots shows a 1-D noise signal (black line) with unit variance from which we extract $M=4$ segments of $N=16$ samples (colored shaded areas), and 4 bar-plots that represent the ratio $\sqrt{v_i/v_{i,j}^{x_1,\ldots,x_M}}$, with $j=1,\ldots,4$. From top to bottom we have: white Gaussian noise with 1) non-overlapping and 2) overlapping segments; correlated Gaussian noise with 3) non-overlapping distant segments, 4) non-overlapping close segments, and 5) overlapping segments. Only for the topmost case $v_i$ is exact, while in general it is very different from $v_{i,j}^{x_1,\ldots,x_M}$.

To appreciate the difference between (14) and (8) we report a set of 1-D examples in which we compare the conventional expression against the real standard deviation of the coefficients. In Figure 4 we report two scenarios, one in which we consider a 1-D signal (the black line) composed of white i.i.d. noise and one composed of correlated noise. For each signal we extract 4 non-overlapping and 4 overlapping

segments of length 16 (the colored segments), that will each constitute a group. We then compute the transform of each group of segments using two 1-D orthonormal transforms (DCT of each segment followed by the Haar transform applied across the segments), equivalent to $\mathcal{T}^{\mathrm{NL}}$ and $\mathcal{T}^{d\mathrm{D}}$ in the $d$-D case. Finally, in the bar-plots we report each row $j$ of $\sqrt{v_i/v_{i,j}^{x_1,\dots,x_M}}$, that is, the ratio between the standard deviations of the noise spectrum computed using (8) and (14), respectively. Note how the conventional formula (8) is exact only in the first case at the top of the figure, where we have non-overlapping segments of i.i.d. noise.

### D. Patch matching

Nonlocal methods compare patches around a moving reference patch at $x_R$. Each potential patch at $x_j$ is evaluated through an ordering function $L_{x_R}(x_j)$ in order to select the best $M$ (possibly overlapping) patches to construct the group $\{z_{x_1}, \dots, z_{x_M}\}$. In practice, the goal is often to find the patches which are most similar to the reference patch in terms of the underlying noise-free content. As $L$ operates on noisy patches, the difference between noise-free content can only be estimated or approximated. One may use directly the $\ell^2$-norm of the difference of the noisy patches

$$L_{x_R}(x_j) = \left\| z_{x_R} - z_{x_j} \right\|_2^2, \tag{15}$$

difference of their spectra $L_{x_R}(x_j) = \left\| s_i^{x_R} - s_i^{x_j} \right\|_2^2$, or other, non-Euclidean distance measures [38]. To reduce the effects of noise, many methods (e.g., [8], [39]) include a scaling of the difference by the respective standard deviations of the coefficients such that

$$L_{x_R}(x_j) = \left\| \frac{s_i^{x_R} - s_i^{x_j}}{\sqrt{v_i}} \right\|_2^2. \tag{16}$$

This aims to give less weight to those coefficients which are presumed to be particularly noisy.

As illustrated by Figure 4, in the case of correlated noise, noise in nearby patches may be strongly correlated with noise in the reference patch. In this case, ranking provided by the above equations may be problematic, as correlation may guide the matching to prioritize patches where the noise is the most similar. It is thus useful to consider the variances between the two patches compared in the patch matching phase to consider the effects of correlation.

The difference between two noisy patches $z_{x_R}$ and $z_{x_j}$ can be written as

$$\left\| z_{x_R} - z_{x_j} \right\|_2^2 = 2 \sum_{i=1}^N \left\langle [s_i^{x_R}, s_i^{x_j}], b_2^{\mathrm{NL}} \right\rangle^2 = 2 \sum_{i=1}^N (s_{i,2}^{x_R,x_j})^2, \tag{17}$$

where $b_2^{\mathrm{NL}} = \frac{1}{\sqrt{2}}[1,-1]$ and $s_i^{x_R}, s_i^{x_j}$ are spectra produced by an arbitrary orthonormal transform $\mathcal{T}^{d\mathrm{D}}$.

We note that $(s_{i,2}^{x_R,x_j})^2$ is a non-central chi-squared random variable with one degree of freedom and with mean and variance

$$\mathrm{E}\left\{ (s_{i,2}^{x_R,x_j})^2 \right\} = v_{i,2}^{x_R,x_j} + \mathrm{E}^2\left\{ s_{i,2}^{x_R,x_j} \right\}, \tag{18}$$

$$\mathrm{var}\left\{ (s_{i,2}^{x_R,x_j})^2 \right\} = 2(v_{i,2}^{x_R,x_j})^2 + 4v_{i,2}^{x_R,x_j} \mathrm{E}^2\left\{ s_{i,2}^{x_R,x_j} \right\}, \tag{19}$$
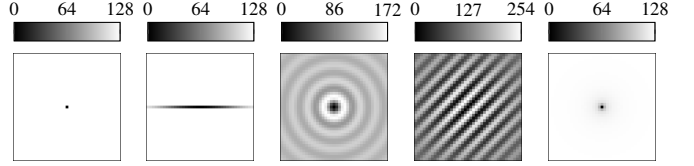


Figure 5. Left to right: maps of $\sum_{i=1}^N v_{i,2}^{x_R,x_j}$ as a function of $x_R - x_j$ for the kernels $g_w$, $g_1$, $g_2$, $g_3$, and $g_4$ shown in Figure 1 with $\|g\|_2^2 = 1$. Note how the center pixels are black as $v_{i,2}^{x_R,x_j} = 0$, $i = 1, \dots, N$, when $x_R = x_j$. The map is invariant to swapping $x_R$ with $x_j$ as seen from (14), where $|\mathscr{F}[\tilde{b}_j^{\mathrm{NL}}]|^2$ is invariant to translation and sign change.

where $v_{i,2}^{x_R,x_j}$ can be calculated with (14) for the corresponding $\mathcal{T}^{d\mathrm{D}}$ and $\mathcal{T}^{\mathrm{NL}}$ transforms. Noting that $2\sum_{i=1}^N \mathrm{E}^2\left\{ s_{i,2}^{x_R,x_j} \right\} = \left\| \mathrm{E}\left\{ z_{x_R} - z_{x_j} \right\} \right\|_2^2$ and by (18), we can express the expectation of (17) as

$$\mathrm{E}\left\{ \left\| z_{x_R} - z_{x_j} \right\|_2^2 \right\} = \left\| \mathrm{E}\left\{ z_{x_R} - z_{x_j} \right\} \right\|_2^2 + 2\sum_{i=1}^N v_{i,2}^{x_R,x_j}, \tag{20}$$

which quantifies the positive bias in the patch comparison and shows that this bias depends exclusively on the noise through $v_{i,2}^{x_R,x_j}$, $i, \dots, N$. The expectation can thus vary with the relative position between patches, i.e. with $x_R - x_j$, as illustrated in Figure 5. For the correlation kernels of this example, noise in patches closer to the reference patch or aligned in the direction of the correlation tends to correlate more with the reference patch, hence lower variances in the patch difference and consequently lower value of $2\sum_{i=1}^N v_{i,2}^{x_R,x_j}$.

By subtracting $2\sum_{i=1}^N v_{i,2}^{x_R,x_j}$ from $\left\| z_{x_R} - z_{x_j} \right\|_2^2$, we get an unbiased estimate of $\left\| \mathrm{E}\{ z_{x_R} - z_{x_j} \} \right\|_2^2$:

$$\mathrm{E}\left\{ \left\| z_{x_R} - z_{x_j} \right\|_2^2 - 2\sum_{i=1}^N v_{i,2}^{x_R,x_j} \right\} = \left\| \mathrm{E}\{ z_{x_R} - z_{x_j} \} \right\|_2^2, \tag{21}$$

which can be used to formulate a bias-corrected ordering function as

$$L_{x_R}(x_j) = \left\| z_{x_R} - z_{x_j} \right\|_2^2 - 2\sum_{i=1}^N v_{i,2}^{x_R,x_j}. \tag{22}$$

### IV. APPLICATION TO IMAGE DENOISING: BM3D

We demonstrate the use of the procedures described in Section III across various components of the BM3D denoising algorithm.

### A. Block matching for BM3D

The first step of BM3D consists of a block-matching procedure where, to produce a group of $M$ similar blocks, a noisy *reference* block is compared against all noisy blocks located within a surrounding finite region (search window). Ideally, the matched blocks would have similar underlying noise-free content, i.e. such that $\left\| \mathrm{E}\left\{ z_{x_R} - z_{x_j} \right\} \right\|_2^2$ is small. However, as shown by (20), this cannot be assessed accurately from a comparison of noisy blocks $\left\| z_{x_R} - z_{x_j} \right\|_2^2$.
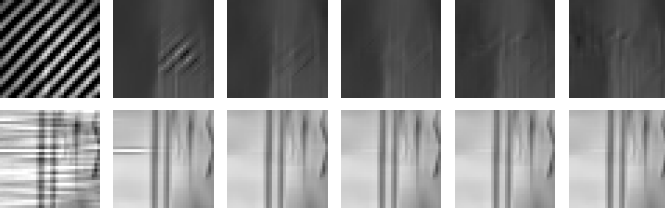
Figure 6. *Lena*'s hair with noise generated by kernel $g_3$, and a view of *Couple* with noise generated by $g_1$, both with $\|g\|_2^2 = 0.02$, and the respective denoising results for hard-thresholding with $\gamma = 0, 1, 2, 3, 4$.
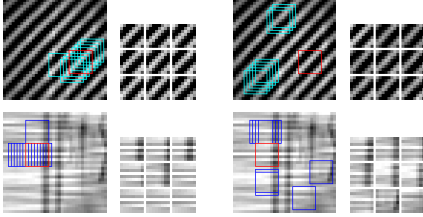


Figure 7. Demonstrations of the block matching process for the images of Figure 6 for $\gamma = 0$ (left) and $\gamma = 3$ (right), showing positions of the reference block and the 8 most similar matches as well as the contents of these 9 blocks. Note how in the case of $\gamma = 0$, blocks are mainly matched following a strong noise pattern, which is still seen in the denoised image shown in Figure 6.

Compared to only subtracting the bias as described by (21), our experiments indicate that the denoising quality is further improved by ranking potential matches according to

$$L_{x_R}(x_j) = \left\| z_{x_R} - z_{x_j} \right\|_2^2 - 2\gamma \sum_{i=1}^{N} v_{i,2}^{x_R, x_j}, \qquad (23)$$

with $\gamma > 1$ (instead of $\gamma = 1$ in (21)), as demonstrated in Figure 6 and Figure 7. All experiments reported in the remainder of this paper are produced with $\gamma = 3$. An interpretation for the increased denoising quality is provided in Section VIII-A, suggesting that $\gamma > 1$ facilitates the inclusion in the group of noisy blocks which differ from the noisy reference block mainly due to the variance of (17).

The common design of BM3D [7] includes a second denoising stage where the block-matching is carried over the denoised estimate of the first stage. Because this estimate can be treated as a noise-free approximation of the clean image, the second block-matching is performed without any bias subtraction, i.e. $\gamma = 0$.

### B. Shrinkage

The core of BM3D is shrinkage performed on the $\mathcal{T}^{(d+1)\mathrm{D}}$ spectrum of the grouped noisy blocks. For a given transform-domain coefficient of the group, a generic shrinkage can be expressed as

$$s_{i,j}^{x_1,\dots,x_M} \longmapsto \alpha_{i,j} s_{i,j}^{x_1,\dots,x_M}, \qquad (24)$$

where $\alpha_{i,j}$ is a shrinkage attenuation factor which depends on $s_{i,j}^{x_1,\dots,x_M}$, the noise statistics, and possible other priors. Various shrinkage functions can be used for this purpose. In particular, the common design of BM3D [7] uses hard-thresholding in the first denoising stage, followed by an empirical Wiener filter in the second stage.

In hard-thresholding, the shrinkage is performed by setting spectrum coefficients smaller than a threshold $\sqrt{v_{i,j}^{x_1,\dots,x_M}}\lambda$ to zero, as they are mostly composed of noise:

$$\alpha_{i,j}^{\mathrm{HT}} = \begin{cases} 1 & \text{if } \left| s_{i,j}^{x_1,\dots,x_M} \right| \geq \sqrt{v_{i,j}^{x_1,\dots,x_M}}\lambda \\ 0 & \text{otherwise,} \end{cases} \qquad (25)$$

where $\lambda \geq 0$ is a fixed constant. In Wiener filtering, the attenuation coefficients of the transfer function are computed from the previous estimate, used as pilot signal, and from the variance of the noise spectrum coefficients as

$$\alpha_{i,j}^{\mathrm{wie}} = \frac{\left\| \left\langle \left[ \hat{y}_{x_1}^{\mathrm{HT}}; \cdots; \hat{y}_{x_M}^{\mathrm{HT}} \right], b_i^{d\mathrm{D}} \otimes b_j^{\mathrm{NL}} \right\rangle \right\|^2}{\left\| \left\langle \left[ \hat{y}_{x_1}^{\mathrm{HT}}; \cdots; \hat{y}_{x_M}^{\mathrm{HT}} \right], b_i^{d\mathrm{D}} \otimes b_j^{\mathrm{NL}} \right\rangle \right\|^2 + \mu^2 v_{i,j}^{x_1,\dots,x_M}}, \qquad (26)$$

where $\hat{y}^{\mathrm{HT}}$ is the estimate of $y$ obtained from the hard-thresholding stage (note the similarity to (6)), and $\mu^2$ is a scaling factor included due to aggregation to influence the bias-variance ratio we wish to minimize through the Wiener filter.

As seen from both (25) and (26), both shrinkage operations depend crucially on the correct calculation of $v_{i,j}^{x_1,\dots,x_M}$.

### C. Aggregation

After calculating the attenuation factors of the group, they can be applied to the $\mathcal{T}^{(d+1)\mathrm{D}}$ spectra to obtain estimates for the grouped blocks:

$$\hat{y}_{x_j} = Q^{d\mathrm{D}} \left\{ \left\langle \alpha_{i,j}^{x_1,\dots,x_M} s_{i,j}^{x_1,\dots,x_M}, q_j^{\mathrm{NL}} \right\rangle \right\}, \qquad (27)$$

where $Q^{d\mathrm{D}}$ is the inverse transform of $\mathcal{T}^{d\mathrm{D}}$, and $q_j^{\mathrm{NL}}$ is the $j$-th transform basis function of the inverse of $\mathcal{T}^{\mathrm{NL}}$. We aggregate the obtained estimates into a buffer using aggregation weights, where blocks that have less residual noise get a larger weight. The general idea behind weighted aggregation is to improve quality of the final estimate and reduce the visibility of artifacts [40], [41]. Originally, BM3D used a unique weight to aggregate each block of a group $\{z_{x_1}, \dots, z_{x_M}\}$. The weight was computed from the inverse of the sum of the variances of the surviving coefficients in hard-thresholding, or from the inverse of the sum of the Wiener coefficients in Wiener filtering: a small weight for a group with large residual variance, and a large weight for a group with small residual variance.

With the proposed noise analysis we can calculate residual noise variance for blocks of the group $\{z_{x_1}, \dots, z_{x_M}\}$ and compute new, block-specific aggregation weights

$$w_j^{\mathrm{agg}} = \left( \sum_{i=1}^{N} \left\langle \left( v_{i,j}^{x_1,\dots,x_M} \alpha_{i,j}^{x_1,\dots,x_M} \right), \left| q_j^{\mathrm{NL}} \right|^2 \right\rangle \right)^{-1}, \qquad (28)$$

where $w_j^{\mathrm{agg}}$ is the aggregation weight of the block $j$, and $\alpha_{i,j}^{x_1,\dots,x_M}$ are the shrinkage attenuation factors, either for Wiener or hard-thresholding.

In other words, in (28) we first compute the sum of the residual variance within each block $j$ after shrinkage by applying an inverse transform to the residual of $v_{i,j}^{x_1,\dots,x_M}$, and then we invert the value to obtain the aggregation weight.

## D. Fast implementation

The increased complexity compared to the conventional methods for approximating the variances can be seen from the new global convolution in (13) or the multiplication and a Fourier transform in (14) compared to (5). Moreover, (14) has to be recomputed for each group, whereas (8) can be computed once at the beginning of the algorithm and reused for every group. Even using fast algorithms, computing either (13) or (14) for every group is not feasible for any reasonably sized image, as the global Fourier transform of $\mathscr{F}[\tilde{b}_j^{\mathrm{NL}}]$ needs to be recomputed for every $j$ of every group. Thus, we introduce several ways to approximate (14) for practical calculation times without significant sacrifices in accuracy.

First, we use a linear interpolation to downscale the PSD $\Psi$ to a $N_f \times \cdots \times N_f$ $d$-dimensional array. Consistent with the downscaling of $\Psi$, we periodically fold $\tilde{b}_j^{\mathrm{NL}}$ to a $N_f \times \cdots \times N_f$ array, reducing its size while preserving the energy of the transform $\mathcal{T}^{\mathrm{NL}}$. Second, since $\tilde{b}_j^{\mathrm{NL}}$ is sparse, the individual cascaded 1-dimensional FFTs required for the separable computation of $\mathscr{F}[\tilde{b}_j^{\mathrm{NL}}]$ are also sparse. For instance, when $d = 2$, instead of $2N_f$ 1-D FFTs of length $N_f$, we need at most $N_f + M_{\mathrm{col}}$, where $M_{\mathrm{col}} \leq M$ is the number of distinct column (or row) components of the spatial locations $x_1, \ldots, x_M$. Third, exploiting the symmetries of the Fourier spectrum, the amount of operations for computing (14) can be almost halved.

It is also possible to calculate the exact variances only for some $s_{i,j}^{x_1,\ldots,x_M}$. Although the misestimation of the noise variance can cause both noise artifacts and oversmoothed areas, it is often relatively safe to approximate the variances for coefficients where the signal is small. Specifically, if we calculate the variances $v_{i,j}^{x_1,\ldots,x_M}$ exactly for coefficients $s_{i,j}^{x_1,\ldots,x_M}$ larger than a threshold $\tau$ and the shrinkage follows (24) with $0 \leq \alpha_{i,j} \leq 1$, then $\tau$ bounds the shrinkage error arising from approximating $v_{i,j}^{x_1,\ldots,x_M}$ for a coefficient such that $\left| s_{i,j}^{x_1,\ldots,x_M} \right| \leq \tau$. This can be used to reduce the computational burden without significant compromises in accuracy, as there are typically few large-magnitude $s_{i,j}^{x_1,\ldots,x_M}$ due to signal compaction in $\mathcal{T}^{(d+1)\mathrm{D}}$ domain.

In the particular case when $v_{i,j}^{x_1,\ldots,x_M}$ are calculated exactly only for some $s_{i,j}^{x_1,\ldots,x_M}$ with $j = j_1,\ldots,j_K$, $K \leq M$ (i.e. for $K$ select planes of the 3-D spectrum, see Figure 2), we can approximate the remaining $(M-K)N$ variances for

| | | | $N_f$ | | |
|---|---|---|---|---|---|
| | | 8 | 16 | 32 | 64 |
| | 1 | 5.47 | 5.64 | 6.52 | 8.96 |
| | 4 | 5.57 | 6.04 | **8.13** | 15.74 |
| $K$ | 8 | 5.69 | 6.65 | 10.53 | 25.66 |
| | 12 | 5.77 | 7.07 | 12.40 | 33.86 |
| | 16 | 5.82 | 7.48 | 14.39 | 42.52 |

$j \notin \{j_1, \ldots, j_K\}$ as

$$v_{i,j}^{x_1,\ldots,x_M} \approx \frac{M v_i - \sum_{k=1}^{K} v_{i,j_k}^{x_1,\ldots,x_M}}{M - K}. \qquad (29)$$

This approximation follows from the energy-preservation identity $\sum_{j=1}^{M} v_{i,j}^{x_1,\ldots,x_M} = M v_i$ that holds for an orthonormal $\mathcal{T}^{\mathrm{NL}}$ and is convenient as $v_i$ (5) needs not be recomputed for each group. Note that (29) can be treated as an interpolator between (8) and (14), because if $K = 0$ then (29) approximates all variances with (5).

In our implementation, we presume that most of the signal is compacted in the first few coefficient planes of $s_{i,j}^{x_1,\ldots,x_M}$, and thus compute exact variances $v_{i,j}^{x_1,\ldots,x_M}$ only for $j = 1,\ldots,K$ and rely on (29) for the rest. The impact of the two approximation parameters $N_f$ and $K$ on the average CPU execution times for single-threaded denoising and PSNR of the estimate are reported in Table II and in Figure 8.

## V. INHERENT LIMITATIONS OF TRANSFORM-DOMAIN BLOCK FILTERING

Ideally, sparsity-based methods transform the signal in a domain where the clean image can represented by few high-magnitude coefficients, whereas the noise is spread over many coefficients that remain of small magnitude. These methods become ineffective when noise also gets compacted into few
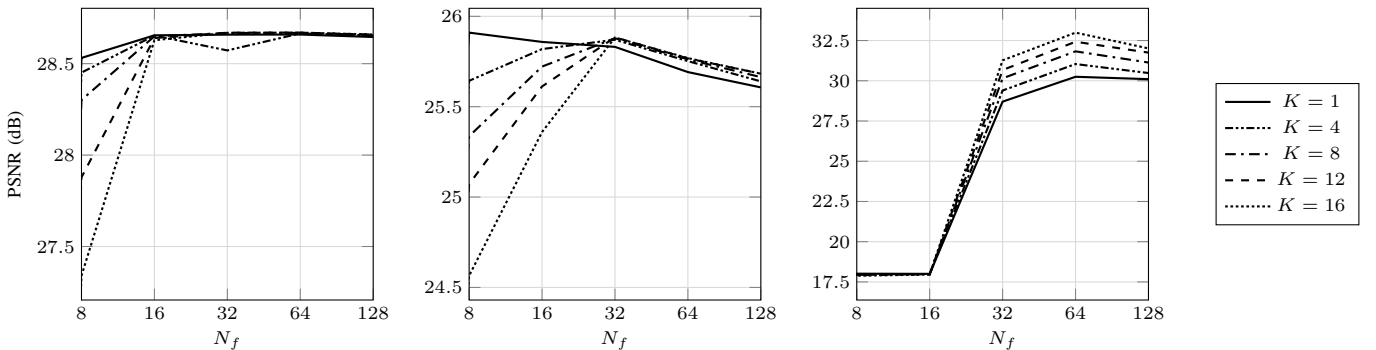


Figure 8. Average PSNR (dB) of denoising a set of images specified in Section VII corrupted by white noise as well as correlated noise defined by $g_1$ and $g_3$ (left to right) with $\|g\|_2^2 = 0.02$ with varying $K$ as a function of $N_f$. Note that although denoising of white noise is not affected by the PSD resizing, the folding of $\tilde{b}_j^{\mathrm{NL}}$ still causes inaccuracies with small values of $N_f$. The effects of $N_f$ and $K$ on execution time are demonstrated in Table II.

coefficients that are among those that capture significant energy of the clean signal – under such scenario (exemplified by some of the PSDs in our experiments), it is not possible to remove noise without at the same time oversmoothing the estimate even with the exact variances (14), as demonstrated in Figure 9. For instance, a $\mathcal{T}^{d\mathrm{D}}$ like the 2-D DCT lacks directional selectivity and cannot differentiate between diagonal and antidiagonal components, making it impossible to attenuate diagonal noise without oversmoothing antidiagonal image components. The small size of the patches further limits the frequency resolution of $\mathcal{T}^{d\mathrm{D}}$.

### A. Global Fourier thresholding and refiltering

An estimate of the details lost due to these systemic factors may be obtained by comparing the global FFT spectrum of the residual $\Delta = z - \hat{y}$, where $\hat{y}$ is the denoised image, against the noise PSD:

$$\hat{\Delta} = \mathscr{F}^{-1}[\mathscr{F}[\Delta]\, H[\alpha_\Delta]], \qquad (30)$$

where $\alpha_\Delta$ is a three-sigma test

$$\alpha_\Delta = \begin{cases} 1 & \text{if } |\mathscr{F}[\Delta]| > 3\sqrt{\Psi} \\ 0 & \text{otherwise,} \end{cases} \qquad (31)$$

and $H$ is a morphological filter to dilate the result of the test, and thus reduce the effects of noise. A new noisy image $z^{\mathrm{GFT}}$ with PSD $\Psi^{\mathrm{GFT}} = \Psi H[\alpha_\Delta]$ is defined as

$$z^{\mathrm{GFT}} = \hat{y} + \hat{\Delta}. $$

Whenever the collaborative filter suppresses noise without systematic oversmoothing, we have $\alpha_\Delta \approx 0$ and thus also $H[\alpha_\Delta] \approx 0$; in this situation $z^{\mathrm{GFT}} \approx \hat{y}$ and its noise PSD is zero, hence the estimate produced by refiltering is not going to substantially differ from the initial $\hat{y}$. Otherwise, with systematic oversmoothing, the spectrum of $\Delta$ contains coefficients



z (17.12 dB)          $\hat{y}^{\mathrm{HT}}$ (27.89 dB)

z (17.09 dB)          $\hat{y}^{\mathrm{HT}}$ (27.50 dB)

Figure 9. View of denoising of *Lena* corrupted by correlated noise defined by $g_3$ (top) and $g_7$ of (4) (bottom), both with $\|g\|_2^2 = 0.02$. On left, the noisy images $z$, and on right the corresponding hard-thresholding estimates $\hat{y}^{\mathrm{HT}}$.

large with respect to $\Psi$; for such frequencies $\alpha_\Delta = 1$, and $z^{\mathrm{GFT}}$ captures the oversmoothing with the associated noise components described by $\Psi H[\alpha_\Delta]$, as demonstrated in Figure 10 and Figure 11. By filtering $z^{\mathrm{GFT}}$ with the collaborative filter, we can restore some of the details lost when filtering $z$, as shown in Figure 12.

The refiltering process is done separately with the estimates of hard-thresholding and Wiener filtering immediately after the respective steps, thus resulting in two hard-thresholding steps followed by two Wiener-filtering steps. The entire denoising
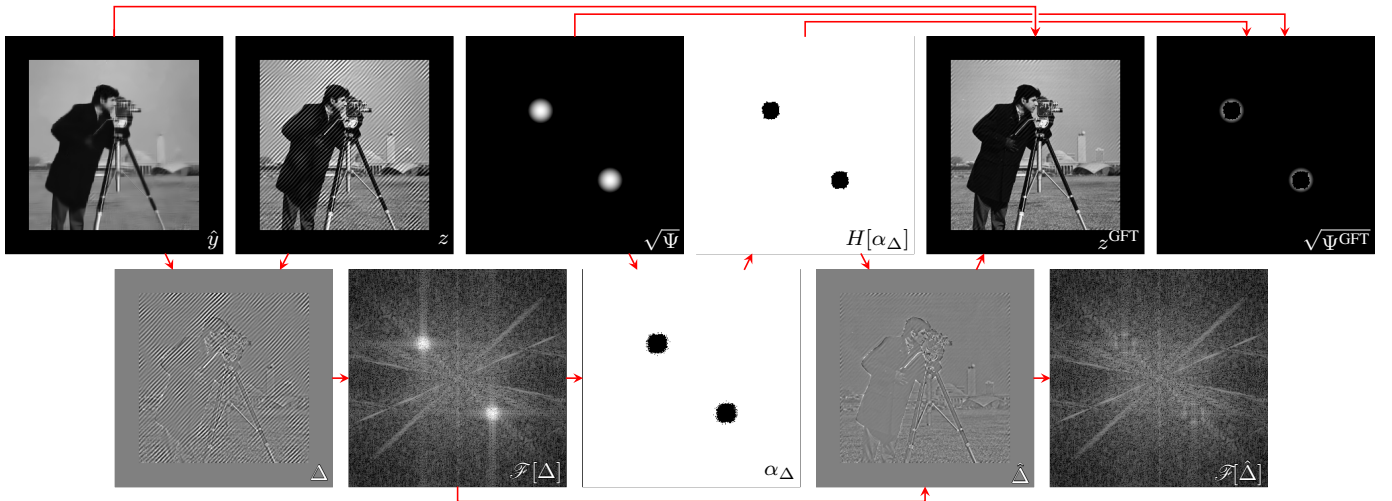


Figure 10. Demonstration of the global Fourier thresholding process applied to the result of hard-thresholding of *Cameraman* corrupted by noise modeled by kernel $g_3$ with $\|g\|_2^2 = 0.02$. First, we obtain the residual $\Delta$ as the difference of the noisy image $z$ and the hard-thresholding stage estimate $\hat{y}$. Then, we perform the three-sigma test (31) on the Fourier spectrum of the residual $\mathscr{F}[\Delta]$ against $\sqrt{\Psi}$ to obtain $\alpha_\Delta$, which is then dilated to $H[\alpha_\Delta]$. With $H[\alpha_\Delta]$, we filter the residual $\Delta$ as in (30) to obtain $\hat{\Delta}$, which is combined with $\hat{y}$ to obtain the new noisy image $z^{\mathrm{GFT}}$ with PSD $\Psi^{\mathrm{GFT}} = \Psi H[\alpha_\Delta]$. For comparison, we also show the Fourier spectrum $\mathscr{F}[\hat{\Delta}]$ of $\hat{\Delta}$. Due to the lack of periodicity in the noise, $z$ and $\hat{y}$ are zero-padded by the correlation kernel size. Fourier spectra and root PSDs are visualized by their log-magnitude.
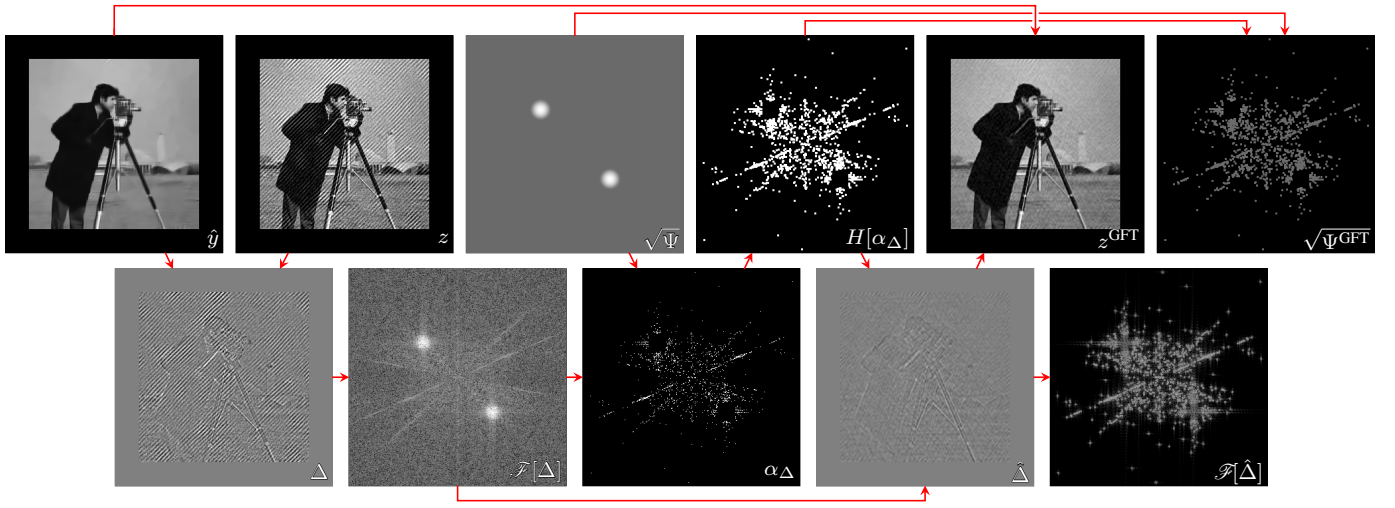
Figure 11. Demonstration of the global Fourier thresholding process as in Figure 10 with the correlation kernel $g_7$ of (4) and $\|g\|_2^2 = 0.02$. The correlated noise produced through this kernel includes a white noise component, which is reflected by the gray background of $\sqrt{\Psi}$; as such, the signal cannot be separated from the residual $\Delta$ as clearly as in Figure 10. The refiltering process manages nevertheless to recover a considerable amount of details, which are preserved through a second application of collaborative hard-thresholding as can be seen from comparing Figure 9(bottom) and Figure 12(bottom).



$z^{\text{GFT}_{\text{HT}}} = \hat{y}^{\text{HT}} + \hat{\Delta}$ (39.58 dB)  $\hat{y}^{\text{HT}_{\text{GFT}}}$ (39.51 dB)

$z^{\text{GFT}_{\text{HT}}} = \hat{y}^{\text{HT}} + \hat{\Delta}$ (29.68 dB)  $\hat{y}^{\text{HT}_{\text{GFT}}}$ (30.95 dB)

Figure 12. Global Fourier thresholding and refiltering of hard-thresholding results of Figure 9. On left, the noisy images $z^{\text{GFT}_{\text{HT}}}$ produced by the global Fourier thresholding; on right, the refiltered results $\hat{y}^{\text{HT}_{\text{GFT}}}$. For the noise represented by $g_3$ (top), $z^{\text{GFT}_{\text{HT}}}$ is almost noise-free and $\hat{y}^{\text{HT}_{\text{GFT}}}$ retains a very significant amount of detail. The noise of $g_7$ (bottom) includes a white noise component, due to which the signal cannot be separated from the residual as clearly, but nevertheless the refiltering provides a significant improvement both visually and with regard to the PSNR compared to the result of Figure 9.



Figure 13. Flowchart of the entire denoising process. The process takes as input the noisy image $z$ as well as the noise PSD $\Psi$. The first step of the algorithm performs hard-thresholding with these arrays as input, resulting in an estimate $\hat{y}^{\text{HT}}$. If refiltering is included, $\hat{y}^{\text{HT}}$ is then processed through the global Fourier thresholding, which yields a new noisy image $z^{\text{GFT}_{\text{HT}}}$ and its corresponding PSD $\Psi^{\text{GFT}_{\text{HT}}}$. These are then filtered through another BM3D hard-thresholding step, resulting in a new estimate $\hat{y}^{\text{HT}_{\text{GFT}}}$. This is used as a pilot estimate for the following Wiener filter, which outputs the first Wiener estimate $\hat{y}^{\text{Wie}}$. This estimate is processed again though the global Fourier thresholding filter, the output of which is a noisy image $z^{\text{GFT}_{\text{Wie}}}$ and its PSD $\Psi^{\text{GFT}_{\text{Wie}}}$, the noisy input to a second Wiener filter with $\hat{y}^{\text{Wie}}$ as a pilot. The output of this Wiener filter is the final estimate $\hat{y}^{\text{Wie}_{\text{GFT}}}$.

process with the refiltering steps is illustrated in Figure 13. The effectiveness of the procedure depends largely on the original noise structure – sometimes the image features removed by the shrinkage cannot be completely separated from noise even through further processing.

It should be noted that as (30) processes the noise in the global Fourier domain, it presumes the noise periodic, i.e. such that the convolution in (2) is circular. Most inaccuracies caused
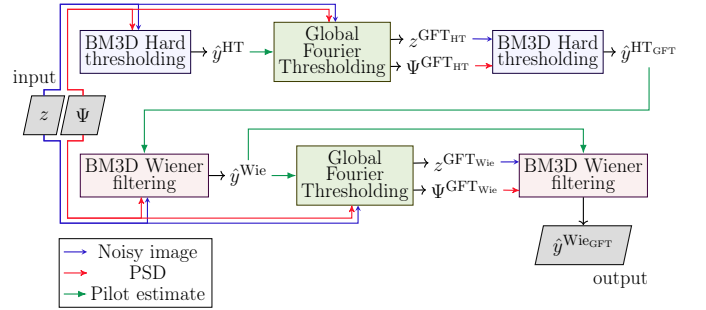
by this mismatch can be mitigated by zero-padding the noisy image and the denoised estimate by the support size of the correlation kernel $g$ and then cropping the resulting $z^{\text{GFT}}$. Corresponding padding and cropping operations are made on $g$ in order to define $\Psi$ for (31) and $\Psi^{\text{GFT}}$ for the subsequent collaborative filtering. Although the zero-padding does not completely mitigate the inaccuracies caused by the lack of periodicity, the remaining artifacts are mainly found at the edges of the image. Such artifacts for the considered cases of correlated noise are demonstrated in Figure 14. Strategies for non-circular deconvolution, such as [42], [43], could perhaps be leveraged to further reduce these artifacts.
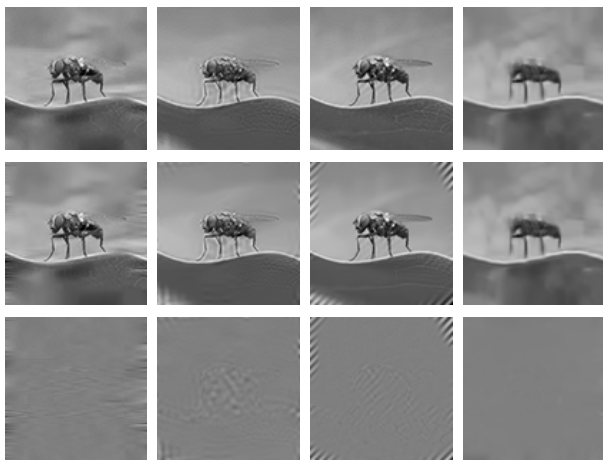
Figure 14. Demonstration of the edge artifacts created by the refiltering process. On the top two rows, denoising results of $100\times100$ noisy images subject to correlated noise by kernels $g_1$, $g_2$, $g_3$, and $g_4$ ($\|g\|_2^2 = 0.02$ in all cases). On top, the noisy images were generated with periodic noise (such that the convolution in (2) is circular). Below, the noise was generated with a non-circular convolution (as used in all other experiments in the paper). On the bottom row, differences between the two estimates. The underlying white noise realizations in the visible areas are identical in both cases, and differences at the boundaries arise after convolution due to the circularity or the lack thereof. Although the estimates are different due to differences in noise at the boundaries, they are largely of similar quality in the central parts of the image.



Figure 15. Relative changes in resulting average PSNR values of denoising a set of images ($\|g\|_2^2 = 0.02$) with changing values of $\lambda$. Only the hard-thresholding stage of BM3D is employed. We note that the best value of $\lambda$ largely depends on the correlation kernel, and may be larger than that of white noise (kernel $g_w$).



Figure 16. Resulting PSNR values of denoising a set of images with correlated noise defined by $g_w$ (top) and $g_1$ (bottom) with $\|g\|_2^2 = 0.02$ while changing the $\lambda$ and $\mu$ parameters. Note the different values of $\mu$.

## VI. PARAMETER SELECTION

A common diagonal shrinkage threshold $\lambda$ for (25) is the so-called universal threshold $\lambda = \sqrt{2\log(MN)}$, where $MN$ is the sample size transformed by $\mathcal{T}^{(d+1)\mathrm{D}}$ [44]. This threshold value follows the statistics of the sample maximum of $\left|\left\langle[\eta_{x_1};\cdots;\eta_{x_M}], b_i^{d\mathrm{D}}\otimes b_j^{\mathrm{NL}}\right\rangle\right|/\sqrt{v_{i,j}^{x_1,\ldots,x_M}}$ when these standardized samples are independently distributed; it is not valid when $\left\langle[\eta_{x_1};\cdots;\eta_{x_M}], b_i^{d\mathrm{D}}\otimes b_j^{\mathrm{NL}}\right\rangle$ are correlated, which can happen either because of correlation in the noise or because of block overlap. For the sake of simplicity and based on empirical tuning, BM3D had used a fixed shrinkage threshold independent of $N$, specifically $\lambda = 2.7$ for white noise [7] and a slightly larger $\lambda = 2.9$ for correlated noise [8].

Figure 15 demonstrates how the optimal $\lambda$ largely depends on the correlation kernel used, and that it may be either smaller or larger than that of white noise. Hence the practice regretfully deviates from the theoretical guidelines from [3], which indicated $\lambda$ for white noise as an upper bound for the other cases, noting that the probability of the sample maximum to exceed a given threshold is highest when standardized samples are independent.

As the Wiener filtering is highly dependent on the pilot signal, the choice of $\mu^2$ is also affected by $\lambda$. Changes in $\mu^2$ can then, up to a point, mitigate the effects of under- or over-filtering of the hard-thresholding estimate, as demonstrated in Figure 16. We note that BM3D had always adopted $\mu^2 = 1$ [7], [8].

The adaptive selection of good $\lambda$ and $\mu^2$ parameters for any given PSD is based on a set of 500 PSDs (randomly generated and excluding any of the tested PSDs excep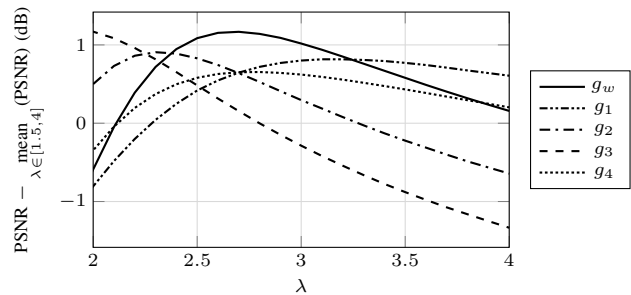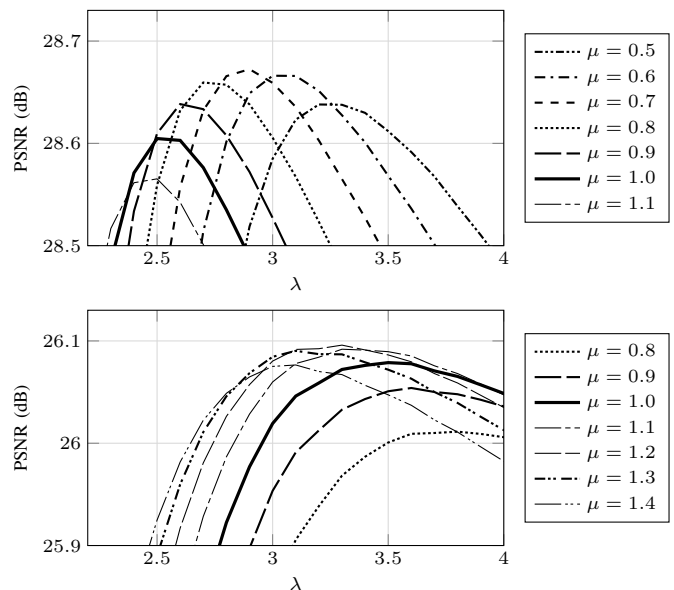t $g_w$) for which we have preliminarily obtained the best $\lambda$ and $\mu^2$ from a grid of test values, defined as those which on average across 5 test images result in the best PSNR (similar to the maxima of Figure 16). Specifically, we find the 20 closest PSDs within this set based on the Mahalanobis distance of the integral projections along the principal axes of the PSDs. The adaptive $\lambda$ and $\mu^2$ are obtained as the weighted average of the respective best parameters for the closest PSDs, with weights reciprocal to the distances. A second pair of $\lambda$ and $\mu^2$ is produced similarly for the refiltering steps.

For the simplification parameters, we select $K = 4$ and $N_f = 32$, which provide a significantly improved runtime without serious compromises in denoising quality, as demonstrated by Table II and Figure 8. As the refiltering procedure performs two stages of both hard-thresholding and Wiener (see Figure 13), its cost is double of that without refiltering.

## VII. EXPERIMENTS

To assess the performance gain from the algorithmic improvements described in Section IV, we consider denoising of images with added white noise as well as the correlated

Table III

AVERAGE PSNR FOR DENOISING OF WHITE NOISE AND EIGHT CORRELATED NOISE CASES OVER *Barbara*, *Boat*, *Cameraman*, *Couple*, *House*, *Lena*, *Man*, AND *Peppers* WITH 10 INDEPENDENT REALIZATIONS FOR EACH COMBINATION OF IMAGE, NOISE CASE, AND VARIANCE LEVEL. WE MARK IN BOLD ALL THE RESULTS THAT CANNOT BE DIFFERENTIATED FROM THE BEST ONE THROUGH A WELCH'S T-TEST [45] WITH $p = 0.05$.

| $g$ | $\|g\|_2^2$ | noisy | new BM3D | new BM3D (refilter) | old BM3D | old BM3D (refilter) | BLS-GSM | BLS-GSM (refilter) | NLMC | NLMC (refilter) | Noise Clinic | Noise Clinic (refilter) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $g_w$ | 0.001 | 30.00 | **35.73** | **35.74** | **35.75** | **35.74** | 35.07 | 35.09 | 35.11 | 35.03 | 34.74 | 35.05 |
| | 0.01 | 20.00 | **30.40** | **30.38** | **30.39** | **30.32** | 29.49 | 29.46 | 29.34 | 29.25 | 27.57 | 28.63 |
| | 0.02 | 16.99 | **28.79** | **28.78** | **28.77** | **28.68** | 27.88 | 27.84 | 27.54 | 27.42 | 25.38 | 26.63 |
| $g_1$ | 0.001 | 29.96 | 36.17 | **36.50** | 35.98 | **36.22** | 35.47 | 35.68 | 30.38 | 31.12 | 32.05 | 32.79 |
| | 0.01 | 19.96 | 28.40 | **28.82** | 26.80 | 26.97 | **28.84** | **28.94** | 21.10 | 21.43 | 21.57 | 22.94 |
| | 0.02 | 16.95 | 25.90 | 26.33 | 23.39 | 23.52 | **26.92** | **26.99** | 18.24 | 18.48 | 18.14 | 19.22 |
| $g_2$ | 0.001 | 30.00 | 35.23 | **36.95** | 34.40 | 35.02 | 35.26 | 35.70 | 30.61 | 31.37 | 33.84 | 34.88 |
| | 0.01 | 20.00 | 28.67 | **31.29** | 25.42 | 25.88 | 29.74 | 30.40 | 21.71 | 22.02 | 24.87 | 26.36 |
| | 0.02 | 16.99 | 26.98 | **29.59** | 22.52 | 22.91 | 28.37 | 28.99 | 18.94 | 19.15 | 22.38 | 23.65 |
| $g_3$ | 0.001 | 30.10 | 37.21 | 42.53 | 32.54 | 33.39 | 40.32 | **44.95** | 29.99 | 30.76 | 32.60 | 34.77 |
| | 0.01 | 20.10 | 31.15 | 41.16 | 21.93 | 22.41 | 35.87 | **42.43** | 20.52 | 20.87 | 20.55 | 20.99 |
| | 0.02 | 17.09 | 29.66 | 40.26 | 18.83 | 19.32 | 34.67 | **41.35** | 17.61 | 17.87 | 17.44 | 17.68 |
| $g_4$ | 0.001 | 29.99 | **34.27** | **34.27** | 34.39 | 34.38 | 33.74 | 33.76 | 32.63 | 33.02 | 33.94 | 34.07 |
| | 0.01 | 19.99 | **27.58** | **27.56** | **27.67** | **27.65** | 27.18 | 27.16 | 24.78 | 25.28 | 26.98 | 27.25 |
| | 0.02 | 16.98 | **25.47** | **25.46** | **25.52** | **25.49** | 25.27 | 25.25 | 22.57 | 23.14 | 24.79 | 25.14 |
| $g_5$ | 0.001 | 30.00 | **35.26** | **35.40** | 35.09 | 35.12 | 34.69 | 34.76 | 32.39 | 32.59 | 32.22 | 32.76 |
| | 0.01 | 20.00 | **28.11** | **28.24** | 27.47 | 27.46 | **28.21** | **28.21** | 25.93 | 26.34 | 22.46 | 23.69 |
| | 0.02 | 16.99 | 25.82 | 25.95 | 25.00 | 24.99 | **26.35** | **26.32** | 24.35 | 24.58 | 19.08 | 20.19 |
| $g_6$ | 0.001 | 30.08 | 35.62 | 37.49 | 32.69 | 32.94 | 37.19 | **38.28** | 32.69 | 32.83 | 32.75 | 32.75 |
| | 0.01 | 20.08 | 29.72 | 31.87 | 22.72 | 22.82 | 31.72 | **32.37** | 28.58 | 28.75 | 21.48 | 21.70 |
| | 0.02 | 17.07 | 28.22 | 30.26 | 19.71 | 19.81 | 30.31 | **30.69** | 27.97 | 27.98 | 18.47 | 18.61 |
| $g_7$ | 0.001 | 30.00 | 34.59 | **35.21** | 34.15 | 34.26 | 34.49 | 34.66 | 32.87 | 32.93 | 33.84 | 34.00 |
| | 0.01 | 20.00 | 28.38 | **29.11** | 26.38 | 26.42 | 28.78 | **28.99** | 28.37 | 28.36 | 25.70 | 26.55 |
| | 0.02 | 16.99 | 26.77 | **27.43** | 23.84 | 23.86 | 27.31 | **27.47** | 27.12 | 26.99 | 23.36 | 24.21 |
| $g_8$ | 0.001 | 29.99 | **34.51** | **34.50** | **34.61** | **34.60** | 33.95 | 33.97 | 32.93 | 33.31 | 34.09 | 34.24 |
| | 0.01 | 19.99 | **28.00** | **27.97** | **28.09** | **28.06** | 27.51 | 27.49 | 25.26 | 25.70 | 27.21 | 27.58 |
| | 0.02 | 16.98 | **25.95** | **25.93** | **26.01** | **25.97** | 25.63 | 25.61 | 23.01 | 23.55 | 25.04 | 25.52 |

noise cases defined in Table I and through (4), respectively associated to kernels $g_w$, $g_1, \ldots, g_4$, and $g_5, \ldots, g_8$. For each kernel, we test three different levels of variance $\|g\|_2^2 = 0.001, 0.01, 0.02$ on a set of 8 natural images (*Cameraman*, *House*, *Peppers* of size $256 \times 256$, and *Barbara*, *Boat*, *Couple*, *Lena*, *Man* of size $512 \times 512$).

We compare denoising results of BM3D with the proposed improvements ("new BM3D") versus BM3D for correlated noise ("old BM3D") [8], NLMeans-C [46], BLS-GSM [47], and Noise Clinic[5] [48]. In the comparison, we include also the refiltering procedure applied to new BM3D as detailed in Section V, as well as to each of the four other algorithms, for which it is implemented as a global Fourier thresholding between two passes of the same algorithm.

The results are reported in Table III and illustrated in Figure 17 and Figure 18. All reported PSNR values are computed after trimming off a 16-pixel border in order to ignore possible edge artifacts caused by the refiltering.

Furthermore, to demonstrate the effects of our improvements over the old BM3D applied to real data, we show a comparison of stripe removal of Terra MODIS satellite data [49] in Figure 19. This long-wavelength infrared imagery is characterized by crosstalk [50], which we here model as

stationary using a horizontal line kernel, ignoring out-of-band contributors.

By comparing results of old and new BM3D, we can see that the exact variances play a crucial role in enabling effective denoising of strongly correlated noise. Furthermore, refiltering provides a significant improvement in most cases of correlated noise, without reduction in quality in cases where further details cannot be recovered.

Notably, the benefits of refiltering are not limited to BM3D; other algoritms benefit from the applied refiltering step in several cases, particularly BLS-GSM for kernel $g_3$, as can be seen from both Table III and Figure 18. However, it should be noted that the refiltering is designed to restore oversmoothed details, not to improve the noise attenuation. Thus in some cases (such as the old BM3D for correlated noise), it does not provide a notable advantage due to the original estimate being very noisy.

## VIII. DISCUSSION

The presented method allows the calculation of exact transform-domain noise variances for any number of dimensions (such as video or volumetric data) or patch shape, applicable to any collaborative filter, such as BM3D, V-BM3D [51], BM4D [37], and RF3D [19]. Previously, all versions of BM3D, apart from the video denoising algorithm RF3D, approximated the $\mathcal{T}^{(d+1)\mathrm{D}}$ noise variance simply by replicating the $\mathcal{T}^{d\mathrm{D}}$ variances.

In RF3D, the noise is modeled as a combination of random and fixed-pattern noise in time, each described by a 2-D

[5]We acknowledge the comparison to Noise Clinic as somewhat unfair, considering that Noise Clinic is a blind denoising algorithm that estimates the noise model and PSD from noisy image. In an attempt to provide sufficient noisy data for the PSD estimation, we pad $y$ of Noise Clinic with a large smooth gradient.
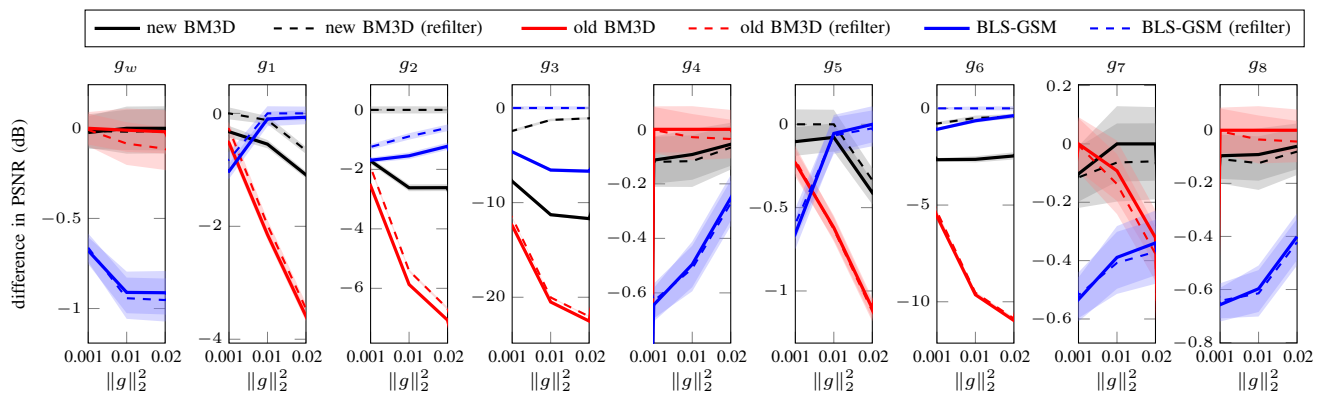
Figure 17. Denoising results of Table III for the three best algoritms, shown as the average PSNR difference to the best result, which is always displayed at 0. The shaded areas visualize the standard deviation of the average PSNR values.

PSD. Hence the overall 3-D noise power spectrum $\Psi$ has two distinct components: a temporal-DC plane equal to the sum of the 2-D PSDs, and a complementary temporal-AC volume composed of stacked replicas of the 2-D PSD of the random noise. Blocks are grouped from different frames into a 3-D array transformed as in BM3D, thus producing a spectrum $s_{i,j}^{x_1,\dots,x_M}$ with $x_1,\dots,x_M$ from $M$ consecutive frames. The variances $v_{i,j}^{x_1,\dots,x_M}$ are computed exactly only for two extreme cases: 1) blocks share the same spatial position ($x_1 = \dots = x_M$), and 2) blocks are sufficiently separated such that the fixed-pattern noise acts as another random noise. Intermediate cases are interpolated from these two extremes based on the proportion of blocks sharing a common spatial coordinate. We note that such approximation is effective only because of the special structure of the 3-D PSD $\Psi$ and of the typical regular arrangement of grouped blocks along motion trajectories in a video. The variance calculation method of Section III-C is significantly more general, even in its fast approximate forms of Section IV-D.

It is interesting to observe how a classical method like BLS-GSM is able to overcome BM3D in a few cases, particularly without refiltering. This is due to the highly redundant steerable pyramid transform adopted by this algorithm, which offers much finer granularity than the simple block transforms used by BM3D.

The refiltering procedure described in Section V is useful applied not only to BM3D, but also other denoising methods. The usability of the simple procedure shows that even with perfect modeling of the noise, the algorithms may be bound by the limitations brought by the used transforms or scales that lack the granularity needed for compacting the clean signal into few high-magnitude coefficients and spread most of the noise energy onto the remaining coefficients. We note that experiments on further iterations of refiltering did not give evidence of significant additional improvement.

In the remainder of this section, we expand on three particular aspects of this work: an interpretation of the empirical preference for a value of $\gamma$ larger than what would simply compensate the bias in (23); the limitations of the stationary correlated noise model (1)–(3) and how to overcome them in practical applications; and the inversion of (14), i.e. how one

might estimate the noise PSD $\Psi$ from the $\mathcal{T}^{(d+1)\mathrm{D}}$ spectrum variances.

### A. Effect of $\gamma > 1$ on block-matching

To provide an interpretation of the effect of $\gamma > 1$, we consider the variance of (17), which by (19) can be written as

$$\mathrm{var}\left\{\left\|z_{x_R} - z_{x_j}\right\|_2^2\right\} \approx 8 \sum_{i=1}^{N} \left(v_{i,2}^{x_R,x_j}\right)^2 +$$
$$16 \sum_{i=1}^{N} v_{i,2}^{x_R,x_j} \mathrm{E}^2\left\{s_{i,2}^{x_R,x_j}\right\}, \quad (32)$$

where the approximation follows from omitting the covariances. Focusing on the cases where the factor $\sum_{i=1}^{N} v_{i,2}^{x_R,x_j}$ plays a significant role in (23), we assume $\sum_{i=1}^{N} v_{i,2}^{x_R,x_j} \mathrm{E}^2\{s_{i,2}^{x_R,x_j}\} \leq \sum_{i=1}^{N} (v_{i,2}^{x_R,x_j})^2$. Thus from (32) we get

$$\mathrm{var}\left\{\left\|z_{x_R} - z_{x_j}\right\|_2^2\right\} \leq 24 \sum_{i=1}^{N} \left(v_{i,2}^{x_R,x_j}\right)^2, \quad (33)$$

and, noting that $\sum_{i=1}^{N} \left(v_{i,2}^{x_R,x_j}\right)^2 \leq \left(\sum_{i=1}^{N} v_{i,2}^{x_R,x_j}\right)^2$,

$$\mathrm{std}\left\{\left\|z_{x_R} - z_{x_j}\right\|_2^2\right\} \leq \sqrt{24} \sum_{i=1}^{N} v_{i,2}^{x_R,x_j}. \quad (34)$$

Hence, we speculate that a $\gamma > 1$ essentially provides an extra subtraction of $\frac{\gamma-1}{\sqrt{6}}\mathrm{std}\left\{\left\|z_{x_R} - z_{x_j}\right\|_2^2\right\}$ from (23), facilitating the inclusion in the group of noisy blocks which differ from the noisy reference block mainly due to the variance of (17).

### B. Overcoming the limitations of the observation model

Although we focus on the case of non-blind denoising, the proposed improvements are fairly robust to limited PSD over- or underestimation. Thus, they are directly applicable to the blind case as well, provided an estimate $\hat{\Psi}$ of the PSD $\Psi$ by an estimator such as [52]–[55]. The impact of scalar misestimation is relatively small and can be appreciated from Figures 6, 15, and 16, as changes in $\gamma$, $\lambda$, and $\mu$ are analogous to scaling the PSD. For example for $g_1$ and variance 0.02, in

Figure 18. A view of *Peppers* with white noise (kernel $g_w$, $\|g\|_2^2 = 0.01$), *Couple* with horizontal noise (kernel $g_1$, $\|g\|_2^2 = 0.01$), and *Man* with diagonal pattern noise (kernel $g_3$, $\|g\|_2^2 = 0.02$). Denoising results of BM3D with the proposed improvements, old BM3D, BLS-GSM, NLM-C and Noise Clinic, as well as these methods with the applied refiltering procedure. Note how the new BM3D yields significant improvements compared to the old BM3D in both cases of correlated noise, while keeping the same performance in white noise. Refiltering restores visible details for both correlated noise cases, without affecting the performance for white noise. Furthermore, especially with the noise generated by $g_3$, refiltering restores significant detail also for the BLS-GSM algorithm.
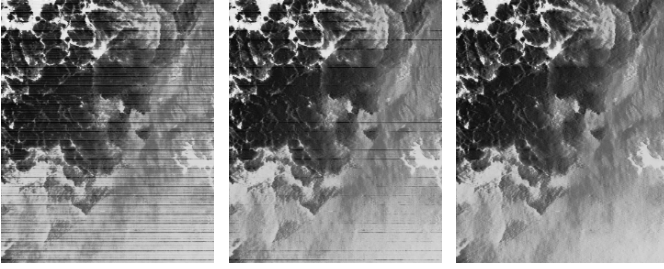
Figure 19.  Stripe removal of Terra MODIS band 30 data (9.580–9.880 $\mu m$) using BM3D. Left to right: noisy image, denoising results with old BM3D, denoising results with new BM3D. Old BM3D fails to remove the stronger stripes, which are successfully removed with the proposed improvements. Images are histogram-equalized for visualization purposes.

Figure 16 the parameter variations lead to loss in the tenths of dB, whereas in Table III the difference between old and new BM3D is more than 2dB.

The model also presumes the noise stationary; however, in BM3D all the filtering operations take place on quite small limited regions. Thus, global stationarity can be relaxed, applying the filter separately on smaller region where stationarity holds with a relevant local model for $\Psi$. In particular, for all the reported experiments, block-matching is done within a $39 \times 39$ search-neighborhood, and the downscaled $\Psi$ as well as the region used for computing the variances have only a $32 \times 32$ support. Likewise, the global Fourier thresholding can be applied on smaller image sections (e.g., $100 \times 100$).

Furthermore, the assumption of additive noise is not fundamentally restrictive. These algorithms are applicable, for example, to signal-dependent noise with spatial correlation [55] by means of variance-stabilizing transformations and suitable renormalization of $\Psi$, as shown in [56].

### C. Potential for estimating the noise PSD $\Psi$

In this work, we assume that the noise PSD $\Psi$ is known, and it is used to calculate the variances $v_{i,j}^{x_1,\ldots,x_M}$ of the spectrum coefficients of any group of patches with block coordinates $x_1,\ldots,x_M$ via (14). The same formula can be leveraged to compute an estimate $\hat{\Psi}$ of $\Psi$ given sample estimates $\hat{v}_{i,j}^{x_1,\ldots,x_M}$ of $v_{i,j}^{x_1,\ldots,x_M}$. Specifically, let us consider a collection of $R$ group coordinates $\left\{x_1^{[r]},\ldots,x_{M^{[r]}}^{[r]}\right\}_{r=1}^R$, and the associated sample estimates $\left\{\hat{v}_{i,j}^{x_1^{[r]},\ldots,x_{M^{[r]}}^{[r]}}\right\}_{r=1}^R$, which could be computed using methods like those in [19], [48], [55], [57]. Following (14), the error in the sample estimates given $\Psi$ can be quantified for the $r$-th group coordinates $x_1^{[r]},\ldots,x_{M^{[r]}}^{[r]}$ as

$$U_r(i,j,\Psi) = \hat{v}_{i,j}^{x_1^{[r]},\ldots,x_{M^{[r]}}^{[r]}} - \left\| \left| X \right|^{-2} \Psi \left| \mathscr{F}\left[b_i^{\mathrm{dD}}\right] \right|^2 \left| \mathscr{F}\left[\tilde{b}_j^{\mathrm{NL}}\right] \right|^2 \right\|_1 .$$

Minimizing this error over the entire collection and spectral indices yields an estimate of $\Psi$ as

$$\hat{\Psi} = \underset{\Psi \geq 0}{\operatorname{argmin}} \sum_{r=1}^R \sum_{j=1}^{M^{[r]}} \sum_{i=1}^N \left| U_r(i,j,\Psi) \right|^q . \qquad (35)$$

where $q=2$ yields a quadratic equation as $U_r(i,j,\Psi)$ is linear on $\Psi$ for a nonnegative $\Psi$, while $q=1$ provides a robust estimate. A smoothness or regularity prior on $\Psi$ can be used to

define a penalty that disambiguates underdetermined solutions. A weighting term $\omega_r$ or $\omega_r(i,j)$ can be included within the summation in (35) to balance the contribution of different sample estimates $\hat{v}_{i,j}^{x_1^{[r]},\ldots,x_{M^{[r]}}^{[r]}}$ to the final estimate $\hat{\Psi}$. Note that (35) can be solved iteratively, e.g., as in [55].

### IX. CONCLUSIONS

We presented a method which allows for both the exact computation and effective approximations of the noise spectrum variance in nonlocal collaborative transforms in any number of dimensions by taking into account the relative positions of the matched blocks and their interplay with the noise power spectrum. The exact variances allow us to both calculate more accurate shrinkage thresholds and to avoid matching blocks that are strongly correlated in noise rather than similar in the underlying image.

The experiments show that the presented method can yield significant improvements in BM3D denoising both visually and in terms of PSNR. The greatest improvements can be noted when the noise is very strongly correlated. We also show that a simple global Fourier thresholding and refiltering step can be used to recover a significant amount of detail with several cases of correlation, as the denoising may be limited due to block size or the used transforms.

A significant improvement for various cases of correlated noise can be noted even when the variances are approximated for a majority of the spectrum coefficients as described in Section IV-D; thus the proposed methods do not entail a computational burden.

### REFERENCES

[1] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Am. Stat. Assoc.*, vol. 90, no. 432, pp. 1200–1224, 1995.

[2] A. Foi, 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, 2007.

[3] I. M. Johnstone and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," *J. R. Stat. Soc. Series B*, vol. 59, no. 2, pp. 319–351, 1997.

[4] R. Neelamani, H. Choi, and R. Baraniuk, "ForWaRD: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Trans. Signal Proces.*, vol. 52, no. 2, pp. 418–433, 2004.

[5] J. S. De Bonet, "Noise reduction through detection of signal redundancy," Rethinking Artificial Intelligence, MIT AI Lab, 1997.

[6] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Sim.*, vol. 4, no. 2, pp. 490–530, 2005.

[7] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, 2007.

[8] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," in *Proc. SPIE Electronic Imaging '08, vol. 6812, no. 681207, San Jose (CA), USA*, 2008, p. 681207.

[9] A. Foi, K. Dabov, V. Katkovnik, and K. Egiazarian, "Shape-adaptive DCT for denoising and image reconstruction," in *Proc. SPIE Electronic Imaging 2006*, vol. 6064, 2006, pp. 203–214.

[10] G. Lubberts, "Random noise produced by x-ray fluorescent screens," *J. Opt. Soc. Am. A*, vol. 58, no. 11, pp. 1475–1483, 1968.

[11] G. Hajdok, J. J. Battista, and I. A. Cunningham, "Fundamental x-ray interaction limits in diagnostic imaging detectors: Frequency-dependent Swank noise," *Med. Phys.*, vol. 35, no. 7, pp. 3194–3204, 2008.

[12] S. J Riederer, N. Pelc, and D. Chesler, "The noise power spectrum in computed x-ray tomography," *Phys. Med. Biol.*, vol. 23, pp. 446–54, 1978.

[13] J. Siewerdsen, L. Antonuk, Y. El-Mohri, J. Yorkston, W. Huang, and I. Cunningham, "Signal, noise power spectrum, and detective quantum efficiency of indirect-detection flat-panel imagers for diagnostic radiology," *Med. Phys.*, vol. 25, no. 5, pp. 614–628, 1998.

[14] E. Samei, "Image quality in two phosphor-based flat panel digital radiographic detectors," *Med. Phys.*, vol. 30, no. 7, pp. 1747–1757, 2003.

[15] A. E. Burgess, "Mammographic structure: Data preparation and spatial statistics analysis," in *Proc. SPIE 3661, Medical Imaging 1999: Image Process.*, 1999, pp. 642–653.

[16] J. Zheng, J. A. Fessler, and H.-P. Chan, "Detector blur and correlated noise modeling for digital breast tomosynthesis reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 1, pp. 116–127, 2017.

[17] H. V. Kennedy, "Modeling noise in thermal imaging systems," in *Proc. SPIE 1969, Infrared Imaging Systems: Design, Analysis, Modeling, and Testing IV*, 1993, pp. 66–77.

[18] J. M. Mooney and F. D. Shepherd, "Characterizing IR FPA nonuniformity and IR camera spatial noise," *Infrared Phys. Technol.*, vol. 37, no. 5, pp. 595–606, 1996.

[19] M. Maggioni, E. Sánchez-Monge, and A. Foi, "Joint removal of random and fixed-pattern noise through spatiotemporal video filtering," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4282–4296, 2014.

[20] J. Chen, Y. Shao, H. Guo, W. Wang, and B. Zhu, "Destriping CMODIS data by power filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2119–2124, 2003.

[21] L. Gómez-Chova, L. Alonso, L. Guanter, G. Camps-Valls, J. Calpe, and J. Moreno, "Correction of systematic spatial noise in push-broom hyperspectral sensors: application to CHRIS/PROBA images," *Appl. Opt.*, vol. 47, no. 28, pp. F46–F60, 2008.

[22] S. Abramov, O. Rubel, V. Lukin, R. Kozhemiakin, N. Kussul, A. Shelestov, and M. Lavreniuk, "Speckle reducing for Sentinel-1 SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Sym. (IGARSS)*, 2017, pp. 2353–2356.

[23] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, 2018.

[24] T. Ehret, A. Davy, J.-M. Morel, G. Facciolo, and P. Arias, "Model-blind video denoising via frame-to-frame training," in *Proc. IEEE C. Comput. Vis. Pat. Rec. (CVPR)*, 2019, pp. 11 369–11 378.

[25] J. Batson and L. Royer, "Noise2Self: Blind denoising by self-supervision," in *Proc. 36th Int. C. Mach. Learn.*, ser. Proc. Mach. Learn. Res., K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97, 2019, pp. 524–533.

[26] N. Eslahi and A. Foi, "Anisotropic spatiotemporal regularization in compressive video recovery by adaptively modeling the residual errors as correlated noise," in *Proc. IEEE Image, Video, Multidim. Signal Process. W. (IVMSP)*, 2018, pp. 1–5.

[27] A. Davy, "Detection of rare and small events via complex background modeling," Ph.D. dissertation, École Normale Supérieure Paris-Saclay, 11 2019, http://www.theses.fr/s162210, http://dev.ipol.im/~adavy/DAVY_MANUSCRIPT.pdf.

[28] A. Davy, T. Ehret, J.-M. Morel, P. Arias, and G. Facciolo, "A Non-Local CNN for Video Denoising," in *Proc. IEEE Int. C. Image Process. (ICIP)*, 2019, pp. 2409–2413.

[29] A. Davy and T. Ehret, "GPU acceleration of NL-means, BM3D and VBM3D," *J. Real-Time Image Process.*, pp. 1–18, 2020.

[30] D. Wang, J. Xu, and K. Xu, "An FPGA-based hardware accelerator for real-time Block-Matching and 3D filtering," *IEEE Access*, 2020.

[31] "HiSilicon Kirin 990 5G specifications," 2019. [Online]. Available: http://www.hisilicon.com/en/Products/ProductList/Kirin

[32] Y. Mäkinen, L. Azzari, and A. Foi, "Exact transform-domain noise variance for collaborative filtering of stationary correlated noise," in *Proc. IEEE Int. C. Image Process. (ICIP)*, 2019, pp. 185–189.

[33] B. Goossens, "NLMeans denoising algorithm, Matlab code," 2008. [Online]. Available: https://quasar.ugent.be/bgoossen/download_nlmeans/

[34] M. Weissman, "$1/f$ noise and other slow, nonexponential kinetics in condensed matter," *Rev. Mod. Phys.*, vol. 60, no. 2, p. 537, 1988.

[35] M. Matrecano, G. Poggi, and L. Verdoliva, "Improved BM3D for correlated noise removal," in *Proc. Int. Conf. Comput. Vision Theor. App. (VISAPP)*, 2012, pp. 129–134.

[36] R. Yin, T. Gao, Y. Lu, and I. Daubechies, "A tale of two bases: Local-nonlocal regularization on image patches with convolution framelets," *SIAM J. Imag. Sci.*, vol. 10, no. 2, pp. 711–750, 2017.

[37] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, 2012.

[38] A. S. Rubel, V. V. Lukin, and K. O. Egiazarian, "Metric performance in similar blocks search and their use in collaborative 3d filtering of grayscale images," in *Proc. SPIE 9019, Image Process.: Algorithms and Systems XII*, 2014, p. 901909.

[39] A. Rubel, V. Lukin, and K. Egiazarian, "Block matching and 3D collaborative filtering adapted to additive spatially correlated noise," *Proc. Int. W. Video Process. Qual. Metrics Consum. Electr. (VPQM), Scottsdale, USA*, 2015.

[40] O. G. Guleryuz, "Weighted averaging for denoising with overcomplete dictionaries," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 3020–3034, 2007.

[41] K. Egiazarian, V. Katkovnik, and J. Astola, "Local transform-based image de-noising with adaptive window size selection," in *Proc. SPIE*, vol. 4170, 2001, p. 14.

[42] M. S. Almeida and M. Figueiredo, "Deconvolving images with unknown boundaries using the alternating direction method of multipliers," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3074–3086, 2013.

[43] S. J. Reeves, "Fast image restoration without boundary artifacts," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1448–1453, 2005.

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

[45] B. L. Welch, "The generalization of "Student's" problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1/2, pp. 28–35, 1947.

[46] B. Goossens, H. Luong, A. Pizurica, and W. Philips, "An improved nonlocal denoising algorithm," in *Proc. Int. W. Local Non-Local Approx. Image Process. (LNLA)*, 2008, pp. 143–156.

[47] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, 2003.

[48] M. Lebrun, M. Colom, and J.-M. Morel, "The noise clinic: a blind image denoising algorithm," *Image Process. On Line*, vol. 5, pp. 1–54, 2015.

[49] "Terra & Aqua Moderate Resolution Imaging Spectroradiometer (MODIS)." [Online]. Available: https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/modis/

[50] T. Wilson, A. Wu, X. Geng, Z. Wang, and X. Xiong, "Analysis of the electronic crosstalk effect in Terra MODIS long-wave infrared photovoltaic bands using lunar images," in *Image and Sig. Process. for Remote Sens. XXII*, vol. 10004. SPIE, 2016, pp. 445 – 459.

[51] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, 2012.

[52] B. Kleiner, R. D. Martin, and D. J. Thomson, "Robust estimation of power spectra," *J. Royal Stat. Soc. B (Meth.)*, vol. 41, no. 3, pp. 313–338, 1979.

[53] A. D. Chave, D. J. Thomson, and M. E. Ander, "On the robust estimation of power spectra, coherences, and transfer functions," *J. Geophys. Res. Sol. Ea.*, vol. 92, no. B1, pp. 633–648, 1987.

[54] B. Aiazzi, L. Alparone, and S. Baronti, "Unsupervised assessment and pyramidal filtering of colored speckle," in *SAR Image Analysis, Modeling, and Techniques II*, F. Posa, Ed., vol. 3869, Int. Soc. Opt. and Phot. SPIE, 1999, pp. 9–20.

[55] L. Azzari, L. R. Borges, and A. Foi, "Modeling and estimation of signal-dependent and correlated noise," in *Denoising of Photographic Images and Video: Fundamentals, Open Challenges and New Trends, M. Bertalmio (Ed.)*. Springer, 2018, pp. 1–36.

[56] L. Azzari and A. Foi, "Variance stabilization in Poisson image deblurring," in *2017 IEEE 14th Int. Symp. Biomed. Imaging (ISBI 2017)*, 2017, pp. 728–731.

[57] A. Danielyan and A. Foi, "Noise variance estimation in nonlocal transform domain," in *Proc. Int. W. Local Non-Local Approx. Image Process. (LNLA)*, 2009, pp. 41–45.