SUPPLEMENTARY MATERIAL FOR ULTRASOUND IMAGE DENOISING WITH MONTE-CARLO RISK ESTIMATION

Ashutosh Gupta[1,4] *Chandra Sekhar Seelamantula*[2] *Thierry Blu*[3] *Himanshu Shekhar*[4], *Nitant Dube*[1] *Shanmuganathan Raman*[4]

[1] Space Applications Centre, ISRO, Ahmedabad, India
[2] Department of Electrical Engineering, Indian Institute of Science (IISc.), Bengaluru, India
[3]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
[4] Indian Institute of Technology (IIT), Gandhinagar, India

1. NOTATIONS

Throughout the paper and the supplementary material, we use the following notations:

1. $\mathbf{Y}$ to denote m dimensional multivariate (vector) random variables.
2. $\mathbf{y}$ to denote m dimensional samples from the corresponding random variables.
3. Y to denote scalar random variables.
4. y to denote samples from the corresponding scalar random variable or a constant depending on the context.
5. Y_i is i^{th} scalar random variable of a vector random variable $\mathbf{Y}$.
6. y_i is i^{th} pixel represented as realization of a scalar random variable Y.
7. $\mathbb{E}$ is expectation operator (underlying random variable is clear from context or mentioned explicitly), $\odot$ is element-wise (Hadamard) product.

2. DETAILS ON UNBIASED RISK ESTIMATION

2.1. Oracle MSE

Our aim is to obtain an estimate of $\mathbf{x}$ (a realization of $\mathbf{X}$), given the measurement $\mathbf{y}$, which is a realization of random image variable $\mathbf{Y}$. We denote this estimate as a function of observables, $\hat{\mathbf{x}} = \mathbf{f}(\mathbf{Y})$. In general, $\mathbf{f}$ may be any linear or non-linear, parametric or non-parametric function. The criterion which we choose to minimize is the ensemble-averaged mean-square error (or *risk*) between $\mathbf{x}$ and $\hat{\mathbf{x}}$

$$\zeta(\mathbf{f}) = \frac{1}{m}\mathbb{E}\{\|\mathbf{f}(\mathbf{Y}) - \mathbf{x}\|^2\} = \frac{1}{m}\sum_{i=1}^{m}(x_i - f_i(\mathbf{Y}))^2, \tag{1}$$

which requires knowledge of the ground-truth reflectance $\mathbf{x}$. Consider the expansion of Eq. (1) (without the factor $1/m$),

$$\zeta(\mathbf{f}) = \|\mathbf{x}\|^2 + \mathbb{E}\{\|\mathbf{f}(\mathbf{Y})\|^2\} - 2\sum_{i=1}^{m}\mathbb{E}\{x_i f_i(\mathbf{Y})\}, \tag{2}$$

where $f_i(\mathbf{Y})$ denotes the i^{th} entry of the denoised image. Since the optimization is carried out with respect to $\mathbf{f}$, the deterministic (but unknown) factor $\|\mathbf{x}\|^2$ does not play a role (in contrast with the Bayesian framework, where a prior is assumed on $\mathbf{x}$). On the other hand, the term $\mathbb{E}\{x_i f_i(\mathbf{Y})\}$ depends on the unknowns x_i and hence a direct optimization is not possible without knowledge of the ground truth image. Throughout our work, we call this version of the cost the *Oracle MSE estimate*.

2.2. Proof of Corollary 1.1 (our result from main paper)

To estimate the Oracle MSE without the ground truth, unbiased risk estimation methods can be applied. Seelamantula and Blu [1] first presented a surrogate risk for the case of multiplicative Gamma distributed noise model called Multiplicative Unbiased Risk Estimate (MURE). In this section, We present a detailed proof of corollary 1.1 from the main paper.

Theorem 1. *(Multivariate version) Let $\mathbf{Y} = \mathbf{x}\mathbf{N}$, where $\mathbf{x} \in \mathbb{R}_+^{\mathbf{m}}$ is deterministic but unknown reflectance image. Let $\mathbf{Y}, \mathbf{N} \in \mathbb{R}_+^m$, and $\mathbf{N} \sim \Gamma(k, k)$ with independent entries, then, the vector random variable*

$$\hat{\zeta}(\mathbf{f}) = \frac{k}{k+1}\|\mathbf{Y}\|^2 - 2\mathbf{Y}^{\mathrm{T}}\mathcal{M}\mathbf{f}(\mathbf{Y}) + \|\mathbf{f}(\mathbf{Y})\|^2 \tag{3}$$

is an unbiased estimator of the MSE, $\zeta(\mathbf{f}) = \mathbb{E}_{\mathbf{N}}\{\|\mathbf{f}(\mathbf{Y}) - \mathbf{x}\|^{\mathbf{2}}\}$, where $\mathbb{E}$ is the expectation operator. For a scalar function $f(Y)$, the operator $\mathcal{M}$ is defined as $\mathcal{M}f(Y) = k\int_0^1 s^{k-1} f(sY)ds$. This notation is extended straightforwardly to multivariate vector functions $\mathbf{f}(\mathbf{Y}) = [f_1(\mathbf{Y}), f_2(\mathbf{Y}), \ldots, f_m(\mathbf{Y})]^{\mathrm{T}}$ according to $\mathcal{M}\mathbf{f}(\mathbf{Y}) = [\mathcal{M}_1 f_1(\mathbf{Y}), \mathcal{M}_2 f_2(\mathbf{Y}), \ldots, \mathcal{M}_m f_m(\mathbf{Y})]^{\mathrm{T}}$, where $\mathcal{M}_i \mathbf{f_i}(\mathbf{Y})$ applies the operator $\mathcal{M}$ to the i^{th} input component of $\mathbf{f}(\mathbf{Y})$ only.

Corollary 1.1. *(Multivariate, series version of MURE) Let $\mathbf{Y} = \mathbf{x}\mathbf{N}$, where $\mathbf{x} \in \mathbb{R}_+^{\mathbf{m}}$ is deterministic but unknown reflectance image. Let $\mathbf{Y}, \mathbf{N} \in \mathbb{R}_+^m$, and $\mathbf{N} \sim \Gamma(k, k)$ with independent entries, then, the vector random variable*

$$\hat{\zeta}(\mathbf{f}) = \frac{k}{k+1}\|\mathbf{Y}\|^2 + \|\mathbf{f}(\mathbf{Y})\|^2 - 2\sum_{i=1}^{m}\sum_{p=0}^{\infty}(-1)^p \frac{k!}{(k+p)!} Y_i^{p+1} \frac{\partial \mathbf{f}_i^{(p)}(\mathbf{Y})}{\partial Y_i}, \tag{4}$$

is an unbiased estimator of the MSE, $\zeta(\mathbf{f}) = \mathbb{E}_{\mathbf{N}}\{\|\mathbf{f}(\mathbf{Y}) - \mathbf{x}\|^{\mathbf{2}}\}$, where $\mathbb{E}$ is the expectation operator. p is the order of partial derivative of $\mathbf{f}_i(\mathbf{Y})$ w.r.t. $\mathbf{Y}_i$, the i^{th} pixel of the input noisy image.

Proof. We first prove the case for a scalar function $f : \mathbb{R} \to \mathbb{R}$. We have:

$$\mathbb{E}\{(f(Y) - x)^2\} \quad = \quad \mathbb{E}\{f(Y)^2\} + \mathbb{E}\{x^2\} - 2\mathbb{E}\{xf(Y)\}.$$

Expanding term by term,

- $\mathbb{E}\{Y^2\} = \mathbb{E}\{x^2 n^2\} = \frac{k+1}{k}\mathbb{E}\{x^2\}$, and hence

$$\mathbb{E}\{x^2\} = \frac{k}{k+1}\mathbb{E}\{Y^2\}. \tag{5}$$

- We have, $\mathbb{E}\{xf(Y)\}$

$$\begin{aligned}
&= \int_{0^+}^{\infty} x\ f(Y)\ f_N(n) dn \\
&= \int_{0^+}^{\infty} x\ f(Y)\ \frac{k^k}{\Gamma(k)} n^{k-1} e^{-kn} dn \\
&= \int_{0^+}^{\infty} kx\ f(Y)\ \frac{n^k k^k}{\Gamma(k)} \left[\frac{e^{-kn}}{nk}\right] dn \\
&= \int_{0^+}^{\infty} kx\ f(Y)\ \left[\ \int_{0^+}^{1} \frac{n^k k^k}{\Gamma(k)}\ \frac{1}{s^2} e^{-kn/s} ds\right] dn \\
&= \int_{0^+}^{\infty} kx\ \left[\ \int_{0^+}^{1}\ f(xn)\ \frac{ns^{k-1}}{s^2} f_N(n/s) ds\right] dn,
\end{aligned}$$

changing the order of integration and substituting $p = n/s$,

$$
\begin{aligned}
& \int_{0^+}^{\infty} kx \left[\int_{0^+}^{1} f(xn) \, \frac{ns^{k-1}}{s^2} f_N(n/s) ds \right] dn \\
= & \int_{0^+}^{1} \int_{0^+}^{\infty} kxp \; f(xps) \; s^{k-1} f_N(p) \; dp \; ds \\
= & \int_{0^+}^{\infty} Y \left[\; k \int_{0^+}^{1} f(sY) \; s^{k-1} ds \right] \; f_N(p) dp, \\
= & \; \mathbb{E}\{Y \mathcal{M} f(Y)\}. \qquad (6)
\end{aligned}
$$

where we applied change of order of integration (assuming conditions of Fubini's theorem to be true and that the limit exists at 0) again, and substituted $Y = xp$. Using (5) and (6), we have, $\mathbb{E}\{(f(Y) - x)^2\}$

$$
\begin{aligned}
= & \; \mathbb{E}\left\{ f(Y)^2 - 2Y\mathcal{M}f(Y) + \frac{k}{k+1} Y^2 \right\}. \\
= & \; \mathbb{E}\{\hat{\zeta}(f)\}.
\end{aligned}
$$

Which shows that the MURE cost $\hat{\zeta}(f)$ is an unbiased estimator of the oracle cost $\zeta(f)$. For a scalar function $f(Y)$, the operator $\mathcal{M}$ is defined as $\mathcal{M}f(Y) = k\int_0^1 s^{k-1} f(sY) ds$. This notation is extended straightforwardly to multivariate vector functions $\mathbf{f}(\mathbf{Y}) = [f_1(\mathbf{Y}), f_2(\mathbf{Y}), \ldots, f_m(\mathbf{Y})]^{\mathrm{T}}$ according to $\mathcal{M}\mathbf{f}(\mathbf{Y}) = [\mathcal{M}_1 f_1(\mathbf{Y}), \mathcal{M}_2 f_2(\mathbf{Y}), \ldots, \mathcal{M}_m f_m(\mathbf{Y})]^{\mathrm{T}}$, where $\mathcal{M}_i \mathbf{f_i}(\mathbf{Y})$ applies the operator $\mathcal{M}$ to the i^{th} input component of $\mathbf{f}(\mathbf{Y})$ only. Hence, the multivariate result is straightforward to obtain by applying the scalar version of the estimator in Theorem 1 of the main paper to the individual components of the cost function in (1). Thus the cost for a vector function can be written as

$$
\hat{\zeta}(\mathbf{f}) = \frac{k}{k+1} \|\mathbf{Y}\|^2 + \|\mathbf{f}(\mathbf{Y})\|^2 - 2\mathbf{Y}^{\mathrm{T}} \mathcal{M}\mathbf{f}(\mathbf{Y}) \qquad (7)
$$

For the series approximation, we note that the cross term operator $\mathcal{M}$ for a vector to vector function can be expanded by applying the integration-by-parts operation:

$$
\begin{aligned}
\mathcal{M}_i f_i(\mathbf{Y}) &= \mathcal{M}_i f_i(Y_1, Y_2, ..., Y_i, ..., Y_m) \\
&= k \int_0^1 s^{k-1} f_i(Y_1, Y_2, ..., sY_i, ..., Y_m) \, \mathrm{d}s, \\
&= k \int_0^1 s^{k-1} f_i(\mathbf{S}_i \mathbf{Y}) \, \mathrm{d}s, \\
&= s^k f_i(\mathbf{S}_i \mathbf{Y}) \Big|_0^1 - \int_0^1 Y_i f_i^{'}(\mathbf{S}_i \mathbf{Y}) \; s^k \; \mathrm{d}s, \\
&= f_i(Y_i) - \int_0^1 Y_i f_i^{'}(\mathbf{S}_i \mathbf{Y}) \; s^{k+1} \; \mathrm{d}s, \\
&= f_i(\mathbf{Y}) - \left(\frac{s^{k+1}}{k+1} Y_i f_i^{'}(\mathbf{S}_i \mathbf{Y}) \Bigg|_0^1 - \int_0^1 Y_i f_i^{'}(\mathbf{S}_i \mathbf{Y}) \; s^{k+1} \; \mathrm{d}s \right), \\
&= f_i(\mathbf{Y}) - \frac{1}{k+1} Y_i f_i^{'}(\mathbf{Y}) + \frac{1}{k+1} \int_0^1 Y_i f_i^{'}(\mathbf{S}_i \mathbf{Y}) \; s^{k+1} \; \mathrm{d}s), \\
&= \sum_{p=0}^{\infty} (-1)^p \, \frac{k!}{(k+p)!} \; Y_i^p \; \frac{\partial f_i^{(p)}(\mathbf{Y})}{\partial Y_i}, \qquad (8)
\end{aligned}
$$

where $\mathbf{S}_i$ is a matrix constructed by replacing i^{th} diagonal element of identity matrix by s. Finally, writing all terms together for the vector version, we have:

$$\begin{aligned}
\mathbb{E}_{\mathbf{N}}\{\hat{\zeta}(\mathbf{f})\} &= \mathbb{E}_{\mathbf{N}}\left\{\frac{k}{k+1}\|\mathbf{Y}\|^2 + \|\mathbf{f}(\mathbf{Y})\|^2 - 2\sum_{i=1}^{m}\sum_{p=0}^{\infty}(-1)^p \frac{k!}{(k+p)!} Y_i^{p+1} \frac{\partial \mathbf{f}_i^{(p)}(\mathbf{Y})}{\partial Y_i}\right\} \\
&= \mathbb{E}_{\mathbf{N}}\left\{\frac{k}{k+1}\|\mathbf{Y}\|^2 + \|\mathbf{f}(\mathbf{Y})\|^2 - 2\mathbf{Y}^{\mathrm{T}}\mathcal{M}\mathbf{f}(\mathbf{Y})\right\} \\
&= \mathbb{E}_{\mathbf{N}}\left\{\|\mathbf{x}\|^2 + \|\mathbf{f}(\mathbf{Y})\|^2 - 2\sum_{i=1}^{m} x_i f_i(\mathbf{Y})\right\} \\
&= \mathbb{E}_{\mathbf{N}}\{\zeta(\mathbf{f})\}.
\end{aligned}$$

Thus, our proposed series-version of MURE is an unbiased estimator of the Oracle cost in Eq. 1. This completes the proof. □

For all our experiments $n = K = 1$ yields reasonably good approximation with low computational complexity.

3. REFERENCES

[1] Chandra Sekhar Seelamantula and Thierry Blu, "Image denoising in multiplicative noise," in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 1528–1532.