EVALUATION OF NOISE REMOVAL IN SIGNALS BY LULU OPERATORS

I N Fabris-Rotelli, K van Oldenmark and P J van Staden

Department of Statistics, University of Pretoria, 0002, Pretoria, South Africa
inger.fabris-rotelli@up.ac.za

ABSTRACT

The LULU smoothers are effective nonlinear smoothers for signals. We investigate their ability to remove different noise types, namely symmetrical and skewed, as well as heavy- and light-tailed, thereby uncovering the true underlying signal. These smoothers prove very effective in removing noise originating from the same distribution as that noise which originally contaminated the signal.

1. INTRODUCTION

The LULU smoothers for signals (sequences) have been developed over the last three decades by Rohwer and his collaborators, (Rohwer, 2005). For a signal $x = (x_i)_{i=-N}^{N}$, the LULU operators L_n and U_n act at position i in the signal and for $n = 1, 2, 3...$ as follows:

$$(L_n(x))_i = \max\{\min\{x_{i-n}, ..., x_i\}, ..., \min\{x_i, ..., x_{i+n}\}\}, \text{ and}$$

$$(U_n(x))_i = \min\{\max\{x_{i-n}, ..., x_i\}, ..., \max\{x_i, ..., x_{i+n}\}\}.$$

The LULU operators are nonlinear but have very useful properties to their name, that is, they are separators, are total variation preserving and fully trend preserving as defined in Rohwer (2005). However, since $L_n(x) \leq x \leq U_n(x)$ the two operators will produce slightly biased results when used individually, namely, L_n smoothes the signal from above and U_n smoothes from below. We thus use the two together as either $L_n \circ U_n$ or $U_n \circ L_n$. These compositions are also biased, but to a far lesser degree. The Discrete Pulse Transform (DPT) of x, $DPT(x) = (D_1(x), D_2(x), ..., D_N(x))$, is obtained as the iterative application of $L_n \circ U_n$ or $U_n \circ L_n$ for $n = 1, 2, ..., N$. The components D_i are obtained as follows, $D_1(x) = (I - P_1)(x), D_n(x) = (I - P_n) \circ Q_{n-1}(x)$, $n = 2, ..., N$, where $P_n = L_n \circ U_n$ or $P_n = U_n \circ L_n$ and $Q_n = P_n \circ ... \circ P_1$, $n \in \mathbb{N}$. The DPT can be seen as the recursive peeling off of pieces of information of width n - we first remove isolated information of width 1, then of width 2, and so on. For some n the remaining signal is considered sufficiently smoothed (denoised). This optimal n is determined by tracking the total variation removed at each step. The total variation of a signal x is defined as, $TV(x) = \sum_{i=-N}^{N} |x_i - x_{i-1}|$. Since our LULU operators are total variation preserving ($TV(x) = TV(Px) + TV((I - P)x)$ where P is either $L_n \circ U_n$ or $U_n \circ L_n$), we can easily track how much variation remains in the smoothed signal, $TV(Px)$, and how much we remove with each iteration or in total, $TV((I - P)x)$, since no variation is lost at any step. Once the optimal n is decided upon, say n_{opt}, the immediate question to ask is how well has the signal been smoothed or equivalently, how well does that which we have removed, $(I - P_{n_{opt}})x$, represents the noise present in the original signal x? It turns out that the DPT is quite effective in removing impulsive noise. One explanation for this is that linear smoothers aren't well suited to removing noise which arises from a long-tailed probability distribution, (Velleman, 1977), which is characteristic when there are outliers present, nor noise which is signal dependent, (Conradie et al., 2005), whereas the LULU operators are nonlinear thereby avoiding these complications. Here we investigate the ability of the DPT to remove imposed noise and uncover the underlying signal effectively. More specifically, by imposing noise chosen from various distributions, see Table 1, we shall determine if the removed noise $(I - P_{n_{opt}})x$ accurately represents the noise initially imposed.

I N Fabris-Rotelli is the corresponding and presenting author. Thanks to Department of Statistics and Statomet at the University of Pretoria for funding.

2. THE GENERALIZED LAMBDA DISTRIBUTION

In order to simulate noise with various distributional shape properties, we use a parameterization of the Generalized Lambda Distribution (GLD) introduced by van Staden & Loots (2009) and defined through its quantile function (QF) by

$$Q(p) \quad = \quad \begin{cases} \alpha + \beta\big((1-\delta)\big(\frac{p^{\lambda}-1}{\lambda}\big) - \delta\big(\frac{(1-p)^{\lambda}-1}{\lambda}\big)\big) & \text{if } \lambda \neq 0 \\ \alpha + \beta\big((1+\delta)\ln p - \delta \ln(1-p)\big) & \text{if } \lambda = 0 \end{cases}$$

where $0 \leq p \leq 1$, α is a location parameter, $\beta > 0$ is a spread parameter and $0 \leq \delta \leq 1$ and λ are shape parameters. The GLD can be characterized through its first four L-moments, that is, the L-location, L_1 , the L-scale, L_2 and the L-skewness and L-kurtosis ratios, $\tau_3 = L_3/L_2$ and $\tau_4 = L_4/L_2$. As shown in Table 1 and Figure 1, we selected eight distributions from the GLD with different distributional shapes by choosing appropriate values for τ_3 and τ_4 and calculating the corresponding parameter values. All selected distributions were standardized and/or shifted so that $L_1 = 0$ and $L_2 = 1$.

Distribution Shape	**L-moments:** $(L_1, L_2, \tau_3, \tau_4)$	**Parameters of the GLD:** $(\alpha, \beta, \delta, \lambda)$
1. Symmetric, Uniform distribution[1]	$(0, 1, 0, 0)$	$(0, 6, 0.5, 1)$
2. Symmetric, short-tailed	$(0, 1, 0, \frac{1}{12})$	$(0, 2.9989, 0.5, 0.3025)$
3. Symmetric, Normal distribution[2]	$(0, 1, 0, \frac{30}{\pi}\tan^{-1}(\sqrt{2})-9)$	$(0, 0.2449, 0.5, 0.1416)$
4. Symmetric, Logistic distribution (heavy-tailed)[1]	$(0, 1, 0, \frac{1}{6})$	$(0, 2, 0.5, 0)$
5. Symmetric, truncated distribution	$(0, 1, 0, \frac{1}{6})$	$(0, 42, 0.5, 5)$
6. Skewed, Rayleigh distribution[2]	$(0, 1, \frac{3\sqrt{2}+2\sqrt{6}-9}{3(\sqrt{2}-1)}, \frac{20\sqrt{6}-9(4+\sqrt{2})}{6(\sqrt{2}-1)})$	$(-1.0173, 2.6641, 0.7305, 0.2071)$
7. Skewed, Gumbel distribution[2]	$(0, 1, \frac{\ln(9/8)}{\ln 2}, \frac{2\ln(256/243)}{\ln 2})$	$(-1.1157, 2.1486, 0.7723, 0.0487)$
8. Skewed, Exponential distribution (J-shaped)[1]	$(0, 1, \frac{1}{3}, \frac{1}{6})$	$(-2, 2, 1, 0)$

[1] Distribution is special case of the GLD [2] Distribution approximated by the GLD

Table 1. Distributions Chosen to Simulate Noise

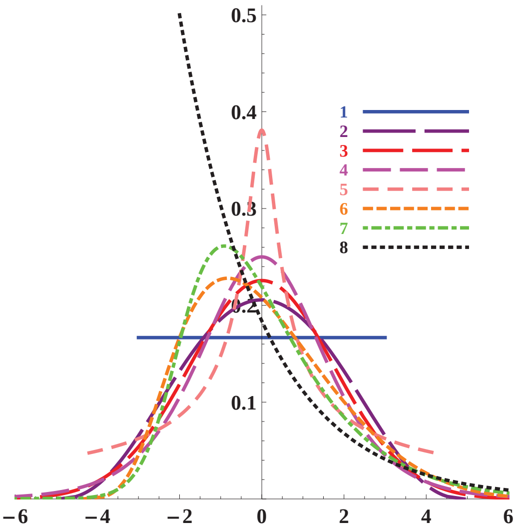

Figure 1. Probability Density Functions of the Noise Distributions in Table 1

The fact that we simulate the noise from a family of distributions, the GLD, with a single functional form as defined through its QF, is important for our investigation as it enables a strongly justified comparison amongst the noise types.

3. SIMULATION

The underlying true signal used was $(s_i) = (a\cos(wi) + b\sin(wi))$ where the parameters a and b are chosen in order to obtain a weak, medium and strong signal respectively with respect to the noise. The period was chosen as 100 throughout, and the frequency w was then calculated through the formula $2\pi/100$. The length of the signal was taken to be 100, that is, 100 data points. For this study we thus for simplicity use the subscripts $1, 2, ..., 1000$ instead of $-N, ..., N$. The amplitude of such a signal is $\sqrt{a^2 + b^2}$, (Foerster, 2002). This signal is periodic and thus has an obvious cyclical trend which enables easy detection of the true signal.

Typically the signal-to-noise ratio (SNR), (Parrish et al., 2000), is used to measure the strength of a signal. The three signals were chosen to have SNR 1, 5 and 9 respectively, which correspond to a weak, medium and strong signal according to the Rose Criterion (Watanabe et al., 2002). As suggested by its name, the SNR is defined as the signal relative to the noise, (see for instance Donoho & Johnstone (1994)). To calculate the SNR, it is common practice to use a measure of location for the signal and a measure of spread for the noise. For example, the SNR can be calculated as the mean signal relative to the standard deviation of the noise (Parrish et al., 2000), and is given by $\text{SNR} = \frac{\text{mean signal}}{\text{std. dev(noise)}}$. We used L_2 as measure of spread for the noise. Recall that we set $L_2 = 1$ for all eight GLDs used to simulate the noise. Since our signals are periodic with zero mean levels, we decided to measure each signal by its amplitude. Hence we calculated the SNR with $\text{SNR} = \frac{\text{amplitude of signal}}{L\text{-scale of noise}}$. So, given $L_2 = 1$ and SNR equal to 1, 5 and 9 respectively, it then follows that the signal parameters a and b are given by 0.5, 4.5 and 8.5, and 0.866, 2.179 and 2.958, respectively, for the weak, medium and strong signals. These three signals are shown in Figure 2.

The DPT was then applied to $(s_i^j + n_i^k), j = 1, 2, 3,\ k = 1, 2, ..., 8$ where (s_i^j) is the j^{th} underlying signal and (n_i^k) is the k^{th} noise signal. In Figure 3, some of these contaminated signals are illustrated. The strength of the signals for the various SNRs can be seen clearly. The DPT was applied in four different ways in order to fully investigate the noise removal and any bias due to the ordering, namely for (1) $L_n \circ U_n \circ L_{n-1} \circ U_{n-1} \circ ... \circ L_1 \circ U_1$, (2) $U_n \circ L_n \circ U_{n-1} \circ L_{n-1} \circ ... \circ U_1 \circ L_1$, (3) $U_n \circ L_n \circ L_{n-1} \circ U_{n-1} \circ ... \circ U_1 \circ L_1$, and (4) $L_n \circ U_n \circ U_{n-1} \circ L_{n-1} \circ ... \circ L_1 \circ U_1$. We shall use the notation LULU, ULUL, LUUL and ULLU for these. The last two options are called the *alternating bias operators* since they alternately swop between the two basic choices $L_n \circ U_n$ and $U_n \circ L_n$. See Jankowitz (2007) for other possibilities of reducing the bias.

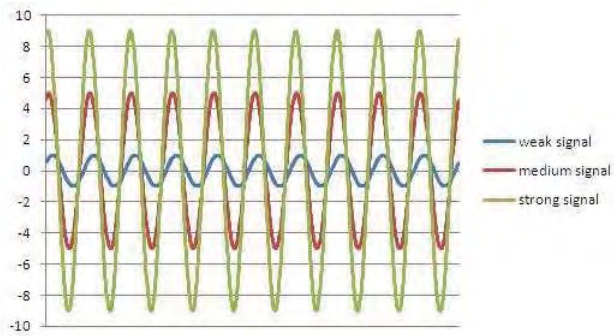

Figure 2. Original Signals with SNR 1 (weak), 5 (medium) and 9 (strong)

For the three signals we thus apply the DPT with respect to (1)-(4) for the 8 different noise types. The total variation is tracked throughout the DPT and the cumulative noise removed for (a) n where half of the added total variation has been removed, and for (b) n_0 where all the added total variation has been removed, is investigated. We investigate (a) as it is understood that the most disruptive noise occurs in the first levels of the DPT, and (b) because this is where it is naturally thought that the original signal should be uncovered. The respective true noise distribution for $k = 1, 2, ..., 8$ is fitted to these noise samples using method of L-moment estimation, (van Staden & Loots, 2009), to investigate if the noise removed up to the two respective points is distributed similarly to the original noise imposed.

Noise Type	1	2	3	4	5	6	7	8
SNR = 1	2050	2066	2072	2078	2075	2056	2056	2039
SNR = 5	2055	2070	2077	2083	2082	2061	2062	2047
SNR = 9	2070	2087	2094	2101	2101	2078	2079	2071

Table 2. Total Variation (rounded) of the 24 Contaminated Signals

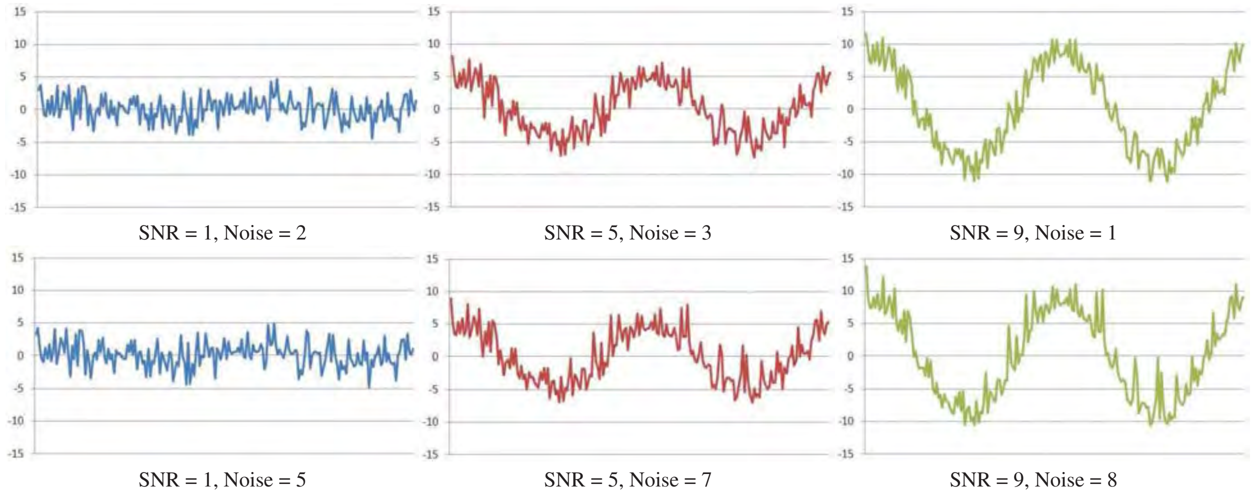

Figure 3. A Sample of the 24 Different Contaminated Signals

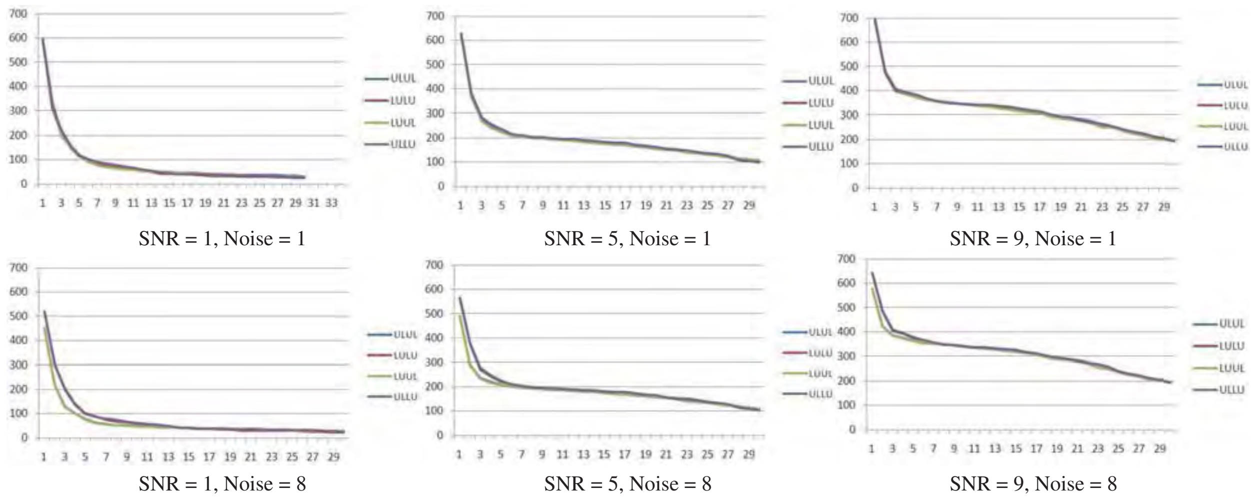

Figure 4. Total Variation Removed at each Level of the DPT for Noise Type 1 (Noise Types 1-5 are similar) and 8 (Noise Types 6-8 are similar)

4. LULU SMOOTHING OF THE CONTAMINATED SIGNALS

Due to the total variation preservation of the DPT, total variation is a good measure to track the smoothing process over n, i.e. from level to level of the DPT. When the total variation removed with each n stabilizes, i.e. doesn't change significantly from n to $n+1$, the added noise has been removed effectively. From the results of the 24 different contaminated signals, the total variation removed at each level remains similar whichever combination, LULU, ULUL, LUUL or ULLU, is used to obtain the DPT, and for each of the three SNRs. A slight difference can only be seen in the three skewed noise types, namely types 6, 7 and 8. It can be seen in Figure 4 how the total variation progresses through the DPT levels. The differences seen between the different SNRs are due to the fact that the weak, medium and strong original signals have a total variation of 40.491, 204.487 and 368.423 respectively, thus the smoothing (decrease in total variation) occurs sharply up until that point and then stabilizes. The contaminated signals have total variation as indicated in Table 2. Comparing Table 2 with Figure 4 it can be seen a huge proportion of the total variation is removed in the first level of the DPT as the remaining total variation drops to around 600 in all cases. The stabilization of the total variation removal varies for the three SNRs investigated. For the weak signal (SNR $= 1$) half the total variation is removed at around $n = 3$ and all the added total variation (i.e. at stabilization) is removed by around $n = 14$ for LULU and ULLU and by around $n = 19$ for LUUL and ULUL.

For the medium signal (SNR $= 5$) half the total variation is removed at around $n = 3$ as well and all the added total variation (i.e. at stabilization) is removed by around $n = 6$. For the strong signal (SNR $= 9$) half the total variation is removed at around $n = 2$ and all the added total variation (i.e. at stabilization) is removed by around $n = 5$. It would thus seem that the stronger the signal (or the weaker the noise) the quicker and more effective the noise removal, and also that the bias between the four L_n and U_n combinations decreases.

To investigate whether the smoothed signal obtained when the total variation stabilizes does in fact resemble the original uncontaminated signal, the MSE measure was used to calculate the differences between the smoothed signal through the DPT levels and the original signal. The MSE is calculated as

$$MSE(x) = \frac{\sum_{i=1}^{1000}(x_i - \bar{x})^2}{1000}.$$

For SNR $= 1$ the combination ULLU provides the lowest MSE from the beginning of the smoothing process. The combination ULUL gives the highest MSE although the differences between the combinations are not drastic. For SNR $= 5$ and 9 the same is seen. See Figure 5 for the MSE for noise type 1. The medium and strong signal give very interesting results for the MSE measurements. It can be seen in Figure 5 that the MSE starts to increase from level 14 of the DPT onwards. This indicates that from this point onwards the smoothing process begins to smooth out the uncovered original signal instead of the noise. As the SNR increases the MSE in the beginning levels of the DPT is more similar for the four L_n and U_n combinations, see Figure 5.

In Figure 6, the smoothed signals can be visually analysed. The higher the SNR the more effective the noise removal, i.e. the more the smoothed signal resembles the original signal.

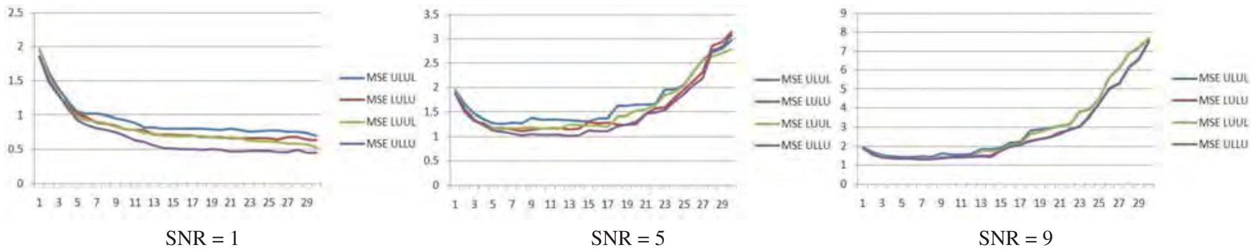

Figure 5. MSE at each Level of the DPT for Noise Type 1 (other noise types are similar)

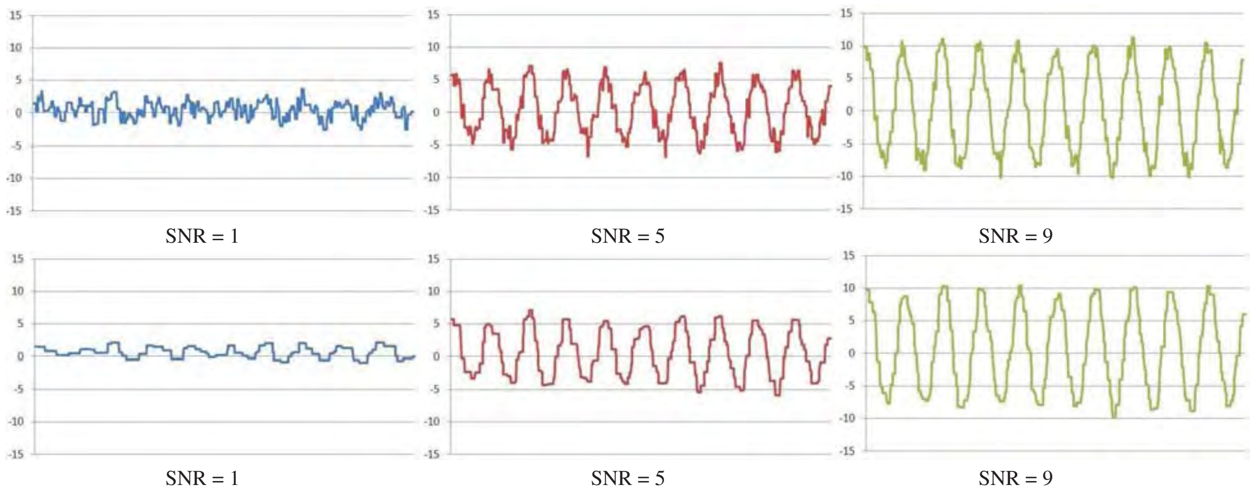

Figure 6. The Smoothed Signals for Noise Type 1 and using LULU (all noise types and L_n, U_n combinations are visually similar): First row indicates the smoothed signal when half the added TV has been removed, Second row indicates the smoothed signal when all the added TV has been removed

5. NOISE REMOVAL OF THE LULU SMOOTHERS

As discussed in Section 3, the cumulative noise removed when (a) half of the added total variation has been removed, and (b) when all the added total variation has been removed, is investigated. The original noise distributions were fitted to this cumulative removed noise. The L-moments and four parameters of the GLD were then compared to evaluate the fit of the removed noise. The results for the L-location, L_1, and the L-skewness ratio, τ_3, are given in Tables 3 and 4. The following observations can be made:

- From Table 3 we see that LULU and ULLU result in a negative shift in location for each SNR, although the shift seems to reduce from noise type 1 through to 8. The shift in location is of course due to the biasedness of the various smoothers, already discussed in Section 1. For ULUL and LUUL the shift in location is seen to be positive for each SNR, but again generally increases from noise type 1 through to 8. An interesting phenomenon can be seen for noise types 6 and 7 for ULLU. For noise type 7 we see a shift of 0 for SNR = 1 and 5, but a negative shift for SNR = 9. Noise distributions 6 and 7 are very similar and thus behave similarly. It is thus evident that the smaller SNRs result in poorer removal of noise type 7, i.e. do not as effectively remove the same noise that was imposed.

	Noise Type	1	2	3	4	5	6	7	8
SNR = 1	LULU	-0.70	-0.60	-0.56	-0.50	-0.43	-0.37	-0.23	0
	LUUL	0.40	0.35	0.30	0.27	0.23	0.53	0.58	0.80
	ULLU	-0.45	-0.38	-0.36	-0.34	-0.29	**-0.16**	**0**	0.27
	ULUL	0.68	0.56	0.51	0.45	0.37	0.73	0.77	1
SNR = 5	LULU	-0.69	-0.62	-0.59	-0.58	-0.54	-0.40	-0.27	0
	LUUL	0.44	0.40	0.38	0.36	0.36	0.60	0.68	0.95
	ULLU	-0.43	-0.41	-0.40	-0.39	-0.37	**-0.20**	**0**	0.22
	ULUL	0.62	0.56	0.52	0.49	0.47	0.74	0.80	1.06
SNR = 9	LULU	-0.76	-0.69	-0.66	-0.63	-0.62	-0.46	-0.33	0
	LUUL	0.49	0.46	0.45	0.44	0.45	0.64	0.72	1.02
	ULLU	-0.57	-0.53	-0.51	-0.49	-0.48	**-0.32**	**-0.20**	0.10
	ULUL	0.67	0.62	0.60	0.57	0.57	0.79	0.85	1.14

Table 3. L-Location Moments of the Fully Removed Noise (Grey: Negative Change, White: Positive Change, In Bold: interesting case)

- The shift in L-location must be considered simultaneously with the change in the L-skewness ratio, as seen in Table 4, since a change in the level of skewness of a distribution will result in a shift in the location of that distribution. There we see a decrease in the L-skewness ratio for LULU and ULLU. This decrease becomes less prominent as SNR increases however. See Figure 7 for the fitted and original distributions for noise types 3 and 7. The shift in location and change in skewness can be seen. The changes are due to the fact that U_n is applied first in LULU and ULLU. The operator U_n removes negative pulses and thus the removed noise favours slightly the negative direction. For ULUL and LUUL there is a general increase in the L-skewness ratio for noise types 1 to 5, and the trend becomes stronger as SNR increases. The changes are due to the fact that L_n is applied first in ULUL and LUUL. The operator L_n removes positive pulses and thus the removed noise favours slightly the positive direction. For noise types 6 to 8, a decrease in the L-skewness ratio is still observed. This is due to the fact that these noise distributions are already positively skewed and thus contain more negative pulses from the start. Although the above discussed change in the L-skewness ratio is clear, the change is very slight, as can be seen by the values in Table 4.

- The fitted distributions fit the original distribution very well when half the added total variation has been removed. Furthermore, in general the fit improves towards the full removal of the added total variation.

- The L-scale does not vary significantly at all for any of the fits and for each SNR investigated. This is an important result as it indicates that the removed noise has very similar spread to the noise which was initially imposed on the signals and almost none of its variation has been left in the smoothed signal.

- The L-kurtosis ratio also does not vary significantly at all for any of the fits and for each SNR investigated, indicating the mass in the tails and centre for each distribution of the imposed noise has been removed intact.

	Noise Type	1	2	3	4	5	6	7	8
	True Skewness Value	**0**	**0**	**0**	**0**	**0**	**0.114**	**0.17**	**0.333**
SNR = 1	LULU	-0.001	-0.001	-0.003	-0.005	-0.006	0.105	0.161	0.301
	LUUL	0.005	0.003	0.002	-0.002	0.003	0.104	0.154	0.29
	ULLU	-0.003	-0.003	-0.004	-0.005	-0.002	0.102	0.157	0.297
	ULUL	0.009	0.006	0.005	0.002	0.004	0.113	0.168	0.315
SNR = 5	LULU	-0.006	-0.0004	0.002	0.005	0.003	0.093	0.148	0.256
	LUUL	0.001	-0.0001	-0.0001	-0.002	0.003	0.086	0.139	0.271
	ULLU	-0.004	-0.002	0.001	0.003	0.003	0.09	0.145	0.256
	ULUL	0.003	0.0004	-0.0002	-0.002	0.004	0.087	0.14	0.268
SNR = 9	LULU	0.00008	0.002	0.006	0.006	0.007	0.092	0.143	0.24
	LUUL	-0.001	0.004	0.005	0.007	0.009	0.087	0.144	0.255
	ULLU	0.001	0.005	0.008	0.007	0.008	0.086	0.137	0.232
	ULUL	0.002	0.003	0.003	0.004	0.006	0.086	0.144	0.259

Table 4. L-Skewness of the Fully Removed Noise (Grey: decrease in skewness, White: increase in skewness)

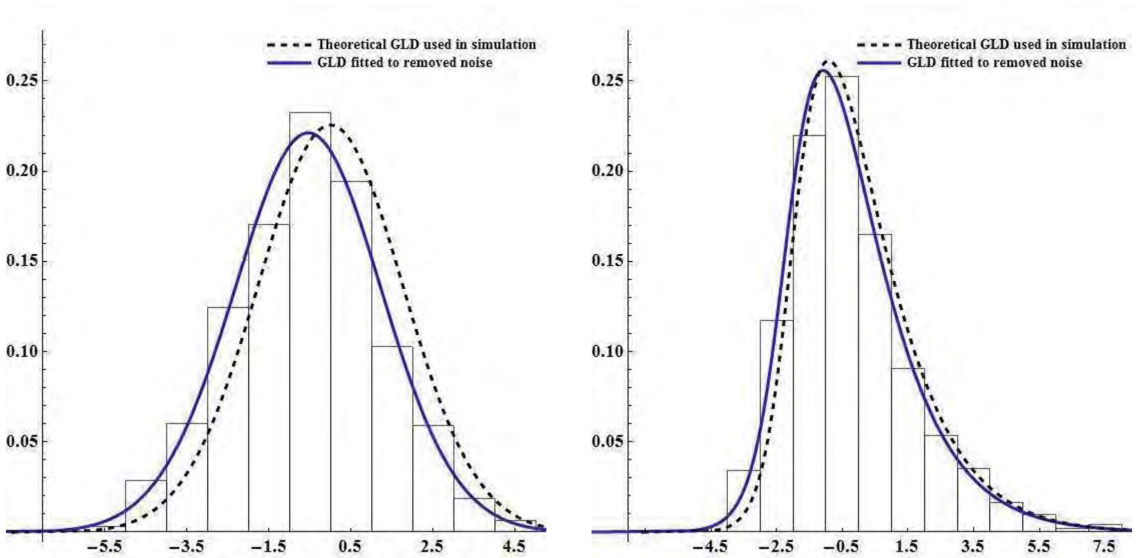

Figure 7. Theoretical and Fitted Noise Distributions

6. CONCLUSION

The ability of the LULU smoothers to remove the different noise types is very effective and from the results we see that the noise removed is distributed similarly to the noise originally imposed. The underlying smoothed signal is also effectively uncovered when the total variation removed at each step begins to stabilize. The effect of the different combinations of L_n and U_n produce interesting results as indicated in Section 5. Future work will look at implementing more effective combinations of L_n and U_n to reduce the bias, such as those in Jankowitz (2007). As expected the fit of the noise removed improves as n increases towards the optimum n_0.

A further possibility for this study is to investigate using as a measure of smoothing the number of pulses removed at each level of the DPT, i.e. at each n, and to compare this with using the total variation as a measure for this purpose.

REFERENCES

Conradie, W. J., de Wet, T., & Jankowitz, M. D. (2005). An overview of LULU smoothers with application to financial data. *Journal for Studies in Economics and Econometrics*, 29(1), 97–121.

Donoho, D. L. & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3), 425–455.

Foerster, P. A. (2002). *Precalculus with Trigonometry: Concepts and Connections*. Birkhuser.

Jankowitz, M. D. (2007). *Some Statistical Aspects of LULU Smoothers*. PhD thesis, Department of Statistics and Actuarial Science, Stellenbosch University, South Africa.

Parrish, T. B., Gitelman, D. R., LaBar, K. S., & Mesulam, M.-M. (2000). Impact of signal-to-noise on functional MRI. *Magnetic Resonance in Medicine*, 44, 925–932.

Rohwer, C. H. (2005). *Nonlinear Smoothers and Multiresolution Analysis*. Birkhäuser.

van Staden, P. J. & Loots, M. T. (2009). Method of L-moment estimation for the generalized lambda distribution. In *Proceedings of the Third Annual ASEARC Conference* Newcastle, Australia.

Velleman, P. F. (1977). Robust nonlinear data smoothers: Definitions and recommendations. *Proceedings of the Natural Academy of Sciences*, 74(2), 434–436.

Watanabe, M., Williams, D. B., & Tomokiyo, Y. (2002). Comparison of detection limits for elemental mapping by EF-TEM and STEMXEDS. *Microscopy and Microanalysis*, 8, 1588–1589.