2015 International Symposium on Nonlinear Theory and its Applications
NOLTA2015, Kowloon, Hong Kong, China, December 1-4, 2015

# Improvements to local projective noise reduction through higher order and multiscale refinements

Jack Murdoch Moore*†‡, Ali Karrech§ and Michael Small‖

* jack.moore@research.uwa.edu.au
† School of Earth and Environment, The University of Western Australia
‡ Mineral Resources, CSIRO
§ School of Civil, Mining and Environmental Engineering, The University of Western Australia
‖ School of Mathematics and Statistics, The University of Western Australia

**Abstract**—Local projection is a particularly effective method of reducing noise in a nonlinear signal while preserving its structure. [1] Kantz and Schreiber [2] established a second order, single scale refinement of the first order, single scale local projective filter of [3, 4]. Their method better accommodates the geometry of a strange attractor underlying data and is implemented in the nonlinear time series analysis package TISEAN. [2] This work extends the refinement to higher order and multiple scales, thereby providing an analytic motivation for the original refinement. Additionally, statistical analysis explains the original refinement's high efficacy. The novel refinements identified in this work can achieve more effective noise reduction than established local projective filters.

## 1. Introduction

Linear filters obscure the structure underlying nonlinear signals. [5] The filtering of nonlinear signals requires techniques specifically developed for that purpose: nonlinear filters, reviewed by [6]. A useful class of these are local projective filters, reviewed by [7]. Local projective filters correct the location of points in phase space informed by the locations of neighbouring points. Since their inception by Cawley and Hsu [3]; Sauer [4], local projective filters have found continuous employment in spheres ranging from the supernally immediate one containing our neurons [8] to the study of light curves from distant stars within the discipline which encompasses our entire universe [9].

Kantz and Schreiber [2] achieved more precise results by better accommodating attractor geometry. They chose as the origin of the projective subspace a weighted sum of the first and second order centres of mass of local phase space neighbourhoods (see figure 2.1).

This paper describes an analytic motivation for the refinement of [2] from which further higher order and multiple scale refinements naturally follow. In many cases, the extensions can reduce noise more effectively than established local projective filters.

Section 2 describes established forms of local projection, including the second order refinement of [2]. Section 3 derives the higher order and multiscale filters proposed in this paper. It goes on to explore the way in which considerations of independently distributed measurement errors may be incorporated into previous multiscale and higher order results. Section 4 introduces the classical nonlinear systems used to assess the filters. Section 5 summarises and suggests ways in which results could be extended.

The paper omits some details of the local projective noise reduction process. For a complete description see [10], which also includes an application to experimental data.

## 2. Background

### 2.1. Local projection

Consider a series of $N$ scalar measurements $s(1), s(2), \ldots, s(N) \in \mathbb{R}$ from a chaotic system. The $N-(m-1)T$ vectors $\mathbf{x}_i \triangleq (s(i), s(i+T), \ldots, s(i+(m-1)T)^T \in \mathbb{R}^m$ are called $m$-histories and can be used to reconstruct phase space [11, 12, 13].

Local projection begins with the transformation of the $m$-histories $\mathbf{x}_i$ via an invertible matrix $\mathbf{R}$ according to

$$\mathbf{z}_i \triangleq \mathbf{R}\mathbf{x}_i. \tag{1}$$

Local projection of an $m$-history $\mathbf{x}_i$ involves replacing $\mathbf{z}_i$ with $\mathbf{z}_i^o + \mathbf{P}_i^{(q)}\left(\mathbf{z}_i - \mathbf{z}_i^o\right)$ (or, equivalently, $\mathbf{x}_i$ with $\mathbf{x}_i^o + \mathbf{R}^{-1}\mathbf{P}_i^{(q)}\left(\mathbf{R}\left(\mathbf{x}_i - \mathbf{x}_i^o\right)\right)$), where $\mathbf{P}_i^{(q)}$ is the matrix representing projection onto the subspace spanned by the eigenvectors corresponding to the $m-q$ largest eigenvalues of the covariance matrix $\mathbf{C}_i \triangleq \sum_{j=1}^{\nu_i}\left(\mathbf{z}_i^{(j)} - \mathbf{z}_i^o\right)\left(\mathbf{z}_i^{(j)} - \mathbf{z}_i^o\right)^T$, $\mathbf{x}_i^{(1)}, \mathbf{x}_i^{(2)}, \ldots, \mathbf{x}_i^{(j)}$ denote the $j$ nearest neighbours of $\mathbf{x}_i$, $\mathbf{z}_i^{(j)} \triangleq \mathbf{R}\left(\mathbf{x}_i^{(j)}\right)$, $\nu_i$ is the number of nearest neighbours of $\mathbf{x}_i$ considered, $\mathbf{x}_i^o$ is the origin of the coordinate sytem in which the projection is performed, and $\mathbf{z}_i^o \triangleq \mathbf{R}\left(\mathbf{x}_i^o\right)$.

Cawley and Hsu [3]; Sauer [4] chose as the origin of the coordinate system $\mathbf{x}_i^o = \overline{\mathbf{x}}_i \triangleq \frac{1}{\nu_i}\sum_{j=1}^{\nu_i}\mathbf{x}_i^{(j)}$, the centre of mass of the nearest neighbours of $\mathbf{x}_i$. Other choices of $\mathbf{x}_i^o$ can yield superior results in local projection; these lead to the higher order and multiple scale filters discussed in this paper.

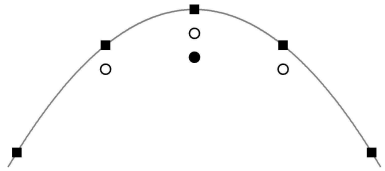

Figure 2.1: Following figure 10.3 of [2]. The centres of mass of local neighbourhoods (∘) do not, in general, lie exactly on the attractor. The second order centres of mass (•) tend to aberrate roughly twice as much.

### 2.2. A second order refinement

Note that the second order centre of mass $\overline{\overline{\mathbf{x}}}_i$ refers to the mean over the $\nu_i$ nearest neighbours of $\overline{\mathbf{x}}_i$ among the neighbourhood centres of mass $\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2, \ldots, \overline{\mathbf{x}}_{N-(m-1)T}$. Kantz and Schreiber [2] observed that the difference between $\mathbf{x}_i$ and the mean over nearest neighbours, $\mathbf{x}_i - \overline{\mathbf{x}}_i$, is roughly half the difference between $\mathbf{x}_i$ and the second order centre of mass, $\mathbf{x}_i - \overline{\overline{\mathbf{x}}}_i$ (see figure 2.1). They deduced that $2\overline{\mathbf{x}}_i - \overline{\overline{\mathbf{x}}}_i$ is an estimate of $\mathbf{x}_i$ in the absence of noise, and hence can be a better point from which to apply local projection than is $\overline{\mathbf{x}}_i$. Figure 2.1 illustrates this observation.

## 3. Analytic motivation

### 3.1. Higher order refinements

For real numbers $\delta > 0$, define the continuous moving average operator $I_\delta$ by $(I_\delta f)(x) = \dfrac{1}{2\delta}\displaystyle\int_{x-\delta}^{x+\delta} f(t)dt$. Kantz and Schreiber [2] achieve a frequently more precise estimate of the $m$-history corresponding to a clean signal by considering a linear combination of centres of mass of first and second order. In analogy, require

$$f_n(0) = a_1 \left(I_\delta f_n\right)(0) + a_2 \left({I_\delta}^2 f_n\right)(0) \tag{2}$$

to hold for all monomials $f_n(x) \triangleq \begin{cases} 1, & n = 0 \\ x^n, & n \neq 0 \end{cases}$ with $0 \leq n \leq n_{\max}$, and with $n_{\max}$ as large as possible. For (positive) odd $n$, both the left and right of (2) are zero. Setting $n = 0$ and $n = 2$ in (2) yields a system of two linear equations with the unique solution $a_1 = 2$ and $a_2 = -1$. This leads to the second order refinement of [2].

For $k = 1, 2, \ldots, 20$, the unique solution to the system of $k$ linear equations

$$f_n(0) = \sum_{i=1}^{k} a_i \left({I_\delta}^i f_n\right)(0) \qquad n = 0, 2, \ldots, 2(k-1) \tag{3}$$

is $a_i = (-1)^{i-1}\binom{k}{i}$, for $i = 1, 2, \ldots, k$.

### 3.2. Multiscale refinements

Requiring that

$$f_n(0) = a_2 \left({I_\delta}^2 f_n\right)(0) + b_1 \left(I_{\beta\delta} f_n\right)(0) \tag{4}$$

holds for integers $n$ such that $0 \leq n \leq 5$ (with $\beta \geq 0$) gives a solution $a_2 = 4, b_1 = -3, \beta = \sqrt{8/3}$.

As will be illustrated in subsection 4.2, the multiscale filter corresponding to (4) for $0 \leq n \leq 5$ often performs worse than the second order, single scale refinement of [2], and rarely performs considerably better. This is surprising, since the multiscale filter should be better able to approximate the geometry of the underlying attractor. The next section explains the result by showing that the second order, single scale filter of [2] attenuates errors particularly effectively.

### 3.3. Attenuating error

The choice of coefficients ($a_2 = -1, b_1 = 2$) and scale ($\beta = 1$) corresponding to the refined filter of [2] not only preserves the component of the geometry of the underlying attractor described by a multivariate Taylor polynomial of degree up to three in each variable, but also attenuates errors particularly effectively.

For non-negative integer $d$ define the discrete moving average operator $J_d$ by $(J_d f)(j) = \dfrac{1}{2d+1}\displaystyle\sum_{i=-d}^{d} f(j+i)$. Let $e(i) \in \mathbb{R}$ be independent random variables with mean zero and unit variance. The $e(i)$ can be thought of as errors; in this section is sought a choice of $a_2, b_1, \beta$ for which they are attenuated by the operator $a_2 {J_d}^2 + b_1 J_{\beta d}$.

To preserve the geometry of the underlying attractor, require (4) to hold for $0 \leq n \leq 3$, yielding $a_2 = A(\beta) \triangleq \frac{\beta^2}{-2+\beta^2}$ and $b_1 = B(\beta) \triangleq 1 - \frac{\beta^2}{-2+\beta^2}$. Consider the expectation of the square of the filtered error at $j = 0$, $S_{a_2,b_1,\beta,d} \triangleq \left\langle \left( \left( \left( a_2 {J_d}^2 + b_1 J_{\beta d} \right) e \right)(0) \right)^2 \right\rangle$, with these geometric constraints on $a_2, b_1, \beta$.

The desirable case of a dense embedding corresponds to a large window $d$. However, in the limit $d \to \infty$, $S_{a_2,b_1,\beta,d} \to 0$. Therefore, normalise the expectation $S_{a_2,b_1,\beta,d}$ by that corresponding to prototypical local projection. This gives $T_{a_2,b_1,\beta} \triangleq \lim\limits_{d\to\infty} \dfrac{S_{a_2,b_1,\beta,d}}{S_{1,0,\beta,d}}$, called the normalised expected error.

As figure 3.3 shows, $T_{A(\beta),B(\beta),\beta}$ has a local minimum at $\beta = \beta^* \approx 0.96215$, for which $a_2 = a_2^* \triangleq A(\beta^*) \approx -0.86174, b_1 = b_1^* \triangleq B(\beta^*) \approx 1.86174$. These values are suggestively close to the coefficients of the refined local projective filter of [2], namely $a_2 = -1, b_1 = 2, \beta = 1$. Indeed, as figure 3.3 shows, the ratio of normalised expected errors is close to unity, with $T_{2,-1,1}/T_{a_2^*,b_1^*,\beta^*} \approx 1.004$.

## 4. Results and Discussion

Following [14, 2], the matrix $\mathbf{R}$ of (1) is diagonal with $R_{11} = R_{mm} = 10^3$ and other entries of the main diagonal set to 1. Again following [14, 2], the highest magnitude corrections are assumed deleterious and rescaled to the mean magnitude. Specifically, the largest 5% are rescaled. Fuller

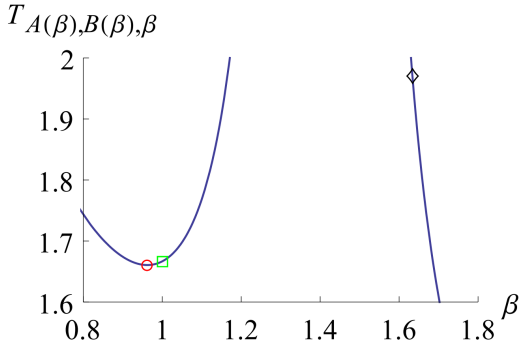


Figure 3.1: Normalised expected error as a function of scale factor $\beta$ for local projective filters satisfying (4) for $0 \leq n \leq 3$. (b) The local minimum at $\beta \approx 0.96215$, marked ○, is close to $\beta = 1$ (corresponding to [2]), marked □. The normalised expected error corresponding to $\beta = \sqrt{8/3}$ is marked ◇.

details of parameter choices and their motivations, and a more complete description of some steps of the local projection process, may be found in [10].

### 4.1. Benchmarks

Following [7], the filters considered in this paper are benchmarked using time series from $x$-coordinates of the [15] system $\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} \sigma(y-x) \\ -y + rx - xz \\ -bz + xy \end{pmatrix}$ with $\sigma = 10$, $r = 28$ and $b = 8/3$ and a time step of $\Delta t = 0.2$ (the large value helps to avoid the advantages of temporal correlation and ensure that methods rely on attractor geometry); the Henon map $\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 1 - ax_t^2 + y_t \\ bx_t \end{pmatrix}$ with $a = 1.4$ and $b = 0.3$; and the Ikeda map $\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 1 + 0.9\,(x_t \cos\phi_t - y_t \sin\phi_t) \\ 0.9\,(x_t \sin\phi_t + y_t \cos\phi_t) \end{pmatrix}$, where $\phi_t = 0.4 - 6/\left(1 + x_t^2 + y_t^2\right)$. Each time series is normalised to unity standard deviation, whereafter Gaussian noise is added.

### 4.2. Results of benchmarking

Table 4.2 presents the results of filtering noisy nonlinear systems with local projective filters of a single scale and orders 1 (corresponding to the prototypical local projection of [3, 4]), 2 (corresponding to the refinement of [2]), up to 5, as well as dual scale local projective filters derived from purely geometric and from simultaneous geometric and statistical considerations. The four most frequently optimal filters are the single scale filters of orders 1, 2 and 3 and the filter derived via considerations of the normalised expected error.

As seen in table 4.2 the multiscale filter corresponding to (4) for $0 \leq n \leq 5$ often performed worse than the second

| | | Order: | 1 | 2 | 3 | 4 | 5 | 2 | 2 |
|---|---|---|---|---|---|---|---|---|---|
| | | Scales: | 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| System | $N$ | $n_{\max}$: | 1 | 3 | 5 | 7 | 9 | 5 | 3 |
| Henon | 5000 | 100%[a] | **1.72** | 1.53 | 1.35 | 1.22 | 1.30 | 1.53 | *1.59* |
| | | 30% | 2.27 | *2.31* | 2.23 | 1.99 | 1.82 | 2.30 | **2.37** |
| | | 10%[a] | 3.19 | 3.46 | 3.45 | 3.39 | 3.26 | **3.53** | *3.50* |
| | | 3% | 3.23 | *3.73* | 3.70 | 3.69 | 3.68 | 3.53 | **3.74** |
| | | 1%[a] | 2.92 | **3.43** | 3.42 | 3.39 | 3.27 | 2.97 | *3.43* |
| | 20000 | 10% | 3.28 | *3.74* | 3.56 | 3.37 | 3.20 | 3.69 | **3.74** |
| | | 1%[a] | 3.46 | **4.10** | *4.07* | 4.01 | 3.88 | 3.84 | 4.04 |
| Ikeda | 5000 | 100% | **1.95** | 1.89 | 1.63 | 1.38 | 1.43 | *1.94* | 1.91 |
| | | 30% | *1.65* | 1.49 | 1.28 | 1.36 | 1.49 | **1.67** | 1.65 |
| | | 10% | 1.84 | 2.03 | **2.10** | 2.03 | 1.98 | 1.87 | *2.05* |
| | | 3% | 1.52 | 1.73 | **1.75** | 1.69 | 1.61 | 1.49 | *1.73* |
| | | 1% | 1.22 | **1.31** | 1.28 | 1.22 | 1.22 | 1.14 | *1.31* |
| | 20000 | 10%[a] | 2.34 | 2.63 | **2.80** | 2.56 | 2.38 | 2.45 | *2.66* |
| | | 1%[a] | 1.87 | *2.31* | **2.36** | 2.25 | 2.09 | 1.82 | 2.31 |
| Lorenz (1963) | 5000 | 100% | **1.69** | 1.37 | 1.24 | 1.39 | *1.53* | 1.40 | 1.46 |
| | | 30% | 1.96 | *2.11* | 1.95 | 1.64 | 1.72 | 1.73 | **2.14** |
| | | 10%[a] | 1.83 | 2.04 | *2.05* | **2.06** | 1.99 | 1.88 | 2.04 |
| | | 3% | 1.64 | 1.80 | **1.85** | *1.82* | 1.75 | 1.66 | 1.81 |
| | | 1% | 1.40 | *1.57* | **1.57** | 1.52 | 1.49 | 1.41 | 1.56 |
| | 20000 | 10% | 1.98 | 2.17 | **2.20** | 2.16 | 2.07 | 1.98 | *2.18* |
| | | 1% | 1.70 | 1.94 | **1.97** | 1.89 | 1.81 | 1.74 | *1.95* |

[a] Considered by [7] (see their table I).

Table 4.1: Proportional increase in signal to noise ratio over eight iterations, $\epsilon_0/\epsilon_8$. Higher order filters have coefficients satisfying (3), and multiple scale filters have coefficients satisfying (4), for $0 \leq n \leq n_{\max}$. The final column corresponds to the filter of subsection 3.3, with coefficients chosen to the minimise normalised expected error. The results corresponding to the most and second most effective filters are, respectively, **bold** and *italic*.

order, single scale refinement of [2], and rarely performed considerably better. However, table 4.2 also shows that the multiscale filter identified through simultaneous geometric and error attenuation considerations (with parameters $a_2 = a_2^*, b_1 = b_1^*, \beta = \beta^*$) almost always performed similarly to, but usually marginally better than, the second order, single scale refinement of [2] (with parameters $a_2 = -1, b_1 = 2, \beta = 1$). This is consistent with the filters' equivalent geometric motivations and similar normalised expected errors (see subsection 3.3).

Table 4.2 shows that as noise level increases from the lowest level considered, 1%, higher order filters become relatively more effective. This trend terminates at a higher noise level, from which the filter of [2] and the filter derived in subsection 3.3 via considerations of error tend to be optimal. The highest noise level considered, 100%, favours prototypical local projective filters. The pattern can be rationalised by considering the relative importance of geometry preservation and error reduction at different noise levels. At low noise levels, geometry is critical and so filters of order two or greater are more effective. Higher order filters become relatively more effective as the relatively low level of noise increases. This may be because, for a given neighbourhood size, higher order centre of mass

measurements incorporate more observations. At intermediate noise levels both error attenuation and geometry are significant. Hence, the filters derived from geometric considerations which also correspond to a point near the minimum of the normalised expected error $T_{A(\beta),B(\beta),\beta}$ of subsection 3.3 are optimal. At high levels of noise when geometry is less important, prototypical local projection, which corresponds to the still lower value of normalised expected error $T_{0,1,\beta} = 1$ (compare with figure 3.1), is most effective.

## 5. Conclusion

This paper examined the refined local projective filter of [2], leading to series of filters involving different numbers of length scales and orders of centre of mass operators. The new filters were compared to established ones. Higher order filters were relatively more effective for data of greater length or corresponding to a higher underlying attractor dimension.

When reformulated for multiple scales, the geometric considerations which led to effective higher order filters did not consistently improve upon existing filters. However, the same geometric criteria, augmented by statistical analysis designed to minimise error, led to filters usually superior to established filters. In addition, the multiscale statistical and geometric considerations explicate the efficacy of the refinement of [2].

As noise level increased from a low value (1%), higher order filters initially became relatively more effective. At intermediate noise levels (∼30%), geometrically motivated filters corresponding to high error attenuation (including a new, dual scale filter) began to dominate until, at the highest noise level (100%), prototypical local projective filters prevailed.


## Acknowledgements

JMM thanks Thomas Poulet, Klaus Regenauer-Lieb, Manolis Veveakis, Jürg Hauser, and the anonymous reviewers. JMM is supported by the Prescott Postgraduate Scholarship of the University of Western Australia and the OCE PhD Scholarship of CSIRO. MS is supported by an Australian Research Council Future Fellowship (FT110100896) and Discovery Project (DP140100203).


## References


[1] D. Chelidze, "Smooth local subspace projection for nonlinear noise reduction." *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 24, no. 1, p. 013121, 2014.

[2] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. New York: Cambridge University Press, 2004.

[3] R. Cawley and G. H. Hsu, "Local-geometric-projection method for noise reduction in chaotic maps and flows." *Physical Review A*, vol. 46, no. 6, pp. 3057–3082, 1992.

[4] T. Sauer, "A noise reduction method for signals from nonlinear systems." *Physica D: Nonlinear Phenomena*, vol. 58, no. 1, pp. 193–201, 1992.

[5] J. Theiler and S. Eubank, "Don't bleach chaotic data." *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 3, no. 4, pp. 771–782, 1993.

[6] E. J. Kostelich and T. Schreiber, "Noise reduction in chaotic time-series data: A survey of common methods." *Physical Review E*, vol. 48, no. 3, pp. 1752–1763, 1993.

[7] P. Grassberger, R. Hegger, H. Kantz, C. Schaffrath, and T. Schreiber, "On noise reduction methods for chaotic data." *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 3, no. 2, pp. 127–141, 1993.

[8] X. Wan, K. Iwata, J. Riera, T. Ozaki, M. Kitamura, and R. Kawashima, "Artifact reduction for EEG/fMRI recording: Nonlinear reduction of ballistocardiogram artifacts." *Clinical Neurophysiology*, vol. 117, no. 3, pp. 668–680, 2006.

[9] N. Jevtić, P. Stine, and J. Schweitzer, "Nonlinear time series analysis of Kepler Space Telescope data: Mutually beneficial progress." *Astronomische Nachrichten*, vol. 333, no. 10, pp. 983–986, 2012.

[10] J. Moore, M. Small, and A. Karrech, "Improvements to local projective noise reduction through higher order and multiscale refinements." *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2015. Forthcoming.

[11] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980*. Verlag, 1981, pp. 366–381.

[12] J. A. Sauer, T. Yorke and M. Casdagli, "Embedology." *Journal of Statistical Physics*, vol. 65, no. 3-4, pp. 579–616, 1991.

[13] L. Noakes, "The Takens embedding theorem." *International Journal of Bifurcation and Chaos*, vol. 1, no. 04, pp. 862–872, 1991.

[14] H. Kantz, T. Schreiber, I. Hoffmann, T. Buzug, G. Pfister, L. G. Flepp, J. Simonet, R. Badii, and E. Brun, "Nonlinear noise reduction: A case study on experimental data." *Physical Review E*, vol. 48, no. 2, pp. 1529–1538, 1993.

[15] E. N. Lorenz, "Deterministic nonperiodic flow." *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.