Estimating stochastic noise using in-situ measurements from a linear wavefront slope sensor

NAZIM ALI BHARMAL[1,*] **AND ANDREW P REEVES**[1]

[1] *Centre for Advanced Instrumentation, Department of Physics, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, United Kingdom*
[*] *Corresponding author: n.a.bharmal@durham.ac.uk*

Compiled January 14, 2016

It is shown how the solenoidal component of noise from the measurements of a wavefront slope sensor can be utilized to estimate the total noise: specifically, the ensemble noise variance. It is well known that solenoidal noise is orthogonal to the reconstruction of the wavefront under conditions of low scintillation (absence of wavefront vortices). Therefore it can be retrieved even with a non-zero slope signal present. By explicitly estimating the solenoidal noise from an ensemble of slopes, it can be retrieved for any wavefront sensor configuration. Furthermore, the ensemble variance is demonstrated to be related to the total noise variance via a straight-forward relationship. This relationship is revealed via the method of the explicit estimation: it consists of a small (4), heuristic set of constants that do not depend on the underlying statistics of the incoming wavefront. These constants seem to apply to all situations—data from a laboratory experiment as well as many configurations of numerical simulation—so the method is concluded to be generic.

OCIS codes: *(110.1080) Active or Adaptive Optics; (110.4280) Noise in imaging systems; (110.3925) Metrics*

http://dx.doi.org/10.1364/ol.XX.XXXXXX

The most common type of wavefront sensor (WFS) in current use[1] is the wavefront slope sensor, typified by the Shack-Hartmann (SH). The SH uses positions of focii from lenslets (defining sub-apertures over a pupil) and from each it estimates two co-located wavefront slopes in orthogonal directions. The number of slope measurements exceeds the number of reconstructed wavefront points[2] and the excess allows for deriving terms that are unphysical for a continuous wavefront[3]. Non-continuous wavefronts, such as those containing vortices[4], are not considered here. This work extends previous developments, which have demonstrated how these unphysical signals can be interpreted as noise[5, 6], towards estimation of the ensemble noise variance when it is inhomogeneous and in the presence of non-zero and temporally-changing measurements.

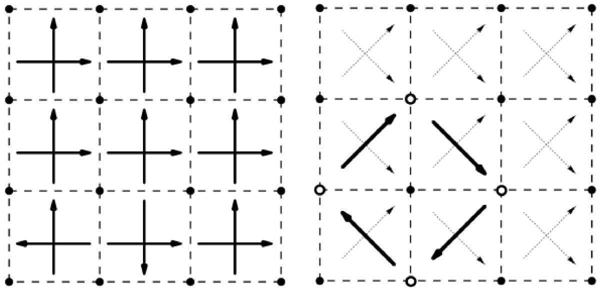

Fig. 1. (Left) A synthetic interaction matrix is defined using Fried geometry: a grid of wavefront points coincide with sub-aperture (dashed lines) corners, whereas the sub-aperture slopes lie between the grid points. (Right) A loop integration matrix is formed by rotated sub-aperture slopes that directly connect groups of four wavefront points into loops; the example loop has slopes in bold and the points as hollow.

In general measurement noise can be divided into two categories: non-solenoidal[6] and solenoidal[5]. The inversion of a linear model is used for retrieving solenoidal noise, and the model's construction permits both homogeneous and inhomogeneous noise to be obtained by appropriate regularization. It is demonstrated that solenoidal noise variance can be converted to ensemble noise variance: this is the wavefront sensor slope noise. The slope noise can be used for estimating propagated wavefront noise[7] or constraining wavefront reconstruction[2]. Alternative methods to measure slope noise[8] can rely on underlying wavefront slope statistics, whereas here no assumptions beyond wavefront continuity are required. To be useful, noise variance is required for all slopes individually and the fidelity of the estimation in this work is demonstrated with slopes from computer simulations and laboratory data.

The definition of solenoidal noise can be made from the reconstruction problem: $\mathbf{G}w = s$, where w represents a wavefront and s the wavefront slopes. The latter are noise-free slopes measured by a linear wavefront sensor. Solenoidal and Non-Solenoidal noise is written as $\eta = \eta_s + \eta_{ns}$, and from noisy slopes the wavefront estimate becomes w_{est}. The matrix $\mathbf{G}$ can be described as a gradient operator. In Fig. 1,left, the relationship is shown between the wavefront grid and the sub-aperture slopes for a synthetic $\mathbf{G}$. Then, using $\mathbf{H}$ (where $\mathbf{HG} = \mathbf{I}$),

Table 1. Parameters for the simulation of a SH-WFS. It is scale (Δ_{sa}=sub-aperture size) and wavelength (λ) independent.

Pupil shape	circular
N (sub-aperture resolution scale)	16
pixels/sub-aperture focus	4
Focal length	$4 \times \Delta_{sa}/\lambda$
Sub-ap. fractional illumination	$\geq 50\,\%$
Illumination total/sub-aperture	100 counts
RON-equivalent/sub-aperture	5
SVD singular value elimination	$10^{-8}\times$ largest value

$$w_{est} = \mathbf{H}(s + \eta_s + \eta_{ns}) = \mathbf{H}(s + \eta_{ns}). \tag{1}$$

In other words η_s is in the null space of matrix **H** which implies the estimation step,

$$\eta_s = (\mathbf{I} - \mathbf{G}.\mathbf{H})\,(s + \eta). \tag{2}$$

This is the first method for estimating η_s. However, this method relies on no other terms also lying in the null space of **H**. Generally **G** represents an interaction matrix which is not square, and **H** is either a pseudo-inverse of this matrix or requires regularization in its inversion. Here, in order to avoid unwanted terms in the null space, **H** is evaluated from a singular-value decomposition of a synthetic interaction matrix from which only machine-noise level singular values are eliminated.

An explicit method to estimate η_s is to construct sums of slope loops from rotated slopes[5, 9], as illustrated in Fig. 1,right. The sum of each slope loop is zero unless solenoidal noise is present. Therefore the following relationship can be written,

$$\mathbf{L}(s + \eta_s + \eta_{ns}) = \mathbf{L}\eta_s, \tag{3}$$

where **L** is described as a discrete loop (or curl) operator. Then inverting this matrix leads to an estimate of η_s,

$$\eta_s = \mathbf{L}^{\dagger}\mathbf{L}(s + \eta). \tag{4}$$

To form $\mathbf{L}^{\dagger}$, which is non-square, a least-squares method is used with regularization,

$$\mathbf{L}^{\dagger} = (\mathbf{L}^T\mathbf{L} + \mathbf{R})^{-1}\mathbf{L}^T. \tag{5}$$

The regularization term **R** is diagonal, as each slope's solenoidal noise is assumed zero-mean and uncorrelated. Treating R_{ii} as an *a priori* constraint of $\langle \eta^2_{s;i} \rangle$ for each slope i suggests,

$$R_{ii} = \alpha + 0.25 \times I(i)/\langle I \rangle + \begin{cases} \beta.x(i)^2 & \text{for xslopes and,} \\ \beta.y(i)^2 & \text{for yslopes.} \end{cases} \tag{6}$$

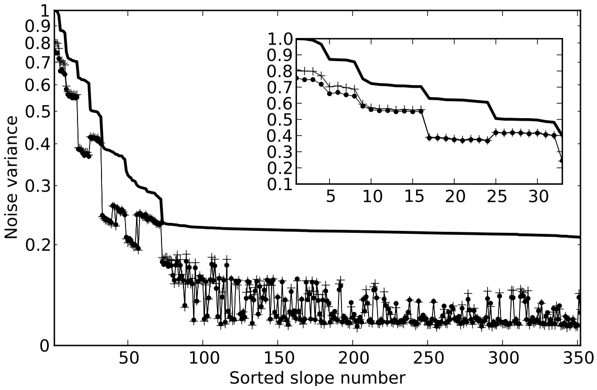

Fig. 2. The normalised variance of noise (bold line), sorted in descending amplitude, from simulated slopes. The lines with plus or point symbols represent the residual variance when the solenoidal noise estimate from gradient reconstruction or loop summation, respectively, is removed. The inset displays the 33 slopes with largest noise.

Table 2. The conversion scaling, m, between $\langle \eta_s \rangle$ and $\langle \eta \rangle$ given how many times an associated sub-aperture is referred to in **L**.

No. of references in **L** to sub-aperture	$\langle \eta_s^2 \rangle \to \langle \eta^2 \rangle$ scaling
1	14/2
2	14/4
3	14/5
4	14/6

As the problem is under-determined (fewer loops than sub-aperture slopes), α constrains the solution. The normalised intensities, $I(i)/\langle I \rangle$, are per sub-aperture associated with the i-th slope, which allows for photon noise. (If slopes are organised as XY-pairs, then this is sub-aperture $i \div 2$.) The final term with normalised pupil coordinates (x or y equals one at the edge of the pupil) allows for perspective elongation effects e.g. observing a laser guide star from off-axis sub-apertures. For the following results, $\alpha = 10^{-3}$ and $\beta = 0$ which means that the regularization is arbitrary, that intensity effects are modulated by 1/4th because of the CoG spot position algorithm employed, and that elongation effects are not relevant.

To compare the estimation of η_s using the gradient reconstruction method (via **H**) and loop summation (via **L**), noisy slope data from a $N \times N$ SH-WFS simulation is utilized. The simulation parameters are shown in Table 1. The resulting comparison is shown in Fig. 2. This plot shows $\langle (\eta - \eta_{s;H})^2 \rangle$ and $\langle (\eta - \eta_{s;L})^2 \rangle$ (the variance of noise minus solenoidal noise via, respectively, **H** and **L**) in descending order of the noise variance. As required, the variance of the residual is less than that of the variance of the noise which implies a partial estimation of noise variance is possible.

Alone, the utility of η_s is limited. First, this term is implicitly rejected during wavefront reconstruction. Second, for analyses of slopes it represents only part of η (which remains unknown). Now we consider $\langle \eta_s^2 \rangle$ in more detail. As η_{ns} and η_s are both random variables originating from the same source, a hypothesis is that the covariance of these term's ensemble variance is non-zero. In other words, noise variance is correlated for the same sub-aperture between the solenoidal and non-solenoidal terms.

A normalised covariance of 0.40 is found from the simulation data. This suggests the following relationship $\langle \eta^2 \rangle = \langle \eta_s^2 \rangle + 2\langle \eta_s \eta_{ns} \rangle + \langle \eta_{ns}^2 \rangle$=$(1+\gamma)\langle \eta_s^2 \rangle$ and, using the simulation data, γ is found to take one of four values. Our conclusion is that ensemble noise variance can be estimated as $\langle \eta_i^2 \rangle = M(i,m) \times \langle \eta^2_{s;i} \rangle$.

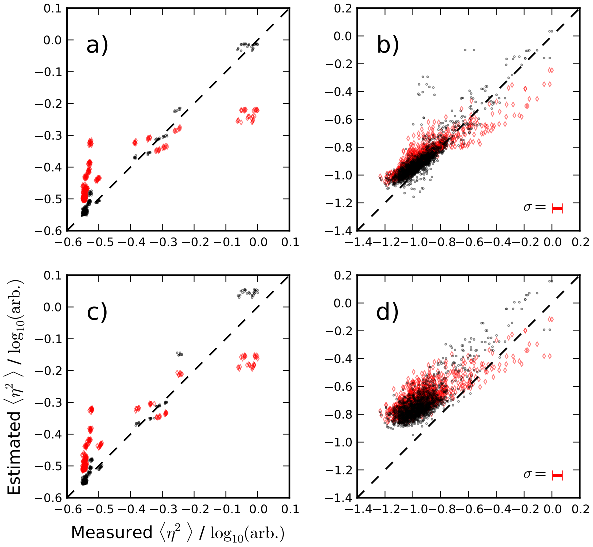

Fig. 3. The estimated vs. determined $\langle\eta^2\rangle$ for WFS slopes that are (LEFT) simulated or (RIGHT) data from DRAGON, and when there is (TOP) no slope signal, $s = 0$, or (BOTTOM) a non-zero signal, $s \neq 0$. The estimate of $\langle\eta^2\rangle$ with open diamonds (red) uses **H** while the points (black) uses **L** and regularization. For the data, the estimated variance uncertainty is shown in the bottom right.

The function M selects the value from m, whose values are shown in Table 2, given slope number i by counting how many times the associated sub-aperture is referred to in **L**. This is between zero to four, inclusive. (If zero, then no retrieval of $\langle\eta^2\rangle$ is possible for either slope in the associated sub-aperture.) The scaling introduced here differs from previous work[5] where alternative wavefront sensor geometries were considered without scaling.

The complete ensemble noise estimation method can therefore be written as,

$$\langle\eta^2\rangle = (\mathbf{M}m) \times \langle\eta_s^2\rangle = (\mathbf{M}m) \times \left\langle \left(\mathbf{L}^\dagger\mathbf{L}(s+\eta)\right)^2 \right\rangle, \qquad \textbf{(7)}$$

where $\mathbf{M}m$ is a vector of length the number of slopes and $s+\eta$ is noisy slope data. In Fig. 3 a comparison of the ensemble noise estimation methods–using **H** or **L**, and then the scaling–is made with pre-determined noise. A corresponding summary of the accuracy is shown in Table 3.

The benefit of using regularization for $\mathbf{L}^\dagger$ is evident: for the simulation results, the dispersion for low noise data (the bulk of the estimate) is much reduced and even the high noise data benefits, and when underlying slope signal is either zero or non-zero ($s = 0$ or $s \neq 0$). For the experimental data, from the 30 × 30 on-axis WFS of the DRAGON test-bench[10], first the signal-free results are discussed. The dispersion (uncertainty) is reduced and becomes consistent with the expected uncertainty. There is a (small) positive bias in the estimate for low noise which is understood as a lack of homogeneity in the underlying WFS implementation.

When the measurements with time-varying slopes ($s \neq 0$) are analysed, it is clear that there appears to be a significant bias in estimating $\langle\eta^2\rangle$. With regularization this bias remains a constant multiple of the pre-determined noise. It is necessary to assume that the pre-determined noise is identical for both sets of measurements i.e. $\langle\eta^2\rangle(s \neq 0) \equiv \langle\eta^2\rangle(s = 0)$. An interpretation is that when there are time- & space-varying slope signals, the SH spots intensities are negatively affected and this introduces further noise. As partial corroboration, an increase in sub-aperture intensity standard deviation by up to ×10 was found in this case. Then the first order estimate of increase in noise becomes $\sim \times 1.6$, which is consistent with the observed bias.

Table 3. Comparison of the noise variance using four different data sets and the two estimation methods, summarising the plots in Fig. 3.

Input	Estimated:measured $\langle\eta^2\rangle$	
	Explicit	Loop sums
a) simulation, $s = 0$	1.1 ± 0.2	1.02 ± 0.03
b) data, $s = 0$	1.1 ± 0.3	1.1 ± 0.2
c) simulation, $s \neq 0$	1.1 ± 0.2	1.02 ± 0.07
d) data, $s \neq 0$	1.8 ± 0.4	1.7 ± 0.3

For a precision of 1,5, or 20% in $\langle\eta^2\rangle$ if treated as a sample of a stationary population, there are required to be 20000, 800, or 51 samples of η_s. The measurements consist of 2500 sequential data and from this we can predict the uncertainty of the measured $\langle\eta^2\rangle$ as being 3%, which is consistent with the uncertainty bars in the figure.

We note here that the values of m are derived from Eqn. 7 by utilising using random, uncorrelated values as a substitute for $s+\eta$. Then m was used with both simulated and measured data, and no significant discrepancy was found in the estimation process. Therefore m is assumed a set of universal constants, although it is not known why they take their established values.

The previous analysis operated under the assumption that $\langle\eta^2\rangle$ was homogeneous and so the noise variance originating from either slope in each sub-aperture was identical (excepting total illumination variation). Now, the regularization will account for inhomogeneous noise, for example when SH spots are elongated.

An example of elongated SH spots is shown in the top left of Fig. 4. In the context of Eqn. 6, the term β is now set to 0.1. To separate the effect of partial illumination of sub-apertures (which causes further elongation), two categories are used: either partially or fully illuminated and shown in the top right of Fig. 4 as either dark or pale shading. The comparison of estimated $\langle\eta^2\rangle$ with actual variance, in the bottom plots of the figure, also follows this division; left/right for the noise originating from fully/partially illuminated sub-apertures.

In Table 4 is shown a summary of the noise comparison, and it can be seen that the most precise estimate for fully illuminated sub-apertures arises from using loop integration with elongation regularization. Using implicit estimation of solenoidal noise results in large scatter, and poor accuracy for partially illuminated sub-apertures.

The loop summation methods with intensity regularization result in much greater accuracy. However, for fully illuminated sub-apertures when elongation regularization is used, a small bias in estimation can be seen. Without this regularization there is instead an error which is correlated with the noise. This correlated error masks any fixed bias and returns an apparently

Table 4. Summary as normalised residuals of the data presented in Fig. 4. The situation describes the two illumination states considered (see the text for details) and the algorithm used, and then whether elongation regularization was applied.

Situation	Regularization?	Normalized residual / %
Full H		17±7
-"- L	N	2.4±4.7
-"- L	Y	3.3±2.8
Partial H		−1±34
-"- L	N	3.3±13
-"- L	Y	−9±15

more accurate estimate. A similar effect can be seen in Fig. 3 when no regularization is used in all sub-plots.

When partially illuminated sub-apertures are considered, their noise is not well estimated. This is not surprising since no appropriate consideration of these sub-aperture's elongation is made. However, consideration of the per sub-aperture illumination (intensity) leads to a substantial improvement in dispersion and reduction in correlated error. Returning to the bias in noise estimates when using elongation regularization, its origin is probably due to including data from partially illuminated sub-apertures whose noise is incorrectly constrained in the construction of $\mathbf{L}^{\dagger}$ in Eqn. 5.

The work presented in this paper is an extension for the method of estimating noise under conditions of non-zero space–time-variable wavefront sensor signals. The principal points deriving from our work are:

1. The existence of solenoidal and non-solenoidal noise is noted and their statistical relationship is derived here in terms of their covariance, and
2. To obtain the scaling from solenoidal noise variance to total noise variance requires the construction of **L** and this scaling is new, and
3. Estimation of $\langle\eta^2\rangle$ is demonstrated with simulated and measured slopes from SH WFSs, and
4. The method accounts for partially illuminated sub-apertures via a constraint term, and
5. The constraints are extended to inhomogeneous noise e.g. variably extended spots.

Due to the nature of the algorithm, it is expected to be usable with linear slopes from any first derivative-based WFS, such as the Pyramid or Fourier gradient filter types.

NAB acknowledges STFC funding from grants ST/L002213/1 and ST/L00075X/1, supporting the DRAGON laboratory experiment and this work. NAB and APR thank all those who have contributed to the development and integration of DRAGON. The data from DRAGON and example implementations of the algorithm, including the simulation code used herein, are available from NAB.

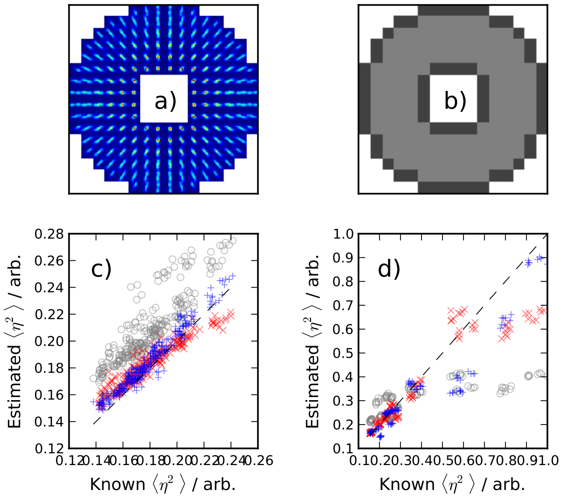

Fig. 4. The estimation of $\langle\eta^2\rangle$ versus the known variance for elongated SH spots, for three algorithms. Plot a) shows the SH spots , and b) whether spot's associated sub-apertures are fully or partially illuminated (pale or dark shading). Plot c)/d) shows the comparison of noise variance for slopes from fully/partially illuminated sub-apertures. The circle, cross, and plus symbols represent the explicit, intensity-only regularised loop summation, or intensity-and-elongation regularised loop summation algorithms respectively.

REFERENCES

1. G. Rousset, Adaptive optics in Astronomy, **1**, 91, (1999).
2. B. Ellerbroek and C. Vogel, Inv.Prob., **25:6**, 063001 (2009).
3. M. Chen, F.S. Roux, and J.C. Olivier, J.Opt.Soc.Am.A, **24**, 1994 (2007).
4. J.F. Nye and M. V. Berry, Proc.Roy.Soc. London. A, **336**. 165-90 (1974).
5. J. Herrmann, J.Opt.Soc.Am., **70**, 28 (1980).
6. G.A. Tyler, J.Opt.Soc.Am. A, **17**, 1828-1839 (2000).
7. T. Fusco, G. Rousset, D. Rabaud, É Gendron, D. Mouillet, F. Lacombe, G. Zins, P-Y. Madec, A-M. Lagrange, J. Charton, D. Rouan, N. Hubin, and N. Ageorges, J.Opt.Soc.Am. A, **6**,585, (2004).
8. F. Vidal, É. Gendron, G. Rousset, T. Morris, A. Basden, R. Myers, M. Brangier, F. Chemla, N. Dipper, D. Gratadour, D. Henry, Z. Hubert, A. Longmore, O. Martin, G. Talbot and E. Younger, A&A, **569**, A16 (2014).
9. M. Hattori and S. Komatsu, J.Mod.Opt., **50**, 1705 (2003).
10. A.P. Reeves, R.M. Myers, T.J. Morris, A.G. Basden, N.A. Bharmal, S. Rolt, D.G. Bramall, N.A. Dipper, and E.J. Younger, Proc.SPIE, **8447**, 84474Y (2012).

1. ADDITIONAL MATERIAL AS REQUESTED BY OSA

A. Complete bibliography

REFERENCES

1. G. Rousset, “Wave-front sensors.” Adaptive optics in Astronomy, **1**, 91, (1999).
2. B. Ellerbroek and C. Vogel, “Inverse Problems in Astronomical Adaptive Optics.” Inv.Prob., **25:6**, 063001 (2009).
3. M. Chen, F.S. Roux, and J.C. Olivier, “Detection of phase singularities with a Shack-Hartmann wavefront sensor.” J.Opt.Soc.Am.A, **24**, 1994 (2007).
4. J.F. Nye and M. V. Berry, “Dislocations in Wave Trains”, 1974 Proc. R. Soc. London. A, (336). 165–90. Retrieved from http://www.jstor.org/stable/78498.
5. J. Herrmann, “Least-squares wave front errors of minimum norm.” JOSA, 70(1), 28-35 (1980).
6. G.A. Tyler, “Reconstruction and assessment of the least-squares and slope discrepancy components of the phase”, J. Opt. Soc. Am. A 17, 1828-1839 (2000).
7. T. Fusco, G. Rousset, D. Rabaud, É Gendron, D. Mouillet, F. Lacombe, G. Zins, P-Y. Madec, A-M. Lagrange, J. Charton, D. Rouan, N. Hubin, and N. Ageorges, “NAOS on-line characterization of turbulence parameters and adaptive optics performance.” J.Opt.A, **6**,585, (2004).
8. F. Vidal, É. Gendron, G. Rousset, T. Morris, A. Basden, R. Myers, M. Brangier, F. Chemla, N. Dipper, D. Gratadour, D. Henry, Z. Hubert, A. Longmore, O. Martin, G. Talbot and E. Younger, “Analysis of on-sky MOAO performance of CANARY using natural guide stars.” A&A, **569**, A16 (2014).
9. M. Hattori and S. Komatsu, “An exact formulation of a filter for rotations in phase gradients and its applications to wavefront reconstruction problems.” J.Mod.Opt., 50:11, 1705-1723 (2003).
10. A.P. Reeves, R.M. Myers, T.J. Morris, A.G. Basden, N.A. Bharmal, S. Rolt, D.G. Bramall, N.A. Dipper, and E.J. Younger, “DRAGON: a wide-field multipurpose real time adaptive optics test bench.” Proc.SPIE, **8447**, 84474Y (2012).