Reaching lambda/100 resolution in static fringes interferometry using linear prediction.

Manuel Mestre [a], Didier Pasquelin [b], Peter Flug [c]

[a] Satie, École Normale Supérieure de Cachan, 61 av. Pst Wilson, F-94235 Cachan, France
[b] Optics-Concept, 12 av des Prés, Montigny, F-78059 St-Quentin, France
[c] Opticsfab, Frankfurter Straße 20, D-35625 Hüttenberg, Germany

ABSTRACT :

Both in industrial close-to-production quality control and in laboratory metrology, measuring optical components and systems with high precision and resolution (typically lambda/100 ptv) is currently achieved by phase-shifting interferometry devices. The main drawbacks of such devices compared to static fringes systems lie in a higher cost, and a greater the sensitivity to the environment, both vibration and air turbulence; the latter becomes unacceptable for large components and large cavity interferometers.

Conversely, static fringes metrology usually lacks precision and resolution. Particularly, the lateral resolution is an issue, due to the sampling theorem. This paper shows how a linear prediction of a random function (with a Bayesian approach) makes it possible to tackle a lambda/100 resolution for the estimated wavefront, being the mathematical expectation of the prediction, i.e. the most probable form with respect to the fringe data. Incidentally, the prediction increases robustness by detecting and correcting aberrant fringe data with a high reliability.

Furthermore, a Monte-Carlo simulation performed on the whole conditional probability density of the wavefront, provides a stochastic sub-fringe-spacing interpolation. As a result, confidence intervals for any parameter of interest (such as ptv, rms, ptv of slopes...) can be estimated over the whole aperture, which is novel worldwide. These algorithms have also been adapted to wavefront reconstruction from gradient data for Shack-Hartmann and for moiré devices.
Examples of implementing these algorithms to industrial software will be shown.

Keywords: Interferometry, Static fringes, Linear prediction, Monte-Carlo simulation, Confidence intervals.

1 INTRODUCTION

The metrology of wavefronts is crucial for Optical Engineering, particularly for testing the optical surfaces of components (lenses, mirrors, prisms...) Usual shop floor measurement devices, most of which were designed by pioneer XIXth century researches, are based on various physical phenomena: Interferometers, Foucault, moiré... Among them, for either technical or historical reasons, Fizeau and Twyman-Green interferometers are widely used. As a result, the ISO standard 10110-5 entitled "Optics and optical instruments. Preparation of drawings for optical elements and systems" introduces Part 5 "Surface form tolerances" by a note: "The terminology of interferometry is used for the specification of tolerances, and in particular for the units in which the tolerances are to be specified [...] Other, non-interferometric methods may be used if the results are converted to the units specified here."

Besides, a dramatic increase in precision is demanded to high-technology components, for space, astronomy, observation, telecommunications, defence... Both the polishing and the measuring processes must keep up. New concepts such as MRF polishing follow this evolution, needing an increased metrology precision. Presently, only high-quality phase-shifting interferometers seem to reach the mythic "lambda/100 ptv".

---

[a] manuel.mestre@free.fr, +33608533675
[b] didier.pasquelin@optics-concept.com, +33130640562
[c] p.flug@opticsfab.com, +4964412104090

However, the drawbacks of phase-shifting devices compared to static fringes systems lie in a greater the sensitivity to the environment, both vibration and air turbulence [1] ; the latter becomes unacceptable for large components and large cavity interferometers. Furthermore, high quality phase-shifting interferometers are more expensive, and difficult to afford for many small or medium-size firms, particularly when the investment goes along with that of a new CNC polishing machine.

The goal of this paper is to show how a careful image and signal processing can reach this "lambda/100 ptv" resolution and precision from static fringes interferograms, provided that the interferometer is in good condition , and operated in a suitable experimental environment.

As stressed by the 2005 edition of the reference "Interferogram analysis for optical testing" by Malacara et al. [2], reconstructing the phase from fringe data is a key step in the fringe analysis process. According to this review, two main types of phase reconstruction are currently used: the simplest one to implement is fitting the data by a projection on a vector space of functions (typically polynomials, such as Zernike, or Legendre). This method is known to show poor accuracy, since the least squares fit is not an interpolation: the reconstructed phase does not reach the data.

The other main class of phase detection techniques, initiated by Takeda [3], is based on a harmonic analysis phase unwrapping either for one single fringe pattern (with Fourier Transform), or for a set of phase shifted interferograms [4]. Numerous papers describe algorithms for correcting some limitations of FFT phase unwrapping techniques: sensitivity to noise, difficulty in dealing with closed fringes [5] or strongly deformed wavefronts.
The spatial (x,y) resolution is half the fringe spacing. As regards z resolution, Burnett et al.[6] consider that "As multiple images are used the resolution of the phase information is an order of magnitude higher than that of the Fourier transform technique, images being resolved typically up to 100th of a wavelength"; in other words, a single fringe pattern can yield only a lambda/10 resolution.
In short, either the FFT based techniques are simple but show low robustness, or they can be powerful at the price of complex optimisation algorithms. Using phase-shifting devices helps, with an additive cost as well. Moreover, unwrapping closed fringes due to strongly deformed wavefronts is feasible, yet not straightforward.

Our approach, though classic in Signal Processing [7], is novel in the field of wavefront interferometry software: we implement an estimation of the wavefront probability density function, conditioned by the fringe data. There are many advantages: the reconstructed wavefront is the Best Linear Predictor (i.e. the most probable interpolation of the data); being linear, the method is highly robust to noise; it provides a detection and a correction of aberrant data; it deals with surface form errors ranging from several wavelengths to a few nanometers; it provides confidence intervals for the wavefront and for any parameter of interest estimated from the wavefront; the algorithms are fairly simple.

This paper will first recall the Bayesian prediction method, showing a brief example. The role of variogram estimation will be stressed. Using the predicted p.d.f. in Monte-Carlo simulation will yield robust estimation and confidence intervals of any given parameter of interest. We will conclude on the industrial interest of static fringes interferometry.

## 2   LINEAR PREDICTION OF A RANDOM SIGNAL: CLASSIC RESULTS

1.1. General issue : performing the best estimation of a continuous process from discrete data
From a simple geometrical point of view, the peak and valley lines of an interferometric fringe pattern (Fig. 1, left) are level curves of the wavefront under test, just as the topography on a map (Fig. 1, centre). Thus the wavefront is clearly oversampled along the lines, and undersampled across them (Fig. 1, right). .

Fig. 1   The fringes peaks and valleys are level curves of the wavefront.

The main issue of the phase reconstruction is shared by most metrology processes: estimating a continuous function from discrete data [7]. In terms of probability, this issue can be expressed as follows: Calling Y the wavefront function of a 2-dimensional spatial variable X, sampled at sites (S_i). For any X, what is the information on $Y(X)$ conditioned by the data $Y(S)$, denoted as $Y(X)|Y(S)$? What estimations can be done for any parameter of interest such as a Peak-to-Valley (PV), a Root Mean Square (RMS), etc.? A connected issue is related to an optimal number and location of the data (controlled by the fringe spacing) and the subsequent precision on the estimation (Fig. 2).
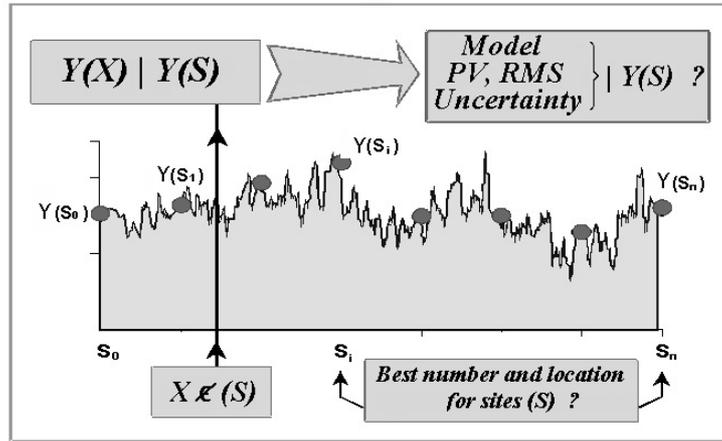
Fig. 2   Main issue of estimating continuous processes from discrete data: finding the conditional p.d.f.

The following section describes a classic answer to this issue: the linear (Bayesian) prediction of Y.

1.2. Linear prediction based on the autocorrelation function (a.c.f.)

In this paper, a material surface is considered as a particular realisation of a random function $Y(\omega, X)$, defined on a subset $D \subset \mathbb{R}^2$ for the spatial variable X, and on a field Ω of events ω.

Function Y is assumed to be the sum of a deterministic function M (usually referred to as "model") and a random function Z with zero mean: for all X in D and all ω in Ω, $\qquad Y(\omega, X) = M(X) + Z(\omega, X)$

The simplest choice for the deterministic M is the class of models linear in a parameter α: $M_\alpha(X) = f(X).\alpha$

where $f(X) = (f_1(X), \ldots, f_k(X)) \in \mathbb{R}^k$ is the general form of the model   (**Eq. 1**)

and $\alpha = (\alpha_1, \ldots, \alpha_k)^T \in \mathbb{R}^k$ is a vector parameter of the same dimension. Denote $Y(\omega, X)$ simply as $Y(X)$ and $Z(\omega, X)$ as $Z(X)$, understating the event variable ω.

In Eq.1 the choice for the dimension k of the model and the form of the component functions (f_i) is a prior based on the user's knowledge on the physical phenomena to estimate. For a wavefront, the (f_i) are typically a set of Zernike polynomials, and k is chosen as a low order. The wavefront is sampled at sites $S = (S_1, \ldots, S_n) \in D^n$, and the measured data is $Y(S) = (Y(S_1), \ldots, Y(S_n))^T \in \mathbb{R}^n$. Let $X = (X_1, \ldots, X_p)$ be a set of sites where we want to estimate Z. For our application, the data consists of sampled fringes peaks or valleys. At that stage, we assume that the reconstruction process has already extracted from the interferogram a 3-dimensional dataset.

In a Bayesian approach, the goal is to estimate the joint probability density function (p.d.f.) of the random vector *Y(X)* conditioned by the known data *Y(S)*, denoted as $(Y(X)|Y(S))$. The conditional expectation $E[Y(X)|Y(S)]$ will stand for the wavefront reconstruction. It is optimal, being the "most probable position" of $Y(X)$ with respect to the data $Y(S)$.

A usual choice is to restrain to a linear case, by assuming Z to be Gaussian, since the Gaussian subset of the random functions forms a vector space. The choice of a zero mean Z prompts to estimate the best model M_α that yields a minimum variance Z. Under these assumptions, the conditional expectation $E[Z(X)|Z(S)]$ is the Best Linear Unbiased Predictor (BLUP). This method, first developed by Matheron [8] and aimed at geostatistics applications, is known as Kriging [9].

Since it has zero mean, the statistical properties of Z are entirely defined by its $\mathbb{R}^2 \to \mathbb{R}$ covariance function

$$Cov(X_1, X_2) = E[Z(X_1).Z(X_2)]$$

Estimating M_p and Z from one single dataset requires other hypotheses : Z is supposed to be stationary, isotropic and ergodic. Then the covariance $Cov(X_1, X_2)$ depends only on $\|X_1 - X_2\|$; it can be written as $Cor(\|X_1 - X_2\|)$. The $\mathbb{R}^+ \to \mathbb{R}$ function Cor is called autocorrelation function (a.c.f.) of Z. $Cor(0)$ is the surface variance, usually denoted as σ^2.

The Bayesian prediction applied to Z yields its first two conditional moments, which entirely defines the Gaussian Z :

$$\left. \begin{array}{l} E_c[Z(X)] = C_{XS}.C_{SS}^{-1}.Z(S) \\ V_c[Z(X)] = C_{XX} - C_{XS}.C_{SS}^{-1}.C_{SX} \end{array} \right\} \qquad \text{(Eq. 2)}$$

with the following covariance matrices : $C_{SS} = \left(Cor(S_i, S_j) \right)_{(i,j) \in [1\cdots n]^2}$ $C_{XX} = \left(Cor(X_i, X_j) \right)_{(i,j) \in [1\cdots p]^2}$

$$C_{XS} = \left(Cor(X_i, S_j) \right)_{(i,j) \in [1\cdots p] \times [1\cdots n]} \qquad C_{SX} = C_{XS}^T \qquad \text{(Eq. 3)}$$

The predicted residual is $Z(S) = Y(S) - f(S).\alpha$, with $f(S) = \left(f(S_1), \ldots, f(S_n) \right)^T = \left(f_j(S_i) \right)_{(i,j) \in [1\cdots n] \times [1\cdots k]}$

Fig. 3 illustrates a 2D prediction of a function Y of a single variable X. The predicted signal (thin line), unknown by the predictive algorithm, is superimposed to the prediction.
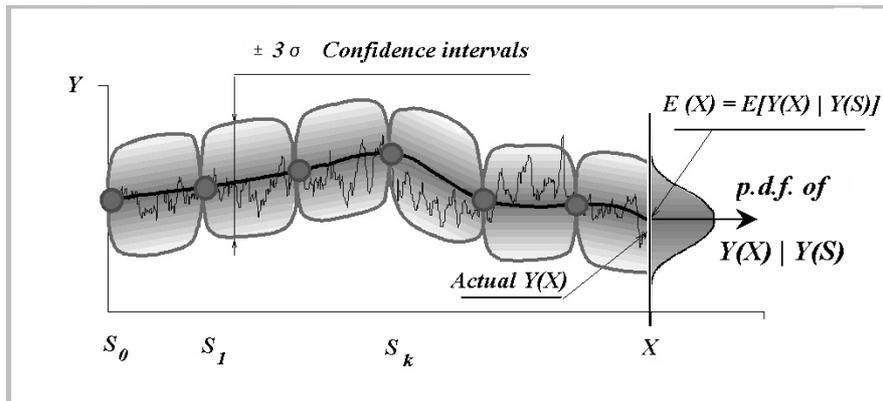
Fig. 3   Example of a 2D profile predicted from a sample.

Another 3D example is a function predicted over a square with data defined at the points of a $[0..3]^2$ grid. The conditional expectation is an actual interpolation : The expectation $E_c\left[Y\left(S_i\right)\right]$ of Y at site S_i is equal to the known $Y\left(S_i\right)$: the reconstructed phase does reach the data (Fig. 4, left). With a decreasing a.c.f., the conditional variance $V_c\left[Y\left(X\right)\right]$ is minimum for $X \in \left(S\right)$ and increases as X gets away from the closest site (Fig. 4, right).
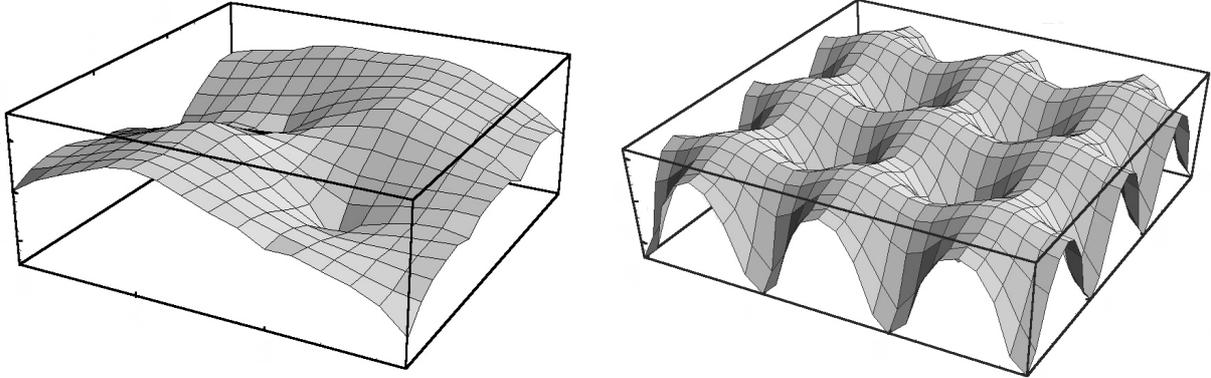

Fig. 4   Example of a surface predicted from 16 data on a square grid. Left: expectation. Right: variance

1.3. Estimating the model and the a.c.f. from the data.

The kriging [9] method estimates the parameter α that defines the model M_α in order to minimise the prediction variance, and proceeds to a *global* prediction. When the surface a.c.f. is unknown, α has to be estimated through a maximum likelihood optimisation, which involves a covariance matrix of the order of the number of data [7].[10] Besides, as discussed in Section 1.5, we perform a *local* prediction, at a single spot X, conditioned by a few surrounding data. But the predicted Y must be continuous: estimating local models would result in a patchwork of surface pieces.

Hence our approach is slightly different from kriging. Furthermore, according to some authors [11], optimising the model in not crucial: being linear, the prediction turns out to be fairly robust to non-stationarity. As a consequence, in our case, a least squares best fit of low-order Zernike over all the fringe data provides quite satisfactory predictions.
In the next section, we will focus on the prediction of $Z(X) = Y(X) - f(X).\alpha$, where f is the set of low-order Zernike we chose as a basis of models, and α is the vector of the least squares coefficients. Based on the assumed ergodicity of Z, and due to the high number of data, a naive though efficient way of estimating the a.c.f. is

computing the spatial average
$$\overset{\wedge}{Cor}_Z\left(T\right) = \left\langle Z\left(X_1\right).Z\left(X_2\right)\right\rangle_{\|X_1 - X_2\| = T} \qquad \textbf{(Eq. 4)}$$

then fitting a given function, typically
$$\overset{\wedge}{Cor}_Z\left(T\right) = a.e^{-b.T}$$

1.4. Linear prediction based on the variogram : predicting the increments of Z.

A robust approach applies linear prediction to the *increments* of Z : Define the functions of the increments of Z :
$$\begin{aligned} \delta Z : D^2 &\to & \mathbb{R} \\ (X, \tau) &\to & Z(X+\tau) - Z(X) \end{aligned}$$
. $\delta Z(X, \tau)$ is the increment of Z at X with step length τ.

Assume that δZ is Gaussian, stationary and ergodic. Now consider the "variance function" or " variogram [12] " of Z:
$$\begin{aligned} Var : D &\to & \mathbb{R}^+ \\ \tau &\to & E\left[\delta Z(X, \tau)^2\right] \end{aligned}$$
. In the stationary and isotropic case, the variogram is stationary and

reduced to a $\mathbb{R}^+ \to \mathbb{R}$ function, related to the a.c.f. by: $Var(T) = 2\left(Cor(0) - Cor(T)\right)$ for $T \in \mathbb{R}^+$. **(Eq. 5)**

Notice that $Var(0) = 0$: the variance of increments with null step length is null.

Choose an origin S_0 and denote the increments $\delta Z(S_0, \tau_1)$ and $\delta Z(S_0, \tau_2)$ as δ_1 and δ_2. The covariance of δ_1 and δ_2 is the expectation of their product $Cov(\delta_1, \delta_2) = E[\delta_1 . \delta_2]$ and can be expressed in terms of variogram :

$$Cov(\delta_1, \delta_2) = 1/2 . \left(Var\left(\|\delta_1\|\right) + Var\left(\|\delta_2\|\right) - Var\left(\|\delta_1 - \delta_2\|\right) \right).$$ **(Eq. 6)**

Let $S = (S_0, ..., S_n)$ be a set of n+1 sites of D, $\delta Z(S) = \left(\delta Z(S)_i\right)_{i \in [1 \cdots n]}$ the known increments of Z with origin S_0 and step lengths $\left(S_i - S_0\right)_{i \in [1..n]}$, $X = (X_1, ..., X_p)$ a set of p sites where we want to estimate Z. Instead of predicting $Z(X)$, we can predict the increments $\delta Z(X)$ with origin S_0 and step lengths $\left(X_j - S_0\right)_{j \in [1..p]}$, conditioned by the known data $\delta Z(S)$. The prediction equations are identical to Eq.1, just replacing Z by δZ and Cor by Cov.

The covariance matrices coefficients can be expressed with the variogram :

$$Cov\left(S_i, S_j\right) = 1/2 . \left(Var\left(\|S_i - S_0\|\right) + Var\left(\|S_j - S_0\|\right) - Var\left(\|S_i - S_j\|\right) \right)$$ and so forth.   **(Eq. 7)**

Estimating the variogram from a large dataset will compute the spatial average

$$\hat{Var}(T) = \left\langle \left(Z(X_1) - Z(X_2)\right)^2 \right\rangle_{\|X_1 - X_2\| = T}$$ then fit a given function, typically $\hat{Var}(T) = c.\left(1 - e^{-d.T}\right)$ **(Eq. 8)**

1.5. Comparing predictive methods : Z versus δZ. Robustness.

When both Z and δZ conform to the hypotheses of being stationary and isotropic, both predictive methods yield identical results. However, real life data is seldom stationary. First, the stationarity of Z implies that of δZ, whereas δZ may be stationary although Z is not. Secondly, when estimating either the a.c.f. Cor or the variogram Var from the data, sophisticated methods such as maximum likelihood turn out to be unstable. In our application to fringe analysis, we take advantage of the high number of data by expressing $Cor(T)$ or $Var(T)$ as a spatial average (as in Eq.4 or 8) , under the hypothesis of ergodicity, as mentioned above.

The main difference between both methods is due to the fact that we perform a *local* prediction, using only a few surrounding data. This choice proceeds from the order n of the data covariance matrix C_{SS} that has to be inverted. This matrix tends to be ill conditioned as its order rises, which just ruins the prediction. Remember that the whole dataset has typically one thousand points : gathering all the data in one single covariance matrix is not relevant. Now, as regards a local prediction, note that the variogram has a fixed value : by definition, $Var(0) = 0$. If non-stationarity disturbs the estimation of Var, its values for step lengths T in the neighbourhood of 0 will be less influenced that its values for large T. In contrast, $Cor(0)$ is equal to σ^2, the surface variance. Indeed, non-stationarity badly impacts the estimation of this parameter.

Predicting the increments δZ instead of Z is rather unusual in the field of applied Signal Processing such as geostatistics. The idea was first applied by Mandelbrot [13] for interpolating fractals, whose variograms follow a power law $Var(T) = T^{2H}$. Due to the self-similarity of those random processes, their variance is not defined, so that only the increments can be predicted. The fact that engineering surfaces (even polished ones) are fractal at a nanometric scale, is another incentive for choosing this approach.

As a conclusion for this section, predicting Z with the a.c.f. or δZ with the variogram are equivalent if the signal is stationary, but if not, predicting δZ is more robust.

3   DETECTING AND CORRECTING ABERRANT DATA

Before performing the phase reconstruction, the predictive algorithm can be applied to the fringe peaks and valleys data itself, in order to detect and discard aberrant data: The p.d.f. $Y\left(S_j\right)\mid Y\left(S_{i\neq j}\right)$ of each data conditioned by some of its closest neighbours is computed, providing confidence intervals (Fig. 5). If $Y\left(S_j\right)$ lies out of the confidence interval with a given confidence rate, the data is discarded. It is not necessary to replace it by its conditional expectation, since this operation would add no information. The phase prediction will just compensate for this missing data.
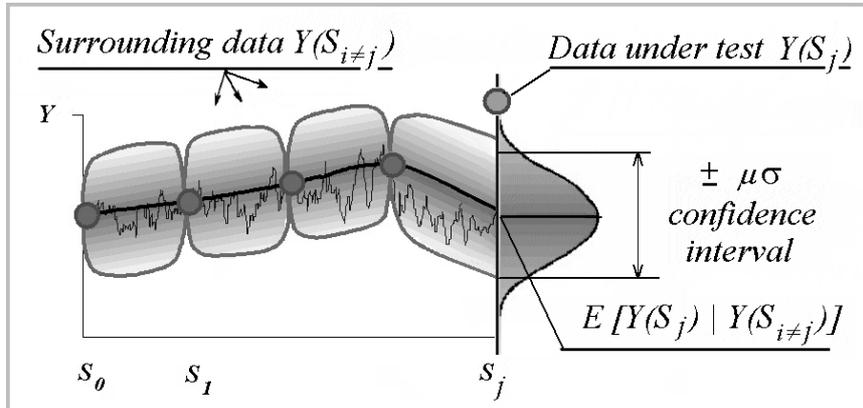

Fig. 5   Testing each data vs its confidence interval conditioned by the surrounding data.

The following example shows a spot on an interferogram resulting in a crater on a bright fringe. The peaks of the crater are detected as fringe data. If not corrected, a bad bump shows up on the phase (Fig. 6). The aberrant data algorithm described above wipes off the defect (Fig. 7). This algorithm makes the phase reconstruction highly robust to noise.
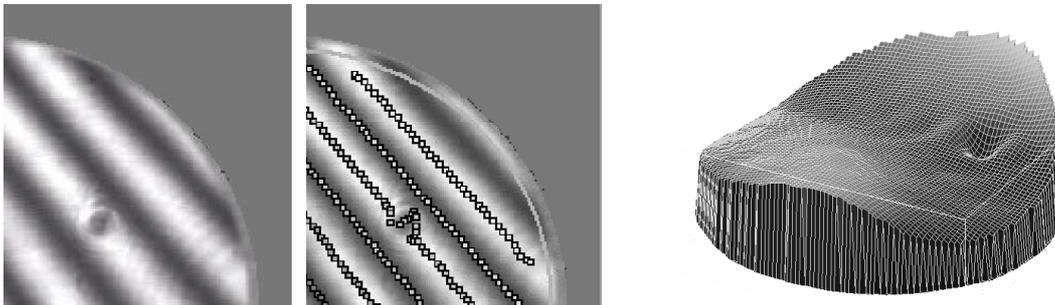
**Fig. 6**   Typical example of aberrant fringe data on an interferogram.
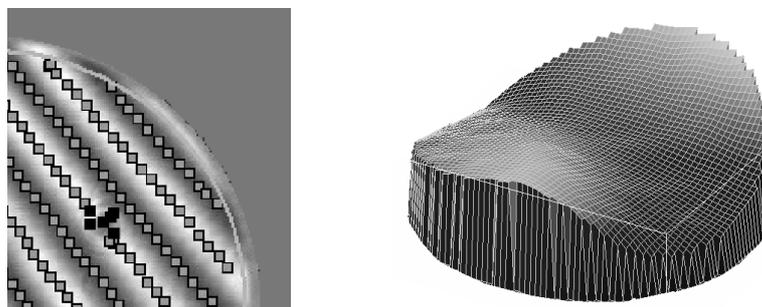
Fig. 7   Detection and rejection of the aberrant data.

4   REACHING LAMBDA/100 RESOLUTION

Indeed, the aberrant data correction algorithm helps dealing with very noisy interferograms. However, if we want to tackle high precision and resolution, we certainly need high quality data. Fig. 8 (left) is a static fringes interferogram of a part, measured on a carefully refurbished interferometer fitted with a gorgeous reference mirror.

After predicting the phase and subtracting its first Zernike components, the residuals on the 2D phase map show evidence of the polishing process, namely Magneto-Rheological Finishing with its typical revolution grooves (Fig. 8, centre). Note the scale on the 3D graph (Fig. 8, right) : the central peak is about 7 nm high, i.e. λ/100 ptv. This value does not stand for the measuring system resolution, it is the amplitude of the detected signal. The resolution is necessarily an order of magnitude smaller, actually nanometric.

When reaching this level of accuracy, it is necessary to gather the best experimental conditions of low air turbulence and solid vibrations, temperature stability... The need for tilting the wavefront is known to create aberrations, mainly astigmatism, which may be solved by averaging two wavefronts with opposite tilts.
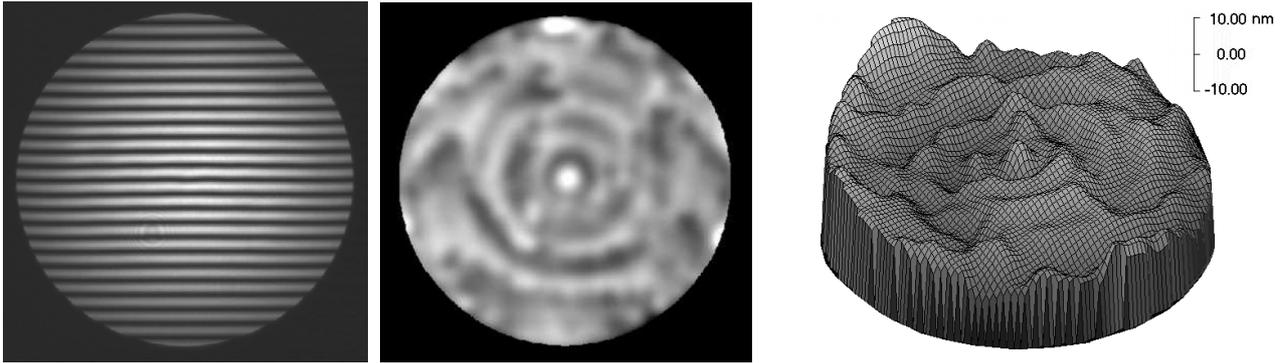

Fig. 8   Reconstructing the phase from a single static fringes interferogram:. λ/100 details are visible .

5   ESTIMATING CONFIDENCE INTERVALS FOR PARAMETERS OF INTEREST

Whatever accurate the reconstructed phase, the measuring process will miss its goal if the parameters of interest, bounded by industrial specifications, are not carefully estimated. A current way of estimating a PV or a RMS is to compute those of the reconstructed phase. However, provided we consider the phase only as estimation from an underlying random function, we need to consider as well any parameter extracted from the wavefront as a random variable. The expectation of the phase PV is *not* the PV of the phase expectation. A specification such as "The mirror PV should be less than x nm" needs to be rephrased in stochastic terms: "According to the data acquired on the mirror, what is the probability that its actual PV be greater than x nm ? " Or in other words, "What is the risk, whose estimation will help the user make a decision on the part (accept or discard)?"

When a p.d.f. is available, it is common to perform a Monte-Carlo simulation [14] with any number of random trials, consistent with this p.d.f. Namely, each trial will be one of the likely surfaces, reaching the data, and whose variogram is that of the prediction. For example, we can extract from each trial # k : $PV_k = \max_{X \in D} Y_k(X) - \min_{X \in D} Y_k(X)$ and build a histogram of these PV, for thousands of trails. The histogram converges towards the p.d.f. of the PV (Fig. 9).
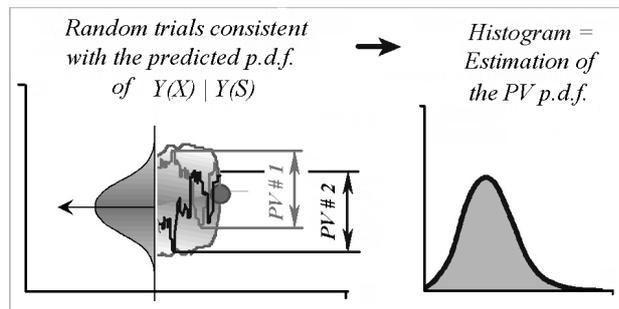
Fig. 9   Estimating the p.d.f. of the PV by Monte-Carlo simulation performed on the predicted p.d.f. of Y.

The next step uses this histogram to estimate the risk that the part PV be greater than a given specification PVmax (Fig. 10).

Estimated PV p.d.f

Confidence =
Probability (PV<PVmax)

Risk =
Probability (PV>PVmax)

PV max
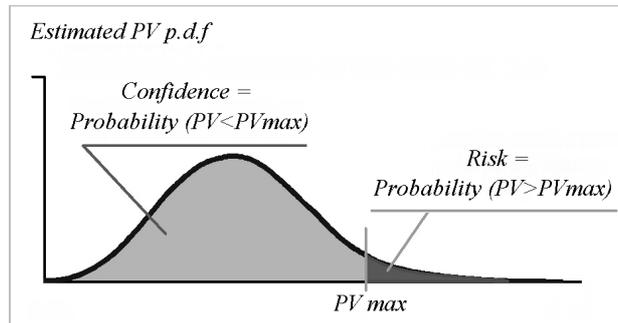
Fig. 10   Using the estimated PV p.d.f. for risk evaluation :
What is the probability for the PV to be out of tolerance ?.

Note that the z precision of these estimates relies on the fringe spacing : the smaller the inter-site distance, the smaller the conditional variance, the better the z precision. (Other approaches for extracting the wavefront RMS variance -and FTM in a forthcoming paper- from the predicted p.d.f. are described in [15].)

But the lateral (x,y) resolution of the phase reconstruction is a different issue : For example, some laser applications specify a PV of slopes averaged over an area d mm in diameter. Requiring a fringe spacing smaller than d, or even a pixel resolution, seems a straightforward idea. In fact, the prediction takes into account the expected variations of the wavefront at any step length, even tending towards zero. As a result, the lateral (x,y) resolution of these estimations is not restricted by the sampling theorem : since the statistical properties of the predicted function are held in the variogram, we do perform an actual statistical sub-fringe-spacing interpolation between the data.

Before concluding this paper, let us mention that these algorithms have also been adapted to wavefront reconstruction from gradient data for Shack-Hartmann, moiré devices, or shearing interferometers. A space and time linear prediction has been applied to optimising dynamic Adaptive Optics fitted with a Shack-Hartmann sensor [16].

6   CONCLUSION

Static fringes interferometry is a traditional and simple technique for measuring wavefront deformations. Industrial shops are fitted with numerous interferometers, most of which having beautiful reference mirrors. Urged by an ever increasing need in optical parts precision, the metrology of wavefronts has turned to phase-shifting devices, believed be the only solution to grant the mythic "lambda/100 resolution".

Yet, a painstaking image and signal processing makes it possible to extract a maximum information from a single interferogram. The linear (Bayesian) prediction of the wavefront provides an optimal reconstruction: the expectation is the "most probable surface" with respect to the fringe data, with a "vertical" z resolution close to that of a phase-shifting device. Applying a prediction to the data itself increases robustness by detecting and correcting aberrant fringe data with a high reliability. Furthermore, by means of Monte-Carlo simulations, the whole predicted p.d.f. yields confidence intervals of any parameter of interest, which is novel in this field. As regards estimating those parameters, the stochastic interpolation between the fringe data provides a precision close to that of phase-shifting, coupled with an effective "lateral" resolution (x,y) smaller than the fringe spacing.

As a conclusion, the method described in this paper can help upgrading interferometers at a significantly lower cost than buying a new system. Besides, the metrology of large components for which air turbulence makes phase shifting unstable, can also be enhanced by software based on linear prediction algorithms [17].

<div align="center">***</div>

REFERENCES

1   J. Wyant, "Vibration Insensitive Interferometric Optical Testing" http://www.optics.arizona.edu/jcwyant

2. D. Malacara et al., "Interferogram analysis for optical testing", CRC Press, London, 2005.

3   M. Takeda et al., "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry" J. Opt. Soc. Am. 72, p. 156-160, 1982

4   T.Judge et al., "A Review of Phase Unwrapping Techniques in Fringe Analysis" Optics and Lasers in Engineering, 21-4, p.199-239, 1994.

5   G. Zongtao, A. Takeda et al., "Coordinate transform method for closed fringe analysis", Proc. SPIE Int. Soc. Opt. Eng., 4101, pp. 38-46, Bellingham, 1981.

6   M. Burnett, "The Automatic Analysis of Interferometric Data", Software documentation, University of Warwick, 1996, http://www.eng.warwick.ac.uk/OEL/courses/undergrad/lec9/FRAN%20Instructions.pdf

7   E. Walter et al., "Identification of Parametric Models", Springer-Verlag, London, 1997.

8   G. Matheron, "Principles of geostatistics", Econ. Geol., 58, p. 1246-1266, 1963.

9   J-P Chiles et al., "Modeling spatial uncertainty", Geostatistics, Wiley series in probability and statistics, 1999.

10 C. R. Dietrichand and M. Osborn, "Estimation of covariance parameters in kriging via restricted maximum likelihood", Mathematical Geology, 23-1, 1991.

11 J.Sacks et al.. "Design and Analysis of Computer Experiments". Statistical Science, p. 409-435. 1989.

12 N. Cressie, "Variogram", Encyclopedia of Statistical Sciences. 9, Wiley, New York, p. 489-491, 1988.

13 B. Mandelbrot and Van Ness 1968 "Fractional Brownian motion, fractional noises and applications". SIAM Review, 10-4, 1968.

14 C.P. Robert and G. Casella. "Monte Carlo Statistical Methods", New York, Springer-Verlag, 2004.

15 M. Mestre & J.P. Rozelot, "Least squares best fit of a deterministic model to experimental data using linear prediction. Estimating confidence intervals of the R.M.S. with prediction.Application to dimensional metrology of mirrors". Journal of Optics A : Pure and Applied Optics. 4, p. 1-10, 2002.

16 J-P. Gaffard and M. Mestre, "Enhancing images in Adaptive Optics through recent developments in predictive methods", ANRT conference, Paris, 1999.

17 ClaraLuna interferometry software, www.optics-concept.com

✳✳✳