ilmedia

Technische Universität ILMENAU

Bridges, Angus; Yacoot, Andrew; Kissinger, Thomas; Tatam, Ralph P.

Multiple intensity reference interferometry for the correction of sub-fringe displacement non-linearities

Multiple intensity reference interferometry for the correction of sub-fringe displacement non-linearities

Angus Bridges[1,2] , **Andrew Yacoot**[1,2] , **Thomas Kissinger**[2] and **Ralph P Tatam**[2,*]

[1] National Physical Laboratory, Teddington, Middlesex TW11 0LW, United Kingdom
[2] Centre for Engineering Photonics, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom

E-mail: r.p.tatam@cranfield.ac.uk

CrossMark

Abstract

Displacement measuring interferometers, commonly employed for traceable measurements at the nanoscale, suffer from non-linearities in the measured displacement that limit the achievable measurement uncertainty for microscopic displacements. Two closely related novel non-linearity correction methodologies are presented here that allow for the correction of non-linearities in cases where the displacement covers much less than a full optical fringe. Both corrections have been shown, under ideal conditions, to be capable of reducing all residual non-linearity harmonics to below the 10 pm level.

Keywords: interferometry, non-linearity, dimensional metrology

(Some figures may appear in colour only in the online journal)

1. Introduction

Optical interferometry provides a direct route for traceability to the SI metre on length scales for which direct time of flight measurement is impractical [1]. For measurements made at the nanoscale, optical interferometry therefore provides the primary route to traceability. Whilst a number of other technologies, including capacitive sensors [2], grating based encoders [3] and x-ray interferometry [4] are able to make high accuracy displacement measurements at the nanoscale, all ultimately rely upon optical interferometry for traceability [1]. Improving the performance of optical interferometry therefore serves two purposes. Firstly, optical interferometry is used directly in a wide range of research and industrial applications, for example metrological atomic force microscopy [5]. Secondly, due to the straightforward path to traceability offered by optical interferometry, improved optical interferometers are desirable for the calibration or characterisation of other nanoscale measurement devices.

Non-linearities may dominate the measurement uncertainty of homodyne displacement measuring interferometers under conditions for which refractive index uncertainties do not dominate, for example, when measuring displacements covering ranges of the order of hundreds of micrometres or less, in some cases limiting instrument performance [6]. In particular, in cases where the displacement covers less than a full optical fringe, as may be the case, for example, for the out of plane measurement axis of a scanning probe microscope, or when evaluating the out of plane motion of a nanopositioning stage [7], ellipse fitting Heydemann style corrections [8] may result in large residual non-linearities, as ellipse fitting of noisy data is unreliable when only partial ellipses are available [9]. Such non-linearities, errors in the measured displacement that are periodic with some harmonic of the illuminating wavelength, are introduced by errors in the DC offsets, gain ratio and quadrature phase of the interferometer output signals [8], or by unwanted multiple reflections within the interferometer optics

[10, 11]. A wide variety of methods have been described in the literature to correct for non-linearities [8, 9, 12–15], including methods suitable for real time implementation [16–18], and methods designed for sub-fringe displacements [9, 15]. Existing sub-fringe correction methods proposed in the literature are however limited to residual non-linearities of approximately 1 nm where displacements of less than one half of an optical fringe are to be corrected [9, 15], far in excess of the tens of picometre non-linearities that can be achieved by optical interferometry in the multi-fringe case [19], although the performance limits may be expected to be dependent on the exact interferometer system to which the correction is applied.

A non-linearity correction methodology is presented in this work that is capable of correction of non-linearities in homo-dyne optical interferometers for sub-fringe displacements, out-performing the Heydemann correction for displacements of less than approximately 120 nm. Two closely related techniques are proposed, both based upon measurement of the optical power returned from each arm of the interferometer, following a related methodology to that of [15] with the goal of reducing residual non-linearities to tens of picometres. First, the theory behind the methodology will be presented, along with the theory behind the calibration procedure developed in this work. The methodologies used to investigate the performance of the correction will then be described, followed by experimental results validating the correction. The limitations and benefits of the correction will then be considered and compared with the Heydemann [8] and beam shuttering [15] techniques.

2. Theory

Quadrature fringe counting in homodyne displacement measuring interferometry relies upon the interferometer producing two signals, separated by a 90° phase difference in the ideal case [12]. A phase-quadrature coating [20] based interferometer is depicted in figure 1 with quadrature outputs I_{sin} and I_{cos}, and intensity reference outputs I_{mes} and I_{ref}. The phase-quadrature coating consists of absorbing thin films of the order of ten nanometres in thickness, and due to the absorbing nature of the films a phase shift is produced between the reflected and transmitted beams. In the configuration employed in this work, the coating produces a phase difference of approximately 90° between the I_{sin} and I_{cos} outputs of the interferometer. Two closely related non-linearity correction methodologies are presented in this section, an overview of which is given in figure 2.

Assuming an optical intensity after the polariser of I_0, the optical intensity signals at the four outputs will be given by

$$I_{cos} = R_Q T_Q T_P \left(T_{mes} + T_{ref} + 2\sqrt{T_{mes}T_{ref}}\cos(\varphi) \right) I_0, \quad (1)$$

$$
\begin{aligned}
I_{sin} = T_P \big(&T_{ref}R_Q^2 + T_{mes}T_Q^2 \\
&+ 2R_Q T_Q \sqrt{T_{mes}T_{ref}}\sin(\varphi + \delta) \big) I_0,
\end{aligned}
\quad (2)
$$
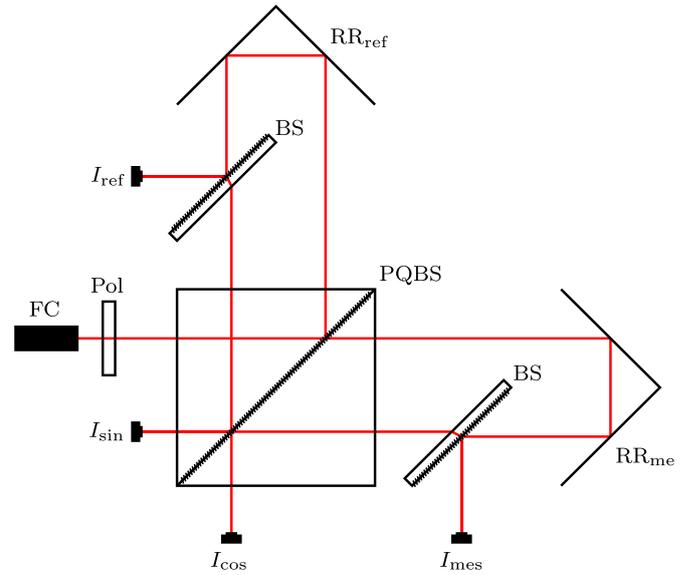
Figure 1. Schematic of the phase-quadrature coating based homodyne interferometer. FC, fibre collimator; PQBS, phase-quadrature beam splitter; Pol, polariser; BS, 50:50 beam splitter; RR$_{mes}$ and RR$_{ref}$, measurement and reference retroreflectors; I_{mes}, measurement beam reference intensity; I_{ref}, reference beam reference intensity; I_{sin}, sine quadrature output; I_{cos}, cosine quadrature output.

$$I_{ref} = T_{ref}R_Q R_P I_0, \quad (3)$$

$$I_{mes} = T_{mes}T_Q R_P I_0, \quad (4)$$

where R_Q and T_Q are the power reflectivity and transmissivity coefficients of the phase-quadrature beam splitter (PQBS) coating, R_P and T_P are the power reflectivity and transmissivity coefficients of the non-polarising beam splitter (BS) coatings, T_{ref} and T_{mes} are the power transmissivity coefficients of the reference and measurement arm beam paths, φ is the interferometric phase, determined by the illuminating wavelength, refractive index, and relative path lengths of the reference and measurement arms of the interferometer, and δ is the quadrature phase error. The parameters T_{ref} and T_{mes} are required to account for losses in the arms of the interferometer, which may arise for example due to the non-ideal reflectivity of metallic mirror or retroreflector coatings.

Equations (1) and (2) assume perfect interference between the beams returned from the reference and measurement arms, however in reality wavefront or polarisation aberrations may result in a reduction in fringe contrast at the sine and cosine outputs, without affecting the intensity measured at the reference and measurement intensity outputs. Such effects will alter the amplitude of the measured quadrature signals, but will not affect the offsets of the quadrature signals, a fact that will be exploited later in the correction process.

The optical intensity signals are not measured directly, and are instead commonly detected and amplified to give a voltage
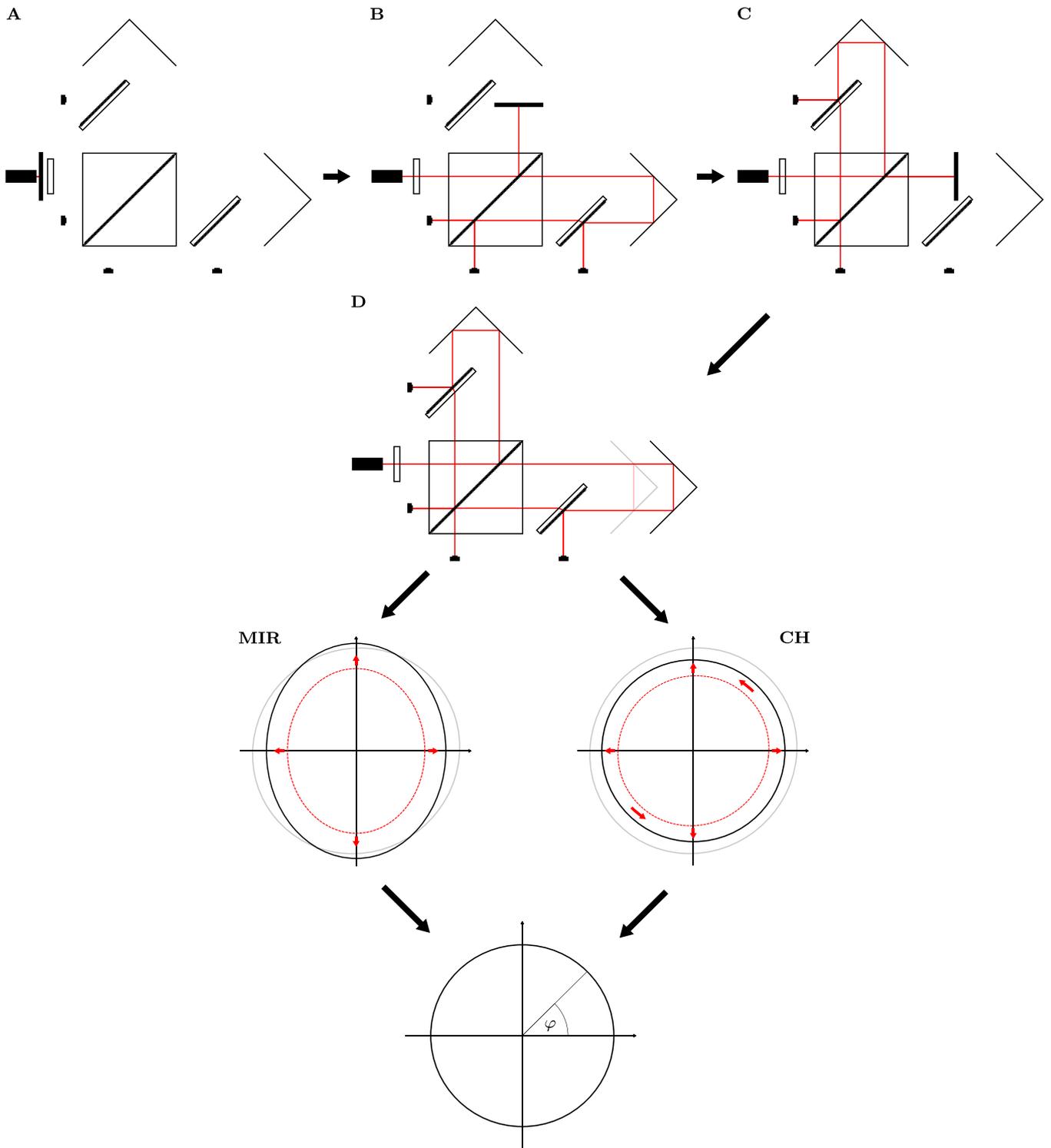
Figure 2. Flowchart demonstrating the multiple intensity reference (MIR) and constrained Heydemann (CH) non-linearity correction methodologies. First, the input beams are blocked and dark voltages measured (A). The reference (B) and measurement (C) arms are then blocked in turn, and calibration measurements taken. Following calibration, a displacement measurement is taken (D), and equations (15) and (16) are applied to centre the quadrature signals, and to attempt to correct the signal amplitudes. For the MIR approach, the sum and difference of the centred signals are taken, and an ellipse fit is made, centred on the origin, and with the ellipse axes aligned to the data axes (MIR). For the CH approach, the centred signals are fitted with an ellipse again constrained to be centred on the origin, but free to rotate about the origin (CH). Finally for both approaches the fit coefficients are used to correct the quadrature signals, and the phase is recovered from the now circular Lissajous figure.

signal with a generally unknown gain and offset for each signal, resulting in

$$V_{\cos} = G_{\cos}I_{\cos} + D_{\cos}, \tag{5}$$

$$V_{\sin} = G_{\sin}I_{\sin} + D_{\sin}, \tag{6}$$

$$V_{\text{ref}} = G_{\text{ref}}I_{\text{ref}} + D_{\text{ref}}, \tag{7}$$

$$V_{\text{mes}} = G_{\text{mes}}I_{\text{mes}} + D_{\text{mes}}, \tag{8}$$

where the coefficients G represent (positive) gain, and the coefficients D represent the voltage offset that is present in the absence of laser illumination, introduced by both background illumination and electronic effects. For simplicity of notation in the following equations, offset free voltage signals will be introduced as

$$V'_{\cos} = V_{\cos} - D_{\cos}, \tag{9}$$

$$V'_{\sin} = V_{\cos} - D_{\sin}, \tag{10}$$

$$V'_{\text{ref}} = V_{\text{ref}} - D_{\text{ref}}, \tag{11}$$

$$V'_{\text{mes}} = V_{\text{mes}} - D_{\text{mes}}. \tag{12}$$

This system of equations may be solved for the zero centred and equal amplitude quadrature signals $S_{\cos}$ and $S_{\sin}$, independently of illumination intensity and losses in the interferometer arms, as

$$
\begin{aligned}
S_{\cos} &= \cos(\varphi), \\
&= -\frac{\begin{array}{c}G_{\text{ref}}G_{\cos}T_{\text{P}}R_{\text{Q}}V'_{\text{mes}} + G_{\text{mes}}G_{\cos}T_{\text{P}}T_{\text{Q}}V'_{\text{ref}} \\ -G_{\text{mes}}G_{\text{ref}}R_{\text{P}}V'_{\cos}\end{array}}{2G_{\cos}T_{\text{p}}\sqrt{G_{\text{mes}}G_{\text{ref}}R_{\text{Q}}T_{\text{Q}}V'_{\text{mes}}V'_{\text{ref}}}},
\end{aligned} \tag{13}
$$

$$
\begin{aligned}
S_{\sin} &= \sin(\varphi + \delta), \\
&= \frac{\begin{array}{c}G_{\text{ref}}G_{\sin}T_{\text{P}}T_{\text{Q}}V'_{\text{mes}} + G_{\text{mes}}G_{\sin}T_{\text{P}}R_{\text{Q}}V'_{\text{ref}} \\ -G_{\text{mes}}G_{\text{ref}}R_{\text{P}}V'_{\sin}\end{array}}{2G_{\sin}T_{\text{p}}\sqrt{G_{\text{mes}}G_{\text{ref}}R_{\text{Q}}T_{\text{Q}}V'_{\text{mes}}V'_{\text{ref}}}},
\end{aligned} \tag{14}
$$

however this form requires the optical and electronic properties of the interferometer system to be measured directly.

Considering equations (13) and (14), some simplifications can be made. Homodyne interferometers recover the interferometric phase using the four quadrant arctangent function [12], which is insensitive to scaling factors common to both quadrature signals. As a consequence, the signals can be arbitrary scaled by a common factor, allowing equations (13) and (14) to be written in terms of gain ratios, rather than absolute gains, and reflectivity to transmission ratios, rather than

absolute reflectivity and transmission coefficients. The cosine signal may also be inverted without affecting the relative phase measurement, allowing both signals to be rewritten with common signs for each component. Combining these changes results in

$$S'_{\cos} = \Gamma_{\text{Q}}V'_{\text{mes}} + G_{\text{m,r}}V'_{\text{ref}} - G_{\text{m,c}}\Gamma'_{\text{P}}V'_{\cos}, \tag{15}$$

$$S'_{\sin} = V'_{\text{mes}} + G_{\text{m,r}}\Gamma_{\text{Q}}V'_{\text{ref}} - G_{\text{m,s}}\Gamma'_{\text{P}}V'_{\sin}, \tag{16}$$

where

$$G_{\text{m,r}} = \frac{G_{\text{mes}}}{G_{\text{ref}}}, \tag{17}$$

$$G_{\text{m,c}} = \frac{G_{\text{mes}}}{G_{\cos}}, \tag{18}$$

$$G_{\text{m,s}} = \frac{G_{\text{mes}}}{G_{\sin}}, \tag{19}$$

$$\Gamma_{\text{Q}} = \frac{R_{\text{Q}}}{T_{\text{Q}}}, \tag{20}$$

$$\Gamma'_{\text{P}} = \frac{R_{\text{P}}}{T_{\text{P}}T_{\text{Q}}}. \tag{21}$$

2.1. Calibration procedure

In order to recover $S'_{\cos}$ and $S'_{\sin}$ the gain and reflectivity to transmission ratios must be measured. Whilst direct measurement would be possible, it is desirable to perform this calculation *in-situ*, as a calibration step. This may be achieved through beam blocking techniques, in a method related to that proposed in [15]. First, the dark voltage coefficients required to calculate the offset free voltage signals V' from the measured voltage signals V may be measured directly by blocking the input beam to the interferometer. By blocking first the reference, and then the measurement beam, separately, eight voltages may be recorded, some of which can be used to calibrate the required ratios. Blocking the reference beam results in voltages

$$V^{(\text{R})}_{\cos} = G_{\cos}R_{\text{Q}}T_{\text{Q}}T_{\text{P}}T_{\text{mes}}I_0, \tag{22}$$

$$V^{(\text{R})}_{\sin} = G_{\sin}T^2_{\text{Q}}T_{\text{P}}T_{\text{mes}}I_0, \tag{23}$$

$$V^{(\text{R})}_{\text{ref}} = 0, \tag{24}$$

$$V^{(\text{R})}_{\text{mes}} = G_{\text{mes}}T_{\text{Q}}R_{\text{P}}T_{\text{mes}}I_0. \tag{25}$$

Similarly, blocking the measurement beam results in voltages

$$V^{(\text{M})}_{\cos} = G_{\cos}R_{\text{Q}}T_{\text{Q}}T_{\text{P}}T_{\text{ref}}I_0, \tag{26}$$

$$V_{\text{sin}}^{(M)} = G_{\text{sin}} R_{\text{Q}}^2 T_{\text{P}} T_{\text{ref}} I_0, \tag{27}$$

$$V_{\text{ref}}^{(M)} = G_{\text{ref}} R_{\text{Q}} R_{\text{P}} T_{\text{ref}} I_0, \tag{28}$$

$$V_{\text{mes}}^{(M)} = 0. \tag{29}$$

From equations (22) to (29) the required ratios and products of ratios may then be determined as

$$G_{\text{m,r}} = \frac{V_{\text{mes}}^{(R)}}{V_{\text{ref}}^{(M)}} \sqrt{\frac{V_{\text{sin}}^{(M)} V_{\text{cos}}^{(M)}}{V_{\text{sin}}^{(R)} V_{\text{cos}}^{(R)}}}, \tag{30}$$

$$\Gamma_{\text{Q}} = \sqrt{\frac{V_{\text{sin}}^{(M)} V_{\text{cos}}^{(R)}}{V_{\text{cos}}^{(M)} V_{\text{sin}}^{(R)}}}, \tag{31}$$

$$G_{\text{m,c}} \Gamma_{\text{P}}' = \frac{V_{\text{mes}}^{(R)}}{V_{\text{cos}}^{(R)}} \Gamma_{\text{Q}}, \tag{32}$$

$$G_{\text{m,s}} \Gamma_{\text{P}}' = \frac{V_{\text{mes}}^{(R)}}{V_{\text{sin}}^{(R)}}. \tag{33}$$

From equations (15), (16) and (30) to (33) equal amplitude zero centred quadrature signals can be calculated, however the quadrature phase error, δ is still present. Two approaches are possible to correct for the quadrature error, which will be outlined in the following subsections.

2.2. Multiple intensity reference technique

The first approach, termed in this work the multiple intensity reference (MIR) technique, assumes the signals have ideally equal amplitudes, and takes the sum and difference of the quadrature signals, yielding

$$S_{\text{c+s}} = S_{\text{cos}}' + S_{\text{sin}}', \tag{34}$$

$$= 2 \cos\left(\frac{1}{2}\delta\right) \cos\left(\varphi + \frac{1}{2}\delta\right), \tag{35}$$

$$S_{\text{c-s}} = S_{\text{cos}}' - S_{\text{sin}}', \tag{36}$$

$$= 2 \sin\left(\frac{1}{2}\delta\right) \sin\left(\varphi + \frac{1}{2}\delta\right), \tag{37}$$

where the quadrature error, introduced by the non-ideal performance of the phase-quadrature coating, has been eliminated at the cost of reintroducing a difference between the amplitudes of the signals. A constrained ellipse fit may then be employed to fit an ellipse centred on the origin, with the major and minor axes aligned with the data axes, in order to correct for the amplitude difference introduced by the sum and difference process. This approach is however vulnerable to

differences in the signal amplitudes before taking the sum and difference.

An ellipse fit constrained as described may be made by applying least squares fit of the form

$$1 = \cos^2\left(\varphi + \frac{1}{2}\delta\right) + \sin^2\left(\varphi + \frac{1}{2}\delta\right), \tag{38}$$

$$= \left(\frac{S_{\text{c+s}}}{2\cos\left(\frac{1}{2}\delta\right)}\right)^2 + \left(\frac{S_{\text{c-s}}}{2\sin\left(\frac{1}{2}\delta\right)}\right)^2, \tag{39}$$

$$= A S_{\text{c+s}}^2 + B S_{\text{c}-s}^2. \tag{40}$$

Equation (40) may be written in matrix form as

$$\boldsymbol{M}_\pm \left(\begin{array}{c} A \\ B \end{array} \right) = \left(\begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right), \tag{41}$$

where $\boldsymbol{M}_\pm$ is the $2 \times N$ data matrix

$$\boldsymbol{M}_\pm = \left(\begin{array}{cccc} S_{\text{c+s},0}^2 & S_{\text{c+s},1}^2 & \cdots & S_{\text{c+s},N}^2 \\ S_{\text{c-s},0}^2 & S_{\text{c-s},1}^2 & \cdots & S_{\text{c-s},N}^2 \end{array} \right) \tag{42}$$

and the right hand side is a $N \times 1$ vector of ones for a fit made to N data points. Solving this matrix equation following the ordinary least squares method results in closed form solutions for the fit coefficients A and B given by

$$A = \frac{\left(\sum S_{\text{s-c}}^2\right)\left(\sum S_{\text{s-c}}^2 S_{\text{s+c}}^2\right) - \left(\sum S_{\text{s+c}}^2\right)\left(\sum S_{\text{s-c}}^4\right)}{\left(\sum S_{\text{s-c}}^2 S_{\text{s+c}}^2\right)^2 - \left(\sum S_{\text{s+c}}^4\right)\left(\sum S_{\text{s-c}}^4\right)}, \tag{43}$$

$$B = \frac{\left(\sum S_{\text{s+c}}^2\right)\left(\sum S_{\text{s-c}}^2 S_{\text{s+c}}^2\right) - \left(\sum S_{\text{s-c}}^2\right)\left(\sum S_{\text{s+c}}^4\right)}{\left(\sum S_{\text{s-c}}^2 S_{\text{s+c}}^2\right)^2 - \left(\sum S_{\text{s+c}}^4\right)\left(\sum S_{\text{s-c}}^4\right)}, \tag{44}$$

where the index of sums, which is not shown for clarity of notation, runs over the full range of the data points to be included in the fit. Equivalently, the sums may be replaced with averages, for example to avoid fixed point overflows, or with moving averages covering a fixed displacement length where the technique is to be applied to displacements covering multiple fringes. Both fit coefficients have a common divisor, which may be neglected as the absolute radius of the ellipse does not effect the phase measurement process, and as a result the fit coefficients may be expressed as

$$A' = \left(\sum S_{\text{s+c}}^2\right)\left(\sum S_{\text{s}-c}^4\right) \tag{45}$$

$$- \left(\sum S_{\text{s}-c}^2\right)\left(\sum S_{\text{s}-c}^2 S_{\text{s+c}}^2\right), \tag{46}$$

$$B' = \left(\sum S_{\text{s}-c}^2\right)\left(\sum S_{\text{s+c}}^4\right) \tag{47}$$

$$- \left(\sum S_{\text{s+c}}^2\right)\left(\sum S_{\text{s}-c}^2 S_{\text{s+c}}^2\right). \tag{48}$$

The sum and difference signals may then be corrected as

$$C_{s+c} = \sqrt{A'} S_{s+c}, \tag{49}$$

$$C_{s-c} = \sqrt{B'} S_{s-c}. \tag{50}$$

2.3. Constrained Heydemann technique

The second possible quadrature phase correction approach, termed here the constrained Heydemann (CH) technique, deals with the unequal amplitude problem by again making a constrained ellipse fit, fitting an ellipse with variable signal amplitudes and quadrature phase errors, but constrained to be centred on the origin.

The CH approach makes a least squares fit of the form

$$1 = C S_{\cos}^2 + E S_{\sin}^2 + F S_{\cos} S_{\sin}. \tag{51}$$

The form of equation (51) may be derived following the same approach as for equation (40), solving

$$S_{\cos} = \alpha \cos(\varphi), \tag{52}$$

$$S_{\sin} = \beta \sin(\varphi + \delta), \tag{53}$$

for the sine and cosine of φ, and taking the sum of the square of sine and cosine to be equal to one. A quadrature error term, δ, has again been included to allow for the correction of non-ideal quadrature phase shifts. Coefficients α and β have been introduced to account for the non-ideal signal amplitudes.

Closed form solutions may also be found for the fit coefficients in this case, however the form is more complex and will not be described here. Instead, the fit coefficients C, E and F will be calculated following the ordinary least squares method as

$$\begin{pmatrix} C \\ E \\ F \end{pmatrix} = \left(M^{\mathrm{T}} M \right)^{-1} M^{\mathrm{T}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \tag{54}$$

where

$$M = \begin{pmatrix} S_{\cos,0}^2 & S_{\cos,1}^2 & \cdots & S_{\cos,N}^2 \\ S_{\sin,0}^2 & S_{\sin,1}^2 & \cdots & S_{\sin,N}^2 \\ S_{\cos,0} S_{\sin,0} & S_{\cos,1} S_{\sin,1} & \cdots & S_{\cos,N} S_{\sin,N} \end{pmatrix}. \tag{55}$$

Following the methodology of Heydemann [8] whilst accounting for the fact that the signals are now constrained to be centred on the origin, the amplitude ratio of the quadrature signals and the sine and tangent of the quadrature phase error may be calculated from the fit coefficients as

$$G_{xy} = \sqrt{\frac{E}{C}}, \tag{56}$$

$$\tan(\delta) = \frac{-F}{\sqrt{4CE - f^2}}, \tag{57}$$

$$\sin(\delta) = \frac{-2\sqrt{CE}}{\sqrt{4CE - f^2}}, \tag{58}$$

and the corrected quadrature signals may be recovered as

$$C_{\cos} = S_{\cos}, \tag{59}$$

$$C_{\sin} = G_{xy} \sin(\delta) S_{\sin} + \tan(\delta) S_{\cos}. \tag{60}$$

3. Experimental methods

In order to test both the MIR and CH techniques, an interferometer was constructed following the schematic representation shown in figure 1. Laser illumination was provided by a fibre coupled frequency stabilised helium neon (HeNe) laser (REO Model 32734) operating at approximately 632.8 nm. The fibre coupled laser was then collimated and polarised with a Glan-Thompson polariser (extinction ratio 100 000:1), ensuring a stable polarisation state at the input to the interferometer. A custom PQBS [20] cube formed the core of the interferometer, splitting the beam into the measurement and reference arms of the interferometer, and introducing an approximately 90° quadrature phase shift between the interferometer outputs. The quadrature phase shift introduced by the PQBS is polarisation dependent, and as sub-fringe fluctuations in the quadrature phase shift are not corrected for with the techniques described here, the high extinction ratio input polariser is required to minimise non-linearities in the interferometer. Silver coated solid glass cube corner retroreflectors were employed in both arms of the interferometer, with the arm lengths approximately balanced to within millimetres. The fixed phase change that takes place upon reflection from the absorbing silver coatings adds a fixed offset to the interferometric phase, and therefore does not affect the proposed correction methods.

The measurement arm retroreflector, RR$_{\mathrm{mes}}$, was mounted on a Queensgate Instruments NPS-X-15A nanopositioning stage, with a manufacturer quoted position noise of 50 pm RMS, allowing precise path length changes to be generated. Plate BSs were used in both the reference and measurement arms, with a 50:50 reflectivity to transmission ratio, and BSs marked as being from the same coating batch were selected as both the MIR and CH techniques assume identical BS properties in each arm of the interferometer. All optical components with planar optical surfaces were tilted slightly to minimise any multiple reflection effects, whilst maintaining parallelism between the measurement arm beams and stage displacement axis. In order to ensure consistent power collection at each of the four detectors, the output beams of the interferometer were focused onto the photodiode detectors (Hamamatsu S2386-18K/L) with 40 mm focal length lenses (not shown in figure 1). Apertures were placed before each lens set to a diameter that did not visible clip the beams in order to minimise the detection of stray light. The photodiodes were operated in photovoltaic mode to minimise dark currents, with the photocurrent amplified by custom electronics. The amplified signals were digitised with 16 bit analogue to digital converters (National

Instruments NI7833R) operating at a sampling frequency of 200 kHz, and, for the purposes of this work, were logged for post processing.

For both the MIR and CH corrections identical calibration steps are required. These were performed in three steps. First, a beam block was placed between the input polariser and PQBS, and the dark voltage signals ($D_{\cos}$, $D_{\sin}$, D_{ref} and D_{mes}) recorded. The beam block was then moved to between the reference retroreflector (RR_{ref}) and the reference arm BS, and the reference blocked voltage signals ($V_{\cos}^{(R)}$, $V_{\sin}^{(R)}$, $V_{\mathrm{ref}}^{(R)}$ and $V_{\mathrm{mes}}^{(R)}$) were recorded. Finally, the beam block was moved to between the measurement arm retroreflector and the measurement arm BS, and the measurement blocked voltage signals ($V_{\cos}^{(M)}$, $V_{\sin}^{(M)}$, $V_{\mathrm{ref}}^{(M)}$ and $V_{\mathrm{mes}}^{(M)}$) were recorded. For each calibration step 100 000 data points were recorded synchronously for all four signals. Ratios and products of ratios were then calculated according to equations (30)–(33) using the full 100 000 data points. Averages were then taken for each ratio, and the averaged ratios used to corrected the measured signals following equations (15) and (16). By taking averages only after the ratios were calculated, the calibration process was made insensitive to laser power fluctuations between each calibration step, and any bias introduced by taking ratios of averaged random variables was eliminated.

At this point, the two proposed techniques diverge. Considering first the MIR technique, sum and difference signals were calculated from equations (34) and (37). Fit coefficients were then calculated according to equations (45) and (48), and the sum and difference signals were corrected using equations (49) and (50). In the case of the CH technique, fit coefficients C, E and F were calculated from equation (54), and the correction was applied following equations (56)–(60).

Following signal correction, the interferometric phase was, for both techniques, calculated using the four quadrant arctangent function as

$$\varphi_{\mathrm{CH}} = \mathrm{atan2}\left(C_{\sin}, C_{\cos}\right) \qquad (61)$$

and

$$\varphi_{\mathrm{MIR}} = \mathrm{atan2}\left(C_{\mathrm{s+c}}, C_{\mathrm{s-c}}\right) \qquad (62)$$

for the CH and MIR corrections respectively. The corresponding displacement was then calculated for both techniques as

$$x = \frac{\lambda}{4\pi}\varphi. \qquad (63)$$

For the purposes of this work any dead path effects [21] introduced by path length imbalances at the start of the measurement were not considered, as path length imbalances were minimal, and any dead path effects present do not affect the non-linearity of the interferometer, which is the focus of this work. Additionally, the effects of refractive index fluctuations [22], were not considered, as such effects do not affect the ability of the interferometer system described here to resolve non-linearities, instead having only a negligible scaling effect on the observed non-linearity.

In order to test the correction techniques, the stage was stepped through the full 15 μm closed loop range in uniform steps. At each stage position $V_{\cos}$, $V_{\sin}$, V_{ref} and V_{mes} were recorded, with a 100 point moving average applied to reduce the effects of noise and higher frequency mechanical vibrations. Residual non-linearities were evaluated by comparison to the stage set position, with the displacement calculated from the stage set positions scaled to cover the same range as the interferometric displacement measurements, minimising the effects of angular misalignments between the stage and interferometer measurement axes on the non-linearity measurement. The differences between the interferometer displacements and the stage displacements were then calculated. In order to separate residual non-linearities of tens of picometres from the combined effects of air turbulence, mechanical noise, and electrical noise, a Fourier transform of the calculated displacement differences was employed. The calculated displacement differences were windowed with a Hann window function [23] to reduce spectral leakage, and zero padded, with a number of zeros equal to the number of data points adjoined to each end of the data. To compensate for the amplitude scaling effects of the window function, the amplitudes resulting from the Fourier transform were multiplied by a scaling factor calculated as the number of displacement difference data points divided by the sum over all values in the window function. For comparison, displacements corrected with the Heydemann correction [8] as implemented by Birch [12] were also calculated, along with displacements corrected with the optical shuttering technique of [15] in some cases. In order to minimise the impact of laser intensity fluctuations the quadrature signals were normalised to reference arm intensity reference signal before applying the Heydemann correction.

4. Results

4.1. Full-fringe case

First, the full-fringe case (the case for which a full ellipse of data is available for correction) was considered, for a displacement covering a range of 15 μm in 15 nm steps. A comparison of the Heydemann [8], beam shuttering [15], MIR and CH algorithms is shown in figure 3. Residual twice per fringe non-linearities (having a spatial period of 158 nm) can be seen for the MIR algorithm, with an amplitude of approximately 9 pm. For both the MIR and CH algorithms a small improvement in performance can be seen over the Heydemann algorithm at a spatial period of 316.4 nm. This effect appears to be a result of the correction for variable in-arm losses made for both corrections, whereas the Heydemann correction can only compensate for laser intensity fluctuations affecting both arms. Away from harmonics of 632.8 nm, the Heydemann, MIR and CH algorithms overlap exactly. The beam shuttering technique [15] yields residual twice per fringe non-linearities with an amplitude of 65 pm for the interferometer used in this work. These residual non-linearities are in this case a result of the shuttering technique not correcting for quadrature phase shifts. A small non-linearity component with an amplitude of 8 pm can be seen at a spatial period of 105.5 nm for all
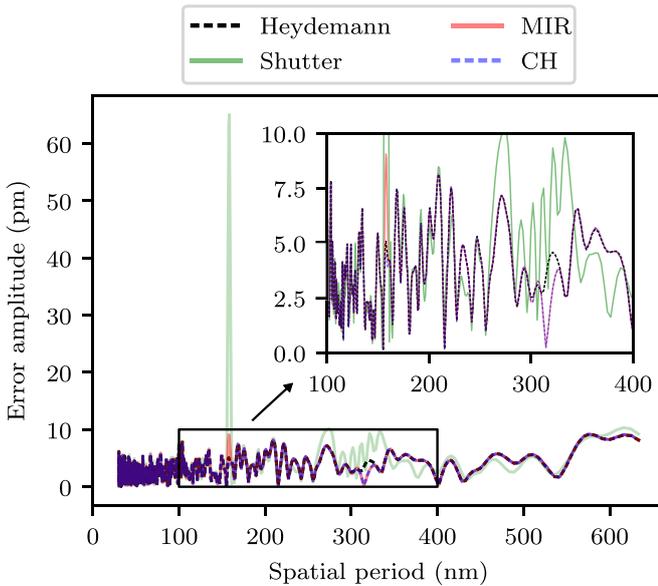
Figure 3. Comparison of the full-fringe performance of the Heydemann [8], beam shuttering [15], MIR and CH non-linearity correction algorithms when applied to a displacement covering 15 µm with a sample spacing of 15 nm. Inset shows region of interest. The Heydemann, MIR, and CH techniques overlap almost exactly away from harmonics of 632.8 nm.

Figure 4. Comparison of the full-fringe performance of the Heydemann [8], beam shuttering [15], MIR and CH non-linearity correction algorithms in the presence of wavefront distortion, when applied to a displacement covering 15 µm with a sample spacing of 15 nm. All techniques plotted overlap almost exactly away from harmonics of 632.8 nm.

Figure 5. Comparison of the full-fringe performance of the Heydemann [8], MIR and CH non-linearity correction algorithms in the presence of wavefront distortion, when applied to a displacement covering 15 µm with a sample spacing of 15 nm. Inset shows region of interest. All techniques plotted overlap almost exactly away from harmonics of 632.8 nm.

algorithms, corresponding to $\lambda/6$, which in this interferometer system appears to be introduced by internal reflections within the solid glass corner cube retroreflector. Smaller higher order harmonics are also likely to be present due to this multiple reflection effect, however they cannot be separated from the noise floor in this system.

The effects of wavefront distortions on the MIR algorithm, motivating the development of the CH algorithm, are demonstrated in figures 4 and 5. In this case some clipping was introduced into the reference arm of the interferometer, with the reference beam falling on a joint between the reflecting surfaces in the reference arm corner cube retroreflector. The high sensitivity of the beam shuttering technique to wavefront distortions can be seen. The large non-linearities seen in this case may also be a result of the sensitivity of the beam shuttering technique to changes in laser intensity between shuttering steps, and between the shuttering steps and measurement taking place. For this reason, the beam shuttering technique is not included in further comparisons. A non-linearity term can also be seen to arise with a spatial period of 158 nm and an amplitude of approximately 50 pm for the MIR algorithm, resulting from imperfect amplitude correction by equations (15) and (16). Similar increases in non-linearity were regularly observed whilst testing the MIR algorithm in cases where wavefront distortions were introduced in either arm of the interferometer before recombination of the beams.

The full-fringe case was also utilised to investigate the stability of the calibration process. A single set of calibration measurements were made followed by a series of 15 µm displacements, again made in 15 nm steps, with the displacement measurement repeated every 2.5 min. The previously
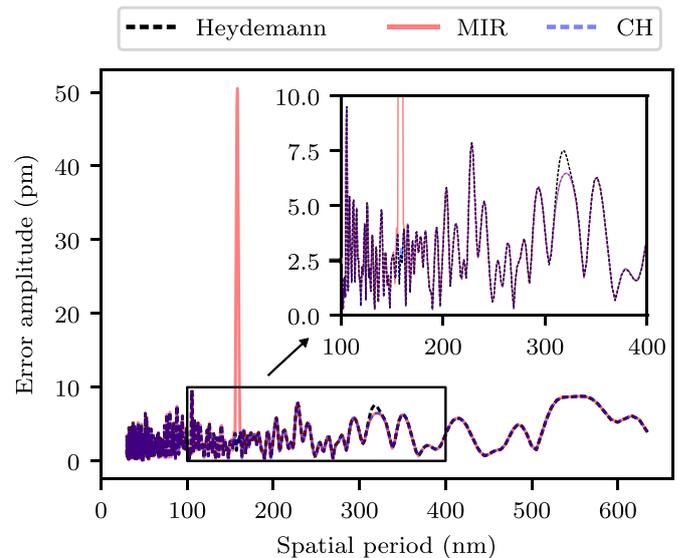
described Fourier method was then applied, and the values of the Fourier amplitudes at spatial periods of 316.4 and 158.2 nm (corresponding to $\lambda/2$ and $\lambda/4$) added for the Heydemann, MIR and CH techniques at each time point. The amplitude sums were then doubled to give an estimate of the peak-to-peak non-linearity, assuming perfect constructive interference between the amplitude terms, with the results shown in figure 6, along with a 1 h moving average of the resulting non-linearity estimates. Moving averages were calculated based upon a 1 h window centred on the plot time point,
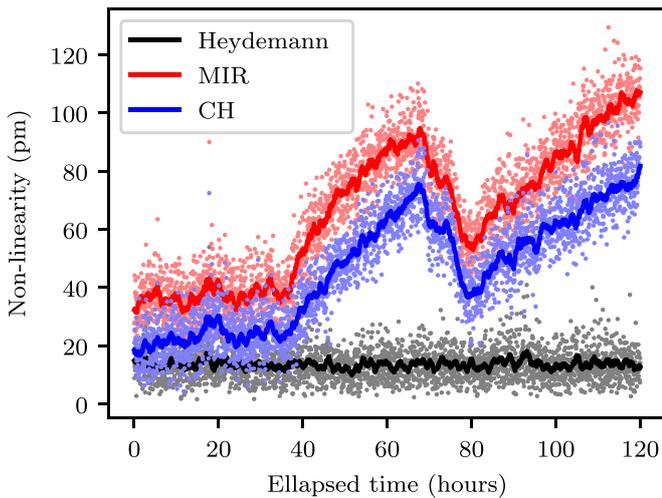
Figure 6. Time stability of MIR and CH calibration process, compared to the Heydemann correction in the full-fringe case. Both individual measurements taken every 2.5 min and 1 h moving averages are shown. Estimated peak to peak non-linearities are plotted, considering only non-linearity components with spatial periods of 316.4 and 158.2 nm.

Figure 7. Sub-fringe performance of the Heydemann, MIR and CH algorithms. Estimated peak to peak non-linearities are shown, considering only non-linearity components with spatial periods of 316.4 and 158.2 nm calculated as described in the text. Error bars indicate standard uncertainties.

with any portion of the window falling outside of the data set neglected.

4.2. Sub-fringe case

The development of the MIR and CH algorithms was driven by the desire for improved non-linearity corrections in the case where displacements cover less than a full optical fringe. As such, the sub-fringe performance of the MIR and CH algorithms was compared to that of the Heydemann correction. To test the sub-fringe performance, a displacement through a 15 μm range was performed in steps of 1.5 nm. A number of data points corresponding to the displacement range indicated on the *x* axis of figure 7 were taken from the start of the full 15 μm range, and the correction coefficients (*A* and *B* for the MIR algorithm; *C*, *E* and *F* for the CH algorithm and the full ellipse fit for the Heydemann algorithm) calculated based on this restricted data set. The full 15 μm displacement was then corrected based upon the calculated coefficients, and the previously described Fourier technique used to estimate non-linearities, as for figure 6. This process was then repeated 100 times, with randomly generated offsets applied to the restricted data sets for each repeat such that the restricted data sets were taken from different parts of the full displacement range. Averages and standard errors were then calculated for each fit range, with the results plotted in figure 7. This approach was required as the Fourier technique used to measure non-linearities in this work requires multiple fringes to give reliable results, and cannot be applied to sub-fringe measurements directly. An important consequence of this approach is that the stability of the non-linearities is also implicitly measured, along with the performance of the algorithms.

It should be noted that the data presented in figure 7 shows a higher large displacement non-linearity limit for all three
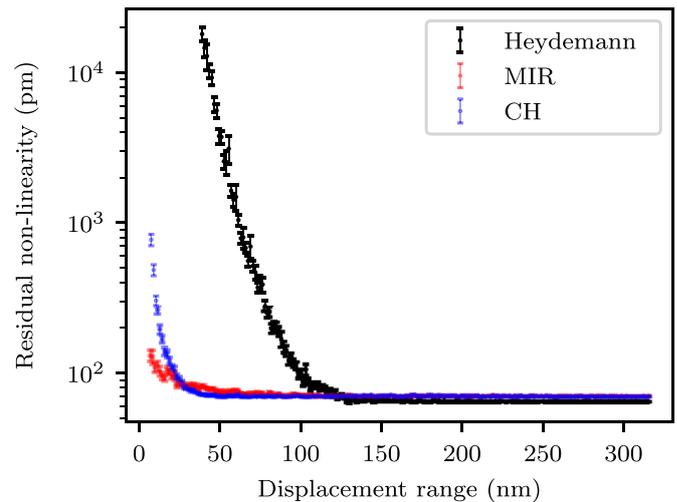
algorithms than may be expected from figure 3. This is a consequence of two effects. Firstly, the decreased displacement step size required, which, in combination with the poor mechanical stability of this proof of concept interferometer, and the Fourier non-linearity measurement technique, resulted in a higher 'noise floor', introduced primarily by mechanical drifts. Secondly, the stability of the ellipse fit parameters across the full displacement range also influences the measured residual non-linearity.

5. Discussion

In this work, the MIR and CH algorithms have been compared to the commonly employed ellipse fitting Heydemann correction for both full-fringe and sub-fringe displacements. Considering first the full fringe results, figure 3 demonstrates that the CH algorithm is capable of reaching the same non-linearity performance as the Heydemann correction, with a small increase in performance for once per fringe non-linearities possible. The MIR algorithm has been shown to have residual non-linearities with a twice per fringe spatial period, in this case having an amplitude of 9 pm. As demonstrated in figure 5, the MIR algorithm is highly sensitive to any effect that results in incomplete interference between the beams returned from each arm, with twice per fringe non-linearities rising to 50 pm in the presence of wavefront distortions from a solid glass corner cube seam. This sensitivity to incomplete interference is a result of the assumption made by the MIR technique that the amplitudes of the quadrature signals may be determined from the intensity reference outputs, an assumption that is only valid in the ideal case of identical wavefronts being returned by each interferometer arm, with identical polarisation states, and with both beams being fully coherent. As a result, some residual non-linearities are likely to be present after applying the MIR correction in most practical

interferometer designs, where incomplete interference may be introduced by wavefront distortions in the optics, misalignment between the interfering beams, or polarisation aberrations. In the interferometer described in this work the residual twice per fringe non-linearities after applying the MIR algorithm seen in figure 3 are likely to be a result of the cumulative effect of small angular alignment errors between the interfering beams introduced by imperfect cube corner retroreflectors, and polarisation aberrations introduced by reflections from the silver coated surfaces of the retroreflectors. Intensity losses, in the absence of any wavefront distortion, would not result in non-linearity, as variable losses in the arms of the interferometer are accounted for in equations (13) and (14) by T_{mes} and T_{ref}. Whilst in this case residual non-linearities introduced by imperfect interference were reduced to 9 pm in amplitude through careful alignment, this may not be possible in a system where non-linearities cannot be evaluated *in-situ*, and as a result, these highly sensitive residual non-linearity sources motivated the development of the CH approach.

Figure 3 demonstrates the best case performance achieved with this system, with the displacement measurement made immediately after calibration. The stability of the calibration process is also important if the MIR or CH techniques are to be used in practical applications. From figure 6 it can be seen that the calibration remained stable over time scales of a few hours, with the residual non-linearities beginning to increase for both the MIR and CH algorithms after 35 h. Some caution should be employed when drawing general conclusions from the data presented in figure 6, as the stability of the calibration is dependent on a range of factors that may vary between interferometer systems, and between laboratories. The first effect to be considered is the mechanical stability of the optical system. Whilst variable losses in each arm of the interferometer are accounted for by both the MIR and CH algorithms, variable losses between the pick off BSs and intensity reference detectors, or between the PQBS and the quadrature detectors are not corrected. In part, this may be compensated by ensuring the optical path lengths from the fibre collimator to the four detectors are equalised, however changes in optical alignment that affect only one detector will still result in an increase in residual non-linearities. The interferometer used in this work was assembled as a proof of concept, and as such greater mechanical stability may be expected from an interferometer designed for practical measurement tasks. Mechanical stability will also be influenced by the thermal stability of the environment. In this case the temperature remained within $\pm 0.25\ °C$ over the course of the measurements presented in figure 6. In addition to the mechanical stability of the system, the stability of the amplification electronics may affect the residual non-linearities. In particular, both the MIR and CH algorithms assume the voltage offsets in the absence of laser illumination are constant, and changes to these offsets, or changes in the amplifier gains, will result in increased residual non-linearities. The intensity stability of the laser source is not critical for either the MIR or CH algorithms, as both algorithms correct for any change in source intensity.

Whilst the influence of longer-term drifts in the calibration parameters is quantified, under the experimental conditions observed in this work, in figure 6, the effects of initial errors in the calibration voltages should also be considered. As the calibration voltages are measured over long integration times (100 000 data points over 0.5 s in this work for each calibration step), and with one arm of the interferometer blocked such that no interference takes place, there is very little influence of mechanical vibration or laser intensity fluctuations, leading to negligible short-term uncertainties in the calibration voltages as compared to the displacement measurement data. As dark voltages are subtracted as part of the calibration process, offset errors due to background light, or electronic effects, do not influence the measurement as long as the offset remains fixed. Varying offsets, for example due to varying levels of background illumination, may contribute to the longer-term drifts observed in figure 6.

In comparison to the beam shuttering technique previously proposed in the literature [15] both the MIR and CH techniques have been demonstrated in figure 4 to be less sensitive to incomplete interference effects. Additionally, the beam shuttering technique does not correct for quadrature phase shift errors, limiting the ultimate performance for the interferometer system described in this work, as shown in figure 3. The beam shuttering technique as described in [15] also does not include an intensity reference, and if implemented in this manner the correction is sensitive to changes in the laser intensity, both with regards to high frequency fluctuations and low frequency drift.

The main advantage of the MIR and CH algorithms over the Heydemann correction is demonstrated in figure 7; for displacements ranges of less than approximately 120 nm (for a single pass interferometer operating at 632.8 nm) residual non-linearities are reduced. In this work, an ordinary least squares fit following the methodology described in [12] was employed by the Heydemann correction, which may not be optimal for fitting small portions of an ellipse [9], however, even where enhanced ellipse fitting algorithms are employed, residual non-linearities in excess of 100 pm have been shown to arise for displacements covering less than one half of an optical fringe [9]. Similarly, ordinary least squares fits were used by both the MIR and CH algorithms when making constrained ellipse fits, and performance improvements may be possible with more advanced fitting algorithms.

Whilst in this work the CH algorithm has been demonstrated to be more robust than the MIR algorithm, in applications where near complete interference can be guaranteed, it can be seen from figure 7 that the MIR algorithm outperforms the CH algorithm for displacements of less than approximately 25 nm. This cross over point is likely to be highly dependent on noise levels in the interferometer system, however in general, for applications where displacements cover only a few nanometres, the MIR algorithm is likely to outperform both the CH and Heydemann algorithms. The relative computation simplicity of the fit coefficient calculations in the case of the MIR algorithm is also a significant advantage, and future work may investigate the possibility of implementing this algorithm in real time on a field programmable gate array (FPGA) device.

6. Conclusion

Two novel non-linearity correction algorithms have been proposed, both aiming to compensate for non-linearities where homodyne interferometers are applied to displacements covering less than one full optical fringe. The first, the MIR technique, is relatively mathematically simple, and may be suitable for real time implementation, however the performance has been shown to be limited in cases where incomplete interference occurs. The second approach, the CH technique, was developed for applications where incomplete interference does occur, and has been demonstrated to be capable of achieving residual non-linearities comparable to that of the Heydemann correction, whilst remaining effective down to small fractions of an optical fringe, in this work outperforming the Heydemann correction for displacements of less than approximately 120 nm. For the interferometer described here, non-linearity peaks at all harmonics of the illuminating wavelength (632.8 nm) were reduced to below 10 pm in amplitude by both techniques under ideal conditions. Whilst in this case the algorithms have been applied to a phase-quadrature coating [20] based homodyne interferometer, the approach is equally applicable to interferometers employing polarisation based techniques to achieve the required quadrature phase shift between the sine and cosine outputs.

Both the MIR and CH techniques have been compared to a beam shuttering technique previously described in the literature [15], and have been found to be less sensitive to incomplete interference and quadrature phase shift errors. The MIR and CH techniques do however require additional optical components and detection electronics as compared to the beam shuttering approach, and this may be a disadvantage in cost sensitive applications.

Future work may implement the MIR technique in real time on an FPGA device, and explore the potential for the CH algorithm to be applied to atomic force microscope cantilever measurements. Further work may also compare the performance of the algorithms to the NPL x-ray interferometer [24], permitting sub-fringe non-linearities to be evaluated directly, without recourse to a Fourier technique.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: 10.17862/cranfield.rd.16460430.

ORCID iDs

Angus Bridges ⬡ https://orcid.org/0000-0002-5861-043X
Andrew Yacoot ⬡ https://orcid.org/0000-0001-6740-821X
Thomas Kissinger ⬡ https://orcid.org/0000-0003-1832-7143
Ralph P Tatam ⬡ https://orcid.org/0000-0001-9599-3639

References

[1] 2019 Mise en pratique for the definition of the metre in the SI, Appendix 2 - SI Brochure Consultative Committee for Length, Bureau international des poids et mesures 9th edn

[2] Zhu F, Spronck J W and Heerens W C 1991 A simple capacitive displacement sensor *Sens. Actuators* A **26** 265–9

[3] Yacoot A and Cross N 2003 Measurement of picometre non-linearity in an optical grating encoder using x-ray interferometry *Meas. Sci. Technol.* **14** 148–52

[4] Yacoot A, Bosse H and Dixson R 2020 The lattice parameter of silicon: a secondary realisation of the metre *Meas. Sci. Technol.* **31** 121001

[5] Danzebrink H U, Koenders L, Wilkening G, Yacoot A and Kunzmann H 2006 Advances in scanning force microscopy for dimensional metrology *Ann. CIRP* **55** 841–78

[6] Misumi I, Gonda S, Kurosawa T and Takamasu K 2003 Uncertainty in pitch measurements of one-dimensional grating standards using a nanometrological atomic force microscope *Meas. Sci. Technol.* **14** 463–71

[7] Yacoot A, Klapetek P, Valtr M, Grolich P, Dongmo H, Lazzerini G M and Bridges A 2019 Design and performance of a test rig for evaluation of nanopositioning stages *Meas. Sci. Technol.* **30** 035002

[8] Heydemann P L M 1981 Determination and correction of quadrature fringe measurement errors in interferometers *Appl. Opt.* **20** 3382–4

[9] Požar T and Možina J 2011 Enhanced ellipse fitting in a two-detector homodyne quadrature laser interferometer *Meas. Sci. Technol.* **22** 085301

[10] Hu P, Wang Y, Fu H, Zhu J and Tan J 2017 Nonlinearity error in homodyne interferometer caused by multi-order Doppler frequency shift ghost reflections *Opt. Express* **25** 3605–12

[11] Bridges A, Yacoot A, Kissinger T and Tatam R P 2020 Polarization-sensitive transfer matrix modeling for displacement measuring interferometry *Appl. Opt.* **59** 7694–704

[12] Birch K P 1990 Optical fringe subdivision with nanometric accuracy *Precis. Eng.* **12** 195–8

[13] Wu C M, Su C S and Peng G S 1996 Correction of nonlinearity in one-frequency optical interferometry *Meas. Sci. Technol.* **7** 520–4

[14] Köning R, Wimmer G and Witkovský V 2014 Ellipse fitting by nonlinear constraints to demodulate quadrature homodyne interferometer signals and to determine the statistical uncertainty of the interferometric phase *Meas. Sci. Technol.* **25** 115001

[15] Zhu J, Hu P and Tan J 2015 Homodyne laser vibrometer capable of detecting nanometer displacements accurately by using optical shutters *Appl. Opt.* **54** 10196–9

[16] Hu P, Zhu J, Guo X and Tan J 2015 Compensation for the variable cyclic error in homodyne laser interferometers *Sensors* **15** 3090–106

[17] Eom T B, Kim J Y and Jeong K 2001 The dynamic compensation of nonlinearity in a homodyne laser interferometer *Meas. Sci. Technol.* **12** 1734–8

[18] Kim J A, Kim J W, Kang C S, Eom T B and Ahn J 2009 A digital signal processing module for real-time compensation of nonlinearity in a homodyne interferometer using a field-programmable gate array *Meas. Sci. Technol.* **20** 017003

[19] Pisani M *et al* 2012 Comparison of the performance of the next generation of optical interferometers *Metrologia* **49** 455–67

[20] Raine K W and Downs M J 1978 Beam-splitter coatings for producing phase quadrature interferometer outputs *Opt. Acta* **25** 549–58

[21] Stone J and Phillips S 1996 Corrections for wavelength variations in precision interferometric displacement measurements *J. Res. Natl Inst. Stand. Technol.* **101** 671–4

[22] Edlén B 1953 The dispersion of standard air *J. Opt. Soc. Am.* **43** 339–44

[23] Harris F J 1978 On the use of windows for harmonic analysis with the discrete Fourier transform *Proc. IEEE* **66** 51–83

[24] Yacoot A and Downs M J 2000 The use of x-ray interferometry to investigate the linearity of the NPL differential plane mirror optical interferometer *Meas. Sci. Technol.* **11** 1126–30