New Journal of Physics

The open access journal at the forefront of physics

PAPER

OPEN ACCESS

Towards probing for hypercomplex quantum mechanics in a waveguide interferometer

S Gstir[1,*], E Chan[1,2], T Eichelkraut[3,5], A Szameit[4], R Keil[1] and G Weihs[1]

1   Institut für Experimentalphysik, Universität Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria
2   École Normale Supérieure de Lyon, 65 Allée d'Italie, 69007 Lyon, France
3   Institute of Applied Physics, Abbe Center of Photonics, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany
4   Institut für Physik, Universität Rostock, Albert-Einstein-Str. 23, 18059 Rostock, Germany
*   Author to whom any correspondence should be addressed.
5   Present address: Janova GmbH, Mühlhenstraße 96, 07745 Jena, Germany.

E-mail: sebastian.gstir@uibk.ac.at

Abstract

We experimentally investigate the suitability of a multi-path waveguide interferometer with mechanical shutters for performing a test for hypercomplex quantum mechanics. Probing the interferometer with coherent light, we systematically analyse the influence of experimental imperfections that could lead to a false-positive test result. In particular, we analyse the effects of detector nonlinearity, input-power and phase fluctuations on different timescales, closed-state transmissivity of shutters and crosstalk between different interferometer paths. In our experiment, a seemingly small shutter transmissivity in the order of about 2×10^{-4} is the main source of systematic error, which suggests that this is a key imperfection to monitor and mitigate in future experiments.

1. Introduction

Quantum mechanics, like any theory in physics, is built upon a certain set of axioms. Even though they are by design not provable, they can be checked for consistency with nature by experiments. If the outcome does not falsify the axiom, it is most desirable to compare the experimental results with quantitative predictions of alternative theories, such that these alternatives can be ruled out or their parameter space restricted. For example, the Bell inequality [1, 2] bounds correlations between pairs of entangled particles based on the assumption of local realism (i.e. the presence of hidden variables governing these correlations). Experimental violations of this bound [3], especially when found in absence of loopholes [4–8] or with very high statistical significance [9, 10], allow ruling out hidden-variable models of quantum mechanics with high confidence.

Even on the more fundamental level of single-particle quantum mechanics, there is ongoing research probing the connection between wave functions and the probability as their measurable, physical quantity, as postulated by the axiom of Born's rule [11]. These experiments test for the existence of higher-order interferences [12–20], which would contradict Born's rule or linearity and would require adaptations of quantum mechanics and/or alternative theories.

An axiom, which only very few experiments were able to probe so far, is the representation of quantum mechanical states by rays in a Hilbert space based on complex numbers. In theory, states could be described using Hilbert spaces based on other fields, such as real, quaternion, octonion and possibly other hyper-complex numbers. Two notable theories using a different number system are real quantum mechanics [21, 22] and quaternionic quantum mechanics (QQM) [23]. Both theories are mathematically consistent, but only quaternionic and (standard) complex quantum mechanics (CQM) are able to correctly describe the dynamics of a two-level system like the polarization state of a photon. In order to determine

whether **CQM** is sufficient or whether **QQM** would better describe nature, specialized experiments are needed.

In 1979, Peres proposed a test to distinguish between quantum theories based on different number systems [24]. This test is based on the dimension of the phase. While complex numbers can be represented by a single real phase parameter, three phase components are needed for quaternions. This allows distinguishing between complex and hypercomplex quantum mechanics in a scattering experiment with variable presence of one to three independent scatterers. In the same publication, he also proposed a specialized version based on the commutation of phase shifts. Only in the case of **CQM** do such shifts commute. This version of the test has so far been implemented via the interference of neutrons [25] and of single photons [26], both showing no trace of quaternions whatsoever. The proposed scattering experiment can also be recast as an interference scenario with three individually switchable amplitudes. Due to its intrinsic sensitivity to a variety of experimental imperfections, such an interferometric test requires a well-characterised high-precision setup and has so far not been implemented in any system.

In this work, we designed and set up an integrated optical waveguide interferometer, to be used in a Peres three-path interference test with photons. In order to access the required amplitude combinations, each path can be individually blocked by mechanical shutters. Compared to free-space setups [17] or partially integrated solutions [27], the present setup provides much better phase stability, which is a prerequisite to obtain meaningful results. For an in-depth analysis of potential systematic errors, which would bias a Peres test with single photons, we probed the interferometer with bright coherent light with negligible shot noise. While this will not let us distinguish between **CQM** and **QQM**, it enables identification of the dominant error mechanisms to be addressed before a Peres test with single photons can be run.

2. Concept of the Peres test

In the standard formulation of quantum mechanics, a three-path interferometer with the paths A, B and C is fully described by three complex amplitudes with real phases $0 \leqslant \phi_i \leqslant 2\pi$ (with $i \in \{A, B, C\}$). Assuming full coherence, their respective phase differences $\Delta\phi_{ij} = \phi_i - \phi_j$ ($i \neq j$) must add to zero:

$$\Delta\phi_{AB} + \Delta\phi_{BC} + \Delta\phi_{CA} = 0. \tag{1}$$

Applying the cosine function to (1) and using trigonometrical transformations results in the so-called Peres parameter

$$F \equiv \alpha^2 + \beta^2 + \gamma^2 - 2\alpha\beta\gamma = 1, \tag{2}$$

with $\alpha \equiv \cos(\Delta\phi_{BC})$, $\beta \equiv \cos(\Delta\phi_{CA})$ and $\gamma \equiv \cos(\Delta\phi_{AB})$ describing the normalized interference terms. The description of quantum mechanical waves with hypercomplex numbers increases the dimension of their phase. In the case of quaternions the phase becomes three-dimensional, such that (2) does not necessarily hold, which is the starting point for the formulation of the interferometric Peres test. Deviations from the expected value $F = 1$ are, according to [24], interpreted as

$$F = \begin{cases} < 1 & \text{hypercomplex numbers are admissible} \\ > 1 & \text{superposition principle is violated.} \end{cases} \tag{3}$$

In order to connect this parameter with directly measurable probabilities, the photon detection rates P_A, P_B, P_{AB} are measured at one of the interferometer's outputs for various path transmission settings (A open, B open, A and B open, etc) and constant input flux. In the case of a two-path interferometer with paths A and B, the measured photon rate P_{AB} is given by

$$P_{AB} = P_A + P_B + 2\sqrt{P_A P_B}\,\gamma. \tag{4}$$

This equation can be used to express the normalized interference term as

$$\gamma = \frac{P_{AB} - P_A - P_B}{2\sqrt{P_A P_B}}. \tag{5}$$

The equations for α and β can be derived analogously and give

$$\alpha = \frac{P_{BC} - P_B - P_C}{2\sqrt{P_B P_C}} \quad \text{and} \tag{6}$$

$$\beta = \frac{P_{CA} - P_A - P_C}{2\sqrt{P_A P_C}}. \tag{7}$$

In order to obtain all normalized interference terms and calculate F, all shutter combinations with a single open path or two open paths need to be measured.

Various experimental imperfections like input-power fluctuations, phase fluctuations and non-zero transmissivity of the shutter can lead to deviations from $F = 1$. For a meaningful interpretation of the Peres experiment, it is therefore necessary to characterise or measure all these imperfections of the interferometer as precisely as possible (see section 5).
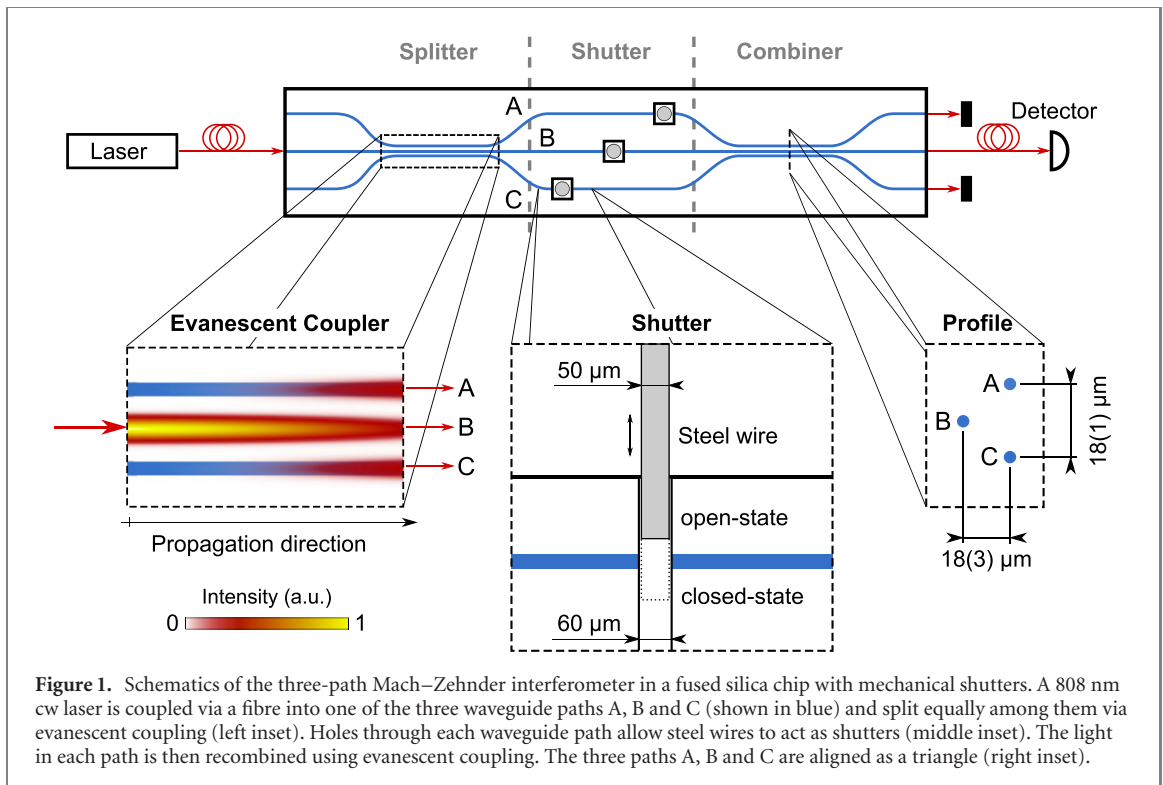
3. Experimental setup

To perform the Peres test we used a three-path Mach–Zehnder interferometer integrated in a fused silica glass chip, with external mechanical shutters, as can be seen in figure 1. The waveguides were directly inscribed using femtosecond laser pulses, which enables three-dimensional path designs [28, 29], including three-path interferometers [30]. In the present work, the three waveguides were aligned in a triangle, which was continuously rescaled in a form-preserving manner to produce the various functional elements of the chip. At minimal size (base length = 18(1) μm), the reduced spacing between each waveguide leads to evanescent coupling (see the left inset of figure 1). The length of the evanescent coupler was chosen to yield for each input an approximately balanced output splitting between the three ports [31]. Note that a perfectly balanced splitting ratio is not critical for the Peres test. One such unit opens the interferometer (henceforth termed splitter), while a mirror-inverted copy (combiner) closes it. In the region between splitter and combiner, the triangle has a maximal baselength of 201(1) μm, minimizing cross-coupling between the separate waveguides. These straight paths are intersected by rectangular 60 μm × 60 μm holes, produced by laser-inscription of nanocracks (in the same fabrication step as the waveguides) and subsequent selective chemical etching with HF-acid [32–34]. Steel wires with 50 μm diameter were inserted into these holes to act as mechanical shutters, whose open/closed state depends on the vertical position of the wire. Each wire is mounted in a ferrule and connected to a stepping motor via an eccentric. This allows to set the various shutter configurations fully automatised during the Peres test. In order to reduce losses from refraction and reflection for the open shutter position, the holes were filled with an index-matching oil, with a refractive index difference of $\Delta n \approx -5 \times 10^{-3}$ at 23 °C in comparison with the fused silica. The transmission through a hole with index-matching oil was independently measured at 635 nm to be >95%, within the accuracy of the measurement. This serves as a lower bound for the transmission at 808 nm, as longer wavelengths are less affected by the sidewall roughness of the holes.

The in- and output paths of the chip were connected to optical fibres via fiber-arrays with refractive index matching gel bridging the gap between the fibre end and the chip surface. The gel has a refractive index difference of $\Delta n \approx -3 \times 10^{-3}$ with respect to the fused silica. On the input we used a polarization-maintaining fibre and on the output side a single-mode fibre, in order to reduce the detected stray-light in comparison to using a multimode or free-space imaging. For our measurements we used port B at the in- and output side, as shown in figure 1. The light of the output fibre is collimated onto a photodiode (Physimetron A139-001). As light source we used an 808 nm laser diode (Thorlabs LD808P030), which is temperature, power stabilised. The power stabilisation is performed by a liquid crystal noise eater (Thorlabs LCC3112), which results in a relative power stability (standard deviation) of 0.32% over 9.2 h.

The close proximity of the waveguides in the evanescent couplers can lead to a rotation of the polarization, which depends on the orientation of the waveguides in the triangle relative to the orientation of the elliptical waveguide profiles [35]. As the detector is polarization insensitive and not every path experiences the same rotation, a polarizer was placed in front of the detector to project the output light into a single polarization and, thereby, prevent systematic errors in our measurements (see appendix A.4).

The whole fused silica chip was enclosed in an aluminium housing. The bottom and sides of the housing are insulated with polystyrene sheets, with gaps for the shutter holes and the in- and output ports of the photonic circuit (see figure C1). The housing was temperature stabilized so that the largest observable fluctuations were smaller than 10^{-3} K by a proportional–integral-controlled resistive heater at the bottom of the housing, centred below the shutter holes. The temperature sensor was mounted on the outside of the housing and was almost completely covered by the insulation. As the top part of the housing was not

Figure 1. Schematics of the three-path Mach–Zehnder interferometer in a fused silica chip with mechanical shutters. A 808 nm cw laser is coupled via a fibre into one of the three waveguide paths A, B and C (shown in blue) and split equally among them via evanescent coupling (left inset). Holes through each waveguide path allow steel wires to act as shutters (middle inset). The light in each path is then recombined using evanescent coupling. The three paths A, B and C are aligned as a triangle (right inset).

insulated, the heater created a vertical temperature gradient. Due to the triangular arrangement of the three interferometer paths, this led to a relative phase change between path B (tip of the triangle) and paths A and C (base of the triangle). Therefore, it was possible to slightly tune the phase differences $\Delta\phi_{AB}$ and $\Delta\phi_{BC}$ simultaneously by changing the housing temperature. They changed by roughly -2.5×10^{-2} rad K^{-1} in the range from 25 °C to 35 °C.

The housing was mounted with nylon screws onto machinable glass-ceramic posts, to reduce the heat flow at the bottom of the housing and, thereby, increase the strength of the temperature gradient. Furthermore, these posts reduced the influence of temperature changes on the in- and output coupling efficiency of the chip to the fibre arrays, thanks to their lower thermal expansion coefficient in comparison to stainless steel posts. We inserted indium foil between the chip and its housing to eliminate air gaps and increase the thermal conductance.

4. Experimental results

We measured the transmitted optical power for all eight shutter combinations in random order and repeated these cycles multiple times (each time with a new independent random ordering to average over possible long-term drifts). The configuration with all three shutters closed P_0 represents the experimental background, e.g. stray-light or detector dark-current. We subtracted the measured background from the other combinations of the same cycle to correct for these influences. The shutter configuration of all paths open P_{ABC} is only needed to analyse specific systematic errors as detailed in section 5.4. We performed the Peres test for two different temperature settings of the chip housing, 23 °C and 30 °C, respectively. The temperature stability is 3 mK to 9 mK (standard deviation) for timescales in the order of the presented measurements. The two different temperatures correspond to different phase differences between the waveguides at the base and the one at the tip of the triangular path arrangement. The results of both measurements were filtered for shutter malfunctions. The durations of the two measurements were 9.2 h and 13.2 h for 23 °C and 30 °C, respectively. Figure 2 shows the result of the Peres test for each cycle at the two temperature settings of 23 °C and 30 °C, respectively. The mean of the measured Peres parameter F is

$$\langle F_{23°C}\rangle = 0.9553(4) \quad \text{and} \quad \langle F_{30°C}\rangle = 0.9683(9), \tag{8}$$

where the indicated uncertainties are the standard errors of the mean corrected for the autocorrelation according to [36]. Both results show a significant deviation in the order of -10^{-2} from the expected value of $F = 1$, which we attribute to systematic errors of the experiment. We will go into more detail about possible sources of such errors in section 5.
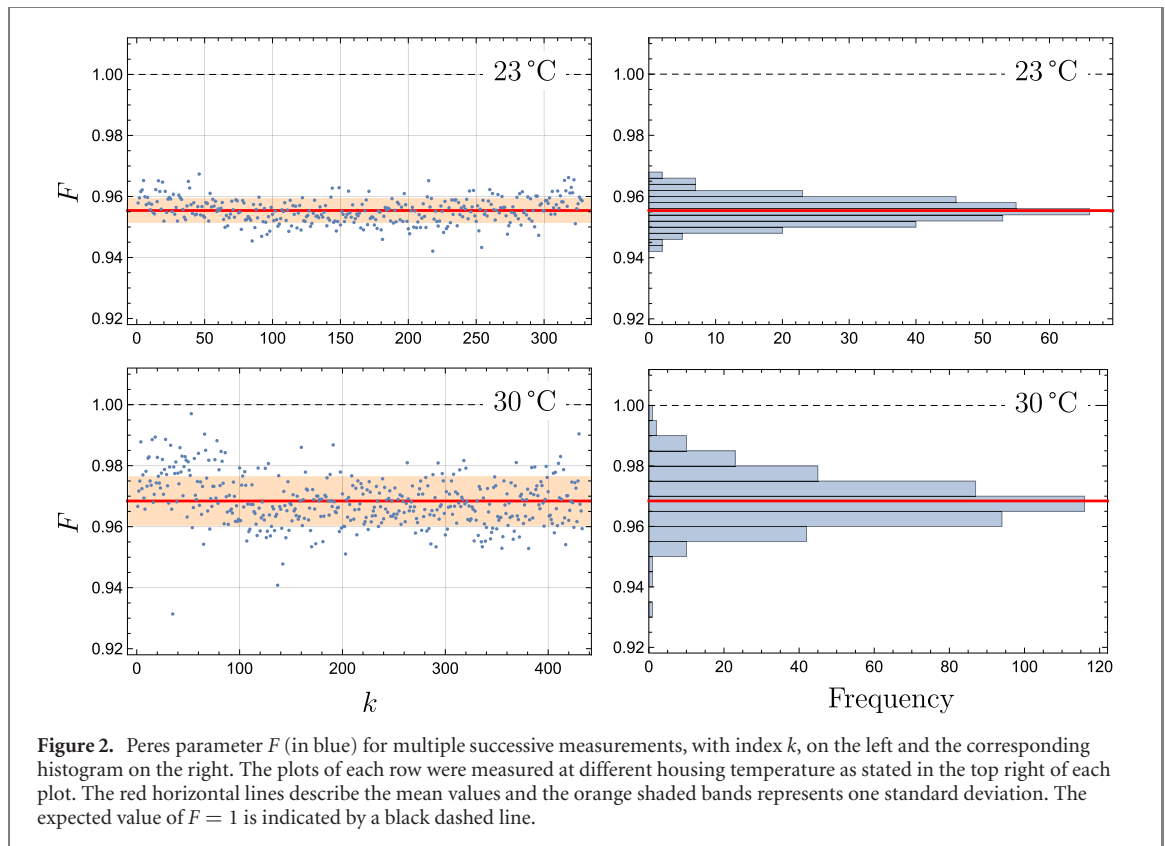
Figure 2. Peres parameter F (in blue) for multiple successive measurements, with index k, on the left and the corresponding histogram on the right. The plots of each row were measured at different housing temperature as stated in the top right of each plot. The red horizontal lines describe the mean values and the orange shaded bands represents one standard deviation. The expected value of $F = 1$ is indicated by a black dashed line.

Table 1. Directly reconstructed and corrected sets of normalized interference terms $\{\alpha, \beta, \gamma\}$ averaged over all cycles for the two housing temperatures T_h. The corrected sets satisfy (1) with minimal Euclidean distance to the directly reconstructed sets. All given errors are the standard deviation over all cycles.

T_h		$\langle\alpha\rangle$	$\langle\beta\rangle$	$\langle\gamma\rangle$
23 °C	Reconstructed	−0.765(16)	0.961(3)	−0.664(15)
	Corrected	−0.806(15)	0.941(5)	−0.612(15)
30 °C	Reconstructed	−0.405(19)	0.980(5)	−0.355(16)
	Corrected	−0.449(18)	0.989(3)	−0.310(20)

5. Systematic error analysis

The Peres parameter is influenced by various experimental imperfections, e.g. fluctuations of the input power and cross-influence of shutters. How much any of these effects influence the result depends on the prevailing phase differences of the interferometer paths. These can be directly reconstructed from the measured combinations using (5)–(7). However, due to the very same error sources influencing the Peres test, the total of the so-obtained phase differences is not zero. In order to estimate the original uninfluenced set of phases for each temperature setting, we calculated the closest set of phases, that is consistent with (1) (see appendix A.1 for details). With this set of corrected phases we can then apply various models of experimental imperfections and determine their influence on the measured Peres parameter. The directly reconstructed and the corrected phase differences are listed as normalized interferences terms (cosine of the phase difference) in table 1.

In the following, we will analyse a number of effects that lead to a systematic shift of the Peres parameter. In general these effects will also lead to random errors.

5.1. Detector nonlinearity
In tests searching for higher-order interferences, any nonlinearity in the experiment can lead to a false-positive detection of higher-order interferences [17, 37]. Similarly, nonlinearities also have a systematic effect on the Peres test, because they bias the addition of the rates in (5)–(7). In our experiment, which uses low-power continuous-wave illumination, the photodetector and its subsequent electronics are the

Table 2. Influence on the Peres parameter F caused by power and phase fluctuations, respectively. The mean of the distribution calculated with a Monte Carlo simulation is given as a deviation from the expected value $\Delta F = F - 1$. The standard deviation is described by σ_F. The results are listed for the two datasets at different housing temperatures T_h.

T_h	Power fluctuations		Phase fluctuations	
	ΔF^{Pow}	$\sigma_{F\text{Pow}}$	ΔF^{Phase}	$\sigma_{F\text{Phase}}$
23 °C	5.1×10^{-4}	1.8×10^{-2}	0.8×10^{-4}	1.0×10^{-2}
30 °C	1.4×10^{-4}	1.4×10^{-2}	-8.1×10^{-4}	1.4×10^{-2}

dominant sources of nonlinearity, which need to be accounted for. Therefore, we measured the nonlinearity of the detector in independent beam combination experiments, following the procedure outlined in [38]. Correcting the measured data of the Peres experiment for this detector nonlinearity results in a correction of F by

$$\Delta F^{\text{NL}}_{23°\text{C}} = -4.7 \times 10^{-5} \quad \text{and} \quad \Delta F^{\text{NL}}_{30°\text{C}} = -9.4 \times 10^{-5}. \tag{9}$$

These deviations are negligible compared to the observed deviation. Thus, we can rule out nonlinearity as a relevant source of error in our experiment.

5.2. Fluctuations

The measurements shown in this work were taken over periods of 9 h and 13 h and temporal effects can occur on various timescales. The effect of slow drift in power or phase difference on a timescale much longer than a measurement cycle ($\sim$1.75 min) is reduced by measuring all eight path combinations in random order. Power fluctuations occurring on timescales shorter than the measurement duration for one shutter combination ($\sim$13 s) are averaged over multiple successive data points. Phase fluctuations on that timescale, however, reduce the interference contrast, which is considered in section 5.3. In this section, we concentrate on power and phase fluctuations on the intermediate timescale of one measurement cycle.

In a first scenario, we assumed fluctuations of the input power during a cycle and overall perfect phase stability. To simulate this, we expressed the power P_i of each single path combination $i \in \{A, B, C\}$ as

$$P_i = P_{\text{in},i} \cdot T_i. \tag{10}$$

The variable $P_{\text{in},i}$ is the average power coupled into the chip during the measurement of the combination i and T_i the fraction passing through the open shutter i and ending up at the path of the detector, effectively depending on the overall splitting factor of both beam splitters of the photonic circuit and the total transmission. For the two-path combinations $i, j \in \{A, B, C\}$, with $i \neq j$, the power P_{ij} is given by

$$P_{ij} = P_{\text{in},ij} \left(T_i + T_j + 2\sqrt{T_i T_j} \cos(\Delta\phi_{ij}) \right). \tag{11}$$

As the total transmission of the chip factors out in the analysis, it is sufficient to determine the normalized splitting factors $t_i = T_i/(T_A + T_B + T_C)$, which were measured in a separate experiment as $t_A = 0.26$, $t_B = 0.52$ and $t_C = 0.22$. While the phase differences were calculated from the corrected normalized interference terms (see table 1). Each input power is treated as an independent normally distributed random variable with mean and standard deviation extracted from the measured data (for details see appendix A.2). Based on the measurement data we can determine the standard deviation of the fluctuations on either the timescale of a single shutter setting or the timescale of the whole measurement, but not directly on the scale of one cycle. The standard deviation on the timescale of the whole measurement is larger and thus leads to more deviations in the Peres parameter. We therefore use it as an upper-bound estimation for the effect of input-power fluctuations. We calculated the deviations of the Peres parameter from the expected value $\Delta F = F - 1$ and its spread σ_F with a Monte Carlo simulation. The results in table 2 show that power fluctuations mainly lead to a random uncertainty of the measured Peres parameter and only a small systematic shift.

To estimate the impact of phase fluctuations, we considered fluctuations on a timescale slower than one shutter combination. Any faster fluctuations would lead to a reduction of the interference contrast, which we discuss in section 5.3. To model the slower phase fluctuations we used equations (10) and (11) and assumed perfect stability of the input power for each combination $P_{\text{in},ij} = P_{\text{in},i} = P_{\text{in}}$. The phase differences of each two-path combination are modelled by independent Gaussian distributed numbers with mean values given by the corrected phase differences (see table 1) and the standard deviation given by the observed fluctuations on the timescale of the whole measurement (see appendix A.2 for details). We again

performed a Monte Carlo simulation and obtained the values for ΔF and σ_F as given in table 2. As it turns out, the phase fluctuations have a similar impact as the power fluctuations, i.e. they mainly lead to a larger random uncertainty and only to a small systematic shift of the Peres parameter. The latter is for both types of fluctuations two orders of magnitude smaller than the observed ΔF and therefore negligible.

5.3. Interference contrast

Phase fluctuations that are faster than the duration of one shutter setting lead to a reduction of the normalized interference terms $\chi \in \{\alpha, \beta, \gamma\}$. Our model of this effect describes the measured normalise interference terms $\tilde{\chi}$ as

$$\chi \longrightarrow \tilde{\chi} = C_\chi \cdot \chi, \tag{12}$$

where $C_\chi < 1$ is the reduced interference contrast. This model predicts a strong dependence of the Peres parameter on the interference contrast C_χ of each pair of paths. Assuming the same interference contrast for all three interference terms ($C_\chi = 1 - \Delta C$), the deviation of the Peres parameter is calculated as follows:

$$\Delta F^C = 2 \left(\alpha\beta\gamma - 1 \right) \Delta C - (4\alpha\beta\gamma - 1) \Delta C^2 + 2\alpha\beta\gamma \cdot \Delta C^3. \tag{13}$$

To estimate the reduction of the interference contrast, we used the corrected phase fluctuations on the timescale between a shutter setting ($\sim$13 s) and a single measurement ($\sim$0.11 s). How these were extracted from the measured data is outlined in appendix A.2. We used these fluctuations to calculate the interference contrasts with a Monte Carlo simulation around the mean phase difference $\Delta\phi = 0$ for each interference term. Due to the maximum of the cosine function at this point, phase fluctuations lead to a unidirectional reduction of the power, such that the resulting interference contrast serves as an estimate at all possible mean phases. However this method does not consider the effect of fluctuations on a timescale below a single measurement (above $\sim$9 Hz). As worst a case estimate, we used the smallest of these interference contrasts ($C = 0.999\,998$) for each normalized interference term $\tilde{\chi}$ resulting in a deviation from the Peres parameter ΔF^C in the order 10^{-6} for both measurements. Compared to the measured deviations this effect is negligible.

We also separately measured the interference contrast of our setup for one path combination by thermally tuning the phase difference, which gives a contrast of $C_\alpha = 1.000(29)$. The measurement and its evaluation is described in appendix A.3. Note that the number of subsequent measurements per shutter setting is only 1/10 of the number in the main experiment. This leads to reduced duration of each shutter setting and therefore reduces the range of fluctuation frequencies affecting the contrast. Hence this interference contrast probably underestimates the degradation of contrast from fluctuations.

5.4. Crosstalk

Ideally, the state of each shutter should only influence the path it is inserted to. However, experimentally one cannot exclude the possibility that closing one shutter also affects the light in the other two paths. The moving steel wires could for example lead to bending of the chip and with it to a change of the phase of the other paths. In theory, the shutter could also influence the transmitted power of other paths, however this seems less likely than an influence on the phase and is, therefore, not considered here.

In the following model of phase crosstalk, the closed shutter of path i changes the phase of another path j by $\Delta\phi_j^i$ as opposed to the open state in path i. This is shown for our triangular path arrangement in figure 3. We constrained these six parameters by symmetry considerations: Due to the slightly elliptical waveguide profiles, which are common for femtosecond laser writing, the symmetry axis of the interferometer runs vertically through path B (cf figure 3). Therefore, it seems very plausible to impose this mirror symmetry also on the crosstalk terms. That is, we assumed that the crosstalk originating from path A is equal to the one from path C and the crosstalk originating from B is the same for both of the other paths:

$$\Delta\phi_C^A = \Delta\phi_A^C \equiv \Delta\phi_H,$$
$$\Delta\phi_B^A = \Delta\phi_B^C \equiv \Delta\phi_D^{CA}, \tag{14}$$
$$\Delta\phi_A^B = \Delta\phi_C^B \equiv \Delta\phi_D^B.$$

We further assume the two diagonal phase crosstalks $\Delta\phi_D^B$ and $\Delta\phi_D^{CA}$ to be of same magnitude

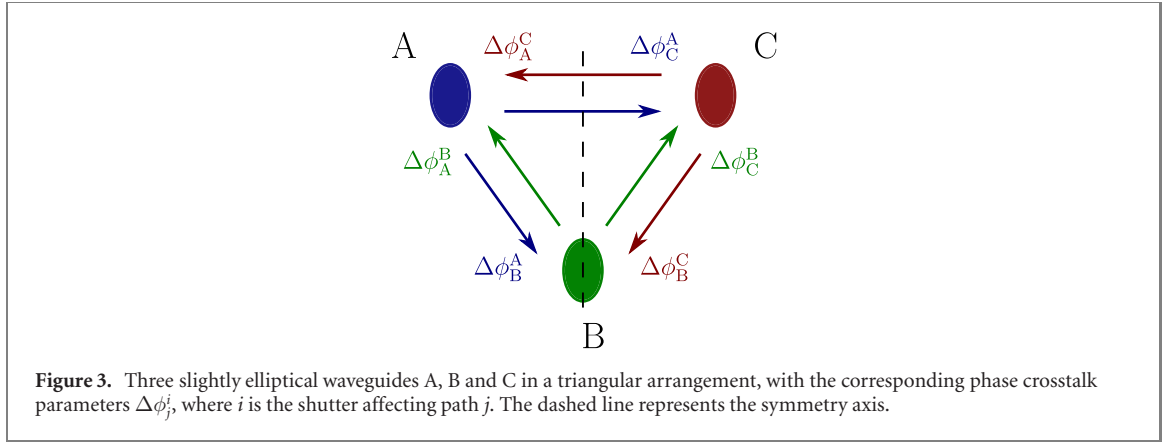$$\Delta\phi_D^B = \Delta\phi_D^{CA} \equiv \Delta\phi_D. \tag{15}$$

Figure 3. Three slightly elliptical waveguides A, B and C in a triangular arrangement, with the corresponding phase crosstalk parameters $\Delta\phi_j^i$, where i is the shutter affecting path j. The dashed line represents the symmetry axis.

Table 3. Calculated deviation of the Peres parameter from its expected value $\Delta F^{\mathrm{CT}} = F^{\mathrm{CT}} - 1$ due to phase crosstalk with a phase shift of $\Delta\phi_{\mathrm{DH}}$. The results are given for the two measurements at different housing temperatures T_{h}.

T_{h}	$\Delta\phi_{\mathrm{DH}}$ $(10^{-2}\ \mathrm{rad})$	ΔF^{CT} (10^{-3})
23 °C	−1.7	−9.0
30 °C	−1.4	6.6

These simplifications lead to the following equations for the normalized interference terms

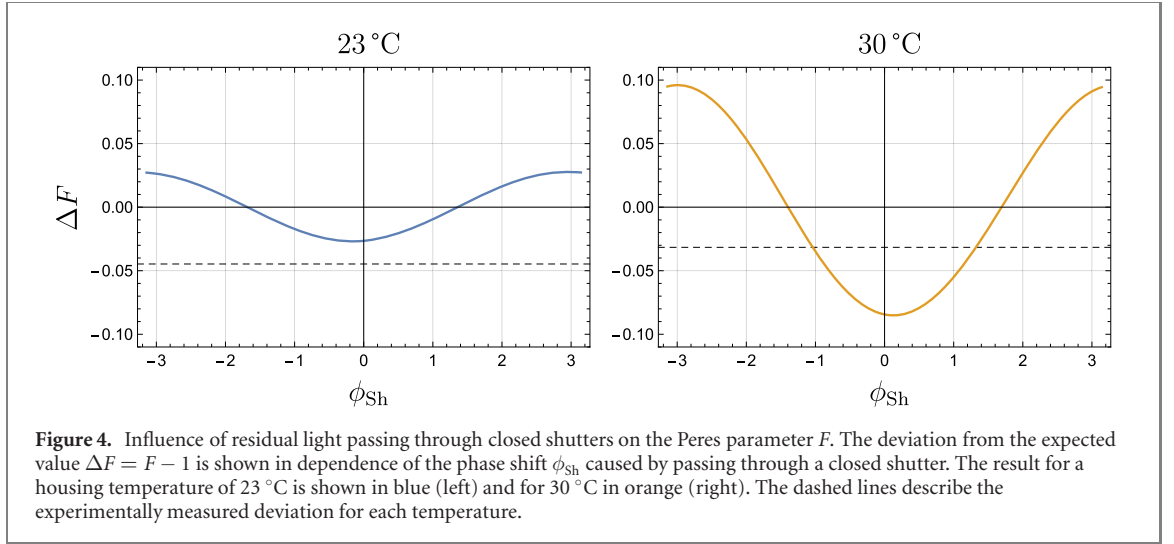$$\alpha = \cos(\Delta\phi_{\mathrm{BC}} + \Delta\phi_{\mathrm{DH}}),$$
$$\beta = \cos(\Delta\phi_{\mathrm{CA}}), \tag{16}$$
$$\gamma = \cos(\Delta\phi_{\mathrm{AB}} - \Delta\phi_{\mathrm{DH}}),$$

with $\Delta\phi_{\mathrm{DH}} = \Delta\phi_{\mathrm{D}} - \Delta\phi_{\mathrm{H}}$. The effects on the normalized interference term β cancel each other out, leaving only the parameter $\Delta\phi_{\mathrm{DH}}$ to be quantified. For this purpose, we additionally performed the Sorkin test, which probes for higher-order interferences in quantum mechanics (see [12–20] and appendix B), on the same set of data used for the Peres test. The Sorkin test has the advantage that of all effects considered in this chapter only crosstalk and nonlinearity (which can here be excluded due to its miniscule effect on the measured data; cf section 5.1) can lead to a systematic shift in its result, ϵ [17]. Therefore, the deviation from the expected value $\epsilon = 0$ can be used to estimate the magnitude of the crosstalk present in our setup. The measured ϵ (see appendix B) deviate significantly from $\epsilon = 0$ for both temperatures. Using our crosstalk model one can then calculate the crosstalk strength $\Delta\phi_{\mathrm{DH}}$ required to produce the observed deviation. Based on these results we then calculated the expected deviation of the Peres parameter ΔF as listed together with $\Delta\phi_{\mathrm{DH}}$ in table 3. Note that we also tested an asymmetric model, where all phase crosstalk arises from a single path (e.g. from one wire moving under stronger friction than the others). This delivered consistently smaller $|\Delta F|$. The results of this section indicate that some crosstalk is present in our experiment. It could well lead to a non-negligible bias on the Peres test, yet the size of this effect seems to fall short of the observed deviations $|\Delta F|$ by one order of magnitude.

5.5. Residual light

Unfortunately, our wire-based shutters exhibit finite transmissivity in their closed state. Therefore we assign to the shutter in each path i a closed-state transmissivity τ_i, due to residual light passing through the wire-blocked hole. This additional light can interfere with the light from open paths and influence the measured Peres parameter. To simplify things, we consider the same transmissivity for each shutter ($\tau_i = \tau \forall i \in \{\mathrm{A, B, C}\}$) and the phase of the residual light to be shifted with respect to the phase in the open-state by ϕ_{Sh}. With this model the measured background P_0 (all shutters closed), the single path P_i and the two path combinations P_{ij} are described by

$$P_0 = \tilde{P}_0 + P_{\mathrm{in}}\tau \left(\sum_i T_i + \sum_{i,j} (1 - \delta_{i,j}) \sqrt{T_i T_j} \cos(\Delta\phi_{ij}) \right) \tag{17}$$

Figure 4. Influence of residual light passing through closed shutters on the Peres parameter *F*. The deviation from the expected value $\Delta F = F - 1$ is shown in dependence of the phase shift ϕ_{Sh} caused by passing through a closed shutter. The result for a housing temperature of 23 °C is shown in blue (left) and for 30 °C in orange (right). The dashed lines describe the experimentally measured deviation for each temperature.

$$P_i = \tilde{P}_0 + P_{\text{in}} \left(T_i + \tau \sum_{j \neq i} T_j + 2\sqrt{\tau T_i} \sum_{j \neq i} \sqrt{T_j} \cos(\Delta\phi_{ij} + \phi_{\text{Sh}}) + 2\tau\sqrt{T_k T_l} \cos(\Delta\phi_{kl}) \right) \tag{18}$$

$$P_{ij} = \tilde{P}_0 + P_{\text{in}} \left(T_i + T_j + \tau T_k + 2\sqrt{T_i T_j} \cos(\Delta\phi_{ij}) + 2\sqrt{\tau T_i T_k} \cos(\Delta\phi_{ik} + \phi_{\text{Sh}}) + 2\sqrt{\tau T_j T_k} \cos(\Delta\phi_{jk} + \phi_{\text{Sh}}) \right), \tag{19}$$

with $i, k, l \in \{A, B, C\}$ and $i \neq k \neq l$. The input power P_{in} already includes in- and out-coupling efficiencies. The transmissivities T_i are the percentage of light passing through path *i* and reaching the detector. The power $\tilde{P}_0$ describes a generic shutter-independent incoherent background, e.g. from straylight or detector dark current. This term subsequently drops out, when calculating the normalized interference terms, as the measured background is subtracted from the single- and two-path combinations.

We estimated the shutter transmissivity based on the measured background signal P_0. For this calculation we assumed $P_i = \tilde{P}_0 + P_{\text{in}} T_i$, i.e. neglecting the contribution of τ compared to T_i. Furthermore we assumed that the shutter-independent background is only given by the dark voltage of the detectors $\tilde{P}_0 = P_{\text{dark}}$. Any additional shutter-independent background would further decrease the shutter transmissivity and, therefore, the result would act as an upper bound. Applying these assumptions to (17), results in

$$\tau = \frac{P_0 - \tilde{P}_0}{(\bar{P}_A + \bar{P}_B + \bar{P}_C + 2\sqrt{\bar{P}_A \bar{P}_B}\gamma + 2\sqrt{\bar{P}_A \bar{P}_C}\beta + 2\sqrt{\bar{P}_B \bar{P}_C}\alpha)}, \tag{20}$$

where $\bar{P}_i = P_i - \tilde{P}_0$, describing the experimentally measured single path terms P_i with the shutter-independent background $\tilde{P}_0$ already subtracted. For the set of normalized interferences α, β and γ, we used the corrected points in phase space. Evaluation of this equation for each measured cycle gives an average shutter transmissivity of

$$\langle \tau_{23°\text{C}} \rangle = 2.20(2) \times 10^{-4} \quad \text{and} \quad \langle \tau_{30°\text{C}} \rangle = 2.707(4) \times 10^{-4}, \tag{21}$$

with the uncertainty being the standard error of mean resulting from variations of the corrected points in phase space, as well as variations of the measured single path and background signals during the experiment. Applying these shutter transmissivities to our model results in the data shown in figure 4 for each housing temperature. The deviation of the Peres parameter from its expected value can be negative as well as positive and strongly depends on the phase shift ϕ_{Sh} of the residual light. The fact that the bias is stronger for the 30 °C measurement is not due to the slightly larger shutter transmissivity, but due to its point in phase space causing an increased sensitivity to the effect. Due to the oscillations in ΔF with the unknown phase ϕ_{Sh}, we cannot reliably estimate a specific figure for the impact of the non-zero shutter transmissivity on our measured results, but can only provide a range given by its extrema. This range is wide enough to contain our experimental result (30 °C) or at least in the same order of magnitude as the observed deviation (23 °C).

5.6. Conclusion
The estimated effect of the various imperfections discussed and modelled in this work on the Peres parameter *F* is summarized in table 4. The results suggest that the only imperfection resulting in deviations

Table 4. The calculated deviation of the Peres parameter F from $F = 1$ for different models and each of the two measurements at a housing temperature of 23 °C and 30 °C, respectively. Both fluctuation models only provide an upper bound for their effect and have a lower bound of 0. The residual light lists the deviation as an upper and lower bound, which corresponds to a phase shift of the transmitted light of $\phi_{sh} \approx \pm\pi$ and $\phi_{sh} \approx 0$, respectively. The last two rows list the total as an upper and lower bound and the experimentally measured deviation for each measurement.

| Model | $F - 1$ | | Notes |
	23 °C	30 °C	
Detector nonlinearity	-4.7×10^{-5}	-9.4×10^{-5}	
Power fluctuations	$< \quad 4.1 \times 10^{-4}$	$< \quad 1.4 \times 10^{-4}$	
Phase fluctuations	$< \quad 1.8 \times 10^{-4}$	$> -8.1 \times 10^{-4}$	
Interference contrast	-2.5×10^{-6}	-4.2×10^{-6}	
Crosstalk phase	-6.4×10^{-3}	6.6×10^{-3}	
Residual light	$< \quad 2.8 \times 10^{-2}$	$< \quad 9.6 \times 10^{-2}$	$\phi_{Sh} \approx \pm\pi$
	$> -2.7 \times 10^{-2}$	$> -8.7 \times 10^{-2}$	$\phi_{Sh} \approx 0$
Total	$< \quad 2.2 \times 10^{-2}$	$< \quad 10.2 \times 10^{-2}$	
	$> -3.3 \times 10^{-2}$	$> -8.1 \times 10^{-2}$	
Measured	$-4.47(4) \times 10^{-2}$	$-3.16(9) \times 10^{-2}$	

of the same order of magnitude as experimentally measured is the residual light transmission through closed shutters. While in the case of the 30 °C measurement this effect could have caused the entire deviation, for the 23 °C measurement, it falls somewhat short of the observed deviation, even if all other considered effects are included. One has to keep in mind, though, that such a sum of all deviations does not take into consideration the interplay of multiple error sources and that each model rests upon certain assumptions and simplifications. For example, if one assumes that the residual light is only caused by shutter C ($\tau_A = \tau_B = 0$) instead of equally by all shutters ($\tau_A = \tau_B = \tau_C$), the lower bound on the deviation of the Peres parameter is reduced by 1.1×10^{-2} to -3.8×10^{-2}. This results in a total deviation of -4.4×10^{-2}, which would almost eliminate the observed difference. All considered potential systematic error sources, besides residual light, lead to effects at least one order of magnitude below the measured deviation. The different deviations for both housing temperatures show that the effect of each error source can be minimized by choosing an optimal set of phase differences.

In order to narrow down the range resulting from the residual light model, phase modulators in each of the paths would be necessary. This could be achieved by placing three heaters on the surface of the chip in between both splitters [39–41]. Arranging them in a similar triangle as the paths allows to create temperature gradients between each of the paths and with that tuning of the phase differences between them. Using phase modulators to sweep the phase of each path would allow determining ϕ_{Sh} and thus increase the accuracy of this analysis. Alternatively, one could cancel out the interference effect (relative magnitude $\sqrt{\tau}$) by scrambling the phase in all closed channels, which would reduce the effect of residual light transmission (to order τ). Control of each path's phase, would furthermore allow to use a specific feature of the three-amplitude Peres test: sweeping one phase leads to a modulation of the Peres parameter F in the case of QQM. This would permit to isolate the fingerprint of a quaternionic phase from experimental imperfections with different phase dependencies. Furthermore, this modulation technique would allow to identify cases, which otherwise would lead to negative results [26]. Even though the actual value of ϕ_{Sh} and, with it, the actual impact on F cannot be determined from our experiment, these results clearly show that shutters with a low residual transmission in the closed-state and, ideally, full phase control are critical for any interferometric test of hypercomplex quantum mechanics.

Acknowledgments

Data availability statement

Data underlying the results presented in this paper are available in [43].

Appendix A. Simulation details

A.1. Reconstruction of the point in phase space
The results of each systematic error model presented in this work depend on the point in phase space, where the Peres test is performed, i.e. the set of phase differences $\{\Delta\phi_{BC}, \Delta\phi_{CA}, \Delta\phi_{AB}\}$. For conventional interference of complex amplitudes, by definition, the sum of this set must be zero (see (1)). However experimental imperfections, like the ones analysed in this work (cf section 5), can lead to a deviation from zero and to seemingly non-physical phase differences. The goal here is to estimate the most probable original physical point in phase space, with phase differences summing to zero and therefore also $F = 1$. This provides a starting point to estimate the effect of each individual experimental influence on the Peres parameter. The following section shows how we determined this point in detail.

The phase space is three-dimensional, as shown in figure A1. The physically allowed phases $\Delta\phi_{BC} + \Delta\phi_{CA} + \Delta\phi_{AB} = 2\pi n$, with $n \in \mathbb{Z}$, form a periodic set of planes in phase space. From the measured set of phase differences we find the closest (using Euclidean distance) physical point in phase space by normal projection onto the planes. Due to the sign ambiguity in the cosine function (e.g. $\alpha = \cos(\Delta\phi_{BC})$), the measured set $\{\alpha, \beta, \gamma\}$ corresponds to eight points. As a global sign flip of all phase differences has no effect on any of the models considered, this reduces to two sets of four inequivalent points, whose projections are shown in orange and blue, respectively. We therefore only consider one of the two sets e.g. the one depicted in blue. Within each set, two points project onto the $n = 0$ plane and two points onto the $\pm 2\pi$-planes. The latter ones are excluded, as they would require much larger shifts, most likely due to inconsistent sign assignments. From the two remaining points we chose the one closer to the $n = 0$ plane for each temperature. This results in the corrected phase differences listed in table 1. Note that choosing the second-closest point to the $n = 0$ plane gives results in the same order of magnitude. The only exception is the phase crosstalk at 30 °C, where the point farther from the $n = 0$ plane leads to no real solution for $\Delta\phi_{DH}$ (see section 5.4).

A.2. Power and phase fluctuations
To simulate the effect of power and phase fluctuations on the Peres parameter (see section 5.2) the magnitude of the respective fluctuations has to be determined. The signal of a single open path P_i is only influenced by fluctuations of the input power and we therefore used its standard deviation to describe this fluctuation. In the case of two open paths, the measured signal fluctuations are a superposition of input-power and phase fluctuations. To estimate the strength of the power fluctuations we calculated the standard deviation of the two-path signals based on the single-path power fluctuations using Gaussian error propagation. The signal of two open paths i and j is described by

$$P_{ij} = P_i + P_j + 2\sqrt{P_i P_j}\,\cos(\Delta\phi_{ij}) \tag{A.1}$$

andtherefore its input-power fluctuations $\sigma_{P_{ij}^{pow}}$ as:

$$\sigma_{P_{ij}^{pow}} = \sqrt{\left(\sqrt{\frac{P_j}{P_i}}\cos(\Delta\phi_{ij}) + 1\right)^2 (\sigma_{P_i})^2 + \left(\sqrt{\frac{P_i}{P_j}}\cos(\Delta\phi_{ij}) + 1\right)^2 (\sigma_{P_j})^2}, \tag{A.2}$$

where σ_{P_i} and σ_{P_j} are the standard deviation of the single-path signal for path i and j respectively. We assumed phase and power fluctuations to be independent of each other and therefore the measured standard deviation of a two-path term is described by

$$\sigma_{P_{ij}} = \sqrt{(\sigma_{P_{ij}^{pow}})^2 + (\sigma_{P_{ij}^{ph}})^2}. \tag{A.3}$$

With this equation we calculated the phase fluctuation $\sigma_{P_{ij}^{ph}}$ using the experimentally measured fluctuation $\sigma_{P_{ij}}$ for the open paths i and j.

Figure A1. Phase space of a three-path interferometer, which is defined by its three phase differences. The black points are calculated from the measurement data averaged over all cycles. The planes describe the relation $\Delta\phi_{BC} + \Delta\phi_{CA} + \Delta\phi_{AB} = 2\pi n$ with $n \in \mathbb{Z}$. The case $n = 0$ is shown in orange, $n = 1$ in blue and $n = -1$ in green. The orange and blue points represent the closest points on the plane to the measured points (black). The dashed lines connect each measured point with its point on the plain and indicate the direction of projection.

A.3. Measuring the interference contrast

In order to measure the interference contrast independently from estimations of the magnitude of phase fluctuations, we reduced the interferometer to a two-path one by keeping one path closed. Sweeping the phase difference over a range including fully constructive or destructive interference allows to directly measure the interference contrast. By creating a vertical temperature gradient along the triangular path arrangement, only the phase differences between path B (tip of the triangular path arrangement) and path A or C (base of the triangle) can be tuned. We decided to closed path A and sweep the phase difference between path B and C. To tune this temperature gradient we heated the sample to about 34.9 °C and shut off the heater afterwards. We measured all four shutter combinations of path B and C in random order about 70 times, during the thermalization of the setup with its surrounding. To increase the time resolution we only took five successive measurements per combination. Each combination took about 5.71 s including 5.16 s for changing the shutter combination. Each cycle takes about 23.25 s with 0.41 s for logging additional parameters like input power and housing temperature.

The thermalization of the housing temperature T_h can be described by an exponential function
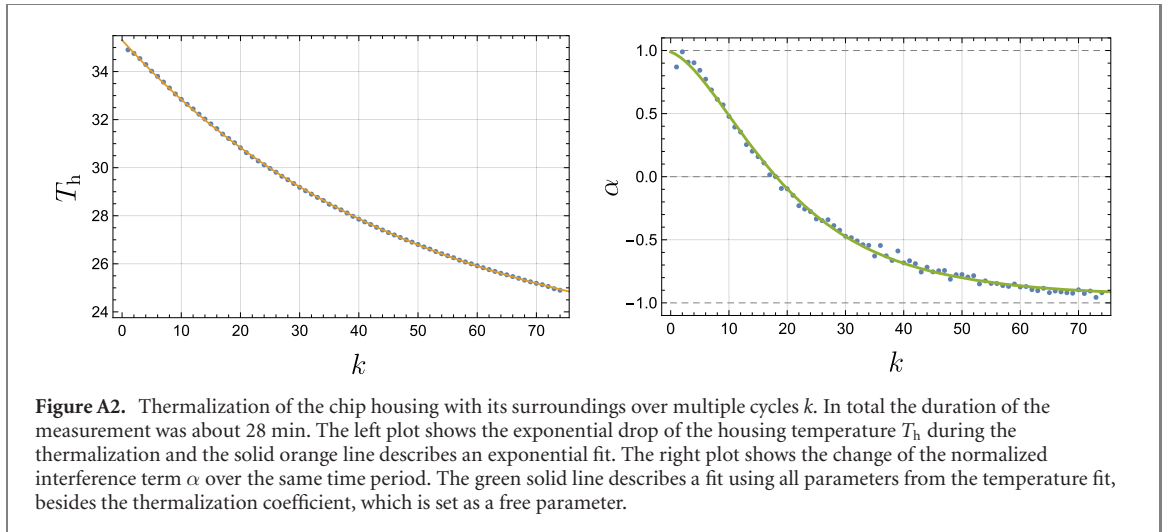
$$T_h = T_0 + \Delta T \exp(-\kappa k), \tag{A.4}$$

where k is the cycle index, T_0 the room temperature, ΔT the range of the temperature change and κ the thermalization coefficient. The recorded temperature is shown together with an exponential fit ($T_0 = 22.04(3)$ C, $\Delta T = 13.269(28)$ C and $\kappa = 0.020\,55(10)$/cycle) in figure A2. This description fits the data quite well and therefore the influence on the normalized interference can be modelled by

$$\alpha = C_\alpha \cos(\Delta\phi_0 + \eta \cdot \Delta T \exp(-\kappa k)), \tag{A.5}$$

where C_α describes the interference contrast, $\Delta\phi_0$ the phase difference at room temperature and η the change of the phase difference per temperature change. This model, with the parameters $\kappa = 0.0415(13)$/cycle, $\Delta\phi_0 = 3.44(7)$ rad, $\eta = 0.203(9)$ rad °C^{-1} and $C_\alpha = 1.000(29)$, fits the measured data very well as shown in figure A2. The fit parameter of C_α shows that the interference contrast of these two paths is, within the experimental precision of this measurement, quite high.

Note that the parameters of the temperature fit were not used for fitting the normalized interference, due the position of the temperature sensor. It is mounted on the outside of the housing, between the insulation and the housing. Therefore the measured housing temperature does not perfectly describe the change of the temperature gradient and with it the change of the phase difference. Also the phase change per temperature η, resulting from the fit of the left plot in figure A2, can not be compared with the one given in section 3, which described a fully thermalized situation and here we have a dynamic thermal process.

Figure A2. Thermalization of the chip housing with its surroundings over multiple cycles k. In total the duration of the measurement was about 28 min. The left plot shows the exponential drop of the housing temperature T_h during the thermalization and the solid orange line describes an exponential fit. The right plot shows the change of the normalized interference term α over the same time period. The green solid line describes a fit using all parameters from the temperature fit, besides the thermalization coefficient, which is set as a free parameter.

A.4. Polarization

The three-dimensional coupling between the waveguides in the fused silica chip leads to polarization rotations. These changes in polarization depend on the orientation of the adjacent waveguides in regards to the orientation of their elliptical waveguide profile. Therefore, the two waveguides at the base of the triangle are affected differently than the one at the tip. This effect leads to a superposition of light of different polarizations in the output ports of the chip. Without a polarizer in front of the detector the combined signal of both orthogonal polarizations is measured. For example if only one path i is open this can be described using the horizontal H and vertical V polarization basis as

$$P_i = P_i^H + P_i^V. \tag{A.6}$$

In case of a two open paths i and j the same principle can be applied leading to

$$P_{ij} = P_i^H + P_j^H + 2\sqrt{P_i^H P_j^H} \cos(\Delta\phi_{ij}^H) + P_i^V + P_j^V + 2\sqrt{P_i^V P_j^V} \cos(\Delta\phi_{ij}^V). \tag{A.7}$$

A polarizer in front of the detector can be set to filter all H-components resulting in $F = 1$, which is trivial to show using (2) and (5)–(7).

Without a polarizer the resulting equation for F is quite long and does not reduce to 1. This can be shown for the case of equal powers for both polarizations $P_i^H = P_i^V = P_i/2$. This reduces P_{ij} to

$$P_{ij} = P_i + P_j + \sqrt{P_i P_j}(\cos(\Delta\phi_{ij}^H) + \cos(\Delta\phi_{ij}^V)). \tag{A.8}$$
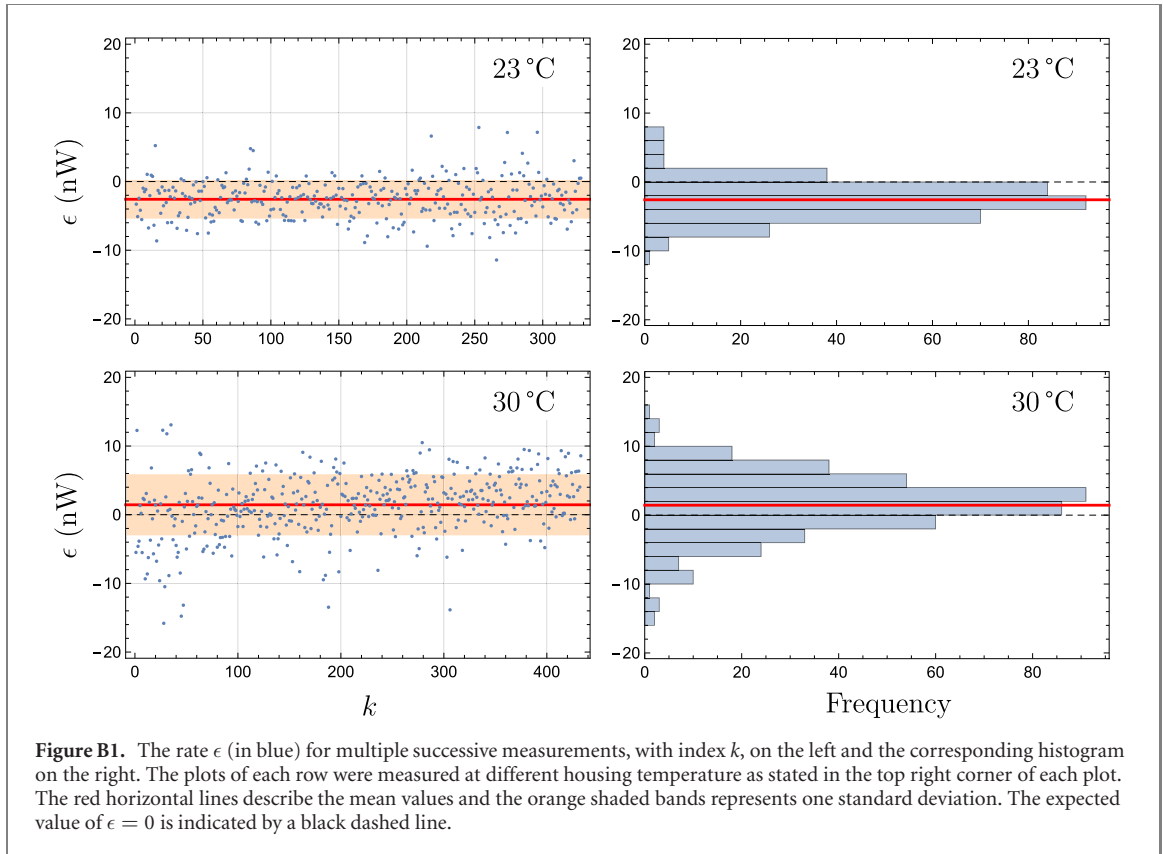
Calculating F for this case results in

$$F = \frac{1}{2}\left(1 + \alpha^H \alpha^V + \beta^H \beta^V + \gamma^H \gamma^V - \frac{1}{2}\sum_{i,j,k\in\{H,V\}} \alpha^i \beta^j \gamma^k + \alpha^H \beta^H \gamma^H + \alpha^V \beta^V \gamma^V\right). \tag{A.9}$$

It is easy to find normalized interference terms α^i, β^i and γ^i, which fulfill $F^i = 1$ for each polarization i, but result in $F \neq 1$. This shows that a measurement of a combination of orthogonal polarizations can lead to a deviation in the Peres parameter. In contrast, by projecting to a single polarization, this effect can be avoided.

Appendix B. Determination of crosstalk via Sorkin test

To determine the phase crosstalk $\Delta\phi_{DH}$ in our experiment we utilize the Sorkin test [12, 42], usually applied to probe the presence of higher-order interferences, which would violate Born's rule [13–20]. We resorted to this test as it is unbiased by most error sources, except by nonlinearity in the setup and crosstalk between the modes. For a three-path interferometer, as used in our experiment, the output rate for all shutters open can be obtained via Born's rule as

$$P_{ABC} = |\psi_A + \psi_B + \psi_C|^2 = P_{AB} + P_{BC} + P_{CA} - P_A - P_B - P_C. \tag{B.1}$$

Figure B1. The rate ϵ (in blue) for multiple successive measurements, with index k, on the left and the corresponding histogram on the right. The plots of each row were measured at different housing temperature as stated in the top right corner of each plot. The red horizontal lines describe the mean values and the orange shaded bands represents one standard deviation. The expected value of $\epsilon = 0$ is indicated by a black dashed line.

Any experimental deviation from this result can be subsumed in the rate ϵ:

$$\epsilon \equiv P_{\mathrm{ABC}} + P_{\mathrm{A}} + P_{\mathrm{B}} + P_{\mathrm{C}} - P_{\mathrm{AB}} - P_{\mathrm{BC}} - P_{\mathrm{CA}}. \tag{B.2}$$

According to Sorkin, deviations from $\epsilon = 0$ indicate higher-order interference and would violate Born's rule. To foster a comparison across different experimental platforms, a normalized version of this rate (called Sorkin parameter) is used. A common variant is the normalization of the mean of ϵ in regards to the mean sum of the absolute values of the pairwise interference terms:

$$\langle \kappa \rangle \equiv \frac{\langle \epsilon \rangle}{\langle |P_{\mathrm{AB}} - P_{\mathrm{A}} - P_{\mathrm{B}}| + |P_{\mathrm{BC}} - P_{\mathrm{B}} - P_{\mathrm{C}}| + |P_{\mathrm{AC}} - P_{\mathrm{A}} - P_{\mathrm{C}}| \rangle}. \tag{B.3}$$

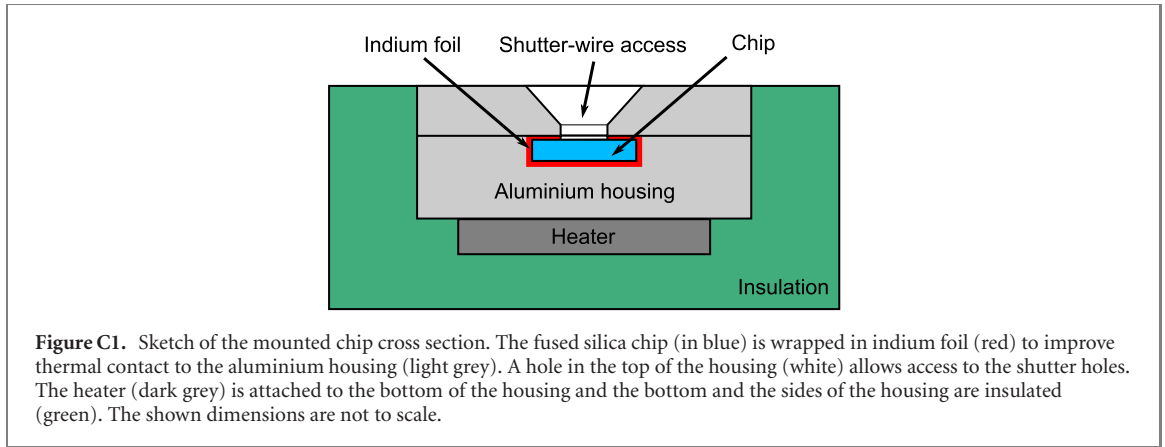This normalization makes κ independent of input power and interference contrast.

Performing the Sorkin test with the same measurement data, supplemented by data of the rate P_{ABC}, resulted in the data as shown in figure B1 and the following mean values for each housing temperature:

$$\langle \epsilon_{23°\mathrm{C}} \rangle = -2.58(16) \text{ nW} \quad \text{and} \quad \langle \epsilon_{30°\mathrm{C}} \rangle = 1.5(5) \text{ nW} \tag{B.4}$$

$$\langle \kappa_{23°\mathrm{C}} \rangle = -14.0(8) \times 10^{-4} \quad \text{and} \quad \langle \kappa_{30°\mathrm{C}} \rangle = 11(4) \times 10^{-2}. \tag{B.5}$$

The given error is the standard error of mean corrected by the autocorrelation according to [36]. Using this method for the error calculation, both housing temperatures show a significant deviation from the expected value $\epsilon = \kappa = 0$. When comparing these values with the literature one has to keep in mind, that this test is usually performed with an interferometer tuned to fully constructive or destructive interference, where phase deviations have smaller effects. As this is not the case in our experiment, the systematic deviations from 0 are larger than they would be at the extremal points.

As higher-order interferences have been experimentally ruled out in other experiments for coherent light illumination to at least $|\kappa| < 3 \times 10^{-5}$ [17], any apparent higher-order interference appearing in the present experiment can be attributed to nonlinearity or crosstalk. We determined the nonlinearity to be negligible by independent measurements (see section 5.1), such that effectively the Sorkin test turns into a measure of crosstalk. We can therefore associate systematic deviations from the predicted $\epsilon = 0$ to cross-channel influences of shutter settings between the waveguides (see discussion in section 5.4). To calculate the phase crosstalk $\Delta \phi_{\mathrm{DH}}$ based on the measured ϵ, we used the normalized interference terms modified by the phase

Figure C1. Sketch of the mounted chip cross section. The fused silica chip (in blue) is wrapped in indium foil (red) to improve thermal contact to the aluminium housing (light grey). A hole in the top of the housing (white) allows access to the shutter holes. The heater (dark grey) is attached to the bottom of the housing and the bottom and the sides of the housing are insulated (green). The shown dimensions are not to scale.

crosstalk $\Delta\phi_{\mathrm{DH}}$ (as described in (16)) and inserted them into the corresponding two-path rates P_{ij} in (B.2). This results in the following equation connecting the phase crosstalk $\Delta\phi_{\mathrm{DH}}$ with the rate ϵ:

$$\epsilon = 2P_{\mathrm{in}} \left(\sqrt{T_A T_B} \left(\gamma - \cos(\Delta\phi_{\mathrm{AB}} - \Delta\phi_{\mathrm{DH}}) \right) + \sqrt{T_B T_C} \left(\alpha - \cos(\Delta\phi_{\mathrm{BC}} + \Delta\phi_{\mathrm{DH}}) \right) \right). \qquad \text{(B.6)}$$

For each temperature, this equation leads to two solutions for $\Delta\phi_{\mathrm{DH}}$, which differ by one order of magnitude. We chose the solution with the lower crosstalk strength (given in table 3), as we expect it to be a small effect.

Appendix C. Setup details

Figure C1 shows the cross-section of the mounted chip in detail.

ORCID iDs

S Gstir ⓘ https://orcid.org/0000-0002-9096-4895
A Szameit ⓘ https://orcid.org/0000-0003-0071-6941
R Keil ⓘ https://orcid.org/0000-0003-1198-9961
G Weihs ⓘ https://orcid.org/0000-0003-2260-3008

References

[1] Bell J S 1964 *Phys. Phys. Fiz.* **1** 195–200
[2] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880–4
[3] Aspect A, Dalibard J and Roger G 1982 *Phys. Rev. Lett.* **49** 1804–7
[4] Hensen B *et al* 2015 *Nature* **526** 682–6
[5] Giustina M *et al* 2015 *Phys. Rev. Lett.* **115** 250401
[6] Shalm L K *et al* 2015 *Phys. Rev. Lett.* **115** 250402
[7] Rauch D *et al* 2018 *Phys. Rev. Lett.* **121** 080403
[8] Li M H *et al* 2018 *Phys. Rev. Lett.* **121** 080404
[9] Poh H S, Joshi S K, Cerè A, Cabello A and Kurtsiefer C 2015 *Phys. Rev. Lett.* **115** 180408
[10] Meraner S, Chapman R J, Frick S, Keil R, Prilmüller M and Weihs G 2021 *SciPost Phys.* **10** 17
[11] Born M 1926 *Z. Phys.* **37** 863–7
[12] Sorkin R D 1994 *Mod. Phys. Lett.* A **9** 3119–27
[13] Sinha U, Couteau C, Jennewein T, Laflamme R and Weihs G 2010 *Science* **329** 418–21
[14] Hickmann J M, Fonseca E J S and Jesus-Silva a J 2011 *Europhys. Lett.* **96** 64006
[15] Park D K, Moussa O and Laflamme R 2012 *New J. Phys.* **14** 113025
[16] Jin F *et al* 2017 *Phys. Rev.* A **95** 012107
[17] Kauten T, Keil R, Kaufmann T, Pressl B, Brukner Č and Weihs G 2017 *New J. Phys.* **19** 033017
[18] Cotter J P, Brand C, Knobloch C, Lilach Y, Cheshnovsky O and Arndt M 2017 *Sci. Adv.* **3** e1602478
[19] Barnea A R, Cheshnovsky O and Even U 2018 *Phys. Rev.* A **97** 023601
[20] Pleinert M O, Rueda A, Lutz E and von Zanthier J 2020 Testing quantum theory with higher-order interference in many-particle correlations (arXiv:2006.09496)
[21] Birkhoff G and Neumann J V 1936 *Ann. Math.* **37** 823–43
[22] Stueckelberg E C G 1960 *Helv. Phys. Acta* **33** 727–52
[23] Adler S L 1995 *Quaternionic Quantum Mechanics and Quantum Fields* (Oxford: Oxford University Press)
[24] Peres A 1979 *Phys. Rev. Lett.* **42** 683–6
[25] Kaiser H, George E A and Werner S A 1984 *Phys. Rev.* A **29** 2276–9
[26] Procopio L M, Rozema L A, Wong Z J, Hamel D R, O'Brien K, Zhang X, Dakić B and Walther P 2017 *Nat. Commun.* **8** 15044

[27] Keil R, Kaufmann T, Kauten T, Gstir S, Dittel C, Heilmann R, Szameit A and Weihs G 2016 *APL Photon.* **1** 081302

[28] Szameit A and Nolte S 2010 *J. Phys. B: At. Mol. Opt. Phys.* **43** 163001

[29] Meany T, Gräfe M, Heilmann R, Perez-Leija A, Gross S, Steel M J, Withford M J and Szameit A 2015 *Laser Photon. Rev.* **9** 363–84

[30] Chaboyer Z, Meany T, Helt L G, Withford M J and Steel M J 2015 *Sci. Rep.* **5** 9601

[31] Spagnolo N, Vitelli C, Aparo L, Mataloni P, Sciarrino F, Crespi A, Ramponi R and Osellame R 2013 *Nat. Commun.* **4** 1606

[32] Marcinkevičius A, Juodkazis S, Watanabe M, Miwa M, Matsuo S, Misawa H and Nishii J 2001 *Opt. Lett.* **26** 277–9

[33] Hnatovsky C, Taylor R S, Simova E, Rajeev P P, Rayner D M, Bhardwaj V R and Corkum P B 2006 *Appl. Phys.* A **84** 47–61

[34] Crespi A, Lobino M, Matthews J C F, Politi A, Neal C R, Ramponi R, Osellame R and O'Brien J L 2012 *Appl. Phys. Lett.* **100** 233704

[35] Heilmann R, Gräfe M, Nolte S and Szameit A 2014 *Sci. Rep.* **4** 4118

[36] Zhang N F 2006 *Metrologia* **43** S276–81

[37] Rozema L A, Zhuo Z, Paterek T and Dakić B 2020 Higher-order interference in quantum physics (arXiv:2003.11048)

[38] Kauten T, Pressl B, Kaufmann T and Weihs G 2014 *Rev. Sci. Instrum.* **85** 063102

[39] Flamini F *et al* 2015 *Light Sci. Appl.* **4** e354

[40] Chaboyer Z, Stokes A, Downes J, Steel M J and Withford M J 2017 *Opt. Express* **25** 33056–65

[41] Dyakonov I V, Pogorelov I A, Bobrov I B, Kalinkin A A, Straupe S S, Kulik S P, Dyakonov P V and Evlashin S A 2018 *Phys. Rev. Appl.* **10** 044048

[42] Sorkin R D 1997 Quantum measure theory and its interpretation *Quantum Classical Correspondence: Proc. 4th Drexel Symp. Quantum Nonintegrability* (Cambridge, MA: International Press) pp 229–51

[43] Gstir S 2021 Datasets for 'towards probing for hypercomplex quantum mechanics in a waveguide interferometer' https://doi.org/10.5281/zenodo.5497283