Confined Contextuality in Neutron Interferometry: Observing the Quantum Pigeonhole Effect

Mordecai Waegell,1,2 Tobias Denkmayr,3 Hermann Geppert,3 David Ebner,1 Tobias Jenke,4 Yuji Hasegawa,3 Stephan Sponar,3 Justin Dressel,1,2 and Jeff Tollaksen1,2

Institute for Quantum Studies, Chapman University, Orange, CA 92866, USA
2Schmid College of Science and Technology, Chapman University, Orange, CA 92866, USA
3Atom Institut, TU Wien, Stadionallee 2, 1020 Vienna, Austria
4Institut Laue-Langevin 6, Rue Jules Horowitz, 38042 Grenoble Cedex 9, France

Previous experimental tests of quantum contextuality based on the Bell-Kochen-Specker (BKS) theorem have demonstrated that not all observables among a given set can be assigned noncontextual eigenvalue predictions, but have never identified which specific observables must fail such assignment. Using neutron interferometry, we remedy this shortcoming by showing that BKS contextuality can be confined to particular observables in the form of anomalous weak values, which can be directly witnessed through weak measurements. We construct a confined contextuality witness from weak values, which we measure experimentally to obtain a 5σ average violation of the noncontextual bound, with one contributing term violating an independent bound by more than 99σ. This experimentally measured confined contextuality confirms the quantum pigeonhole effect, wherein eigenvalue assignments to contextual observables apparently violate the classical pigeonhole principle.

Introduction.— Quantum contextuality, as introduced by Bell, Kochen and Specker (BKS) [1,2], forbids all observable properties of a system from being predefined independently from how they are observed. This phenomenon is one of the most counterintuitive aspects of quantum mechanics, and finds itself at the heart of recent quantum information processing applications [3–8]. The BKS theorem is proved by a BKS-set of observables [9,10], which contains geometrically related and mutually commuting subsets (or measurement contexts) that result in a logical incompatibility: Attempting to pre-assign eigenvalues globally to the entire BKS-set (i.e., noncontextually) results in a contradiction with the predictions of quantum mechanics. Previous contextuality experiments [11,13] have demonstrated that at least one observable property in a special BKS-set of possible observations cannot be predefined without the values depending on the choice of measurement context; however, these experiments have never isolated exactly where this contextuality occurs.

In this Letter, using recently developed weak measurement techniques in neutron interferometry [14–20], and a generalization of BKS-contextuality to weak measurements [21], we experimentally demonstrate that we can isolate which measurement context within a BKS-set must produce the contradiction, essentially confining the contextuality [22]. Leveraging recent results [23,27] that link contextuality to weak values [28], we construct and measure contextuality witnesses composed of weak values. The preparations, weak measurements, and postselections for each weak value are in contexts that belong to the BKS-set. These boundary conditions confine the contradiction to specific observables, in the form of anomalous weak values that lie outside their spectral range.

In our experiment, we witnessed the BKS-contextuality of neutron spins. We measured the spin by performing path-dependent spin rotations, making the path a weakly-coupled meter for the spin; conditioning the path measurements on spin postselections then reveals the desired weak values [19]. We measured seventeen independent ensembles of neutron spins, and tested two different types of weak-value-based contextuality-witnesses [23,26,28] for each odd number of spin ensembles from 3 to 17. These witnesses violated their noncontextuality bounds by 99σ, showing that the contextuality was indeed confined. One particular 5-spin weak value exceeded its bound by more than 99σ.

Confining BKS contextuality.—BKS contextuality confinement follows from the Aharonov-Bergmann-Lebowitz (ABL) formula [29], which gives the probability of obtaining a particular strong measurement outcome i, between a preparation |ψ⟩ and a postselection ⟨φ|. The ABL formula can be expressed in terms of weak values ⟨Π⟩w = ⟨φ|Π|ψ⟩/⟨φ|ψ⟩ [22].

$$P_{ABL}(\Pi_i = 1 | \psi, \phi, B) = \frac{|\langle \Pi_i \rangle^w|^2}{\sum_{k \in B} |\langle \Pi_k \rangle^w|^2}. \quad (1)$$

As shown in Ref. [24], the ABL formula constrains any noncontextual hidden variable theory (NCHVT) since a projection with an ABL probability of 1 must also be assigned a value of 1 in any NCHVT. It then follows from

$$\sum_{k \in B} |\langle \Pi_k \rangle^w| = 1 \quad \implies \quad P_{ABL}(\Pi_i = 1 | \psi, \phi, B) = 1$$

Furthermore, if B contains only two outcomes, then the converse also follows: ⟨Π⟩w = 1 implies

$$P_{ABL}(\Pi_i = 1 | \psi, \phi, B) = 1.$$
For our specific case of \( N > 2 \) spins, consider a product of any two spin operators \( ZZ \), with spectral decomposition \( ZZ = (+1)\Pi_{\text{even}}^{} + (-1)\Pi_{\text{odd}}^{} \) in terms of the rank-2 parity projectors, \[ \Pi_{\text{even}} = \Pi_+ \otimes \Pi_+ + \Pi_- \otimes \Pi_- \quad \Pi_{\text{odd}} = \Pi_+ \otimes \Pi_- + \Pi_- \otimes \Pi_+ \] with \( \Pi_\pm = |\pm \rangle \langle \pm | = (1 \pm Z)/2 \). Given \( |\psi \rangle \) and \( |\phi \rangle \) defined above, \( (\Pi_{\text{even}}^\psi)_{w} = 0 \) and \( (\Pi_{\text{odd}}^\phi)_{w} = 1 \), and thus \( (ZZ)_{w} = -1 \). The ABL rule in Eq. (1) then implies \( ZZ = -1 \) for all pairs of spins in any NCHVT, as illustrated in Fig. 1.

This pairwise constraint is the quantum pigeonhole effect [30]. To see this, let the spin eigenstates \( |\pm \rangle \) correspond to two boxes in which pigeons may be placed. The projectors in Eq. (2) describe definite numbers of pigeons in each box, up to an exchange of boxes; i.e., \( \Pi_{\text{even}}^{} \) denotes two pigeons in one box, while \( \Pi_{\text{odd}}^{} \) denotes one pigeon in each. The pigeonhole principle states that if \( N > 2 \) pigeons are placed in two boxes, then at least one box must contain multiple pigeons. However, the constraint \( ZZ = -1 \) for all pairs implies that, regardless of how many pigeons are placed in the two boxes, no two pigeons are ever in the same box!

Witnessing BKS contextuality.—Following the pigeon analogy, all NCHVT assignments of definite numbers of pigeons to each box are forbidden. The projectors corresponding to such forbidden assignments for \( N = 3 \) are

\[ \Pi_{1}^{(3)} = \Pi_+ \otimes \Pi_+ \otimes \Pi_+ + \Pi_- \otimes \Pi_- \otimes \Pi_- \]
\[ \Pi_{2}^{(3)} = \Pi_+ \otimes \Pi_- \otimes \Pi_+ + \Pi_- \otimes \Pi_+ \otimes \Pi_- \]
\[ \Pi_{3}^{(3)} = \Pi_- \otimes \Pi_+ \otimes \Pi_+ + \Pi_+ \otimes \Pi_- \otimes \Pi_- \]
\[ \Pi_{4}^{(3)} = \Pi_- \otimes \Pi_- \otimes \Pi_+ + \Pi_+ \otimes \Pi_+ \otimes \Pi_- \]

which are the invariant eigenspaces of the first row of Fig. 1. (\( \Pi_{1}^{(3)} \) indicates all three pigeons in one box, while \( \Pi_{2,3,4}^{(3)} \) are the permutations of two and one). We call complete sets of forbidden projectors like this contextual bases for brevity.

Crucially, such forbidden BKS value assignments manifest as anomalous projector weak values (with real part outside the range \([0, 1]\) in the contextual basis of Eq. (3) [22]. This appearance of anomalous weak values directly connects the ABL-confined BKS contextuality to a generalized form of contextuality that admits weak measurements [21][23]. Thus, measuring anomalous weak values in a contextual basis constitutes a direct witness of confined BKS contextuality.

As an explicit example, the weak value of \( \Pi_{1}^{(3)} \) is

\[ (\Pi_{1}^{(3)}{w}) = \prod_{n=1}^{3} \frac{1 + Z_{w}}{2} + \prod_{n=1}^{3} \frac{1 - Z_{w}}{2} = -\frac{1}{2} \]  \[ (4) \]

with \( |\psi \rangle \) and \( |\phi \rangle \) above, where each \( n \) is a distinct spin, and \( Z_{w} = (|+Y| Z |+X\rangle /|+Y\rangle |+X\rangle \) is a purely imaginary single-spin weak value. This example also illustrates a subtle point about the connection between anomalous weak values and contextuality. A projector weak value for a separable composite system with a negative real part is a witness of contextuality [23]. However, the projector weak values in each term above are \( (\Pi_{1}^{\pm})_{w} = (1 \pm Z_{w})/2 = e^{\pm i\pi/4}/\sqrt{2} \), which have positive real parts. Nevertheless, the product of three such weak values, \( \Pi_{12}^{(3)}{w} = (\Pi_{1}^{+})_{w} (\Pi_{2}^{+})_{w} = e^{\pm i\pi/4}/\sqrt{8} \), has a negative real part, enabling the contextuality witness. In this sense, the observation of a non-zero complex phase for a projector weak value on a single system already implies a witness for contextuality on a larger composite system.

The logic for the above construction for \( N = 3 \) generalizes to odd \( N > 3 \) (see the family of Wheel BKS-sets [10]). We first define the ordered set \( x_{i}^{(N)} \) of all \( i \in 1 \ldots 2^{N-1} \) binary sequences of \( N - 1 \) digits, with a 0 appended to the end of each sequence, e.g., \( x_{1}^{(3)} = (0, 0, 0) \), \( x_{2}^{(3)} = (0, 1, 0) \), etc. The weak values for the contextual basis of \( N \) spins are then

\[ (\Pi_{i}^{(N)})_{w} = \prod_{n=1}^{N} \frac{1 + (-1)^{x_{i,n}^{(N)}} Z_{w}}{2} + \prod_{n=1}^{N} \frac{1 - (-1)^{x_{i,n}^{(N)}} Z_{w}}{2} \]  \[ (5) \]

where \( x_{i,n}^{(N)} \) is the \( n \)th digit of \( x_{i}^{(N)} \). As in Eq. (4), these projector weak values are determined by the single-spin \( Z_{w} \). This great simplification enables us to construct all forbidden projectors for any number of spins by measuring single-spin \( Z_{w} \) from independent ensembles of experimental runs. All projector weak values then evaluate to \( (\Pi_{i}^{(N)})_{w} = \pm 2^{-(N-1)/2} \), with a sign depending on the index.

Finally, we construct an unbiased contextuality witness \( C_{(N)} \) using all \( 2^{N-1} \) rank-2 projectors in an \( N \)-spin contextual basis that aggregates the contextuality of the entire basis,

\[ C_{(N)} = \text{Re} \left[ I - \sum_{i=1}^{2^{N-1}} s_{i} \Pi_{i}^{(N)} \right], \]  \[ (6) \]

with \( s_{i} = \text{sign} \text{Re}(\Pi_{i}^{(N)})_{w} \). Regardless of the signs \( s_{i} \), if all \( 0 \leq \text{Re}(\Pi_{i}^{(N)})_{w} \leq 1 \), then \( C_{(w)} \geq 0 \). Observing \( C_{(w)} < 0 \) is
thus an experimental witness of confined BKS-contextuality. This choice of the signs $s_i$ optimizes $C_w^{(N)}$ by accumulating anomalous parts of the weak values (below 0 or above 1), producing the ideal values $C_w^{(N)} = 1 - 2(N-1)/2$.

Experimental procedure.—In our experiment, we measure the weak value $Z_w$ of the neutron spin in the $z$-direction using an interferometer. The neutron’s path degree of freedom (DOF) is used as a pointer to measure both the real and imaginary parts of $Z_w$. This approach has already been successfully used to completely determine weak values of massive systems [19]. The experiment was conducted at the instrument S18 at the high flux reactor reactor of the Institute Laue-Langevin (ILL) in Grenoble, France. A graphical illustration of the experimental setup is depicted in Fig. 2.

A silicon channel-cut perfect crystal selects neutrons with a wavelength of $\lambda_0 = 1.91 \text{ Å} (\lambda/\lambda_0 \approx 0.02)$ by a 3-fold [220] Bragg reflection from a white neutron beam. Between the monochromator and the interferometer crystal, two magnetically birefringent prisms split the unpolarized beam in two subbeams, one with the neutron spin aligned parallel to the positive $z$-direction and one aligned antiparallel. Even though the angular separation is just four seconds of arc, only the beam with spin up component fulfills Bragg’s condition at the interferometer’s first plate. The degree of polarization is above 99% with the neutron spin state given by $|+Z\rangle$. The other beam passes through unaffected and does not further contribute to the experiment.

A DC coil in front of the interferometer generates a constant magnetic field $B_y$ in $y$-direction. After entering the coil, the neutron experiences a non-adiabatic field change and its spin starts to precess around $B_y$. If the magnetic field magnitude is adjusted accordingly, the neutron spin will turn by exactly $\pi/2$ in the coil. This changes the initial spin state from $|+Z\rangle$ to $|+X\rangle$. The first DC coil completes the spin preselection.

At the first plate of the interferometer, the beam is coherently split by amplitude division. In each beam one small coil in a Helmholtz configuration produces a weak magnetic field in the $±z$ direction. To prevent thermal stress on the interferometer, the coils are water cooled. The weak magnetic fields lead to path-dependent spin rotations around the field axis in the $xy$-plane, causing (weak) entanglement between the spin and path DOF of each neutron. For all measurements, the angle of rotation was set to $\alpha = 15^\circ$. Since an angle of 90° would be a strong measurement, this $\alpha$ corresponds to a relative interaction strength of $\sim 16\%$.

Between the second and final interferometer plate, a sapphire phase shifter is inserted. A phase shifter in combination with a Cadmium beam block mounted on a rotational stage provides full control over the neutron’s path DOF for the pointer readout. The phase shifter can change the path qubit state in the equatorial plane of the Bloch sphere, while the beam block permits access to the path eigenstates at the poles.

At the final plate of the interferometer, the two beams are recombined. A second DC coil in combination with a CoTi supermirror array enables postselection on arbitrary spin states. Of the two outgoing beams of the interferometer, only the forward direction (O-detector) is spin analyzed to postselect $|+Y\rangle$. The neutrons are detected by $^3\text{He}$ counter tubes.

All measurements were performed in the IN/OUT style. That is, the spin rotation is turned on, then the intensity is recorded for a fixed phase shifter position. The spin rotation is then turned off, and the intensity is again measured for the same phase shifter position. This makes it possible to determine the phase $\chi$ of the empty interferometer. Also, this method allows us to maximally reduce the influences of a phase drift of the interferogram (due to unavoidable instability of the apparatus) during the determination of the ensemble’s weak value. With every phase shifter scan, all states that lie on the equator of the path qubit’s Bloch sphere are evaluated. Therefore it is possible to extract all necessary data to obtain the real part of the spin weak value from one such phase shifter scan. To gain information about the imaginary part, it is necessary to access the neutron path eigenstates. Experimentally this is done by blocking one beam path at a time. If an intensity is recorded while path $II$ is blocked, a postselected measurement on the state $|I\rangle$ is performed, and vice versa. One measurement set consists of two interferograms (IN/OUT) to determine $\text{Re} [Z_w]$ and two single intensities to get $\text{Im} [Z_w]$. In addition, background measurements are performed, which are later subtracted from the recorded data. The purpose of the OUT measurement, i.e., the one without a spin rotation inside the interferometer, is to find the phase values of $\chi = \pi/2$ and $\chi = 3\pi/2$ of the empty interferometer. In order to calculate the intensities corresponding to the states $(|I\rangle \pm |II\rangle)/\sqrt{2}$, the phase values obtained from the OUT measurements are
inserted into the IN measurement’s fit functions. The weak value’s real part causes a phase shift in the pointer system. Since it is zero for our particular choice of pre and post selected states, no phase shift between the IN and OUT interferogram is observed. In contrast to that, the weak value’s imaginary part shifts the pointer state away from the equatorial plane towards the poles of the Bloch sphere. Physically this results in the fact that the relative amplitudes of the path eigenstates are changed.

In the experiment the weak values $Z_w$ of 17 individual ensembles were determined with high precision (see Supplementary Material). To extract $Z_w$ for one ensemble of neutrons, two $\chi$-scans were recorded, as well as two single intensities. Together with the required background measurements a total measurement time of $\sim 10000$ seconds was necessary to determine the real and imaginary part of each $Z_w$.

**Results.**—Figure 3 shows the final results that violate the noncontextuality bound $C^{(N)}_w \geq 0$. Using the recorded data for each $Z_w$, these unbiased contextuality witness $C^{(N)}_w$ were calculated using Eq. (5) for all odd numbers of spins from 3 to 17. Representative projectors for odd $N$ are also shown (upper left), which make equal use of the IN and OUT interferograms. In contrast to that, the weak value’s real part causes a phase shift in the pointer system. The weak values $\hat{\Pi}^{(5)}_w$ and $\hat{\Pi}^{(9)}_w$ violate the classical boundary by $\sim 7\sigma$, and $\hat{\Pi}^{(5)}_w = -2.85 \pm 0.41$, $\hat{\Pi}^{(9)}_w = -0.2508 \pm 0.0025$, violating the bound by $\sim 99\sigma$.

**Concluding remarks.**—We have experimentally demonstrated the confinement of contextuality within a BKS-set of observables to a particular measurement context, using modern techniques in neutron interferometry. This confinement explicitly connects three previously distinct concepts: BKS-contextuality [2], the ABL rule [29], and the equivalence between anomalous projector weak values and a generalized notion of contextuality [21]. This connection has allowed us to witness the contextuality for BKS-sets of odd numbers of neutron spins between three and seventeen. Notably, this witness does not require entangled preparations or measurements, or indeed any interaction between the different spins, unlike existing approaches to demonstrating BKS-contextuality [11]. The entangled measurement contexts that would usually be required have values that are forced by the ABL rule according to the geometry of the BKS-set itself, so they need not be measured. In this way, confining the contextuality serves to simplify its experimental observation.

The confinement of contextuality to a particular measurement context that we demonstrated here not only raises interesting foundational questions [30], but may also suggest future quantum information processing applications [3, 6].

This research was supported (in part) by the FWF and No. P24973-N20. T.D., H.G., D.E., T.J., S.S., and Y.H. per-
formed the experiment and analyzed the data. T.D., M.W. and J.D. performed the error analysis and generated figures. T.D., M.W., and J.D. co-wrote the manuscript. M.W. and J.T. developed the original theory.

[1] J. S. Bell, Rev. Mod. Phys. 38, 447 (1966).
[2] S. Kochen and E. Specker, J. of Math. and Mech. 17, 59 (1967).
[3] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature 510, 351 (2014).
[4] H. Bechmann-Pasquinucci and A. Peres, Phys. Rev. Lett. 85, 3313 (2000).
[5] A. Cabello, V. D’Ambrosio, E. Nagali, and F. Sciarrino, Phys. Rev. A 84, 030302 (2011).
[6] A. A. Abbott, C. S. Calude, J. Conder, and K. Svozil, Phys. Rev. A 86, 062109 (2012).
[7] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Phys. Rev. Lett. 102, 010401 (2009).
[8] M. Waegell, Phys. Rev. A 89, 012321 (2014).
[9] M. Waegell and P. K. Aravind, J. Phys. A: Math. and Theor. 45, 405301 (2012).
[10] M. Waegell and P. K. Aravind, Phys. Rev. A 88, 012102 (2013).
[11] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C. F. Roos, Nature 460, 494 (2009).
[12] H. Bartosik, J. Klepp, C. Schmitzer, S. Sponar, A. Cabello, H. Rauch, and Y. Hasegawa, Phys. Rev. Lett. 103, 040403 (2009).
[13] V. D’Ambrosio, I. Herbauts, E. Amselem, E. Nagali, M. Bourennane, F. Sciarrino, and A. Cabello, Phys. Rev. X 3, 011012 (2013).
[14] H. Rauch and S. Werner, “Neutron Interferometry, Clarendon,” (2000).
[15] H. Rauch, H. Lemmel, M. Baron, and R. Loidl, Nature 417, 630 (2002).
[16] Y. Hasegawa, R. Loidl, G. Badurek, M. Baron, and H. Rauch, Nature 425, 45 (2003).
[17] J. Klepp, S. Sponar, and Y. Hasegawa, PTEP 2014, 082A01 (2014).
[18] T. Denkmayr, H. Geppert, S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa, Nat. Commun. 5 (2014).
[19] S. Sponar, T. Denkmayr, H. Geppert, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa, Phys. Rev. A 92, 062121 (2015).
[20] T. Denkmayr, H. Geppert, H. Lemmel, M. Waegell, J. Dressel, Y. Hasegawa, and S. Sponar, preprint arXiv:1604.04102 (2016).
[21] R. W. Spekkens, Phys. Rev. A 71, 052108 (2005).
[22] M. Waegell and J. Tollaksen, preprint arXiv:1505.00098 (2015).
[23] M. F. Pusey, Phys. Rev. Lett. 113, 200401 (2014).
[24] M. F. Pusey and M. S. Leifer, EPTCS 195, 295 (2015).
[25] M. S. Leifer and R. W. Spekkens, Phys. Rev. Lett. 95, 200405 (2005).
[26] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nat. Commun. 7, 11780 (2016).
[27] F. Piacentini, A. Avella, M. P. Levi, R. Lassana, F. Villa, A. Tosi, F. Zappa, M. Gramegna, G. Brida, I. P. Degiovanni, and M. Genovese, Phys. Rev. Lett. 116, 180401 (2016).
[28] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. 60, 1351 (1988).
[29] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Phys. Rev. 134, B1410 (1964).
[30] Y. Aharonov, F. Colombo, S. Popescu, I. Sabadini, D. C. Struppa, and J. Tollaksen, PNAS 113, 532 (2016).
Table I. Experimentally determined weak values for 17 different ensembles.

| Set Number | Re $[Z_w]$   | Im $[Z_w]$   |
|------------|--------------|--------------|
| # 1        | $-0.024 \pm 0.044$ | $0.970 \pm 0.094$ |
| # 2        | $-0.005 \pm 0.044$ | $1.050 \pm 0.098$ |
| # 3        | $0.018 \pm 0.045$ | $1.002 \pm 0.095$ |
| # 4        | $-0.103 \pm 0.045$ | $0.925 \pm 0.092$ |
| # 5        | $-0.032 \pm 0.044$ | $0.979 \pm 0.094$ |
| # 6        | $-0.097 \pm 0.045$ | $1.024 \pm 0.096$ |
| # 7        | $0.002 \pm 0.049$ | $0.912 \pm 0.099$ |
| # 8        | $-0.041 \pm 0.049$ | $0.985 \pm 0.102$ |
| # 9        | $0.101 \pm 0.051$ | $0.920 \pm 0.099$ |
| # 10       | $-0.020 \pm 0.050$ | $0.931 \pm 0.099$ |
| # 11       | $0.022 \pm 0.050$ | $1.037 \pm 0.104$ |
| # 12       | $-0.070 \pm 0.049$ | $0.874 \pm 0.095$ |
| # 13       | $0.011 \pm 0.049$ | $0.790 \pm 0.092$ |
| # 14       | $0.062 \pm 0.050$ | $0.910 \pm 0.096$ |
| # 15       | $-0.084 \pm 0.051$ | $1.039 \pm 0.105$ |
| # 16       | $-0.121 \pm 0.052$ | $0.973 \pm 0.103$ |
| # 17       | $-0.079 \pm 0.050$ | $1.003 \pm 0.103$ |

Measurement Procedure.

To determine the weak value $Z_w$, three interference fringes are recorded:

1. The OUT curve without any interaction, to evaluate the phase of the empty interferogram.
2. The IN curve with a weak spin rotation of $\alpha = 15^\circ$ in each of the interferometer’s arms, which yields $I_{y \pm}$.
3. One interference fringe with orthogonal pre and postselected spin states, which is then subtracted from the IN/OUT curve as an effective background.

Additionally, two single intensities with one or the other beam blocked are recorded ($I_{z \pm}$). Also for those two intensities a background measurement with orthogonal pre and postselected spin states is performed. The recorded background intensities are then subtracted from the signal. Figure 4 shows a typical IN and OUT curve of one experimental run.

The measurement procedure was repeated until altogether 17 different ensembles were recorded. For each ensemble the weak value is extracted. The results are listed in Table I.

It is also noteworthy that the errors of set 1 to 6 are smaller than the others. This is due to a change in reactor power. While the first six interferograms were recorded at a power of $\sim 58$ MW, for the last eleven a power of $\sim 43$ MW was available. The increase in reactor power leads to an increase in neutron flux. A higher count rate offers better statistics and reduces the uncertainty of the recorded values. We used the first N spin weak values in Table I for our analysis of $N$-qubit contextuality witnesses in the main text.

Supplementary Material.—To determine the weak value of the Pauli spin operator $Z$ the spin degree of freedom is weakly coupled to the path degree of freedom [19]. As described in the main body of the paper the weak value’s real part is then inferred from an interference fringe, while two single intensity measurements are necessary to determine the weak value’s imaginary part.

Since the weak value’s real part is zero no phase shift is seen between the IN and the OUT curve. $I_{y \pm}$ are obtained by two single intensity measurements. Background has already been subtracted.

Figure 4. (Color Online) Obtained interferogram for one ensemble: Since the weak value’s real part is zero no phase shift is seen between the IN and the OUT curve. $I_{z \pm}$ are obtained by two single intensity measurements. Background has already been subtracted.