Supplementary Information: Observing High-Dimensional Nonlocality using Multi-Outcome Spectral Measurements

Kiki Dekkers,[1] Laura Serino,[2] Nicola D'Alessandro,[3] Abhinandan Bhattacharjee,[2]
Benjamin Brecht,[2] Armin Tavakoli,[3] Christine Silberhorn,[2] and Jonathan Leach[1,*]

[1] *School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK*
[2] *Paderborn University, Integrated Quantum Optics, Institute for Photonic Quantum Systems (PhoQS), Paderborn, Germany*
[3] *Physics Department and NanoLund, Lund University, Box 118, 22100 Lund, Sweden*
[*] *j.leach@hw.ac.uk*

CGLMP INEQUALITY

The CGLMP Bell inequality [1] is a facet of the local polytope in the scenario in which Alice and Bob have binary inputs $x, y \in \{1, 2\}$ and respective outcomes $a, b \in \{0, \ldots, d-1\}$, for some integer $d \geq 2$. For any local hidden variable theory, it holds that

$$I_d = \sum_{k=0}^{\lfloor d/2 \rfloor - 1} \left(1 - \frac{2k}{d-1}\right) \big[P(a_1 = b_1 + k) + P(b_1 = a_2 + k + 1) + P(a_2 = b_2 + k) + P(b_2 = a_1 + k) \\ - P(a_1 = b_1 - k - 1) - P(b_1 = a_2 - k) - P(a_2 = b_2 - k - 1) - P(b_2 = a_1 - k - 1)\big] \leq 2, \tag{S1}$$

where we have defined

$$P(a_x = b_y + z) = \sum_{a,b=0}^{d-1} P(a, b|x, y) \delta_{a_x,\ b_y + z \bmod d}, \tag{S2}$$

and

$$P(b_y = a_x + z) = \sum_{a,b=0}^{d-1} P(a, b|x, y) \delta_{a_x + z \bmod d,\ b_y}. \tag{S3}$$

For the special case of $d = 2$ the CGLMP inequality reduces to the CHSH inequality.

To violate the CGLMP inequality, Alice and Bob can share the maximally entangled state

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle_A |j\rangle_B , \tag{S4}$$

The best-known measurement selections correspond to the following bases

$$|K_{k|x}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(i \frac{2\pi}{d} j(k + \alpha_x)\right) |j\rangle_A \text{ and} \tag{S5}$$

$$|L_{l|y}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(i \frac{2\pi}{d} j(-l + \beta_y)\right) |j\rangle_B , \tag{S6}$$

where $\alpha_1 = 0, \alpha_2 = 1/2, \beta_1 = 1/4$ and $\beta_2 = -1/4$. These are the eigenvectors from the Fourier and inverse Fourier bases as defined in the original CGLMP inequality [1]. This gives rise to a joint probability of

$$P(k, l|x, y) = \frac{1}{d^3} \frac{\sin^2\left[\pi(k + \alpha_x) + \pi(-l + \beta_y)\right]}{\sin^2\left[\frac{\pi}{d}(k + \alpha_x) + \frac{\pi}{d}(-l + \beta_y)\right]}. \tag{S7}$$

Here we have grouped Alice's $(k + \alpha_x)$ and Bob's $(-l + \beta_y)$ terms together. The ks and ls are the measurement outcomes, given a choice of basis $x = 1, 2$ for Alice and $y = 1, 2$ for Bob.

In our work, both Alice and Bob both perform measurements in the Fourier basis, with measurements corresponding to

$$|\phi\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(-ij\phi\right) |j\rangle \,. \tag{S8}$$

Note that in the original CGLMP work, the Fourier transform is defined without the negative sign. The complex overlap is then

$$\langle\phi|_A \langle\phi|_B |\psi\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(ij\phi_A\right) \langle j|_A \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(ij\phi_B\right) \langle j|_B \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle_A |j\rangle_B \tag{S9}$$

$$= \frac{1}{\sqrt{d^3}} \sum_{j=0}^{d-1} \exp\left(ij(\phi_A + \phi_B)\right). \tag{S10}$$

The joint probability parameterised by ϕ_A and ϕ_B that we measure is given by

$$P(\phi_A, \phi_B) = |\langle\phi|_A \langle\phi|_B |\psi\rangle|^2 \tag{S11}$$

$$= \frac{1}{d^3} \left| \sum_{j=0}^{d-1} \exp\left(ij(\phi_A + \phi_B)\right) \right|^2 \tag{S12}$$

$$= \frac{1}{d^3} \frac{\sin^2\left[d(\phi_\mathrm{A} + \phi_\mathrm{B})/2\right]}{\sin^2\left[(\phi_\mathrm{A} + \phi_\mathrm{B})/2\right]} \tag{S13}$$

$$= \frac{1}{d^3} \frac{\sin^2\left[\frac{d}{2}\phi_\mathrm{A} + \frac{d}{2}\phi_\mathrm{B}\right]}{\sin^2\left[\frac{1}{2}\phi_\mathrm{A} + \frac{1}{2}\phi_\mathrm{B}\right]}. \tag{S14}$$

In practice, the finite resolution of our measurement system means that we record discrete phases. If, in any 2π window, an experiment records N outcomes, these can be identified as the d outcomes from M measurement bases, i.e. $N = M \times d$.

The experimentally measured coincidences at the phases ϕ_A and ϕ_B can be mapped to the conditional probabilities required for the CGLMP inequality via

$$\phi_A \to \frac{2\pi}{d}\left(a + \frac{x}{M}\right) \text{ and } \phi_B \to \frac{2\pi}{d}\left(\mathrm{mod}(-b, d) + \frac{y}{M} - \frac{1}{4}\right), \tag{S15}$$

where $a = \{0, 1, ...d-1\}$ is the measurement outcome index for Alice, and $x = \{0, 1, ..., M-1\}$ is the measurement bases index for Alice, and Bob's measurement is indexed by b and y. This gives rise to the conditional probability

$$P(a, b|x, y) = \frac{1}{d^3} \frac{\sin^2\left[\pi\left((a + \frac{x}{M}) + (\mathrm{mod}(-b, d) + \frac{y}{M} - \frac{1}{4})\right)\right]}{\sin^2[\frac{\pi}{d}\left((a + \frac{x}{M}) + (\mathrm{mod}(-b, d) + \frac{y}{M} - \frac{1}{4})\right)]}, \tag{S16}$$

where the measurement vectors are

$$|A_{a|x}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(-i\frac{2\pi}{d} j \left(a + \frac{x}{M}\right)\right) |j\rangle_A \text{ and} \tag{S17}$$

$$|B_{b|y}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(-i\frac{2\pi}{d} j \left(\mathrm{mod}(-b, d) + \frac{y}{M} - \frac{1}{4}\right)\right) |j\rangle_B \,. \tag{S18}$$

We see that the outcomes k and l from the CGLMP inequality are the outcomes a and b in our work, and the additional contributions to the phases α_x and β_y required for the basis choices are the terms x/M and $y/M - 1/4$. Note that in the CGLMP work x, y are the binary inputs $x, y \in \{1, 2\}$, whereas in our work we measure in multiple bases corresponding to $x, y \in \{0, 1, ..., M-1\}$. Whether the Fourier transform is defined with or without the negative sign results in the same joint conditional probability, equation S16.

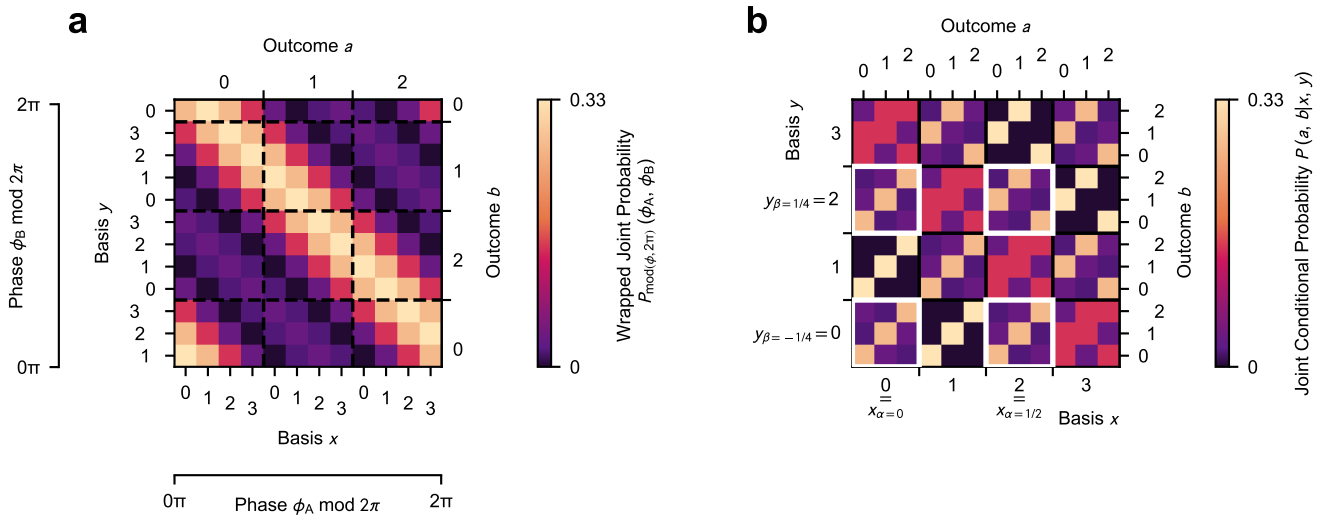

FIG. 1. **Visualisation of relationship between, ϕ_A (ϕ_B), and, a (b) and x (y) for $d = 3$ and $M = 5$. a**, Theoretical wrapped joint probability $P_{\mathrm{mod}(\phi,\ 2\pi)}(\phi_A, \phi_B)$. **b**, Joint probability distribution sorted by measurement bases.

DATA ANALYSIS

Mapping the joint spectral intensity $C(\Delta\nu_A, \Delta\nu_B)$ to the wrapped joint probability $P_{\mathrm{mod}(\phi,\,2\pi)}(\phi_A, \phi_B)$

Whenever a function has a finite extent, its Fourier transform is modulated by an envelope. For example, single-slit diffraction modulates the double-slit interference pattern in Young's double slit experiment. Similarly, in our system, the joint spectral intensity is modulated by a Gaussian envelope due to the finite widths σ_t of our Gaussian time bins. Now consider the case where we are interested in the joint probability distribution $P(\phi_A, \phi_B)$ in the Fourier domain of our entangled state. The amplitude of the joint probability distribution $P(\phi_A, \phi_B)$ decreases as we move away from the distribution centre. Therefore, the modulated distribution over one phase period does not accurately represent the sought after unmodulated $P(\phi_A, \phi_B)$. Li and Zhao [2] showed that when only using one phase period of the modulated $P(\phi_A, \phi_B)$ to extract the necessary the CGLMP measurements, the maximal theoretically possible Bell parameter decreases - in some cases to even below the LHVM limit. They propose mitigating this effect by carefully selecting experimental parameters to suppress the envelope modulation of $P(\phi_A, \phi_B)$. However, we recognise that the unmodulated $P(\phi_A, \phi_B)$ distribution can be recovered by mapping the modulated $P(\phi_A, \phi_B)$ onto the wrapped joint probability $P_{\mathrm{mod}(\phi,\,2\pi)}(\phi_A, \phi_B)$. This ensures that no outcomes are discarded.

This mapping of the joint spectral intensity $C(\Delta\nu_A, \Delta\nu_B)$ onto $P_{\mathrm{mod}(\phi,\,2\pi)}(\phi_A, \phi_B)$ is achieved as follows. The boundaries of the 2π phase periods are found by fitting three diagonal Gaussians to the principal interference fringes in the $C(\Delta\nu_A, \Delta\nu_B)$ data. Their separation was restricted to an integer multiple of d to ensure an integer number of measurement bases M. We infer $C(\phi_A,\ \phi_B)$ from the best fit of the principal fringes separation parameter. Now that the cyclic 2π square regions have been identified, we stack these regions, i.e. sum them, to obtain the wrapped $C_{\mathrm{mod}(\phi,\,2\pi)}(\phi_A, \phi_B)$. This coincidence distribution is normalised to $P_{\mathrm{mod}(\phi,\,2\pi)}(\phi_A, \phi_B)$ such that for each combination of measurement basis x and y, the total probability for all measurement outcome combinations a and b sums to 1, i.e. $\sum_{a,b=0}^{d-1} P(a, b|x, y) = 1$. The relationship between ϕ_A (ϕ_B), and a (b) and x (y) is given by equation 3 (4). For a visualisation of these relationships, see Fig. 1.

Poisson statistics

Assuming the joint spectral intensity coincidence detection events C follow Poisson statistics, its standard deviation σ is given by $\sigma = \sqrt{C}$.

The standard deviation on the Bell parameters I_d in Fig. 5b in the main paper was determined as follows for

each dimension d. A Poisson distribution was created where the expected number of events occurring was set by the wrapped coincidence distribution $C_{\mathrm{mod}(\phi,\,2\pi)}(\phi_{\mathrm{A}},\phi_{\mathrm{B}})$. The Poisson distribution was then normalised to $P_{\mathrm{mod}(\phi,\,2\pi)}(\phi_{\mathrm{A}},\phi_{\mathrm{B}})$ and its Bell parameter was extracted. This was repeated 50 times. The standard deviation on each Bell parameter I_d of the experimental data was calculated based on the set of the Bell parameters of its respective 50 Poisson distributions. It was checked that 50 repeats were sufficient as increasing the number of Poisson distribution repeats did not impact the magnitude of the standard deviation significantly.

Statistical significance

The estimate of the Bell parameter indicates a violation of the CGLMP inequality in dimensions up to eight. Here, we discuss the statistical significance of these violations and show that they all come with vanishingly small p-values. To quantify the statistical significance, we need to evaluate the probability that a local hidden variable model (LHVM) could reproduce the measured finite-statistics values of the CGLMP parameter. The probability of statistical fluctuations in the LHVM explaining the measured violations is called the p-value. Thus, a small p-value corresponds to a high confidence in observing nonlocality.

To estimate the p-values for the violations obtained in the experiment, we use the method described in [3]. We begin considering a general Bell inequality in the form

$$\mathcal{B} = \sum_{x,y,a,b} s_{ab}^{xy} P(ab|xy) \leq \beta \tag{S19}$$

and estimate the probability of selecting settings (x, y) from the photon coincidences counts $N_{a,b,x,y}$ detected when obtaining outcome (a, b)

$$P(x,y) = \frac{\sum\limits_{a,b} N_{a,b,x,y}}{\sum\limits_{a,b,x,y} N_{a,b,x,y}}. \tag{S20}$$

The inequality can then be rewritten as

$$\mathcal{B} = \sum_{x,y,a,b} s_{abxy} P(x,y) P(ab|xy). \tag{S21}$$

In this form, s_{abxy} can be viewed as the weighted score that the observers obtain when they answer a and b to the question x and y. Any round of the Bell game then corresponds to evaluating a cost function whose value is between $s_{max} = \max\limits_{a,b,x,y} s_{abxy}$ and $s_{min} = \min\limits_{a,b,x,y} s_{abxy}$. The total score that the observers obtain after n rounds of the experiment will then be

$$c = \sum_{i=1}^{n} s_{a_i b_i x_i y_i} \tag{S22}$$

This form is favourable because the data can now be interpreted in terms of an n-round game in which a certain score is awarded in every round. Since $s_{abxy} = s_{ab}^{xy}/P(x,y)$, this score is influenced by the settings probability distribution. In order to bound the p-value associated with the measured score, we use McDiarmid's inequality [4]

$$p\text{-value} \leq \left(\left(\frac{s_{\max} - \beta}{s_{\max} - c/n} \right)^{\frac{s_{\max} - c/n}{s_{\max} - s_{\min}}} \left(\frac{\beta - s_{\min}}{c/n - s_{\min}} \right)^{\frac{c/n - s_{\min}}{s_{\max} - s_{\min}}} \right)^{n}. \tag{S23}$$

We apply this bound to the experiment's data. Due to the high number of photon coincidences obtained with our setup and the comparatively large estimated inequality violations, all p-values up to dimension 6 are all below 10^{-200}. For dimensions 7 and 8, the smaller estimated inequality violations lead to larger p-values but these are still vanishingly small. Specifically, they are $8.7 \cdot 10^{-81}$ and $9.7 \cdot 10^{-25}$ for $d = 7$ and $d = 8$ respectively.

More nonlocality from the same data?

The experiment provides measurement statistics in many local bases. It is therefore natural to ask whether it is reasonable to focus on the CGLMP inequality given that this test only uses two bases per party.

In general, the nonlocality of any given bipartite probability distribution can be assessed by means of convex programming techniques. In principle, this offers an exact solution. If the distribution is found to be nonlocal, one can compute a suitable distance measure to the set of correlations realisable with LHVM. One can also use the duality theory of convex programming to extract a Bell inequality that is violated by the target distribution [5]. However, all this rapidly becomes expensive as the number of inputs and outputs grows, which is the case in our experiment.

To overcome this obstacle, we have used efficient approximation methods for determining whether the distribution is nonlocal, and if so extracting the relevant Bell inequality. Our approach is based on the Frank-Wolfe algorithm [6, 7] and closely follows the method developed in [8].

As shown in Tab. I, by constructing Bell inequalities tailored to our measured data, we obtain sharper bounds in scenarios with two measurement settings per party, improving the statistical significance beyond the CGLMP result. Furthermore, when extending the method to more than two measurements settings, we observe no practical benefit. In fact, the noise tolerance remains practically the same, only decreasing from $v_2 = 0.7423$ to $v_{30} = 0.7411$ when moving to 30 settings in dimension 2. A similar behaviour was observed in higher dimensions. For example, in dimension 8 the critical resistance to noise remains unchanged, up to numerical resolution, when increasing the settings from 2 to 9. Additionally, the statistical penality of adding more settings to the test overshadows those minor gains, ultimately increasing the p-value in all cases.

A natural question could be whether the lack of improvements stems from using the theoretical probability distribution. To investigate this, we tested Bell inequalities obtained when using as target probability distribution the no-singaling point closest to the experimental data. Since the results were nearly identical to those obtained with the theoretical distribution, we do not elaborate further on this analysis.

d	v_{CGLMP}	v_{opt}	p-value$_{\mathrm{CGLMP}}$	p-value$_{\mathrm{opt}}$
2	0.7499	0.7482	≈ 0	≈ 0
3	0.7489	0.7423	≈ 0	≈ 0
4	0.7963	0.7800	≈ 0	≈ 0
5	0.8239	0.8129	≈ 0	≈ 0
6	0.8616	0.8615	≈ 0	≈ 0
7	0.9346	0.9331	$8.72 \cdot 10^{-81}$	$3.04 \cdot 10^{-83}$
8	0.9692	0.9640	$9.74 \cdot 10^{-25}$	$1.35 \cdot 10^{-34}$

TABLE I. Critical visibilities (v_{CGLMP}, v_{opt}) and corresponding p-values (p-value$_{\mathrm{CGLMP}}$, p-value$_{\mathrm{opt}}$), obtained when testing the CGLMP ineqality and the optimal two settings inequality found numerically for dimensions $d = \{2, \ldots, 8\}$.

EXPERIMENTAL DATA

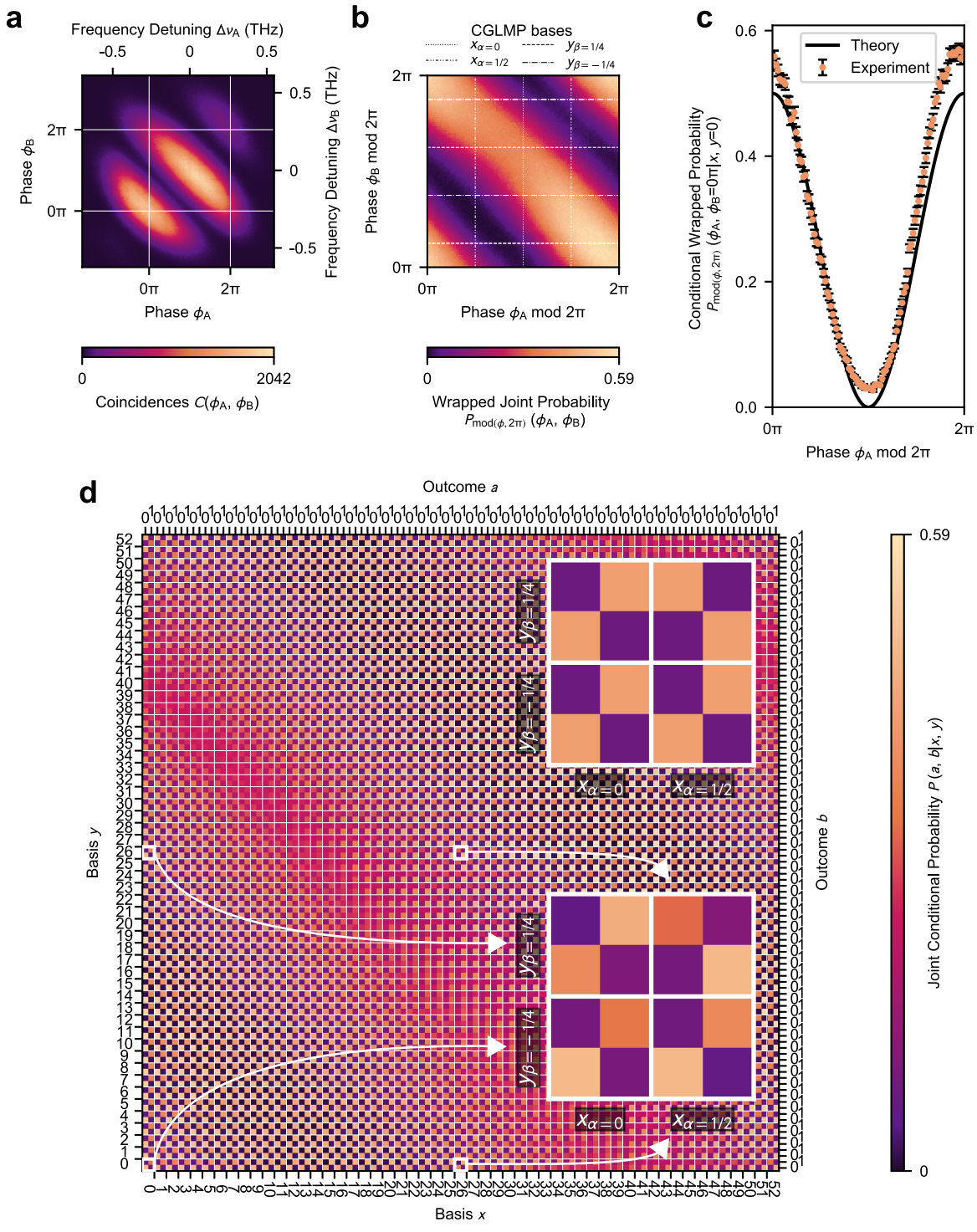

FIG. 2. Data for $d = 2$.

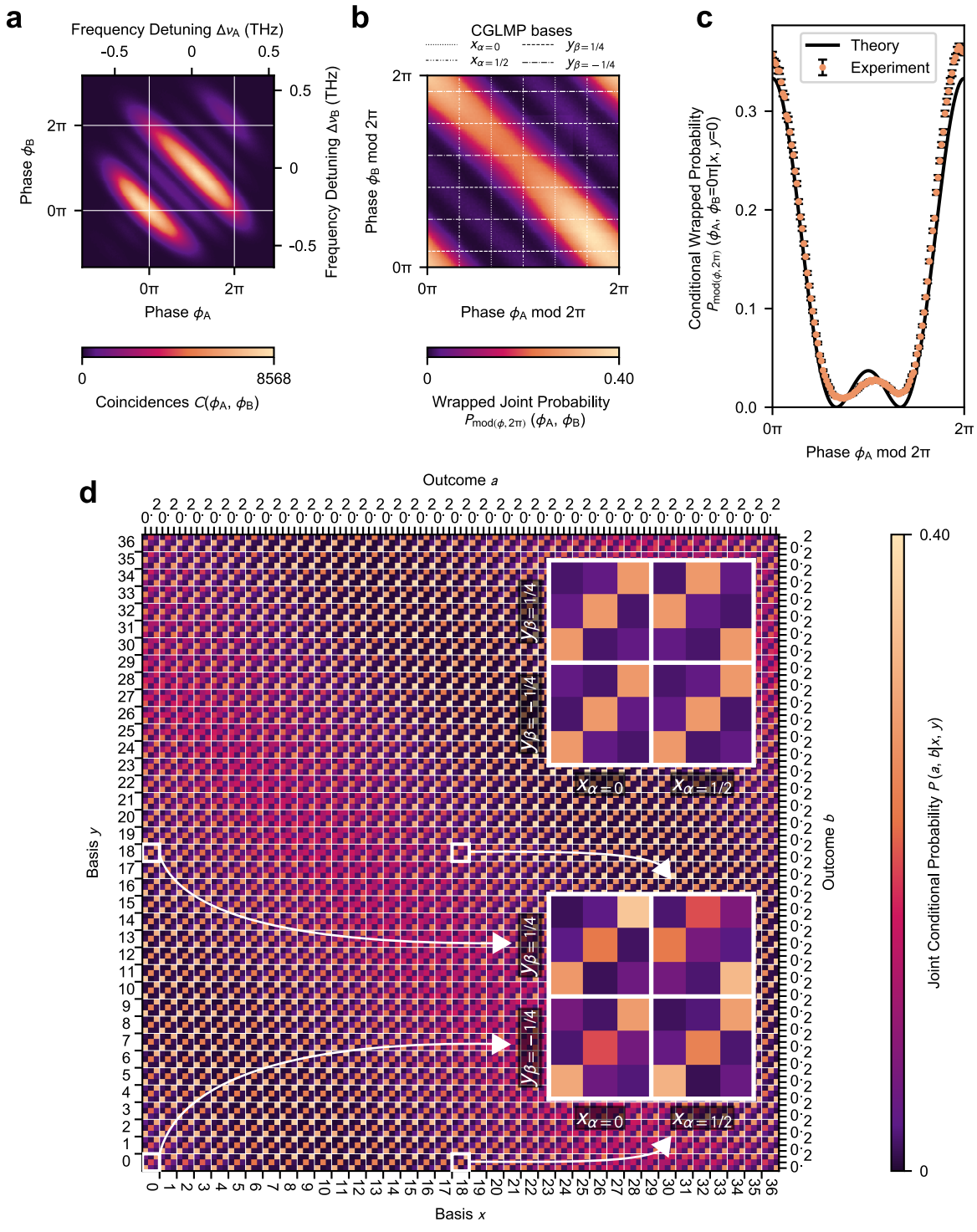

FIG. 3. Data for $d = 3$.

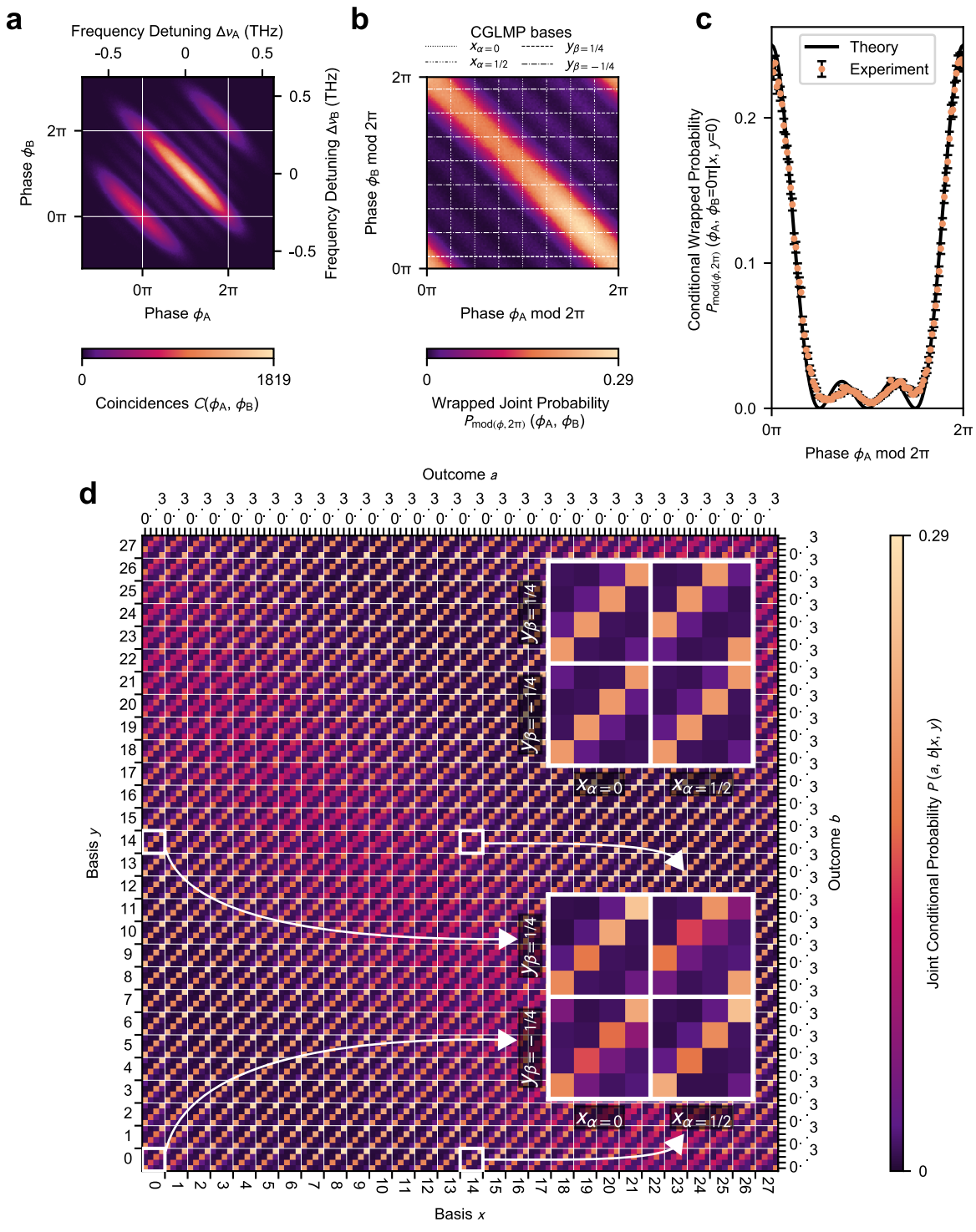

FIG. 4. Data for $d = 4$.

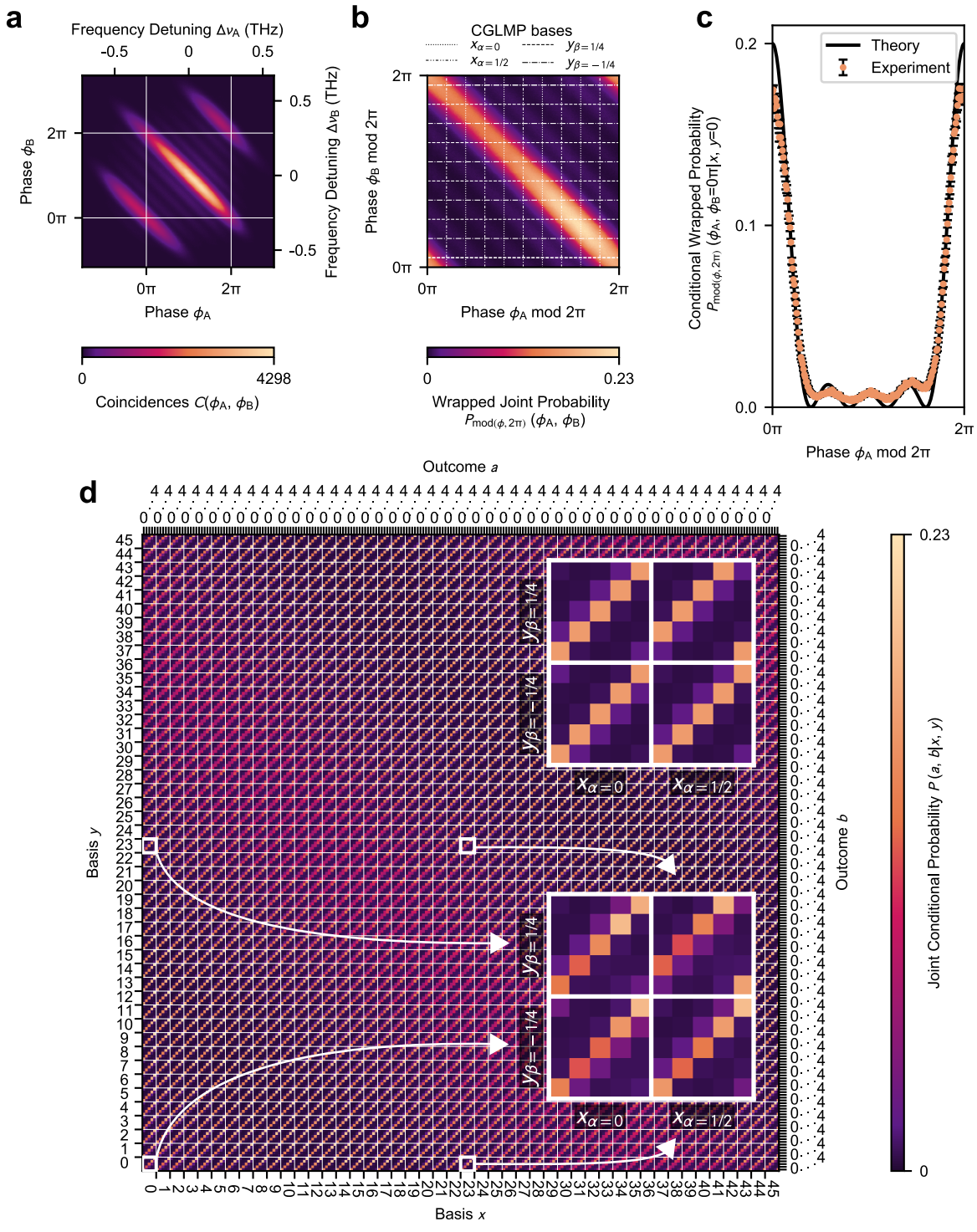

FIG. 5. Data for $d = 5$.

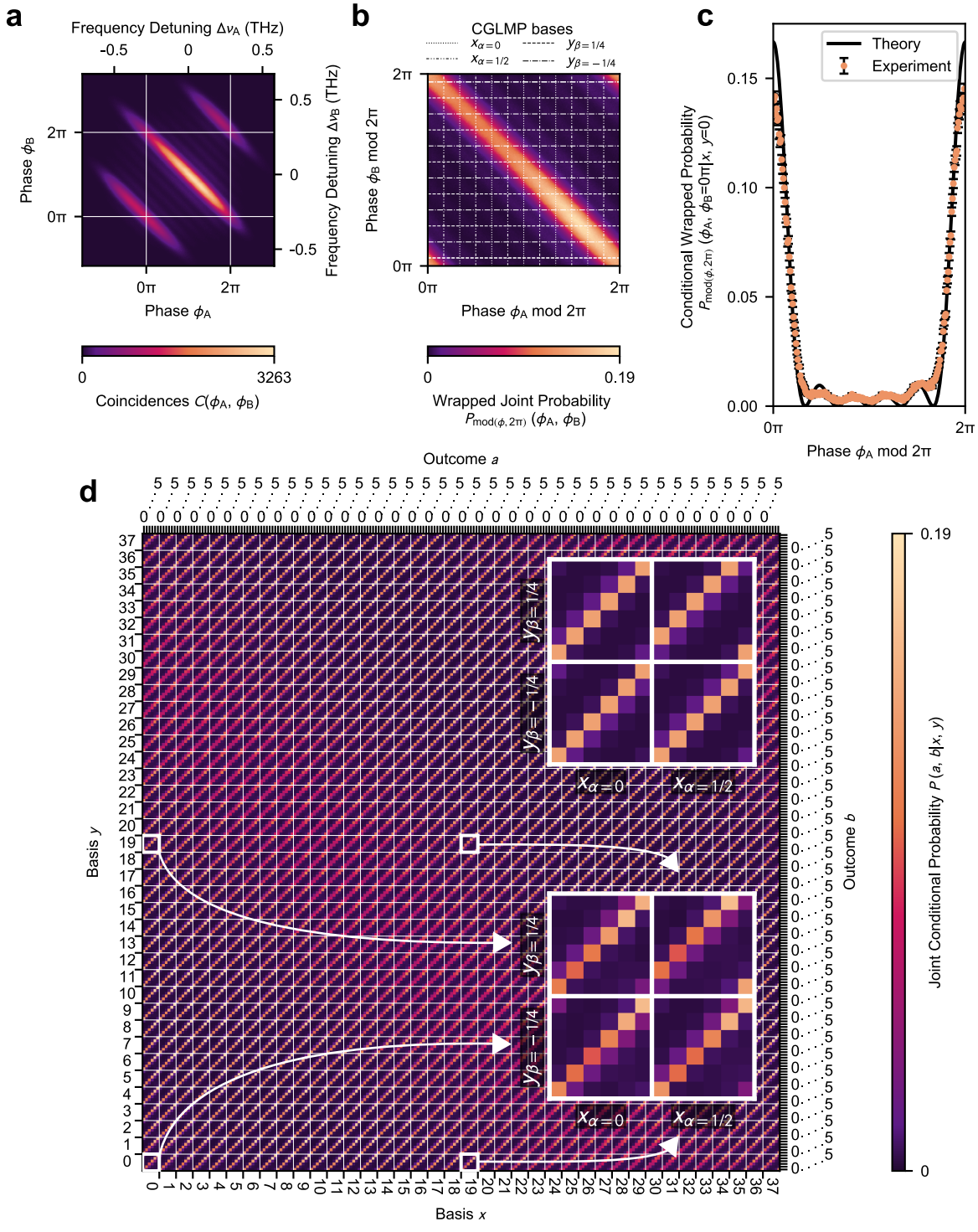

FIG. 6. Data for $d = 6$.

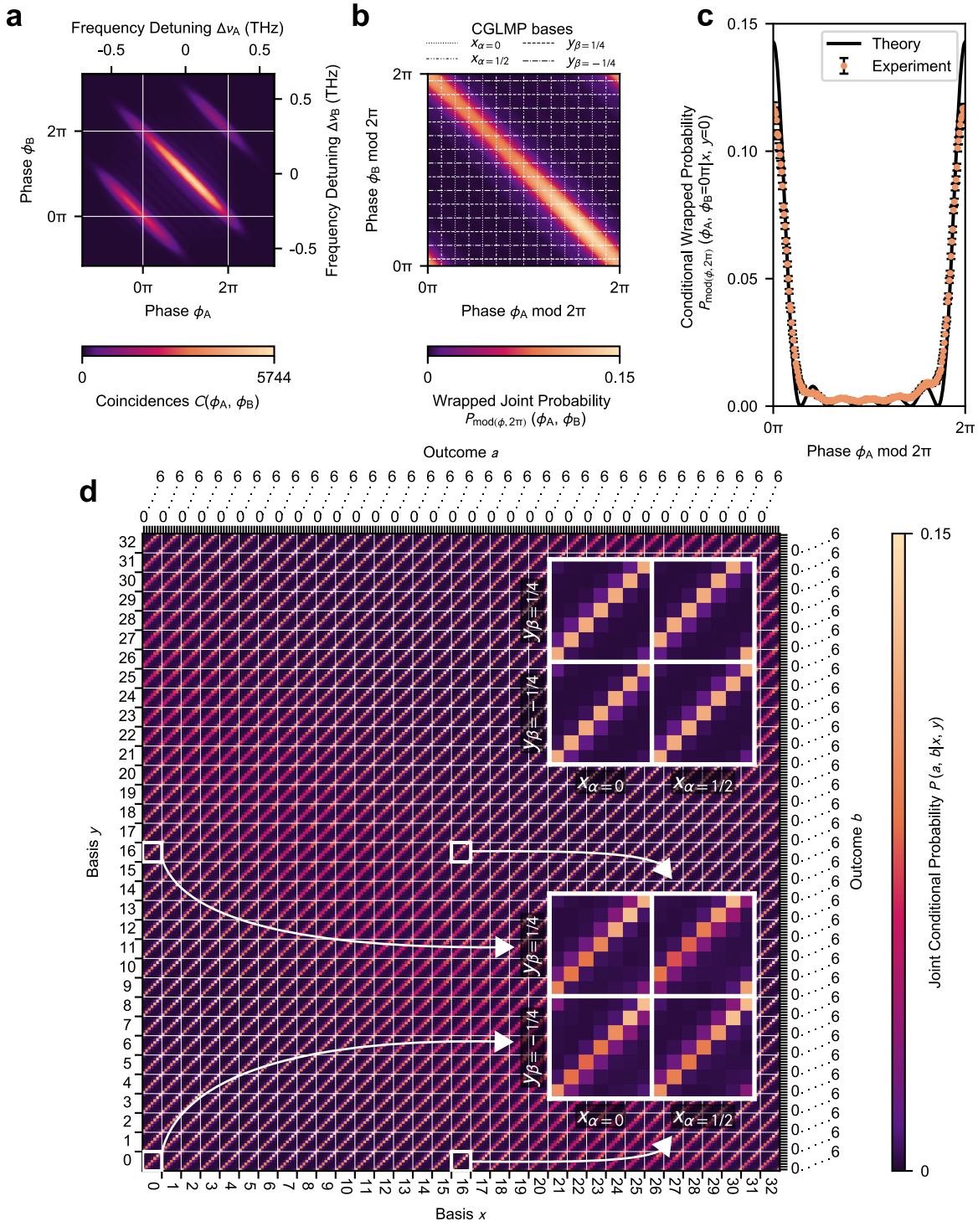

FIG. 7. Data for $d = 7$.

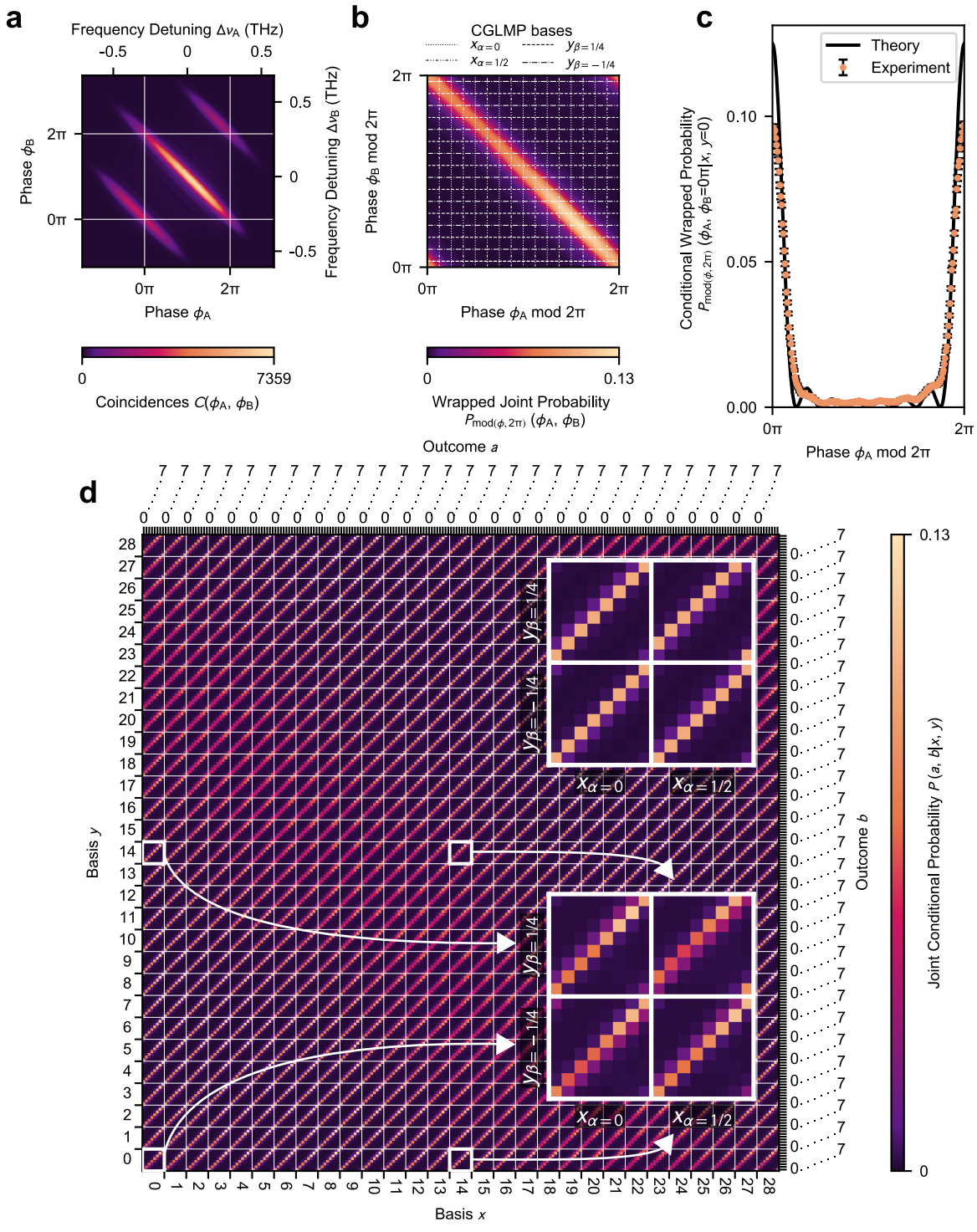

FIG. 8. Data for $d = 8$.

REFERENCES

[1] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. "Bell Inequalities for Arbitrarily High-Dimensional Systems". In: *Physical Review Letters* 88.4 (2002), p. 040404.

[2] W. Li and S. Zhao. "Bell's inequality tests via correlated diffraction of high-dimensional position-entangled two-photon states". In: *Scientific Reports* 8.1 (2018), p. 4812.

[3] D. Elkouss and S. Wehner. "(Nearly) optimal P values for all Bell inequalities". In: *npj Quantum Information* 2.1 (2016).

[4] C. McDiarmid. "Surveys in Combinatorics, 1989: On the method of bounded differences". In: 1989.

[5] A. Tavakoli, A. Pozas-Kerstjens, P. Brown, and M. Araújo. "Semidefinite programming relaxations for quantum correlations". In: *Reviews of Modern Physics* 96.4 (2024).

[6] G. Braun et al. *Conditional Gradient Methods*. 2023.

[7] I. M. Bomze, F. Rinaldi, and D. Zeffiro. *Frank-Wolfe and friends: a journey into projection-free first-order optimization methods*. 2021.

[8] S. Designolle et al. "Improved local models and new Bell inequalities via Frank-Wolfe algorithms". In: *Physical Review Research* 5.4 (2023).