Supporting Information

Krenn et al. 10.1073/pnas.1402365111

SI Text

Laguerre–Gauss Modes

Laguerre–Gauss (LG) modes are solutions of the paraxial wave equation (which, for example, describes the propagation of laser beams). They are described by two mode numbers: the radial mode number n and the angular mode number l. The LG modes are orthogonal to each other, and form a complete basis. They can be written as

$$LG_{n,l}(r,\phi,z=0)=\frac{N_{n,l}}{w_0}\left(\frac{r\sqrt{2}}{w_0}\right)^{|l|}\exp\left(-\frac{r^2}{w_0^2}\right)\mathrm{L}_n^{|l|}\left(\frac{2r^2}{w_0^2}\right)\exp(il\phi), \quad [S1]$$

where $N_{n,l}$ is a normalization constant, w_0 is the beam waist at $z=0$, and $L_n^{|l|}$ are Laguerre polynomials. The l number corresponds to the orbital angular momentum (OAM) of the mode, and n stands for the number of nodes in the intensity profile in radial direction.

Mathematical Formulation of Our Result

With d-dimensional entanglement we understand correlations that can never be explained by convex combinations of pure states with a Schmidt rank lower than d. Notice that we indeed have entanglement in a 34,596 (186×186)-dimensional Hilbert space between almost all possible involved dimensions. However, even if all levels are entangled, it might still be that a convex combination of qubit states would still be enough to explain these correlations. This is what we explicitly exclude in deriving the nonlinear witness for the Schmidt number, where an entanglement dimensionality of 103 signifies that no combination of $d=102$ states could explain the observed correlations, thus proving genuine high-dimensional entanglement.

Bounding the Schmidt Number from Normalized Subspace Correlations

The goal of this section is to the determine lower bounds on the Schmidt number from the sum of all two-dimensional normalized subspace correlations in a D dimensional system. It is organized as follows:

i) We first introduce the general correlation criterion for normalized subspaces.
ii) We continue by deriving tight lower bounds for the same correlations in nonnormalized subspaces.
iii) Finally we derive the global maximum of this correlation function for perfectly correlated modes.

Using the following abbreviation

$$|LG_{n,l}\rangle=|k\rangle, \quad [S2]$$

where we count all LG modes of $\{n,l\}$ [i.e., all Laguerre-Gauss modes with mode number n (radial modes) and l (OAM modes)], we can represent the performed measurements via the following operators:

$$\begin{aligned}\sigma_x^{kl}&:=|k\rangle\langle l|+|l\rangle\langle k|,\\ \sigma_y^{kl}&:=i|k\rangle\langle l|-i|l\rangle\langle k|,\\ \sigma_z^{kl}&:=|k\rangle\langle k|-|l\rangle\langle l|.\end{aligned} \quad [S3]$$

To lower bound the dimensionality of entanglement in normalized subspaces we use the following correlation function

$$C(\rho)=\sum_{k<l}^{D-1}\sum_{l=1}^{D-1}g\left(\rho^{kl}\right), \quad [S4]$$

where the sum is taken over all ρ^{kl}, the normalized subspace density matrices, where all but two degrees of freedom on both sides are ignored, i.e.,

$$\rho^{kl}:=\frac{(|k\rangle\langle k|+|l\rangle\langle l|)\otimes(|k\rangle\langle k|+|l\rangle\langle l|)\rho(|k\rangle\langle k|+|l\rangle\langle l|)\otimes(|k\rangle\langle k|+|l\rangle\langle l|)}{N_{kl}}, \quad [S5]$$

where N_{kl} is the normalization, such that $\mathrm{Tr}(\rho^{kl})=1$, and

$$g\left(\rho^{kl}\right)=\mathrm{Tr}\left(\left(\sigma_z^{kl}\otimes\sigma_z^{kl}-\sigma_y^{kl}\otimes\sigma_y^{kl}+\sigma_x^{kl}\otimes\sigma_x^{kl}\right)\rho^{kl}\right). \quad [S6]$$

Comparing these to the correlations on the total state, i.e.,

$$f_{kl}(\rho)=\mathrm{Tr}\left(\left(\sigma_z^{kl}\otimes\sigma_z^{kl}-\sigma_y^{kl}\otimes\sigma_y^{kl}+\sigma_x^{kl}\otimes\sigma_x^{kl}\right)\rho\right), \quad [S7]$$

we can write

$$g(\rho_{kl})=\frac{f_{kl}}{N_{kl}}, \quad [S8]$$

and thus

$$C(\rho)=\sum_{k<l}^{D-1}\sum_{l=1}^{D-1}\frac{f_{kl}}{N_{kl}}. \quad [S9]$$

It is important to note that of course we only consider contributions from subspaces with a nonzero contribution, i.e., if $N_{kl}=0$ we set $g_{kl}=0$.

Lower Bounds from Nonnormalized Subspaces. The first step of the witness construction is the determination of the maximal value of $f(\rho):=\sum_{k<l}\sum_{l=1}^{D-1}f_{kl}(\rho)$ for d-dimensional states.

$$\sum_{k<l}^{D-1}\sum_{l=1}f_{kl}\leq\max_{|\psi_d\rangle}f(|\psi_d\rangle\langle\psi_d|), \quad [S10]$$

where $|\psi_d\rangle=\sum_{i=0}^{d-1}\lambda_i|i\rangle_A|i\rangle_B$. The first step to achieve this maximization is to realize that

$$f(|\psi\rangle\langle\psi|) \leq \mathrm{Tr}\left[\left(2D|\phi_D\rangle\langle\phi_D| + (D-3)\sum_{i=0}^{D-1}|ii\rangle\langle ii|\right)|\psi\rangle\langle\psi|\right], \quad \textbf{[S11]}$$

where $|\phi_D\rangle := (1/\sqrt{D})\sum_{i=0}^{D-1}|ii\rangle$ and we got an upper bound via setting all negative contributions to 0. It can be derived (1,2) that the maximal overlap of a D-dimensional state $|\phi_D\rangle$ with a d-dimensional state $\mathrm{Tr}[|\psi_d\rangle\langle\psi_d||\phi_D\rangle\langle\phi_D|])$ is achieved by $|\psi_d\rangle := (1/\sqrt{d})\sum_{i=0}^{d-1}|ii\rangle$. This state also maximizes $\mathrm{Tr}[((D-3)\sum_{i=0}^{D-1}|ii\rangle\langle ii|)|\psi\rangle\langle\psi|]$ such that we can infer that indeed this state achieves the global maximum for $\max_{|\psi_d\rangle} f(|\psi_d\rangle\langle\psi_d|)$ and thus we can directly calculate

$$f(\rho) \leq (2d + (D-3)). \quad \textbf{[S12]}$$

Due to the linearity of $f(\rho)$, the maximum is a pure state, thus **[S12]** is an upper limit for pure and mixed states.

The Structure of the Global Maximum. Before we proceed to bound the maximum of $C(\rho)$ as a function of D and d we require a few observations:

$$g_{kl} = \frac{4\Re e[\langle kk|\rho|ll\rangle] + \langle kk|\rho|kk\rangle + \langle ll|\rho|ll\rangle - \langle kl|\rho|kl\rangle - \langle lk|\rho|lk\rangle}{\langle kk|\rho|kk\rangle + \langle ll|\rho|ll\rangle + \langle kl|\rho|kl\rangle + \langle lk|\rho|lk\rangle} \leq \frac{4\Re e[\langle kk|\rho|ll\rangle] + \langle kk|\rho|kk\rangle + \langle ll|\rho|ll\rangle}{\langle kk|\rho|kk\rangle + \langle ll|\rho|ll\rangle}, \quad \textbf{[S13]}$$

i.e., monotonically decreasing in the elements $\langle kl|\rho|kl\rangle + \langle lk|\rho|lk\rangle$ for all k and l. We now argue that for our physical system it is sufficient to maximize over states that are perfectly correlated, i.e., states that can be written in the form of $\rho = \sum_{k,l} c_{kl}|kk\rangle\langle ll|$ with $\mathrm{Tr}(\rho) \leq 1$. This is physically motivated and discussed in *Effect of Deviation from Perfect Correlation*. This particular form of the maximizing density matrices implies that all Schmidt decompositions can be made in a computational basis and we can write every vector in the decomposition of the maximizing ρ as $|\psi_\alpha\rangle = \sum_{k\in\alpha}\lambda_k^\alpha|kk\rangle$.

This further implies we can decompose the maximizing density matrix as

$$\rho_{max} = \sum_i \sum_{\alpha_i} p_{\alpha_i}|\psi_{\alpha_i}\rangle\langle\psi_{\alpha_i}|, \quad \textbf{[S14]}$$

where $\alpha \subset \{0, 1, (\cdots), D-1\}$ with $|\alpha| \leq d$ (where $|\alpha|$ is the number of elements in α) denotes the set of dimensions in which the decomposition element is entangled. Without loss of generality we consider the case $|\alpha| = d$, as every density matrix that can be decomposed into Schmidt-rank $d' < d$ states is strictly contained in this definition (for some $\lambda_k = 0$). We write **[S13]** as

$$g_{kl}(\rho) = 4\frac{\sum_{\alpha_i:(k\in\alpha_i)\wedge(l\in\alpha_i)}\sum i p_{\alpha_i}\lambda_k^{\alpha_i}\lambda_l^{\alpha_i}}{\sum_{\alpha_i:(k\in\alpha_i)\vee(l\in\alpha_i)}\sum i p_{\alpha_i}\left(\left(\lambda_k^{\alpha_i}\right)^2 + \left(\lambda_l^{\alpha_i}\right)^2\right)} + 1. \quad \textbf{[S15]}$$

Here, λ_k can be chosen to be real (as it is standard for the Schmidt decomposition), as any complex phases would only decrease the value of $Re[\lambda_k\lambda_l]$. Using the abbreviation $\tilde{\lambda}_k^{\alpha_i} := \sqrt{p_{\alpha_i}}\lambda_k^{\alpha_i}$ we can take the partial derivatives with respect to $\tilde{\lambda}_k^{\alpha_i}$ to find general conditions for all extremal points:

$$\frac{\partial}{\partial\tilde{\lambda}_k^{\alpha_i}}C(\rho) = \sum_{l\in\alpha}\frac{4\tilde{\lambda}_l^{\alpha_i}}{N_{kl}} - 2\tilde{\lambda}_k^{\alpha_i}\sum_l\frac{g_{kl}-1}{N_{kl}} = 0. \quad \textbf{[S16]}$$

Thus

$$\sum_l\frac{g_{kl}-1}{N_{kl}} = \sum_{l\in\alpha}\frac{4\tilde{\lambda}_l^{\alpha_i}}{N_{kl}2\tilde{\lambda}_k^{\alpha_i}}. \quad \textbf{[S17]}$$

Since the left-hand side is symmetric in k and l we can conclude that $\tilde{\lambda}_l^{\alpha_i} = \tilde{\lambda}_k^{\alpha_i}$. It then directly follows that

$$\sum_l\frac{g_{kl}-1}{N_{kl}} = \sum_{l\in\alpha}\frac{4}{2N_{kl}}. \quad \textbf{[S18]}$$

And since only the right-hand side depends on α, every partial sum of $1/N_{kl}$ is equal for $k, l \in \cup\alpha_i$. This implies that all N_{kl} that appear in the union of all α_i will be equal to some N. Thus, we can write

$$C(\rho_{max}) \leq \sum_{k,l\in\cup\alpha_i}\frac{f_{kl}}{N} + \sum_{k\in\cup\alpha_i, l\notin\cup\alpha_i}\frac{N_{kl}}{N_{kl}}, \quad \textbf{[S19]}$$

where $N = 2/\cup\alpha_i$ and we know from Eq. **S12** that $\sum_{k<l}\sum_{l=1}^{x-1}f_{kl} \leq 2d + (x-3)$ for all $x \geq d$. We thus get an upper bound

$$C(\rho) \leq Dd + \frac{D}{2}(D-3). \quad \textbf{[S20]}$$

Now we know that all extremal points (of course including all maxima) of the function are upper bounded by this value. At the same time we know that on every boundary (given by a partial set of $|\psi_{\alpha_i}\rangle$) all extremal points are upper bounded by an even lower value. Together this implies that indeed we have found the global maximum.

Maximizing Quantum State. It is worth pointing out that in general for every $d < D$ this bound is tight, i.e., there exists a mixed state which saturates this bound

$$\rho_d = \frac{1}{\binom{D}{d}}\sum_{\alpha\subset\{0,1,\ldots,D-1\}}|\phi_\alpha^d\rangle\langle\phi_\alpha^d|, \quad \textbf{[S21]}$$

where $|\phi_\alpha^d\rangle := (1/\sqrt{d})\sum_{k\in\alpha}(|kk\rangle)$ and $|\alpha| = d$.

When we compare the maximizing state with the expected state in our experiment (nonmaximally entangled pure state), we see that we most likely will underestimate the Schmidt number for photons created in down-conversion, thus we might have an even higher number of entangled dimensions.

Effect of Deviation from Perfect Correlation

For the derivation of the bounds **S20**, we assumed perfect correlations between all involved modes. This is physically motivated. The first quantum number corresponds to the angular momentum of light, which is known to be conserved in the down-conversion process (3–5). The second quantum number corresponds is the radial momentum, for which the limit of perfect correlation (so-called quasi-Schmidt modes) can be realized experimentally with very good approximation (6, 7), as can be seen in our experimental data in Fig. 3*B*.

Deviations from this assumption have been analyzed numerically. We have analyzed how nonperfect correlated states or nonorthogonal projections influence our quantity **S9**. We have found that for small deviations such as observed in our experiment, the quantity **S9** can only decrease, thus the resulting bounds still hold. Therefore, the application of our entanglement dimensionality criterion is perfectly justified.

Observed Visibilities

For the visualization of the observed visibilities, please see Fig. S2.

Considered Modes vs. Entanglement Dimensionality

Fig. S3 visualizes the considered modes and corresponding entanglement dimensionality for the measured results in our experiment.

Example: To explain this behavior more, we give a simple example. Considering the state

$$|\psi\rangle = N(0.5|0,0\rangle + 0.07|1,-1\rangle + 0.01|2,-2\rangle + 0.01|3,-3\rangle),$$

where N is the normalization, we can calculate the subspace visibilities. The sum of the x, y, and z visibilities in every subspace ($SV = V_x + V_y + V_z$) gives

$$SV^{0,1} = 1.55^{0,1}, \quad SV^{0,2} = 1.08, \quad SV^{0,3} = 1.08, \quad SV^{1,2} = 1.56,$$
$$SV^{1,3} = 1.56, \quad \text{and} \quad SV^{2,3} = 3.$$

Calculating the sum of all visibilities (according to Eq. **1** in the main text) gives $W = 9.723$. Comparing with our inequality from **[2]** in the main text, using $D = 4$, we can confirm a two-dimensional entanglement (bounds are 6, 10, and 14 for two-, three-, and four-dimensional entanglement, respectively).

If we only consider modes 1, 2, and 3 (and do not consider mode 0), we get $W = 6.12$, which confirms a three-dimensional entanglement (bounds are 3 and 6 for two- and three-dimensional entanglement, respectively). This example shows how considering a smaller number of modes can verify a higher-dimensional entanglement with our method.

Computer-Generated Holograms

The holograms on the spatial light modulators (SLMs) are calculated using a plane-wave approximation. This could lead to nonorthogonal projective measurements. However, for our system, this effect can only reduce the visibilities and thus reduce the observed dimensionality (Fig. S1). We restrict ourselves to two-dimensional subspaces, as this leads to simpler holograms on the SLMs and increases the mode transformation accuracy due to the finiteness of the pixels. Furthermore this method allows us to treat nonmaximally entangled state (Fig. 3*A*) directly; therefore we do not need to perform any entanglement concentration.

Statistical Uncertainty

The detected photon numbers are assumed to be Poisson distributed, which leads to asymmetric distribution especially for low count rates. Analytical treatment of error propagation for such a large number of measurements was not feasible. Therefore, all confidence intervals have been calculated using Monte Carlo simulations. The statistical uncertainty is very small because it depends on 200,000 measurements with a relatively small uncertainty.

1. Gühne O (2004) Detecting quantum entanglement: Entanglement witnesses and uncertainty relations. PhD thesis (Univ of Hannover, Hannover, Germany).
2. Fickler R, et al. (2014) Interface between path and OAM entanglement for high-dimensional photonic quantum information. arXiv:1402.2423.
3. Mair A, Vaziri A, Weihs G, Zeilinger A (2001) Entanglement of the orbital angular momentum states of photons. *Nature* 412(6844):313–316.
4. Torres JP, Alexandrescu A, Torner L (2003) Quantum spiral bandwidth of entangled two-photon states. *Phys Rev A* 68:050301.
5. Dada AC, Leach J, Buller GS, Padgett MJ, Andersson E (2011) Experimental high-dimensional two-photon entanglement and violations of generalized Bell inequalities. *Nat Phys* 7:677680.
6. Miatto FM, Yao AM, Barnett SM (2011) Full characterization of the quantum spiral bandwidth of entangled biphotons. *Phys Rev A* 83:033816.
7. Salakhutdinov VD, Eliel ER, Löffler WL (2012) Full-field quantum correlations of spatially entangled photons. *Phys Rev Lett* 108(17):173604.

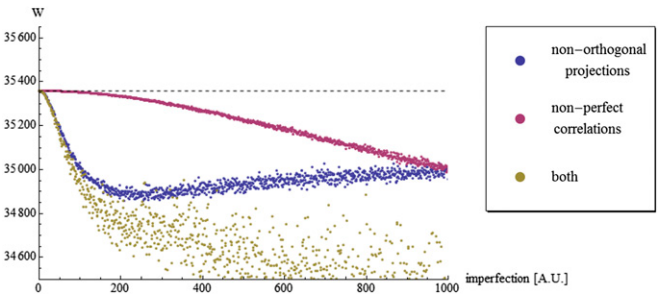

Fig. S1. A nonmaximally entangled state with $D = 186$ dimensions is considered, similar to the state we expect from our experiment. It is considered how nonorthogonal projections and other measurement-induced errors (blue), nonperfect correlations (red), and both effects simultaneously (yellow) influence the value of the quantity W [or $C(\rho)$ (Eq. **S9**)]. For each type, 1,000 cases are calculated. The imperfections are introduced randomly, in each step the introduced imperfections increase. The dashed black line shows the value calculated without imperfections. It can be seen that any introduction of nonperfect correlations or nonorthogonal projections only decreases W. These results guarantee that the experimentally observed deviation from the perfectly correlated state cannot artificially increase the observed dimensionality (in contrast, it decreases the quantity W, thus the observed Schmidt number).

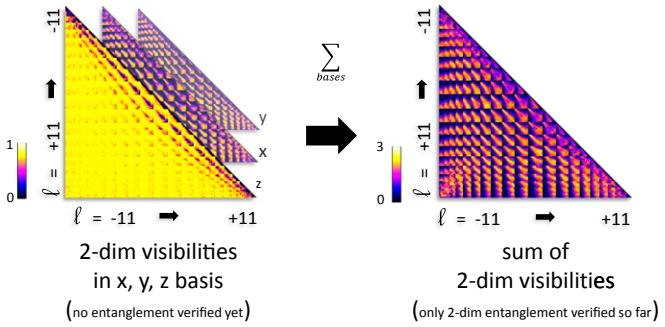

Fig. S2. (*Left*) The visibilities of all of the two-dimensional subsets in all of the three bases are shown. The *z* visibility is usually larger, because nonmaximal entanglement (Fig. 3*A*) reduces visibility in the *x* and *y* basis only. When we sum up the three visibilities (*Right*), we can see that some subsets are more entangled than others. Every subspace with a value larger than 1 is two-dimensional entangled; this is true for most of the 17,000 subspaces. Modes with similar count rates have high visibilities, whereas modes with a very different count rate have very low visibilities in the *x* and *y* bases. To reveal information about the global, high-dimensional entanglement, we have to sum up all visibilities of all subsets and calculate quantity *W* in Eq. **1**.

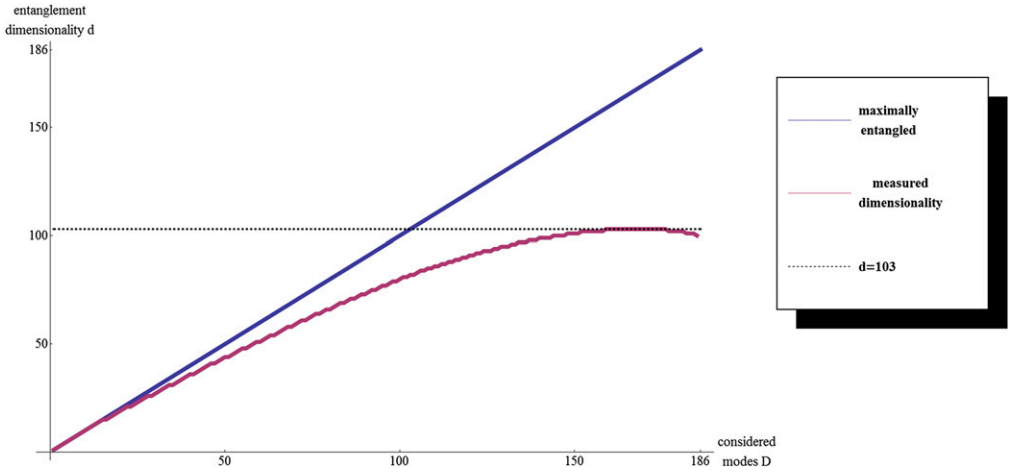

Fig. S3. The entanglement dimensionality *d* depends on the number of modes considered for $\langle W \rangle$, as described in inequality **2** in the main text. We measured all two-dimensional subspace correlations in *x*, *y*, and *z* bases for 186 different modes. If the state would be maximally entangled, one could extract as much entanglement dimensionality as one considers modes (blue line). However, as it can be seen in Fig. 4*A*, there are modes which contribute stronger to $\langle W \rangle$ (yellow region), and some modes that contribute less to the witness (for instance the Gaussian mode $LG_{0,0}$). One can remove certain modes (i.e., not considering all two-dimensional subspace measurement that contains these specific modes). We find a maximal detectable entanglement dimensionality *d*, if we only consider the 167 strongest contributing modes (red curve). This leads to 103-dimensional entanglement (black dashed line). We also observe a set of 15 modes which are entangled in 15 dimensions, which could be significant in special protocols in quantum communication.