

Topological implications in Quantum tomography

Teiko Heinosaari, Michael M. Wolf, *talk presented by:* David Reeb

Niels Bohr Institute, 2100 Copenhagen, Denmark

(Dated: January 13, 2011)

How many measurement settings or outcomes are necessary in order to identify a quantum system which is constrained by prior information? We show that if the latter restricts the system to a set of lower dimensionality, then topological obstructions can increase the required number of outcomes (or binary settings) by a factor of two over the number of real parameters needed to characterize the system. Conversely, almost every measurement becomes informationally complete with respect to the constrained set if the number of outcomes exceeds twice the Minkowski dimension of the set.

We provide a general analysis of topological implications in quantum tomography and apply the ideas to determine the minimal cardinalities of measurements which are informationally complete w.r.t arbitrary rank constraints. Explicit constructions for such measurements are given.

Quantum tomography aims at identifying quantum systems. In order to achieve this, we use measurement data typically supplemented by prior information. In this work we consider cases where prior information effectively reduces the dimensionality, i.e., the number of parameters which are necessary to characterize the state of a system. Physically, one may think of scenarios of interferometry, process tomography or parameter estimation, where one prepares the initial state which then evolves depending on a certain number of unknown parameters before one measures the final system. Effective reductions of the number of parameters can also be due to a constraining symmetry or fixed energy or particle number.

We are interested in identifying the state by using as little measurement data as possible. Clearly, if states in the considered set are parameterized by a number, say $d_{\mathcal{M}}$, of independent real parameters, then we need at least this number of measurement outcomes or binary measurements in order to pinpoint the state. As an example take the manifold of pure states in a d -dimensional Hilbert space. Their description requires $2d-2$ real parameters, as opposed to d^2-1 real parameters needed to describe an arbitrary density matrix. So if we want to determine a pure state by a single measurement with m outcomes (or, equivalently, m binary measurements), how large has m to be? Is $m \sim 2d$ sufficient as counting parameters suggests, or do we need $m \sim d^2$ since after all the set of pure state density matrices spans the entire state space? This particular question has been addressed in a number of publications [Wei92, AW99, FSC05], but the answer regarding the optimal scaling of m has remained somewhat elusive, so far.

A related problem has been addressed based on *compressed sensing* ideas, where it has been shown [GLF⁺10] that for $d \times d$ density matrices of rank r , $m = O(dr \log(d)^2)$ binary measurements are sufficient in order to identify the state with high probability. In this light we emphasize that our focus lies on schemes which identify the system unambiguously and deterministically. We should also stress that in all discussed scenarios “measurement” always refers to a statistical experiment rather than to a single-shot experiment.

The present work contains two complementary parts. The first one, discussed in Sec.I, is based on the observation that any measurement which is informationally complete when supplemented by prior information is a mapping into the space

of measurement outcomes which preserves topological invariants. Together with linearity of quantum mechanics this imposes non-trivial constraints on the minimal number of measurement settings or outcomes which are needed to complete the prior information. This point of view enables us for instance to show that for pure states any informational complete measurement requires $m \sim 4d$ up to an additive logarithmic correction.

In the second part, discussed in Sec.II, we then provide upper bounds on the required number of measurement settings or outcomes. We show that if the number of settings or outcomes exceeds twice the Minkowski dimension of the set consistent with prior information, then almost every measurement will suffice to ultimately identify the system unambiguously. We will then provide explicit constructions of measurement schemes which in particular show that the above mentioned $m \sim 4d$ scaling for pure states, and more generally $m \sim 4dr$ for states with rank bounded by r , can indeed be achieved. Technical proofs are given in an appendix.

I. TOMOGRAPHY FROM A TOPOLOGICAL PERSPECTIVE

Tomography preserves topology. Let $\mathcal{M} \in \mathbb{C}^{d \times d}$ be a manifold in the set of $d \times d$ density matrices. Examples of manifolds are the set of density matrices of fixed rank[Dit95], states with given spectrum or other unitary orbits such as those arising in interferometry or parameter estimation schemes.

We will write $d_{\mathcal{M}}$ for the *real* dimension of the manifold and think of \mathcal{M} as the set of density matrices constrained by prior information, and $d_{\mathcal{M}} \leq d^2 - 1$ the number of real parameters which are required for characterizing an element in \mathcal{M} .¹

We will consider two related ways of identifying a state $\rho \in \mathcal{M}$ by measurements: (i) we perform a single measurement (POVM) with, say $m + 1$, outcomes from whose prob-

¹ Of course, we may need an extra discrete parameter, which enumerates the charts. The number of the latter is however finite for any compact manifold.

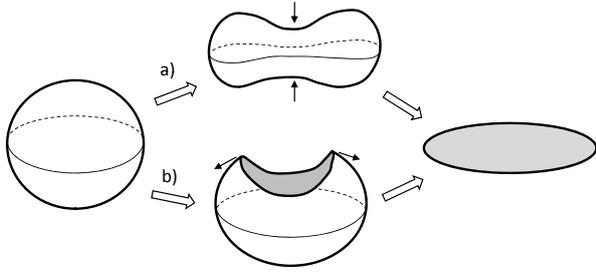


FIG. 1: A mapping from the sphere in \mathbb{R}^3 to the plane \mathbb{R}^2 is either (a) not injective, i.e., it identifies initially different points, or (b) not continuous. When regarding the sphere as the manifold of pure qubit states and the mapping as a measurement with 3 outcomes (of which only two can have independent probabilities), this simple picture implies that a pure state informationally complete (=injective) measurement requires 4 outcomes.

abilities we want to determine ρ , or alternatively (ii) we perform m different measurements and use their expectation values in order to identify ρ . Both scenarios can yield m independent real numbers so that we can view them as a mapping $h : \mathcal{M} \rightarrow \mathbb{R}^m$ from the initial manifold into Euclidean space. In mathematical terms, we get

$$h(\rho)_i = \text{tr}[\rho A_i], \quad i = 1, \dots, m, \quad (1)$$

where the A_i 's are Hermitian matrices which, in scenario (i), have to satisfy the additional constraints $A_i \geq 0$ and $\sum_{i=1}^m A_i \leq \mathbb{1}$. Since the two scenarios can be treated on the same footing, we will often not distinguish between them and just talk about *measurement schemes*, meaning either of them. If h allows to identify any state $\rho \in \mathcal{M}$ unambiguously, we will call the scheme *informationally complete* (w.r.t. \mathcal{M}) meaning that for any pair $\rho_1, \rho_2 \in \mathcal{M} : h(\rho_1) = h(\rho_2)$ implies that $\rho_1 = \rho_2$, i.e., h is injective. Note that the two described scenarios are mathematically equivalent in the sense that m Hermitian operators can always be transformed into POVM elements via $A_i \mapsto A'_i = \alpha A_i + \beta \mathbb{1}$ for suitable $\alpha, \beta \in \mathbb{R}$ so that informational completeness is preserved.²

The crucial point is now that a measurement scheme is informationally complete if and only if h is a *topological embedding*, i.e., it preserves topological properties. The latter means that h is (a) injective, (b) continuous, and (c) has a continuous inverse on its image. While (a) is indeed a reformulation of “informational completeness”, (b) and (c) are consequences of the linearity of the measurement process (see Prop.1 for details).

Hence, for m to admit an informationally complete measurement scheme, it is necessary that there exists a topological embedding $\mathcal{M} \rightarrow \mathbb{R}^m$. This, however, depends not only on the dimension $d_{\mathcal{M}}$ of the manifold but also on its topological invariants.

² Another transformation of interest is orthonormalization in the space of Hermitian matrices, i.e., the fact that we can always assure that $\text{tr}[A_i A_j] = \delta_{ij}$ in case of scenario (ii).

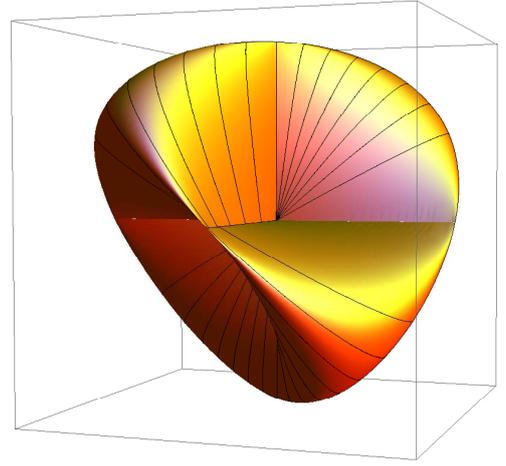


FIG. 2: The *Roman surface* is a self-intersecting surface in \mathbb{R}^3 obtained by a mapping from the real projective plane $\mathbb{R}P^2$. In our context it arises when the manifold of three-dimensional real, pure states is mapped onto the expectation values of three σ_x -type observables. Since the manifold is not orientable, every three-outcome measurement has to be non-injective—here reflected by the self-intersections of the surface.

Example—pure qubit states: Consider the unit sphere in \mathbb{R}^3 which may represent the manifold of pure qubit states on the Bloch sphere (see Fig.1). Take any measurement with 3 outcomes leading to a vector of probabilities in \mathbb{R}^3 . Since probabilities have to sum up to one, all information is already contained in the first 2 components, so the measurement may effectively be regarded as a map from the sphere into the plane \mathbb{R}^2 . For a measurement which is informationally complete different initial states (=points on the sphere) must not be mapped onto the same final point. Due to linearity of quantum mechanics a discontinuous mapping (option (b) in Fig.1) cannot arise from a measurement. Every continuous map into the plane, however, necessarily identifies different points of the sphere ((a) in Fig.1). Consequently, there cannot be a pure state informationally complete measurement with 3 outcomes. More precisely:

Example 1 (Pure qubit states) Let \mathcal{M} be the manifold of pure qubit states and $h : \mathcal{M} \rightarrow \mathbb{R}^m$ with $h(\rho)_i = \text{tr}[\rho^{\otimes n} A_i]$ corresponding to a measurement scheme. The latter is informationally complete w.r.t. \mathcal{M} iff it is so for all qubit density matrices. Moreover, if $m = 2$ then there exist orthogonal pure states ψ and ϕ such that $h(\psi) = h(\phi)$.

The last statement is an application of the Borsuk-Ulam theorem. The following example requires even more measurement outcomes despite $d_{\mathcal{M}} = 2$.

Example—pure states in \mathbb{R}^3 : Consider now \mathcal{M} as the manifold of pure states in \mathbb{C}^3 with real amplitudes, i.e., $|\psi\rangle = \sum_{i=1}^3 x_i |i\rangle$ with $x \in \mathbb{R}^3$. Due to normalization every such state can be represented by a unit vector $x \in S^2$. Since x and $-x$, however, represent the same state, we have to identify antipodes so that $\mathcal{M} = \mathbb{R}P^2$ becomes the *real projective plane*. The latter is known to be a non-orientable 2-manifold so that

there is no topological embedding in \mathbb{R}^3 . In fact, all closed non-orientable surfaces (i.e., topological 2-manifolds) without boundary cannot be embedded in \mathbb{R}^3 , but they do embed into \mathbb{R}^4 . In our context, the map $x \mapsto (x_1x_2, x_2x_3, x_3x_1, x_1^2 - x_2^2)$ is an embedding of \mathbb{RP}^2 in \mathbb{R}^4 which can be realized by a measurement scheme: the first three components can be obtained by σ_x -type measurements and the fourth via a σ_z -type measurement. Hence we obtain:

Example 2 (Pure states in \mathbb{R}^3) Let \mathcal{M} be the manifold of pure states in \mathbb{R}^3 and $h : \mathcal{M} \rightarrow \mathbb{R}^m$ with $h(\rho)_i = \text{tr}[\rho^{\otimes n} A_i]$ corresponding to a measurement scheme. Then informational completeness requires $m \geq 4$ and $m = 4$ can be achieved already for $n = 1$.

For $m = 3$ and $n = 1$ a measurement of the three σ_x -type operators gives rise to the *Roman surface* displayed in Fig.2. The failure of informational completeness is reflected by self-intersections of the surface.

Obstructions from differential topology. Manifolds of interest in quantum tomography often arise from unitary orbits, so that they have a differentiable structure—they are *smooth manifolds* [Lee02]. In such a case we may resort to differential topology which imposes more restrictive conditions on the existence of *smooth* embeddings. Before we apply these to the concrete cases of pure states and states with general rank constraints we provide some general background:

Although for most smooth manifolds the minimal embedding dimension m is not known exactly, quite narrow intervals have been determined for many cases of interest (see [Jam71, Ada93] for an overview). One tool to derive lower bounds on the minimal m is Chern's result [Che48] that a smooth embedding into \mathbb{R}^m requires that the dual Stiefel-Whitney classes vanish $\bar{W}(\mathcal{M})_i = 0$ for all $i \geq m - d_{\mathcal{M}}$. Other bounds can be obtained from an index theorem due to Atiyah and Hirzebruch [AH59] and similar ideas in [May65, Sug79]. On the positive side a general upper bound is due to Whitney [Whi44] who showed that a smooth embedding $\mathcal{M} \rightarrow \mathbb{R}^m$ always exists if $m \geq 2d_{\mathcal{M}}$ (actually $m \geq 2d_{\mathcal{M}} - 1$ unless m is a power of 2). Whitney's bound is known to be optimal, i.e., in the worst case the dimension of the Euclidean space has to be twice the dimension of the manifold.

In order to apply lower bounds on dimensions for smooth embeddings to tomography we have as before to show that any informational complete h preserves topological properties—but now in the smooth category, i.e., such that \mathcal{M} is diffeomorphic to its image. This appears to be more subtle as before and we postpone the proof which covers the cases discussed below to Prop.1 in the appendix.

Obstructions on states with flat spectrum. We will now generalize the qubit example to pure states of arbitrary dimension. The manifold of pure states in \mathbb{C}^d can be identified with the *complex projective space* \mathbb{CP}^{d-1} which has real dimension $d_{\mathcal{M}} = 2d - 2$. The map from \mathbb{CP}^{d-1} to Hermitian rank-one projections is itself a smooth embedding. This, together with Prop.1 implies that non-embedding results for \mathbb{CP}^{d-1} equally apply to the tomography of pure quantum states. The probably best non-embedding result in this case can be found

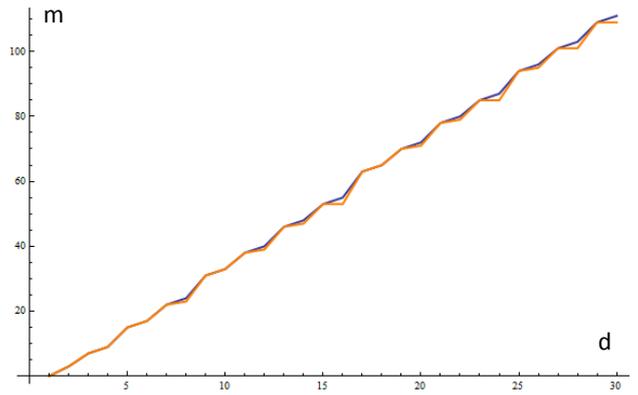


FIG. 3: Upper and lower bound for the minimal number m for which informational completeness for pure states in \mathbb{C}^d can be achieved. Note that the bounds coincide for $d < 8$ and that they differ by at most 2 until $d = 30$.

in [May65] which states that an embedding $\mathbb{CP}^{d-1} \rightarrow \mathbb{R}^m$ requires

$$m > \begin{cases} 2d_{\mathcal{M}} - 2\alpha & \forall d > 1, \\ 2d_{\mathcal{M}} - 2\alpha + 2 & d \text{ odd, and } \alpha = 3 \pmod{4} \\ 2d_{\mathcal{M}} - 2\alpha + 1 & d \text{ odd, and } \alpha = 2 \pmod{4}, \end{cases} \quad (2)$$

where α denotes the number of 1's in the binary expansion of $d - 1$, i.e., in particular $\alpha \leq \log_2(d)$. Note that this is almost $2d_{\mathcal{M}} = 4d - 4$, the worst case according to Whitney's general embedding result. For \mathbb{CP}^{d-1} the latter can be slightly improved [Ste70, Muk81] to the extent that embeddings are known with

$$m = \begin{cases} 2d_{\mathcal{M}} - \alpha & \forall d > 1, \\ 2d_{\mathcal{M}} - \alpha - 1 & \text{for even } d > 2. \end{cases} \quad (3)$$

A priori, there is no guarantee that such embeddings have a representation in the form of Eq.(1), i.e., that they can be realized by quantum measurements. Fortunately, they can as we will see in appendix VII. The upper bounds on m as stated in Eq.(3) then come with explicitly, albeit rather cumbersome, constructed observables (see Fig.3). Simpler constructions can be obtained for slightly weaker bounds as discussed in Sec.II.

Before we come to explicit constructions of measurement schemes, we will discuss the set of $d \times d$ density matrices which are proportional to a projection of dimension r , i.e., states which are maximally mixed within a subspace of dimension r . This set forms a smooth manifold of real dimension $d_{\mathcal{M}} = 2r(d - r)$ which is isomorphic to the *complex Grassmannian* manifold $G(r, d - r)$ [Dim96]. Again Prop.1 assures that non-embedding results carry over to measurement schemes and from [Sug79] we obtain a bound for embeddings $G(r, d - r) \rightarrow \mathbb{R}^m$ in the form of Eq.(2) but now with $d_{\mathcal{M}} = 2r(d - r)$ and $\alpha = \sum_{j=1}^r \beta(d - j) - \beta(j - 1)$ where $\beta(n)$ is the number of ones in the binary expansion of n . For $r = 1$ this coincides with the aforementioned bound

for pure states. Note that this provides a lower bound for informationally complete measurement schemes w.r.t. all sets which includes such a ‘‘Grassmannian manifold’’, like the set of density matrices with rank bounded by r .

II. UPPER BOUNDS AND EXPLICIT CONSTRUCTIONS

So far we discussed lower bounds on the number of measurement outcomes or settings. In this section will provide upper bounds and explicit constructions which show that the bounds are essentially tight.

We will in the following regard the space of Hermitian matrices in $\mathbb{C}^{d \times d}$ as a real vector space and identify it with \mathbb{R}^{d^2} . To start with a general positive result which is reminiscent of Whitney’s embedding theorem we allow to go beyond the framework of manifolds and let \mathcal{M} be any compact set of density matrices (regarded as a subset of \mathbb{R}^{d^2}) and assign a fractal dimension to it. The *Minkowski dimension* $D_{\mathcal{M}}$ is obtained by considering the minimal number $N_{\epsilon}(\mathcal{M})$ of ϵ -balls needed to cover \mathcal{M} and taking the limit

$$D_{\mathcal{M}} := \limsup_{\epsilon \rightarrow 0} \frac{\log(N_{\epsilon}(\mathcal{M}))}{\log(1/\epsilon)}. \quad (4)$$

If \mathcal{M} is a smooth manifold of real dimension $d_{\mathcal{M}}$, as all the sets discussed so far, then $D_{\mathcal{M}} = d_{\mathcal{M}}$. By Mane’s theorem [Man81, HK99] almost any (in the Lebesgue measure sense) linear map from \mathbb{R}^{d^2} into \mathbb{R}^m is injective on \mathcal{M} if $m > 2D_{\mathcal{M}}$. Viewing such a map as a real $m \times d^2$ matrix, we can identify a Hermitian matrix $A_i \in \mathbb{C}^{d \times d}$ with each of the m rows. Hence, if $m > 2D_{\mathcal{M}}$, then almost any measurement scheme is informationally complete w.r.t. \mathcal{M} . In principle, this bound can be refined to $m > \delta_{\mathcal{M}}$, where $\delta_{\mathcal{M}}$ is the *Hausdorff dimension* of the set $\mathcal{M} - \mathcal{M} = \{M_1 - M_2 | M_i \in \mathcal{M}\}$ [Rob09]. This bound is generally better since $\delta_{\mathcal{M} - \mathcal{M}} \leq D_{\mathcal{M} - \mathcal{M}} \leq 2D_{\mathcal{M}}$, but it may be more difficult to handle. We also mention that for the inverse mappings Hölder continuity can be proven and the respective constants can be bounded [BAEFN93, HK99].

For the case of pure states and states with more general rank constraints we will now improve on this and provide explicit constructions.

We will prove in the appendix that the following set of $m = 4d - 5$ operators is informationally complete for the manifold of pure state density operators: consider two types of matrices X_{α} and Y_{β} which we label by integers $\alpha = 1, \dots, 2d - 2$ and $\beta = 1, \dots, 2d - 3$ respectively. The X_{α} ’s are taken to be such that $(X_{\alpha})_{kl} = \delta_{k+l, \alpha+1}$, i.e., there are 1’s along the α ’th anti-diagonal and zeroes elsewhere. The Y_{β} ’s are similarly

defined with non-zero entries solely along the anti-diagonals, in this case $(Y_{\beta})_{kl} = 0$ unless $k + l = \beta + 2$. The entries are chosen such that the matrices are anti-symmetric with entries i below the diagonal. The set $\{A_i\} := \{X_{\alpha}, Y_{\beta}\}$ then forms a set of $4d - 5$ measurements whose expectation values allow to identify any pure state unambiguously. Similarly, after a suitable affine transformation we obtain a POVM with $4d - 4$ possible outcomes which is informationally complete for the manifold of pure states. A better but more cumbersome construction can be obtained from the work of Milgram [Mil67] as discussed in appendix VII.

Now consider the set \mathcal{M}_r of $d \times d$ density matrices with $\text{rank}(\rho) \leq r$. In this case we will follow a more indirect route in order to obtain an informationally complete measurement scheme. We are looking for a set of operators $\{A_i\}_{i=1, \dots, m}$ such that $\forall i : \text{tr}[A_i(\rho_1 - \rho_2)] = 0$ implies that $\rho_1 = \rho_2$ if both are elements of \mathcal{M}_r . The set of differences $(\rho_1 - \rho_2)$, however, equals up to a rescaling the set

$$\mathcal{S} := \{X = X^{\dagger} \in \mathbb{C}^{d \times d} | \text{tr}[X] = 0, \text{rank}(X) \leq 2r\}. \quad (5)$$

Suppose we have a linear subspace \mathcal{B} of matrices with the property that $\text{rank}(X) \geq 2r + 1$ for any nonzero $X \in \mathcal{B}$. Then the orthogonal complement \mathcal{B}^{\perp} of \mathcal{B} has exactly the property that $\text{tr}[A(\rho_1 - \rho_2)] = 0 \forall A \in \mathcal{B}^{\perp}$ only if $\rho_1 = \rho_2$. We will prove in the appendix that it is possible to construct \mathcal{B} with the required property and with $\dim \mathcal{B} = (d - 2r)^2$. With $\dim \mathcal{B}^{\perp} = d^2 - \dim \mathcal{B}$ and using that $\mathbb{1} \in \mathcal{B}^{\perp}$ we finally obtain that for any $r < d/2$

$$m = 4r(d - r) - 1, \quad (6)$$

is sufficient for \mathcal{M}_r . Note that for fixed r , the scaling in d again matches that of the lower bound up to an additive logarithmic term.

III. CONCLUSION

Regarding a measurement scheme supplemented by prior information as a mapping between topological spaces seems to be a new perspective which links optimality results in quantum tomography to non-trivial theorems in algebraic and differential topology. Clearly, this can be applied beyond the outlined cases of rank constraints. We note that the provided analysis focused on pointing out the limits and is thus complementary to other approaches which focus more on the efficiency of the reconstruction [GLF⁺10] or allow for cases of failure [FSC05].

[Ada93] Masahisa Adachi. *Embeddings and Immersions*. American Mathematical Society, 1993.

[AH59] M. F. Atiyah and F. Hirzebruch. Quelques theoremes de nonplongement pour les varietes differentiables. *Bull. Soc. Math. France*, 87:383–396, 1959.

[AW99] Jean-Pierre Amiet and Stefan Weigert. Reconstructing a pure state of a spin s through three stern-gerlach measurements. *J. Phys. A*, 32:2777 – 2784, 1999.

[BAEFN93] A. Ben-Artzi, A. Eden, C. Foias, and B. Nicolaenko. Hölder continuity for the inverse of mane’s projection. *Journal*

of *Math. Ann. Appl.*, 178:22–29, 1993.

- [Che48] Shiing-Shen Chern. On the multiplication in the characteristic ring of a sphere bundle. *Ann. Math.*, 49:362–372, 1948.
- [CMW08] Toby Cubitt, Ashley Montanaro, and Andreas Winter. On the dimension of subspaces with bounded schmidt rank. *J. Math. Phys.*, 49:022107, 2008.
- [Dim96] I. Dimitric. A note on equivariant embeddings of grassmannians. *Publ. Inst. Math. (Beograd)*, 59:131–137, 1996.
- [Dit95] Jochen Dittmann. On the Riemannian metric on the space of density matrices. *Rep. Math. Phys.*, 36:309–315, 1995.
- [FSC05] Steven T. Flammia, Andrew Silberfarb, and Carlton M. Caves. Minimal informationally complete measurements for pure states. *Foundations of Physics*, 35:1985–2006, 2005.
- [GLF⁺10] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, 2010.
- [HK99] Brian R. Hunt and Vadim Y. Kaloshin. Regularity of embeddings of infinite-dimensional fractal sets into finite-dimensional spaces. *Nonlinearity*, 12:12631275, 1999.
- [Jam59] I.M. James. Some embeddings of projective spaces. *Mathematical Proceedings of the Cambridge Philosophical Society*, 55:294–298, 1959.
- [Jam71] I.M. James. Euclidean models of projective spaces. *Bull. London Math. Soc.*, 3:257–276, 1971.
- [Lee02] John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2002.
- [Man81] R. Mane. *Dynamical Systems and Turbulence ar maps*, volume 898 of *Lecture Notes in Mathematics*, chapter On the dimension of the compact invariant sets of certain non-line, pages 230–242. Springer, 1981.
- [May65] K.H. Mayer. Elliptische Differentialoperatoren und Ganzahligkeitssätze für Charakteristische Zahlen. *Topology*, 4:295–313, 1965.
- [Mil67] R. James Milgram. Immersing projective spaces. *Ann. Math.*, 85:473–482, 1967.
- [Muk81] Mukherjee. Embedding complex projective spaces in euclidean space. *Bull. London Math. Soc.*, 13:323–324, 1981.
- [Rob09] James C Robinson. Linear embeddings of finite-dimensional subsets of banach spaces into euclidean spaces. *Nonlinearity*, 22(4):711, 2009.
- [Ste70] B. Steer. On the embedding of projective spaces in euclidean space. *Proc. London Math. Soc.*, 21:489–501, 1970.
- [Sug79] Tamio Sugawara. Non-immersion and non-embedding theorems for complex grassmann manifolds. *Proc. Japan Acad.*, 55:59–64, 1979.
- [Wei92] S. Weigert. *Phys. Rev. A*, 45:7688, 1992.
- [Whi44] H. Whitney. The self intersection of a smooth n-manifold in 2n-space. *Ann. of Math.*, 45:220, 1944.

IV. APPENDIX: MANIFOLDS AND THEIR EMBEDDINGS

Here we will discuss conditions under which the existence of a (smooth) embedding is implied by informational completeness, i.e., injectivity of the map $h : \mathcal{M} \rightarrow \mathbb{R}^m$ defined in Eq.(1). We use notions from differential topology as defined in [Lee02] and in the following understand embeddings and manifolds to be smooth unless otherwise stated. Embeddings which are not necessarily smooth are called *topological embeddings*. We will throughout suppose that \mathcal{M} is a *compact embedded submanifold* of \mathbb{R}^{d^2} where we identify the latter with the space of Hermitian matrices in $\mathbb{C}^{d \times d}$. With a slight

abuse of notation we will often write \mathcal{M} for both, the manifold and its inclusion in \mathbb{R}^{d^2} and similarly we write h for the map from \mathcal{M} as well as for the extended map from \mathbb{R}^{d^2} . We denote by $T_p(\mathcal{M})$ the tangent space of \mathcal{M} at $p \in \mathcal{M}$ and by $h_* : T_p(\mathcal{M}) \rightarrow T_{h(p)}(h(\mathcal{M}))$ the derivative, which is a linear map between the tangent spaces (sometimes call push-forward). The following cone will play an important role:

$$\Delta(\mathcal{M}) := \{X \in \mathbb{R}^{d^2} \mid X = \lambda(M_1 - M_2) \text{ for some } M_i \in \mathcal{M}, \lambda > 0\}. \quad (7)$$

Proposition 1 *Let \mathcal{M} be a compact embedded submanifold of \mathbb{R}^{d^2} , where we identify the latter with the space of Hermitian matrices in $\mathbb{C}^{d \times d}$, and define a map $h : \mathcal{M} \rightarrow \mathbb{R}^m$ as in Eq.(1). Then*

1. *h is a topological embedding iff it is injective,*
2. *if h is injective and for all $p \in \mathcal{M}$: $T_p(\mathcal{M}) \subseteq \Delta(\mathcal{M})$, then h is a smooth embedding,*
3. *$T_p(\mathcal{M}) \subseteq \Delta(\mathcal{M})$ holds for all $p \in \mathcal{M}$ if $\mathcal{M} = G(r, d - r)$ is the complex Grassmannian manifold understood as the submanifold in the space of $d \times d$ Hermitian matrices which consists of all orthogonal projections of rank r .*

Proof. 1. By definition a topological embedding is an injective continuous map which has a continuous inverse on its image. So in particular injectivity is implied by h being a topological embedding. For the converse note that h is linear on \mathbb{R}^{d^2} and thus continuous. Moreover, by assumption $h : \mathcal{M} \rightarrow h(\mathcal{M})$ is a continuous bijection and as such has a continuous inverse since \mathcal{M} is supposed to be compact.

2. We have to show that h is (i) a topological embedding, (ii) smooth and (iii) has an injective derivative everywhere. Clearly h is smooth and according to 1. it is a topological embedding if it is injective. Due to linearity of h on \mathbb{R}^{d^2} we have $h_* = h$ but we have to be careful with the domains in order to argue that injectivity of h (as a mapping from \mathcal{M}) implies injectivity of h_* (as a set of mappings from $T_p(\mathcal{M})$ for any $p \in \mathcal{M}$). By assumption, for any $p \in \mathcal{M}$ and $X \in T_p(\mathcal{M})$ we have $X \in \Delta(\mathcal{M})$. Then indeed $h_*(X) = 0$ together with injectivity of h implies $X = 0$ since $h_*(X) = \lambda(h(M_1) - h(M_2))$ is zero only if $M_1 = M_2$.

3. Let us first identify the tangent space at an arbitrary point $P \in \mathcal{M}$ which is now a Hermitian projector with $\text{tr}[P] = r$. Considering a curve within \mathcal{M} through P given by the unitary orbit $c(t) := e^{iHt} P e^{-iHt}$ for some Hermitian matrix H and $t \in \mathbb{R}$. The derivative $\partial_t c(t)|_{t=0} = i[H, P]$ is an element of $T_p(\mathcal{M})$ and in fact, such derivatives span the entire tangent space

$$T_p(\mathcal{M}) = \{X = X^\dagger \mid X = i[H, P] \text{ for some } H = H^\dagger\}. \quad (8)$$

In order to see this we have to show that they span a vector space which has the same dimension as the manifold (for which $d_{\mathcal{M}} = 2r(d - r)$). To this end, note that there is a one-to-one relation between commutators and block off-diagonal

matrices in the sense that we can always write

$$i[H, P] = \begin{pmatrix} 0 & C \\ C^\dagger & 0 \end{pmatrix}, \quad C \in \mathbb{C}^{r \times (d-r)}, \quad (9)$$

in the basis where $P = \mathbb{1} \oplus 0$. So the dimensions match, which verifies Eq.(8).

In a suitable basis any element $X \in T_P(\mathcal{M})$ is such that

$$X = \left[\bigoplus_{i=1}^r \begin{pmatrix} 0 & c_i \\ c_i & 0 \end{pmatrix} \right] \oplus 0_{d-2r}, \quad c_i \geq 0, \quad (10)$$

since Eq.(9) allows us to work with the singular values $\{c_i\}$ of C by transforming $X \mapsto (U \oplus V)X(U \oplus V)^\dagger$ with appropriate unitaries U and V . Setting $\lambda := \max_i c_i$ equal to the operator norm of X we complete the proof if we show that every 2×2 matrix of the form $c\sigma_x$ with $c \in [0, 1]$ is a difference of two projections. This can be seen to be true by taking the difference of two pure qubit states whose Bloch vectors are parameterized by $(c, \pm\sqrt{1-c^2}, 0)$. \square

V. APPENDIX: JAMES' CONSTRUCTION

Here we prove that the mentioned set of $4d - 5$ Hermitian operators $\{A_i\} := \{X_\alpha, Y_\beta\}$ is such that $\forall i : \text{tr}[A_i \rho_1] = \text{tr}[A_i \rho_2]$ implies that $\rho_1 = \rho_2$ if both are pure states. The construction is inspired by the embedding given in [Jam59] where the following type of auxiliary matrices appear. Consider a set of upper-triangular matrices $\{C_\gamma \in \mathbb{C}^{d \times d}\}_{\gamma=2, \dots, 2d}$ which are such that $(C_\gamma)_{kl} = 0$ if $k + l > \gamma$ and $(C_\gamma)_{kl} \neq 0$ if $k + l = \gamma$.

Lemma 1 *If for two vectors $x, y \in \mathbb{C}^d$ we have*

$$\forall \gamma : \langle x | C_\gamma | x \rangle = \langle y | C_\gamma | y \rangle, \quad (11)$$

then $y = e^{i\varphi} x$ for some $\varphi \in \mathbb{R}$.

Proof. Assume that the n 'th component of x is the first with a non-zero entry $x_n \neq 0$. Then $\langle x | C_{2n} | x \rangle = (C_{2n})_{n,n} |x_n|^2$ which by Eq.(11) implies that there is a $\varphi \in \mathbb{R}$ so that $y_l = e^{i\varphi} x_l$ holds for all $l \leq n$ (since $x_l = y_l = 0$ for $l < n$ and $|x_n| = |y_n|$). By induction we can now prove that the same proportionality has to hold for all other components. So assume for some $m \geq n$ that $y_l = e^{i\varphi} x_l$ holds for all $l \leq m$. Then for $\gamma = m + n + 1$

$$\begin{aligned} \langle x | C_\gamma | x \rangle &= \sum_{i+j \leq \gamma} (C_\gamma)_{ij} \bar{x}_i x_j \\ &= (C_\gamma)_{n, m+1} e^{i\varphi} \bar{y}_n x_{m+1} + \sum_{i+j \leq m+n} (C_\gamma)_{ij} \bar{y}_i y_j, \end{aligned} \quad (12)$$

where we replaced $x_l \rightarrow e^{-i\varphi} y_l$ for all $l \leq m$ and exploited that C_γ is upper triangular and that $x_l = 0$ for all $l < n$. Together with the hypothesis in Eq.(11) and the assumption that $(C_\gamma)_{n, m+1} \neq 0$ this implies indeed that $y_{m+1} = e^{i\varphi} x_{m+1}$. \square

As a consequence we obtain:

Proposition 2 *Consider the set of $4d - 5$ Hermitian operators $\{A_i\} := \{X_\alpha, Y_\beta\}$ as defined above. If for two vectors $x, y \in \mathbb{C}^d$ with $\|x\| = \|y\|$ we have $\langle x | A_i | x \rangle = \langle y | A_i | y \rangle$ for all i , then $|x\rangle\langle x| = |y\rangle\langle y|$.*

Proof. We use Lemma 1 for specific matrices which we construct as $C_{2d} = \mathbb{1}$ and for $\gamma = 2, \dots, 2d - 1$ as

$$(C_\gamma)_{kl} = \begin{cases} \delta_{k+l, \gamma}, & k < l, \\ 1/2, & k = l = \gamma/2 \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

Now note that $C_\gamma = (X_{\gamma-1} + iY_{\gamma-2})/2$ for $\gamma = 3, \dots, 2d - 1$ and $C_2 = X_1/2$. Hence, the identity in Eq.(11) for $\gamma = 2, \dots, 2d - 1$ is guaranteed by $\forall i : \langle x | A_i | x \rangle = \langle y | A_i | y \rangle$ and for $\gamma = 2d$ it holds due to $\|x\| = \|y\|$. \square

VI. APPENDIX: RANK- r CONSTRUCTION

In this appendix we outline the construction of a subspace \mathcal{B} of $d \times d$ matrices with the properties that

- $\mathcal{B}^\dagger = \mathcal{B}$,
- $\text{tr}[T] = 0$ for every $T \in \mathcal{B}$,
- $\dim \mathcal{B} = (d - 2r)^2$,
- $\text{rank}(T) \geq 2r + 1$ for every nonzero $T \in \mathcal{B}$.

The main part of our construction follows [CMW08] to which we refer for further details. The following fact will be needed. Let M be a totally nonsingular $m \times m$ -matrix with real entries. It can be, for instance, a Vandermonde matrix of the form

$$M = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{m-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_m & \alpha_m^2 & \cdots & \alpha_m^{m-1} \end{pmatrix}$$

with $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_m$. As explained in Lemma 9 in [CMW08], any linear combination of ℓ columns of M contains at most $\ell - 1$ zero elements.

For each $2r + 1 \leq k \leq d - 1$, we build up $k - 2r$ matrices as follows. We choose $k - 2r$ columns from a totally nonsingular $k \times k$ -matrix and we put them to the k th diagonal. Any linear combination of these $k - 2r$ matrices has at least $2r + 1$ nonzero elements, hence the rank is at least $2r + 1$. We also take all transposes of these matrices to our spanning set of matrices.

For the main diagonal we take, again, $d - 2r$ columns from a totally nonsingular $d \times d$ -matrix. Let v_1, \dots, v_{d-2r} denote these column vectors. We want to guarantee that the resulting matrices are traceless and therefore we modify these vectors. We choose a real vector u which has no zero entries and which is orthogonal to every v_1, \dots, v_{d-2r} . The vector u can for instance be taken as the last row vector of the inverse of

the Vandermonde matrix; this satisfies orthogonality by construction. Moreover, since the Vandermonde matrix is totally non-singular (i.e., all its minors are non-vanishing), its inverse has no zero entry because the latter are just multiples of minors (cofactors).

The new vectors $\tilde{v}_1, \dots, \tilde{v}_{d-2r}$ are then the entrywise products of v_1, \dots, v_{d-2r} with u .

In total, we have build up $d - 2r + 2 \sum_{k=2r+1}^{d-1} (k - 2r) = (d - 2r)^2$ linearly independent matrices.

VII. APPENDIX: MILGRAM'S CONSTRUCTION

Here we discuss how and why the bound in Eq.(3) can be realized via a proper measurement scheme. The construction of the observables is rather cumbersome and based on the work of Milgram [Mil67]. So will only argue why this corresponds to a proper measurement scheme rather than reproducing the construction.

For m as in Eq.(3) Milgram constructed a set of m bilinear maps, i.e., matrices $A_j \in \mathcal{M}_d(\mathbb{C})$, $j = 1, \dots, m$ which have the following properties:

(i) *Vanishing real inner product* in the sense that for all $x \in \mathbb{C}^d$ we have $\langle x, A_j x \rangle_{\mathbb{R}} = 0$ for the real inner product $\langle x, y \rangle_{\mathbb{R}} = \Re \langle x, y \rangle$. That is, each A_j has to be skew-symmetric w.r.t. the real inner product and thus anti-Hermitian w.r.t. to the standard complex inner product. In order to see the latter note that $\langle x, A_j y \rangle_{\mathbb{R}} + \langle y, A_j x \rangle_{\mathbb{R}} = 0 \forall x, y \in \mathbb{C}^d$ can be written as

$$\Re \langle \psi | \begin{pmatrix} 0 & A_j \\ A_j & 0 \end{pmatrix} | \psi \rangle = 0, \quad \forall \psi \in \mathbb{C}^{2d}. \quad (14)$$

Hence, the Hermitian part of the anti-diagonal block matrix has to vanish which indeed means that $A_j^\dagger = -A_j$.

(ii) *Completeness*. As noted by Mukherjee [Muk81] we can define matrices $T_j = iA_j$, which according to (i) are Hermitian, such that the map $f : \mathbb{C}^d \rightarrow \mathbb{R}^m$ defined via $f(x)_j = \langle x, T_j x \rangle$ has the property that $f(x) = f(y)$ implies that x is proportional to y .

The set of m Hermitian matrices T_j therefore leads to a measurement scheme which is informationally complete w.r.t. the set of pure states.