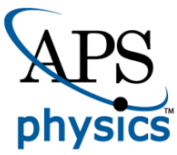

Efficient Measurement of Quantum Dynamics via Compressive Sensing

A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White

Phys. Rev. Lett. **106**, 100401 — Published 7 March 2011

DOI: 10.1103/PhysRevLett.106.100401

Efficient measurement of quantum dynamics via compressive sensing

A. Shabani,[1] R. L. Kosut,[2], M. Mohseni[3], H. Rabitz[1], M. A. Broome[4], M. P. Almeida[4], A. Fedrizzi[4], and A. G. White[4]

[1]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA,*
[2]*SC Solutions, Sunnyvale, California 94085, USA,* [3]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA,*
[4]*Centre for Engineered Quantum Systems and Centre for Quantum Computer and Communication Technology, School of Mathematics and Physics, The University of Queensland, QLD 4072, Australia*

The resources required to characterise the dynamics of engineered quantum systems—such as quantum computers and quantum sensors—grow exponentially with system size. Here we adapt techniques from compressive sensing to exponentially reduce the experimental configurations required for quantum process tomography. Our method is applicable to dynamical processes that are known to be nearly-sparse in a certain basis and it can be implemented using only single-body preparations and measurements. We perform efficient, high-fidelity estimation of process matrices on an experiment attempting to implement a photonic two-qubit logic-gate. The data base is obtained under various decoherence strengths. We find that our technique is both accurate and noise robust, thus removing a key roadblock to the development and scaling of quantum technologies.

Understanding and controlling the world at the nanoscale—be it in biological, chemical or physical phenomena—requires quantum mechanics. It is therefore essential to characterize and monitor *realistic* complex quantum systems that inevitably interact with typically uncontrollable environments. One of the most general descriptions of open quantum system dynamics is a quantum map—typically represented by a *process matrix* [1]. Methods to identify the process matrix are collectively known as *quantum process tomography* (QPT) [1, 2]. For a d-dimensional quantum system, they require $\mathcal{O}(d^4)$ experimental configurations: combinations of input states, on which the process acts, and a set of output observables. For a system of n of the simplest quantum objects, namely qubits —two-level quantum systems—$d{=}2^n$. The required physical resources hence scale *exponentially* with system size. In principle, a single generalized measurement is sufficient for full process tomography in a extended Hilbert space relying on highly nonlocal many-body measurements that are physically unavailable [3]. Recently, a number of alternative methods have been developed for efficient and selective estimation of quantum processes [4]. However, full characterization of quantum dynamics of comparably small systems, such as a recently demonstrated 8-qubit ion trap [5], would still require over a billion experimental configurations, clearly a practical impossibility. So far, process tomography has therefore been limited by experimental, and—to a lesser extent—by off-line computational resources, to systems of 2 and 3 qubits [6–8].

Here we adapt techniques from *compressive sensing* (CS) to develop an *experimentally efficient* method for QPT. It requires only $\mathcal{O}(s\log d)$ configurations if the process matrix is *s-compressible* in some known basis, *i.e.*, it is *nearly sparse* in that it can be well approximated by an s-sparse process matrix. This is usually the case, because engineered quantum systems aim to implement a *unitary* process which is maximally-sparse in its eigenbasis. In practice, as observed in QPT experiments in liquid-state NMR [9], photonics [6, 10, 11], ion traps [12], and superconducting circuits [7], a near-unitary process will still be nearly-sparse in this basis, and still compressible. The near sparsity emanates from decoherence originating in few dominant system-environment interactions. This is more apparent for weakly decohering systems [13].

We experimentally demonstrate our algorithm by estimating the 240 real parameters of the process matrix of a canonical photonic two-qubit gate, Fig. 1, from a reduced number of configurations. For example, from just 18 and 32 configurations, we obtain fidelities of 94% and 97% with process matrices obtained from an overcomplete set of all 576 available configurations.

Compressive sensing provides methods for compression of information carried by a large-size signal into a significantly smaller one along with efficient convex optimization algorithms to decipher this information [14, 15]. Originally developed to exploit compressible features of natural audio and video signals, applications of compressive sensing have recently found their way to quantum tomography: Simulations of compressive sensing for QPT [16], application to ghost-imaging [17], and quantum state tomography for *low-rank* density matrices [18]. The latter provides a quadratic reduction of physical resources compared to standard state tomography, *i.e.*, for a density matrix of rank r, $\mathcal{O}(rd\log^2 d)$ vs. standard d^2 settings, and it also has the main advantage that rank is basis independent. Very recently, the low-rank matrix completion technique has been used in an efficient method of quantum state tomography that in particular can be applied to one-dimensional physical systems whose states are well approximated by matrix product states [19].

Under reasonable assumptions, a quantum map on a d-dimensional space has the general representation [1],

$$\mathcal{S}(\rho){=}\sum_{\alpha,\beta=1}^{d^2}\chi_{\alpha\beta}\Gamma_\alpha\rho\Gamma_\beta^\dagger \qquad (1)$$

where χ, the $d^2\times d^2$ process matrix, is positive semidefinite, $\chi{\geq}0$, and trace preserving, $\sum_{\alpha,\beta}\chi_{\alpha\beta}\Gamma_\beta^\dagger\Gamma_\alpha{=}I_d$, with $\{\Gamma_\alpha\}$

an orthonormal matrix basis set, $\mathbf{Tr}(\Gamma_\beta^\dagger\Gamma_\alpha)=\delta_{\alpha\beta}$. Note that sparsity is a property of the map representation not the map itself. Data is collected by preparing an ensemble of identical systems in one of the states $\{\rho_1,...,\rho_k\}$, inputting them to the process χ, and then measuring an observable M from the set $\{M_1,...,M_\ell\}$. For a pair (ρ,M), the outcome will be $y_{M,\rho}=\mathbf{Tr}(\mathcal{S}(\rho)M)$. If the experiment is repeated for all configurations, *i.e.*, $(\rho_i,M_i), i=1,\ldots,m=k\ell$, the relation between the vector of outcomes $y=[y_{M_1,\rho_1},\ldots,y_{M_m,\rho_m}]^T$ and the true process matrix, denoted by χ_0, can be represented by a linear map $y=\Phi\vec{\chi}_0$, where $\vec{\chi}_0$ is the vectorized form of the process matrix χ_0 and Φ is an $m\times d^4$ matrix of coefficients of the form $\mathbf{Tr}(\Gamma_\alpha\rho_i\Gamma_\beta^\dagger M_i)/\sqrt{m}$.

In general, estimating a sparse process matrix with an unknown sparsity pattern from an underdetermined set of linear equations ($m<d^4$) would seem highly unlikely. Compressive sensing, however, tells us that this can be done by solving for χ from the convex optimization problem:

$$\text{minimize } \|\vec{\chi}\|_{\ell_1} \text{ subject to } \|y-\Phi\vec{\chi}\|_{\ell_2}\leq\varepsilon, \qquad (2)$$

and positive-semidefinite and trace-preserving conditions as defined above. The parameter ε quantifies the level of uncertainty in the measurements, that is, we observe $y=\Phi\chi_0+w$ with $\|w\|_{\ell_2}\leq\varepsilon$. From [15, 21], recovery via (2) is ensured if (i) the matrix Φ satisfies the *restricted isometry property*:

$$1-\delta_s\leq\frac{\|\Phi\vec{\chi}_1(s)-\Phi\vec{\chi}_2(s)\|_{\ell_2}^2}{\|\vec{\chi}_1(s)-\vec{\chi}_2(s)\|_{\ell_2}^2}\leq 1+\delta_s \qquad (3)$$

for all s-sparse $\chi_1(s)$, $\chi_2(s)$ process matrices; (ii) the *isometry constant* $\delta_{2s}<\sqrt{2}-1$ and (iii) the number of configurations $m\geq C_0 s\log(d^4/s)$. Under these conditions, the solution $\chi^\star$ of (2) satisfies,

$$\|\vec{\chi}^\star-\vec{\chi}_0\|_{\ell_2}\leq\frac{C_1}{\sqrt{s}}\|\vec{\chi}_0(s)-\vec{\chi}_0\|_{\ell_1}+C_2\varepsilon \qquad (4)$$

where $\chi_0(s)$ is the best s-sparse approximation of χ_0 and C_0, C_1, C_2 are constants on the order of $\mathcal{O}(\delta_s)$, see [20]. The restricted isometry property states that two s-sparse process matrices $\chi_1(s)$ and $\chi_2(s)$ can be distinguished if their relative distance is nearly preserved after the measurements, *i.e.*, under transformation by Φ. If the measurements are noise free, $\varepsilon=0$, and the process matrix is actually s-sparse, $\chi_0=\chi_0(s)$, then the right hand side of (4) is zero leading to perfect recovery, $\chi^\star=\chi_0$. Otherwise the solution tends to the best s-sparse approximation of the process matrix plus the additional term due to measurement error ε. If for an n-qubit QPT with $d=2^n$ the conditions of the above analysis are satisfied, then the number of experimental configurations m scales *linearly* with sn, specifically, $m\geq C_0 s(4n\log 2-\log s)=\mathcal{O}(sn)$. In [20], using the measure concentration properties of random matrices, following the arguments in [15, 21], we show that if Φ is constructed from random input states $\{\rho_i\}$, and random observables $\{M_i\}$, then the restricted isometry in (3) holds with high probability. Also a test is presented to certify the sparsity assumption.

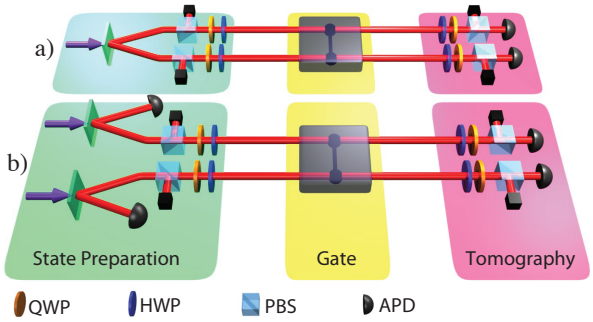

FIG. 1: Experimental scheme. Two-photon inputs were prepared with either (a) a high-rate, non-scalable, two-photon source or (b) a low-rate, scalable, four-photon source. The qubits are encoded using polarisation, as described in the text. The quantum process is a photonic entangling-gate. A measurement configuration is defined as some combination of state preparation and tomography, implemented here with quarter- and half- waveplates (QWP, HWP), polarizers (PBS), and photon detectors (APD). For details see [20].

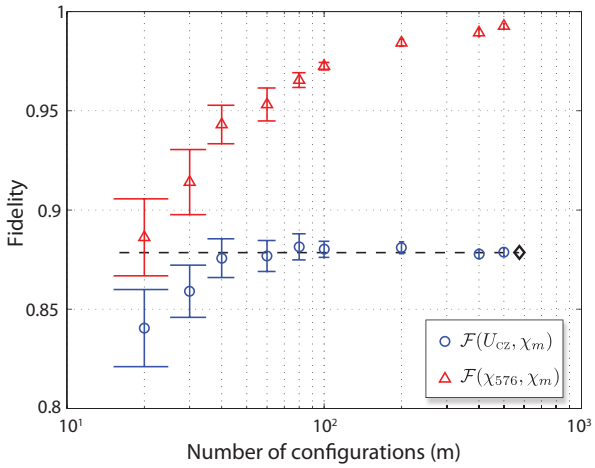

FIG. 2: Process fidelities vs. number of input-output configurations, for each compressive QPT estimate, χ_m, in the gate-basis of the ideal CZ-gate for the lowest measured noise level, $\mathcal{P}=0.91$. The dashed line shows the fidelity of the full estimate $\mathcal{F}(U_{CZ},\chi_{576})=0.89$ (black diamond). Error bars are obtained by solving (2) for 50 different random combinations of m inputs and observables.

A nearly-sparse process matrix can thus be encoded into an *exponentially smaller* number of measurement outcomes, which can be recovered to within the bounds of (4) by solving (2). We now test our algorithm experimentally against standard QPT on a two-qubit gate under a range of decoherence—and thus sparsity—conditions. We used a photonic controlled-phase, CZ, gate, Fig. 1 where the qubits are encoded in orthogonal polarization states of single photons ($|H\rangle$, horizontal, and $|V\rangle$, vertical). We performed full process tomography [6, 10, 11] of the gate with both 2-photon and 4-photon arrangements, preparing 16 pair-wise combinations of the 4 input states $\{|H\rangle,|V\rangle,|D\rangle,|R\rangle\}$ and, for each input, measuring 36 two-qubit combinations of the observables $\{|H\rangle,|V\rangle,|D\rangle,|A\rangle,|R\rangle,|L\rangle\}$,

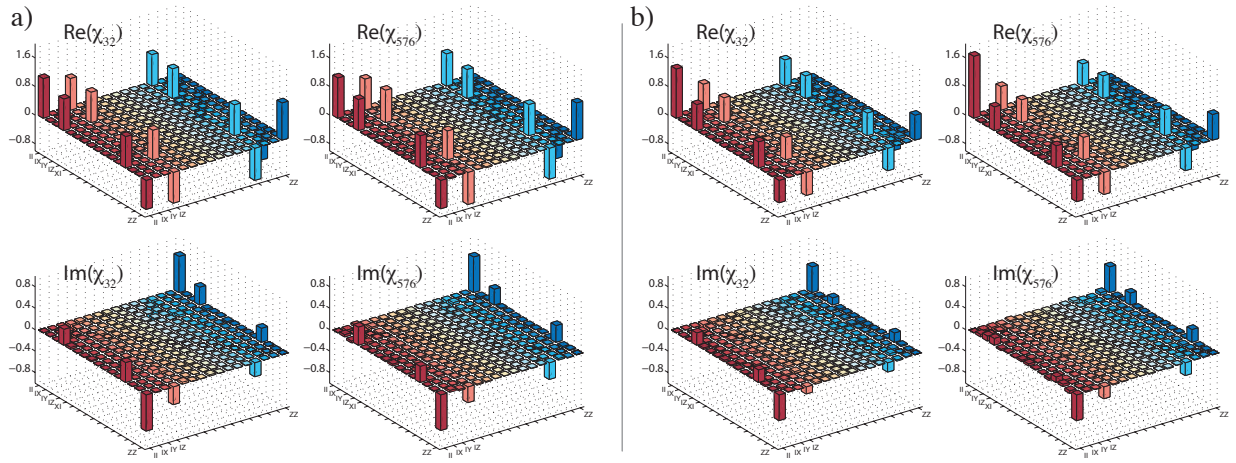

FIG. 3: Real and imaginary process matrix elements in the Pauli basis for the CQPT estimate χ_{32}, 32 configurations (left) vs. full data estimate χ_{576}, 576 configurations (right) for A) a low noise, 2-photon experiment, $\mathcal{P}$=0.91, and B) a high-noise, 4-photon experiment, $\mathcal{P}$=0.62. The CQPT reconstructions have fidelities, $\mathcal{F}(\chi_{576},\chi_{32})$, of 95% and 85% respectively. The CQPT estimation accuracy is excellent for low noise, and reliable even for high noise, see [20] for more details.

where $|D\rangle$=($|H\rangle$+$|V\rangle$)/$\sqrt{2}$, $|A\rangle$=($|H\rangle$−$|V\rangle$)/$\sqrt{2}$, $|R\rangle$=($|H\rangle$+$i|V\rangle$/$\sqrt{2}$, and $|L\rangle$=($|H\rangle$−$i|V\rangle$/$\sqrt{2}$. These 576 input-output configurations represent an overcomplete set which allows the best possible estimate of the quantum process, denoted χ_{576} [6].

The compressed quantum process tomography (CQPT) estimate of the 16×16 process matrix, denoted χ_m, is obtained by solving (2) with y=$C_{\rm sel}p$ and Φ=$C_{\rm sel}G$ where p is the 576 × 1 experimental probabilities corresponding to each of the 576 configurations, G is the 576×256 matrix obtained from all the configurations and the basis set [Γ_α in (1)],and $C_{\rm sel}$ is the m×576 matrix corresponding to taking a selection of $m \leq 576$ of all possible configurations. The basis set is obtained from the singular-value decomposition of the ideal CZ-gate: the process matrix in this basis is maximally sparse with a *single* non-zero 1,1-element which equals 4. The measurement error bound ε in (2) is chosen to be just slightly larger than $\sqrt{m}\sigma$, where σ is the minimum feasible root-mean-square level obtained from (2) using all configurations, *i.e.*, with $C_{\rm sel}$=I_{576}. We quantify decoherence using the *process purity*, $\mathcal{P}$=$\mathbf{Tr}(\chi_m^2/d^2)$, which varies from 0 for a completely decohering channel, to 1 for a unitary process: in our experiment we used six decoherence levels (see [20] for details), giving purities of $\{0.62, 0.74, 0.77, 0.79, 0.86, 0.91\} \pm 0.01$.

Figure 2 shows, for the lowest decoherence level, the *process fidelities* [6] versus the number of randomly-selected configurations, m. Each process matrix, $\{\chi_m\}$, is obtained by solving (2). We use the fidelity between (i) the compressive measurement and the ideal, $\mathcal{F}(U_{\rm CZ},\chi_m)$; and (ii) the compressed and optimal measurements, $\mathcal{F}(\chi_{576},\chi_m)$. Note that as m increases the fidelity with the ideal converges to the value of 0.89 obtained from χ_{576}; likewise, the fidelity with the full estimate converges to unity. Similar plots exist for every level of decoherence, with fidelities reduced accordingly.

We have so far used random selections of probabilities from the full data set, which allows us a comprehensive test of compressive sensing theory. Experiments, however, don't yield probabilities but physical quantities, e.g. count rates. To date, algorithms for more efficient state [18] or process tomography have assumed probabilities as a starting point. Since normalization is an issue to some extent in all physical architectures, it will be necessary to investigate the robustness and scalability of algorithms for real-world experiments.

For our photonic two-qubit gate, which is lossy and intrinsically probabilistic, the probabilities were obtained by normalising counts using a full basis set of observables extracted from all measurements, I_{576}. Having sufficient configurations to allow for normalisation necessarily imposes limits on CQPT efficiency: for low m, we are restricted in how random our selections can be. (Details and some permissible configurations in [20]). As an example, Fig. 3 shows process matrices reconstructed via CQPT from just one of these configurations compared to the respective full data estimates. We used 32 combinations of the 16 inputs $\{|H\rangle,|V\rangle,|D\rangle,|R\rangle\}$ and 2 observables $\{|R\rangle|I\rangle,|I\rangle|R\rangle\}$, where I is the identity. The agreement is excellent as one can see from the fidelities and the correct reproduction of imaginary elements—which are ideally zero. Another striking feature is that we obtain highly faithful reconstructions of a non-local process using only *local* measurements [3].

A further crucial test is whether CQPT enables us to locate errors and implement necessary corrections: a common example is identifying local rotations that move the process closer to the ideal. By optimising $\mathcal{F}(U_{\rm CZ},\chi_{32})$, we calculated local corrections to χ_{32}; applying them to the full estimate χ_{576}, $\mathcal{F}(U_{\rm CZ},\chi_{576})$ improved, on average, over all decoherence levels, by 4.1%. This is very close to the average 4.9% improvement obtained by calculating and applying local corrections *directly* to χ_{576}. Even a low-configuration CQPT estimate of a noisy process therefore enables improvements.

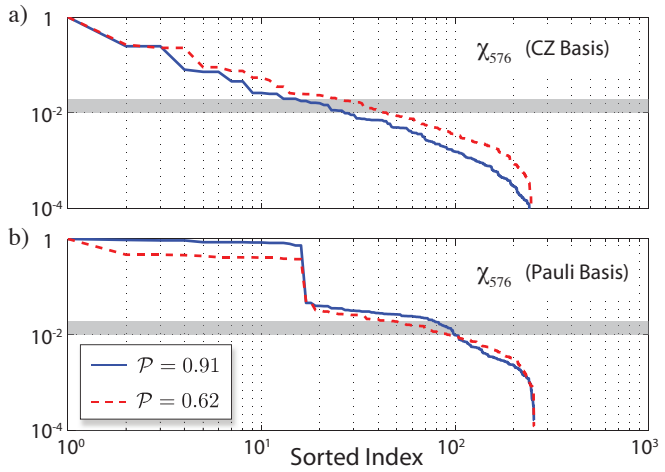

FIG. 4: Absolute values of the 256 process matrix elements of χ_{576} for our lowest and highest noise level, sorted by relative magnitude (with respect to the 1,1-element) in the CZ basis (top) and the Pauli basis (bottom). The error threshold, which indicates the required number of configurations, is shown in grey.

That high-fidelity estimates are obtained by CQPT can be understood from the error bound (4) which shows that the CQPT estimate tends towards the best s-sparse approximation of the true process, in this case our best estimate χ_{576}. Fig. 4 shows the process matrix elements, sorted by relative magnitude, for low and high noise levels, in two basis sets. The s-sparse approximation levels indicated in (4) are reached where the matrix elements drop below the error threshold $(0.01$–$0.02)$. For the corresponding m, we can therefore expect a successful, high-fidelity, CQPT reconstruction. In the CZ-basis, the plots show that for low noise, $s\in[20, 30]$, which correlates well with the fidelities in Fig. 2; for high noise $s\in[40, 60]$. Although the process matrix is still somewhat sparse in the Pauli-basis (Fig. 3), the corresponding plots in Fig. 4 indicate that $\sim$100 configurations are needed to obtain an estimate of comparable quality. Furthermore, the sorted magnitude values in the CZ basis decay exponentially, which is sufficient to declare the process matrix s-compressible, *e.g.*, [22, 23]. Intriguingly, this exponential decay is a signature of *model-based compressive sensing* where the scaling goes from $m{=}\mathcal{O}(s\log(d/s))$ to $m{=}\mathcal{O}(s)$ [23]. This demands further investigation, since it appears that QPT fits this framework, particularly when the process matrix is expanded in the ideal basis corresponding to the unitary design goal.

Our experimental results are supported numerically by simulations of a 2-qubit process as well as simulation studies for 3 and 4 qubit systems which show the same type of compressibility, see [20]. Applying CQPT to larger systems will require careful attention to classical post-processing which—as in QPT—scales exponentially. The standard software we used here (see [20]), can easily handle 2 and 3 qubit CQPT systems. For larger systems, more specialized software can increase speed by orders of magnitude, *e.g.*, [22].

A number of research directions arise from this work: incorporating knowledge of model structure properties; tightening the bounds on scaling laws; understanding how near-sparsity s and rank r vary with system dimension, d; pursuing highly efficient convex-computational algorithms; and selection of optimal configurations. Compressive tomography techniques can also be applied to quantum metrology and Hamiltonian parameter estimation: for example, estimating selective properties of biological or chemical interest in molecular systems and nanostructures with typically sparse Hamiltonians [24].

We thank J. Romberg and S. Jafarpour for discussions. We acknowledge funding by the ARC Discovery and Federation Fellow programs and an IARPA-funded US Army Research Office contract. RLK and HR are supported by DARPA Grant FA9550-09-1-0710 administered through AFOSR.

[1] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).
[2] J. B. Altepeter *et al.*, Phys. Rev. Lett. **90**, 193601 (2003);M. Mohseni and D. Lidar, Phys. Rev. Lett. **97**, 170501 (2006); M. Lobino *et al.*, Science **322**, 563 (2008).
[3] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Phys. Rev. A **77**, 032322 (2008).
[4] J. Emerson *et al.*, Science **317**, 1893 (2007);A. Bendersky *et al.* , Phys. Rev. Lett. **100**, 170501 (2006); M. Branderhorst *et al.*, New Journal of Physics **11**, 115010 (2009).
[5] H. Haffner *et al.*, Nature **438**, 643 (2005).
[6] J. O'Brien *et al.*, Phys. Rev. Lett. **93**, 080502 (2004).
[7] R. Bialczak *et al.*, et al., Nature Physics **6**, 409 (2010).
[8] T. Monz *et al.*, Phys. Rev. Lett. **102**, 040501 (pages 4) (2009).
[9] A. M. Childs *et al.*, Phys. Rev. A **64**, 012314 (2001);N. Boulant *et al.* , Phys. Rev. A **67**, 042322 (2003); Y. Weinstein *et al.*, J. Chem. Phys. **121**, 6117 (2004).
[10] M. W. Mitchell *et al.*, Phys. Rev. Lett. **91**, 120402 (2003).
[11] Y. Nambu and K. Nakamura, Phys. Rev. Lett. **94**, 010404 (2005).
[12] M. Riebe *et al.*, Phys. Rev. Lett. **97**, 220407 (2006).
[13] M. Mohseni and A. T. Rezakhani, Phys. Rev. A **80**, 010101 (2009);A. G. Kofman and A. N. Korotkov , Phys. Rev. A **80**, 042103 (2009).
[14] E. Candes and M. Wakin, IEEE Sig. Proc. Mag. **25**, 21 (2008).
[15] E. J. Candes and T. Tao, IEEE Inform. Theory News. **Dec. 14** (2008).
[16] R. L. Kosut, arXiv:0812.4323v1[quant-ph] (2008).
[17] O. Katz *et al.*, Applied Physics Letters **95**, 131110 (2009).
[18] D. Gross *et al.*, Phys. Rev. Lett. **105**, 150401 (2010).
[19] M. Cramer *et al.*, Nat. Commun. **1**, 149 (2010).
[20] Online material,
[21] E. J. Candes, Compte Rendus de l'Academie des Sciences, Paris, Serie I **346**, 589 (2008).
[22] D. Needell and J. A. Tropp, Appl. Comp. Harmonic Anal. pp. 301–321 (2008).
[23] R. G. Baraniuk *et al.*, IEEE Transactions on Information Theory **56**, 1982 (2010).
[24] J. Yuen-Zhou *et al.*, Arxiv preprint arXiv:1006.4866 (2010).
[25] R. Baraniuk *et al.*, Constructive Approximation **28**, 253 (2008).
[26] R. Okamoto *et al.*, Phys. Rev. Lett. **95**, 210506 (2005).

[27] N. Kiesel *et al.*, Phys. Rev. Lett. **95**, 210505 (2005).
[28] M. Barbieri *et al.*, Journal of Modern Optics **56**, 209 (2009).