A quantum algorithm to efficiently sample the work distribution and to estimate the free energy of quantum systems

Augusto J. Roncaglia,1,2 Federico Cerisola,1 and Juan Pablo Paz1,2

1Departamento de Física, FCEyN, UBA, Ciudad Universitaria Pabellón 1, 1428 Buenos Aires, Argentina
2IFIBA CONICET, UBA, FCEyN, UBA, Ciudad Universitaria Pabellón 1, 1428 Buenos Aires, Argentina

(Dated: 12 September 2014 – submitted)

We present a new method to measure work and to efficiently sample its probability distribution with fixed precision. The method can be used to estimate free energies on a quantum computer. It is based on the fact that work measurement is a generalized measurement.–

**Introduction.** – One of the most interesting results of recent decades in statistical mechanics is the derivation of fluctuation theorems [1, 2]. Notably, Jarzynski and Crooks identities relate non-equilibrium work and free energies differences between equilibrium states. These theorems are very useful to evaluate free energies at the nanoscale [3]. For quantum systems the definition of work is rather subtle as it is not represented by a hermitian operator valued measure (POVM) as they form a set of positive operator valued measures (POVM) [6]. We present a new method to measure work at a single time on an enlarged system. (i) Work can be estimated using a variant of the “phase estimation algorithm” which is such that work is detected as the outcome of the single time measurement with probability \( P(w) \). (ii) The efficient sampling of \( P(w) \) can be combined with fluctuation theorems to estimate differences between the free energy of quantum states.

PACS numbers: 05.70.Ln, 03.67.Ac

The measurement of the final probability of work \( w = E_f - E_i \) is typically larger than the dimension of the space of states. Therefore, a Hermitian operator representing work cannot exist. This does not imply, of course, that work is not measurable. On the contrary, a natural strategy to measure work performed on a system that initially has a density matrix \( \rho(t_0) \) and is driven from an initial Hamiltonian \( H = H(t_f) \) to a final one \( \tilde{H} = H(t_f) \) is the following [5]: We can measure the energy at two times \( t_i \) and \( t_f \). Work is then defined as \( w = \tilde{E}_m - E_n \), where \( \tilde{E}_m \) and \( E_n \) are eigenvalues of \( H \) and \( \tilde{H} \) (i.e., they satisfy \( H|\phi_n\rangle = E_n|\phi_n\rangle \) and \( \tilde{H}|\tilde{\phi}_m\rangle = \tilde{E}_m|\tilde{\phi}_m\rangle \)). Thus, work \( w \) is measured with probability

\[
P(w) = \sum_{n,m} p_n p_{m,n} \delta(w - (\tilde{E}_m - E_n)),
\]

where \( p_n = \langle \phi_n | \rho(t_0) | \phi_n \rangle \) is the probability to initially obtain the energy \( E_n \) and \( p_{m,n} = |\langle \phi_n | U_{f,i} | \tilde{\phi}_m \rangle|^2 \) is the transition probability between the energy eigenstates when the system is driven by the evolution operator \( U_{f,i} = U(t_f, t_0) \). From the above expression for \( P(w) \), the Jarzynski equality [11] can be derived as follows: The characteristic function \( \chi(s) = \int dw P(w) \exp(-iws) \), is such that \( \chi(-i\beta) = \langle \exp(-\beta w) \rangle \). In turn, using that initially the state is thermal \( \rho(t_0) = \exp(-\beta H)/Z_0 \), \( \chi(-i\beta) \) can be computed and the remarkable identity \( \langle \exp(-\beta w) \rangle = \exp(-\beta \Delta F) \) can be proved, where \( F \) is the Helmholtz free energy. Similarly, from eq. (1) one can obtain Crooks identity [6] which involves the work distribution \( P_F(w) / P_B(w) \) for a forward (backward) process. It reads as: \( P_F(w) / P_B(w) = \exp(\beta(w - \Delta F)) \).

*Work measurement is a generalized measurement.* – We can rewrite eq. (1) as \( P(w) = \text{Tr}[\tilde{\rho} W(w)] \) where

\[
W(w) = \sum_{n,m} p_{m,n} \delta(w - E_{m,n}) |\phi_n\rangle \langle \phi_n|,
\]

with \( E_{m,n} \equiv (\tilde{E}_m - E_n) \). Operators \( W(w) \) define a positive operator valued measure (POVM) as they form a set...
of non-negative operators which decompose the identity as \( \int dw W(w) = I \) (the operators \( W(w) \) are not orthogonal since the number of values that \( w \) can take is larger than dimension of the Hilbert space). As any POVM \([14]\) can be realized as a projective measurement on an enlarged system, it is always possible to design an apparatus such that: (i) it produces an output \( w \) with probability \( P(w) \); (ii) when \( w \) is recorded, the system is prepared in a state \( \rho_w \) (that depends on \( \rho, w \), and on the measurement implementation). For instance, a simple method to measure work is to couple the system \( S \) with an auxiliary system \( A \) in such a way that \( A \) keeps a record of the energy of \( S \) at two times. To do this, \( S \) and \( A \) must interact twice through a measurement interaction (MI). This interaction is described by the Hamiltonian \( H_{MI} = \lambda \hat{H} \otimes \hat{p} \), where, \( \lambda \) is a constant and \( \hat{p} \) is the generator of translations between the pointer states \( |w \rangle \) of \( A \). In the simplest case we can consider \( A \) with a continuous degree of freedom, where \( \{ |w \rangle, w \in \mathbb{R} \} \) is a basis of its states of states. The MI has an evolution operator \( U_{MI} = \exp(-i H_{MI} t) \) which is such that

\[
U_{MI}(|\phi_n\rangle \otimes |w = 0\rangle) = |\phi_n\rangle \otimes |w = E_n\rangle. \tag{3}
\]

Then, if \( w = E_n \) is recorded in \( A \) we know that the energy was \( E_n \) (this is achieved by calibrating the apparatus choosing appropriately both \( \lambda \) and the interaction time). To measure work we use \( A \) twice and let the system evolve in between with the operator \( U_E = U_{f,i} = U(t_f, t_i) \). Thus, applying the MIEMI sequence of operations, defined as the composition \( U_{MI}^1 \rightarrow U_E \rightarrow U_{MI} \) (with \( U_{MI} = \exp(-i \lambda \hat{H} \otimes \hat{p} t) \)) an initial state \( |\Psi(t_0)\rangle = |\phi_0\rangle \otimes |w = 0\rangle \) evolves into the final state

\[
|\Psi_f\rangle = \sum_{n,m} \langle \hat{\phi}_m | U_E | \phi_n \rangle \langle \phi_n | \phi_0 \rangle |\hat{\phi}_m\rangle \otimes |w = E_{m,n}\rangle. \tag{4}
\]

Then, the probability to find \( A \) in the state \( |w\rangle \), which is computed as \( P(w) = \langle \Psi_f | (I \otimes |w\rangle \langle w|) |\Psi_f\rangle \), is precisely given by eq. (1). The state of the system after detecting work \( w \) is \( \rho_w = A_w \rho A_w^\dagger / P(w) \), where \( A_w \) is such that \( W(w) = A_w^\dagger A_w \), and

\[
A_w = \sum_{n,m} \delta(w - E_{m,n}) \langle \hat{\phi}_m | U_E | \phi_n \rangle |\hat{\phi}_m\rangle \langle \phi_n |. \tag{5}
\]

In summary, we have described a simple method to measure work, which is such that the outcome \( w \) is generated with probability \( P(w) \), giving rise to the set of non-orthogonal states \( \rho_w \) for the system. In fact, although work is not a Hermitian operator, it can be measured with an ordinary POVM.

**Measuring work in a Stern Gerlach type experiment.** The work measurement we proposed can be realized in a conceptually simple experiment. In fact, the required interaction between \( S \) and \( A \) is precisely the one that takes place in a Stern Gerlach experiment where a particle with spin interacts with an inhomogeneous magnetic field. The momentum of the particle is shifted in two different directions depending on the projection of the spin along the magnetic field. The magnitude of the shift depends on the magnetic field gradient and on the duration of the interaction. The use of Stern Gerlach type interactions to create coherent superpositions of wave packets with different momenta of an atomic beam was recently demonstrated using an atom chip \([15]\). In this remarkable experiment, a gradient pulse, generated by coils in the atom chip, is applied on a \(^{87}\text{Rb}\) cloud obtained from a BEC. The gradient acts as a beam splitter as the atoms follow different trajectories depending on its internal (Zeeman splitting) state. Thus, the atomic cloud split in two coherent pieces. The internal state of the atoms can be manipulated using radio frequency pulses. This setup enables measuring work as follows: a) apply a gradient pulse during a time \( T_1 \), after releasing the atomic cloud from the trap; b) enforce an arbitrary evolution of the internal state with r.f. fields; c) apply a second gradient pulse during a time \( T_2 \). The final state, as shown in \([15]\), has four components whose intensities are determined by the work distribution. The angle of splitting determines the different values of work. Therefore the error in the determination of the work probability is equivalent to the error in determine the mean value of energy. Thus, the experiment in \([15]\) can be interpreted as a simple work measurement (in this case the initial and final hamiltonians are proportional to \( T_1 \) and \( T_2 \)).

**Work estimation through phase estimation.** The above method naturally translates into a quantum algorithm to efficiently sample \( P(w) \) up to fixed precision. The algorithm would run on a quantum computer which could be used to efficiently estimate moments of the work distribution. The method is a variant of the phase estimation algorithm \([16]\), that plays a central role in many quantum algorithms. We consider an \( N \)-qubit system \( S \) (\( D_S = 2^N \)) and an \( M \)-qubit ancilla \( A \) (\( D = 2^M \)) determines the precision of the sampling, as described below. We assume for simplicity that the Hamiltonians \( H \) and \( \tilde{H} \) have bounded spectra that take values between \( \pm E_{M/2} \) (this condition can be relaxed).

The algorithm we describe gives an output \( x \) (which is an \( m \)-bit string) with a probability \( P_D(x) \), which is a coarse-grained version of the work distribution given in \([1]\). Each value of \( x \) identifies a certain amount of work through the identity \( w = 4E_M x/D \). Positive (negative) values of \( w \) correspond to \( 0 < x \leq D/4 \) (\( 3D/4 \leq x \leq D - 1 \)). The quantum algorithm for sampling \( P(w) \), shown in Fig. \([1]\) has six elementary steps: (i) prepare the initial state \( |x = 0\rangle \) for \( A \) and \( \rho \) for \( S \); (ii) apply a quantum Fourier transform (QFT) on \( A \) mapping \( |x\rangle \) onto its conjugate state \( |\tilde{x}\rangle = U_{QFT} |x\rangle = \frac{1}{\sqrt{D}} \sum_{l=0}^{D-1} e^{i2\pi lx} |l\rangle \); (iii) apply the controlled operator \( U_{MI} = \sum_{l=0}^{D-1} |l\rangle \langle l| \otimes U^l \), where \( U^l = \exp(-i \pi H t/4E_M) \); (iv) apply the unitary driv-
ing $U_{f,t}$ over $S$; (v) apply another controlled operation $\hat{U}_{MI} = \sum_{t=0}^{D-1} |t\rangle \langle t| \otimes \hat{U}^t$, with $\hat{U} = \exp(-i\tau \hat{H} t/4E_M)$; (vi) apply the inverse QFT in $A$ and measure its state in the $|x\rangle$ basis. This algorithm applies nothing but the MIEMI sequence described above as the phase estimation subroutine is nothing but a standard measurement interaction.

$$\rho_A \xrightarrow{U_{QFT}} U^\dagger \rho U_{QFT} \xrightarrow{U^t}$$

FIG. 1: Quantum circuit for the estimation of work probability distribution. The initial state of the auxiliary $M$-qubit system is a pure state $\rho_A = |\psi_A\rangle \langle \psi_A|$ with $|\psi_A\rangle = |x = 0\rangle$. When the initial state of the system is pure given by $|\phi\rangle = \sum_{n=1}^N c_n |\phi_n\rangle$ then the state of $S$ and $A$ just before the measurement is $|\Psi\rangle_{S,A} = \frac{1}{\sqrt{N}} \sum_{n,m=1}^N c_n |\phi_m\rangle |U_E| |\phi_n\rangle \sum_{t,x} e^{i/4 \Delta t(x, \frac{E_m - E_n}{4E_M})}|w\rangle \otimes |\tilde{\phi}_m\rangle$.

The operators $A_D(x)$ defining the POVM, which are such that $P_D(x) = \text{Tr}(\rho A_D(x))$, are also a convolution between the exact expression given in (5) and a filter function, i.e. $A_D(x) = \int dw' F_D(4E_M \frac{x}{D} - w') A(w')$.

It is simple to show that $P_D(x)$ rapidly approaches $P_{cg}(x)$, defined as the convolution of $P(w)$ with a rectangular function (which is unity for $|w| \leq 2E_M/D$ and zero otherwise). In fact, if $P(w)$ is bounded, then it is straightforward to show that $\|P_{cg} - P_D\|_\infty = O(\|P\|_\infty/D)$. Therefore, the difference between $P_D(x)$ and $P_{cg}(x)$ decreases exponentially with the size of $A$. In Figure 2 we compare $P_{cg}(w)$ and $P_D(w)$ for two random Hamiltonians $H$ and $\hat{H}$. As $N = 10$, the number of different values of $w$ is $2^{20}$. It is clear that even for a small $A$ (with $M = 5$ qubits), the sampling of the coarse-grained work distribution is highly accurate.

Estimating the free energy by sampling over the work distribution. The fluctuation theorems (in the form of Jarzynski or Crooks identities) are used as a tool to evaluate thermodynamic functions in classical systems [3].

To do this, we need to sample the work distribution $P(w)$ and compute the expectation value $\int dw P(w) \exp(-\beta w)$, which is related to the free energy difference through the Jarzynski identity. The above quantum algorithm enables us to follow the same strategy with a quantum system. Thus, by sampling $P_D(x)$ we can efficiently estimate averages such as $\langle w \rangle$, etc. The accuracy of this estimation depends on the number of samplings $K$ as $1/\sqrt{K}$. So, for fixed precision (independent of the dimensionality of the Hilbert space of $S$) this method is efficient. In Figure 2 we show the dependence of the estimated $\Delta F$ with the number of times the distribution $P_D(x)$ is sampled (for two random Hamiltonians of a system of $N = 10$ qubits). However, as it is the case for classical systems, this strategy is not always efficient. In fact, efficiency depends on the properties of $P(w)$, because negative values of work, for which $\exp(-\beta w)$ is large are typically under represented in the sampling process (a situation that becomes worse when $\beta$ at low temperatures).

An alternative approach is to use Crooks identity. For this, have to sample two distributions, $P_F(w)$ and $P_B(w)$, which correspond to the forward and backwards
process (where the system evolves with $U_f,i$ and $U_i,f$ respectively). In fact, Crooks identity implies that $e^{-\Delta F} \int dw P_D(-w) = \int dw e^{-\beta w} P_F(w)$, where the the interval $I$ is arbitrary. Estimating $\Delta F$ in this way may be a good strategy when Jarzynski identity is inefficient. Also, as $\Delta F$ must be independent of the arbitrary interval $I$, this version of Crook’s identity provides us with a criterion to ensure the convergence of the method.

Summary and comparison with previous work.— We presented a new method to measure work using a standard POVM measurement. The method not only inspires new interpretations of existing experiments [15] but also a new quantum algorithm to efficiently sample over a coarse grained version of the work distribution $P(w)$. The algorithm, that could run in a quantum computer, produces an $M$-bit output $x$ with a probability $P_D(x)$ which is such that $P_D(x) = P(w \in I_x)$ with an accuracy that grows exponentially with $M$. Here, $w \in I_x$ if $|w \mp 4E_M x/D| \leq 2E_M/D$ (where the $\mp$ sign respectively corresponds to the cases $0 \leq x \leq D/4$ and $3D/4 \leq x \leq D - 1$). This efficient sampling can be used to estimate the free energy of quantum states. The method that evaluates the characteristic function $\chi(s)$ \cite{10,11} requires estimating the expectation value of a single qubit operator for each value of $s$. In this way, the work average could be obtained efficiently as it involves estimating the derivative of $\chi(s)$ at the origin. Using this method to sample $P(w)$ requires reconstructing this function from its Fourier transform $\chi(s)$. To achieve the same precision we attain with $M$ qubits in $\mathcal{A}$ this would require evaluating $\chi(s)$ in $2^M$ points. Also, free energy estimation is more involved than in our case since it requires the complete function $\chi(s)$ to analytically continue it to compute $\chi(-i\beta)$. Instead, as we can directly sample coarse-grained versions of $P(w)$ we can efficiently estimate the free energy at least for certain families of Hamiltonians.

Finally, we stress that in order to evaluate free energies our method requires to prepare initially a thermal equilibrium state $\rho = \exp(-\beta H)/Z_\beta$ (the same is required in \cite{10,11}). However, this resource is not necessary if we use the recently proposed quantum Metropolis algorithm that enables the efficient sampling over the Gibbs ensemble [17]. If we only use the algorithm to estimate the forward distribution $P_F(w)$ we need either the thermal state as a resource or to use the quantum Metropolis to sample over an equilibrium state with the initial Hamiltonian. However, if we use Crooks identity, where forward and backward distributions appear, extra resources are needed since we would need to have access to the thermal state of the final Hamiltonian or to use Metropolis to sample over those equilibrium states.

This work was partially supported by grants from Antcyt (PICT 02843) and Ubacyt.

\begin{thebibliography}{10}
\bibitem{1} “Nonequilibrium Equality for Free Energy Differences”, C. Jarzynski, Phys. Rev. Lett. \textbf{78}, 2690 (1997).
\bibitem{2} “Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences”, G.E. Crooks, Phys. Rev. E \textbf{60}, 2721 (1999).
\bibitem{3} “Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies”, D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, C. Bustamante, Nature \textbf{437}, 231 (2005).
\bibitem{4} “Colloquium: Quantum fluctuation relations: Foundations and applications”, M. Campisi, P. Hanggi, and P. Talkner, Rev. Mod. Phys. \textbf{83}, 771 (2011).
\bibitem{5} “A Quantum Fluctuation Theorem”, J. Kurchan, arXiv:cond-mat/0007360.
\bibitem{6} “Jarzynski Relations for Quantum Systems and Some Applications”, H. Tasaki, arXiv:cond-mat/0009244 (2000).
\bibitem{7} “Fluctuation theorems: Work is not an observable”, P. Talkner, E. Lutz, and P. Hanggi, Phys. Rev. E \textbf{75}, 050102(R) (2007).
\bibitem{8} “Employing Trapped Cold Ions to Verify the Quantum Jarzynski Equality”, G. Huber, F. Schmidt-Kaler, S. Deffner, and E. Lutz, Phys. Rev. Lett. \textbf{101}, 070403 (2008).
\bibitem{9} “Crooks Relation in Optical Spectra: Universality in Work Distributions for Weak Local Quenches”, M. Heyl and S. Kehrein, Phys. Rev. Lett. \textbf{108}, 190601 (2012).
\bibitem{10} “Extracting Quantum Work Statistics and Fluctuation Theorems by Single-Qubit Interferometry”, R. Dorner, S. R. Clark, L. Heaney, R. Fazio, J. Goold, V. Vedral, Phys. Rev. Lett. \textbf{110}, 230601 (2013).
\bibitem{11} “Measuring the Characteristic Function of the Work Distribution”, L. Mazzola, G. De Chiara, M. Paternostro, Phys. Rev. Lett. \textbf{110}, 230602 (2013).
\bibitem{12} “Power of One Bit of Quantum Information”, E. Knill, R. Laflamme, Phys. Rev. Lett. \textbf{81}, 5672 (1998).
\bibitem{13} “Experimental reconstruction of work distribution and verification of fluctuation relations at the full quantum level”, T. Batalhão, A. M. Souza, L. Mazzola, R. Aucuaise, I. S. Oliveira, J. Goold, G. De Chiara, M. Paternostro, R. M. Serra, arXiv:1308.3241.
\bibitem{14} Quantum Theory: Concepts and Methods, A. Peres, (Springer, 1993).
\bibitem{15} “Coherent SternGerlach momentum splitting on an atom chip”, S. Machluf, Y. Japha, R. Folman, Nat. Commun. \textbf{4}, 2424 (2013).
\bibitem{16} Quantum Information and Computation, I. Chuang and M. Nielsen, (Cambridge University Press, Cambridge, UK, 2001).
\bibitem{17} “Quantum Metropolis sampling”, K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, F. Verstraete, Nature \textbf{471}, 87 (2011).
\end{thebibliography}