Probabilistic Metrology Attains Macroscopic Cloning of Quantum Clocks

B. Gendra\(^1\), J. Calsamiglia\(^1\), R. Muñoz-Tapia\(^1\), E. Bagan\(^1,2\) and G. Chiribella\(^3\)

\(^1\)Departamento de Física de la Ciencia de la Computación, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain
\(^2\)Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore
\(^3\)Department of Applied Physics, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

It has been recently shown that probabilistic protocols based on postselection boost the performances of state estimation and the replication of quantum clocks. Here we demonstrate that the improvements in these two tasks have to match exactly in the macroscopic limit where the number of clones grows to infinity, preserving the equivalence between asymptotic cloning and estimation for arbitrary values of the success probability. Remarkably, the cloning fidelity depends critically on the number of rationally independent eigenvalues of the clock Hamiltonian. We also prove that probabilistic metrology can simulate cloning in the macroscopic limit for arbitrary sets of states, provided that the performance of the simulation is measured by testing small groups of clones.

High-resolution measurements and new sensors powered by quantum effects are among the most appealing gadgets promised by the field of quantum technologies \(^1\) \(^2\). Consequently, intense effort is being devoted to the design of prototype setups that achieve quantum-enhanced precision and sensitivity \(^3\) \(^4\). In recent years, probabilistic setups based on postselection have attracted a great deal of attention, coming in different variants such as weak value amplification \(^5\) \(^12\) and Probabilistic Metrology (PM) \(^13\) \(^20\). These schemes are based on filters that herald the occurrence of a favourable event, conditional to which the precision is enhanced far beyond the usual limits — e. g. with a scaling upgraded from the Standard Quantum Limit (SQL) to the Heisenberg Limit (HL). Typically, the more dramatic is the improvement, the smaller is the probability of the favourable event, with a trade-off curve between precision and probability that can be quantified explicitly in several interesting cases \(^15\) \(^17\). Note that PM maintains its appeal even in the regime where the probability of favourable events is small: Indeed, in this regime one can design the filter so that, when the unfavourable event occurs, the state of the system is approximately unchanged, thanks to the so-called Gentle Measurement Lemma \(^21\). As a result, PM offers the experimenter the bonus of hitting the HL from time to time — and knowing when this favourable event happens — without compromising optimality of the average scaling.

The advantages of probabilistic filters are not limited to metrology. Instead, they affect a variety of tasks, including cloning \(^18\) \(^22\) \(^23\) and amplification \(^24\) \(^30\). Very recently, it has been shown that for quantum clocks the use of a filter can lead to the phenomenon of super-replication \(^18\), allowing to convert \(n \gg 1\) synchronized clocks into \(m \ll n^2\) replicas, whose joint state appears to be exponentially close to the ideal target of \(m\) perfect copies. Achieving such a replication rate with high fidelity is impossible without filtering, because a deterministic machine that produces more than \(O(n)\) nearly perfect replicas would lead straight into a violation of the SQL.

Considering the striking difference in performance, it is natural to ask whether deterministic and probabilistic cloning machines differ in other, more fundamental features. The most fundamental feature of all is arguably the asymptotic equivalence with state estimation \(^31\) \(^34\), i. e. the fact that, in the macroscopic limit \(m \to \infty\), the optimal performance of quantum cloning can be achieved by measuring the input copies and preparing the clones in a state that depends on the measurement outcome. The no-cloning theorem itself \(^37\) can be considered as a particular instance of this equivalence: two states that can be cloned perfectly by a deterministic machine can also be cloned perfectly in the macroscopic limit and therefore they can be distinguished perfectly by a deterministic estimation strategy, which means that they must be orthogonal to one another. For deterministic machines, the cloning-estimation equivalence has been proved in full generality when the performance of cloning is assessed on small groups of \(k \ll m\) clones \(^33\) \(^34\) and has been recently conjectured to hold even when the \(m\) clones are examined collectively \(^35\) \(^36\). However, nothing is known in the probabilistic case, where the tradeoff between performance and probability of success adds a new twist to the problem. Here, proving the equivalence requires showing that for every cloning machine there is a protocol based on state estimation that, in the macroscopic limit, achieves the same fidelity with the same probability. But is the enhanced precision of PM sufficient to keep up with the highly-increased performance of probabilistic cloning machines?

In this Letter we answer the question in the affirmative, showing that postselection does not challenge the fundamental equivalence between cloning and estimation. We first work out explicitly the example of quantum clocks, where the performance enhancements are the most prominent. We consider clock states \(\hat{\psi}_t = e^{-it\hat{H}}\hat{\psi}_0\) generated from an arbitrary input state \(\hat{\psi}_0\)
by time evolution with arbitrary Hamiltonian acting on a \(d\)-dimensional Hilbert space \(H\) and we exhibit PM protocols that achieve the performances of the optimal cloning machine for every desired value of the success probability. In this comparison, we use the most restrictive criterion, namely the global fidelity between the clones and \(m\) perfectly synchronized replicas of the original clock. We evaluate the fidelity explicitly and discover that its value depends critically on the number of rationally independent eigenvalues of the Hamiltonian. The result is derived using new techniques, based on the Smith normal form [2], which we expect to be useful for other problems in quantum metrology and optimal quantum information processing. Furthermore, we analyze the scenario where the performances of cloning are judged from groups of \(k \ll m\) clones, establishing the equivalence between probabilistic cloning and estimation for arbitrary sets of input states and for arbitrary values of the success probability. This result extends the validity of equivalence to all points of the optimal performance-probability tradeoff.

Let us start from the concrete example of quantum clocks. With a suitable choice of basis, the state of a clock at time \(t = 0\) can be written as \(|\psi_n\rangle = \sum_{j=0}^{d-1} \sqrt{p_j}|j\rangle\), where \(H|j\rangle = e_j|j\rangle\) and \(p_j\) is the probability that a measurement of energy gives outcome \(e_j\). Without loss of generality, we assume that all probabilities \(\{p_j\}\) are non-zero, that the eigenvalues \(\{e_j\}\) are distinct, and that \(c_0 = 0\). In the case of \(n\) identical synchronized clocks, we denote the state at time \(t\) as \(|\Psi_n^t\rangle := |\psi_t\rangle \otimes_n |\psi_t\rangle\). The values of the total energy can be labeled by the partitions of \(n\) into \(d\)-non-negative integers \((n_0, n_1, \ldots, n_{d-1})\). Denoting by \(\text{n} \in \mathbb{Z}_{(d-1)}\times \{e_1, \ldots, e_{d-1}\}\) the 1 column integer \(1\) row real matrix \(\text{n} = (n_0, n_1, \ldots, n_{d-1})\) we express the corresponding energy as \(E_n := \sum_{j=0}^{d-1} e_j n_j = e_n\). The spectrum of the total Hamiltonian will be denoted by \(\mathcal{S}_p = \{E_n\}_{n \in \mathcal{P}_n}\), where \(\mathcal{P}_n\) is the lattice of vectors \(n\) satisfying \(n_j \geq 0\) for every \(j\) and \(\sum_{j=0}^{d-1} n_j \leq n\).

By collecting the vectors that lead to the same energy \(E\), we define the set \(\mathcal{P}_n^E := \{n \in \mathcal{P}_n : E_n = E\}\), so that \(|\mathcal{P}_n^E|\) is the degeneracy of \(E\). Then, the state of the \(n\) clocks can be written as

\[
|\Psi_n^t\rangle = \sum_{E \in \mathcal{S}_p} e^{-iEt} \sqrt{\rho_{E,n}} |E,n\rangle, \quad p_{E,n} := \sum_{n \in \mathcal{P}_n^E} p_{n,n},
\]

where \(p_{n,n}\) is the multinomial distribution \(p_{n,n} := n! \prod_{j=0}^{d-1} p_j^{n_j}/n!\) and

\[
|E,n\rangle := \frac{1}{\sqrt{\rho_{E,n}}} \sum_{n \in \mathcal{P}_n^E} \sqrt{\rho_{n,n}} |n,n\rangle.
\]

In the metrology scenario, the goal is to estimate the time \(t\) as accurately as possible. The estimate, denoted by \(\hat{t}\), can be used to produce \(m\) approximate clones, by preparing a state \(\hat{\Psi}_m^t = \sum_{E \in \mathcal{S}_p} e^{-iEt} \sqrt{\rho_{E,m}} |E,m\rangle\), so that, averaging over all possible values of \(\hat{t}\), the output state resembles \(|\Psi_m^t\rangle\). As a performance measure, we adopt the worst-case fidelity

\[
F = \inf_{t \in \mathbb{R}} \int dt p(t) \langle \hat{\Psi}_m^t | \hat{\Psi}_m^t \rangle^2,
\]

where \(p(t)\) is the probability of estimating \(\hat{t}\) when the true value is \(t\). For covariant families of states, such as the quantum clocks under consideration, the worst-case fidelity is equal to the average of the fidelity with respect to the uniform prior. In the case of PM the estimation strategy does not provide an estimate all the times, but sometimes declares “failure”, in which case one abstains from producing copies. Hence, \(p(t)\) has to be understood as \(p(t,t,\text{succ})\), the probability of estimating \(\hat{t}\) conditional to the fact that the strategy succeeded in producing an estimate. To take this into account, one can equivalently replace the input state \(|\Psi_1^t\rangle\) by a state the form \(\Pi^{1/2}|\Psi_1^t\rangle/||\Pi^{1/2}|\Psi_1^t\rangle||\), where \(\Pi\) is suitable operator satisfying \(0 \leq \Pi \leq 1\). By the symmetry of the figure of merit in Eq. (3), it is easy to see that \(\Pi\) can be chosen to be diagonal in the energy basis, namely \(\Pi = \sum_{E \in \mathcal{S}_p} \pi_E |E,n\rangle \langle E,n|\). With this choice the probability of success is independent of \(t\) and is given by \(P_{\text{succ}} := \langle \Psi_1^t | \Pi |\Psi_0^t\rangle\). Fixing the coefficients of the guess state and of the filter, we show that the supremum of the fidelity over all quantum measurements is [39]

\[
F = \sum_{\mathcal{E}} \left( \frac{\sum_{E \in \mathcal{S}_p} \sqrt{\rho_{E,pE+\epsilon,m} \pi_E / P_{\text{succ}}}}{\sum_{E \in \mathcal{S}_p} \sqrt{\rho_{E,pE+\epsilon,m} \pi_E / P_{\text{succ}}}} \right)^2,
\]

where the outer sum runs over the set \(\{\mathcal{E} = E_m - E_n | E_m \in \mathcal{S}_p, E_n \in \mathcal{S}_p\}\) and the prime (‘) means that the sum is restricted to those \(E \in \mathcal{S}_p\) such that \(E + \epsilon \in \mathcal{S}_p\). For a fixed filter \(\Pi\), we denote by \(F^{\Pi}_{\text{PM}}\) the supremum of the fidelity in Eq. (4) over all possible guess states. The question now is whether, suitably choosing the filter \(\Pi\), the value of \(F^{\Pi}_{\text{PM}}\) can reach the fidelity of the optimal quantum copy machine in the macroscopic limit \(m \to \infty\). Let us examine now the cloning scenario. Here, one can produce copies coherently using a general quantum operation (completely positive trace-non-increasing map) \(\mathcal{C}_{n,m}\) that maps states on \(\mathcal{H}^{\otimes n}\) to states on \(\mathcal{H}^{\otimes m}\). The quantum operation \(\mathcal{C}_{n,m}\) can always be written as a composition of a probabilistic filter \(\Pi\) followed by a trace preserving map \(\mathcal{D}_{n,m}\) [18]. Again, the symmetry of the figure of merit allows one to choose without loss of generality a filter of the form \(\Pi = \sum_{E \in \mathcal{S}_p} \pi_E |E,n\rangle \langle E,n|\). Our first step is to upper bound the maximum fidelity achievable for a fixed filter, denoted by \(F_{\text{CL}}\), as [39]

\[
F_{\text{CL}} \leq \max_{E \in \mathcal{S}_p} p_{E,m} \left( \sum_{E \in \mathcal{S}_p} \sqrt{\rho_{E,pE} \pi_E / P_{\text{succ}}} \right)^2.
\]
The next step is to show that the fidelity $F_{PM}^{\Pi}$ achieves the upper bound in the limit of large $m$. For the sake of illustration, we start from the simple case where all the energies $\{e_j\}$ are commensurable, i.e., $e_j = k_j \varepsilon$ where $k_j$ is an integer and $\varepsilon$ is a fixed unit of energy. Then, the energy eigenvalues for $m$ clocks can be written as $E_m = \varepsilon k m$. For sufficiently large $m$, it is easy to see that the minimum spacing between two consecutive energies is given by $\Delta E^* = \varepsilon \gcd\{k_m\}$, with $\gcd$ denoting the greatest common divisor. This fact is an immediate consequence of Bezout’s identity in number theory [40]. Now, suppose that $m$ is asymptotically large. In this limit, the multinomial distribution $p_{m,m}$ approaches the multivariate normal distribution $\mathcal{N}(\mu p, \Sigma)$, where $p := (p_1, \ldots, p_{d-1})$ and the covariance matrix $\Sigma$ has entries $\Sigma_{ij} = mp_j(1-p_j)$, $\Sigma_{il} = -mp_j p_l$, $j \neq l$. Clearly, this implies that the probability distribution $p_{E,m}$ in Eq. (1) is concentrated in a window of size $O(\sqrt{m})$ centred around the mean value $\langle H \rangle$. Within this window, every two consecutive energies differ by the minimum amount $\Delta E^*$ [41]. Finally, note that by dimensional arguments the degeneracy of a typical energy grows as $|\beta E| \sim m^{d-2}$. Thus, the sum over $m \in \mathcal{P}_m$ that defines $p_{E,m}$ in Eq. (1) can be approximated by the integral of $\mathcal{N}(mp, \Sigma)$ over a $(d-2)$-dimensional domain. As a result, $p_{E,m}$ is approximated by the discrete Gaussian distribution

$$
p_{E,m} \approx \Delta E^* \frac{e^{-\frac{(E - \mu p)^2}{2 \Var(H)}}}{\sqrt{2\pi m \Var(H)}}, \quad \Delta E^* = \varepsilon \gcd\{k_m\}, \quad (6)
$$

where $\Var(H) = \langle H^2 \rangle - \langle H \rangle^2$ is the variance of $H$.

Thanks to Eq. (6) we are now in position to show that $F_{PM}^{\Pi}$ approaches $F_{CL}^{\Pi}$ in the macroscopic limit. Indeed, since for $m \gg n$ the probability $p_{E,m}$ is almost constant over every interval of size $O(n)$, we can pull out a factor $\max_{E \in \mathcal{S}_m} \langle p_{E,m} \rangle$ from both sums in Eq. (4) introducing an error that vanishes in the asymptotic limit. Moreover, we can choose a guess state with probabilities $p_{E,m}$ given by a discrete Gaussian with width growing as $\sqrt{m^{1-\eta}}$, $0 < \eta < 1$. Since the width is much larger than $O(n)$, we have $p_{E+\varepsilon,m} \approx p_{E,m}$ for every $E \in \mathcal{S}_n$, and we can pull out the term $p_{E,m}$ from the sum over $\mathcal{S}_n$. Since the width is much smaller than $\sqrt{m}$, we have $\sum_{E} p_{E,m} \approx 1$. Hence, our choice of guess state attains the upper bound in Eq. (5) and we have [39]

$$
F_{PM}^{\Pi} \approx F_{CL}^{\Pi} \approx \frac{\Delta E^*}{\sqrt{2\pi m \Var(H)}} \left( \sum_{E \in \mathcal{S}_n} \sqrt{p_{E,n}} \frac{\pi E}{P_{\text{suc}}} \right)^2. \quad (7)
$$

The reader should not be misled by the simplicity of Eq. (6), which superficially may seem an application of the Central Limit Theorem (CLT). The CLT gives an approximation of the cumulative distribution of $p_{E,m}$, not of the probability mass itself. In fact, those who believe that $p_{E,m}$ should converge to a Gaussian are in for a surprise in the case where the eigenvalues of $H$ are not commensurable. In this case, the eigenvalues $\{e_j\}$ can be expressed as integer linear combinations of a minimal number $r$ of rationally independent units of energy $\{\epsilon_i\}_{i=1}^r$. Rational independence means that, for every set of integer coefficients $\{c_i\}$, the relation $\sum c_i \epsilon_i = 0$ implies $c_i = 0$ for every $l$. In terms of the units $\{\epsilon_i\}$, we expand each eigenvalue as $e_j = \sum_i \epsilon_i k_{ij}$ where $\{k_{ij}\}$ are integer coefficients, uniquely defined thanks to the rational independence of the units. Using the decomposition, we express the energy of $m$ clocks as $E_m = \varepsilon \tilde{m}$ where $\varepsilon = (\epsilon_1, \ldots, \epsilon_r)$, and $\tilde{m} = (\tilde{m}_1, \ldots, \tilde{m}_r)$ is the 1 column matrix with components $\tilde{m}_i = \sum_j k_{ij} m_j$, or, more compactly, $\tilde{m} = K \bar{m}$ where $K$ is the $r \times (d-1)$ matrix with entries $\{k_{ij}\}$. The matrix $K$ maps the lattice $\mathcal{P}_m \subset \mathbb{Z}^{d-1}$ into a new lattice $\mathcal{P}_m \subset \mathbb{Z}^r$ with the special feature that the points in $\mathcal{P}_m$ are into one-to-one correspondence with the energies in $\mathcal{S}_m$ (again, due to the rational independence of the units). Hence, instead of the probability distribution $p_{E,m}$ we can consider the probability distribution $p_{\bar{m},m} := \sum_{m} K_{m} \bar{m} \times p_{\bar{m}}$. Now, for large $m$ this probability distribution is concentrated in a volume of size $O(m^{r/2})$ centred around the mean $mK p$. The typical vectors $m \in \mathcal{P}_m$ associated to points in this volume form a regular Bravais lattice and the number of vectors associated to a point $\tilde{m}$ grows as $m^{d-r-1}$ [39]. By the same arguments as in the paragraph after Eq. (5), we can approximate the sum in the expression of $p_{\bar{m},m}$ with an integral, thus obtaining [39]

$$
p_{\bar{m},m} \approx \Delta V^* \exp \left\{ -\frac{\varepsilon^2}{2 \det \Sigma} \right\}, \quad (8)
$$

where $\bar{m}_0 = mK p$, $\Sigma = K E K^t$ and $\Delta V^*$ is the volume of a minimal cell of the lattice $\mathcal{P}_m$. Using the Smith normal form of $K$ one can show that $\Delta V^* = \gcd\{|K_1|\}$, where $\{K_1\}$ is the set of all minors of $K$ of order $r$ [39]. Eq. (8) has two major consequences: First, the probability mass $p_{E,m}$ does not converge to a Gaussian, as one would naively expect from a misapplication of the CLT. Instead, it converges to the non-continuous function $p_{\tilde{m}(E),m}$, where $\tilde{m}(E)$ is the value of $\tilde{m}$ such that $E = E_m$. Second, all the steps of the proof for commensurable energies can now be reproduced by replacing the energy $E \in \mathcal{S}_p$ with the corresponding vector $\tilde{m}(E)$. In this way we obtain $F_{PM}^{\Pi} \approx F_{CL}^{\Pi} \approx F_{\Pi}$, with

$$
F_{\Pi} := \frac{\Delta V^*}{\sqrt{(2\pi)^r \det \Sigma}} \left( \sum_{\bar{m} \in \mathcal{P}_m} \sqrt{p_{\bar{m},n}} \frac{\pi \bar{m}}{P_{\text{suc}}} \right)^2. \quad (9)
$$

Note that, since $\det \Sigma$ scales as $m$, the fidelity scales as $m^{-r/2}$. Hence, the asymptotic behaviour depends dramatically on the number of rationally independent units.
Quite remarkably, this means that the fidelity is not continuous in the Hamiltonian: Although one can approximate arbitrarily well the Hamiltonian $H$ with another Hamiltonian $H'$ that has commensurable energies, the corresponding fidelities are not going to be close. The origin of the discontinuity is that the value of the fidelity depends on the closure of the set of clock states, due to the infimum in Eq. (3). When the energy eigenvalues are commensurable, the time evolution is periodic and the orbit $\{\Psi^n_t, t \in \mathbb{R}\}$ is a close one-dimensional curve. But when the energies are combinations of $r$ rationally independent units, the orbit is dense in an $r$-dimensional submanifold of the manifold of pure states. This phenomenon is the quantum analog of a classic feature of integrable Hamiltonian systems \[12\], where rationally independent frequencies lead to ergodic time evolutions in phase space. Here the fidelity is discontinuous because it depends on the long-time behaviour of the time evolution, during which the quantum clock can probe a higher dimensional manifold. Note that for finite times the differences between both families of clock-states generated by $H$ and $H'$ can be arbitrary small and that continuity is retrieved if we restrict the infimum in Eq. (3) to a fixed time interval $[T_1, T_2]$.

Having proven the asymptotic equivalence between probabilistic metrology and cloning of quantum clocks, we now give closed expressions for optimal fidelity, maximized over all possible filters. We focus on the large $n$ limit under the condition $n \ll \sqrt{m}$ and consider two relevant regimes: First, we allow arbitrarily low probability of success, showing that the ultimate fidelity is \[39\] \[43\]

$$F = \left[(2\pi)^r \det \Sigma\right]^{-1/2} |\tilde{T}_n| \Delta V^*,$$

(10)

where $|\tilde{T}_n|$ is the number of sites in the lattice $T_n$. Since $|\tilde{T}_n|$ scales as $n^r$, the fidelity scales as $F \sim (n/\sqrt{m})^r$.

Second, we consider the case where the probability of success is high, i.e. $P_{\text{succ}} = 1 - \eta$ for some small $\eta$. Here the optimal fidelity acquires the particularly simple form

$$F = \left(4n/m\right)^{1/2} \left[1 + \eta(1 - 2^{-r/2})\right] + O(\eta^2).$$

(11)

Quite surprisingly, $F$ does not depend on the coefficients $\{p_j\}$ of the input state, but only on the number of rationally independent units $r$.

We conclude by discussing the equivalence between probabilistic metrology and cloning for arbitrary sets of states. Here we assess the performance of cloning by looking at a random subset of $k \ll m$ clones, evaluating the global fidelity between the state of the $k$ clones and the state of $k$ ideal copies. Clearly, since the $k$ clones are picked at random, one can assume that the optimal cloner is invariant under permutation of the $m$ output systems. Technically, this means that the cloner is described by a quantum operation $\mathcal{E}_{n,m}$ such that, for every permutation $\pi$, one has $\mathcal{Z}_\pi \mathcal{E}_{n,m} = \mathcal{E}_{n,m}$, where $\mathcal{Z}_\pi$ is the permutation map defined by $\mathcal{Z}_\pi(\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_m) = \rho_{\pi(1)} \otimes \rho_{\pi(2)} \otimes \cdots \otimes \rho_{\pi(m)}$. Using a de Finetti-type argument, we prove the following result \[39\].

**Theorem 1.** For every quantum operation $\mathcal{E}_m$ with input in $\mathcal{H}_m$ and permutationally invariant output in $\mathcal{H}_m$ there exists a PM protocol, described by a quantum operation $\mathcal{E}_m$, such that i) $\mathcal{E}_m$ and $\mathcal{E}_m$ have the same success probability and ii) the error probability in distinguishing between $\mathcal{E}_m$ and $\mathcal{E}_m$ by inputting a state $\rho$ and measuring $k$ output systems is lower bounded by $\rho_{\text{err}} \leq \frac{1}{2} + (kd^2)/(2mP_{\text{succ}}(\rho))$, where $P_{\text{succ}}(\rho) := \text{tr}[\mathcal{E}_m(\rho)]$.

In the case of cloning, the result implies that the $k$-copy fidelity of an arbitrary $n$-to-$m$ cloner on a generic input state $|\psi_x\rangle^{\otimes n}$ can be achieved by PM, up to an error of size $k/[mP_{\text{succ}}(|\psi_x\rangle\langle\psi_x|^{\otimes n})]$ \[39\]. Hence, the error vanishes as $k/m$ for every process with success probability larger than a given finite value for every possible input. This result extends the equivalence between cloning and estimation to every point of the tradeoff curve between fidelity and success probability. The result holds also in the Bayesian scenario where the input state $|\psi_x\rangle^{\otimes n}$ is given with prior probability $p_x$ and one considers average fidelities and average success probabilities. Furthermore, we show that the average success probability of the optimal $n$-to-$m$ cloner is lower bounded by a finite value independent of $m$ and therefore the best asymptotic cloner can be simulated by PM \[39\].

In conclusion, we proved that probabilistic protocols empowered by postselection do not challenge the fundamental equivalence between cloning and estimation. We worked out explicitly the case of quantum clocks, where the performance enhancements for both tasks are most dramatic, and developed a technique to evaluate the optimal asymptotic fidelity. We found out that the asymptotic fidelity depends critically on the number of rationally independent units generating the spectrum of the Hamiltonian, due to an effect that is analog to ergodicity of classical dynamical systems. Finally, we discussed the case of arbitrary families of states, establishing an equivalence between probabilistic metrology and cloning when the performance is quantified by the fidelity of $k \ll m$ randomly chosen clones.

We thank Elio Ronco-Bonvelli for his collaboration at early stages of this work. We acknowledge support by the European Regional Development Fund (ERDF), the National Basic Research Program of China (973) 2011CB000300 (2011CB000301), the Spanish MICINN (project FIS2008-01236) with FEDER funds, the Generalitat de Catalunya CIRIT (project 2009SGR-0985), the National Natural Science Foundation of China through Grants 11350110207, 61033001, and 61061130540, and by the Foundational Questions Institute through the large grant “The fundamental principles of information dynamics”. GC is supported by the 1000 Youth Fellowship Program of China.
REFERENCES

[1] J.P. Dowling and G.J. Milburn, Phil. Trans. R. Soc. Lond. A 361, 1655-1674 (2003).
[2] J.L. O’Brien, A. Furusawa, and J. Vuckovic, Nature Photon. 3, 687-695 (2009).
[3] V. Giovannetti, S. Lloyd, and L. Maccone, Science 306, 1330-1336 (2004).
[4] V. Giovannetti, S. Lloyd, and L. Maccone, Nature Photon. 5, 222-229 (2011).
[5] Y. Aharonov and L. Vaidman, Phys. Rev. A 41, 11 (1990).
[6] O. Hosten and P. Kwiat, Science 319, 787 (2008).
[7] P. B. Dixon, D. J. Starling, A. N. Jordan, and J. C. Howell, Phys. Rev. Lett. 102, 173601 (2009).
[8] D.J. Starling, P. B. Dixon, A. N. Jordan, and J. C. Howell, Phys. Rev. A 80, 041803 (2009).
[9] S. Wu and Y. Li, Phys. Rev. A 83, 052106 (2011).
[10] S. Massar and S. Popescu, Phys. Rev. A 84, 052106 (2011).
[11] J. Fiurášek, New. J. Phys. 8, 192 (2006).
[12] S. Massar and S. Popescu, Phys. Rev. A 84, 052106 (2011).
[13] B. Gendra, E. Ronco-Bonvehi, J. Calsamiglia, R. Munoz-Tapia, and E. Bagan, New. J. Phys. 14, 105015 (2012)
[14] B. Gendra, E. Ronco-Bonvehi, J. Calsamiglia, R. Munoz-Tapia, and E. Bagan, Phys. Rev. Lett. 110, 100501 (2013)
[15] G. Chiribella, Y. Yang, and A.C.-C. Yao, Nature Commun. 4, 2915 (2013).
[16] M. Marcus and H. Minc, A survey of matrix theory and matrix inequalities (Dover, New York, 1964).
[17] A. Katok and B. Hasselblatt, Introduction to the modern theory of dynamical systems, Cambridge (1996).
[18] G. Chiribella, G. M. D’Ariano, Phys. Rev. Lett. 97, 250503 (2006)
[19] G. Chiribella, Lecture Notes in Computer Science, 6519, 9 (2011).
[20] J. Bae and A. Acín, Phys. Rev. Lett. 97, 030402 (2006)
[21] G. Chiribella, G. M. D’Ariano, Phys. Rev. Lett. 97, 250503 (2006)
[22] G. Chiribella, Lecture Notes in Computer Science, 6519, 9 (2011).
[23] Y. Yang and G. Chiribella, LIPICS 22, 220 (2013).
[24] G. Chiribella and Y. Yang, arXiv:1404.0990
[25] W. K. Wootters and W. H. Zurek, Nature 299, 802 (1982).
[26] M. Marcus and H. Minc, A survey of matrix theory and matrix inequalities (Dover, New York, 1964).
[27] Supplemental Material.
[28] G. E. Andrews, Number theory, Courier Dover Publications (1994).
[29] The difference may be larger at the ends of the spectrum, where the density of points is generally more rarefied.
[30] A. Katok and B. Hasselblatt, Introduction to the modern theory of dynamical systems, Cambridge (1996).
[31] This expression is also correct for small $n$. For large $n$, we can replace $|\tilde{P}_n|\Delta V^*$ by $V_n$, the volume of the minimal polytope that contains the lattice $\tilde{P}_n$.
[32] A.S. Holevo, Probabilistic And Statistical Aspects Of Quantum Theory, North-Holland Series In Statistics And Probability (North-Holland, Amsterdam 1982).
SUPPLEMENTAL MATERIAL

Fidelity of probabilistic metrology

In this section, we derive Eq. (4) for the PM fidelity. Since later we will show that asymptotically the r.h.s. of Eq. (4) achieves the optimal cloning fidelity, here it is easy to show that the r.h.s. of Eq. (4) can be achieved by some suitable measurement.

For every $\sigma > 0$, consider the operators $\Phi^n_{t,\sigma} = |\Phi^n_{t,\sigma}\rangle\langle \Phi^n_{t,\sigma}|$, where

$$|\Phi^n_{t,\sigma}\rangle := \sqrt{p_\sigma(\hat{t})} \sum_{E \in \mathcal{S}_n} e^{-iE\hat{t}}|E,n\rangle,$$  \hspace{1cm} (12)

and $p_\sigma(t)$ is a suitable probability distribution. The latter is chosen as follows: For commensurable energies, when the evolution is periodic, $p_\sigma(t)$ is the uniform distribution $p_\sigma(t) = 1/T$ over the period $T$ [$p_\sigma(t) = 0$ for $t < 0$ and $T < t$]. For incommensurable energies, $p_\sigma(t)$ is the Gaussian distribution $p_\sigma(t) = (2\pi\sigma^2)^{-1/2} e^{-t^2/(2\sigma^2)}$. Now, for commensurable energies, the operators $\{\Phi^n_{t,\sigma}\}_{t \in R}$ define a quantum measurement: indeed, it is immediate to check the normalization condition $\int dt \Phi^n_{t,\sigma} = 1$, where $1_n$ denotes the identity on the subspace containing the state of the $n$ input copies. For incommensurable energies, the operators $\{\Phi^n_{t,\sigma}\}_{t \in R}$ form an “approximate measurement”, satisfying the approximate normalization condition $\int dt \Phi^n_{t,\sigma} = 1 + O(e^{-\sigma^2})$. Note that in the limit $\sigma \to \infty$ the approximate measurement becomes arbitrarily close to a legitimate measurement, as the normalization defect disappears in such limit.

Let us denote by $F_\sigma$ the value obtained by inserting the approximate measurement $\{\Phi^n_{t,\sigma}\}_{t \in R}$ in Eq. (3). Since our set of approximate measurements becomes closer and closer to the set of allowed measurements as $\sigma \to \infty$, the limit value $F_* := \lim_{\sigma \to \infty} F_\sigma$ is an achievable value of the fidelity. We now show that $F_*$ is equal to the r.h.s. of Eq. (4): recalling the expansion of $|\Psi^n_{t}\rangle$ in Eq. (1) and the definition of $|\Psi^n_{t}\rangle$ after Eq. (2), we obtain

$$F_* = \inf_{t} \int dt p_\sigma(\hat{t}) \left| \sum_{E \in \mathcal{S}_n} e^{-iE(\hat{t}-t)} f_E \right|^2,$$  \hspace{1cm} (13)

with

$$f_E := \sum_{E \in \mathcal{S}_n} \sqrt{\frac{\pi E P_{E,n} P_{E+E,m} P_{E+E,m}}{P_{\text{succ}}}}.$$

Integrating over $\hat{t}$ we then obtain

$$F_* = \lim_{\sigma \to \infty} \inf_{t} \sum_{E,E'} e^{-\frac{\pi E P_{E,n} P_{E+E,m} P_{E+E,m}}{P_{\text{succ}}}} f_E^2,$$

which coincides with the r.h.s. of Eq. (4).

Cloning fidelity and its upper bound

In this section we derive the upper bound (5) to the probabilistic cloning fidelity [1]. For the sake of self-completeness, we also derive the fidelity itself. We start from the definition of the worst case cloning fidelity:

$$F = \inf_{t} \frac{\langle \Psi^n_{t} | \mathcal{C}_{n,m}(\Psi^n_{t}) | \Psi^n_{t} \rangle}{P_{\text{succ}}}$$  \hspace{1cm} (15)

where in full generality the quantum operation $\mathcal{C}_{n,m}$ [see paragraph before Eq. (5)] has been decomposed as a probabilistic filter followed by a deterministic (trace preserving) map, i.e., as $\mathcal{C}_{n,m} = \mathcal{D}_{n,m} \circ \Pi$, and the probability of success is $P_{\text{succ}} = \langle \Psi^n_{t} | \Pi | \Psi^n_{t} \rangle$. The covariance of quantum clocks enables us to drop the infimum in the last equation and consider without loss of generality only invariant filters of the form $\Pi = \sum_{E \in \mathcal{S}_n} \pi_E |E,n\rangle\langle E,n|$, as well as covariant maps, which satisfy

$$\mathcal{D}_{n,m}(U^n_t \cdot U^n_{t'}) = U^n_{t'} \mathcal{D}_{n,m}(\cdot) U^n_{t},$$

where $U^n_t = \sum_{E \in \mathcal{S}_n} e^{-iE(f_t |E,n\rangle\langle E,n|}$ is the time evolution operator. Using the Choi-Jamiolkowski isomorphism, Eq. (16) is equivalent to

$$\mathcal{D} = U^n_t \otimes U^n_{t'}^* \mathcal{D} (U^n_t \otimes U^n_{t'}^*),$$

which in turn implies the direct sum decomposition $\mathcal{D} = \sum_{\epsilon} \mathcal{D}_\epsilon$, where

$$\mathcal{D}_\epsilon := \sum_{E,E' \in \mathcal{S}_n} d_{E,E'}^\epsilon |E+\epsilon,m\rangle\langle E'+\epsilon,m| \otimes |E,n\rangle\langle E',n|,$$

and the ‘primed’ sum include only those terms for which $E+\epsilon, E'+\epsilon \in \mathcal{S}_m$. Taking the above into account, the probabilistic cloning fidelity can be cast as

$$F = \frac{1}{P_{\text{succ}}} \sum_{E,E'} \sum_{\epsilon} \sqrt{\pi E P_{E,n} P_{E+E,m} P_{E+E,m}} d_{E,E'}^\epsilon,$$

where $d_{E,E'}^\epsilon$ is the probability of success of the probabilistic filter followed by a deterministic (trace preserving) map.

An upper bound to the fidelity [19] can be obtained by pulling the maximum value of $d_{E,E'}^\epsilon$ out of the sums. Recalling the positivity of the Choi-Jamiolkowski operator, $\mathcal{D} \geq 0$, one has $\mathcal{D}_\epsilon \geq 0$ and $|d_{E,E'}^\epsilon|^2 \leq d_{E,\epsilon,E'}^\epsilon d_{E,\epsilon,E'}^\epsilon$. Thus,

$$F \leq \max_{E \in \mathcal{S}_n} \left\{ \sum_{E,E' \in \mathcal{S}_n} \sum_{\epsilon} \sqrt{\pi E P_{E,n} P_{E+E,m} P_{E+E,m}} \frac{\pi E P_{E,n} P_{E',m} P_{E',m}}{P_{\text{succ}}} \right\},$$  \hspace{1cm} (20)
where we have interchanged the order of summation. We can now use the Schwarz inequality to write

$$\sum_{\xi} \sqrt{d_{E,E}^\xi} \geq \left( \sum_{\xi} d_{E,E}^\xi \right)^{1/2} \left( \sum_{\xi} d_{E,E}^{\xi'} \right)^{1/2}.$$

(21)

The first (second) sum on the right hand side of the last equation can be extended to values of $\xi$ for which $E + \xi E \in SP_n$ ($E + \xi E \in SP_n$), as $d_{E,E}^{\xi} \geq 0$ ($d_{E,E}^{\xi} \geq 0$). Then, each of these sums is unity since $\mathcal{P}_{n,m}$ is trace preserving. Substituting in (20), we obtain the bound in Eq. (5).

**Geometry of the problem and Smith variables**

Rationally independent energy units and lattices

We recall that as a set of points the partitions $\{n_n, n\}$, where $n \in \mathcal{P}_n$, is a regular lattice on the simplex $\Delta_n^{d-1} = \{x_0, \ldots, x_{d-1}\} \in \mathbb{R}^d : x_j \geq 0, j = 0, \ldots, d - 1$ and $\sum_{j=0}^{d-1} x_j = n$, of edge length $n$. Each site of this lattice is of the form $(n - \sum_{j=1}^{d-1} n_j, n)$. The vectors (1 column integer matrices) $n \in \mathcal{P}_n$ form themselves also a regular lattice, defined by the inequalities $0 \leq n_j \leq n - \sum_{j=1}^{d-1} n_j, j = 1, \ldots, d - 1$, inside the corner of a $(d - 1)$-dimensional cube of side length $n$: $\Delta_n^{d-1} = \{(x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1} : 0 \leq x_j \leq n - \sum_{j=1}^{d-1} x_j, j = 1, \ldots, d - 1\}$. Note that if $n \leq m$, one has the inclusion $\mathcal{P}_n \subset \mathcal{P}_m$.

Recall also that the spectrum of $H^\mathcal{P}$, is given by $\text{Sp}_n = \{E_n = \text{e} \text{n} : n \in \mathcal{P}_n\}$. All vectors $n$ that give rise to a particular value $E$ of the energy, i.e., those in the set $\mathcal{P}_n^E = \{n \in \mathcal{P}_n : E_n = E\}$, necessary lie in the affine hyperplane obtained by translating the hyperplane orthogonal to $\mathcal{e}$. Since $\mathcal{P}_n$ is a regular lattice, it is not clear a priori how many of its sites fall on the hyperplane defined by $\mathcal{e}$; the actual degeneracy of $E$ strongly depends on the commensurability of the energies $\{e_j\}$ in the spectrum of $H$. To identify all distinct values of $E_n \in \text{Sp}_n$ we use the fact that the energies of the Hamiltonian $H$ can always be written as a linear combination with integer coefficients of a minimal set of rationally independent 'energy units' $\{e_i\}_{i=1}^r$, $r \leq d - 1$, namely

$$e_j = \sum_{i} k_{ij} e_i \quad k_{ij} \in \mathbb{Z}.$$  

(22)

By minimal we mean that no subset of $\{e_i\}_{i=1}^r$ is sufficient to write every energy in the spectrum of $H$ as in Eq. (22). Note that the rational independence of $e_i$ implies that $k_{ij}$ is fixed once the choice of energy units is made. With this, the energies of $\text{Sp}_n$ can be written as $E_n = \sum_{i=1}^r e_i \tilde{n}_i$, where $\tilde{n}_i = \sum_{j=1}^{d-1} k_{ij} n_j$. It is then useful to introduce the vector notation $\tilde{n} = \text{K} n$, where $\text{K}$ is the $r \times (d - 1)$ matrix whose integer entries are $k_{ij}$, and $E_n = \text{e} \tilde{n}$, where $\text{e} = (\varepsilon_1, \ldots, \varepsilon_r)$. Because of the rational independency of the energy units, we conclude that there is a bijection between the distinct energies in $\text{Sp}_n$ and the points in the set $\mathcal{P}_n = \{\tilde{n} = \text{K} n : n \in \mathcal{P}_n\}$.

We view each column $\text{K}_l = (k_{1l}, \ldots, k_{rl})^t \in \mathbb{Z}^r$ of $\text{K}$ as a set of vectors that span the (infinite) Bravais lattice

$$\mathcal{P}_\infty = \big\{\tilde{n} = \sum_{j=1}^d n_j \text{K}_l, n_j \in \mathbb{Z}\big\}.$$  

(23)

We have the obvious inclusion $\mathcal{P}_n \subset \mathcal{P}_\infty$. For finite $n$, $\mathcal{P}_n$ departs from the Bravais lattice $\mathcal{P}_\infty$ in two ways: i) the lattice $\mathcal{P}_n$, defined by the linear transformation $\text{K}$ acting on $\mathcal{P}_n$, inherits its boundaries and hence lies inside the convex $r$-dimensional polytope $\Delta_n^r := \text{K} \Delta_n^{d-1}$; ii) near the boundaries of $\Delta_n^r$, some points of $\mathcal{P}_\infty$ are missing in $\mathcal{P}_n$ (since none of the corresponding inverse images satisfy the constrains that define $\mathcal{P}_n$, given in the first paragraph of this section), so we typically have $\mathcal{P}_n \subseteq \mathcal{P}_\infty \cap \Delta_n^r$. These boundary related issues have a minor effect in our analysis for asymptotically large $n$, as we argue below.

Smith vectors/variables. Volume of primitive cell

For $r < d - 1$, the vectors $\text{K}_l$ are not linearly independent and, therefore, they cannot be a minimal set of primitive vectors of the lattice $\mathcal{P}_\infty$. On the other hand, having a minimal set of primitive vectors is necessary in order to compute the volume $\Delta V^* \text{ of the unit cell [see, e.g., Eqs. (8) and (9)]}$. To this purpose, we use the Smith normal form of $\text{K}$ [2]:

$$\text{K} = \text{TAPS},$$  

(24)

where $\text{T} \in \mathbb{Z}^{r \times r}$ and $\text{S} \in \mathbb{Z}^{(d-1) \times (d-1)}$ are unimodular matrices, i.e. invertible matrices over the integers with $\det \text{T} = \det \text{S} = \pm 1$; $\text{A} \in \mathbb{Z}^{r \times r}$ is a diagonal matrix with entries

$$\text{(A)}_{ll} = \frac{\gcd\{\text{K}_{ll}\}}{\gcd\{\text{K}_{ll-1}\}}.$$  

(25)

where $\gcd$ stands for greatest common divisor and $\{\text{K}_{ll}\}$ is the set of all minors of $\text{K}$ of order $l$; $\text{P}$ is the $r \times (d - 1)$ matrix with entries

$$\text{(P)}_{ll} = 1, \quad 1 \leq l \leq r,$$  

(26)

and zero otherwise. Using (24) we can define the lattice $\mathcal{P}_n$ in terms of the new vectors/variables

$$s = (s_1, \ldots, s_r)^t := \text{PSn},$$  

(27)

which we coin Smith vectors/variables. Given a matrix $\text{K}$ the Smith form guarantees that the choice of Smiths variables is unique, up to linear combinations within the degenerate subspaces of $\text{A}$ (if any). Note that since the
matrix $S$ (and analogously $T$) is unimodular, the gcd of each of its rows and columns is unity. Indeed, for the first column (similarly for the other columns/rows) one has $S_1 = \gcd\{(S)_{j,1}\} a_1$, for some $a_1 \in \mathbb{Z}^{d-1}$. Then, $1 = \det S = \gcd\{(S)_{j,1}\} \det S$, where $S$ is obtained by substituting $a_1$ for the first column of $S$. Since $\det S \in \mathbb{Z}$, necessarily $\det S = \gcd\{(S)_{j,1}\} = 1$. Being the case that the gcd of each row of $S$ is unity, Bezout lemma ensures that each component of $s$, $s_i$, will take all integer values.

That is, in stark contrast to the variables $\tilde{n}_i$, the Smith variables $n_i$ take values independently of each other, with unit spacing between consecutive values. This means that the Bravais lattice $\tilde{\mathcal{P}}_\infty$ is defined in terms of the Smith variables as

$$\tilde{\mathcal{P}}_\infty = \{ \tilde{n} = \sum_{l=1}^r s_l v_l, s_l \in \mathbb{Z} \}, \quad (28)$$

where $\{v_l\}_{l=1}^r$ are a set of (linear independent) primitive vectors of the lattice defined by each of the $r$ columns of $TA$. It follows that the volume of the unit cell can be computed as

$$\Delta V^* = |\det(v_1, v_2, \ldots, v_r)| = \det A = \gcd(\{ |K|_r \}), \quad (29)$$

where we have used Eq. (25).

**Parametrizing the energy in terms of the Smith variables**

Both the vectors $\tilde{n}$ and the Smith vectors $s$ are in one-to-one correspondence with the distinct energies in $Sp_n$, and thus they are interchangeable in all our arguments. As a matter of fact, they would coincide had we chosen the energy units as

$$\varepsilon^S := \varepsilon T A \quad (30)$$

(the script $S$ stands for Smith), so that the total energy becomes $E_n = \varepsilon^S s$. Note that, if $\varepsilon$ is minimal, then also $\varepsilon^S$ must be minimal, since the two vectors are related by an invertible matrix with integer entries.

Despite the one-to-one correspondence, the Smith variables $s$ are more convenient than the variables $\tilde{n}$. Indeed, they define a cubic lattice with unit spacing, which facilitates, e.g., taking the continuum limit. Note that one has

$$\det \tilde{\Sigma} = \det(\varepsilon^S K S^T K') = \det \left[ (TA^T P S) \Sigma (TA^T P S)^t \right] = \det^2 A \det \Sigma^S \quad \Sigma^S := PS \Sigma (PS)^t$$

which implies the relation

$$\Delta V^* \left( \det \tilde{\Sigma} \right)^{-\frac{1}{2}} = \left( \det \Sigma^S \right)^{-\frac{1}{2}}. \quad (31)$$

Using this fact, the volume of the unit cell in Eqs. (9) and (10) can be absorbed into the matrix $\Sigma^S$, which is the covariance matrix of the probability distribution of Smith variables $p_{s,n}$.

**Equivalence of different choices of energy units**

It has already been mentioned that the choice of rationally independent energy units that define $K$ is not unique. Here we show that such ambiguity has no physical implications. Our strategy is to show that different choices of energy units lead to Smith variables that are related by unimodular matrices.

Let $\varepsilon$ and $\varepsilon'$ be two minimal sets of rationally independent energy units spanning the spectrum of $H$. In the following we will use the same notation of the previous section, attaching primes to all the matrices and quantities defined in terms of $\varepsilon'$. Note that, by minimality, there must be an invertible transformation $R$ such that $\varepsilon = \varepsilon' R$. Comparing the two relations $E_n = \varepsilon K' n$ and $E_n = \varepsilon K n$ and using the rational independence of the units $\varepsilon'$ we obtain the relation

$$K'n = RKn$$

$$= RTA^T P S n$$

$$= RTA s$$

$$= Ms, \quad M := RTA, \quad (32)$$

the second and third equalities coming from the Smith form in Eq. (24) and the definition of the Smith vectors in Eq. (27), respectively. Now, recall that $K'$ is a matrix of integers, and, therefore $K'n$ is a vector of integers for every $n \in \mathbb{Z}^{d-1}$. Since the Smith variables $s_i$ are independent and take all possible integer values, by choosing $s = (1, 0, \ldots, 0)^t$, the relation $K'n = Ms$ implies that the first column of $M$ must have integer entries. By the same argument, all columns of $M$ must have integer entries, i.e. $M$ is an integer matrix. Note that $M$ is invertible (since it is defined as the product of three invertible matrices), although its inverse needs not be a matrix with integer entries.

Let us write $M$ in the Smith form $M = UBV$, where $U$ and $V$ are unimodular and $B$ is an invertible diagonal matrix. Inserting this expression in Eq. (32), we obtain

$$K'n = UBV s$$

$$= UBVPSn,$$

having used Eq. (27). Since the relation holds for arbitrary integer vectors, we obtained $K' = UBVPS$. Note that, by definition of the matrix $P$ in Eq. (20), we have

$$VP = PW, \quad (33)$$
where $W$ is the $(d - 1) \times (d - 1)$ matrix defined as, e.g., $W = V \oplus 1_{d-1-r}$. Hence, we have $K' = UBWPS$. Now, comparing this equation with the Smith form $K' = T'\Lambda'P'S'$ we obtain $T' = U$, $\Lambda' = B$, $P' = P$ and $S' = W S$. In conclusion, the Smith variables defined by $K'$ are related to the Smith variables defined by $K$ through the relation

$$s' = P' S' n
= P \Lambda W S n
= V P S n
= V s,$$

having used Eq. (33) in the third equality. Clearly, the covariance matrix of $s'$ is given by $\Sigma'^S = V \Sigma S V'$, where $\Sigma S$ is the covariance matrix of $s$, and, thanks to the unimodularity of $V$, we have $\det \Sigma'^S = \det \Sigma S$. Using Eq. (31) we conclude that

$$\Delta V^*(\det \Sigma)^{-1/2} = (\det \Sigma S)^{-1/2} = (\det \Sigma'^S)^{-1/2} = \Delta V^*(\det \Sigma')^{-1/2}.$$

This proves that physically relevant quantities, such as the most probable energy in Eq. (8) or the fidelity (9), do not depend on the choice of energy units used to represent the spectrum of $H$.

**Boundary defects**

Ideally, one would like to have $\mathcal{T}_n = \mathcal{T}_\infty \cap \Delta^n$, however, near the boundary of the polytope $\Delta^n$, some sides are missing, giving rise to loss of regularity, as already mentioned. We argue here that the thickness of the defective region does not scale with $n$. This ensures that in the asymptotic limit of large $n$, e.g., the sums over $\mathcal{T}_\infty$ can always be safely approximated by integrals over $\Delta^n$, even when the probability distribution $p_n$ is flat, as required to obtain Eq. (10). The thickness of the defective region depends solely on the intrinsic properties of the Hamiltonian $H$, including the number of rationally independent energy units, through the matrix $K$.

To simplify the argument let us assume that $\mathcal{T}_\infty$ is a cubic lattice with unit spacing. It has been shown above that this assumption does not entail any loss of generality, as it just requires choosing energy units as in (30). Let $K' = (k_{1,1}, \ldots, k_{d-1,1})$ stand for the row $l$ of the matrix $K$. Then, for each $l = 1, \ldots, r$

$$\tilde{n}_l = k'_l \cdot x, \quad x \in \mathbb{R}^{d-1},$$

where $\tilde{n}_l$ is the $l$-th component of $\tilde{n} \in \mathcal{T}_\infty \cap \Delta^n$, defines a discrete family of hyperplanes that are orthogonal to $K'$. The kernel of $K \in \mathbb{Z}^{r \times (d-1)}$ can always be spanned by a set $\{u_k\}_{k=1}^{d-r-1}$ of independent vectors of $\mathbb{Z}^{d-1}$ that are orthogonal to $\{K'_l\}_{l=1}^{r}$. Upper bounds, $b$, to the minimum length of these (integer) vectors (high-dimensional extensions of the so-called Siegel’s bound) can be found in [34] and depends entirely on the matrix $K$ ($b$ is thus independent of $n$). We note that $\{u_k\}_{k=1}^{d-r-1}$ define a lattice within $\ker K$ and any $(d - r - 1)$-dimensional ball or radius $b$ (or larger) contains at least one site of it.

A site $\tilde{n} \in \mathcal{T}_\infty \cap \Delta^n$ belongs to $\mathcal{T}_n$ iff there is at least one site $n \in \mathcal{T}_n$ that satisfies (34) for all $l$. The Smith normal form and Bezout lemma (see previous subsection) ensure that there are infinitely many vectors $n$ in $\mathcal{T}_\infty$ satisfying (34) for a given $\tilde{n}$ (but they are not necessarily in $\mathcal{T}_n$). The difference between any two such vectors, $n - n'$ belongs to $\ker K$, i.e., satisfies $n - n' = \sum_{k=1}^{d-r-1} \eta_k u_k$, $\eta_k \in \mathbb{Z}$. Therefore, if a given $\tilde{n}$ is such that the intersection of the polytope $\Delta^n$ with the hyperplanes defined by (34) contains a ball of radius $b$, it is ensured that $\tilde{n} \in \mathcal{T}_n$. This is so because at least one site in $n_0 + \ker K$, where $n_0 \in \mathcal{T}_\infty$ is a solution (any of the infinitely many solutions) of (34), will be contained in this ball, as argued above. Since $\Delta^n$ is convex, by increasing $n$ (the length of its edge) the volume of its intersection with the hyperplanes (34), which is also a convex polytope, grows as $n^{d-r-1}$ and it will eventually contain a ball of radius $b$. Only sides whose distance to the boundary is kept fixed and is small enough can escape from this fate and may thus not be in $\mathcal{T}_n$. This concludes our proof.

By the same argument, since the volume of the intersection of $\Delta^n$ with the hyperplanes (34) grows with $n$ (but at a fixed distance from the boundary), the number of balls or radius $b$ it will eventually contain grows as $n^{d-r-1}$ and so does the number of point in $\ker K$, i.e., the numbers of points in $\mathcal{P}_n$ associated to a given $\tilde{m}$, as required to obtain, e.g., Eq. (8).

As a final remark, we note also that the pattern of defects near the boundary of $\Delta^n$ is exactly the same for every $n$, provided that $n$ is sufficiently large. The reason is that all the lattices $\mathcal{P}_n$ are similar (i.e., have the same shape but different size). Hence, the regions of them that are at a fixed distance from the boundary are identical and, as argued above, these regions determine the pattern of defects of $\mathcal{T}_n$.

**Equivalence between PM and macroscopic cloning of quantum clocks**

In this section we show that the PM fidelity attains the upper bound [5] to the cloning fidelity, and thus prove the equivalence between PM and macroscopic cloning of quantum clocks. The proof uses the one-to-one correspondence between $Sp_m$ and $\mathcal{T}_m$, as well as the fact that the probability distribution $p_m$ approaches a mul-
tivariate Gaussian distribution as $m$ goes to infinity. The former, enables us to write Eq. (4) as

$$F_{PM}^\Pi = \sum_{\mu} \left( \sum_{n} \chi_{\nu} \sqrt{P_{n+m, \mu} P_{n+m, \mu}} \right)^2,$$

where $\chi_{\nu} := \sqrt{P_{n+m, \nu} / P_{\text{suc}}}$ and the outer sum runs over $\mathcal{P}_{n+m} := \{ \mu = \bar{n} - \bar{n} | \bar{n} \in \mathcal{P}_n, \bar{n} \in \mathcal{P}_m \}$. To keep our notation as uncluttered as possible we suppress the tildes throughout the rest of this section. We further assume that the multivariate normal distribution $p_{m,m}$ peaks at $m = 0$, and so does $p_{n,m}$, but we make no other assumption on the form of $p_{n,m}$ (it could, e.g., be flat, in which case any point in $\mathcal{P}_n$ could be chosen to be 0). This may require shifting the vectors in $\mathcal{P}_m$ by a fixed $m_0 \in \mathcal{P}_m$ (similarly, by a fixed $n_0 \in \mathcal{P}_n$ for those in $\mathcal{P}_n$). The primed summation is restricted to vectors $\mu$ such that $\mu + n \in \mathcal{P}_m$.

For any $\rho > \nu > 0$, we define the set $\mathcal{R}_\rho = \{ \mu : \forall n \in \mathcal{P}_m, \mu + n \in \mathcal{P}_m \}$ and $|\mu + n| \leq \rho \}$. Then

$$F_{PM}^\Pi > \sum_{\mu \in \mathcal{R}_\rho} \left( \sum_{n} \chi_{\nu} \sqrt{P_{n+m, \mu} P_{n+m, \mu}} \right)^2.$$

Note that we can drop the prime in the last sum over $n$. We have

$$F_{PM}^\Pi > p_{0,m} e^{-\rho^2 / 2\sigma^2} \sum_{\mu \in \mathcal{R}_\rho} \left( \sum_{n} \chi_{\nu} \sqrt{P_{n+m, \mu}} \right)^2,$$

where $m \sigma^2$ is the smallest eigenvalue of the covariance matrix of $p_{m,m}$. Here, we explicitly display the $m$ dependence of the covariance matrix eigenvalues; thus $\sigma^2$ does not scale with $m$. Let us choose the ‘guessed’ distribution as

$$\hat{p}_{m,m} = \hat{p}_{0,m} e^{-\rho \left( \frac{|m|^2}{2m} \right)}, \sum_{m \in \mathcal{P}_m} \hat{p}_{m,m} = 1.$$  

Then,

$$\hat{p}_{n+m, \mu} = \hat{p}_{0,m} e^{-\rho \left( \frac{|n+m|^2}{2m} \right)} \geq \hat{p}_{0,m} e^{-\rho \left( \frac{|n|^2}{2m} \right)} \geq \hat{p}_{m,m},$$

where we have used that $|\mu| \leq |n| + \rho$ if $\mu \in \mathcal{R}_\rho$. Thus, the following bound holds

$$F_{PM}^\Pi > p_{0,m} e^{-\rho^2 / 2\sigma^2} \left( \sum_{n} \chi_{\nu} \right)^2 \sum_{\mu \in \mathcal{R}_\rho} \hat{p}_{m,m}.$$

We now need to lower bound the last sum. For this, we write

$$\sum_{\mu \in \mathcal{R}_\rho} \hat{p}_{m,m} = 1 - \sum_{\mu \in \mathcal{R}_\rho \cap \mathcal{P}_m} \hat{p}_{m,m}.$$

For $\mu \in \mathcal{R}_\rho \cap \mathcal{P}_m$ one has $\rho < |n + \mu| \leq \nu + |\mu|$, then, recalling that $\rho > \nu$,

$$\sum_{\mu \in \mathcal{R}_\rho \cap \mathcal{P}_m} \hat{p}_{m,m} < 1 - \left( \sum_{\mu \in \mathcal{R}_\rho \cap \mathcal{P}_m} \hat{p}_{m,m} \right).$$

Note that $|\mathcal{P}_m| \leq (m + d - 1)! / \left| m(d-1) ! \right| \sim m^{-d-1}$, for large $m$. With all the above,

$$F_{PM}^\Pi > \left( \max_{m} p_{m,m} \right) e^{-\rho^2 / 2\sigma^2} \left( \sum_{n} \chi_{\nu} \right)^2,$$

for some positive constant $C$. Therefore, if $\rho = m^{1/2}$, and $\zeta = m^d$, with $(1 + \epsilon)/2 < \delta < \epsilon > 0$, then $\rho^2 / m = m^{-\delta}, \zeta / m = m^{-1+\delta}, \zeta \rho / m = m^{1-2\epsilon} + \delta$ and $\zeta \rho / m = m^{1+\epsilon}$. Thus, for large $m$ we have

$$F_{PM}^\Pi > \left( \max_{m} p_{m,m} \right) \left( \sum_{n} \chi_{\nu} \right)^2.$$

This result holds provided $p_{m,m}$ is a multivariate normal distribution picked at some $m_0 \in \mathcal{P}_m$. The actual distribution is multinomial on the points of $\mathcal{P}_m$. However, as $m$ becomes asymptotically large, the induced distribution $p_{m,m}$ becomes arbitrarily close to the multivariate normal assumed in the proof above. Recalling that the right hand side of (44) is also an upper bound to $F_{CL}^\Pi$ and, thus to $F_{PM}^\Pi$, we finally conclude that (we restore the suppressed tildes)

$$F_{PM}^\Pi = F_{CL}^\Pi = \left( \max_{m} p_{m,m} \right) \left( \sum_{n} \chi_{\nu} \right)^2$$

for asymptotically large $m$ and fixed $n$. This leads to Eq. (9), of which Eq. (7) is a particular case for $r = 1$.

Explicit calculations

In this section we give some details of the calculation leading to Eqs. (10) and (11). As already mentioned, Smith vectors $s \in Z^*$ [recall Eq. (27)] are most suited to this purpose because they form a cubic lattice of unit step size, i.e., their minimal cell has volume $\Delta V^* = 1$. In this sense, they are just a particular instance of vectors $\mathbf{m} \in \mathcal{P}_m$. Hence, to avoid further proliferation of notation, we will use here the generic symbols $\Sigma$ and $\mathcal{P}_m$ to refer to the covariance matrix of the multivariate Gaussian distribution $p_{s,m}$ and the lattice of the Smith vectors $s$ respectively. Since $\Sigma$ scales with $m$, we write $\Sigma = m \Sigma_1.$
where \( \Sigma_1 \) is independent of \( m \). Then, the expression for the asymptotic fidelity in (9) becomes

\[
F^\Pi = \frac{1}{\sqrt{(2\pi m)^r \det \Sigma_1}} \left( \sum_{s \in \mathcal{P}_n} \sqrt{p_{s,n}} \frac{\pi_s}{P_{\text{succ}}} \right)^2.
\]

(46)

The last sum can be evaluated in the asymptotic limit of large \( n \) (recall however that we assume \( n \ll m \), as we show below. In this case, also \( p_{s,n} \) approaches a multivariate normal distribution, as that in Eq. (8), and the sum over \( \mathcal{P}_n \) can be approximated by an integral over the polytope \( \Delta_n \).

As a warmup act, we first compute the fidelity \( F^\Pi \) for the deterministic protocol, i.e., when \( P_{\text{succ}} = 1 \). Since no filtering is applied, we have \( \pi_s = 1 \) for all \( s \in \mathcal{P}_n \). The sum in (46) simplifies to

\[
\int_{\Delta_n} ds \sqrt{p_{s,n}} \simeq \int_{\mathbb{R}^r} ds \sqrt{p_{s,n}} = 2^{r/2} \left( \frac{(2\pi)^r}{\det \Sigma} \right)^{1/4} = 2^{r/2} (2\pi)^{-r/4} (\det \Sigma)^{-1/4} \times \sqrt{p_{s,n}}.
\]

(47)

as \( 2^{-r/2} (2\pi)^{-r/4} (\det \Sigma)^{-1/4} \times \sqrt{p_{s,n}} \) is also a properly normalized multivariate normal distribution with covariance matrix \( 2\Sigma \). Substituting in Eq. (46) we obtain

\[
F^\Sigma = \left( \frac{4n}{m} \right)^{r/2}.
\]

(48)

This expression agrees with Eq. (11) in the deterministic limit, when \( n \to 0 \).

In the probabilistic case, \( P_{\text{succ}} < 1 \), we need to optimize the filter parameters \( \{\pi_s\} \). To simplify the notation, let us use the definition of the normalized state \( |\xi\rangle \), with components \( \xi_s := \sqrt{p_{s,n}} \pi_s / P_{\text{succ}} \). Then, the maximum fidelity, \( F = \max_{\pi} F^\Pi \), is obtained when the sum in (46) takes its maximum value:

\[
\max \sum_s \xi_s,
\]

subject to \( \sum_s \xi_s^2 = 1 \)

and \( \xi_s \leq \sqrt{p_{s,n} / P_{\text{succ}}} \), \( s \in \mathcal{P}_n \),

(49)

(50)

(51)

where \( \pi \) is the normalization constraint, and \( \xi \) comes from the positivity and trace preserving requirements on the stochastic filter. To solve (49), (50), and (51), we use Lagrange multipliers and the Karush-Kuhn-Tucker conditions. The problem reduces to solving the stationary conditions

\[
\frac{\partial}{\partial \xi_s} \left( \sum_{s'} \xi_{s'} \right) = \frac{\partial}{\partial \xi_s} \left[ \frac{1}{2} \left( \sum_{s'} \xi_{s'}^2 - 1 \right) \right] + \sum_{s'} \sigma_{s'} \left( \xi_{s'} - \sqrt{p_{s',n} / P_{\text{succ}}} \right),
\]

(52)

where the sums extend to \( s \in \mathcal{P}_n \), and \( 1/(2\zeta) \) and \( \{\sigma_s\}_{s \in \mathcal{P}_n} \) are multipliers. Conditions (50) and (51) are called primal feasibility conditions. In addition, one has to impose that \( \sigma_s \geq 0 \), known as dual feasibility condition, and

\[
\sigma_s \left( \xi_s - \sqrt{p_{s,n} / P_{\text{succ}}} \right) = 0,
\]

(53)

known as complementary slackness condition, both for all \( s \in \mathcal{P}_n \). The latter, implies that at any site of \( \mathcal{P}_n \), either \( \sigma_s = 0 \) or \( \xi_s = \sqrt{p_{s,n} / P_{\text{succ}}} \), in which case we say that \( s \) belongs to the coincidence set \( \mathcal{C} \), i.e., \( \mathcal{C} := \{ s \in \mathcal{P}_n : \xi_s = \sqrt{p_{s,n} / P_{\text{succ}}} \} \). If \( s \notin \mathcal{C} \), Eq. (52) readily gives the constant solution \( \xi_s = \zeta \).

If \( P_{\text{succ}} < \min_{s \in \mathcal{P}_n} p_{s,n} \), then \( p_{s,n} / P_{\text{succ}} > 1 \geq \xi_s^2 \), thus \( \mathcal{C} = \emptyset \). In this case, normalization implies \( \zeta = |\mathcal{P}_n|^{-1/2} \), where \( |\mathcal{P}_n| \) is the number of sites in \( \mathcal{P}_n \). Substituting in (46), we have

\[
F = \frac{|\mathcal{P}_n|}{\sqrt{(2\pi m)^r \det \Sigma_1}}.
\]

(54)

Recalling Eq. (31) and \( m \Sigma = \bar{\Sigma} \), this equation becomes Eq. (10), which holds for any choice of vectors \( \mathbf{n} \). Notice that both, Eqs. (54) and (10), also hold for small \( n \). For large \( n \), \( |\mathcal{P}_n| \Delta V^* \) (recall \( \Delta V^* = 1 \) for Smith vectors) approaches \( V_n \), the volume of the polytope \( \Delta_n \), as the irregularities or defects of the lattice \( \mathcal{P}_n \) can only arise within a finite distance from its boundary (see previous sections). Closed formulas for \( |\mathcal{P}_n| \) or \( V_n \) depend on the Hamiltonian \( H \) and do not seem to generalize easily. In the main text, only the obvious scalings \( |\mathcal{P}_n| \sim V_n \sim n^r \) are used to show that \( F \sim (n/\sqrt{m})^r \).

FIG. 1. Plot of \( \xi_{s,n} \) (z axis) for large \( n \) (the truncated bell-shaped surface). The figure also shows the complement of the coincidence set and \( \sqrt{p_{s,n} / P_{\text{succ}}} \) for \( P_{\text{succ}} > \min_{s \in \mathcal{P}_n} p_{s,n} \).

For \( P_{\text{succ}} > \min_{s \in \mathcal{P}_n} p_{s,n} \), the problem becomes more involved, as the constraint (51) is now non-trivial and \( \mathcal{C} \neq \emptyset \). Since \( p_{s,n} / P_{\text{succ}} \) is bell-shaped, the complement of the coincidence set is the ‘ellipsoid’ \( \mathcal{P}_\alpha := \{ s \in \mathcal{P}_n : (s - s_0)^t \Sigma^{-1} (s - s_0) \leq \alpha^2 \} \), for some \( \alpha \) (Fig. 1).
shows a two-dimensional version of it), and we have the solution
\[ \xi_s = \begin{cases} \sqrt{p_{s,n}/p_{\text{suc}}} & \text{if } s \in \mathcal{C}_\alpha \\ \sqrt{p_{s,n}/p_{\text{suc}}} & \text{if } s \notin \mathcal{C}_\alpha \end{cases}, \tag{55} \]
where we have defined
\[ p_{s,n} := \frac{e^{-\alpha^2/2}}{(2\pi)^r \det \Sigma}, \tag{56} \]
and we note that the parameter \( \alpha \) that gives the size of the ‘ellipsoid’ \( \mathcal{C}_\alpha \) is determined by normalization:
\[ 1 = \sum_s \xi_s^2 = \int d^s \xi_s^2. \]
Note also that the solution \( \xi_s \) is a ‘truncated’ multivariate normal distribution, as shown in Fig. 1. The above integral thus splits into two straightforward ways of the two regions in \( \mathbb{R}^2 \), and we obtain the relation:
\[ P_{\text{suc}} = \frac{1}{\Gamma(\frac{r}{2}+1)} \left( \Gamma\left(\frac{r}{2}, \frac{\alpha^2}{2}\right) + \alpha r e^{-\alpha^2/2} / 2^{r-1} \right), \tag{57} \]
where \( \Gamma(a, x) \) is the upper incomplete Gamma function. Proceeding along the same line, one can compute \( \sum_s \xi_s \) to obtain
\[ F = \frac{1}{\Gamma(\frac{r}{2}+1)} \left( \frac{n}{2m} \right)^{r/2} \left[ 2^{r-1} \Gamma\left(\frac{r}{2}, \frac{\alpha^2}{2}\right) + \alpha e^{-\alpha^2/2} / 2^{r-1} \right]. \tag{58} \]
Eqs. \( 57 \) and \( 58 \) give the solution to our optimization problem in parametric form, in terms of \( \alpha \). Finding the fidelity \( F \) as a explicit function of \( P_{\text{suc}} \) would require inverting the relation \( 57 \), which cannot be done analytically for an arbitrary success probability. We can however obtain an analytic expression for \( F \) for a success probability close to one, i.e., for \( P_{\text{suc}} = 1 - \eta \), \( \eta \ll 1 \), by simply expanding the right hand side of Eq. \( 57 \) to leading order in \( \alpha \). Such expansion reads:
\[ \eta = 1 - P_{\text{suc}} = \frac{\alpha^2}{2^{1+r/2} \Gamma(2+r/2)} + \mathcal{O}(\alpha^{r+3}). \tag{59} \]
Solving for \( \alpha \) and substituting in the expansion (at leading order in \( \alpha \)) of the right hand side of Eq. \( 58 \) one obtains
\[ F = \left( \frac{4n}{m} \right)^{r/2} \left[ 1 + (1 - 2^{-r/2}) \eta \right] + \mathcal{O}(\eta^2). \tag{60} \]
Eqs. \( 57 \) through \( 60 \) hold provided \( P_{\text{suc}} \) does not exponentially vanish with \( n \). If it does, as in \( [10] \), where we assumed that \( P_{\text{suc}} < \min_{\mathbb{R}_n} p_{s,n} \), we obtain the better scaling \( F \sim (n/\sqrt{m})^r \). The reason for this different scaling is that a multivariate normal distribution only approximates \( p_{s,n} \) accurately around its peak, whereas it falls off exponentially with \( n \) at the tails (where \( \min_{\mathbb{R}_n} p_{s,n} \) lies).

### Proof of theorem 1

**Proof.** The proof follows the same lines of the proofs of theorems 4 and 5 in Ref. [5]. Let \( \mathcal{C}_m \) be a quantum operation transforming states on \( H_{\text{in}} \) into states on \( H \otimes m \). First, we suppose that the output states of \( \mathcal{C}_m \) have support contained in the symmetric subspace of \( H \otimes m \). In this case, it is useful to consider the universal measure-and-prepare channel \( \mathcal{M} \) defined by
\[ \mathcal{M}(\rho) = \int d\psi \; d_m^+ \; \text{tr}[\psi \otimes \rho] \; \psi \otimes \rho \]
where \( d\psi \) is the invariant measure over the pure states, \( d_m^+ = (\frac{m+1}{d-1}) \) is the dimension of the symmetric subspace. Note that the map \( \mathcal{M} \) is trace-preserving for all states with support in the symmetric subspace. Moreover, \( \mathcal{M} \) provides a good approximation of the partial trace \( \mathcal{T} \):
\[ \| \text{tr}_{m-k} \circ \mathcal{M} \|_o \leq \frac{2kd}{m}, \tag{61} \]
where \( \| \mathcal{L} \|_o \) denotes the diamond norm of a linear, Hermitian-preserving map \( \mathcal{L} \), defined as
\[ \| \mathcal{L} \|_o := \sup_{r \in \mathbb{N}} \sup_{|\Psi \rangle \in H_{\text{in}} \otimes \mathcal{C}^r} \text{tr}[\mathcal{L}(\mathcal{I}_r)(|\Psi\rangle\langle\Psi|)], \quad \| |\Psi\rangle\| = 1 \]
\( \mathcal{I}_r \) denoting the identity map for an \( r \)-dimensional quantum system. Using Eq. \( 61 \), it is immediate to construct the desired PM protocol. The protocol consists in performing the quantum operation \( \mathcal{C}_m \) and subsequently applying the measure-and-prepare channel \( \mathcal{M} \). Mathematically, it is described by the quantum operation \( \mathcal{C}_m = \mathcal{M} \circ \mathcal{C}_m \). Since \( \mathcal{M} \) is trace-preserving (in the symmetric subspace), one has
\[ \text{tr}[\mathcal{C}_m(\rho)] = \text{tr}[\mathcal{C}_m(\rho)] \]
for every input state \( \rho \), that is, \( \mathcal{C}_m \) and \( \mathcal{C}_m \) have the same success probability. Moreover, one has the bound
\[ \| \text{tr}_{m-k} \circ \mathcal{C}_m - \text{tr}_{m-k} \circ \mathcal{C}_m \|_o \leq \| \text{tr}_{m-k} \circ \mathcal{M} \circ \mathcal{C}_m \|_o \leq \| \text{tr}_{m-k} \circ \mathcal{M} \|_o \| \mathcal{C}_m \|_o \leq \frac{2kd}{m}, \tag{62} \]
having used the fact that \( \| \mathcal{C}_m \|_o \leq 1 \) by definition. In words, the \( k \)-copy restrictions of \( \mathcal{C}_m \) and \( \mathcal{C}_m \) are close to each other provided that \( k \ll m \).

Now, suppose that the output of \( \mathcal{C}_m \) has support outside the symmetric subspace. In this case, the invariance under permutations implies that \( \mathcal{C}_m \) has a Stinespring dilation of the form
\[ \mathcal{C}_m = \text{tr}_{\mathcal{H}_E} \circ \mathcal{K}_m \quad \mathcal{K}(\rho) := K_m \rho K_m^\dagger, \]
where $\mathcal{H}_E = \mathcal{H}^\otimes m$ and $K_m$ is an operator with range contained in the symmetric subspace of $\mathcal{H}^\otimes m \otimes \mathcal{H}_E \simeq (\mathcal{H} \otimes \mathcal{H})^\otimes m$ (for a proof see e.g. [5]). Hence, we can take the measure-and-prepare quantum operation $\mathcal{E}_m := \mathcal{M}_E \circ \mathcal{H}_m$, where $\mathcal{M}_E$ is the universal measure-and-prepare channel on $\mathcal{H}^\otimes m \otimes \mathcal{H}_E$, and we can define $\tilde{\mathcal{E}}_m := \text{tr}_{\mathcal{H}_E} \circ \mathcal{E}_m$. By definition, the success probability of $\tilde{\mathcal{E}}_m$ is equal to the success probability of $\mathcal{E}_m$. Moreover, one has the relation
\[
\left\| \text{tr}_{m-k} \circ \mathcal{E}_m - \text{tr}_{m-k} \circ \tilde{\mathcal{E}}_m \right\|_\infty \\
= \left\| \text{tr}_{m-k} \circ \text{tr}_{\mathcal{H}_E} \circ \left( \mathcal{H}_m - \mathcal{E}_m \right) \right\|_\infty \\
= \left\| \text{tr}_{\mathcal{H}_E} \circ \left( \text{tr}_{m-k} \otimes \text{tr}_{\mathcal{H}_E}^{-1} \right) \circ \left( \mathcal{H}_m - \mathcal{E}_m \right) \right\|_\infty \\
\leq 2d^2 k ,
\]
where $\text{tr}_{\mathcal{H}_E}^k$ denotes the partial trace over $k$ ancillary Hilbert spaces. Hence, one gets the bound
\[
\left\| \text{tr}_{m-k} \circ \mathcal{E}_m - \text{tr}_{m-k} \circ \tilde{\mathcal{E}}_m \right\|_\infty \\
\leq \left\| \text{tr}_{\mathcal{H}_E}^k \right\|_\infty \left\| \left( \text{tr}_{m-k} \otimes \text{tr}_{\mathcal{H}_E}^{-1} \right) \circ \left( \mathcal{H}_m - \mathcal{E}_m \right) \right\|_\infty \\
\leq 2d^2 k ,
\]
having used Eq. (62) with $\tilde{\mathcal{E}}_m$, $\mathcal{E}_m$, and $d$ replaced by $\mathcal{E}_m$, $\mathcal{H}_m$, and $d^2$, respectively.

Finally, the error probability in distinguishing between $\mathcal{E}_m$ and $\tilde{\mathcal{E}}_m$ by inputting a state $\rho$ and measuring $k$ output system is equal to the error probability in distinguishing between the two states
\[
\tilde{\rho}_{m,k} = \frac{\text{tr}_{m-k}[\tilde{\mathcal{E}}_m(\rho)]}{\text{tr}[\tilde{\mathcal{E}}_m(\rho)]} \quad \text{and} \quad \rho_{m,k} = \frac{\text{tr}_{m-k}[\mathcal{E}_m(\rho)]}{\text{tr}[\mathcal{E}_m(\rho)]} ,
\]
respectively. Assuming equal prior probabilities for the two quantum operations, Helstrom theorem gives the bound $p_{\text{err}} = \frac{1}{2} \left[ 1 + \frac{1}{2} \| \tilde{\rho}_{m,k} - \rho_{m,k} \|_1 \right]$ and therefore we have
\[
p_{\text{err}} \leq \frac{1}{2} \left[ 1 + \frac{\| \text{tr}_{m-k} \circ (\mathcal{E}_m - \tilde{\mathcal{E}}_m) \|_\infty}{2P_{\text{succ}}(\rho)} \right] \\
\leq \frac{1}{2} \left[ 1 + \frac{kd^2}{mP_{\text{succ}}(\rho)} \right] ,
\]
where $P_{\text{succ}}(\rho) := \text{tr}[\mathcal{E}_m(\rho)] \equiv \text{tr}[\tilde{\mathcal{E}}_m]$. ■

Approximation of the optimal $k$-copy cloning fidelity

Suppose that we are given $n$ copies of the state $|\psi_x\rangle \in \mathcal{H}$, $x \in X$ and that we want to produce $m$ approximate copies, whose quality is assessed by checking a random group of $k$ output systems. Let $\mathcal{C}_{n,m}$ be the quantum operation describing the cloning process. Since the $k$ systems are chosen at random, we can restrict our attention to quantum operations that are invariant under permutation of the output spaces. Conditional on the occurrence of the quantum operation $\mathcal{C}_{n,m}$ and on preparation of the input $|\psi_x\rangle$, the $k$-copy cloning fidelity is given by
\[
F_{k,x}[\mathcal{C}_{n,m}] = \frac{O_{k,x}[\mathcal{C}_{n,m}]}{P_x[\mathcal{C}_{n,m}]} \\
O_{k,x}[\mathcal{C}_{n,m}] := \langle \varphi_x \otimes \text{tr}_{m-k}[\mathcal{C}_{n,m}(\varphi_x^\otimes n)] | \varphi_x \otimes \mathcal{C}_{n,m} | \varphi_x \rangle,
\]
where $\varphi_x$ denotes the rank-one projector $\varphi_x := |\psi_x\rangle \langle \psi_x|$. Now, constructing the quantum operation $\mathcal{C}_{n,m}$ as in theorem 1 and using Eq. (63) we have
\[
|O_{k,x}[\mathcal{C}_{n,m}] - O[\mathcal{C}_{n,m}]| \leq \frac{2kd^2}{m} ,
\]
and, therefore,
\[
|F_{k,x}[\mathcal{C}_{n,m}] - F_{k,x}[\mathcal{C}_{n,m}]| \leq \frac{2kd^2}{mP_{\text{succ}}(\psi_x^\otimes n)} \quad \forall x \in X .
\]
In conclusion, as long as the probability of success $P_{\text{succ}}(\psi_x^\otimes n)$ is lower bounded by a finite value independent of $m$, the $k$-copy fidelities of the cloning processes $\mathcal{C}_{n,m}$ and $\mathcal{C}_{n,m}$. Since the bound holds for arbitrary quantum operations, in particular it holds for the quantum operation describing the optimal cloner with given probability of success.

It is immediate to extend the derivation to the Bayesian scenario where the input state $|\psi_x\rangle^\otimes n$ is given with probability $p_x$ and one considers the average fidelity and average success probability. Indeed, the average $k$-copy fidelity is given by
\[
F_k[\mathcal{C}_{n,m}] = \frac{O_k[\mathcal{C}_{n,m}]}{P[\mathcal{C}_{n,m}]} \\
O_k[\mathcal{C}_{n,m}] := \sum_x p_x O_{k,x}[\mathcal{C}_{n,m}] ,
\]
and one has the bound
\[
F_k[\mathcal{C}_{n,m}] \leq \frac{2kd^2}{mP_{\text{succ}}} ,
\]
$P_{\text{succ}}$ being the average success probability. Again, for every fixed value of the success probability, the fidelity of the optimal cloner is achieved by a PM protocol.

Lower bound on the average probability of success

We now show that for every fixed $n$, the probability of success of the optimal cloner is lower bounded by a finite value. Precisely, we prove the following
Lemma 1. The quantum operation $\mathcal{E}_{n,m}^*$ corresponding to the $n$-to-$m$ cloner that maximizes the fidelity in Eq. (64) can be chosen without loss of generality to have success probability $P_{\text{succ}}$ equal to $1/\|\tau^{-1}\|_{\infty}$, where $\tau$ is the average input state $\tau := \sum_x p_x \psi_x \otimes \bar{\psi}_x$ and $\|\tau^{-1}\|_{\infty}$ is the maximum eigenvalue of $\tau^{-1}$.

Proof. For a generic quantum operation $\mathcal{E}_{n,m}$, the $k$-copy fidelity can be expressed in terms of its Choi operator $C_{n,m}$ as

$$F_k[\mathcal{E}_{n,m}] = \frac{\text{tr}[\Omega \left(I^\otimes n \otimes \bar{\tau}C_{n,m}\right)]}{\text{tr}[\left(I^\otimes n \otimes \bar{\tau}\right)C_{n,m}]} \quad \Omega := \frac{1}{(m)} \sum_{S} \sum_{x} p_x (\psi^\oplus_x)_S \otimes \bar{\psi}_x^n,$$

where the outer summation runs over all $k$-element subsets $S$ of the output Hilbert spaces, $(\psi^\oplus_x)_S$ denotes the operator $\psi^\oplus_x$ acting on the Hilbert spaces in the set $S$, and $\bar{\tau}$ ($\bar{\psi}_x$) is the complex conjugate of $\tau$ ($\psi_x$). Following the arguments of [3, 4], one can easily show that the maximum fidelity over all quantum operations is given by

$$F_k^* = \left\|\left(I^\otimes n \otimes \bar{\tau}^{-\frac{1}{2}}\right)\Omega \left(I^\otimes n \otimes \bar{\tau}^{-\frac{1}{2}}\right)\right\|_{\infty}.$$

The maximum is achieved by choosing a Choi operator $C^*_{n,m}$ of the form

$$C^*_{n,m} = \gamma \left(I^\otimes n \otimes \bar{\tau}^{-\frac{1}{2}}\right) \langle \Psi \rangle \langle \bar{\Psi} \rangle \left(I^\otimes n \otimes \bar{\tau}^{-\frac{1}{2}}\right)$$

where $\gamma \geq 0$ is a proportionality constant and $\langle \Psi \rangle$ is the eigenvector of $\left(I^\otimes n \otimes \bar{\tau}^{-\frac{1}{2}}\right)\Omega \left(I^\otimes n \otimes \bar{\tau}^{-\frac{1}{2}}\right)$ with maximum eigenvalue. With this choice, the success probability $P[\mathcal{E}^*_{n,m}]$ is given by

$$P[\mathcal{E}^*_{n,m}] = \text{tr}[C^*_{n,m} \left(I^\otimes m \otimes \bar{\tau}\right)] = \gamma.$$

We now show that $\gamma$ can be always chosen to be larger than $1/\|\tau^{-1}\|_{\infty}$. To this purpose, note that the only constraint on $\gamma$ is that the quantum operation $\mathcal{E}^*_{n,m}$ must be trace non-increasing. Now, for a generic state $\rho$ one has

$$\text{tr}[\mathcal{E}^*_{n,m}(\rho)] = \gamma \text{ tr}[C^*_{n,m} \left(I^\otimes m \otimes \bar{\rho}\right)]$$

$$
\leq \gamma \|\bar{\rho}\|_{\infty} \leq \gamma \|\tau^{-1}\|_{\infty} \|\rho\|_{\infty} \leq \gamma \|\tau^{-1}\|_{\infty}.
$$

Hence, the choice $\gamma = 1/\|\tau^{-1}\|_{\infty}$ leads to a legitimate (trace non-increasing) quantum operation. \hfill \blacksquare

[1] G. Chiribella, Y. Yang, and A.C.-C. Yao, Nature Commun. 4, 2915 (2013).
[2] M. Marcus and H. Minc, A survey of matrix theory and matrix inequalities (Dover, New York, 1964).
[3] E. Bombieri and J. Vaaler, Invent. math. 73, 11 (1983).
[4] I. Borosh, M. Flahive, D. Rubin and B. Treybig, Proc. Amer. Math. Soc. 105, 844 (1989).
[5] G. Chiribella, Lecture Notes in Computer Science, 6519, 9 (2011).
[6] J. Fiurášek, Phys. Rev. A 70, 032308 (2004).
[7] G. Chiribella and J. Xie, Phys. Rev. Lett. 110, 213602 (2013).