Ultimate Limits for Quickest Quantum Change-Point Detection

Marco Fanizza,[1,*] Christoph Hirche,[2,†] and John Calsamiglia[1,‡]
[1]*Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain*
[2]*Center for Quantum Technologies, National University of Singapore Zentrum Mathematik, Technical University of Munich, 85748 Garching, Germany*

(Received 8 August 2022; revised 30 May 2023; accepted 5 June 2023; published 14 July 2023)

Detecting abrupt changes in data streams is crucial because they are often triggered by events that have important consequences if left unattended. Quickest change-point detection has become a vital sequential analysis primitive that aims at designing procedures that minimize the expected detection delay of a change subject to a bounded expected false alarm time. We put forward the quantum counterpart of this fundamental primitive on streams of quantum data. We give a lower bound on the mean minimum delay when the expected time of a false alarm is asymptotically large, under the most general quantum detection strategy, which is given by a sequence of adaptive collective (potentially weak) measurements on the growing string of quantum data. In addition, we give particular strategies based on repeated measurements on independent blocks of samples that asymptotically attain the lower bound and thereby establish the ultimate quantum limit for quickest change-point detection. Finally, we discuss online change-point detection in quantum channels.

DOI: 10.1103/PhysRevLett.131.020602

The detection of sudden changes of a stochastic random variable is one of the most fundamental problems in statistical analysis with a wide range of applications. For instance, in the context of industrial process monitoring [1,2], detecting sudden changes in sensor readings can help identify faults or equipment malfunctions that could lead to downtime or accidents if not addressed in a timely manner. In medical sciences [3,4], change-point detection can be used to detect the onset of an infectious disease outbreak or the onset of a medical emergency while monitoring the vital signs of a patient. Similar examples can be found in climate research [5,6], cybersecurity [7], and robotics [8,9] to name a few. Change-point detection has become a field of its own in classical statistical analysis [3,10–12] with large activity on the fundamental side, establishing optimal estimators and trade-off regions, in diverse idealized settings, and on the applied side developing statistical and machine learning techniques that operate with real life data streams.

Recently, the concept of change-point detection has been generalized to the quantum world [13–16]. Here, we have a device outputting quantum states. By default this device will output a certain state ρ, but from a given (random) point of time it will start producing the state σ. The goal is to identify this change point. In [13–15] the problem has been considered as an instance of hypothesis testing, where one collects a fixed number of quantum states and then tries to determine if and where a change point occurred. Since one requires the full sequence, this is usually considered offline change-point detection. For the special case of pure quantum states, solutions where given for the identification of the position of the change, optimizing the mean probability of error [14] and the probability of unambiguous identification [15]. For any practical application, however, one usually cannot wait for the entire sequence to be collected (it could potentially be infinite). Therefore, there is a strong motivation to consider online detection or quickest change-point detection: an algorithm that samples every copy sequentially and fires an alarm as soon as it detects the change. In this scenario, the natural quantities to consider are the time delay in detecting a change point versus the risk of a false alarm, i.e., falsely detecting a change when none has happened. In classical statistical analysis the most studied such algorithm is Page's cumulative sum (CUSUM) algorithm [17]. Apart from its computational simplicity, one of its most important features is its optimality under certain risk criteria, as shown first by Lorden [18] in the asymptotic setting and in [19,20] in the finite regime. In [16] online strategies for quantum change-point detection have been considered in the restricted scenario of pure states where unambiguous (local) identification is possible. In this case, the postchange state σ does not lie in the support of the default state ρ, and one can find a suitable measurement for which one of the outcomes can only be triggered by the postchange state, and thereby guarantees the absence of false alarms while keeping a finite mean detection delay. However, for realistic (mixed) states there is a trade-off between the false alarm time and the detection delay. In order to optimize this trade-off one needs to consider general sequential quantum strategies. This class of strategies has been recently studied in the context of sequential hypothesis of quantum states [21,22] and channels [23].

As the main contribution of this Letter we provide the ultimate quantum limit for quickest change-point detection: we give a lower bound to the mean detection delay that can be reached by the class of most general detection strategies with a given bounded (asymptotically large) expected false alarm time, and we provide a quantum version of the CUSUM algorithm, called QUSUM, and show that it asymptotically attains the aforementioned lower bound. In particular, this algorithm uses only l-local (projective) measurements on the incoming quantum states; still, asymptotically it cannot be outperformed by even a sequence of possibly weak, collective, and adaptive quantum measurements.

This Letter is structured as follows. We first present the problem in its simplest form and state the two main results, which are then proven in dedicated sections. Afterward, we comment on the implications of our results for the problem of change-point detection in sequences of quantum channels. We conclude by mentioning open problems.

Setting and results.—The change-point sequence is a sequence of d-dimensional states $\{\rho^{(n)}\}$, $n = 1, 2, \ldots$, such that if $n \leq \nu$, $\rho^{(n)} = \rho$ and if $n > \nu$, $\rho^{(n)} = \sigma$. At each step n, the algorithm receives a copy of the state $\rho^{(n)}$. The latter is then measured together with the current state of previously received copies by a joint quantum measurement, whose outcome determines whether to continue or to stop at step n and emit an alarm that signals that the change has occurred. Let us call T the random variable corresponding to the alarm time n at which stopping occurs for some given strategy. Let $\mathbb{E}_{\infty/\nu}$ denote expectation values with respect to some measurement acting sequentially on a sequence of copies of ρ (the change point never happens, $\nu = \infty$) or the change-point sequence for a specific finite ν. Similarly, the probability of an event E is denoted by $P_{\infty/\nu}(E)$. We leave the precise algorithm implicit, but it should be clear that it could be any sequential quantum measurement on the sequence $\{\rho^{(n)}\}$, with outcomes described by discrete random variables $\{X_1, \ldots, X_n\}$, where X_i is the outcome of the measurement realized after getting $\rho^{(i)}$. We also use X^n to denote the vector of random variables $\{X_1, \ldots, X_n\}$ and x^n to refer to the vector of values $\{x_1, \ldots, x_n\}$ they assume. We define the mean false alarm time as

$$\bar{T}_{\mathrm{FA}} = \mathbb{E}_\infty[T]. \tag{1}$$

We will consider families of strategies that have $\bar{T}_{\mathrm{FA}}$ larger than a constant. Having a large expected false alarm time is desirable to avoid stopping early, i.e., before the change ($T < \nu$). In addition, in order to quantify the response time or delay ($T - \nu > 0$), we define the so-called worst-worst case mean delay as [12]

$$\bar{\tau}^\star := \sup_{\nu \geq 0} \sup_{\substack{x^\nu \text{ with} \\ P_\infty(X^\nu = x^\nu) > 0}} \mathbb{E}_\nu[T - \nu | T > \nu, X^\nu = x^\nu]. \tag{2}$$

This figure of merit considers the worst mean delay over all possible locations of the change point and over all possible measurement outcomes before the change [24]

$$\bar{\tau} := \sup_{\nu \geq 0} \mathbb{E}_\nu[T - \nu | T > \nu] \leq \bar{\tau}^\star. \tag{3}$$

In the following we always assume $\operatorname{supp} \sigma \subseteq \operatorname{supp} \rho$. Otherwise, as in the pure case discussed above, there exists a projector Π such that $\mathrm{tr}[\Pi \rho] = 0$ and $\mathrm{tr}[\Pi \sigma] = c > 0$; therefore, the change can be detected with high probability in finite time and no false alarms (infinite $\bar{T}_{\mathrm{FA}}$). This also implies that the two entropic quantities that will play a prominent role here, $D(\sigma \| \rho) = \mathrm{tr}[\sigma(\log \sigma - \log \rho)]$ and $D_{\max}(\sigma \| \rho) = \inf\{\lambda \geq 0 : \sigma \leq 2^\lambda \rho\}$, are bounded. For strategies with fixed false alarm time $\bar{T}_{\mathrm{FA}}$, optimal strategies are those that minimize $\bar{\tau}^*$. We show that the asymptotic behavior of the optimal $\bar{\tau}^*$ for large $\bar{T}_{\mathrm{FA}}$ is

$$\bar{\tau}^* \sim \frac{\log \bar{T}_{\mathrm{FA}}}{D(\sigma \| \rho)}. \tag{4}$$

We prove this via two theorems, which respectively provide an upper bound and a lower bound to $\bar{\tau}^\star$.

Theorem 1: Achievability.—Given a change-point problem with two finite-dimensional states ρ and σ, $D(\sigma \| \rho) < \infty$, for any $\epsilon > 0$, and $\bar{T}_{\mathrm{FA}}$ large enough, there is a QUSUM algorithm such that $\bar{\tau}^\star \leq (\log \bar{T}_{\mathrm{FA}})/[D(\sigma \| \rho)(1 - \epsilon)] + O(1)$.

QUSUM is, as the name suggests, a quantum version of the classical CUSUM algorithm; see also [17]. We show that the performance of QUSUM is asymptotically optimal.

Theorem 2: Optimality.—Any algorithm for a change-point problem with two finite-dimensional states ρ and σ, $D(\sigma \| \rho) < \infty$, with expected false alarm $\bar{T}_{\mathrm{FA}}$ must satisfy $\bar{\tau}^\star \geq \bar{\tau} \geq (1 - \epsilon)[\log \bar{T}_{\mathrm{FA}} / D(\sigma \| \rho)][1 + o(1)]$ for any $\epsilon > 0$.

In the following we give the proof of these theorems and discuss some generalizations thereof.

Achievability.—We will prove the achievability in two steps. First, we study the detection delay of a simple algorithm that repeats the same measurement on each individual incoming state and then extend it to the case where repeated measurements are performed on blocks of a fixed number of copies. If a fixed positive operator-valued measure (POVM) $\{M_{x_i}\}$ is applied to the ith state, outcome x_i will appear with probability

$$p(x_i) = \mathrm{tr}[M_{x_i} \rho] \quad \text{or} \quad q(x_i) = \mathrm{tr}[M_{x_i} \sigma], \tag{5}$$

depending on the underlying state. We can define the log-likelihood ratio and their partial sums,

$$Z_i = \log \frac{q(x_i)}{p(x_i)}, \qquad Z_j^n = \sum_{i=j}^n Z_i. \tag{6}$$

It can easily be seen that the mean of the first is given by the relative entropy

$$\mathbb{E}_q[Z_i] = \sum_i q_i \log\frac{p_i}{q_i} =: D(q\|p), \tag{7}$$

where $\mathbb{E}_q$ denote expectation values with respect to the probability distribution q. Note that second quantity in Eq. (6) with $j = \nu + 1$ gives the log-likelihood ratio of a sequence of independent identically distributed (IID) outcomes x^n sampled from either a source with change point at ν or from a source with no change: $\lambda_n^{(\nu)} := \log[P_\nu(X^k = x^k)/P_\infty(X^n = x^n)] = Z_{\nu+1}^n$. Following the CUSUM algorithm we can fix a threshold value h and consider for each possible change point j a stopping time

$$T_j = \min\{n \geq j : Z_j^k \geq h\}, \tag{8}$$

where we define $T_j = \infty$ if $\{n \geq j : Z_j^n \geq h\} = \emptyset$, and given these, we define the CUSUM stopping time

$$T^\star = \min_{j\geq 1} T_j. \tag{9}$$

From the above definition of $\lambda_n^{(\nu)}$ it follows that $T^\star$ can be understood as the first time when, given the current measurement record, the probability of having had a change in the past is e^h times more likely than having no change—see the Supplemental Material [25] (SM) for a commented picture of a classical CUSUM test. We can now use a result by Lorden (Theorem 2 in [18]). In our notation, we have that if $P_\infty(T_1 < \infty) \leq \alpha$,

$$\bar{T}_{\mathrm{FA}} = \mathbb{E}_\infty[T^\star] \geq \frac{1}{\alpha}, \quad \text{and} \quad \bar{\tau}^\star \leq \mathbb{E}_0[T_1]. \tag{10}$$

The above premise holds since

$$\begin{aligned} P_\infty(T_1 < \infty) &= \mathbb{E}_\infty[I_{T_1<\infty}] = \mathbb{E}_0\left[\frac{p(x^{T_1})}{q(x^{T_1})} I_{T_1<\infty}\right] \\ &= \mathbb{E}_0[\mathrm{e}^{-Z_1^{T_1}} I_{T_1<\infty}] \leq \mathrm{e}^{-h} =: \alpha, \end{aligned} \tag{11}$$

where in the second equality we have used the change of measure in order switch the distributions on which the expectation value is computed: $\mathbb{E}_p[f(x)] = \sum_x p(x)f(x) = \sum_x q(x)[p(x)/q(x)]f(x) = \mathbb{E}_q[(p(x)/q(x))f(x)]$, and the last equality holds because at the stopping time $Z_1^{T_1}$ is necessarily larger than h. Finally, using Wald's identity [29] (see also Chap. 3 of [12] and SM), $\mathbb{E}_0[Z_1^{T_1}] = \mathbb{E}_0[\sum_{i=1}^{T_1} Z_i] = \mathbb{E}_0[Z_1]\mathbb{E}_0[T_1]$:

$$\mathbb{E}_0[T_1] = \frac{\mathbb{E}_0[Z_1^{T_1}]}{\mathbb{E}_0[Z_1]} = \frac{h + \mathbb{E}[s]}{D(q\|p)} \to \frac{h}{D(q\|p)}, \tag{12}$$

when $h \to \infty$, where $s := Z_1^{T_1} - h$ is the "overshoot" and the limit holds since $Z_i < \infty$. Putting all the results together we get, using Eq. (12) in Eq. (10),

$$\bar{\tau}^\star \leq \mathbb{E}_0[T_1] = \frac{h}{D(q\|p)} + O(1) \leq \frac{\log \bar{T}_{\mathrm{FA}}}{D(q\|p)} + O(1). \tag{13}$$

Optimizing over all measurements gives us the achievable trade-off for this particular strategy,

$$\bar{\tau}^\star \leq \frac{\log \bar{T}_{\mathrm{FA}}}{D_M(\sigma\|\rho)} + O(1) \text{ as } \bar{T}_{\mathrm{FA}} \to \infty, \tag{14}$$

in terms of the measured relative entropy $D_M(\sigma\|\rho) := \sup_{\{M_i\}\mathrm{POVM}} D(q\|p)$ [30,31]. Note that already projection-valued measurements achieve the measured relative entropy [32].

Now, more generally, instead of measuring each copy of $\rho^{(i)}$ separately, the QUSUM algorithm is based on performing a joint measurement on blocks of l states that are either $\rho^{\otimes l}$ or $\sigma^{\otimes l}$ (assuming the change point happens at a multiple of l; see the SM for the general case). The above trade-off is now easily modified to $\bar{\tau}^\star \leq \log(\bar{T}_{FA}/l)/[D_M(\sigma^{\otimes l}\|\rho^{\otimes l})/l]$.

If ρ and σ are states of a d-dimensional Hilbert space, Hayashi showed (Theorem 2 of [31]) that for any σ and l there is a projection-valued measurement $\{M_{x_i}^{(l)}\}$, depending only on ρ, such that if $p^{(l)}(i) = \mathrm{tr}[M_{x_i}^{(l)}\rho^{\otimes l}]$, $q^{(l)}(i) = \mathrm{tr}[M_{x_i}^{(l)}\sigma^{\otimes l}]$, $\forall\ \sigma$

$$D(\sigma\|\rho) - \frac{(d-1)\log(l+1)}{l} \leq \frac{1}{l} D(q^{(l)}\|p^{(l)}) \leq D(\sigma\|\rho). \tag{15}$$

Choosing l such that $(1/l)D(q^{(l)}\|p^{(l)}) \geq D(\sigma\|\rho)(1-\epsilon)$ we obtain the statement of the theorem [33]. Since we are considering $h \to \infty$ we can choose l arbitrarily large and consider the limit $\lim_{l\to\infty}(1/l)D_M(\sigma^{\otimes l}\|\rho^{\otimes l}) = D(\sigma\|\rho)$. This implies that in the asymptotic limit of large h and l we have

$$\bar{\tau}^\star \leq \frac{\log \bar{T}_{\mathrm{FA}}[1 + o(1)]}{D(\sigma\|\rho)}. \tag{16}$$

In Fig. 1 we illustrate QUSUM tests for qubit states, where we use the measurement of [31], but also the enhanced class of j-angle-optimized measurements, which outperform Hayashi's measurement by construction. These measurements begin with a projection on subspaces with fixed total angular momentum j, followed by a measurement in a product basis, which is fixed in [31]. The generalization consists of optimizing the product basis for each value of j (see SM). In particular, j-angle-optimized measurements,

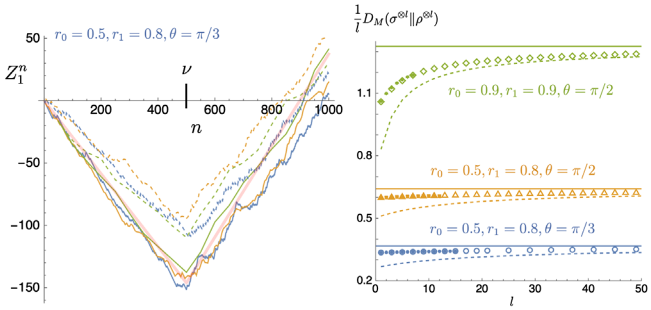

FIG. 1. QUSUM for qubit states with Bloch vector of lengths r_0 (ρ) and r_1 (σ) and relative angle θ. Left: Log-likelihood stochastic trajectory for Hayashi's (dashed) and j-angle-optimized (solid) block-sampling strategies for block lengths $l = 1$, 5, 50 (blue, orange, green). Larger l approach the optimal rate given by the quantum relative entropy (thick red line). Right: Measured relative entropy (per copy) for several strategies: Hayashi (dashed), j-angle-optimized (open markers), optimal from SDP (dots), and asymptotically attainable upper bound given by quantum relative entropy.

as well as Hayashi's, are based on Schur sampling [37–39] and are hence efficiently implementable. The figure also shows the measured relative entropy for different block lengths, which determines the performance of block sampling strategies. The enhanced class shows a noticeably quicker convergence to the relative entropy as the block size increases. The (optimal) measured relative entropy, whose approximation is computed by a semidefinite program (SDP) [32,40], is shown to be very close to the values obtained for the j-angle-optimized measurement.

In addition, since the measurement in [31] achieving this bound does not depend on σ, we can also generalize this achievability result in the case where the state after the change point is unknown and belongs to a finite family of states $\mathcal{S}$. In this case, we get asymptotically

$$\bar{\tau}^\star \leq \frac{\log \bar{T}_{\mathrm{FA}}[1 + o(1)]}{\min_{\sigma \in \mathcal{S}} D(\sigma \| \rho)}. \tag{17}$$

Guarantees in the case of infinite families can be obtained if there exists a suitable discretization of $\mathcal{S}$ (see SM for details).

Optimality.—In this section, we prove Theorem 2, which shows that no strategy can attain a better trade-off between the detection delay and the false alarm time than that given by the quantum relative entropy. To that end we first have to define the considered class of strategies. Since they can make use of general adaptive measurements, it is necessary to specify how the state changes after the measurement, using quantum instruments. A quantum instrument is described by a set of completely positive trace-nonincreasing maps $\{\mathcal{M}_x(\cdot)\}$, with $\sum_x \mathcal{M}_x(\cdot)$ trace preserving. For a fixed measurement outcome x the (normalized) postmeasurement state is $\rho_x = \mathcal{M}_x(\rho)/\mathrm{tr}[\mathcal{M}_x(\rho)]$, where $\mathrm{tr}[\mathcal{M}_x(\rho)]$ corresponds to the probability of obtaining x given the state ρ. In our setting, at each step i, we get a fresh copy of $\rho^{(i)}$, which is either ρ or σ. Let $\rho_{x^{i-1}}$ be the postmeasurement state of the $(i-1)$th step; then we apply the ith quantum instrument $\mathcal{M}^i$ (possibly depending on previous records x^{i-1}) as $\mathcal{M}^i(\rho^{(i)} \otimes \rho_{x^{i-1}})$, receiving a classical output x_i and a new postmeasurement state $\rho_{x^i} = \mathcal{M}^i_{x_i}(\rho^{(i)} \otimes \rho_{x^{i-1}})/\mathrm{tr}[\mathcal{M}^i_{x_i}(\rho^{(i)} \otimes \rho_{x^{i-1}})]$. We denote the postmeasurement states as $\rho^{(\infty)}_{x^i}$ if they originate from a sequence with no change point, and $\rho^{(\nu)}_{x^i}$ if they come from a sequence with change point ν. We denote $p(x_i|x^{i-1}) = \mathrm{tr}[\mathcal{M}^i_{x_i}(\rho \otimes \rho^{\infty}_{x^{i-1}})]$, and $q^{(\nu)}(x_i|x^{i-1}) = \mathrm{tr}[\mathcal{M}^i_{x_i}(\sigma \otimes \rho^{(\nu)}_{x^{i-1}})]$. We now define the local and cumulative log-likelihood ratios at step i for a candidate change point ν:

$$Z_i^{(\nu)} = \log \frac{q^{(\nu)}(x_i|x^{i-1})}{p(x_i|x^{i-1})}, \qquad \lambda_n^{(\nu)} = \sum_{i=\nu+1}^{n} Z_i^{(\nu)}. \tag{18}$$

Note that for a fixed sequence x^i we can always write $p(x^i) = \mathrm{tr}[M_{x^i} \rho^{\otimes i}]$ for a joint measurement $\{M_{x^i}\}$ giving a sequence of outcomes x^i.

While we still get a sequence of classical measurement outcomes as a result, these can now be highly correlated and the usual techniques for IID distributions do not longer apply. In the following we will make heavy use of a result initially stated by Lai [41] and reformulated in [12], which we adapt in a form that is applicable to our case, and give the proof in the SM for completeness [42].

Theorem 3.—For a change-point model with log-likelihoods $Z_i^{(\nu)}$ and $\epsilon > 0$, no strategies can exceed the trade-off given by $\bar{\tau}^* \geq (1-\epsilon)(\log \bar{T}_{\mathrm{FA}}/I)[1 + o(1)]$, for large $\bar{T}_{\mathrm{FA}}$, for any I that satisfies the condition

$$\lim_{n\to\infty} \sup_{\nu \geq 0} \sup_{\substack{x^\nu \text{ with} \\ P_\infty(X^\nu = x^\nu) > 0}} P^*_\nu(x^\nu) = 0 \quad \text{where}$$

$$P^*_\nu(x^\nu) \coloneqq P_\nu\left\{\max_{i \leq n} \lambda^{(\nu)}_{\nu+i} \geq I(1+\epsilon)n \,\middle|\, X^\nu = x^\nu\right\}. \tag{19}$$

In loose terms, the rate I in this theorem has to be such that for any value $I' > I$ arbitrarily close to I, the stochastic trajectories exhibited by $\lambda^{(\nu)}_{\nu+i}$ (equivalent to those in the lhs of Fig. 1) that exceed the value of $I'n$ between the change point ν and $\nu + n$ occur with a vanishing probability as n increases.

The challenging art is to determine the smallest I such that Eq. (19) holds for any $\epsilon > 0$. For IID distributions the relative entropy $I = D(q\|p)$ can be seen to satisfy this criterion. In our case we need to find the rate I considering that the underlying probability distribution can be produced by the most general kind of measurement strategy.

We start by getting rid of the supremum over ν. Note that all states up to position ν will be ρ as in the case that there is no change. At position $\nu+1$ we will therefore try to discriminate between two states: $\sigma\otimes\rho_{x^\nu}$ and $\rho\otimes\rho_{x^\nu}$. It is now easy to see that for any ν and any measurement on the sequence of states, there also exists a measurement in the case of $\nu=0$ that results in the same probability distribution after the change point, consisting of simply preparing the state ρ_{x^ν} and then applying the original strategy. It follows that we can without loss of generality set $\nu=0$ and therefore also omit the essential supremum.

We will now bound P_0^* based on the strong converse for quantum Stein's lemma [43], which states that if a sequence of binary tests $\{M_0^{(n)}, M_1^{(n)}=\mathbb{1}-M_0^{(n)}\}$ is such that $\mathrm{tr}[M_1^{(n)}\rho^{\otimes n}]\leq \mathrm{e}^{-n[D(\sigma\|\rho)+\delta]}$ for some $\delta>0$, then $\lim_{n\to\infty}\mathrm{tr}[M_1^{(n)}\sigma^{\otimes n}]=0$. Let us denote the log-likelihood ratio of the outcome sequence x^i for $\nu=0$ as λ_{x^i}. Define the set $S_i \coloneqq \{x^i|\lambda_{x^i}\geq nI(1+\epsilon), \lambda_{x^j}<nI(1+\epsilon)\ \ \forall\ \ j<i\}$.

We have the following chain of equalities:

$$\begin{aligned}
P^{(0)}\left[\max_{1\leq i\leq n}\lambda_{x^i}\geq nI(1+\epsilon)\right]
&=\sum_{\substack{x^n:\\ \max_{1\leq i\leq n}\lambda_{x^i}\geq nI(1+\epsilon)}} q^{(0)}(x^n)\\
&=\sum_{1\leq i\leq n}\sum_{x^i\in S_i} q^{(0)}(x^i)=\sum_{1\leq i\leq n}\sum_{x^i\in S_i}\mathrm{tr}\left[M_{x^i}^{(i)}\sigma^{\otimes i}\right]\\
&=\sum_{1\leq i\leq n}\sum_{x^i\in S_i}\mathrm{tr}\left[M_{x^i}^{(i)}\otimes \mathbb{1}^{\otimes(n-i)}\sigma^{\otimes n}\right].
\end{aligned} \tag{20}$$

Defining the binary POVM $\{\tilde{M}_0^i, \tilde{M}_1^n=\mathbb{1}-\tilde{M}_0^n\}$, with $\tilde{M}_1^n=\sum_{1\leq i\leq n}\sum_{x^i\in S_i}M_{x^i}^{(i)}\otimes\mathbb{1}^{\otimes(n-i)}$, we get $P^{(0)}[\max_{1\leq i\leq n}\lambda_{x^i}\geq nI(1+\epsilon)]=\mathrm{tr}[\tilde{M}_1^n\sigma^{\otimes n}]$. Also, since $q(x^i)=\mathrm{e}^{\lambda_{x^i}}p(x^i)\geq \mathrm{e}^{nI(1+\epsilon)}p(x^i)\ \forall\ \ x^i\in S_i$,

$$\begin{aligned}
\mathrm{tr}[\tilde{M}_1^n\rho^{\otimes n}]&=\sum_{1\leq i\leq n}\sum_{x^i\in S_i}p(x^i)\\
&\leq \mathrm{e}^{-nI(1+\epsilon)}\sum_{1\leq i\leq n}\sum_{x^i\in S_i}q(x^i)\\
&=\mathrm{e}^{-nI(1+\epsilon)}\mathrm{tr}[\tilde{M}_1^n\sigma^{\otimes n}]\leq \mathrm{e}^{-nI(1+\epsilon)}.
\end{aligned} \tag{21}$$

By the strong converse, this means that if $I\geq D(\sigma\|\rho)$, $\lim_{n\to\infty}P_{(0)}^*=\lim_{n\to\infty}\mathrm{tr}[\tilde{M}_1^n\sigma^{\otimes n}]=0$.

This proves the optimality, $\bar{\tau}^\star\geq\bar{\tau}\geq(1-\epsilon)$ $[\log\bar{T}_{\mathrm{FA}}[1+o(1)]/D(\sigma\|\rho)]$. Even more so, optimality holds also for a collection $\mathcal{S}$ of possible states after the change, with $I=\min_{\sigma\in\mathcal{S}}D(\sigma\|\rho)$.

Change point with channels.—One can define an analogous change-point problem for quantum channels. In this setting, at step n the algorithm receives a black-box use of a channel $\mathcal{N}^{(n)}$, which is $\mathcal{N}$ if $n\leq\nu$ and $\mathcal{M}$ if $k>\nu$, and the most general quantum strategies are allowed, such as using quantum memory, adapting operations between each use of the channel (see SM for a precise definition). In this case, we can leverage the achievability result for states to obtain $\bar{\tau}^\star\leq\log\bar{T}_{\mathrm{FA}}[1+o(1)]/D^\infty(\mathcal{M}\|\mathcal{N})$ where $D^\infty(\mathcal{M}\|\mathcal{N})=\lim_{l\to\infty}\sup_\rho(1/l)D(\mathcal{M}^{\otimes l}(\rho)\|\mathcal{N}^{\otimes l}(\rho))$ (here and in the following the input state in maximization can be any state entangled with an arbitrarily large reference system). On the other hand, we can adapt the lower bound proof using a known strong converse [44], obtaining $\bar{\tau}^\star\geq(1-\epsilon)[\log\bar{T}_{\mathrm{FA}}[1+o(1)]/\tilde{D}_1^\infty(\mathcal{M}\|\mathcal{N})]$, $\forall\ \epsilon>0$, where $\tilde{D}_\alpha^\infty(\mathcal{M}\|\mathcal{N})=\lim_{l\to\infty}\sup_\rho(1/l)\,\tilde{D}_\alpha(\mathcal{M}^{\otimes l}(\rho)\|\mathcal{N}^{\otimes l}(\rho))$, $\tilde{D}_\alpha(\rho\|\sigma)=(1/\alpha-1)\log\mathrm{tr}[\sigma^{(1-\alpha/2\alpha)}\rho\sigma^{(1-\alpha/2\alpha)}]^\alpha$, and $\tilde{D}_1^\infty(\mathcal{M}\|\mathcal{N})=\lim_{\alpha\to1}\tilde{D}_\alpha^\infty(\mathcal{M}\|\mathcal{N})$. The quantities in the two bounds have been conjectured to coincide [45].

Conclusions.—We have showed asymptotic optimality of the QUSUM algorithm, with a trade-off given by the relative entropy, solving the quickest change-point detection problem for quantum states in the asymptotic setting. Our results apply also to the setting when the state after the change is not known. We have proposed a measurement scheme and show numerically that it approaches the optimal trade-off with finite block lengths. It remains unclear how to address the optimality of the quickest detection for finite number of samples. In the asymptotic setting, it would be interesting to find achievability results for non-IID states, especially those for which a strong converse can be found.

We thank Andreas Winter and Milán Mosonyi for useful discussions. This work was supported by the QuantERA grant C'MON-QSENS!, by Spanish MICINN PCI2019-111869-2 and Spanish Agencia Estatal de Investigación, project PID2019–107609GB-I00/AEI/10.13039/501100011033 and cofunded by the European Union Regional Development Fund within the ERDF Operational Program of Catalunya (project QuantumCat, ref. 001-P-001644). C. H. has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Grant Agreement No. H2020-MSCA-IF-2020-101025848. J. C. also acknowledges support from ICREA Academia award. M. F. is supported by Juan de la Cierva—Formaciòn (Spanish MICIN project FJC2021-047404-I), with funding from MCIN/AEI/10.13039/501100011033 and European Union "NextGenerationEU"/PRTR.

*Corresponding author.
marco.fanizza@uab.cat
†Corresponding author.
christoph.hirche@gmail.com
‡Corresponding author.
john.calsamiglia@uab.cat

[1] D. M. Hawkins, P. Qiu, and C. W. Kang, The changepoint model for statistical process control, J. Qual. Technol. **35**, 355 (2003).

[2] T. L. Lai, Sequential changepoint detection in quality control and dynamical systems, J. R. Stat. Soc. Ser. B **57**, 613 (1995).

[3] J. Chen and A. K. Gupta, *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance* (Springer Science & Business Media, New York, 2011).

[4] D. Rosenfield, E. Zhou, F. H. Wilhelm, A. Conrad, W. T. Roth, and A. E. Meuret, Change point analysis for longitudinal physiological data: Detection of cardio-respiratory changes preceding panic attacks, Biol. Psychol. **84**, 112 (2010).

[5] C. Gallagher, R. Lund, and M. Robbins, Changepoint detection in climate time series with long-term trends, J. Clim. **26**, 4994 (2013).

[6] J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, A review and comparison of changepoint detection techniques for climate data, J. Appl. Meteorol. Climatol. **46**, 900 (2007).

[7] M. N. Kurt, Y. Yılmaz, and X. Wang, Distributed quickest detection of cyber-attacks in smart grid, IEEE Trans. Inf. Forensics Secur. **13**, 2015 (2018).

[8] S. Niekum, S. Osentoski, C. G. Atkeson, and A. G. Barto, Online Bayesian changepoint detection for articulated motion models, in *Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, New York, 2015), pp. 1468–1475.

[9] A. Ranganathan, Pliss: Labeling places using online changepoint detection, Auton. Robots **32**, 351 (2012).

[10] M. Basseville, I. V. Nikiforov *et al.*, *Detection of Abrupt Changes: Theory and Application* (Prentice Hall, Englewood Cliffs, 1993), Vol. 104.

[11] H. V. Poor and O. Hadjiliadis, *Quickest Detection* (Cambridge University Press, Cambridge, England, 2008).

[12] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection* (Chapman and Hall/CRC, Boca Raton, 2014).

[13] D. Akimoto and M. Hayashi, Discrimination of the change point in a quantum setting, Phys. Rev. A **83**, 052328 (2011).

[14] G. Sentís, E. Bagan, J. Calsamiglia, G. Chiribella, and R. Munoz-Tapia, Quantum Change Point, Phys. Rev. Lett. **117**, 150502 (2016).

[15] G. Sentís, J. Calsamiglia, and R. Muñoz-Tapia, Exact Identification of a Quantum Change Point, Phys. Rev. Lett. **119**, 140506 (2017).

[16] G. Sentís, E. Martínez-Vargas, and R. Muñoz-Tapia, Online strategies for exactly identifying a quantum change point, Phys. Rev. A **98**, 052305 (2018).

[17] E. S. Page, Continuous inspection schemes, Biometrika **41**, 100 (1954).

[18] G. Lorden *et al.*, Procedures for reacting to a change in distribution, Ann. Math. Stat. **42**, 1897 (1971).

[19] G. V. Moustakides *et al.*, Optimal stopping times for detecting changes in distributions, Ann. Stat. **14**, 1379 (1986).

[20] Y. Ritov, Decision theoretic optimality of the cusum procedure, Ann. Stat. **18**, 1464 (1990).

[21] E. Martinez Vargas, C. Hirche, G. Sentís, M. Skotiniotis, M. Carrizo, R. Muñoz-Tapia, and J. Calsamiglia, Quantum Sequential Hypothesis Testing, Phys. Rev. Lett. **126**, 180502 (2021).

[22] Y. Li, V. Y. Tan, and M. Tomamichel, Optimal adaptive strategies for sequential quantum hypothesis testing, Commun. Math. Phys. **392**, 993 (2022).

[23] Y. Li, C. Hirche, and M. Tomamichel, Sequential quantum channel discrimination, in *Proceedings of the 2022 IEEE International Symposium on Information Theory (ISIT)* (IEEE press, New York, 2022).

[24] If POVMs with nondiscrete outcomes are allowed, then one may consider also $\bar{\tau}^\star \coloneqq \sup_{\nu\geq 0}\operatorname{ess\,sup}\mathbb{E}_\nu[T-\nu|T>\nu, X_1, \ldots, X_\nu]$, where the essential supremum means that the optimization is done a supremum over set of measurement outcomes of nonzero measure, where the measure of a set is given by its probability. For simplicity, in this Letter we consider only measurement with discrete outcomes.

[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.131.020602 for extended discussion of technical derivations and deferred proofs, which includes Refs. [26–28].

[26] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, Amortized channel divergence for asymptotic quantum channel discrimination, Lett. Math. Phys. **110**, 2277 (2020).

[27] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability for quantum channels, Phys. Rev. Res. **1**, 033169 (2019).

[28] K. Fang, O. Fawzi, R. Renner, and D. Sutter, Chain Rule for the Quantum Relative Entropy, Phys. Rev. Lett. **124**, 100501 (2020).

[29] A. Wald, *Sequential Analysis* (Dover Publications, INC., Mineola, New York, 2004).

[30] D. Petz, A variational expression for the relative entropy, Commun. Math. Phys. **114**, 345 (1988).

[31] M. Hayashi, Asymptotics of quantum relative entropy from a representation theoretical viewpoint, J. Phys. A **34**, 3413 (2001).

[32] M. Berta, O. Fawzi, and M. Tomamichel, On variational expressions for quantum relative entropies, Lett. Math. Phys. **107**, 2239 (2017).

[33] Notice also that an upper bound on the sufficient size of the blocks to get the achievability in the theorem is $k = \tilde{O}(d/[\epsilon D(\sigma\|\rho)])$ (hiding logarithmic factors), with an improved dependence on the dimension with respect to optimal tomography in trace distance, which requires $\Theta(d^2)$ copies [34,35]. In the SM we argue that for any fixed pair of states one can in fact choose l depending only on relative entropies of the pair, therefore not explicitly on the dimension, based on [36].

[34] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, Sample-optimal tomography of quantum states, IEEE Trans. Inf. Theory **63**, 5628 (2017).

[35] R. O'Donnell and J. Wright, Efficient quantum Tomography, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2016), pp. 899–912.

[36] K. M. Audenaert, M. Mosonyi, and F. Verstraete, Quantum state discrimination bounds for finite sample size, J. Math. Phys. (N.Y.) **53**, 122205 (2012).

[37] D. Bacon, I. L. Chuang, and A. W. Harrow, Efficient Quantum Circuits for Schur and Clebsch-Gordan Transforms, Phys. Rev. Lett. **97**, 170502 (2006).

[38] A. W. Harrow, Applications of coherent classical communication and Schur duality to quantum information theory, Ph.D. thesis, MIT, 2005.
[39] H. Krovi, An efficient high dimensional quantum Schur transform, Quantum **3**, 122 (2019).
[40] H. Fawzi, J. Saunderson, and P. A. Parrilo, Semidefinite approximations of the matrix logarithm, Found. Comput. Math. **19**, 259 (2019).
[41] T. L. Lai, Information bounds and quick detection of parameter changes in stochastic systems, IEEE Trans. Inf. Theory **44**, 2917 (1998).
[42] Reference [41] shows that the proof applies also when the probability in presence of a change point, conditioned on past events, can depend on the change-point location. The statement of the theorem in [41] has an extra condition about the convergence of the log-likelihood that is not needed for our purposes and expresses the trade-off only as a limit for large $\log \bar{T}_{\mathrm{FA}}$ and $\epsilon \to 0$, while we state it with finite ϵ.
[43] T. Ogawa and H. Nagaoka, Strong converse and Stein's lemma in quantum hypothesis testing, IEEE Trans. Inf. Theory **46**, 2428 (2000).
[44] H. Fawzi and O. Fawzi, Defining quantum divergences via convex optimization, Quantum **5**, 387 (2021).
[45] K. Fang, G. Gour, and X. Wang, Towards the ultimate limits of quantum channel discrimination, arXiv:2110.14842.