Quickest Detection for Human-Sensor Systems Using Quantum Decision Theory

Luke Snow ©, Graduate Student Member, IEEE, Vikram Krishnamurthy ©, Fellow, IEEE, and Brian M. Sadler ©, Life Fellow, IEEE

Abstract—A sensor observes an underlying state of nature in noise, computes a posterior probability of this state, and provides this posterior as a recommendation to a human decision-maker. The human then makes decisions based on this recommendation. By recording these human decisions over time, how can a quickest detector detect a change in the underlying state? This paper addresses the above human-sensor interface problem. This framework generalizes classical quickest detection since the detector only has access to human decisions rather than noisy signals with known statistical properties. We utilize a Quantum Decision Theory model (widely studied in mathematical psychology) to capture peculiarities of human decision making. We provide theoretical guarantees for the quickest detector’s optimal policy, such as a threshold structure, a bound on performance, robustness to variations in the decision-model parameters, and increased performance under ‘rational’ human decisions. Finally, we illustrate a numerical implementation of this quickest detector in the context of the Prisoner’s Dilemma problem, in which it has been observed that Quantum Decision Theory can uniquely model empirically tested violations of the sure-thing principle.

Index Terms—Quickest change detection, quantum decision making, Blackwell dominance, human-sensor interface.

Glossary of Symbols

Quantum Decision Model

| Symbol | Definition |
|--------|------------|
| \( \mathbb{R} \) | real number field |
| \( \mathbb{C} \) | complex number field |
| \( X, \mathcal{E}, K \) | state space, element of, cardinality |
| \( A, \alpha, \lambda, \phi \) | action space, element of, cardinality |
| \( H_X, H_A \) | state and action Hilbert spaces |
| \( H = H_X \otimes H_A \) | joint Hilbert space |
| \( |\psi_j\rangle \in \mathbb{C}^{AK \times 1} \) | state in \( H \) |
| \( \langle\psi_j| \in \mathbb{C}^{1 \times AK} \) | transposed state in \( H \) |

\( \rho_t \) | psychological state |
| \( \mathcal{L}(\alpha, \lambda, \phi) \) | Lindbladian operator |
| \( (\alpha, \lambda, \phi) \) | psychological parameters |
| \( u : X \times A \to \mathbb{R} \) | utility function |
| \( P_n \) | projector onto subspace \( a_i \otimes X \subset H \) |
| \( \Gamma(\cdot) \) | steady-state action distribution |
| \( \text{Tr}(\cdot) \) | Trace operation |

Quickest Change Detection

| Symbol | Definition |
|--------|------------|
| \( n \in \mathbb{N} \) | discrete time |
| \( x_n \in \{1, 2\} \) | state of nature |
| \( B_{a,y} \) | noisy observation density |
| \( \tau^0 \) | jump time of state of nature |
| \( \pi_n \) | posterior formed by QCD at time \( n \) |
| \( \eta_n \) | posterior formed by sensor at time \( n \) |
| \( u_n \) | action taken by QCD |
| \( f, \delta \) | false alarm, delay penalties |
| \( \mu^*(\pi) \) | QCD optimal policy |
| \( p, P \) | transition probability, matrix |
| \( J_\mu \) | expected cost under policy \( \mu \) |

I. INTRODUCTION

THE problem of ‘quickest detection’ involves a detector aiming to declare that an underlying state has changed, by observing noisy measurements of the state and subject to delay and false alarm penalties. This classical problem [1] is fundamental to statistical signal processing [2], [3], [4], and has applications in monitoring power networks [5], sensor networks [6], internet traffic [7], among numerous others. In this paper we construct and analyze a Bayesian sequential quickest detection framework to detect a change in an underlying state by observing human decisions that are influenced by the state. Specifically, a noisy sensor computes the posterior probability of an underlying state, and provides this posterior as a recommendation to a human, who then makes a decision. The quickest detector observes a sequence of human decisions and aims to detect a change in the underlying state. Previous efforts [8], [9], [10] have studied this problem using certain microeconomic models for human decision-making. These models tend to treat the human as perfectly rational with infinite computational resources for determining an optimal action. However, empirical studies [11] reveal that humans routinely behave suboptimally and violate traditional axioms of microeconomic decision-making.

Luke Snow and Vikram Krishnamurthy are with the Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: las474@cornell.edu.)

Brian M. Sadler is with the U.S. Army Research Laboratory, Adelphi, MD 20783 USA.

Digital Object Identifier 10.1109/TSP.2023.3346186
Motivated by recent work in mathematical psychology, we employ ‘Quantum Decision Theory’ to model the human decisions in our framework. The model we employ [12] captures salient properties of human decision making that cannot be modeled by traditional, such as microeconomic, models. Indeed, it has been shown that a model generalizing to axioms resting on quantum probability is sufficient to capture observed violations of traditional decision-making axioms [13], [14], [15]. In this paper we provide structural results to characterize the optimal performance of the detector which attempts to detect an underlying state change by observing human decisions which are produced via Quantum Decision Theory.

A. Human Sensor Based Change Point Detection

The quickest detection framework of this paper is schematically illustrated in Fig. 1. An underlying state changes at a geometrically distributed unknown time. At each time instant, a sensor obtains a noisy measurement of the underlying state, computes the posterior probability of the state, and provides this information to a human decision maker (e.g., as a recommendation). The human uses this information to choose an action at each time instant, whereby the action is dependent on the human’s belief in the particular underlying state. These human decisions are monitored by a Quickest Change Detector, which computes a belief in the underlying state to exploit knowledge of the human decision-model. Based on the computed belief, the Quickest Detector then decides to continue or declares that a change has occurred according to a policy, which maps a belief in the underlying state to this binary decision. If the detector declares ‘change’ before the change has occurred it incurs a false alarm penalty, and if it declares ‘no change’ when the change has occurred it incurs a delay penalty. The optimal policy minimizes the expected total penalty incurred. Note that this model is in contrast to the ‘classical’ quickest change detection problem, where the change detector directly observes the sensor output. Thus we incorporate intermediate human decisions, and investigate the structure of the optimal policy in this generalized framework.

A key assumption here is that the detector has a sufficiently rich model of the human decision-process. Existing structural results for this problem [8], [9], [10] assume some form of microeconomic utility maximization on the part of the human, but this assumed decision-structure is readily violated empirically [11], [13], [16] suggesting that these structural results may not generally apply to real-world environments. Motivated by this, we consider this framework using Quantum Decision Theory [12], in order to capture a wider range of human decision-making phenomena.

B. Why Quantum Decision Theory?

Generative models for human decision making are studied extensively in behavioral economics and psychology. The classical formalisms of human decision making are the Expected Utility models of Von-Neumann and Morgenstern (1953) [17] and Savage (1954) [18]. Despite the successes of these models, numerous experimental findings, most notably those of Kahneman and Tversky [11], have demonstrated violations of the proposed decision making axioms. There have since been subsequent efforts to develop axiomatic systems which encompass wider ranges of human behavior, such as the Prospect Theory [19].

Quantum Decision Theory ([20], [21], [22], [23] and references therein) has emerged as a new paradigm which is capable of generalizing current models and accounting for certain violations of axiomatic assumptions. For example, it has been empirically shown that humans routinely violate Savage’s ‘sure-thing principle’ [13], [24], which is equivalent to violation of the law of total probability, and that human decision making is affected by the order of presentation of information [14], [15] (“order effects”). These violations are natural motivators for treating the decision making agent’s mental state as a quantum state in Hilbert space; The mathematics of quantum probability was developed as an explanation of observed self-interfering and non-commutative behaviors of physical systems, directly analogous to the findings which Quantum Decision Theory (QDT) aims to treat. Indeed, the models of Quantum Decision Theory have been shown to reliably account for violations of the ‘sure-thing principle’ and order effects [20].

The model we utilize in this work, coming from [12], not only accounts for the above generalized decision phenomena, but comes equipped with a psychological parametrization representing a particular individual’s decision process and modeling such phenomena as bounded rationality. This parametrization will serve as a useful tool in that we can gain insight into the dependence of the optimal quickest detector structure on levels of e.g., rationality.
For the sake of brevity here, Appendix H discusses how this quantum decision model relates to other models for human decision-making, such as Prospect Theory and anticipatory decision making.

C. Main Results and Outline

We study the theoretical structural properties of the optimal policy of the human-decision based change detector outlined in Section I-A, utilizing quantum decision theory to model human decisions. In Section II we outline the quantum decision making process [12]. Section III outlines the quickest change detection protocol and the computation of the quickest detector’s optimal policy, with the incorporation of the human decision making model of [12]. Section IV presents our main theoretical results regarding the quickest detector’s performance and optimal policy resulting from our model. Specifically, we derive results for

1) Existence of a threshold optimal policy (Theorem 1): The quickest detector’s optimal policy exhibits a single-threshold. This is in contrast to the multi-threshold optimal policy present in the quickest change detection model of [8].

2) Intermediate human decisions hinder detection performance (Theorem 2): Under the optimal policy, the quickest detector performs strictly worse in expectation than in the classical quickest change detection protocol.

3) Sensitivity of detection performance to psychological parameters (Theorem 3): We provide an upper bound on the expected cost incurred by the quickest detector when only an estimate of the quantum decision maker’s psychological parameters is available. This has immediate practical relevance since the exact psychological processing of decision preferences for different individuals will differ and will seldom be known to an external observer.

4) Detection performance depends on agent rationality (Theorem 4, Theorem 5, computational results): We show that there exist disjoint convex regions in the psychological parameter space which induce performance ordering, i.e. the quickest detector performs strictly better when the decision maker has parameters in one region vs. the other. We provide a numerical simulation which validates this existence and suggests that the quickest detector performs better as the decision maker becomes more rational.

Along with these results, in Section V we provide a numerical example of the quickest detection scheme in the context of the Prisoner’s Dilemma problem. The ability of the quantum model to account for empirically observed violations of the sure-thing principle [25] is illustrated in this context.

II. QUANTUM MODEL FOR HUMAN DECISION MAKING

This section presents the quantum decision model that we will incorporate in the quickest detection scheme. Readers who are unfamiliar with quantum probability may refer solely to the abstracted decision protocol in Section II-A. This will be sufficient for subsequent sections and our results. The details of Sections II-B to A are not necessary for an understanding of the quickest detection procedure, but provide insight into the structure of the quantum decision formulation and the impact of the psychological parameters.

A. Decision-Making Protocol

For each discrete-time step \( n \in \mathbb{N} \) we have:

1) A state of nature \( x_n \in \{1, 2\} = \mathcal{X} \)

2) A state posterior \( \eta_n(x) \), \( x \in \mathcal{X} \), obtained by a Bayesian sensor from noisy observation \( y_n \sim \mathbb{P}(y|x_n) \) and prior \( \eta_{n-1}(x) \).

3) The human decision maker takes action \( a_n \sim \Gamma^{\eta_n} (\cdot) \) (1)

i.e., \( a_n \) is taken probabilistically from action distribution \( \Gamma^{\eta_n} (\cdot) \). This action distribution is derived in Section II-C, and is parametrized by the following:

- state posterior \( \eta_n(x) \)
- decision maker’s utility function \( u : \mathcal{A} \times \mathcal{X} \rightarrow \mathbb{R} \), mapping joint state-action pairs to values
- \((\alpha, \lambda, \phi)\): free parameters governing unique psychological characterization.

Thus, at an abstract level the action distribution encodes a human’s preference over actions given a particular belief \( \eta_n(x) \) in the underlying state, a reward function \( u(a, x) \) specifying the value of taking action \( a \) in state \( x \), and unique psychological parameters \((\alpha, \lambda, \phi)\) which govern e.g., rationality (see Section II-D). The above high-level representation is sufficient for the quickest detection framework presented in Sec. III. A detailed construction of the action distribution \( \Gamma^{\eta_n} (\cdot) \) is given in Section II-C and Appendix A. Next we give some background on quantum theory preliminaries and present the derivation of action distribution \( \Gamma^{\eta_n} (\cdot) \).

B. Background on Quantum Hilbert Space Structure [12]

Here we introduce the quantum theory concepts of Hilbert space, projectors and density matrices, which will be utilized in the construction of the psychological model of [12] in Sections II-C and A. Since we consider finite-dimensional systems, a Hilbert space \( \mathcal{H} \) is a linear space with an inner product \( \langle \psi_1 | \psi_2 \rangle \in \mathbb{C} \) for \( |\psi_1\rangle, |\psi_2\rangle \in \mathcal{H} \). If the state of a system is given by \( |\psi\rangle \in \mathcal{H} \) then we say it is in a pure state. We can form a projector \( P_\psi = |\psi\rangle \langle \psi | \) (outer product of \( |\psi\rangle \) with itself), which acts on \( \mathcal{H} \) as \( P_\psi |\phi\rangle = \langle \psi |\phi \rangle |\psi\rangle \) \( \forall |\phi\rangle \in \mathcal{H} \). Also note that the set of projectors \( \{P_\psi, |\psi\rangle \in \mathcal{H} \} \) has a bijective mapping to states \( \{|\psi\rangle \in \mathcal{H} \} \), so any pure state \( |\psi\rangle \) can be described in terms of its projector \( P_\psi \).

A density operator \( \rho : \mathcal{H} \rightarrow \mathcal{H} \) is such that

i) it is Hermitian: \( \rho^ \dagger = \rho \)
ii) it has trace one: \( \text{Tr}(\rho) = 1 \)
iii) it is positive semi-definite: \( \langle \psi | \rho |\psi \rangle \geq 0 \ \forall |\psi\rangle \in \mathcal{H} \)

Any density operator \( \rho \) can be expressed as \( \sum_j p_j |\psi_j\rangle \langle \psi_j | \), with \( \sum_j p_j = 1 \), \( |\psi_j\rangle \in \mathcal{H} \). Note that if \( p_1 = 1, p_j > 1 / j > 1 \), then \( \rho = P_{\psi_1} \) and \( \rho \) describes the pure state \( |\psi_1\rangle \). Otherwise, \( \rho \) encapsulates the idea that there is some uncertainty about the state of the system, with the probability of the state being \( |\psi_j\rangle \) given by \( p_j \); we call such a state a mixed state. Consider
two Hilbert spaces \( \mathcal{H}_1, \mathcal{H}_2 \). The tensor product \( \mathcal{H}_1 \otimes \mathcal{H}_2 \) is a Hilbert space together with a bilinear mapping \( \phi: \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2 \) such that:

i) For \( x \in \mathcal{H}_1, y \in \mathcal{H}_2 \), the set of all vectors \( \phi(x,y) \) forms a total subset of \( \mathcal{H}_1 \otimes \mathcal{H}_2 \), i.e., its closed linear span is equal to \( \mathcal{H}_1 \otimes \mathcal{H}_2 \).

ii) For \( x_1, x_2 \in \mathcal{H}_1, y_1, y_2 \in \mathcal{H}_2 \), we have \( \langle \phi(x_1, y_1), \phi(x_2, y_2) \rangle = \langle x_1 | x_2 \rangle \langle y_1 | y_2 \rangle \).

C. Modeling Psychological State via Quantum Probability

Suppose there are \( K \) underlying states in the state space \( \mathcal{X} \) (we have established that in our framework \( K = 2 \), see Section II-A), and \( A \) actions in the action space \( \mathcal{A} \). For each state \( i \in \{1, \ldots, K\} \) construct a corresponding unit complex vector \( \mathcal{E}_i \in \mathbb{C}^K \) such that \( \{\mathcal{E}_i\}_{i=1}^K \) are orthonormal. For each action \( i \in \{1, \ldots, A\} \), construct a complex vector \( a_i \in \mathbb{C}^A \) such that \( \{a_i\}_{i=1}^A \) are orthonormal. Denote \( \mathcal{H}_X = \text{span}\{\mathcal{E}_1, \ldots, \mathcal{E}_K\}, \mathcal{H}_A = \text{span}\{a_1, \ldots, a_A\} \), and form the tensor product Hilbert space \( \mathcal{H} = \mathcal{H}_X \otimes \mathcal{H}_A \). Let \( |\psi_j\rangle \in \mathcal{H} \) denote a column vector in this tensor product space, i.e., \( |\psi_j\rangle \in \mathbb{C}^{AK \times 1} \). Let \( |\psi_j\rangle \) denote a row vector, i.e., \( |\psi_j\rangle \in \mathbb{C}^{1 \times AK} \). The agent’s psychological state is represented by a density operator \( \rho_t \) which acts on the Hilbert space \( \rho_t: \mathcal{H} \to \mathcal{H} \). Specifically, \( \rho_t = \sum_j p_j |\psi_j\rangle \langle \psi_j| \) with \( \sum_j p_j = 1 \), \( |\psi_j\rangle \in \mathcal{H} \forall j \).

This construction is referred to as a mixed state in quantum mechanics. A mixed state is a generalization of a pure state to a probability distribution over pure states. A mixed state representation (density operator) is used in quantum decision theory for the sake of generality. The psychological state \( \rho_t \) evolves according to the Lindbladian operator \( \mathcal{L}_{(\alpha,\lambda,\phi)} \) by the ordinary differential equation \( \frac{d\rho_t}{dt} = \mathcal{L}_{(\alpha,\lambda,\phi)}\rho_t \) where \((\alpha,\lambda,\phi)\) are free parameters which govern the evolution, each having a psychological interpretation, see [12]. Implicit in \( \mathcal{L}_{(\alpha,\lambda,\phi)} \) is a belief \( \eta(x) \) in the underlying state \( x \in \{1, \ldots, K\} \) and a utility function \( u: \mathcal{A} \times \mathcal{X} \to \mathbb{R} \). The psychological state \( \rho_t \) encodes a time dependent probability distribution \( \Gamma(a,t) \) over actions \( a \in \mathcal{A} \) in the following way. Let \( P_i := \sum_{j=1}^K |a_i \otimes \mathcal{E}_j\rangle \langle a_i \otimes \mathcal{E}_j| \) be the projector onto \( a_i \otimes \mathcal{X} \), i.e., the subspace spanned by action vector \( a_i \), then \( \Gamma(a_i,t) = \text{Tr}(P_i \rho_t P_i^\dagger) \), where \( P_i^\dagger \) is the adjoint of \( P_i \).

The psychological state at time \( t \), \( \rho_t \), evolves according to (2) and induces a distribution \( \Gamma^0(a_i,t) \) over the action space as \( \Gamma^0(a_i,t) = \text{Tr}(P_i \rho_t P_i^\dagger) \) \( \Gamma^0(a_i,t) \) is the steady-state distribution where \( \alpha,\lambda,\phi \) are free parameters which govern the evolution, each having a psychological interpretation, see [12]. Implicit in \( \Gamma^0(a_i,t) \) is a belief \( \eta(x) \) in the underlying state \( x \in \{1, \ldots, K\} \) and a utility function \( u: \mathcal{A} \times \mathcal{X} \to \mathbb{R} \). The psychological state \( \rho_t \) encodes a time dependent probability distribution \( \Gamma(a,t) \) over actions \( a \in \mathcal{A} \) in the following way. Let \( P_i := \sum_{j=1}^K |a_i \otimes \mathcal{E}_j\rangle \langle a_i \otimes \mathcal{E}_j| \) be the projector onto \( a_i \otimes \mathcal{X} \), i.e., the subspace spanned by action vector \( a_i \), then \( \Gamma(a_i,t) = \text{Tr}(P_i \rho_t P_i^\dagger) \), where \( P_i^\dagger \) is the adjoint of \( P_i \).

The psychological state at time \( t \), \( \rho_t \), evolves according to (2) and induces a distribution \( \Gamma^0(a_i,t) \) over the action space as

\[ \Gamma^0(a_i,t) = \lim_{t \to \infty} \Gamma^0(a_i,t) \]

We assume action \( a_i \) is taken probabilistically according to the steady-state distribution \( \Gamma^0(a) \) which is independent from the initial state \( \rho_0 \). This represents the action choice occurring after deliberation has ended, and the steady state is typically reached relatively quickly\(^{2}\). We can then abstract away from the time dependence to get the map \( \mathcal{L}_{(\alpha,\lambda,\phi)}: (\eta(x), u(x,a)) \to \Gamma^0(a) \)

This map induces the action distribution (1) in the decision-making protocol of Section II-A. The structure of \( \mathcal{L}_{(\alpha,\lambda,\phi)} \) incorporates the agent’s objective utility function \( u(x,a) \), state belief \( \eta(x) \), and parametrization \((\alpha, \lambda, \phi)\) to produce a distribution over actions. For full details on the construction of \( \mathcal{L}_{(\alpha,\lambda,\phi)} \), we refer to Appendix A and [12].

D. Practicality in Human Decision Making

1) Violation of the Sure-Thing Principle: The above quantum model for human decision making accounts for violations of the sure-thing principle (STP), which we now describe. Suppose there exists an action \( a \) and two states \( \mathcal{E}_1, \mathcal{E}_2 \). Suppose \( \Gamma \) is a non-degenerate posterior belief (strictly in the interior of the unit simplex) of the state. The sure-thing principle [18] states that \( P(a|\Gamma) \) will be a convex combination of \( P(a|\mathcal{E}_1) \) and \( P(a|\mathcal{E}_2) \), i.e., \( \exists \epsilon \in (0,1) \) such that

\[ P(a|\Gamma) = \epsilon P(a|\mathcal{E}_1) + (1-\epsilon) P(a|\mathcal{E}_2) \]

Violations of this principle (when there does not exist any \( \epsilon \in (0,1) \) such that (5) holds) are readily observed in empirical human decision-making studies. Pothos and Busemeyer [16] (see also [24]) review empirical evidence for the violation of STP and show how quantum models can account for it by introducing quantum interference in the probability evolution. Note that this violation cannot be accounted for by traditional models which rely on classical probability, as the sure-thing principle follows directly as a consequence of the law of total probability.

2) Time Evolution: QDT is able to model time evolving preference distributions, and [27] gives empirical evidence for the applicability of quantum decision models over traditional dynamic Markovian models [28], [29] investigates empirically observed decision preference evolution phenomena such as preference oscillation and the choice-induced preference change. They conclude that the quantum decision model of [12] provides “the best available single system account of these three characteristics—evolution, oscillation, and choice-induced preference change”.

3) Psychological Parametrization: The parameters \((\alpha, \lambda, \phi)\) also allow for practical psychological interpretation. The parameter \( \alpha \) interpolates between the purely quantum preference evolution and the dissipative Markovian evolution in [29], and thus a higher \( \alpha \) corresponds to increased rationality, in the sense of choosing actions which accord with classical expected utility maximization. \( \lambda \) is a measure for bounded rationality.
as (from (31)) it is a monotonic measure of the ability to 
discriminate between the profitability of different options. The 
interpretation of $\phi$ (in (30)) is more nuanced, but can be thought 
of as the relevance of the formation of a belief in the underlying 
state to the decision making process. See [12], [29] for detailed 
discussion on these interpretations.

III. QUICKEST CHANGE DETECTION WITH QUANTUM 
DECISION MAKER

We now introduce the quickest change detection protocol 
and the formulation of an optimal policy for such a protocol. The 
quickest change detection problem consists of a state of 
nature $x_n \in \{0, 1\}$, indexed by discrete time $n \in \mathbb{N}$, which jump 
changes at some unknown time $\tau^0$. The aim of quickest detection 
is to determine the jump time $\tau^0$ of the state of nature $\{x_n\}$. 
This is accomplished by evaluating a policy $\mu : \Pi \to \{1, 2\}$, 
which maps from a belief $\pi$ (in the unit one-simplex $\Pi$) in the underlying state $x_n$, to a binary decision (declare ‘change’ 
or ‘no change’). The 

optimal policy $\mu^*(\pi)$ is computed as that 

which minimizes the following Kolmogorov-Shiryaev criterion 
(encoding the expected total cost incurred) [1]: 

$$J_{\mu^*}(\pi) = \inf_{\mu} J_{\mu}(\pi),$$ 

$$J_{\mu}(\pi) = dE_{\mu}[(\pi - \tau^0)] + fP_{\mu}(\pi < \tau^0)$$ (6) 

where $\tau = \inf\{n : u_n = 1\}$ is the time at which the global 
decision maker announces the change, $E_{\mu}[(\pi - \tau^0)]$ and 
$P_{\mu}(\pi < \tau^0)$ are the expected detection delay and probability 
of false-alarm under policy $\mu$, respectively. 
The parameters $d > 0$ and $f > 0$ specify the delay penalty and false alarm 
penalty, respectively.

The optimal policy $\mu^*(\pi)$ (6) can be formulated as the solution 
of a stochastic dynamic programming equation. The quickest detection problem (6) is an example of a stopping-time 
partially observable Markov decision process (POMDP) with a stationary optimal policy.

We now introduce some notation, then describe the protocol 
in detail.

i) The state of nature $\{x_n \in \{1, 2\}, n \geq 0\}$ models the change event which we aim to detect. $x_n$ starts in state 

2 and jumps to state 1 at a geometrically distributed random time $\tau^0$ with $E[\tau^0] = \frac{1}{1-p}$ for some $p \in [0, 1)$. So, 

$\{x_n\}$ is a 2-state Markov chain with absorbing transition matrix and initial probability 

$$P = \begin{bmatrix} 1 & 0 \\ 1-p & p \end{bmatrix}, \quad \pi_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$ (7) 

with change time $\tau^0 = \inf\{n : x_n = 1\}$.

ii) The quantum decision agents act sequentially. A sensor 

observe the state of nature $x_n$ in noise and computes a 
Bayesian posterior distribution $\eta(x)$ of the underlying state. This is given to the human, who then makes a decision (action) $a_n$ according to the steady-state action 
distribution $\Gamma^n(a_i)$ induced by the Lindbladian operator $L_{(\alpha, \lambda, \phi)}$ and map (4).

iii) Based on the history of local actions $a_1, \ldots, a_n$, the 
global decision maker chooses action 

$$u_n = \{1 (stop and announce change), 2 (continue)\}$$

iv) Define the detector’s belief $\pi_n$ and private belief $\eta_n$ at time $n$ as the posterior distributions initialized with $\eta_0 = \pi_0 = [0, 1]'$:

$$\pi_n(x) = \mathbb{P}(x_n = x | a_1, \ldots, a_n), \quad x = 1, 2$$

$$\eta_n(x) = \mathbb{P}(x_n = x | a_1, \ldots, a_{n-1}, y_n),$$ (8) 

where $y_n$ is the private observation recorded by agent $n$. 
We have $\pi_n(x), \eta_n(x) \in \Pi$, the unit one-simplex.

A. Change Detection Protocol [8]

We now detail the multi-agent quickest change detection protocol:

1) Human Decision Maker $n$

i) Sensor obtains detector’s belief $\pi_{n-1}$ and signal $u_{n-1}$ from global decision maker. The process only continues if $u_{n-1} = 2$.

ii) Let $\mathcal{Y}$ denote the observation space. The sensor records noisy observation $y_n \in \mathcal{Y}$ of state of nature $x_n$ with conditional density 

$$B_{x,y} = p(y_n = y | x_n = x)$$ (9) 

iii) Private Belief. The sensor evaluates the Bayesian 

private belief $\eta_n = [\mathbb{P}(x_n = 1), \mathbb{P}(x_n = 2)]'$ as 

$$\eta_n = T(\pi_{n-1}, y_n), \quad T(\pi, y) = \frac{B_{y,P}^\pi}{\sigma(\pi, y)},$$ (10) 

$$\sigma(\pi, y) = I \cdot B_y P^\pi, \quad B_y = \text{diag}(B_{1,y} B_{2,y})$$ (11) 

and feeds this to the human agent.

iv) Human Decision. The agent’s private belief $\eta_n$ parameterizes the Lindbladian operator $L_{(\alpha, \lambda, \phi)}$. This 

induces a steady-state action probability distribution 
$\Gamma^n(a_i)$ via the map (4), and the action $a_n$ is taken 
probabilistically according to $\Gamma^n$.

2) Quickest Detector. Based on the decisions $a_n$ of local 
decision maker $n$, the quickest detector: 

i) Updates the detector belief $\pi_n = [\mathbb{P}(x_n = 1), \mathbb{P}(x_n = 2)]'$ from $\pi_{n-1}$ to $\pi_n$ as 

$$\pi_n = T(\pi_{n-1}, a_n)$$

$$\tilde{T}(\pi, a) = \frac{R_{\pi}(a) P^\pi}{\sigma(\pi, a)}, \quad \sigma(\pi, a) = I' R_{\pi}(a) P^\pi$$

$$R_{\pi}(a) = \text{diag}(R_{1,\pi}(a), R_{2,\pi}(a))$$

$$R_{x,\pi}(a_n) = \mathbb{P}(a_n = a | x_n = x, \pi_{n-1})$$ (12) 

The action probabilities $R_{x,\pi}(a)$ are computed as 

$$R_{x,\pi}(a) = \int \tilde{T}^n_{y_n}(a) B_{x,y} dy$$ (13)
where $\Gamma_{y}^{n-1}$ is the human’s induced action distribution (4) given detector’s belief $\pi_{n-1}$ and observation $y$. Specifically, $\Gamma_{y}^{n-1}$ is the output of the map (4), with input $\eta(x) = T(\pi_{n-1}, y), u(x, \alpha)$ and estimated parametrization $(\alpha, \lambda, \phi)$. Observe that here the quickest detector has an estimate of the psychological parametrization $(\alpha, \lambda, \phi)$; later we will investigate the performance sensitivity to this estimate.

ii) Chooses global action $u_n$ using optimal policy $\mu^*$:

$$u_n = \mu^*(\pi_n) \in \{1{\text{stop}}, 2{\text{continue}}\}. \quad (14)$$

iii) If $u_n = 2$, then set $n$ to $n + 1$ and go to step 1. If $u_n = 1$, then stop and announce change.

We assume the global decision maker knows $P$ (7) and the agent’s action $a_n$, and has an estimate $(\alpha, \lambda, \phi)$ of the agent’s psychological parametrization. The global decision maker does not know the observation $y_n$, or the private belief $\eta_n$. We assume all agents have the same psychological parametrization $(\alpha, \lambda, \phi)$ (such as if the same agent acts sequentially), otherwise the optimal detection strategy is non-stationary. The update (12) is where the quantum decision theory enters our quickest detection formulation. In simple terms, the action of the human is a probabilistic function of the noisy measurement of the sensor. So the likelihood of the action given the state enters our computation for the belief state $\pi_n$ in quickest detection.

B. Quickest Detector Optimal Policy [8]

Considering the aim of quickest detection, characterized by (6), we now outline the details of the optimal policy stochastic dynamic programming formulation.

1) Costs: To present the dynamic programming equation we first formulate the false alarm and delay costs (6) incurred by the global decision maker in terms of the detector’s belief $\pi_n$. Recall the objective of the detector’s optimal policy $\pi^*$ is to declare “stop” at a time $\tau$, from observations of sequential actions $\{a_n\}$, such that $\tau$ is close to the true change-point $\tau^0$. The formulation is quantified by having the policy $\pi^*$ minimize a sum of false-alarm and delay costs. The false-alarm cost penalizes the policy for declaring “stop” before the change has occurred, i.e., $\tau < \tau^0$, and the delay cost penalizes the policy for declaring “stop” after the change has occurred, i.e., $\tau > \tau^0$. These penalties, respectively $f$ and $d$, are parameters which govern the respective importance of each type of detection error. Thus, the optimal policy may differ for differing $f, d$, but our results hold for any such choice of $f > 0, d > 0$.

i) False Alarm Penalty: If global decision $u_n = 1$ (stop) is chosen before the change point $\tau^0$, then a false alarm penalty is incurred. The false alarm event $\{x_n = 2, u_n = 1\}$ represents the event that a change is announced before the change happens at time $\tau^0$. Recall (7) the jump change occurs at time $\tau^0$ from state 2 to state 1. Then recalling $f > 0$ is the false alarm penalty in (6), the expected false alarm penalty is

$$f\mathbb{P}_\mu(\tau < \tau^0) = f\mathbb{E}_\mu\{\mathbb{E}[I(x_n = 2, u_n = 1)|G_n]\} \quad (15)$$

where $G_n$ is the $\sigma$-algebra generated by $(a_1, \ldots, a_n)$ and $I$ is the indicator function. Clearly $\mathbb{E}[I(x_n = 2, u_n = 1)|G_n]$ can be expressed in terms of the detector’s belief $\pi(2)$ as

$$C(\pi_n, u_n = 1) = f\mathbb{E}_\pi\{n\}, \quad \text{where} \quad e_2 = [0, 1]^T \quad (16)$$

ii) Delay Cost of Continuing: If global decision $u_n = 2$ is taken then Protocol 1 continues to the next time. A delay cost is incurred when the event $\{x_n = 1, u_n = 2\}$ occurs, i.e. no change is declared at time $n$. The expected delay cost is

$$d\mathbb{E}[\mathbb{E}[I(X_n = 1, u_n = 2)|G_n]|d > 0 \text{denotes the delay cost.}$$

In terms of the detector’s belief, the delay cost is

$$C(\pi_n, u_n = 2) = d\mathbb{E}_\pi\{n\}, \quad \text{where} \quad e_1 = [1, 0]^T \quad (17)$$

Recall that we aim to minimize the total expected cost incurred. This is formalized as minimizing the Kolmogorov-Shiryaev criterion (6), which can be rewritten as

$$J_\mu = \mathbb{E}_\mu\left\{\sum_{n=1}^{\tau-1} C(\pi_n, 2) + C(\pi_{\tau}, 1)\right\} \quad (18)$$

where $\tau = \inf\{n : u_n = 1\}$ is adapted to the $\sigma$-algebra $G_n$. Since $C(\pi, 1), C(\pi, 2)$ are non-negative and bounded for all $\pi \in \Pi$, stopping is guaranteed in finite time.

2) Bellman’s Equation for Quickest Detection Policy: Consider the costs (16), (17) defined in terms of the detector’s belief $\pi$. Then the optimal stationary policy $\mu^*(\pi)$ defined in (6) and associated value function $V(\pi)$ are the solution of Bellman’s dynamic programming functional equation

$$Q(\pi, 1) := C(\pi, 1)$$

$$Q(\pi, 2) := C(\pi, 2) + \sum_{a \in A_d} V(\tilde{T}(\pi, a))\tilde{\sigma}(\pi, a)$$

$$\mu^*(\pi) = \arg\min\{Q(\pi, 1), Q(\pi, 2)\},$$

$$V(\pi) = \min\{Q(\pi, 1), Q(\pi, 2)\} = J_\mu^*(\pi) \quad (19)$$

The detector’s belief update $\tilde{T}$ and normalization measure $\tilde{\sigma}$ were defined in (12). The goal of the detector is to solve for the optimal quickest change policy $\mu^*$ in (19) or, equivalently, determine the optimal stopping set $S$

$$S = \{\pi : \mu^*(\pi) = 1\} = \{\pi : Q(\pi, 1) \leq Q(\pi, 2)\} \quad (20)$$

3) Value Iteration Algorithm: The optimal policy $\mu^*(\pi)$ and value function $V(\pi)$ can be constructed as the solution of a fixed point iteration of Bellman’s equation (19). The resulting algorithm is called the value iteration algorithm.
The value iteration algorithm proceeds as follows: Initialize $V_0(\pi) = 0$ and for iterations $k = 1, 2, \ldots$

\[
V_{k+1}(\pi) = \min_{u \in U} Q_{k+1}(\pi, u), \\
\mu^*_{k+1}(\pi) = \arg\min_{u \in U} Q_{k+1}(\pi, u), \quad \pi \in \Pi, \\
Q_{k+1}(\pi, 1) = C(\pi, 1), \\
Q_{k+1}(\pi, 2) = C(\pi, 2) + \sum_{a \in A_1 \times A_2} V_k(T(\pi, a))\tilde{\sigma}(\pi, a)
\]

(21)

Let $B$ denote the set of bounded real-valued functions on $\Pi$. For any $V, V \in B$ and $\pi \in \Pi$, define the sup-norm metric $\sup V(\pi) - \tilde{V}(\pi)$. Since $C(\pi, 1), C(\pi, 2), \pi \in \Pi$ are bounded, the value iteration algorithm (21) generates a sequence of lower semi-continuous value functions $\{V_k\} \subset B$ that converges pointwise as $k \to \infty$ to $V(\pi) \in B$, the solution of Bellman’s equation.

IV. CHARACTERIZING THE STRUCTURE OF THE QUICKEST DETECTOR

In this section we analyze several structural properties of the quickest detection protocol detailed in Sec. III. Our results in this section are structured as follows: In Section IV-A we prove that the optimal policy (19) has a single threshold structure. In Section IV-B we provide a lower bound on the optimal cost incurred by the quickest detector via the policy of Sec. III. Specifically, this lower bound is given by the optimal cost incurred within the classical quickest change detection protocol, i.e. without intermediate human decisions. The key idea here is to use Blackwell dominance [30] between matrices characterizing the quickest detector observations and the noisy sensor observations. We say a matrix $M_1$ Blackwell dominates matrix $M_2$ (and that $M_2$ is Blackwell dominated) if there exists column stochastic matrix $B$ such that $M_2 = M_1 B$. In Section IV-C we consider the performance sensitivity to the quickest detector’s estimate of the psychological parameterization, and prove an upper bound on the cumulative cost incurred in terms of the cumulative cost incurred given perfect knowledge of the parameterization and a KL Divergence term.

A. Existence of a Threshold Optimal Policy

We will show that, given the quantum decision making quickest change detection protocol detailed in Section III-A, the quickest detector’s optimal policy (6) exhibits a single threshold behavior.

**Theorem 1:** Given the quantum decision making quickest change detection protocol detailed in Section III-A, the quickest detector’s optimal policy $\mu^*$ (6) exhibits a single threshold state $\pi'$ such that

$$\mu^*(\pi) = \begin{cases} 2, & \pi < \pi' \\ 1, & \pi \geq \pi' \end{cases}$$

**Proof:** See Appendix B.

In Section V we numerically implement the value iteration algorithm (21) in the context of a ‘Prisoner’s Dilemma’ quickest detection scheme. In particular, Fig. 5 demonstrates the single threshold behavior of the optimal policy $\mu^*(\pi)$ (19).

This optimal policy structure is in contrast to the multi-threshold policy obtained in [8], in which an anticipatory model was used for the human decision makers. Within a multi-threshold (non-convex stopping region) policy, there exist points where the optimal behavior is to transition from declaring ‘change’ to declaring ‘no change’ as the probability of change increases. This is not only counterintuitive, but makes the design of human-sensor quickest detectors more complex. Thus the single threshold policy exhibited in our case is desirable for intuitive and practical design purposes See Appendix H for comparisons to the anticipatory model of [8].

B. Lower Bound on Performance

We now show that the optimal cost incurred by quickest change detection with quantum agents is greater than that incurred by the classical Bayesian framework. We note that this result is not due to the ‘quantum’ behavior, but holds for intermediate human decisions in general. Nevertheless, this is useful since performance analysis of standard quickest detection [31] applies as a lower bound for quickest detection with quantum agents. Consider the optimal policy and cost of the classical Bayesian quickest change detection. [31]. Similar to (21), the optimal policy $\mu^*(\pi)$ and cost $V(\pi)$ incurred by the classical quickest detection, satisfy the stochastic dynamic programming equation:

$$\mu^*(\pi) = \arg\min_{u \in U} Q(\pi, u), \quad V(\pi) = \min_{\pi \in \Pi} Q(\pi, u),$$

where

$$Q(\pi, 2) = C(\pi, 2) + \sum_{y \in Y} I(\pi, y)\sigma(\pi, y),$$

$$I(\pi, y) = C(\pi, 1), \quad J_{\mu^*}(\pi) = V(\pi) \tag{22}$$

Here $I(\pi, y)$ is the Bayesian filter update defined in (10) and $J_{\mu^*}(\pi)$ is the cumulative cost of the optimal policy starting with initial belief $\pi$. Note that in classical quickest detection, there is no detector’s belief update (12) or interaction between public and private beliefs.

**Theorem 2:** Consider the quantum decision making quickest change detection protocol in Section III-A and the associated value function $V(\pi)$ in (21). Consider also the classical quickest change detection problem with value function $V(\pi)$ in (22). Then for any detector belief $\pi \in \Pi$, the optimal cost incurred by the classical quickest detection is smaller than that of quickest detection with quantum decision agents. That is,

$$V(\pi) \leq V(\pi) \forall \pi \in \Pi.$$

**Proof:** See Appendix C.

Informally, this result can be interpreted by the observation that the intermediate human decision-making process results in loss of information pertaining to the underlying state. Indeed, we use Blackwell Dominance arguments within the proof, which formalize this notion of cascaded information loss. The practical interpretation is that regardless of the human psychological parameterization, i.e. perfectly rational etc., the hierarchical detection structure in which there is an intermediate human decision making process results in decreased
C. Sensitivity of Detection Performance to Psychological Parameters

Recall that the quickest detector uses an estimate of the psychological parameters ($\alpha$, $\lambda$, $\phi$). Thus, we would like to characterize how the quickest detection performance depends on such an estimate. In this section, we quantify this question and provide a bound on the deviation of the performance from that incurred by perfect knowledge of the psychological parameters.

First we begin by defining some notation. Recall the domain of the psychological parameters ($\alpha$, $\lambda$, $\phi$) $D := [0, 1] \times [0, \infty) \times [0, 1] \subset \mathbb{R}^3$. Define $\Lambda$ to be a probability density function in $D$, $\Lambda : D \mapsto \mathbb{R}_+$, $\int_D \Lambda(\gamma) d\gamma = 1$, representing the quickest detector’s probabilistic estimate of the local decision maker’s psychological parameters. Denote the actual psychological parameterization of the local decision maker by $\gamma$ in $D$.

We now reconsider the decision making protocol from Section III-A when the quickest detector only has this probabilistic estimate of the local decision maker’s psychological parameters. The quantum decision maker step remains the same, except let us now denote the steady-state action distribution by $\tilde{\Gamma}(\alpha_i)$ to denote that this is a result of the true parametrization $\alpha_i$ rather than an estimate. In this section, we quantify how this generalized protocol presents a bound on the cumulative cost incurred when the quickest decision maker uses $\tilde{\alpha}_i$ in place of the true $\alpha_i$.

The action probabilities $\tilde{R}_{x,\pi}(a)$ are now computed as

$$\tilde{R}_{x,\pi}(a) = \frac{\tilde{\Gamma}(a)}{\tilde{\sigma}(a, x)} = \tilde{\Gamma}(a) \frac{\tilde{\sigma}(a, x)}{\tilde{\sigma}(a)},$$

where $\tilde{\sigma}(a, x)$ is the QDM’s induced action distribution (4) given detector’s belief $\pi_{n-1}$, observation $y$, and psychological parameterization $\gamma$ in $D$. The quickest detector then chooses action $u_n$ according to (14), where the optimal policy is computed using the value iteration algorithm (21) with this new function $\tilde{T}(\pi, a)$.

We are interested in characterizing how this generalized procedure effects the quickest change performance. We now define some notation which will allow us to reason about this. Notice that quickest change decision making protocol in Section III-A is completely characterized as a two-state partially observed Markov Decision Process (POMDP) with underlying state transition matrix $P$ and observation likelihood $R_{\pi}(a)$. Similarly, the generalized protocol presented immediately above is characterized as a POMDP with transition matrix $P$ and observation likelihood $\tilde{R}_{\pi}(a)$. Notice that in our case the observation likelihoods are functions of $\pi$. We can then denote these POMDPs as $\theta = (P, R_{\pi})$ and $\tilde{\theta} = (P, \tilde{R}_{\pi})$, and their resultant optimal policies $\mu^*(\theta)$ and $\mu^*(\tilde{\theta})$, respectively.

Let $J_{\mu^*(\theta)}(\pi; \theta)$ and $J_{\mu^*(\tilde{\theta})}(\pi; \tilde{\theta})$ denote the discounted cumulative costs incurred by these POMDPs when using policy $\mu^*(\theta)$. Similarly, $J_{\mu^*(\theta)}(\pi; \theta)$ and $J_{\mu^*(\tilde{\theta})}(\pi; \tilde{\theta})$ denote the discounted cumulative costs incurred by these POMDPs when using policy $\mu^*(\tilde{\theta})$. These POMDP formulations have cost $C(x_n, u) = f_1(x_{n+1}) + d_1(x_n-1, a_{n+1})$ and an implicit discount factor $\gamma = 1-p$ (see [32] for details).

Now we can formulate a bound on the cumulative cost incurred when the quickest detector uses $\tilde{\theta}$ in place of the true $\theta$.

**Theorem 3:** Consider the quickest change detection protocols in which the quickest detector uses $\tilde{T}(\pi, a)$ and $\tilde{T}(\pi, a)$ for its belief update. Denoting the corresponding POMDP characterizations by $\theta = (P, R_{\pi})$ and $\tilde{\theta} = (P, \tilde{R}_{\pi})$, respectively, and using the notation defined above, we have the inequality

$$J_{\mu^*(\theta)}(\pi; \theta) \leq J_{\mu^*(\theta)}(\pi; \theta) + 2K\|\theta - \tilde{\theta}\|$$

where $D(R_{\pi,\gamma}(\pi, \gamma)) := \sum \log(R_{\pi,\gamma}(\pi, \gamma)/\tilde{R}_{\pi,\gamma}(\pi))$ denotes the Kullback-Leibler (KL) divergence.

**Proof:** The proof slightly adapts that of Theorem 14.9.1 of [33] and can be found in Appendix D.

Observe that by Corollary 1 (presented after this proof in Appendix B), the KL divergence term $D(R_{\pi,\gamma}(\pi, \gamma))$ is continuous with respect to parameters $(\alpha, \lambda, \phi)$. Thus, the detection performance (given by cumulative cost $J_{\mu^*(\theta)}(\pi; \theta)$) of a quickest detector exploiting an estimate $(\tilde{\alpha}, \tilde{\lambda}, \tilde{\phi})$ of the human psychological parameters $(\alpha, \lambda, \phi)$ is bounded above by a continuous function of the inaccuracy (quantified by an appropriate error norm in parameter space) of the parameter estimate. Informally, a change of $\epsilon$ in the parameter estimate will result in change in detection performance of $O(\epsilon)$. In this sense, the detection performance is robust to inaccuracies of the estimated human psychological parameterization.

D. Blackwell Dominance Properties: Rationality Improves Detection Performance

Here we present two theorems which will be used with our numerical study to reveal the existence of disjoint convex regions of the psychological parameter space which induce detection performance ordering. Theorem 4 states that if one steady-state action distribution Blackwell dominates another, then the value function induced by the former is upper bounded by that of the latter. This allows us to reason about the detection performance (characterized by the value function) by investigating the property of Blackwell dominance between steady-state distributions. Theorem 5 allows us to interpolate this performance ordering for all convex combinations of steady-state distributions which have this Blackwell dominance property. For ease of explanation, we say matrix $M_1$ is Blackwell dominating with respect to matrix $M_2$ (and that $M_2$ is Blackwell dominated) if there exists column stochastic (columns sum to 1) matrix $B$ such that $M_2 = M_1B$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Fig. 2. Sequential Quickest Change Detection with quantum agents. The underlying state of nature $x_n$ jumps changes at time $\tau^0 \sim \text{Geo}(1-p)$, where $p$ is known. At each time $n$ a sensor observes $y_n \sim P(y|x)$, and the detector’s belief signal $\pi_n$ from the previous time point. The sensor outputs a private belief $\eta_n$ (obtained via Bayesian update) in the underlying state to the quantum decision maker. The decision maker’s parameterized psychological Lindbladian $L_{(\alpha, \lambda, \phi)}(\eta_n)$ evolves to steady state $\Gamma$ (4) and an action $a_n$ is taken probabilistically from $\Gamma$. The quickest detector sees $a_n$ and outputs its detector’s belief $\pi_n$ and signal $u_n$ according to (12) and (14).

**Theorem 4:** Let $\Gamma^\pi$ and $\hat{\Gamma}^\pi$ be two steady-state action distributions, resulting from map (4) with prior $\pi$ and different Lindbladian parametrizations $(\alpha, \lambda, \phi)$, $(\hat{\alpha}, \hat{\lambda}, \hat{\phi})$, respectively. Suppose there exists a column stochastic matrix $M^\pi$ such that $\Gamma^\pi = \hat{\Gamma}^\pi M^\pi$. Then, incorporating these distributions in the update (13), the value iteration algorithm (21) yields $V(\pi) \geq \hat{V}(\pi)$, where $V$ and $\hat{V}$ are the value functions resulting from the use of distributions $\Gamma^\pi$ and $\hat{\Gamma}^\pi$, respectively.

**Proof:** See Appendix E.

Somewhat more informally, this Theorem states that if steady-state distribution $\Gamma^\pi$ Blackwell dominates another steady state distribution $\Gamma^\pi$ for all $\pi \in \Pi$, then the quickest detector’s performance (cumulative cost incurred) corresponding to the former distribution is better than that corresponding to the latter distribution.

**Theorem 5:** Suppose there exist probability mass vectors $\Gamma_1, \Gamma_2 \in \mathbb{R}^N (N \in \mathbb{N})$ and stochastic matrices $M_1, M_2 \in \mathbb{R}^{N \times N}$ such that $\Gamma_1 = \Gamma M_1$ and $\Gamma_2 = \Gamma M_2$. Form $\Gamma_3 = \mathbb{R}^N$ as $\Gamma_3(a) = \gamma_3 \Gamma_1(a) + (1 - \gamma_3) \Gamma_2(a)$, $\gamma_3 \in [0, 1]$. For any $\alpha \in \{1, \ldots, N\}$. Then there exists a stochastic matrix $M_3$ such that $\Gamma_3 = \Gamma M_3$.

**Proof:** See Appendix F.

This Theorem states that if two steady-state distributions $\Gamma_1$ and $\Gamma_2$ are Blackwell dominated by a third steady-state distribution $\Gamma$, then any distribution $\Gamma_3$ which is a convex combination of $\Gamma_1$ and $\Gamma_2$ will also be Blackwell dominated by $\Gamma$.

Here we demonstrate a numerical consequence of the preceding Theorems and the numerical verifications of Appendix G. There exist disjoint convex regions $R_1$ and $R_2$ in the parameter space $(\alpha, \lambda, \phi)$, such that the value functions $V_{p_1}$ and $V_{p_2}$, resulting from Lindbladian parameterizations (4) $p_2 = \{\alpha_1, \lambda_1, \phi_1\} \in R_1$, $p_2 = \{\alpha_2, \lambda_2, \phi_2\} \in R_2$ satisfy $V_{p_1}(\pi) \geq V_{p_2}(\pi)$ for all $\pi \in \Pi$, $\forall p_1 \in R_1, p_2 \in R_2$.

Fig. 3 demonstrates two such regions which have been verified numerically (see Appendix G), the green region corresponding to $R_1$ and the blue region corresponding to $R_2$. In words, this result means that for the quickest change detection system of Fig. 2, given local human decision makers acting with a decision making process characterized by the Lindbladian evolution with parameter $p_1 \in R_1$ and another acting with process parameterized by $p_2 \in R_2$, the quickest detector’s optimal cost in the former case is upper bounded by that of the latter case. Intuitively, this means that the quickest change detection system performs strictly better when the human decision maker has psychological parameters in the green region compared to the blue region.

**V. NUMERICAL EXAMPLE**

Here we provide a tutorial numerical example using the Prisoner’s Dilemma problem [12]. We demonstrate the ability of quantum decision theory to account for violations of the sure-thing principle, and provide an implementation of the quickest detector for a psychological parameterization which results in this violation. The key takeaways are that the Lindbladian model (29) can account for violations of the sure-thing principle (which cannot be accounted for by classical models), and that the quickest detector implementing this model still performs reasonably well while this violation is occurring.

**1) Construction of Lindbladian Operator:** Here we illustrate the construction of the Lindbladian operator (29) for the...
Prisoner’s Dilemma example. Suppose that the two underlying states of nature are whether or not the opponent defects (D) or cooperates (C), i.e. \( \mathcal{X} = \{1\text{(cooperate)}, 2\text{(defect)}\} \). The actions of the agent are also to either cooperate or defect, i.e. \( \mathcal{A} = \mathcal{X} \), and this action will depend on the agent’s belief in the underlying state (the opponent’s choice) and the payoff matrix. In this case we have the payoffs \( a = u(C|C), b = u(C|D), c = u(D|D), d = u(D|C) \). We have a four-dimensional space of states \( \mathcal{H} = \mathcal{H}_X \otimes \mathcal{H}_A = \{CC, DC, CD, DD\} \) since two actions (cooperate or defect) are each associated to two states of nature (opponent cooperates or defects). To construct the Lindbladian operator (29), we need to construct the Hamiltonian \( H \) and the Cognitive Matrix \( C(\lambda, \phi) \) (30). Following (32), we build the matrix \( \Pi(\lambda) \) as

\[
\Pi(\lambda) = \begin{bmatrix}
1 - \mu(\lambda) & \mu(\lambda) & 0 & 0 \\
1 - \mu(\lambda) & \mu(\lambda) & 0 & 0 \\
0 & 0 & 1 - \nu(\lambda) & \nu(\lambda) \\
0 & 0 & 1 - \nu(\lambda) & \nu(\lambda)
\end{bmatrix}
\]

(26)

where \( \mu(\lambda) = \frac{\lambda}{\alpha + \lambda^2} \) and \( \nu(\lambda) = \frac{\lambda}{\beta + \lambda^2} \). Suppose the agent has belief in the opponents action (underlying state) given by \( \eta(x) \), such that \( \eta(1) = P(\text{opponent cooperates}) \) and \( \eta(2) = P(\text{opponent defects}) \). Then, following (33) we have

\[
B = \begin{bmatrix}
\eta(1) & 0 & \eta(2) & 0 \\
0 & \eta(1) & 0 & \eta(2) \\
0 & \eta(1) & 0 & \eta(2) \\
0 & 0 & 1 & 1
\end{bmatrix}
\]

(27)

This simple Hamiltonian \( H \) also agrees with [12], as well as those used in quantum rankings of complex networks [34].

2) Violation of the Sure-Thing Principle: The sure-thing principle, as described in Section II-D dates back to Savage [35] and is given in (5). This principle was refuted in an experiment of Tversky and Shafir [36] and this violation has been regularly experimentally reproduced since. Note that this principle (see Section II-D for more formal definition) follows from the axioms of classical probability, namely the law of total probability.

Thus, any classically probabilistic model for human decision making will be unable to account for such violations, hence the need for generalized quantum models.

Busemeyer et. al. [25] investigate experimental violations of the sure-thing principle (STP) in the context of the Prisoner’s Dilemma, with payoff values \( a = 20, b = 5, c = 10, \) and \( d = 25 \). They find a defection rate of 91% when the opponent is known to defect and 84% when the opponent is known to cooperate. The STP is violated since the defection rate drops to 66% when the choice of the opponent is unknown. We use the previous Lindbladian construction to reproduce this violation [12], see Fig. 4.

3) Quickest Detector Implementation: We implement the quickest change detection protocol of Section III-A as well as the optimal policy computation of Section III-B, within the context of this Prisoner’s Dilemma problem. We take the underlying state to be the choice of the opponent, \( x_n \in \{1\text{(cooperate)}, 2\text{(defect)}\} \). We assume this state jump changes from 2 to 1 according to a geometric distribution with parameter \( p = 0.95 \). We use the following simple measurement model:

\[
Y = \{1, 2, 3\}, \quad B = \begin{bmatrix} 0.6 & 0.25 & 0.15 \\ 0.15 & 0.25 & 0.6 \end{bmatrix}
\]

(28)

where recall \( Y \) is the observation space, i.e. there are three possible observations. These observation are state-dependent with conditional probabilities given by \( B_{x,y} = p(y_n = y|x_n = x) \), for \( y_n \in \{1, 2, 3\}, x_n \in \{1, 2\} \). This observation is used in the computation (10) to obtain \( \eta_n(x) \), which is input to the quantum decision map (4).

The quantum decision maker chooses an action probabilistically according to the map (4), with Lindbladian operator constructed as done in Section V-I. In order to incorporate a violation of the STP, we use the parametrization \( (\alpha = 0.812, \lambda = 10.495, \phi = 0.9) \) (Observe from Fig. 4 that this parametrization can result in an STP violation), and for simplicity we assume the quickest detector knows this parametrization. For false alarm penalty \( f = 5 \) and delay penalty \( d = 1 \), the quickest detector computes its optimal policy via the value iteration algorithm.
(21). This results in an optimal decision threshold of \( \pi(1) = 0.834 \) (denote this as \( \pi' \)), as illustrated in the optimal policy \( \mu^*(\pi) \) plot in Fig. 5.

VI. CONCLUSION

At an abstract level, statistical signal processing deals with signal estimation using sensors, while psychology aims to model (understand) human decision making. We have presented a sequential change detection framework involving a human decision maker (modeled via quantum decision theory from recent results in psychology), sensor and a quickest detector. Quantum decision theory captures important features of human decision making such as order effects and violation of the sure-thing principle (total probability rule). The framework of this paper contributes to the area of human-sensor interface design and analysis.

The aim of our quickest detection formulation was to detect a change in the underlying state by observing the human decisions which are influenced by the state. We characterized the structure of the quickest detection policy. We showed that the optimal policy has a single threshold, and that the optimal cost incurred is lower bounded by that of the classical quickest detection framework, suggesting that the intermediate human decisions cannot improve the detection performance. We have also provided an upper bound on the cumulative cost incurred by the analyst when only a probabilistic estimate of the human’s psychological parametrization is available. This upper bound is given in terms of the cumulative cost incurred when the analyst has perfect knowledge of the parameters and the KL-Divergence between action distributions induced by the true and estimated parametrizations. Finally, we also showed that certain humans in the loop are better than others (w.r.t. quantum decision parameters) in terms of the overall cost in quickest detection. We emphasize that these conclusions are only made possible by the utilization of the quantum decision model, since it allows for analysis w.r.t psychological parametrizations and can model violations of traditional decision-making axioms.

APPENDIX A
QUANTUM DECISION THEORY: LINDBLADIAN OPERATOR CONSTRUCTION

We now discuss the operator \( \mathcal{L}_{(\alpha, \lambda, \phi)} \) in more detail. The evolution of the density operator is given by (2) where

\[
\mathcal{L}_{(\alpha, \lambda, \phi)} \rho_t = -i(1 - \alpha) [H, \rho_t] + \alpha \sum_{m,n} \gamma_{(m,n)} \left( L_{(m,n)} \rho_t L^\dagger_{(m,n)} - \frac{1}{2} \{ L^\dagger_{(m,n)} L_{(m,n)}, \rho_t \} \right)
\]

(29)

Here \( [A, B] = AB - BA, \{ A, B \} = AB + BA, A^* \) is complex conjugate of \( A, H = \text{diag}(1_m, \ldots, 1_m) \) of \( m \times m \) matrix of ones and \( L_{(m,n)} = |k\rangle \langle m| \), where \( |k\rangle \) is the \( k \)th basis vector of \( \mathcal{H} \). The coefficient \( \gamma_{(m,n)} \) is given by the \( (m,n)^{\text{th}} \) element of the cognitive matrix \( C(\alpha, \lambda) \):

\[
\gamma_{(m,n)} := [C(\alpha, \lambda)]_{m,n} = [(1 - \phi) \Pi^T(\lambda) + \phi B^T]_{m,n}
\]

(30)

For utility function \( u : A \times \mathcal{X} \to \mathbb{R} \), construct

\[
p(a_j|\mathcal{E}_t) = \frac{u(a_j|\mathcal{E}_t)^\lambda}{\sum_{A} u(a_j|\mathcal{E}_t)^\lambda}
\]

(31)

and define

\[
P(\mathcal{E}_t) := [p(a_1|\mathcal{E}_t) \ p(a_2|\mathcal{E}_t) \cdots p(a_m|\mathcal{E}_t)] \otimes 1_{K \times 1}
\]

\[
\Pi(\lambda) = \text{diag}(P(\mathcal{E}_t), \ldots, P(\mathcal{E}_K))
\]

(32)

where \( 1_{K \times 1} \) is a vector with all ones and, \( A \otimes B \) is the Kronecker product of \( A \) and \( B \). Define \( \eta_n(s) = p(s|u_n, y_n) \) given the noisy observation \( y_n \) and input signal \( u_n \) at time \( n \in \mathbb{N} \), with \( s \in \mathcal{X} \).

We define

\[
B := \left[ \eta_1(\mathcal{E}_1) \ \eta_2(\mathcal{E}_2) \cdots \eta_K(\mathcal{E}_K) \right] \otimes 1_{m \times 1} \otimes 1_{m \times m}
\]

(33)

See [12] for the psychological motivation behind this structure. (29) is the standard form of the Lindblad-Kossakowski ordinary differential equation, which governs the behavior of quantum systems interacting with an external environment, or ‘open’ quantum systems.

APPENDIX B
PROOF OF THEOREM 1

Proof: We first show that the action distribution \( \Gamma_\pi^a(t) := \mathbb{P}(a|\pi, y) \) induced as the unique steady-state distribution for the Lindbladian \( \mathcal{L}_{(\alpha, \lambda, \phi)} \) with parameters \( (\alpha, \lambda, \phi) \) and initial state \( \rho_0 \), via the map (4), is a continuous function of \( \pi \) and \( y \). We then use this within an induction argument in the value iteration algorithm (21) to complete the argument.

The vectorized solution of (2) [12], for any vectorized initial condition \( \bar{\rho}(0) \), is

\[
\bar{\rho}(t) = \exp(\mathcal{L}_{(\alpha, \lambda, \phi)} t) \bar{\rho}(0)
\]

Fix \( t > 0 \) and, examining the structure of the operator \( \mathcal{L}_{(\alpha, \lambda, \phi)} \) (29), consider the map

\[
\Lambda^t_{(m,n)} : \eta_h(\cdot) \to [\exp(\mathcal{L}_{(\alpha, \lambda, \phi)} t) \bar{\rho}(0)]_{(m,n)}
\]

By inspection, each element \( [\mathcal{L}_{(\alpha, \lambda, \phi)}]_{(m,n)} \) of (29) is continuous with respect to \( \gamma_{(m,n)} \) and thus with respect to \( \eta_h(\cdot) \) (by (30) and (33)). Thus the map \( \Lambda^t_{(m,n)} \) is continuous with respect to \( y \) and \( \pi \forall (m,n) \in [1,d]^2, t \in \mathbb{R}_+ \). Also observe that \( \eta_h(\cdot) \) is a continuous function of \( \pi \) (for a fixed observation \( y \)) and \( y \) (for a fixed prior \( \pi \)), as a Bayesian update. So we have that the action distribution at time \( t \)

\[
\Gamma^\pi_y(a,t) = \text{Tr}(P_a \exp(\mathcal{L}_{(\alpha, \lambda, \phi)} t) \bar{\rho}(0) P_a^\dagger)
\]

is continuous with respect to \( y \) and \( \pi \forall (m,n) \in [1,d]^2, t \in \mathbb{R}_+ \). Thus the stationary distribution \( \Gamma^\pi_y(a) = \lim_{t \to \infty} \Gamma^\pi_y(a,t) \) is continuous with respect to \( \pi \) and \( y \). We now use induction on the value iteration algorithm (21). The algorithm begins with \( \mathbb{V}_0(\pi) = 0 \forall \pi \in \Pi \). Thus \( \mathbb{V}_0(\pi) \) is trivially concave. Now assume \( \mathbb{V}_k(\pi) \) is concave for some \( k \in \mathbb{N} \). We have the update \( \mathbb{V}_{k+1}(\pi) = \min_{\pi} \mathbb{C}(\pi, 1), \mathbb{C}(\pi, 2) + \sum_{\pi_1 \in A_1 \times A_2} \mathbb{V}_k(T(\pi, a_1)|\sigma(\pi, a_2)) \). Observe from (21) that \( \mathbb{V}_k(\pi) \) is positively homogeneous; that is, for any \( \alpha > 0 \),
\( V_k(\alpha \pi) = \alpha V_k(\pi) \). Choosing \( \alpha = \sigma(\pi, a) \) yields \( V_{k+1}(\pi) = \text{min}\{C(\pi, 1), C(\pi, 2) + \sum_{a \in A} \sum_{\pi \in \mathcal{P}} V_k(R_k(a)P(\pi, a))\} \) Recall that \( R_k(a) \) is computed via (13), and thus we have that \( R_k(a) \) is a continuous function of \( \pi \), \( \forall a \in A \). Also recall that \( C(\pi, 1) \) and \( C(\pi, 2) \) are linear in \( \pi \). Thus \( V_{k+1}(\pi) \) is concave. This completes the induction step. Now the value iteration algorithm (21) converges pointwise, so the optimal value function \( \bar{V}(\pi) = J^*_\mu(\pi) = \lim_{k \to \infty} V_k(\pi) \) is concave. This immediately implies, by (19), that the optimal policy \( \bar{\pi}(\mu) \) cannot have more than one threshold.

**Corollary 1:** Recall that \( \Gamma^\pi_\mu(a) \) is inherently dependent on a choice of parameters \((\alpha, \lambda, \phi)\). From the Proof of Theorem 1 and by inspection of the Lindbladian structure (29), we have that \( \Gamma^\pi_\mu(a) \) is continuous with respect to parameters \((\alpha, \lambda, \phi)\). Then also \( R_{\pi, \pi}(a) \) (13) is continuous with respect to \((\alpha, \lambda, \phi)\).

**APPENDIX D**

**Proof of Theorem 3**

**Proof:** The proof is adapted from that provided for Theorem 14.9.1 of [33]. We first note that by the reasoning of [32] (Appendix, Proof of Theorem 2), POMDPs \( \theta \) and \( \hat{\theta} \) have implicit discount factor \( \gamma = 1 - p \). The cumulative cost incurred by applying policy \( \mu(\pi) \) to model \( \theta \) satisfies at time \( n \)

\[
J^\mu(\pi; \theta) = C(\mu(\pi; \theta) + \gamma \sum_a J^{\mu - 1}(T(\pi, a, \mu(\pi); \theta))(\sigma(\pi, a, \mu(\pi); \theta) + \gamma). (35)
\]

Therefore, the absolute difference in cumulative cost for models \( \theta, \hat{\theta} \) satisfies

\[
|J^\mu(\pi; \theta) - J^\mu(\pi; \hat{\theta})| \leq \gamma \sum_a \sigma(\pi, a, \mu(\pi); \theta)\sigma(\pi, a, \mu(\pi); \hat{\theta}) + \gamma \sum_a J^{\mu - 1}(T(\pi, a, \mu(\pi); \theta))|\sigma(\pi, a, \mu(\pi); \theta) - \sigma(\pi, a, \mu(\pi); \hat{\theta})| - J^{\mu-1}(T(\pi, a, \mu(\pi); \hat{\theta}))|\sigma(\pi, a, \mu(\pi); \theta) - \sigma(\pi, a, \mu(\pi; \hat{\theta}))|
\]

Therefore, the absolute difference in cumulative cost for models \( \theta, \hat{\theta} \) satisfies

\[
|J^\mu(\pi; \theta) - J^\mu(\pi; \hat{\theta})| \leq \gamma \sum_a \sigma(\pi, a, \mu(\pi); \theta)\sigma(\pi, a, \mu(\pi); \hat{\theta}) + \gamma \sum_a J^{\mu - 1}(T(\pi, a, \mu(\pi); \theta))|\sigma(\pi, a, \mu(\pi); \theta) - \sigma(\pi, a, \mu(\pi; \hat{\theta}))|
\]

Observe that \( \sum_{\pi,a} \sigma(\pi, a, \mu(\pi); \theta) = 1 \). Then evaluating \( \sigma(\pi, a, \mu(\pi; \theta) = 1 \) yields \( \sigma(\pi, a, \mu(\pi; \theta) - \sigma(\pi, a, \mu(\pi); \hat{\theta})| \)

\[
J^\mu(\pi; \theta) - J^\mu(\pi; \hat{\theta}) \leq \gamma \sum_a \sigma(\pi, a, \mu(\pi; \theta) - \sigma(\pi, a, \mu(\pi; \hat{\theta}))| \)
\]

where the last inequality follows from Pinsker’s inequality. Now we also have

\[
\sup_{\pi \in \Pi} J^{\mu - 1}(\pi; \theta) \leq \frac{1}{1 - \gamma} \max_i C(\epsilon_i, u)
\]

We use these bounds in (36) to obtain

\[
\sup_{\pi \in \Pi} |J^{\mu - 1}(\pi; \theta) - J^{\mu - 1}(\pi; \hat{\theta})| \leq \gamma \sup_{\pi \in \Pi} |J^{\mu - 1}(\pi; \theta) - J^{\mu - 1}(\pi; \hat{\theta})| + \sqrt{2}\gamma \max_i C(\epsilon_i, u) \sup_{\pi \in \Pi} \sum_j P_{ij}[D(R_{ij, \pi} || R_{ij, \pi})]^{1/2}
\]

Now starting with \( J^{(0)}(\pi; \theta) = J^{(0)}(\pi; \hat{\theta}) \), unraveling (39) yields

\[
\sup_{\pi \in \Pi} |J^{(0)}(\pi; \theta) - J^{(0)}(\pi; \hat{\theta})| \leq \frac{\sqrt{2}}{1 - \gamma} \max_i C(\epsilon_i, u) \sup_{\pi \in \Pi} \sum_j P_{ij}[D(R_{ij, \pi} || R_{ij, \pi})]^{1/2}
\]
The algorithm begins with $V_0(\pi) = \mathcal{V}_0(\pi) = 0 \ \forall \pi \in \Pi$, so we trivially have $V_0(\pi) \geq \mathcal{V}_0(\pi)$. We also know that the value function $\mathcal{V}_k$ is concave for all $\pi \in \Pi$ (see proof of Theorem 1), so Jensen’s inequality can be invoked to produce

$$\begin{align*}
\mathcal{V}(T(\pi, a)) &= \mathcal{V} \left( \sum_{i=1}^{A} \hat{T}(\pi, i) \frac{\hat{\sigma}(\pi, i)}{\sigma(\pi, a)} M(i, a) \right) \\
&\geq \sum_{i=1}^{A} \hat{\mathcal{V}}(T(\pi, i)) \frac{\hat{\sigma}(\pi, i)}{\sigma(\pi, a)} M(i, a) 
\end{align*}$$

(46)

Thus $\sum_{a=1}^{A} \hat{\mathcal{V}}(T(\pi, a)) \sigma(\pi, a) \geq \sum_{a=1}^{A} \hat{\mathcal{V}}(T(\pi, a)) \sigma(\pi, a)$ and, assuming $\mathcal{V}_k(\pi) \geq \mathcal{V}_k(\pi)$, we have

$$C(\pi, 2) + \sum_{a=1}^{A} \hat{\mathcal{V}}(T(\pi, a)) \sigma(\pi, a) \geq C(\pi, 2) + \sum_{a=1}^{A} \hat{\mathcal{V}}(T(\pi, a)) \sigma(\pi, a)$$

(47)

Thus $\mathcal{V}_{k+1}(\pi) \geq \mathcal{V}_k(\pi)$ and the induction step is complete. The value iteration algorithm (21) converges pointwise, so $\mathcal{V}(\pi) \geq \hat{\mathcal{V}}(\pi)$. 

\[ \square \]

APPENDIX E

PROOF OF THEOREM 4

Proof: Consider the update (12) and define $R_{x, \pi}(a) = \int_{Y} \Gamma_{y}^{\pi-1}(a)B_{x,y}dy$ and $\hat{R}_{x, \pi}(a) = \int_{Y} \hat{\Gamma}_{y}^{\pi-1}(a)B_{x,y}dy$. Using $\Gamma_{y}^{\pi}(a) = \sum_{i=1}^{A} \Gamma_{y}^{\pi}(i)M(i, a)$ (where $A$ is the cardinality of the action space) yields

$$R_{x, \pi}(a) = \int_{Y} \sum_{i=1}^{A} \Gamma_{y}^{\pi-1}(i)M(i, a)B_{x,y}dy$$

$$= \sum_{i=1}^{A} \int_{Y} \Gamma_{y}^{\pi-1}(i)B_{x,y}dyM(i, a)$$

$$= \sum_{i=1}^{A} \hat{R}_{x, \pi}(i)M(i, a)$$

(43)

Now, following (12):

$$T(\pi, a) = \frac{R_{x, \pi}(a)P^{'\pi}}{\sigma(\pi, a)} = \begin{bmatrix} 1 \end{bmatrix} R_{x, \pi}(a)P^{'\pi}$$

$$\hat{T}(\pi, a) = \frac{\hat{R}_{x, \pi}(a)P^{'\pi}}{\hat{\sigma}(\pi, a)} = \begin{bmatrix} 1 \end{bmatrix} \hat{R}_{x, \pi}(a)P^{'\pi}$$

$$R_{x}(a) = \text{diag}(R_{1, \pi}(a), R_{2, \pi}(a))$$

$$\hat{R}_{x}(a) = \text{diag}(\hat{R}_{1, \pi}(a), \hat{R}_{2, \pi}(a))$$

(44)

Now observe that we can manipulate $T(\pi, a)$ in the following way:

$$T(\pi, a) = \frac{R_{x}(a)P^{'\pi}}{\sigma(\pi, a)} = \sum_{i=1}^{A} \hat{R}_{x}(i)M(i, a)P^{'\pi}$$

$$= \frac{\sum_{i=1}^{A} \hat{R}_{x}(i)M(i, a)}{\sigma(\pi, a)} \hat{\sigma}(\pi, a)$$

$$= \sum_{i=1}^{A} \frac{\hat{T}(\pi, i)}{\sigma(\pi, a)} M(i, a)$$

(45)

We now use induction in the value iteration algorithm (21). The algorithm begins with $V_0(\pi) = \mathcal{V}_0(\pi) = 0 \ \forall \pi \in \Pi$, so we also know that $\mathcal{V}_k(\pi)$ is concave for all $\pi \in \Pi$. This allows us to verify the existence of convex regions in parameter space for which the performance of the quickest detector is strictly ordered. The first result, Example 1, verifies that for a specific subset of the parameter space, a convex combination of parameterizations which result in Blackwell dominated distributions also results in a Blackwell dominated distribution. The second result, Example 2, verifies the converse, that a convex combination of parameterizations which result in distributions that are Blackwell dominating is also Blackwell dominating. These two in conjunction reveal that the Blackwell ordering is closed under

\[ \square \]

APPENDIX F

PROOF OF THEOREM 5

Proof: We prove the existence of such a stochastic matrix by construction. First, we have that $\Gamma_{1} = \Gamma_{M_{1}}$ and $\Gamma_{2} = M_{2} \hat{\Gamma}$, so $\Gamma_{1}(a) = \sum_{i=1}^{N} \hat{\Gamma}(i)M_{1}(i, a)$, $\Gamma_{2}(a) = \sum_{i=1}^{N} M_{2}(a, i)\hat{\Gamma}(i)M_{2}(i, 1)$ Then

$$\Gamma_{3}(a) = \gamma_{a}\Gamma_{1}(a) + (1 - \gamma_{a})\Gamma_{2}(a)$$

$$= \gamma_{a}\sum_{i=1}^{N} \hat{\Gamma}(i)M_{1}(i, a) + (1 - \gamma_{a})\sum_{i=1}^{N} \hat{\Gamma}(i)M_{2}(i, a)$$

$$= \sum_{i=1}^{N} \left( \gamma_{a}M_{1}(i, a) + (1 - \gamma_{a})M_{2}(i, a) \right) \hat{\Gamma}(i)$$

(48)

Now simply form matrix $M_{3}$ as $M_{3}(i, a) = \gamma_{a}M_{1}(i, a) + (1 - \gamma_{a})M_{2}(i, a)$ so that $\Gamma_{3} = M_{3} \hat{\Gamma}$. It is also easily verified that $M_{3}$ is stochastic, since $M_{1}$ and $M_{2}$ are stochastic.

\[ \square \]

APPENDIX G

NUMERICAL RESULTS

We now present a series of computational results which allow us to verify the existence of convex regions in parameter space for which the performance of the quickest detector is strictly ordered. The first result, Example 1, verifies that for a specific subset of the parameter space, a convex combination of parameterizations which result in Blackwell dominated distributions also results in a Blackwell dominated distribution. The second result, Example 2, verifies the converse, that a convex combination of parameterizations which result in distributions that are Blackwell dominating is also Blackwell dominating. These two in conjunction reveal that the Blackwell ordering is closed under
convex parameter combinations, and thus we can interpolate this dominance ordering to hold between all points within the convex hulls of computed dominance points; this is studied in Example 3 below.

Denote $\Gamma_{y,1}$ and $\Gamma_{y,2}$ as the map (4) outputs for distinct Lindbladian parameterizations $(\alpha_1, \lambda_1, \phi_1)$ and $(\alpha_2, \lambda_2, \phi_2)$, respectively. Let $(\alpha_3, \lambda_3, \phi_3) = \epsilon (\alpha_1, \lambda_1, \phi_1) + (1 - \epsilon) (\alpha_2, \lambda_2, \phi_2)$ for $\epsilon \in (0, 1)$, and $\Gamma_{y,3}$ be the resultant action distribution (from (4)) for Lindbladian parameterization $(\alpha_3, \lambda_3, \phi_3)$.

**Numerical Verification 1:** We have verified numerically that for $(\alpha_1, \alpha_2, \phi_1, \phi_2) \in [0.1, 0.5]$, $(\alpha_1, \lambda_1, \phi_1) \in [0, 100]$, and $\epsilon \in (0, 1)$, there exists $\{\gamma_\alpha\}, \alpha \in \{1, \ldots, M\}$ such that $\gamma_\alpha \Gamma_{y,1} + (1 - \gamma_\alpha) \Gamma_{y,2}$. In words, within the parameter confines defined, interpolating between two Lindbladian parameterizations via a convex combination results in an action distribution for which each action probability is a convex combination of (lies between) the action probabilities resulting from the initial two parameterizations.

This numerical verification, along with the following two Theorems, will be used to prove our first computational result which reveals that, for certain regions in the parameter space, a Blackwell dominance order is closed under convex combinations.

**Result 1.** Performance dominance is closed under convex combinations of dominated distributions: Consider Lindbladian parameterizations $(\alpha_1, \lambda_1, \phi_1) \in [0.1, 0.5] \times [10, 100] \times [0.1, 0.5]$, $(\alpha_2, \lambda_2, \phi_2) \in [0.1, 0.5] \times [10, 100] \times [0.1, 0.5]$, $(\alpha_3, \lambda_3, \phi_3) \in [0.1, 0.5] \times [10, 100] \times [0.1, 0.5]$ and respective resultant action distributions $\Gamma_{y,1}, \Gamma_{y,2}, \Gamma_{y,3}$ from the map (4) with prior $\pi$. Let $(\bar{\alpha}, \bar{\lambda}, \bar{\phi}) = \epsilon (\alpha_1, \lambda_1, \phi_1) + (1 - \epsilon) (\alpha_2, \lambda_2, \phi_2)$ for $\epsilon \in [0, 1]$, and $\Gamma_{y,\bar{\pi}}$ the resultant action distribution. Denote $V(\pi), V_1(\pi), V_2(\pi), V_3(\pi)$ the value functions resulting from the value iteration algorithm (21) using $\Gamma_{y,1}, \Gamma_{y,2}, \Gamma_{y,3}$ in (12), respectively. Suppose there exists stochastic matrices $M_{1}^\pi, M_{2}^\pi$ such that $\Gamma_{y,1} = \Gamma_{y,2}^\pi M_{1}^\pi$ and $\Gamma_{y,2} = \Gamma_{y,3}^\pi M_{2}^\pi$. Then we have $V(\pi) \geq V_1(\pi), V_1(\pi) \geq V_2(\pi), V_2(\pi) \geq V_3(\pi) \forall \pi \in \Pi$.

**Verification:** By the numerical verification 1, we have that

$$\exists \{\gamma_\alpha\}_{a=1}^{A} : \Gamma_{y}^\pi(a) = \gamma_\alpha \Gamma_{y,1}^\pi(a) + (1 - \gamma_\alpha) \Gamma_{y,2}^\pi(a)$$

Thus, by Theorem 5, there exists a stochastic matrix $M_{y}^\pi$ such that $\Gamma_{y,1} = \Gamma_{y,2} M_{1}^\pi$. Then by invoking Theorem 4 using each equality $\Gamma_{y,1} = \Gamma_{y,2}^\pi M_{1}^\pi$, $\Gamma_{y,2} = \Gamma_{y,3}^\pi M_{2}^\pi$, $\Gamma_{y,3} = \Gamma_{y,2} M_{2}^\pi$, the results follow.

Now we provide another numerical verification and two computational results which allow us to conclude closure of Blackwell dominance orderings in the opposite direction.

Denote $\Gamma_{y,1}$ and $\Gamma_{y,2}$ as the map (4) outputs for distinct Lindbladian parameterizations $(\alpha_1, \lambda_1, \phi_1)$ and $(\alpha_2, \lambda_2, \phi_2)$, respectively. Let $(\alpha_3, \lambda_3, \phi_3) = \epsilon (\alpha_1, \lambda_1, \phi_1) + (1 - \epsilon) (\alpha_2, \lambda_2, \phi_2)$ for $\epsilon \in (0, 1)$, and $\Gamma_{y,3}$ be the resultant action distribution (from (4)) for Lindbladian parameterization $(\alpha_3, \lambda_3, \phi_3)$.

**Numerical Verification 2:** We have verified numerically that for $(\alpha_1, \alpha_2) \in [0.8, 1], \phi_1, \phi_2 \in [0.1, 0.5], \lambda_1, \lambda_2 \in [10, 100]$, and $\epsilon \in (0, 1)$, there exists $\{\gamma_\alpha\}, \alpha \in \{1, \ldots, M\}$ such that $\gamma_\alpha \Gamma_{y,1} + (1 - \gamma_\alpha) \Gamma_{y,2}$.
quantum model with respect to pertinent microeconomic models such as Prospect Theory [19].

1) Comparison with Anticipatory Decision-Making [8]: [8] analyzes a quickest change detection scheme where the human decisions are made via a behavioral economics anticipatory decision model. The anticipatory model of [8], first proposed in [37], aims to model agents who make optimal decisions while taking into account the probabilities of future events. Our model and results differ from [8] in the following ways:

i) Modeling Aims: The anticipatory model of [8] captures the peculiarities of multi-stage human planning, where due to antipatatory emotions, the current decision depends on the probabilities of future decisions. In contrast, QDT models decision anomalies such as violation of the sure thing principle and order effects, which are empirically observed [13], [14], [15]. The two models are complementary and capture different effects of human decision making.

ii) Optimal Decision-Making Policy: In [8], the optimal policy is shown to exhibit multiple thresholds, while we show that the optimal policy utilizing the quantum model is single-threshold.

iii) Psychological Parameter Analysis: The quantum model, in contrast to that of [8], contains a free-parametrization (governing an individual decision-maker’s unique decision processing) which we exploit in our analysis. We analyze the sensitivity of detection performance to knowledge of this parametrization, and show that detection performance is robust to parameter estimation. We also analyze the dependence of detection performance on these parameters, and find that improving agent rationality improves detection performance. Thus, this work provides contrasting insights into this human-machine quickest detection structure, which reveals the relevance of the human decision model for the detector’s performance.

2) Comparison with Prospect Theory: It is worthwhile contrasting the Quantum Decision Theory (QDT) [12] with the two classical microeconomic decision models, Expected Utility Theory (EUT) [17] and Prospect Theory (PT) [19]. PT models a human’s decision as maximizing an expected subjective utility, constructed as the composition of a non-linear weighting function with the objective utility function. Thus the QDT can generalize PT.

i) More General Decision Phenomena: Empirical findings in cognitive psychology and behavioral science exhibit anomalies with respect to classical decision models defined by Kolmogorov’s probability axioms [12], which PT assumes [38]. QDT expands Kolmogorov probability to quantum probability, thereby expanding the range of decision phenomena which can be explained; see [13], [14], [15].

ii) Dynamic Preference Evolution: PT is inherently static – it models decisions made at a single time-instant. In comparison, QDT models time-evolving preference distributions. In this paper, we consider the stationary distribution of the QDT dynamics to generate observed decisions. But the ability of QDT to incorporate aspects of time-evolution such as choice-induced preference change makes it appealing for general human-machine interaction frameworks.

For a more detailed comparison between QDT and PT, the reader is referred to [39] and [40].

REFERENCES

[1] A. N. Shiryaev, “On optimum methods in quickest detection problems,” Theory Probability Appl., vol. 8, no. 1, pp. 22–46, 1963.

[2] A. Tartakovsky and G. Moustakides, “State-of-the-art in Bayesian changepoint detection,” Sequential Anal., vol. 29, no. 2, pp. 125–145, 2010.

[3] Y. Liang, A. G. Tartakovsky, and V. V. Veeravalli, “Quickest change detection with non-stationary post-change observations,” 2021, arXiv:2110.01581.

[4] H. V. Poor and O. Hadjiliadis, Quickest Detection. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[5] Y. C. Chen, T. Banerjee, A. D. Dominguez-Garcia, and V. V. Veeravalli, “Quickest line outage detection and identification,” IEEE Trans. Power Syst., vol. 31, no. 1, pp. 749–758, Jan. 2016.

[6] V. Raghavan and V. V. Veeravalli, “Quickest change detection of a Markov process across a sensor array,” IEEE Trans. Inf. Theory, vol. 56, no. 4, pp. 1961–1981, Apr. 2010.

[7] A. Lakhina, M. Crovella, and C. Diot, “Diagnosing network-wide traffic anomalies,” ACM SIGCOMM Comput. Commun. Rev., vol. 34, no. 4, pp. 219–230, 2004.

[8] V. Krishnamurthy, “Quickest change detection of time inconsistent anticipatory agents. Human-sensor and cyber-physical systems,” IEEE Trans. Signal Process., vol. 69, pp. 1054–1069, 2021.

[9] V. Krishnamurthy, “Quickest detection POMDPs with social learning: Interaction of local and global decision makers,” IEEE Trans. Inf. Theory, vol. 58, no. 8, pp. 5563–5587, Aug. 2012.

[10] V. Krishnamurthy, “How to schedule measurements of a noisy Markov chain in decision making?” IEEE Trans. Inf. Theory, vol. 59, no. 7, pp. 4440–4461, Jul. 2013.

[11] D. Kahneman, P. S. Slovic, P. Slovic, and A. Tversky, Judgment Under Uncertainty: Heuristics and Biases. Cambridge, U.K.: Cambridge Univ. Press, 1982.

[12] I. Martínez-Martínez and E. Sánchez-Burillo, “Quantum stochastic walks on networks for decision-making,” Sci. Rep., vol. 6, no. 1, pp. 1–13, 2016.

[13] D. Aerts, J. Broekaert, M. Czachor, and B. D’Hooghe, “A quantum-conceptual explanation of violations of expected utility in economics,” in Proc. Int. Symp. Quantum Interact., New York, NY, USA: Springer-Verlag, 2011, pp. 192–198.

[14] J. S. Trueblood and J. R. Busseymeyer, “A quantum probability account of order effects in inference,” Cogn. Sci., vol. 35, no. 8, pp. 1518–1552, 2011.

[15] J. R. Busseymeyer, E. M. Pothos, R. Franco, and J. S. Trueblood, “A quantum theoretical explanation for probability judgment errors,” Psychol. Rev., vol. 118, no. 2, p. 193, 2011.
[16] E. M. Pothos and J. R. Busemeyer, “A quantum probability explanation for violations of ‘rational’ decision theory,” Proc. Roy. Soc. B: Biol. Sci., vol. 276, no. 1665, pp. 2171–2178, 2009.

[17] O. Morgenstern and J. Von Neumann, Theory of Games and Economic Behavior, Princeton, NJ, USA: Princeton Univ. Press, 1953.

[18] L. J. Savage, “The theory of statistical decision,” J. Amer. Stat. Assoc., vol. 46, no. 253, pp. 55–67, 1951.

[19] D. Kahneman and A. Tversky, “Prospect theory: An analysis of decision under risk,” in Handbook of the Fundamentals Financial Decision Making. Part I. Singapore: World Scientific, 2013, pp. 99–127.

[20] J. Busemeyer and P. Bruza, Quantum Models of Cognition and Decision. Cambridge, U.K.: Cambridge Univ. Press, 2012. Accessed: Dec. 8, 2021. [Online]. Available: https://books.google.com/books?id=0vxhTG_ZLAC

[21] A. Khrennikov, Ubiquitous Quantum Structure. New York, NY, USA: Springer-Verlag, 2010.

[22] V. I. Yukalov and D. Sornette, “Mathematical structure of quantum decision theory,” Adv. Complex Syst., vol. 13, no. 5, pp. 659–698, 2010.

[23] V. I. Yukalov and D. Sornette, “Quantitative predictions in quantum decision theory,” IEEE Trans. Syst., Man, Cybern. Syst., vol. 48, no. 3, pp. 366–381, Mar. 2018.

[24] A. Y. Khrennikov and E. Haven, “Quantum mechanics and violations of the sure-thing principle: The use of probability interference and other concepts,” J. Math. Psychol., vol. 53, no. 5, pp. 378–388, 2009.

[25] J. Busemeyer, M. Matthew, and Z. Wang, “Quantum game theory explanation of disjunction effect,” in Proc. 28th Annu. Conf. Cogn. Sci. Soc., Erlbaum Mahwah, 2006, pp. 131–135.

[26] D. Manzano, “A short introduction to the Lindblad master equation,” AIP Adv., vol. 10, no. 2, 2020, Art. no. 025106.

[27] J. R. Bussemyer, Z. Wang, and A. Lambert-Mogiliansky, “Empirical comparison of Markov and quantum models of decision making,” J. Math. Psychol., vol. 53, no. 5, pp. 423–433, 2009.

[28] J. T. Townsend et al., Stochastic Modeling of Elementary Psychological Processes. CUP Archive, 1983.

[29] P. D. Kvam, J. R. Busemeyer, and T. J. Pleskac, “Temporal oscillations in preference strength provide evidence for an open system model of constructed preference,” Sci. Rep., vol. 11, no. 1, pp. 1–15, 2021.

[30] D. Blackwell, “Comparison of experiments,” Proc. 2nd Berkeley Symp. on Math. Stat. Probab., vol. 2., Berkeley, CA, USA: Univ. of California Press, 1951.

[31] A. G. Tartakovsky and V. V. Veeravalli, “General asymptotic Bayesian theory of quickest change detection,” Theory Probability Appl., vol. 49, no. 3, pp. 458–497, 2005.

[32] L. A. Johnston and V. Krishnamurthy, “Opportunistic file transfer over a fading channel: A POMDP search theory formulation with optimal threshold policies,” IEEE Trans. Wireless Commun., vol. 5, no. 2, pp. 394–405, Feb. 2006.

[33] V. Krishnamurthy, Partially Observed Markov Decision Processes, Cambridge, U.K.: Cambridge Univ. Press, 2016.

[34] E. Sánchez-Burillo, J. Duch, J. Gómez-Gardenes, and D. Zueco, “Quantum navigation and ranking in complex networks,” Sci. Rep., vol. 2, no. 1, pp. 1–8, 2012.

[35] L. J. Savage, The Foundations of Statistics. Courier Corporation, 1972.

[36] A. Tversky and E. Shafir, “The disjunction effect in choice under uncertainty,” Psychol. Sci., vol. 3, no. 5, pp. 305–310, 1992.

[37] A. Caplin and J. Leahy, “Psychological expected utility theory and anticipatory feelings,” Quart. J. Econ., vol. 116, no. 1, pp. 55–79, 2001.

[38] M. H. Birmbaum, “New paradoxes of risky decision making,” Psychol. Rev., vol. 115, no. 2, pp. 463–501, 2008.

[39] M. Favre, A. Wittwer, H. R. Heinimann, V. I. Yukalov, and D. Sornette, “Quantum decision theory in simple risky choices,” PLoS One, vol. 11, no. 12, 2016, Art. no. e0168045.

[40] G. M. Ferro, T. Kovalenko, and D. Sornette, “Quantum decision theory augments rank-dependent expected utility and cumulative prospect theory,” J. Econ. Psychol., vol. 86, 2021, Art. no. 102417.

Luke Snow (Graduate Student Member, IEEE) received the B.S. (summa cum laude) degree in electrical engineering and a minor in mathematics from Clemson University, in 2021. He is working toward the Ph.D. degree with the Electrical and Computer Engineering Department at Cornell University, where he is advised by Prof. Vikram Krishnamurthy. His research interests include statistical signal processing, human-sensor systems, and inverse reinforcement learning. He received the “W.M. Riggs” Award as the most outstanding senior in electrical engineering. He was a recipient of the NSF Graduate Research Fellowship. He has held Research Intern positions with M.I.T. Lincoln Laboratory and the U.S. Army Research Laboratory.

Vikram Krishnamurthy (Fellow, IEEE) received the B.E. degree in electrical engineering from the University of Auckland, in 1988, and the Ph.D. degree in mathematical systems theory from Australian National University, in 1992. Currently, he is a Professor with the School of Electrical and Computer Engineering, Cornell University. From 2002 to 2016, he was a Professor and the Canada Research Chair of the University of British Columbia, Canada. His research interests include statistical signal processing and stochastic control in social networks and adaptive sensing. He served as a Distinguished Lecturer for the IEEE Signal Processing Society and the Editor-in-Chief of the IEEE Journal on Selected Topics in Signal Processing. In 2013, he was awarded an Honorary Doctorate from KTH (Royal Institute of Technology), Sweden. He is the author of the book titled Partially Observed Markov Decision Processes (Cambridge University Press, 2016).

Brian M. Sadler (Life Fellow, IEEE) received the B.S. and M.S. degrees from the University of Maryland, College Park, and the Ph.D. degree from the University of Virginia, Charlottesville, all in electrical engineering. He is the U.S. Army Senior Research Scientist for Intelligent Systems, and a Fellow of the Army Research Laboratory (ARL), Adelphi, MD, USA. He has been a Distinguished Lecturer for IEEE Signal Processing and IEEE Communications Societies, and an Editor for several journals in Signal Processing, Communications, and Robotics. He received Best Paper Awards from IEEE Signal Processing Society in 2006 and 2010, several ARL and Army R&D awards, and a 2008 Outstanding Invention of the Year Award from the University of Maryland. He received the Presidential Rank Award in 2021. His research interests include multiagent intelligent systems, signal processing, and information science.