Lyapunov based Stochastic Stability of a Quantum Decision System for Human-Machine Interaction

Luke Snow, Shashwat Jain and Vikram Krishnamurthy, Fellow, IEEE

Abstract—In mathematical psychology, decision makers are modeled using the Lindbladian equations from quantum mechanics to capture important human-centric features such as order effects and violation of the sure thing principle. We consider human-machine interaction involving a quantum decision maker (human) and a controller (machine). Given a sequence of human decisions over time, how can the controller dynamically provide input messages to adapt these decisions so as to converge to a specific decision? We show via novel stochastic Lyapunov arguments how the Lindbladian dynamics of the quantum decision maker can be controlled to converge to a specific decision asymptotically. Our methodology yields a useful mathematical framework for human-sensor decision making. The stochastic Lyapunov results are also of independent interest as they generalize recent results in the literature.

Index Terms—Stochastic Lyapunov Stability, Super-martingales, Lindbladian Evolution, Quantum Decision Making

I. INTRODUCTION

Recent studies in mathematical psychology [1], [2], [3], show that the Lindbladian equations from quantum mechanics facilitate modeling peculiar aspects of human decision making. Specifically, such quantum decision models capture order effects (humans perceive $P(H|A \cap B)$ and $P(H|B \cap A)$ differently in decision making), violation of the sure thing principle (human perception of probabilities in decision making violates the total probability rule), and temporal oscillations in decision preferences. Motivated by the design of human-machine interaction systems, this paper addresses the following question: Given a sequence of human decisions over time, how can a controller (machine) adapt the Lindbladian dynamics (of the human decision maker) so as to converge to a specific decision? To investigate this, we develop a novel generalization of recent results involving finite-step stochastic Lyapunov functions. Thus at an abstract level, we study the stochastic stability of a switched controlled Lindbladian dynamic system where the switching occurs due to the interaction of the controller (machine) and decision maker (human) at specific (possibly random) time instants.

Partial results of this paper were submitted to IEEE CDC 2022.

This research was supported in part by the National Science Foundation grant CCF-2112457 and Army Research Office grant W911NF-19-1-0365

Luke Snow las474@cornell.edu, Shashwat Jain sj474@cornell.edu, Vikram Krishnamurthy vikram@cornell.edu are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA

All three authors contributed equally to this work.

A. Decision Making Context

Figure 1 shows our schematic setup. The finite-valued random variable $s \sim \pi_0(\cdot)$ denotes the underlying state of nature, where $\pi_0$ is a known probability mass function. The input signals $y_k$ and $z_k$ are noisy observations of the state with conditional observation densities $p(y|s)$ and $p(z|s)$, respectively. The observation $y_k$ is fed into a Bayesian recommender which recommends humans the most probable underlying state $\alpha_k$, thereby influencing the human’s psychological state $\rho$. This state is represented as a density operator in Hilbert Space, which evolves via the Lindbladian equation parametrized by the recommendations $\alpha_k$ and control input $u_k$. The controller gets its own version of the knowledge $\beta_k$ of the underlying state from the sensor which observes the underlying state as $z_k$. The controller detects the actions $\alpha_k$ and using the sensor input $\beta_k$ it outputs a control signal to $u_k$ to the human. The density operator $\rho$ encodes a probability distribution over actions $\{a_j\}_{j \in \{1,\ldots,m\}}$, and at each time point an action is taken according to this distribution. The machine observes the actions and outputs a feedback control signal $u_k$ to the human.

Examples: Several examples in robotics [4], interactive marketing/advertising [5], recommender systems [6] and control of game-theoretic economic interactions [7] exploit models for human decision making. One example is a machine assisted healthcare system for patients with dementia [8], in which the patient is assisted by a machine (smart watch) to wash his hands. The machine’s sensor detects whether a certain set of sequential actions are followed by the patient, and then sends those results to a controller which gives an audio/video command to the patient. In this context the underlying states ($s$) are the tap water, soap dispenser and towel dispenser, which are partially observed by both sensor and the patient. The patient has a psychological state ($\rho$) and the resultant hand washing actions ($a_k$) are detected by the controller which gives the control input ($u_k$). In our work, we model the psychological state of the patient as a Lindbladian evolution as shown in Figure 1 since this accounts for a wider range of human behavior, such as irrational decisions which could be made by the dementia patient, than classical models.

B. Main Results and Organization

Given the described human-machine decision making system, the question we ask is: Can the human’s decision preference be guided by the input control signals such that a desired target action is eventually taken at every time step? Our results reveal that this indeed is the case. We show this by developing...
Controller

Human-Machine Interaction Model. The state is observed by both Sensor and Bayesian Recommender which send a posterior to both Controller and Human. The Human and Machine interact via action and control feedback, where Human acts based upon the control input given by the controller and the Bayesian Recommender. The machine controller uses input from sensor and actions of the human and gives a control feedback to human after $T$ time-steps.

![Diagram of Human-Machine Interaction Model](image)

C. Literature Review

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) [11] and Savage (1954) [12]. Despite the successes of these models, numerous experimental findings, most notably those of Kahneman and Tversky [13], have demonstrated violations of the proposed decision making axioms. These violations are natural motivators for treating 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.

**Remark.** QDT models in psychology do not claim that the brain is acting as a quantum device in any physical sense. Instead QDT serves as a parsimonious generative blackbox model for human decision making that is backed up by extensive experimental studies [2], [15].

Within Quantum Decision Theory, several recent advances have utilized quantum dynamical systems to model time-evolving decision preferences. The classical model for this type of time-evolving mental state is a Markovian model, but in [22] an alternative formulation based on Schrödinger’s Equation is developed. This model is shown to both reconcile observed violations of the law of total probability via quantum interference effects and model choice-induced preference changes via quantum projection. This is further advanced in [23], and [1] where the mental state is modeled as an open-quantum system. This open-quantum system representation allows for a generalization of the widely used Markovian model of preference evolution, while maintaining these advantages of the quantum framework. Busemeyer et. al. [2] provide empirical analysis which supports the use of open-quantum models and conclude “An open system model that incorporates elements of both classical and quantum dynamics provides the best available single system account of these three characteristics—evolution, oscillation, and choice-induced preference change”.

To prove stability of this human-machine interaction, we generalize a recent result in finite-step Lyapunov stability given in Qin et. al [10], which itself is a generalization of the work in [24] from deterministic to stochastic systems. [10] aims to advance Lyapunov theory for stochastic dynamical systems, in particular for network algorithms for distributed computation, and applies the developments to several distributed computation problems. They show how the results can be directly used to study problems in which one needs to...
prove convergence of inhomogeneous Markov chains, such as in distributed optimization [25] [26], other distributed coordination algorithms [27] [28] [29], and solving linear algebraic equations [30] [31]. Thus our finite-step stochastic Lyapunov stability results can find application in these domains, but we leave this as future work.

**Notation**
1) $|n\rangle \in \mathcal{S}$ is the $n^{th}$ basis vector in a Hilbert space $\mathcal{S}$
2) $\mathcal{H}([\{s_1\}, \ldots, \{s_k\}])$: the Hilbert Space spanned by the orthonormal basis vector set $\{s_1, \ldots, s_k\}$
3) $A^\dagger$: adjoint of $A$
4) $|\langle m|$: outer product of $|n\rangle$ and $|m\rangle$
5) Density operator $\rho$: $\rho = |\Psi\rangle \langle \Psi|$ for some $|\Psi\rangle \in \mathcal{S}$
6) Trace of operator $A$: $\text{Tr}(A) = \sum_{i=1}^{N} |\langle i| A |i\rangle|$ for basis $|i\rangle \in \mathcal{S}$
7) Random events are defined on $(\Omega, \mathcal{F}, \mathbb{P})$
8) $\sigma(x_1, \ldots, x_k)$: sigma algebra generated by $\{x_i\}_{i=1}^k$

**II. Quantum Probability Model for Human Decision Making**

This section presents the open-quantum system model that we will use to represent the decision preference evolution of the human decision maker. We define the evolution of the density operator of the decision maker using the open-system Quantum Lindbladian Equation, proposed in [1] and implemented in [3]. Reference [2] provides empirical evidence which concludes that this open-system structure is the most parsimonious model which can capture observations of dynamical preference evolution such as oscillation and choice-induced preference change. We first illustrate the high-level psychological model in Subsection II-A, the detailed mathematical and psychological derivation of this model is given in Appendix I-B Sec. II-B Sec. II-C and Appendix I-B present the derivation and construction of our Lindbladian psychological model and discuss why this model is useful. The reader may skip these sections and go directly to II-A and then to Sec. III and Sec. IV for the main results.

**A. Lindbladian Preference Evolution Model**

The psychological model consists of the following:
1) A state of nature $s \in \{1, \ldots, n\} = \mathcal{X}$, with probability mass function $\pi_0(s)$
2) A state posterior $\alpha_k$ obtained by a Bayesian sensor from noisy observation $y_k \sim p(y_k|s)$ and prior $\alpha_{k-1}$, $\alpha_k$ parameterizes the human psychological evolution, as quantified in Appendix I-B
3) A state information signal $\beta_k$ which depends on the noisy observation $z_k \sim p(z_k|\alpha_k)$. For example, this could represent the output of a deep learning classifier. This is discussed further in Sec. III-A
4) An action $a_k \in \{1, \ldots, m\} = \mathcal{A}$ by the human at discrete time $k$, for $k = 0, 1, 2, \ldots$
5) An objective state-action utility function $\zeta: \mathcal{X} \times \mathcal{A} \to \mathbb{R}$ (1)

which gives a real-valued utility to each unique state-action pair. We call this the objective utility function because it will be augmented by several psychological parameters to produce a subjective utility function eq. (2), which will influence the human’s action preference evolution, see Appendix I-B

6) Scalar control input $u_k \in [-\bar{u}, \bar{u}], \bar{u} \in \mathbb{R}_+$, generated by the machine every $T$ time steps, where $T$ is a random variable, $T : \Omega \to \mathbb{N}$. The control input $u_k$ is piecewise constant for $T$ time-steps. $T$ has a known probability mass function $\pi_T(\cdot)$. The input $u_k$ controls the parameters of the psychological state evolution (2), and thereby influences the human decision making. $u_k$ models a recommendation signal (for example a posterior probability of the state of nature) by the machine to the human. These must satisfy the constraints in Sec. III-A

7) The discrete time evolution of the psychological state $\rho_{k+T} = M_{a_k,T}^u \rho_k M_{a_k,T}^{u\dagger}$ (2)

where $k = 0, 1, \ldots$ denotes discrete-time. $\alpha_k$ is a $T$-length sequence $\{a_k\}_{j=1}^T$ of random actions $a_k$, taking values in $\mathcal{A}$, and $\rho_k$ is a complex-valued $d \times d$ matrix. We derive $M_{a_k,T}$ in the next subsection as a time discretization of the Lindbladian ordinary differential equation from quantum mechanics.

8) The action probability distribution at time $k$ $\mathbb{P}(a_k) = \text{Tr}(M_{a_k,k}^u \rho_k M_{a_k,k}^{u\dagger}), a_k \in \{1, \ldots, m\}$ (3)

In mathematical psychology, the distribution over actions (3) is called the human’s preference distribution [1]. It can be observed that $\sum_{a_k} \text{Tr}(M_{a_k,k}^u \rho_k M_{a_k,k}^{u\dagger}) = 1$ and $\text{Tr}(M_{a_k,k}^u \rho_k M_{a_k,k}^{u\dagger}) \geq 0 \forall a_k \in \{1, \ldots, m\}$. Equation (2) represents the discrete-time psychological evolution of a human decision maker. It is derived from the time discretization of the quantum mechanical Lindbladian differential equation, which quantifies the evolution of an open-quantum system; that is, a quantum system which interacts with a dissipative external environment. It has been shown in [3] [1] that since this model incorporates aspects of both classical and quantum state evolution, it more accurately describes a wider range of psychological phenomena than classical or quantum models alone.

Note that the induced action choices are inherently probabilistic, with probabilities defined in (3), rather than classically deterministic. The relevance of probabilistic decision making models, in particular quantum models, is discussed in [32]. The human-machine interactions also occur every $T$ time steps, with $T$ a random variable. This is useful because it guarantees convergence of sample paths even for randomly initialized human-machine interaction time scales. This could correspond to uncertainty in the decision/processing speed of a particular human (or between different humans) or of a particular machine. On a practical note, this model incorporates the randomness of human reaction times in decision making.
We prove this convergence with random $T$ in the proof of Theorem 2 and this uses our finite-step Lyapunov result provided as Theorem 3. An interesting technical note is that this proof cannot be extended to the case when $T$ is a nontrivial random process. Our proof induces process subsequences which each converge and exhaust the entire process sequence, therefore implying convergence of the sequence. If $T$ is time-evolving (e.g. a recurrent Markov chain) then the same argument will result in sub-sequences which do not exhaust the original sequence, and therefore overall convergence cannot be shown. However, $T$ can be a random process on the positive integers which either evolves according to a Markov chain with an absorbing state or is non-increasing eventually with probability one. We do not provide proofs for these straightforward generalizations, they may be useful in modeling evolving human-machine interaction time scales.

B. Discrete time representation of Lindblad Equation in terms of Kraus Operators

Our aim here is to derive which is the discrete time version of (4). This yields the Lindbladian operator in terms of Kraus Operators used in Sec. III We will first define the Lindbladian operator and how it evolves. Then in Theorem I we prove a Lindbladian Operator can be reformulated in terms of Kraus Operators as done in [33].

Given a state of nature $s$ which is a random variable that can take on $n$ values, the human decision maker chooses one of $m$ possible actions. The quantum based model for human decision making is governed by the Lindbladian evolution of the psychological state. With $L^u$ denoting the Lindbladian operator, the Lindbladian ordinary differential equation for the dynamics of the psychological state $\rho_t$ over time $t \in [0, \infty)$ is specified as

$$\frac{d\rho_t}{dt} = L^u_{\Phi,\rho_0} \rho_t, \quad \rho_0 = \frac{1}{mn} \text{diag}(1, \ldots, 1)_{mn \times mn}$$

(4)

$$L^u_{\Phi,\rho_0} \rho = -i(1 - \phi_1)|H^u, \rho|$$

$$+ \phi_1 \sum_{k,j} \sum_{\ell(k,j),(\phi_2,\phi_3)} \left( L_{(k,j)} \rho L_{(k,j)}^\dagger - \frac{1}{2} \{L_{(k,j)}^\dagger L_{(k,j)}, \rho \} \right)$$

Here $[A, B] = AB - BA$ is the commutator of $A$ and $B$, $\{A, B\} = AB + BA$ is the anti-commutator of $A$ and $B$, $A^*$ is the complex conjugate of $A$, and $k, j \in \{1, \ldots, mn\}$. Also, $\Phi = (\phi_1, \phi_2, \phi_3) \in [0, 1] \times [0, 1] \times [0, \infty)$ is a vector of free parameters which determine the quantum decision maker’s behavior, as discussed in Appendix I-B. We assume that the machine has full knowledge of these behavioral parameters, as methods for estimating these via training are outside the scope of this paper. The terms $\Phi, L, \gamma$ and $H$ will be defined in Appendix I-B in the context of an appropriate psychological model. Eq. (4) is the canonical form of the Lindbladian operator for open-quantum systems.

The first term $|H^u, \rho|$ is known as the Von-Neumann equation and generalizes Schrödinger’s Equation from the pure state domain to the density operator domain. The second term captures non-reversible Markovian dissipative effects of an interaction with an external environment, making this system “open”. These two terms are weighted by the coefficients $\phi_1$, which interpolates between the purely quantum von Neumann evolution ($\phi_1 = 0$) and the purely dissipative dynamics ($\phi_1 = 1$).

In order to prove convergence of the psychological state $\rho$ under an appropriate control policy, done in Sec. III we need to derive the discrete time representation from the differential equation (4).

Theorem 1: The psychological state update (2) is obtained as the first-order Euler time-discretization of the Lindbladian differential equation (4).

Proof: See Appendix I-A

C. Practicality in Modeling Human Decision Making

The Lindbladian model (4) captures important human decision features such as violation of the sure-thing principle, temporal preference oscillations and order effects, which we now describe. These features cannot be explained by purely Markovian models in a parsimonious way [22].

The violation of the sure-thing principle: The total probability law, also called the Sure Thing Principle (STP), is

$$P(A) = P(B) P(A|B) + P(\bar{B}) P(A|\bar{B})$$

events $A$ and $B$. Violation occurs when $= \bar{B}$ is replaced by either $= B$ or $= \bar{B}$. Pothos and Busemeyer [32] (see also [18]) 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. Here we demonstrate that the Lindbladian model in Appendix I-B can model these violations.

Consider the simple case comprising two actions $a_1, a_2$ and two states $E_1, E_2$. Suppose $\Gamma$ is a non-degenerate posterior belief (strictly in the interior of the unit simplex) of the underlying state. The violation of the sure thing principle occurs when $P(a_2|\Gamma)$ is not a convex combination of $P(a_2|E_1)$ and $P(a_2|E_2)$, i.e.

$$P(a_2|\Gamma) \neq \epsilon P(a_2|E_1) + (1 - \epsilon) P(a_2|E_2) \forall \epsilon \in (0, 1)$$

Equivalently, if $P(a_2|\Gamma) > \max\{P(a_2|E_1), P(a_2|E_2)\}$ or $P(a_2|\Gamma) < \min\{P(a_2|E_1), P(a_2|E_2)\}$ then the sure thing principle has been violated. This latter form is explicitly demonstrated in Fig. 2. The probabilities of $a_2$ $P(a_2|E_1)$ (green surface) $P(a_2|E_2)$ (blue surface), and $P(a_2|\Gamma)$ (red surface) are plotted as functions of free parameters $\phi_1$ and $\phi_2$, with $\phi_2 = 10$ fixed. The regions of the plot where the red surface is outside the space in between the blue and green surfaces correspond to parameterizations under which the sure thing principle has been violated. Fig. 3A shows a cross-section of Fig. 2 at $\phi_1 = 0.29$ and explicitly shows the violation of total probability for $\phi_1 > 0.88$.

Oscillation in preference evolution: Fig. 3 illustrates a sample path of the action preference distribution evolution governed by (3). This sample path in particular demonstrates two essential features of the Lindbladian model:

- Preferences which oscillate over time can be modeled. These oscillations in temporal preference evolution are empirically observed in [2] and cannot be modeled by classical Markovian models.
Quantum decision making facilitates modeling human subjective probabilities that violate the total probability law (sure thing principle). The surface for $P(a_2|\Gamma)$, denoted in red crosses the surfaces of $P(a_2|E_1 \text{ known})$ denoted in blue and $P(a_2|E_2 \text{ known})$ denoted in green. The law is not violated if the red surface lies between the green and the blue surface.

**Fig. 2.** Quantum decision making facilitates modeling human subjective probabilities that violate the total probability law (sure thing principle). The surface for $P(a_2|\Gamma)$, denoted in red crosses the surfaces of $P(a_2|E_1 \text{ known})$ denoted in blue and $P(a_2|E_2 \text{ known})$ denoted in green. The law is not violated if the red surface lies between the green and the blue surface.

- Since the Lindbladian model includes dissipative (Markovian) dynamics, it does not oscillate forever; it converges to a steady-state distribution. This steady-state distribution is commonly used in the literature [3] [1]; Here we use it to generate the probability data in Figs. 2 and 3 which reveals violations of the sure thing principle.

**III. MACHINE CONTROL OF HUMAN DECISION MAKER**

The purpose of this section is to design a stochastic Lyapunov-based controller. By using this controller, our main result (Theorem 2) shows that regardless of the initial psychological state of the human, the machine is able to control the human preference in such a way that the target action is eventually chosen at every time step with probability one. Theorem 2 generalizes the Lyapunov formulation in [9] to demonstrate almost sure convergence when the machine interacts with the human over random time intervals $T$.

We first define some notation: With $d = nm$, let $D$ denote the space of non-negative Hermitian matrices with trace 1:

$$D := \{\rho \in \mathbb{C}^{d \times d} : \rho = \rho^\dagger, \text{Tr}(\rho) = 1, \rho \geq 0\}$$

(5)

Let $\{b_r\}_{r=1}^d$ be a set of orthonormal vectors in $\mathbb{C}^d$ (isomorphic to $\mathcal{H}(A) \otimes \mathcal{H}(X)$), where each $|b_r\rangle$ corresponds to a unique state-action pair. Let $\mathcal{S}$ be the Hilbert space spanned by the orthonormal basis $\{b_r\}_{r=1}^d$. Thus for each $a \in A$ there exists a unique subset $\{b_{r_1}, \ldots, b_{r_a}\} = a \otimes X$, and for each $s \in X$ there exists a unique subset $\{|b_{r_1}\}, \ldots, |b_{r_m}\} = A \otimes s$

**Definition 1:** $|b_r\rangle \in \{b_r\}_{r=1}^d$ is the target state which is chosen by the machine and to which the human is guided via the control input

The difference between target action and target state should be clarified. The target action $a^* \in A$ lies in the action space and corresponds to one particular decision the human can make. The target state $|\bar{b}_r\rangle \in a^* \otimes X$ is one particular state in the Hilbert space $\mathcal{S}$ consisting of both the action space $A$ and the state space $X$. We will show that the psychological state can be guided to any target state, in particular some $|\bar{b}_r\rangle \in a^* \otimes X$, which is equivalent to convergence to the target action $a^*$.

The term ‘Open-loop (super) martingale’ below denotes a (super) martingale when the control input $u_k = 0$ for $k = 0, 1, \ldots$

**A. Machine control**

Recall from Sec. IIA that the machine interacts with the human via a control signal $u_k$. We now introduce additional assumptions on this machine controller:

1) Control Input Constraints

We need the following assumptions which constrain the influence of control $u_k$ on the psychological evolution.

a) For each $u_k$, $\sum_{a_k} M_{a_k}^u a_k = \mathbb{I}$ for actions $a_k$. 

Fig. 4. Time evolution of action probabilities exhibits oscillatory behavior. This behavior cannot be captured by the classical Markovian model, which only exhibits monotonic behavior. Here the parameters are $\phi_1 = 0.5$, $\phi_2 = 10$ and $\phi_3 = 0.5$.

**Fig. 4.** Time evolution of action probabilities exhibits oscillatory behavior. This behavior cannot be captured by the classical Markovian model, which only exhibits monotonic behavior. Here the parameters are $\phi_1 = 0.5$, $\phi_2 = 10$ and $\phi_3 = 0.5$.
b) For $u_k = 0$, the $d \times d$ matrices $M_{a_k}^0$ are diagonal:

\[ M_{a_k}^0 = \sum_{n=1}^{d} c_{a_k, n} |n\rangle \langle n|, c_{a_k, n} \in \mathbb{C} \]

All $M_{a_k}^0$ are $\mathbb{C}^2$ (continuously differentiable) functions of $u_k$

2) The control signal $u_k$ is chosen according to\(^\text{1}\)

\[ u_k := \arg\min_{u \in [-1, 1]} \mathbb{E}[V(\rho_k + T)|\rho_k, u, \beta_k, \zeta] \quad (6) \]

Here $V(\rho)$ will be defined in \(^9\) and is parameterized by the set \(\{\sigma_r\}_{r \in \{1, \ldots, d\}}\) (see Appendix II.C).

**Discussion.** The constraints (a)-(c) are needed as technical assumptions for the quantum control methods of \([9]\), and thus for our proof of Theorem \([2]\) to hold. For our purposes it suffices that $u_k \in [-1, 1]$, see \([9]\) for details. For our particular purposes, we are deriving the structure of $M_{a_k}$ by the psychologically motivated Lindbladian structure given in \([1]\). So in order to utilize the stochastic Lyapunov stability methods of \([9]\), we must adapt the Lindbladian structure (parameterization by $u_k$) such that constraints (a)-(c) are satisfied. Appendix \([\text{IB}]\) discusses how this is accomplished.

Constraint (a) is satisfied since the Lindbladian operator is a Kraus operator (see Sec. \([\text{II-B}]\) Theorem \([1]\)). Constraint (b) implies that Lindbladian operator is a diagonal operator when $u_k = 0$; it is shown by our formulation of the Lindbladian operator which is indeed a diagonal operator when $u_k = 0$, as given in Appendix \([\text{IB}]\). The constrained (c) relates to the continuity of the Lindbladian operator with respect to $u_k$ and is satisfied by our construction given in Appendix \([\text{IB}]\). This continuity property will be used to derive the parameters of the Lyapunov function in Appendix II.C.

We note that the choice of target state (Definition \([1]\)) is arbitrary in the context of our main result, i.e. any particular target state can be chosen by the machine. For instance, we could have the machine determine the target state $|\tilde{b}_r\rangle \in a^* \otimes \mathcal{X}$, with corresponding action $a^*$ as that which optimizes the Bayesian expected objective utility:

\[ a^* = \arg\max_{a \in \{1, \ldots, m\}} \mathbb{E}[\zeta(a, \mathcal{E})|\beta_k] \quad (7) \]

where $\zeta$ is the objective state-action utility function. As the human has a utility which is augmented by subjective biases (see Appendix \([\text{IB}]\) and a psychological evolution which is inherently probabilistic, \([\text{7}]\) would essentially improve the human’s decision making such that it is Bayesian optimal. However, note that this particular choice of which target action to guide the human to does not factor into the proof of our result that the human can be guided to any chosen target action eventually.

**B. Main Result 1. Convergence of Psychological State**

Here we prove the convergence of the controlled psychological state to a desired target action.

**Theorem 2:** Consider the discrete time density operator evolution \((2)\) and any target state $|\tilde{b}_r\rangle$ (see Definition \([1]\)). Then, there exists a control sequence \(\{u_k\}_{k \in \mathbb{N}}\) generated by the machine such that the human psychological state $\rho_k$

\[ \text{converges to} \ |\tilde{b}_r\rangle \langle \tilde{b}_r| \text{ with probability one for any initial} \ \rho_0 \in D, \text{where} \ D \text{ is defined in} \ (5). \]

**Remark:** The convergence to any psychological state $|\tilde{b}_r\rangle \langle \tilde{b}_r| \in a \otimes \mathcal{X}$, implies the convergence to the corresponding action $a \in A$. Since $\forall a \in A$ there exists a set $\{ |b_1\rangle, \ldots, |b_n\rangle \} = a \otimes \mathcal{X}$ (see discussion directly above Definition \([1]\)). Theorem \([2]\) implies that the machine can guide the human to choose any specific target action $a \in A$ asymptotically.

**Proof:** First with $\eta_k = \{a_k, T\}$, rewrite \((2)\) as

\[ \rho_{k+T} = M_{\eta_k}^{u_k} \rho_k = \frac{M_{\eta_k}^{u_k} \rho_k M_{\eta_k}^{u_k} \dagger}{\text{Tr}(M_{\eta_k}^{u_k} \rho_k M_{\eta_k}^{u_k} \dagger)} \quad (8) \]

We define the following Lyapunov function, which forms a supermartingale under both open-loop (zero-input) and closed-loop (feedback control $\rho_k$) conditions for the process \((8)\):

\[ V(\rho) := \sum_{r=1}^{d} \sigma_r |\tilde{b}_r\rangle \langle \tilde{b}_r| - \frac{\epsilon}{2} \sum_{r=1}^{d} |\tilde{b}_r\rangle \langle \tilde{b}_r|^2 \quad (9) \]

where $\sigma_r$ is non-negative $\forall r \in \{1, \ldots, d\}$ and $\epsilon$ is strictly positive. The set \(\{\sigma_r\}_{r=1}^{d}\) and $\epsilon$ are chosen according to \([9]\) such that the Lyapunov function $V(\rho_k) > 0$ for $\rho_k \in D$ and $\rho_k$ converges to the intended subspace $|\tilde{b}_r\rangle \langle \tilde{b}_r|$ with probability 1. By \([9]\) and \([34]\), $|\tilde{b}_r\rangle \langle \tilde{b}_r|$ is an open-loop martingale given the $T$-step density operator evolution \((3)\) (see Appendix \([\text{II-D}]\) for proof). $V(\rho)$ is a concave function of the open-loop martingales $|\tilde{b}_r\rangle \langle \tilde{b}_r|$ and therefore is an open-loop ($u_k = 0$) supermartingale given the process \((8)\):

\[ \mathbb{E}[V(\rho_{k+T})|\rho_k, u_k = 0] - V(\rho_k) \leq 0 \]

The following feedback control mechanism is used

\[ u_k := \arg\min_{u \in [-1, 1]} \mathbb{E}[V(\rho_{k+T})|\rho_k, u] \quad (10) \]

to get $\mathbb{E}[V(\rho_{k+T})|\rho_k, u_k] \leq \mathbb{E}[V(\rho_{k+T})|\rho_k, u_k = 0]$. Here the expectation is taken with respect to $\eta_k$,

\[ \mathbb{E}[V(\rho_{k+T})|\rho_k, u] = \mathbb{E}[V(\mathcal{M}_{\eta_k}^u \rho_k)] = \int_{\Omega} V(\mathcal{M}_{\rho_k(\omega)}^u \rho_k) d\omega \]

where $\Omega$ is the sample space under which the process is induced. Define

\[ \tilde{Q}(\rho_k) := \mathbb{E}[V(\rho_{k+T})|\rho_k, u_k] - V(\rho_k) \leq \mathbb{E}[V(\rho_{k+T})|\rho_k, u_k = 0] - V(\rho_k) \leq 0 \]

$V(\rho)$ is a continuous function, so using \([35, \text{Chapter 8}]\), we have the $T$-step control sequence $\{\rho_{k+T}\}_{k \in \mathbb{N}}$ that converges to the set $D_\infty := \{\rho : \tilde{Q}(\rho) = 0\}$ with probability one.

By Lemma \([1]\) presented immediately after this proof, the set $D_\infty$ is restricted to our desired state $\{ |\tilde{b}_r\rangle \langle \tilde{b}_r| \}$. So now we must show that the entire sequence $\{\rho_k\}_{k \in \mathbb{N}}$ converges to this set.

We now show that $\mathbb{P}[\lim_{k \to \infty} \rho_k = |\tilde{b}_r\rangle \langle \tilde{b}_r|] = 1$, i.e. the entire sequence converges to the target state with probability one. This utilizes Theorem \([3]\) which is developed in Sec. \([\text{IV}]\).
The statement of Lemma 1 can be rewritten as
\[
\mathbb{P}[\lim_{i \to \infty} \rho_{k+iT} = [b_r] \langle b_r \rangle] = 1
\]
where the choice of \( \{\sigma_r\}_{r=1}^d \) from Lemma 2 of [9] and our proof of Lemma 1 ensures that
\[
\mathbb{E}[V_r(\rho_{k+iT}) | \rho_k = [b_r] \langle b_r \rangle, u_k = u]
\]
has a strict local minimum at \( u = 0 \) if and only if
\[
\begin{align*}
\mathbb{E}[V_r(\rho_{k+iT}) | \rho_k = [b_r] \langle b_r \rangle, u_k = u] &= \mathbb{E}[V_r(M^0_{\eta_k} \rho_{k+iT})] \\
&\leq V_r(\rho_{k+iT})
\end{align*}
\]
Now, for \( q \in \{1, \ldots, T\}, \)
\[
\mathbb{E}[V_r(\rho_{k+(i+q)T}) | \rho_k = [b_r] \langle b_r \rangle, u = 0] = \mathbb{E}[V_r(M^0_{\eta_k} \rho_{k+(i+q)T})] \leq V_r(\rho_{k+(i+q)T})
\]
since \( V_r \) is an open-loop supermartingale. This implies \( \exists K \in \mathbb{N} \) such that
\[
\mathbb{E}[V_r(\rho_{k+iT}) | \rho_k, u_k] \leq V_r(\rho_k) \forall k \geq K
\]
Since there exists a unique mapping from elements of \( D \) to elements of \( \mathbb{R}^{d^2} \), applying Theorem 3 yields \( \mathbb{P}[\lim_{k \to \infty} \rho_k = [b_r] \langle b_r \rangle] = 1 \), thus proving the Theorem.

**Lemma 1:** For the evolution \( \hat{\mathbb{P}} \), target state \([b_r] \langle b_r \rangle\) and feedback control \( \hat{\mathbb{P}} \), we have \( \mathbb{P}[\lim_{k \to \infty} \rho_{k+iT} = [b_r] \langle b_r \rangle] = 1 \)

**Proof:** The proof follows immediately from Lemma 2 of [9]. For any target subspace \([b_r] \langle b_r \rangle\), the set \( \{\sigma_r\}_{r=1}^d \) can be chosen in such a way that \( D_{\infty} = [b_r] \langle b_r \rangle\). The idea is the following: A state \( \rho_k \) is in the limit set \( D_{\infty} \) iff for all \( u \in [-1, 1] \),
\[
\mathbb{E}[V_r(\rho_{k+iT}) | \rho_k, u] - V_r(\rho_k) \geq 0
\]
Also, since \( V_r \) is an open-loop supermartingale, \( \forall \rho_k \in D \):
\[
\mathbb{E}[V_r(\rho_{k+iT}) | \rho_k, u = 0] - V_r(\rho_k) \leq 0
\]
By Lemma 2 of [9], presented in Appendix C as Lemma 3, we have
\[
\mathbb{E}[V_r(\rho_{k+iT}) | \rho_k = [b_r] \langle b_r \rangle, u_k = u] \leq V_r(\rho_k) \langle b_r \rangle]
\]
Therefore, by (11), we have that \([b_r] \langle b_r \rangle\) is in the limit set \( l_{\infty} \) if and only if \( r = \bar{r} \).

To summarize, we showed that for the discrete time psychological state evolution of (2), there exists a control policy which allows the machine to guide the human decisions such that a target decision is made eventually, almost surely.

**C. Numerical Example. Controlling Human Decision Evolution**

We now provide numerical examples which illustrate the time evolution of the action distribution (12) and target action convergence. In each case we initialize the psychological state \( \rho_0 \) as a uniform (non-informative) distribution over the action space. The psychological state evolution is given by (2), and the objective state-action utility \( \zeta(a_{\hat{S}}, E_{\hat{S}}) \) in (12) is initialized to have uniformly distributed random values between 0 and 10.

1) Example 1: Suppose there are two underlying states of nature \( E_1 \) and \( E_2 \), and four actions \( a_1, a_2, a_3, a_4 \). The target action is \( a_4 \). We use parameters \( \phi_1 = 0.8 \) and \( \phi_2 = 10 \). We simulated the preference evolution according to the protocol (1)-(8) of Sec. II-A and (1)-(2) of Sec. II-A. The action preference distribution is plotted as a function of time, with actions and controls taken every \( T = 10 \) time points. A typical sample path is shown in Fig. 5.

![Fig. 5](image-url) Simulation illustrating how the machine controls the human psychological state to converge to action 4. In this simulation the human decision maker’s action probabilities converge after \( T = 60 \) time steps.

**IV. Finite Step Stochastic Lyapunov Stability**

Recall in Sec. III our stability proof of the human-machine decision system (Theorem 2) employed a finite-step stochastic
Lyapunov argument, the proof of which was postponed to this section. The purpose of this section is two-fold: to generalize an existing finite-step stochastic Lyapunov result given in [10], and to use a special case of this generalization as this argument which enables the proof of our main result, Theorem 2.

This section presents three theorems:

- Theorem 4 is equivalent to Theorem 1 of [10], but we provide an alternative proof which allows for direct generalization.
- Theorem 5 generalizes Theorem 4 to the case when the finite-step size T is a random variable.
- Theorem 6 follows as a special case of Theorem 5 which is directly applicable to our proof of Theorem 2.

Consider the discrete time stochastic system described by

$$x_{k+1} = f(x_k, y_{k+1}), \ k = 0, 1, 2, \ldots$$  \hspace{1cm} (13)

Here $x_k \in \mathbb{R}^n$, and $\{y_k : k = 0, 1, 2, \ldots\}$ is a $\mathbb{R}^d$ valued stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Consider the filtration (increasing sequence of $\sigma$-fields) defined by

$$\mathcal{F}_0 = \{0, \Omega\}, \quad \mathcal{F}_k = \sigma(y_1, \ldots, y_k) \text{ for } k \geq 1.$$  

We choose $x_0 \in \mathbb{R}^n$ as a constant with probability one. Thus $\{x_n\}$ is an $\mathbb{R}^n$-valued stochastic process adapted to $\mathcal{F}_k$.

**Theorem 3:** For the discrete-time stochastic system (13), let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuous non-negative and radially unbounded function. Suppose there exists a random variable $\Phi : \Omega \rightarrow \mathbb{N}$ such that for any $k$,

$$\mathbb{E}[V(x_{k+T}) | \mathcal{F}_k] - V(x_k) \leq -\varphi(x_k)$$  \hspace{1cm} (14)

where $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous and satisfies $\varphi(x) \geq 0 \ \forall x \in \mathbb{R}^n$.

Then for any initial condition $x_0 \in \mathbb{R}^n$, $x_k$ converges to $\mathcal{D}_1 := \{x \in \mathbb{R}^n : \varphi(x) = 0\}$ with probability one.

**Proof:** The proof follows from Theorems 4 and 5 below.

**Theorem 4:** For the discrete-time stochastic system (13), let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuous non-negative and radially unbounded function. Define the set $\mathcal{Q}_\lambda := \{x : V(x) < \lambda\}$ for some positive $\lambda$, and assume that:

(a) $\mathbb{E}[V(x_{k+1}) | \mathcal{F}_k] - V(x_k) \leq 0$ for any $k$ such that $x_k \in \mathcal{Q}_\lambda$.

(b) There exists an integer $T \geq 1$, independent of $\omega \in \Omega$, such that for any $k$, $\mathbb{E}[V(x_{k+T}) | \mathcal{F}_k] - V(x_k) \leq -\varphi(x_k)$, where $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous and satisfies $\varphi(x) \geq 0$ for any $x \in \mathcal{Q}_\lambda$.

Then for any initial condition $x_0 \in \mathcal{Q}_\lambda$, $x_k$ converges to $\mathcal{D}_1 := \{x \in \mathcal{Q}_\lambda : \varphi(x) = 0\}$ with probability at least $1 - V(x_0)/\lambda$.

**Proof:** From assumption (b),

$$\mathbb{E}[V(x_{k+T}) | \mathcal{F}_k] - V(x_k) \leq -\varphi(x_k) \leq 0, \ \forall x_k \in \mathcal{Q}_\lambda$$  \hspace{1cm} (15)

where $\varphi(x)$ is continuous. Now, by assumption (a) and [35, p.196], if we start with $x_0 \in \mathcal{Q}_\lambda$, then the probability of staying in $\mathcal{Q}_\lambda$ during the entire resultant trajectory is at least $1 - V(x_0)/\lambda$, i.e.

$$\mathbb{P}[\sup_{k \in \mathbb{N}} V(x_k) \geq \lambda] \leq V(x_0)/\lambda$$  \hspace{1cm} (16)

Next construct $T$ subsequences of $\{X_k\}$ as follows:

$$\{X_i^{(0)}\} = \{X_0, X_T, \ldots\}, \quad \{X_i^{(1)}\} = \{X_1, X_{T+1}, \ldots\}, \ldots$$

$$\{X_i^{(T-1)}\} = \{X_{T-1}, X_{2T-1}, \ldots\}$$

Suppose $\varphi(x) \geq 0 \ \forall x$; Then for all $k \in K := \{0, \ldots, T-1\}$, $\{V(X_i^{(k)})\}$ is a non-negative supermartingale process by (15), and thus by Doob’s convergence theorem converges to a limit with probability 1. From (15) we have for all $k \in K$ and $n \in \mathbb{N}$

$$\sum_{l=1}^{n} \mathbb{E}(V(X_i^{(k)})) - \mathbb{E}(V(X_i^{(k-1)})) \leq -\mathbb{E}\left(\sum_{l=0}^{n-1} \varphi(X_i^{(k)})\right)$$

and thus

$$0 \leq \mathbb{E}(V(X_i^{(k)})) \leq \mathbb{E}(V(X_i^{(0)})) - \mathbb{E}\left(\sum_{l=0}^{n-1} \varphi(X_i^{(k)})\right)$$

We use Fatou’s Lemma to obtain

$$\mathbb{E}\left(\sum_{l=0}^{\infty} \varphi(X_i^{(k)})\right) < \infty$$

and by the Borel-Cantelli lemma we have

$$\mathbb{P}[\lim_{n \rightarrow \infty} \varphi(X_i^{(k)}) = 0] = 1 \ \forall k \in K$$

Now suppose $\varphi(x) \geq 0$ only for $x \in \mathcal{Q}_\lambda$. Stop $\{X_i^{(k)}\}$ on first leaving $\mathcal{Q}_\lambda$. Then for $x \notin \mathcal{Q}_\lambda$, $\varphi(x) = 0$ for this stopped set. This stopped process is a supermartingale and the proof is the same as above.

It is now apparent that

$$\lim_{n \rightarrow \infty} \varphi(X_i^{(k)}(\omega)) = 0$$  \hspace{1cm} (17)
\( \forall k \in K \omega \in \hat{\Omega} = \{ \omega \in \Omega : x_n(\omega) \in Q_\lambda \ \forall n \in \mathbb{N} \} \) so we have
\[
P[ \lim_{n \to \infty} \varphi(X_n(\omega)) = 0 ] \geq 1 - V(x_0)/\lambda
\]
by the analysis in Appendix 1B and 1C.

We now present a generalization of Theorem 4 for the case when the finite-step length \( T \) is a random variable.

**Theorem 5:** For the discrete-time stochastic system (13), let \( V : \mathbb{R}^n \to \mathbb{R} \) be a continuous non-negative and radially unbounded function. Define the set \( Q_\lambda = \{ x : V(x) < \lambda \} \) for some positive \( \lambda \), and assume that:

(a) \( E[V(x_{k+1})|F_k] - V(x_k) \leq 0 \) for any \( k \) such that \( x_k \in Q_\lambda \)
(b) There exists a random variable \( T : \Omega \to \mathbb{N} \), such that for any \( k \),
\[
E[V(x_{k+T})|F_k] - V(x_k) \leq -\varphi(x_k)
\]
where \( \varphi : \mathbb{R}^n \to \mathbb{R} \) is continuous and satisfies \( \varphi(x) \geq 0 \) for any \( x \in Q_\lambda \)

Then for any initial condition \( x_0 \in Q_\lambda \), \( x_k \) converges to \( D_1 := \{ x \in Q_\lambda : \varphi(x) = 0 \} \) with probability at least \( 1 - V(x_0)/\lambda \).

**Remark:** Note that Theorem 5 is equivalent to Theorem 4 when the condition that \( T \) is independent of \( \omega \in \hat{\Omega} \) is removed.

**Proof:** Let \( \{ X_t \}_{t \in \mathbb{N}} \) be the original random process corresponding to the stochastic system (13) initialized with \( x_0 \in Q_\lambda \). Construct the random process \( \{ Y_t \}_{t \in \mathbb{N}} \) by inducing the same stopped process on \( \{ X_t \} \) as in the proof of Theorem 4. This simplifies analysis by omitting sample paths which leave the set \( Q_\lambda \). Thus we consider only the sample paths which remain in \( Q_\lambda \) for the entire trajectory and we will impose a bound on the measure of this set below.

Now define a new random process \( \{ Y_{k_0}^k \}_{k \in \mathbb{N}} \) as follows:

Starting with \( Y_{k_0}, k_0 \in \mathbb{N} \), define \( \forall \omega \in \Omega : Y_{k_1}^k(\omega) = Y_{k_0 + T(\omega)}(\omega), Y_{k_2}^k(\omega) = Y_{k_1 + T(\omega)}(\omega), \ldots \)
\[
Y_{k_n}^k = Y_{k_0 + nT(\omega)}(\omega)
\]
i.e. take a subsequence (over multiples of \( T(\omega) \)) of the original sequence initialized at \( Y_{k_0}, k_0 \in \mathbb{N} \). Then by (b) we have:
\[
E[V(Y_{k+1}^k)|F_k] - V(Y_k^k) \leq -\varphi(Y_k^k) \leq 0,
\]
\[
\forall Y_k^k \in Q_\lambda, \ i \in \mathbb{N} \text{ (18)}
\]
where \( F_k = \sigma(\{ Y_n^k \}_{n=0}^k) \). Thus by the same method as in the proof of Theorem 4 we can obtain
\[
0 \leq E(V(Y_{k+1}^k)) \leq E(V(Y_k^k)) - E(\sum_{l=0}^{i} \varphi(Y_{k_l}^k))
\]
so by Borel-Cantelli:
\[
P[ \lim_{i \to \infty} \varphi(Y_{k_i}^k) = 0 ] = 1 \ \forall k_0 \in \mathbb{N} \quad (19)
\]
where recall \( \{ Y_{k_i}^k \}_{i \in \mathbb{N}} \) is the sequence induced by the initial condition \( Y_{k_0} \). Let \( \Omega_{k_0} = \{ \omega : \lim_{i \to \infty} \varphi(Y_{k_i}^k) = 0 \} \). We know from (19) that \( P(\Omega_{k_0}) = 1 \ \forall k_0 \in \mathbb{N} \) and thus
\[
P(\{ \omega : \lim_{i \to \infty} \varphi(Y_{k_i}^k) = 0 \ \forall k_0 \in \mathbb{N} \}) = P( \bigcap_{k_0 \in \mathbb{N}} \Omega_{k_0} ) = 1
\]
\[3\text{This procedure is done implicitly in the proof of Theorem 4.} \]

Define
\[
\Omega = \{ \omega : \lim_{i \to \infty} \varphi(Y_{k_i}^k(\omega)) = 0 \ \forall k_0 \in \{ 1, \ldots, T(\omega) \} \}
\]
Obviously \( \bigcap_{k_0 \in \mathbb{N}} \Omega_{k_0} \subseteq \hat{\Omega} \) and so \( P(\hat{\Omega}) = 1 \), and \( \forall \omega \in \hat{\Omega} \) the set of sub-sequences \( \{ Y_{k_i}^k(\omega) \}_{i=1, k_0=1, \ldots, T(\omega)} \) exhaust the entire sequence \( \{ Y_i(\omega) \}_{i \in \mathbb{N}} \) and thus \( \lim_{i \to \infty} \varphi(Y_i(\omega)) = 0 \ \forall \omega \in \hat{\Omega} \). Denote \( \Omega^* = \{ \omega : \lim_{i \to \infty} \varphi(Y_i(\omega)) = 0 \} \supseteq \hat{\Omega} \).

So \( P(\Omega^*) = 1 \), and observe that
\[
\Omega^* \cap \hat{\Omega} \subseteq \{ \omega : \lim_{i \to \infty} \varphi(Y_i(\omega)) = 0 \}
\]
so
\[
P[ \lim_{n \to \infty} \varphi(Y_n(\omega)) = 0 ] \geq P[\Omega^* \cap \Omega] \geq 1 - V(x_0)/\lambda
\]
where \( \Omega \) is the same as defined in the proof of Theorem 4 and \( P(\hat{\Omega}) \geq 1 - V(x_0)/\lambda \) by assumption (a) and (35, p.196).

**Theorem 4** is an alternative proof of Theorem 1 in [10] and **Theorem 5** builds upon it to generalize \( T \) to a random variable. **Theorem 5** follows from **Theorem 5** in the following way: Start with the assumptions of **Theorem 5**. Now let \( \phi(x) \geq 0 \ \forall x \in \mathbb{R}^n \). Then \( Q_\lambda \in \mathbb{R}^n \). Now since \( V \) is radially unbounded, the only such \( \lambda \) for which this can hold is \( \lambda = \infty \). Then as a result of **Theorem 5** we have: for any initial \( x_0 \in Q_\lambda = \mathbb{R}^n \), \( x_k \) converges to \( D_1 \) with probability at least \( 1 - V(x_0)/\lambda = 1 \), since \( \lambda = \infty \).

To summarize, this section provided a generalization of the finite-step Lyapunov function result of Qin. et. al. [10], by making \( T \) a random variable. We applied this to show the stability of the Lindbladian dynamics to prove almost sure convergence of the density operator (psychological state), but the generalization is of independent interest.

**V. CONCLUSION AND EXTENSIONS**

Quantum decision models have emerged as a novel paradigm in psychology that can parsimoniously model special aspects of human decision making (such as order effects and sure thing principle) which classical Markov models cannot.

In this paper, we formulated a human-machine interaction system as a **controlled** quantum decision making process. We constructed a machine control procedure based on stochastic Lyapunov stability which allows the psychological state of a human to be guided to any arbitrary target state eventually almost surely. This used a random finite-step stochastic Lyapunov function result, which is also of independent interest.

The Lindbladian quantum models that we used, capture subjective biases and violations of rationality which cannot be captured by classical models; If a human’s decision making is suboptimal due to these subjective irrational biases, then the machine can compensate and guide the human to be an optimal Bayesian expected utility maximizer (with respect to the machine’s state measurement).

We made several simplifying assumptions that warrant further investigation. We assumed that the machine knows the parameterization of the human’s Lindbladian operator for the psychological state. It would be interesting to see what the analysis yields when these are only estimates with...
some distribution. We considered the case when there is one specified action for the human to be guided to. Generalization to classes of actions and associated rate of convergence to the target actions are interesting extensions.

APPENDIX I
QUANTUM DECISION MODEL. ADDITIONAL DETAILS

This Appendix discusses additional details of the Quantum Decision model used in the paper. We draw heavily from [1].

A. Open-Quantum System as a Psychological Model

The derivation of the psychological evolution [4] uses a quantum random walk (QRW) over a state-action network, see [1] for details. The purpose of this subsection is to discuss this QRW formulation of [4]. Each node of the network represents a unique state-action pair. So, in our case there are a total $nm$ nodes. The network is decomposed into $m$ sub-graphs, representing underlying states. Each sub-graph has $n$ nodes and each node within this sub-graph has a self loop and is also connected to all the other nodes in the sub-graph bidirectionally, representing the action-action transition with in an underlying state. These edges are weighed according to objective state-action utility function $b$. These $m$ sub-graphs are connected via same actions across all underlying states. An example of this structure can be seen in Fig. 7 for the case of two underlying states and three actions $a_1, a_2, a_3$. This network structure allows for the simultaneous comparison of utilities and elicitation of beliefs in the underlying state.

![State-Action network](image)

Fig. 7. State-Action network over which the Quantum Random Walk takes place. $\mathcal{E}_j$'s are the underlying states and $a_i$'s are the actions.

B. Stochastic Quantum Random Walk. Lindbladian Dynamics

We now relate the QRW in Appendix I-A to the Lindbladian Dynamics in [4]. This forms the basis of the decision making process [2] that we control.

1) The state space $X = \{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_N\}$. The Hilbert space over $X$ is defined as $\mathcal{H}_X := \mathcal{H}(\mathcal{E}_1, \ldots, \mathcal{E}_N)$. This space contains $|\mathcal{E}_i\rangle$, $n$-dimensional complex vectors indexed by the state $i$.

2) The action space $A = \{a_1, a_2, \ldots, a_m\}$. The Hilbert space over $A$ is defined as $\mathcal{H}_A := \mathcal{H}(a_1, \ldots, a_m)$. This space contains $|a_j\rangle$, $m$-dimensional complex vector indexed by the action $j$.

3) The psychological state space $\mathcal{S} := \mathcal{H}_X \otimes \mathcal{H}_A$. This space contains all action-state pairs. Each tuple $|\mathcal{E}_i, a_j\rangle$ represents each node in the graph given in Sec. I-A.

The psychological state $\rho$ is represented as a density operator in the Hilbert Space $\mathcal{S}$, and its evolution is given by [4]. We also define non-negative functions $f(u_k)$ and $g(u_k)$ to be continuously differentiable and equal to zero for $u_k = 0$.

1) The Hamiltonian $H^u$ is defined as the block diagonal $nm \times nm$ symmetric matrix

$$H^u := \begin{bmatrix} \mathbb{H} & \frac{1}{2}f(u)\mathbb{I} & \cdots & \frac{1}{2}f(u)\mathbb{I} \\ \frac{1}{2}f(u)\mathbb{I} & \mathbb{H} & \cdots & \frac{1}{2}f(u)\mathbb{I} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}f(u)\mathbb{I} & \frac{1}{2}f(u)\mathbb{I} & \cdots & \frac{1}{2}f(u)\mathbb{I} \end{bmatrix}$$

with $z = (m + n - 3)f(u) + 1$. $\phi_1$ determines a convex combination of the quantum and Markovian terms in [4]. A larger value of $\phi_1$ will result in more dissipative (less oscillatory) dynamics [1].

1) $L_{k,j} = \langle k | j \rangle$ represents the walk from $k^{th}$ action-state tuple to $j^{th}$ action-state tuple.

2) We define $\gamma^u_{(k,j)}$ as the $(k,j)^{th}$ element in the Cognitive Matrix

$$\gamma^u_{(k,j)} := [C(\phi_2, \phi_3, u)]_{k,j}$$

$$= [(1 - \phi_3)\Pi^T(\phi_2, u) + \phi_3 B^T]_{k,j}$$

where $\Pi(\phi_2, u) := \text{diag}(A_1, A_2, \cdots, A_N)$

Each $A_j$ is the $m \times m$ matrix

$$[A_j]_{(k,l)} = \begin{cases} \frac{\theta_{(k,j)}(u)}{z_{(k,j)}} & k \neq l \\ \frac{\theta_{(k,j)}(1 - g(u))}{\theta_{(k,j)}(1 - g(u))} & k = l \end{cases}$$

Here the agent’s subjective utility of action $a_i$ given state $\mathcal{E}_j$ is defined as

$$\theta_{i,j}(\phi_2) = \frac{\zeta(a_i, \mathcal{E}_j)^{\phi_2}}{\sum_i \zeta(a_i, \mathcal{E}_j)^{\phi_2}}$$

where $\zeta(a_i, \mathcal{E}_j)$ is the objective utility [1].

$$B := [\alpha_k(\mathcal{E}_1) \quad \alpha_k(\mathcal{E}_2) \quad \cdots \quad \alpha_k(\mathcal{E}_n)] \otimes I_{m \times 1} \otimes I_{m \times m}$$

Recall $\alpha_k$ is the posterior belief of the underlying state.
Discussion of (20), (21), (22).

In (22), \( \phi_2 \in [0, \infty) \) models the agent’s ability to discriminate between the profitability of different actions, in terms of the objective utility \( L \). As seen in (22), when \( \phi_2 = 0 \) all actions \( a_j \in A \) have the same probability of being chosen. When \( \phi_2 \rightarrow \infty \), only the dominant action is chosen.

In (21), \( \phi_3 \in (0, 1) \) interpolates between the \( \Pi \) and \( B \) matrices and therefore models how much the human distinguishes between the underlying states \( \{ \varepsilon_1, \ldots, \varepsilon_n \} \). In our framework, \( \phi_3 \) is a function of the control input and is given by \( \phi_3 = u^{M^i}, \quad i = 2, 3, \ldots \) to satisfy the input constraints in Sec. III-A.

\( \Pi(\phi_2, u) \), in (21), acts as a transition probability matrix between nodes within individual connected components (each state grouping) of the State-Action network, as shown as the blue and red sections of the two-state example network Fig. 2. The matrix \( B \) is also a stochastic matrix and represents the transitions between actions. Without loss of generality, we can assume that an external sensor makes measurements of the state, computes the Bayesian posterior, and provides this to the human.

The cognitive matrix \( C(\phi_2, \phi_3, u) \) in (22) and Hamiltonian \( H^u \) (20) act as transition matrices on the State-Action network. Taken together in the structure of (4), these induce a quantum random walk on this network which represents the psychological evolution of state-action preferences. There can be a wide variety of observed preference evolution within the two-dimensional parameter space defined by \( (\phi_1, \phi_2) \). We can exploit the results from [36] to estimate these parameters.

APPENDIX II
ADDITIONAL PROOFS

A. Proof of Theorem II

Proof: A first order Euler discretization of (4) yields:

\[
\rho(t + \Delta t) = \begin{bmatrix} 1 - T \Delta t (i (1 - \phi_1) H^u \\
+ \phi_1 \frac{1}{2} \sum_{m,n} \gamma_m^u \nu_{m,n} L_{m,n}^\dagger \right] \rho(t) \\
[1 - T \Delta t (-i (1 - \phi_1) H^u \\
+ \phi_1 \frac{1}{2} \sum_{m,n} \gamma_m^u \nu_{m,n} L_{m,n}^\dagger \right] \\
+ \phi_1 \Delta t \sum_{m,n} \gamma_m^u \nu_{m,n} L_{m,n} \rho(t) L_{m,n}^\dagger \\
\end{bmatrix}
\]

(23)

where \( \Delta t \) is the interval of discretization and is sufficiently small so that the bracketed terms of (23) remain positive. Denoting \( \rho_{k+1} \) as \( \rho(t + \Delta t) \) and \( \rho_{k+T} \) as \( \rho(t + \Delta tT) \):

\[
\rho_{k+T} = \sum_{\nu} K_{\nu,T}^{u_k} \rho_k K_{\nu,T}^{u_k\dagger}
\]

(24)

where

\[
K_{0,T}^{u_k} = [1 - T \Delta t (i (1 - \phi_1) H^{u_k} \\
+ \phi_1 \frac{1}{2} \sum_{m,n} \gamma_m^{u_k} \nu_{m,n} L_{m,n}^\dagger \right] \\
K_{\nu \neq 0,T}^{u_k} = \sqrt{T \Delta t \phi_1 \gamma_m^{u_k} \nu_{m,n} L_{m,n} \}, \quad T \in \mathbb{N}
\]

Eq. (24) represents the discrete time psychological preference evolution. This representation is used to obtain the form (2), which allows us to prove convergence of the psychological state in Sec. III.

When the human chooses action \( a \), the psychological state is projected into the respective action subspace of the Hilbert Space \( a \otimes \mathcal{X} \), defined in III-A. The psychological state evolves as:

\[
\rho_{k+T} = \frac{P_{a_k} \left( \sum_{\nu} \nu K_{\nu,T} \rho \nu K_{\nu,T}^\dagger \right) P_{a_k}^\dagger}{\text{Tr}(P_{a_k} \left( \sum_{\nu} \nu K_{\nu,T} \rho \nu K_{\nu,T}^\dagger \right) P_{a_k})}
\]

where \( a_k \) is a random variable taking values in \( \{ 1, \ldots, m \} \), with probability given by

\[
\mathbb{P}(a_k) = \text{Tr}(P_{a_k} \left( \sum_{\nu} \nu K_{\nu,T} \rho \nu K_{\nu,T}^\dagger \right) P_{a_k}^\dagger)
\]

from (3). \( P_{a_k} \) represents the projection into the subspace corresponding to the action taken, i.e. the action acts as a quantum measurement on the psychological state. Now letting

\[
M_{a_k}^{u_k} = P_{a_k} \left( \sum_{\nu} K_{\nu,T} \rho \nu K_{\nu,T}^\dagger \right)
\]

we obtain (2).

B. Convergence of constructed subsequences implies convergence of sequence

From (17) we have \( \lim_{n \rightarrow \infty} \varphi(X_n^{(k)}(\omega)) = 0 \, \forall \omega \in \Omega \). Let \( \omega \in \Omega \) and \( \varphi_n^{(k)} \) denote \( \varphi(X_n^{(k)}(\omega)) \) and \( \varphi_n \) denote \( \varphi(X_n(\omega)) \).

We have:\( \forall \epsilon > 0 \exists N_k \) such that \( \varphi_n^{(k)} < \epsilon \, \forall n > N_k \). Take \( N^* = \max_{k \in \{0, \ldots, T-1\}} N_k \). Suppose \( \lim_{n \rightarrow \infty} \varphi_n \neq 0 \); Then \( \exists \epsilon > 0 \) such that \( \forall N \in \mathbb{N} \exists n_0 > N \) with \( \varphi_n > \epsilon \). Since the subsequences are exhaustive, i.e. for any \( \varphi_n \exists k, m \), such that \( \varphi_n = \varphi_m^{(k)} \), we know that for any \( \epsilon > 0 \), \( \varphi_n < \epsilon \, \forall n > N^* \) so such a \( n_0 \) does not exists for \( N^* \) and thus by contradiction we have \( \lim_{n \rightarrow \infty} \varphi_n = 0 \).

C. Construction of \{ \sigma_r \}_{r=1}^d

This subsection explains how we choose the parameters \{ \sigma_r \}_{r=1}^d \) for the Lyapunov function \( V_r(\cdot) \) in (9). The following two Lemmas are provided in [9]; for completeness we list them here. [9] uses these Lemmas to show that the limit set of the process of interest contains only the target state. We provide intuition for this idea in the proof of Lemma II (Section III-B).

Lemma II (9]): Consider the \( d \times d \) matrix \( R \) defined by

\[
R_{n_1,n_2} = \sum_{a_k} \frac{2}{n_1} \left| \langle n_1 | \frac{d M_{a_k}^{u_k}}{du} | n_2 \rangle \right|^2 + 2 \delta_{n_1,n_2} \text{Re} \left( \langle n_1 | \frac{d^2 M_{a_k}^{u_k}}{du^2} | n_2 \rangle \right)
\]

where \( \text{Re}(x) \) denotes the real part of complex valued scalar \( x \), and \( \delta \) denotes the Kronecker Delta. When \( R \) is a nonzero matrix, the non-negative matrix \( P = I - \frac{R_{\text{Tr}(R)}}{R} \) is a stochastic matrix (each row sums to one).
Lemma 3 ([9]): Assume that the directed graph $G$ of the matrix $R$ defined in Lemma 2 is strongly connected. For any $\bar{n} \in \{1, \ldots, d\}$ there exist $d-1$ strictly positive real numbers $e_{\bar{n} n}, n \in \{1, \ldots, d\} \setminus \{\bar{n}\}$, such that

- for any real numbers $\{\lambda_{n}\}, n \in \{1, \ldots, d\} \setminus \{\bar{n}\}$, there exists a unique $\sigma = \{\sigma_{n}\}, n \in \{1, \ldots, d\} \in \mathbb{R}^{d}$ with $\sigma_{\bar{n}} = 0$ such that $R \sigma = \lambda \bar{n}$ and $\lambda_{n} = -\sum_{n \neq \bar{n}} e_{\bar{n} n}$
  - if additionally $\lambda_{n} < 0$ for all $n \in \{1, \ldots, d\} \setminus \{\bar{n}\}$, then $\sigma_{n} > 0 \forall n \in \{1, \ldots, d\} \setminus \{\bar{n}\}$.
- for any solution $\sigma \in \mathbb{R}^{d}$ of $R \sigma = \lambda$, the function $V_{\rho}(\sigma) = \sum_{n=1}^{d} \sigma_{n} \langle \rho | n \rangle |B\rangle |B\rangle$ satisfies

$$d^{2} V_{\rho}(\sum_{a_{k}} M_{a_{k}}^{\dagger} (\langle n | M_{a_{k}}^{n} \rangle |B\rangle |B\rangle |B\rangle |B\rangle)_{|B\rangle |B\rangle}_{|B\rangle |B\rangle} = \lambda_{n}$$

$\forall n \in \{1, \ldots, d\}$

D. Martingale Proof for Density Operator Evolution

As stated in Sec. III-B we prove that under the evolution of $\rho_{\bar{n}}$, $\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\}$ is a martingale, i.e.

$$\mathbb{E}[\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\} | B_{\mathbb{E}}] = \{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\} \forall T \in \mathbb{N}$$

where the filtration $B_{\mathbb{E}} = \sigma(\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\}, \ldots, \{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\})$.

Denote

$$M_{a_{k}} \rho_{\bar{n}} = M_{a_{k}}^{0} \rho_{\bar{n}} \quad \text{where} \quad T \in \mathbb{N}$$

$$M_{a_{k}}^{\dagger} M_{a_{k}}^{n} = \frac{M_{a_{k}}^{\dagger} M_{a_{k}}^{n} M_{a_{k}}^{0} + M_{a_{k}}^{0} M_{a_{k}}^{n} M_{a_{k}}^{\dagger}}{\text{Tr}(M_{a_{k}}^{0} \rho_{\bar{n}} M_{a_{k}}^{\dagger})}$$

Since

$$\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\} = \text{Tr}(\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\})$$

$$\mathbb{E}[\text{Tr}(\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\} | B_{\mathbb{E}})] = \sum_{a_{k} = 1}^{m} \text{Tr}(M_{a_{k}} \rho_{\bar{n}} M_{a_{k}}^{\dagger}) \text{Tr}(\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\})$$

$$\sum_{a_{k} = 1}^{m} \text{Tr}(M_{a_{k}} \rho_{\bar{n}} M_{a_{k}}^{\dagger}) \text{Tr}(\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\} = \text{Tr}(\sum_{a_{k} = 1}^{m} M_{a_{k}} \rho_{\bar{n}} M_{a_{k}}^{\dagger} |b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\))$$

Thus $\{b_{\bar{n}}, \rho_{\bar{n}} | b_{\bar{n}}\}$ is a martingale.

REFERENCES

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

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

[3] J. Busemeyer, Q. Zhang, S. Balakrishnan, and Z. Wang, “Application of quantum—Markov open system models to human cognition and decision,” Entropy, vol. 22, no. 9, p. 990, 2020.

[4] M. Askarpour, D. Mandrioli, M. Rossi, and F. Vicentini, “Formal model of human erroneous behavior for safety analysis in collaborative robotics,” Robotics and computer-integrated Manufacturing, vol. 57, pp. 465–476, 2019.

[5] D. Belanche, C. Flavían, and A. Pérez-Ruiz, “Consumer empowerment in interactive advertising and e-commerce consequences: The pitre model,” Journal of Marketing Communications, vol. 26, no. 1, pp. 1–20, 2020.

[6] L. Liu, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, and T. Zhou, “Recommender systems,” Physics reports, vol. 519, no. 1, pp. 1–49, 2012.

[7] A. R. Hota and S. Sundaram, “Controlling human utilization of failure-prone systems via taxes,” IEEE Transactions on Automatic Control, vol. 66, no. 12, pp. 5772–5787, 2021.
[32] E. M. Pothos and J. R. Busemeyer, “A quantum probability explanation for violations of ‘rational’ decision theory,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 276, no. 1665, pp. 2171–2178, 2009.

[33] P. Pearle, “Simple derivation of the Lindblad equation,” *European Journal of Physics*, vol. 33, no. 4, pp. 805–822, Apr 2012. [Online]. Available: https://doi.org/10.1088/0143-0807/33/4/805

[34] H. Amini, P. Rouchon, and M. Mirrahimi, “Design of strict control-Lyapunov functions for quantum systems with qud measurements,” in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 8193–8198.

[35] H. Kushner, *Introduction to Stochastic Control*, Holt, Rinehart and Winston, 1971. [Online]. Available: https://books.google.com/books?id=vUrzAAAAZMAJ

[36] Z. Xue, H. Lin, and T. H. Lee, “Identification of unknown parameters for a class of two-level quantum systems,” *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1805–1810, 2013.

---

**Luke Snow** Biography text here.

**Shashwat Jain** Biography text here.

**Vikram Krishnamurthy** Biography text here.