2022 IEEE 61st Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Lyapunov based Stochastic Stability of Human-Machine Interaction: A Quantum Decision System Approach

Luke Snow, Shashwat Jain, Vikram Krishnamurthy

2022 IEEE 61st Conference on Decision and Control (CDC) | 978-1-6654-6761-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CDC51059.2022.9992472

***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.**

I. Introduction

Recent studies in mathematical psychology [21], [19], [8], show that the Lindbladian equations from quantum mechanics facilitate modeling peculiar aspects of human decision making. Such quantum decision models capture *order effects* (humans perceive $P(H|A\cap B)$ and $P(H|B\cap A)$ differently in decision making) and violation of the *sure thing principle* (human perception of probabilities in decision making violates the total probability rule). 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.

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 π_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 human's psychological state ρ is represented as a density operator in Hilbert Space, which evolves via the Lindbladian equation parametrized by the observation y_k and input u_k. The density operator ρ encodes a probability distribution over actions $\{a_j\}_{j\in\{1,\dots,m\}}$, 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 [5], interactive marketing/advertising [6] and recommender systems [20] exploit models for human decision making. One example is a machine assisted healthcare system for patients with dementia [12], 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 (ρ) and the resultant hand washing actions (a_k) are sensed by the sensor, then the controller 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.

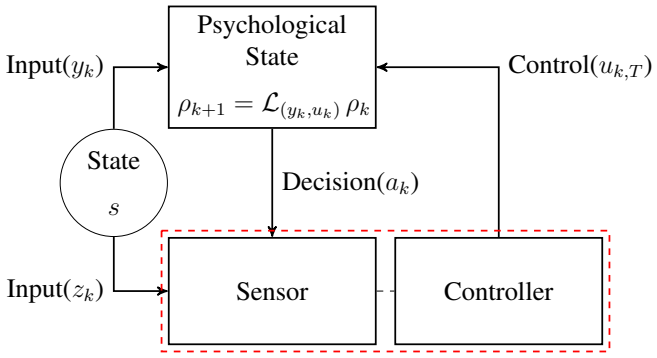

Fig. 1. Human-Machine Interaction Model

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 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 vikramk@cornell.edu are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA

this by developing a novel Lyapunov stability result for a Lindbladian dynamic system.

The main results and organization of the paper is as follows:
(1) Sec. II introduces the open-quantum cognitive model of Martinez et. al. [21] and represents the discretized process in a form that will be mathematically useful for us.
2) Sec. III presents Theorem (1), which is our main result and shows the stochastic stability of our human-machine decision system. The proof uses the methodology of Amini et. al [3], along with Lyapunov techniques and Theorem (2) which we provide in Sec. IV.
3) Sec. IV provides a generalization of a finite-step Lyapunov stability result given in Qin et. al. [25] in Theorem (4), to the case when the finite-step interval T is a random variable. Also Theorem (2) is a modified form of this result which is used to prove Theorem (1) in Sec. III.

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) [22] and Savage (1954) [26]. Despite the successes of these models, numerous experimental findings, most notably those of Kahneman and Tverksy [14], 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 [15]. However, given the complexity of human psychology and behavior it is no surprise that current models still have points of failure. The theory of Quantum Decision Making ([7], [16], [30] 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' [17], [1], 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 [28] [9] ("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.

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 [19], [7].

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 [10] 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 [4], and [21] 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. [19] 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".

Notation

1) $|n\rangle \in \mathcal{S}$ is n^{th} basis vector in a Hilbert space $\mathcal{S}$.
2) $\mathcal{H}(\{|s_1\rangle, \ldots, |s_k\rangle\})$: Hilbert space spanned by the orthonormal basis vector set $\{|s_1\rangle, \ldots, |s_k\rangle\}$
3) $A^\dagger$: adjoint of A
4) $|n\rangle\langle m|$: outer product of $|n\rangle$ and $|m\rangle$
5) Density operator ρ: $\rho = |\Psi\rangle\langle\Psi|$ for some $|\Psi\rangle \in \mathcal{S}$
6) Trace of operator A: $\mathrm{Tr}(A) = \sum_{l=1}^{N} \langle n_l| A |n_l\rangle$ for basis $|n_l\rangle \in \mathcal{S}$
7) Random events are defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

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 as given in [21].

A. Lindbladian Dynamics

Given a state of nature 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 $\mathcal{L}$ denoting the Lindblad operator, the Lindbladian ordinary differential equation for the dynamics of the psychological state ρ_t over time $t \in [0, \infty)$ is specified as

$$\frac{d\rho_t}{dt} = \mathcal{L}_{(\alpha,\lambda,\phi)}\,\rho_t, \quad \rho_0 = \frac{1}{nm}\mathrm{diag}(1, \cdots, 1)_{mn\times mn} \quad (1)$$

Here (α, λ, ϕ) are free parameters which determine the quantum decision maker's behavior, as discussed in Appendix VI-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.

In subsequent sections we will control the evolution of ρ_t and formulate Lyapunov stability conditions. We are interested in the general case when the machine can only observe the human's actions and output a control every T time steps, where T is a random variable. We thus define the

evolution of the density operator in (2), from the controller's perspective, over T-step iterations.

The psychological model comprises the following:

1) A state of nature $s \in \{1, \ldots, n\} = \mathcal{X}$, with probability mass function $\pi_0(s)$
2) An action $a_k \in \{1, \ldots, m\} = \mathcal{A}$ by the human at discrete time k, for $k = 0, 1, 2, \ldots$.
3) Scalar control input $u_k \in [-\bar{u}, \bar{u}]$, $\bar{u} \in \mathbb{R}_+$, from the machine. This controls the parameters of the Lindbladian operator in equation (2), and models a recommendation signal (for example a posterier probability of the state of nature) by the machine to the human.
4) The discrete time evolution of the psychological state

$$\rho_{k+T} = \mathbb{M}^{u_k}_{a_k,T}\,\rho_k = \frac{M^{u_k}_{a_k,T}\,\rho_k\,M^{u_k\dagger}_{a_k,T}}{\mathrm{Tr}(M^{u_k}_{a_k,T}\,\rho_k\,M^{u_k\dagger}_{a_k,T})} \tag{2}$$

where $k = 0, 1, \ldots$ denotes discrete-time. Recall the random variable $T : \Omega \to \mathbb{N}$ specifies the time intervals over which the machine interacts with the human. T has a known probability mass function $\pi_T(\cdot)$. a_k is a T-length sequence $\{a_{k_i}\}_{i=1}^{T}$of random actions a_{k_i} taking values in $\mathcal{A}$. See Appendix VI-C for the definition of $\mathbb{M}^{u_k}_{a_k,T}$.
5) The action probability distribution at time k

$$p(a_k) = \mathrm{Tr}(M^{u_k}_{a_k}\,\rho_k\,M^{u_k\dagger}_{a_k}), a_k \in \{1, \cdots, m\} \tag{3}$$

See Appendix VI-B and VI-C for further model details.

B. Practicality in Modeling Human Decision Making

The above Lindbladian model captures important human decision features such as the sure-thing principle and order effects, which we now describe. These features cannot be explained by purely Markovian models without sacrificing their explanatory power.

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

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

for events A and B. Violation occurs when $=$ is replaced by either $<$ or $>$. Suppose $P(A|B) = 0.6$, $P(A|\bar{B}) = 0.5$. Then if the probability the human decision maker chooses action A is either greater than 0.6 or less than 0.5, then the STP has been violated.

[21] shows that the Lindbladian model accounts for violations of the STP. Pothos and Busemeyer [24] (see also [17]) 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.

Order Effects: It is well-established in psychology [27], [11], [29] that the order of presented information can affect the final judgement of a human [13]. Order effects are not easily accounted for using classical set-theoretic probability axioms, since $P(H|A \cap B) = P(H|B \cap A)$, i.e. the order of presentation of events A and B does not influence the final probability judgement of H. Alternative models of inference have been proposed, such as the averaging model [27] and the belief-adjustment model [13], but these are only heuristic ad-hoc models which lack axiomatic foundations. Quantum probability is a natural axiomatic framework which can account for these effects, see [28], [7], [16], [9] and references therein. Order effects naturally arise from the non-commutative structure of quantum interactions.

III. Machine Control of Human Decision Maker

This section exploits the Lyapunov function formulated in [3] and a generalized finite-step convergence theorem, (that will be proved in Section (IV)), to prove our main result, Theorem 1. This theorem states that regardless of the initial psychological state of the human, the machine is able to control the preference in such a way that the target action is eventually chosen at every time step with probability one.

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

$$D := \{\rho \in \mathbb{C}^{d\times d} : \rho = \rho^\dagger, \mathrm{Tr}(\rho) = 1, \rho \geq 0\} \tag{4}$$

Let $\{|b_r\rangle\}_{r=1}^{d}$ be a set of orthonormal vectors in $\mathbb{C}^d$, where each $|b_r\rangle$ corresponds to a unique state-action pair. Let $\mathcal{S}$ be the Hilbert space formed by taking these $\{|b_r\rangle\}_{r=1}^{d}$ as an orthonormal basis. We consider scalar control inputs $u_k \in \mathbb{R}$ satisfying constraints given in Appendix VI-D. For our purposes it suffices that $u_k \in [-1, 1]$, see [3] for details. The term *'Open-loop (super) martingale'* below denotes a (super) martingale when the control input $u_k = 0$ for $k = 0, 1, \ldots$.

The following is the main result:

Theorem 1: Given the discrete time density operator evolution (2) and any target state $|\bar{b}_r\rangle, r \in \{1, \ldots, d\}$, there exists a control sequence $\{u_k\}_{k\in\mathbb{N}}$ generated by the machine such that the human psychological state ρ_k converges to $|\bar{b}_r\rangle\langle\bar{b}_r|$ with probability one for any initial $\rho_0 \in D$.

Proof: We will follow the formulation developed in [3]. First with $\beta_k = \{a_k, T\}$, rewrite (2) as

$$\rho_{k+T} = \mathbb{M}^{u_k}_{\beta_k}\rho_k = \frac{M^{u_k}_{\beta_k}\,\rho_k\,M^{u_k\dagger}_{\beta_k}}{\mathrm{Tr}(M^{u_k}_{\beta_k}\,\rho_k\,M^{u_k\dagger}_{\beta_k})} \tag{5}$$

We define the following Lyapunov function, which forms a supermartingale under both open-loop (zero-input) and closed-loop (feedback control (u_k)) conditions for the process (5):

$$V_\epsilon(\rho) := \sum_{r=1}^{d} \sigma_r \langle b_r|\,\rho\,|b_r\rangle - \frac{\epsilon}{2}\sum_{r=1}^{d}\langle b_r|\,\rho\,|b_r\rangle^2 \tag{6}$$

where σ_r is non-negative $\forall r \in \{1, \ldots, d\}$ and ϵ is strictly positive. The set $\{\sigma_r\}_{r=1}^{d}$ and ϵ are chosen according to [3] such that the Lyapunov function $V_\epsilon(\rho_k) > 0\,\forall\,\rho_k \in D$ and ρ_k converges to the intended subspace $|\bar{b}_r\rangle\langle\bar{b}_r|$ with probability 1. By [3] and [2], $\langle b_r|\,\rho\,|b_r\rangle$ is an open-loop martingale given the density operator evolution (5) (see Appendix VI-E for proof). V_ϵ is a concave function of the open-loop martingales $\langle b_r|\,\rho\,|b_r\rangle$ and therefore is an open-loop $(u_k = 0)$ supermartingale given the process (5).

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

The following feedback control mechanism is used

$$u_k := \underset{u\in[-1,1]}{\text{argmin}}\ \mathbb{E}[V_\epsilon(\rho_{k+T})|\rho_k, u]$$

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

$$\mathbb{E}[V_\epsilon(\rho_{k+T})|\rho_k, u] = \mathbb{E}[V_\epsilon(\mathbb{M}^u_{\beta_k}\rho_k)] = \int_\Omega V_\epsilon(\mathbb{M}^u_{\beta_k(\omega)}\rho_k)d\omega$$

where Ω is the sample space under which the process is induced. Define $\tilde{Q}(\rho_k) := \mathbb{E}[V_\epsilon(\rho_{k+T})|\rho_k, u_k] - V_\epsilon(\rho_k) \leq \mathbb{E}[V_\epsilon(\rho_{k+T})|\rho_k, u_k = 0] - V_\epsilon(\rho_k) \leq 0$. $V_\epsilon(\rho)$ is a continuous function, so using [18, Chapter 8], we have the T-step control sequence $\{\rho_{k+iT}\}_{i\in\mathbb{N}}$ that converges to the set $D_\infty := \{\rho : \tilde{Q}(\rho) = 0\}$ with probability one.

We will first show that the set D_∞ is restricted to our desired state $\{|\bar{b}_r\rangle\langle\bar{b}_r|\}$, then that the entire sequence $\{\rho_k\}_{k\in\mathbb{N}}$ converges to this set. The former is proved in Lemma 2 of [3]; For any target subspace $|\bar{b}_r\rangle\langle\bar{b}_r|$, the set $\{\sigma_r\}_{r=1}^d$ can be chosen in such a way that $D_\infty = |\bar{b}_r\rangle\langle\bar{b}_r|$. The idea is the following: A state ρ_k is in the limit set D_∞ iff for all $u \in [-1, 1]$,

$$\mathbb{E}[V_\epsilon(\rho_{k+T})|\rho_k, u] - V_\epsilon(\rho_k) \geq 0 \tag{7}$$

Also, since V_ϵ is an open-loop supermartingale, $\forall \rho_k \in D$:

$$\mathbb{E}[V_\epsilon(\rho_{k+T})|\rho_k, u=0] - V_\epsilon(\rho_k) \leq 0 \tag{8}$$

By Lemma 2 of [3], given any $\bar{r} \in \{1, \ldots, d\}$ and fixed $\epsilon > 0$, the weights $\sigma_1, \ldots, \sigma_d$ can be chosen so that V_ϵ satisfies the following property: $\forall r \in \{1, \ldots, d\}, u \mapsto \mathbb{E}[V_\epsilon(\rho_{k+T}) \ |\rho_k = |b_r\rangle\langle b_r|, u_k = u]$ has a strict local minimum at $u = 0$ for $r = \bar{r}$ and a strict local maximum at $u = 0$ for $r \neq \bar{r}$. This combined with (8) ensures that for any $r \neq \bar{r}, \exists\, u \in [-1, 1]$ such that $\mathbb{E}[V_\epsilon(\rho_{k+T}) \ |\rho_k = |b_r\rangle\langle b_r|, u_k = u] - V_\epsilon(|b_r\rangle\langle b_r|) < 0$. Therefore, by (7), we have that $|b_r\rangle\langle b_r|$ is in the limit set l_∞ if and only if $r = \bar{r}$.

We now show that $\mathbb{P}[\lim_{k\to\infty}\rho_k = |\bar{b}_r\rangle\langle\bar{b}_r|] = 1$, i.e. the entire sequence converges to the target state with probability one. This utilizes Theorem 2 which is developed in Section IV. We have shown that $\mathbb{P}[\lim_{i\to\infty}\rho_{k+iT} = |\bar{b}_r\rangle\langle\bar{b}_r|] = \mathbb{P}[\exists K \in \mathbb{N} : \rho_{k+iT} = |\bar{b}_r\rangle\langle\bar{b}_r| \ \forall i \geq K] = 1$. The set $\{\sigma_r\}$ was chosen such that $\mathbb{E}[V_\epsilon(\rho_{k+T}) \ |\rho_k = |b_r\rangle\langle b_r|, u_k = u]$ has a strict local minimum at $u = 0$ for $r = \bar{r}$ so $u_{k+iT} = \underset{u\in[-1,1]}{\text{argmin}}\{\mathbb{E}[V_\epsilon(\rho_{k+(i+1)T})|\rho_{k+iT}, u]\} = 0 \,\forall\, i \geq K$. So, for $q \in \{1, \ldots, T\}$, $\mathbb{E}[V_\epsilon(\rho_{k+(i+q+1)T})|\rho_{k+(i+q)T}, u = 0] = \mathbb{E}[V_\epsilon(\mathbb{M}^0_{\beta_k}\rho_{k+(i+q)T})] \leq V_\epsilon(\rho_{k+(i+q)T})$ since V_ϵ is an open-loop supermartingale. We now know $\exists K \in \mathbb{N}$ such that $\mathbb{E}[V_\epsilon(\rho_{k+T})|\rho_k, u_k] \leq V_\epsilon(\rho_k) \ \forall k \geq K$. Since there exists a unique mapping from elements of D to elements of $\mathbb{R}^{2d^2}$, we apply Theorem 2 to prove $\mathbb{P}[\lim_{k\to\infty}\rho_k = |\bar{b}_r\rangle\langle\bar{b}_r|] = 1$.

Since $\forall a_j \in \mathcal{A}, \exists$ a set $\{|b_{j_1}\rangle, \ldots, |b_{j_n}\rangle\} = a_j \otimes \mathcal{X}$, the convergence to any $|\bar{b}_r\rangle\langle\bar{b}_r|, |\bar{b}_r\rangle \in a_j \otimes \mathcal{X}$, implies the convergence to $a_j \in \mathcal{A}$. ∎

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 asymptotically, almost surely.

IV. Finite Step Stochastic Lyapunov Stability

The purpose of this section is two-fold. First, we generalize Theorem 1 of Qin et. al. [25] to the case when the finite step size can be a random variable. Our main result below is Theorem 2. Recall that we used this result in Section III to prove stability of the human-decision system. Second, Theorem 4 below is of independent interest.

Consider the discrete time stochastic system described by

$$x_{k+1} = f(x_k, y_{k+1}), \quad k = 0, 1, 2, \ldots \tag{9}$$

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 σ-fields) defined by $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_k = \sigma(y_1, \ldots, y_k)$ 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 2: For the discrete-time stochastic system (9), let $V : \mathbb{R}^n \to \mathbb{R}$ be a continuous non-negative and radially unbounded function. Suppose we have the condition:
(a) There exists a random variable $T : \Omega \to \mathbb{N}$ such that for any k, $\mathbb{E}[V(x_{k+T})|\mathcal{F}_k] - V(x_k) \leq -\varphi(x_k)$, where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is continuous and satisfies $\varphi(x) \geq 0$ for any $x \in \mathbb{R}^n$. 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.

This Theorem follows from Theorems 3 and 4; proofs are given for both of these.

Theorem 3: For the discrete-time stochastic system (9), 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 λ, and assume that:
(a) $\mathbb{E}[V(x_{k+1})|\mathcal{F}_k] - V(x_k) \leq 0$ for any k such that $x_k \in Q_\lambda$
(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 $\varphi : \mathbb{R}^n \to \mathbb{R}$ 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 $\mathcal{D}_1 := \{x \in Q_\lambda : \varphi(x) = 0\}$ with probability at least $1 - V(x_0)/\lambda$

Proof: We have

$$\mathbb{E}[V(x_{k+T})|\mathcal{F}_k] - V(x_k) \leq -\varphi(x_k) \leq 0, \forall x_k \in Q_\lambda \tag{10}$$

where $\varphi(x)$ is continuous. Now, Kushner [18, p.196] has shown that if we start with $x_0 \in Q_\lambda$ then the probability of staying in Q_λ during the entire resultant trajectory is at least $1 - V(x_0)/\lambda$, i.e.

$$\mathbb{P}[\sup_{k\in\mathbb{N}} V(x_k) \geq \lambda] \leq V(x_0)/\lambda \tag{11}$$

Next construct T subsequences of $\{X_k\}$ as follows: $\{X_i^{(0)}\} = \{X_0, X_T, \ldots\}, \{X_i^{(1)}\} = \{X_1, X_{T+1}, \ldots\}, \ldots, \{X_i^{(T-1)}\} = \{X_{T-1}, X_{2T-1}, \ldots\}$. 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 (10), and thus by Doob's convergence

theorem converges to a limit with probability 1. From (10) we have for all $k \in K$ and $n \in \mathbb{N}$ $\sum_{l=1}^{n} \mathbb{E}(V(X_l^{(k)})) - \mathbb{E}(V(X_{l-1}^{(k)})) \leq -\mathbb{E}(\sum_{l=0}^{n-1} \varphi(X_l^{(k)}))$ and $0 \leq \mathbb{E}(V(X_n^{(k)})) \leq \mathbb{E}(V(X_0^{(k)})) - \mathbb{E}(\sum_{l=0}^{n-1} \varphi(X_l^{(k)}))$. We use Fatou's Lemma to obtain $\mathbb{E}(\sum_{l=0}^{\infty} \varphi(X_l^{(k)})) < \infty$ and by Borel-Cantelli we have $\mathbb{P}[\lim_{n\to\infty} \varphi(X_n^{(k)}) = 0] = 1 \, \forall \, k \in K$. Now suppose $\varphi(x) \geq 0$ only for $x \in Q_\lambda$. Stop $\{X_n^{(k)}\}$ on first leaving Q_λ. Then for $x \notin Q_\lambda$, $\varphi(x) = 0$ for this stopped set. This stopped process is a supermartingale and the proof is the same as above.

It is now apparent that $\lim_{n\to\infty} \varphi(X_n^{(k)}(\omega)) = 0 \; \forall k \in K$ and $\omega \in \bar{\Omega} = \{\omega \in \Omega : x_n(\omega) \in Q_\lambda \; \forall n \in \mathbb{N}\}$, so we have $\mathbb{P}[\lim_{n\to\infty} \varphi(X_n(\omega)) = 0] \geq 1 - V(x_0)/\lambda$ by the analysis in Appendix VI-A and (11). ∎

Theorem 4: Theorem (3) holds when T is an integer-valued random variable $T : \Omega \to \mathbb{N}$, where Ω is the underlying sample space and T is independent of the process $X_l^{(k)}$.

Proof: This follows from the previous proof, with expectations conditioned on $T(\omega) = \tau$, the set of sequences with interval length $\tau \in \mathbb{N}$: $\sum_{l=1}^{n} \mathbb{E}(V(X_l^{(k)})) - V(X_{l-1}^{(k)})|T = \tau) \leq -\mathbb{E}(\sum_{l=0}^{n-1} \varphi(X_l^{(k)})|T = \tau)$. Applying Fatou's Lemma yields $\mathbb{E}(\sum_{l=0}^{\infty} \varphi(X_l^{(k)})|T = \tau) < \infty$ and so by Borel-Cantelli we have $\mathbb{P}[\lim_{n\to\infty} \varphi(X_n^{(k)}) = 0|T = \tau] = 1 \quad \forall k \in K$. The same arguments from the proof of Theorem (3) yield $\mathbb{P}[\lim_{n\to\infty} \varphi(X_n(\omega)) = 0|T = \tau] \geq 1 - V(x_0)/\lambda$. So, $\mathbb{P}[\lim_{n\to\infty} \varphi(X_n(\omega)) = 0|T = \tau]\mathbb{P}(T = \tau) \geq (1 - V(x_0)/\lambda)\mathbb{P}(T = \tau) \implies \mathbb{P}[\lim_{n\to\infty} \varphi(X_n(\omega)) = 0] = \sum_{\tau\in\mathbb{N}} \mathbb{P}[\lim_{n\to\infty} \varphi(X_n(\omega)) = 0|T = \tau]\mathbb{P}(T = \tau) \geq \sum_{\tau\in\mathbb{N}}(1 - V(x_0)/\lambda)\mathbb{P}(T = \tau) = 1 - V(x_0)/\lambda$ ∎

To summarize, this section provided a generalization of the finite-step Lyapunov function result of Qin. et. al. [25]. We applied this to show 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

Our main result, Theorem (1), showed that for a human-machine decision system modeled as a controlled quantum decision system, there exists an optimal control policy under which the human's action choice can be guided to an arbitrary target action with probability one. This can be useful in tense or stressful decision situations when the human is subject to cognitive bias and irrational preferences; The machine can act as a rational Bayesian expected utility maximizer and control the human's preference to a Bayesian optimal choice. In proving this we have utilized a novel random finite-step Lyapunov function result, which we present and prove in Sec. (IV), and which stands as an independent result. There are several simplifying assumptions we have made, which warrant further investigating. For one, we have assumed that the machine knows the (α, λ, ϕ) parametrization of the human's Lindbladian mental operator. It would be interesting to see what the analysis yields when these are only estimates with some distribution. It is worthwhile extending our results to more general human-machine decision systems.

VI. Appendix

A. Convergence of constructed subsequences implies convergence of sequence

We have $\lim_{n\to\infty} \varphi(X_n^{(k)}(\omega)) = 0 \; \forall \omega \in \bar{\Omega}$. Let $\omega \in \bar{\Omega}$ and $\varphi_n^{(k)}$ denote $\varphi(X_n^{(k)}(\omega))$ and φ_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 $\lim_{n\to\infty} \varphi_n \neq 0$: Then $\exists \epsilon > 0$ such that $\forall \; N \in \mathbb{N} \; \exists \; n_0 > N$ with $\varphi_{n_0} > \epsilon$. 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 $\lim_{n\to\infty} \varphi_n = 0$.

B. Lindbladian Psychological Model Construction

Let psychological state space $\mathcal{S} := \mathcal{H}(\{|\mathcal{E}_l, \mathcal{A}_j\rangle\}_{\substack{1\leq l\leq n \\ 1\leq j\leq m}})$ be a Hilbert space. Here $\mathcal{H}$ is defined in Notation (2). Each $|\mathcal{E}_l\rangle$ is an n-dimensional complex vector indexed by the state l, and $|\mathcal{A}_j\rangle$ is an m-dimensional complex vector indexed by the action j. The evolution of the density operator is given by $\frac{d\rho}{dt} = \mathcal{L}_{(\alpha,\lambda,\phi)} \, \rho$ where

$$\mathcal{L}_{(\alpha,\lambda,\phi)} \, \rho = -i(1-\alpha)[H, \rho] + \alpha \sum_{m,n} \gamma_{(m,n)} \left(L_{(m,n)} \, \rho \, L_{(m,n)}^\dagger - \frac{1}{2}\{L_{(m,n)}^\dagger L_{(m,n)}, \rho\} \right) \tag{12}$$

Here $[A, B] = AB - BA$, $\{A, B\} = AB + BA$, A^* is complex conjugate of A, $H = \text{diag}(1_m, \cdots, 1_m)_{mn\times \, mn}$ with 1_m an $m \times m$ matrix of ones. $L_{(m,n)} = |m\rangle \langle n|$ is the jump operator, which represents the jump from m^{th} cognitive state to n^{th} cognitive state. $\gamma_{(m,n)} := [C(\lambda, \phi)]_{m,n} = [(1 - \phi)\Pi^T(\lambda) + \phi K^T]_{m,n}$. For utility function $u : \mathcal{A} \times \mathcal{X} \to \mathbb{R}$, $p(a_j|\mathcal{E}_l) = \frac{u(a_j|\mathcal{E}_l)^\lambda}{\sum_{j=l}^{m} u(a_j|\mathcal{E}_l)^\lambda}$ we define $P(\mathcal{E}_l) := \begin{bmatrix} p(a_1|\mathcal{E}_l) & p(a_2|\mathcal{E}_l) & \cdots & p(a_m|\mathcal{E}_l) \end{bmatrix} \otimes 1_{n\times 1}$ and $\Pi(\lambda) = \text{diag}(P(\mathcal{E}_1), \cdots, P(\mathcal{E}_n))$ where $1_{n\times 1}$ is a vector with all 1's and, $A \otimes B$ is the kronecker product of A and B. Define $\eta_k(s) = p(s|u_k, y_k)$ given the noisy observation y_k and input signal u_k, with $s \in \mathcal{X}$. We define $K := \begin{bmatrix} \eta_k(\mathcal{E}_1) & \eta_k(\mathcal{E}_2) & \cdots & \eta_k(\mathcal{E}_n) \end{bmatrix} \otimes 1_{m\times 1} \otimes \mathbb{I}_{m\times m}$. $\alpha \in [0, 1]$ represents the amount of quantum behavior in the evolution of the density operator. $\lambda \in [0, \infty)$ can be thought of as the agent's ability to discriminate between the profitability of different actions. $\phi \in (0, 1)$ represents the relevance of discrimination between underlying states $\{\mathcal{E}_1, \ldots, \mathcal{E}_n\}$.

When we consider the closed-loop feedback control mechanism, the scalar control input u_k directly parametrizes the structure of the cognitive matrix, so that we can have $C_{u_k}(\lambda, \phi)$. We leave this as a parameter and do not define the specific effect of u_k on $C_{u_k}(\lambda, \phi)$, as long as u_k satisfies the constraints $(i) - (iv)$ of Section VI-D.

C. Discrete time representation of Lindblad Equation

We time discretize the Lindbladian dynamics by representing it using Kraus operators to get (2), see [23]. We discretize the evolution in equation (12) in multiples of Δt in order to incorporate T time steps and rewrite it as $\rho_{k+T} = \sum_a K_{a_k,T}^{u_k} \rho_k K_{a_k,T}^{u_k\dagger}$ where $K_{0,T}^{u_k} = \left[\mathbb{I} - T\Delta t(i(1-\alpha)H + \frac{1}{2}\sum_{m,n}\alpha\gamma_{m,n}^{u_k} L_{(m,n)}^\dagger L_{(m,n)})\right]$ and $K_{a\neq 0,T}^{u_k} = \sqrt{T\Delta t\alpha\gamma_{m,n}^{u_k}} L_{(m,n)}$, $T \in \mathbb{N}$. When the human chooses action a, the psychological state is projected into the respective action subspace $a \otimes \mathcal{X}$. The psychological state evolves as: $\rho_{k+T} = \frac{P_{a_k}(\sum_\nu K_{\nu,T}\,\rho_k\,K_{\nu,T}^\dagger)P_{a_k}^\dagger}{\mathrm{Tr}(P_{a_k}(\sum_\nu K_{\nu,T}\,\rho_k\,K_{\nu,T}^\dagger)P_{a_k}^\dagger)}$ where a_k is a random variable taking values in $\{1,\ldots,m\}$, with probability given by $\mathbb{P}(a_k) = \mathrm{Tr}(P_{a_k}(\sum_\nu K_{\nu,T}\,\rho_k\,K_{\nu,T}^\dagger)P_{a_k}^\dagger)$ 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 $M_{a_k,T}^u = P_{a_k}(\sum_\nu K_{\nu,T}^u)$ we obtain (2).

D. Control Input Constraints

(i): For each $u_k, \sum_a M_a^{u_k\dagger} M_a^{u_k} = \mathbb{I}$ for general quantum measurements a.
(ii): For $u_k = 0$, all M_a^0 are diagonal in the same orthonormal basis $\{|n\rangle \,|\, n \in \{1,\ldots,d\}\}$: $M_a^0 = \sum_{n=1}^d c_{a,n}\,|n\rangle\langle n|\,, c_{a,n} \in \mathbb{C}$
(iii): For all $n_1 \neq n_2$ in $\{1,\ldots,d\}$, $\exists\, a \in \{1,\ldots,d\}$ such that $|c_{a,n_1}|^2 \neq |c_{a,n_2}|^2$.
(iv): All $M_a^{u_k}$ are C^2 functions of u_k

E. Martingale Proof

We prove that under the evolution of (2), $\langle b_r|\,\rho_k\,|b_r\rangle$ is a martingale, i.e. $\mathbb{E}[\langle b_r|\,\rho_{k+T}\,|b_r\rangle\,|\mathcal{B}_k] = \langle b_r|\,\rho_k\,|b_r\rangle \;\; \forall T \in \mathbb{N}$ where the filtration $\mathcal{B}_k = \sigma(\langle b_r|\,\rho_1\,|b_r\rangle, \ldots, \langle b_r|\,\rho_k\,|b_r\rangle)$.
Denote $\mathbb{M}_{a_k}\rho_k = \mathbb{M}_{a_k,T=1}^0\rho_k = \frac{M_{a_k,1}^0\,\rho_k\,M_{a_k,1}^{0\dagger}}{\mathrm{Tr}(M_{a_k,1}^0\,\rho_k\,M_{a_k,1}^{0\dagger})}$. Since $\langle b_r|\,\rho_k\,|b_r\rangle = \mathrm{Tr}(|b_r\rangle\langle b_r|\,\rho_k)$, $\mathbb{E}[\mathrm{Tr}(|b_r\rangle\langle b_r|\,\rho_{k+1})|\rho_k, u] = \sum_{a_k=1}^m \mathrm{Tr}(M_{a_k}\,\rho_k\,M_{a_k}^\dagger)\,\mathrm{Tr}(|b_r\rangle\langle b_r|\,\mathbb{M}_{a_k}\,\rho_k) = \sum_{a_k=1}^m \mathrm{Tr}(|b_r\rangle\langle b_r|\,M_{a_k}\,\rho_k\,M_{a_k}^\dagger) = \mathrm{Tr}(\sum_{a_k=1}^m M_{a_k}M_{a_k}^\dagger\,|b_r\rangle\langle b_r|\,\rho_k) = \mathrm{Tr}(|b_r\rangle\langle b_r|\,\rho_k)$.
Thus $\langle b_r|\,\rho_k\,|b_r\rangle$ is a martingale.

REFERENCES

[1] AERTS, D., BROEKAERT, J., CZACHOR, M., AND D'HOOGHE, B. A quantum-conceptual explanation of violations of expected utility in economics. In *International Symposium on Quantum Interaction* (2011), Springer, pp. 192–198.

[2] AMINI, H., ROUCHON, P., AND MIRRAHIMI, M. Design of strict control-Lyapunov functions for quantum systems with qnd measurements. In *2011 50th IEEE Conference on Decision and Control and European Control Conference* (2011), pp. 8193–8198.

[3] AMINI, H., SOMARAJU, R. A., DOTSENKO, I., SAYRIN, C., MIRRAHIMI, M., AND ROUCHON, P. Feedback stabilization of discrete-time quantum systems subject to non-demolition measurements with imperfections and delays. *Automatica 49*, 9 (2013), 2683–2692.

[4] ASANO, M., BASIEVA, I., KHRENNIKOV, A., OHYA, M., AND TANAKA, Y. Quantum-like dynamics of decision-making. *Physica A: Statistical Mechanics and its Applications 391*, 5 (2012), 2083–2099.

[5] ASKARPOUR, M., MANDRIOLI, D., ROSSI, M., AND VICENTINI, F. Formal model of human erroneous behavior for safety analysis in collaborative robotics. *Robotics and computer-integrated Manufacturing 57* (2019), 465–476.

[6] BELANCHE, D., FLAVIÁN, C., AND PÉREZ-RUEDA, A. Consumer empowerment in interactive advertising and ewom consequences: The pitre model. *Journal of Marketing Communications 26*, 1 (2020), 1–20.

[7] BUSEMEYER, J., AND BRUZA, P. *Quantum Models of Cognition and Decision*. Quantum Models of Cognition and Decision. Cambridge University Press, 2012.

[8] BUSEMEYER, J., ZHANG, Q., BALAKRISHNAN, S., AND WANG, Z. Application of quantum—Markov open system models to human cognition and decision. *Entropy 22*, 9 (2020), 990.

[9] BUSEMEYER, J. R., POTHOS, E. M., FRANCO, R., AND TRUEBLOOD, J. S. A quantum theoretical explanation for probability judgment errors. *Psychological review 118*, 2 (2011), 193.

[10] BUSEMEYER, J. R., WANG, Z., AND LAMBERT-MOGILIANSKY, A. Empirical comparison of Markov and quantum models of decision making. *Journal of Mathematical Psychology 53*, 5 (2009), 423–433.

[11] FURNHAM, A. The robustness of the recency effect: Studies using legal evidence. *The Journal of General Psychology 113*, 4 (1986), 351–357.

[12] HOEY, J., POUPART, P., VON BERTOLDI, A., CRAIG, T., BOUTILIER, C., AND MIHAILIDIS, A. Automated handwashing assistance for persons with dementia using video and a partially observable markov decision process. *Computer Vision and Image Understanding 114*, 5 (2010), 503–519. Special issue on Intelligent Vision Systems.

[13] HOGARTH, R. M., AND EINHORN, H. J. Order effects in belief updating: The belief-adjustment model. *Cognitive psychology 24*, 1 (1992), 1–55.

[14] KAHNEMAN, D., SLOVIC, S. P., SLOVIC, P., AND TVERSKY, A. *Judgment under uncertainty: Heuristics and biases*. Cambridge university press, 1982.

[15] KAHNEMAN, D., AND TVERSKY, A. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, 2013, pp. 99–127.

[16] KHRENNIKOV, A. *Ubiquitous quantum structure*. Springer, 2010.

[17] KHRENNIKOV, A. Y., AND HAVEN, E. Quantum mechanics and violations of the sure-thing principle: The use of probability interference and other concepts. *Journal of Mathematical Psychology 53*, 5 (2009), 378–388.

[18] KUSHNER, H. *Introduction to Stochastic Control*. Holt, Rinehart and Winston, 1971.

[19] KVAM, P. D., BUSEMEYER, J. R., AND PLESKAC, T. J. Temporal oscillations in preference strength provide evidence for an open system model of constructed preference. *Scientific reports 11*, 1 (2021), 1–15.

[20] LÜ, L., MEDO, M., YEUNG, C. H., ZHANG, Y.-C., ZHANG, Z.-K., AND ZHOU, T. Recommender systems. *Physics reports 519*, 1 (2012), 1–49.

[21] MARTÍNEZ-MARTÍNEZ, I., AND SÁNCHEZ-BURILLO, E. Quantum stochastic walks on networks for decision-making. *Scientific reports 6*, 1 (2016), 1–13.

[22] MORGENSTERN, O., AND VON NEUMANN, J. *Theory of games and economic behavior*. Princeton university press, 1953.

[23] PEARLE, P. Simple derivation of the lindblad equation. *European Journal of Physics 33*, 4 (apr 2012), 805–822.

[24] POTHOS, E. M., AND BUSEMEYER, J. R. A quantum probability explanation for violations of 'rational'decision theory. *Proceedings of the Royal Society B: Biological Sciences 276*, 1665 (2009), 2171–2178.

[25] QIN, Y., CAO, M., AND ANDERSON, B. D. Lyapunov criterion for stochastic systems and its applications in distributed computation. *IEEE Transactions on Automatic Control 65*, 2 (2019), 546–560.

[26] SAVAGE, L. J. The theory of statistical decision. *Journal of the American Statistical association 46*, 253 (1951), 55–67.

[27] SHANTEAU, J. C. An additive model for sequential decision making. *Journal of Experimental Psychology 85*, 2 (1970), 181.

[28] TRUEBLOOD, J. S., AND BUSEMEYER, J. R. A quantum probability account of order effects in inference. *Cognitive science 35*, 8 (2011), 1518–1552.

[29] WALKER, L., THIBAUT, J., AND ANDREOLI, V. Order of presentation at trial. *The Yale Law Journal 82*, 2 (1972), 216–226.

[30] YUKALOV, V. I., AND SORNETTE, D. Mathematical structure of quantum decision theory. *Advances in Complex Systems 13*, 05 (2010), 659–698.