Teaching Control Theory in High School

John Doyle, Yorie Nakahira, Yoke Peng Leong, Emily Jenson, Adam Dai, Dimitar Ho, Nikolai Matni
Control and Dynamical Systems, Caltech

Abstract— **Controls is increasingly central to technology, science, and society, yet remains the "hidden technology." Our appropriate emphasis on mathematical rigor and practical relevance in the past 40 years has not been similarly balanced with technical accessibility. The aim of this tutorial is to enlist the controls community in helping to radically rethink controls education. In addition to the brief 2 hour tutorial at CDC, we will have a website with additional materials, but particularly extensive online videos with mathematical details and case studies. We will also have a booth in the exhibition area at CDC with live demos and engaging competitions throughout the conference.**

OVERVIEW

What follows are some short papers that briefly sketch the motivation, striking progress, and substantial remaining challenges in making control theory much more accessible to broader audiences. This involves both broadening the scope of the theory, but also simplifying the math required to a remarkable extent. The "high school" in the title is untested and is really a stretch goal but is not entirely facetious, as we will show that many of the core concepts in controls can be introduced and key theorems proven using only high school algebra. This makes it possible to accessibly explain the basics about the power and danger of control in uncertain dynamical systems, and the difficulties due to plant instability, actuator saturation, and unstable zero dynamics. Including delay and quantization in communication and computation directly greatly broadens the relevance of the theory and make the proofs longer and more tedious but (again remarkably) not fundamentally more advanced.

Neuroscience provides particularly attractive motivating case studies, as almost everyone is interested in how brains work. Surprisingly, experimental demos and theories that significantly push the state of neuroscience research can be reproduced by audiences and readers with little or no equipment, and quite sophisticated case studies can be done using video games requiring minimal and commodity hardware. While this tutorial will focus on neuro motivation, there are many other domains with equally attractive case studies, and particularly modern commodity robotics and toy (e.g. Lego) hardware and software.

As familiar example, imagine riding a mountain bike down a twisting rocky trail. Vision is used to see the trail ahead, and this advanced warning is used to plan a route that stays on the trail. Even large variations (disturbances) in the trail can be navigated with small error provided there is enough advanced warning relative to speed. As a more extreme example, we can easily thread a needle on a different continent given enough time and money. But the bumps in the trail are handled by a reflex layer that has necessarily delayed reactions, which can easily result in catastrophically large errors and crashes to even small disturbances. These reflexes are in addition to the vestibular-ocular reflex (VOR) which is stabilizing vision despite the bumps. Usually, we readily multiplex the planning control with both (and many other) reflexes with no loss in performance, despite being terrible multitaskers in general.

Thus with enough advanced warning and resources we can be almost infinitely robust, and at the opposite extreme, infinitely fragile. Our nervous system has a layered control architecture that allows us to typically exploit the former and avoid the latter. These tradeoffs are not unique to sensorimotor control or even the nervous system. Due to a vaccine, which exploits the advanced warning capability of the immune system, there have been no smallpox deaths since 1980. In contrast, before the vaccine, smallpox infections caused roughly 400 million deaths worldwide from 1900-1980. For these universal tradeoffs and layered architectures, we need a correspondingly universal theoretical framework.

The papers that follow sketch some key ideas on how to make control theory more motivated and accessible, but with a CDC audience in mind. Thus we will skip lots of basic, elementary background necessary for a "high school" audience that would overwhelm the available space and focus on the results we think are most educationally innovative. We also have lots of obvious (and possibly not obvious) holes throughout that we hope you will help us with.

We will have an exhibition booth at CDC with live demos and competitions using video games that illustrate fundamental tradeoffs in sensorimotor control, and a website with lots of videos and other tutorial materials.

Teaching Controls: Neuroscience Motivation

Yorie Nakahira, Yoke Peng Leong, Nikolai Matni, John Doyle

Abstract— Neuroscience potentially provides excellent motivation for not only control theory but also its integration with computing, communications, and biophysics, and allows for simple but illustrative case studies and demos that can be easily reproduced and understood by readers and audiences with diverse backgrounds. The human nervous system is remarkably robust and efficient in both sensorimotor control and maintaining physiological homeostasis, yet also can have devastating fragilities. While enormous progress is being made in understanding both the components and the total system behavior, less progress has been made in connecting these, and one of our objectives is to close this gap, while keeping the needed mathematics as elementary and accessible as possible. This paper focuses on sensorimotor control to motivate specific theory that will be developed further in later papers in this series.

I. Introduction

Human nervous system behavior is remarkably robust in both sensorimotor control and maintaining physiological homeostasis despite hardware that is extraordinarily efficient and flexible but slow and inaccurate, at least by modern engineering standards. Research progress is increasingly elucidating these hardware speed, accuracy, and efficiency tradeoffs in familiar engineering terms [1]. At the behavior level, a similar idea is emerging that humans are better described by robust than optimal control [2]. While these views appear very compatible, what has been missing is rigorous theory to connect component level constraints and tradeoffs with those at the system level. In retrospect, it is now clear that the right blend of robust and network control theory can close this gap, and do so with remarkably accessible math, at least initially.

Specifically, we focus on examples with extremes of robust and fragile behavior, identify equally extreme heterogeneity in nerve composition, motivate the theory that connects them, and throughout present case studies and demos that can be easily reproduced by the reader with minimal and readily available equipment. The papers that follow this one then explore various details in the theory and experiments.

A. Background in neuroscience

Sensorimotor control. Human sensorimotor control has sensors, computation and communication components, and actuators, in which a movement is realized by layers of feedback loops. Its decision principles – such as why the movement follows a particular trajectory out of many other choices – has been a one of the major concern in sensorimotor control community. Although robust control has been used to explain the result of particular motion experiments, little is known about how the hardware limitations reflect on the fundamental limitations of sensorimotor performance

nor the strategies achieving optimal performance given the hardware constraints.

Nervous system. A minimal and admittedly naive but convenient starting point is to view the central nervous system (CNS) as a collection of grey matter bundles of neurons doing computation connected by communications via nerves, which are white matter bundles of up to millions of axons (see Fig. 1), each of which transmits from one neuron to possibly many thousands of receiving neurons. The CNS is also connected by nerves to sensors and actuators in the peripheral nervous system.

Neurons generate action potentials, which are a short-lasting event in which the electrical membrane potential of a cell rapidly rises and falls and propagates down an axon to receiving neurons (via synapses and dendrites). What we will focus on is that spikes in nerves can have temporal (across time) and/or spatial (across axons in the same nerve) dimensions that contains information (such as spiking rate, timing, or combinations of them). However, this neural signaling is limited in both speed and accuracy, and there exists a tradeoff between speed and accuracy in signaling. The latencies are mainly due to conduction delays (the time taken for the action potential to transmit down the axon), refractory period (the time a cell takes to reset to its initial state after firing an action potential), and integration and synaptic delays (such as synapses between nerves, or in the computation procedures of the soma). Limited accuracy is due to having finite number of axons with a finite rate of action potentials per axon (because of energy constraint, refractory period, and so on).

While there is much remaining to be done to understand exactly how the nervous system uses action potentials to communicate for control, and the tradeoffs between efficiency, speed, and accuracy, enough is known to make simple models and compare them with known anatomy and behav-
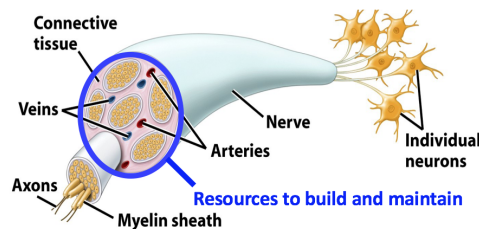

Fig. 1. Nerves are composed of bundles of fibers (called axons) which produce action potentials to transmit information. Given fixed lengths the metabolic cost to build and maintain a nerve is roughly proportional to it cross sectional area.

ior, which are extremely heterogeneous in all dimensions. Such extremes are suggestive of severe constraints and evolutionary pressures, whether they arise from natural selection or engineering design, and thus cry out for a coherent theory, hopefully one consistent with neuroscientists' intuitions. This is what we aim to provide.

B. Why L_1/ℓ_∞ robust control

This collection of papers focuses on controllers that minimize the worst-case $\ell_\infty \to \ell_\infty$ gains of a system, resulting in ℓ_1 robust control problems [3]. Robust and optimal control has increasingly been dominated by frequency domain metrics that quantify bounds on energy, either worst-case ($\mathcal{H}_\infty$) or average-case ($\mathcal{H}_2$). While this is very natural in many traditional engineering applications, and leverages engineers frequency-domain intuitions, we will argue that in the context of sensorimotor control, cell biology, and many other new application areas of interest, this is a less natural space of signals. A dramatic example is a mountain-biker riding down a steep rocky trail. The average or worst-case *energy* transfer to the rider's position caused by bumps is less important than to stay on the trail and not crash, which is more naturally modeled as $\ell_\infty \to \ell_\infty$. Indeed, most biological control is more naturally posed in ℓ_∞ than ℓ_2.

A further reason to study problems with ℓ_∞ bounded signals is that they more naturally and easily allow for quantization and saturation to be incorporated into the control problem, both nonlinear phenomena that are ubiquitous in biological and engineered systems. In fact, as we show in this paper, special instances of such problems even allow for analytic closed form solutions. Finally, to most nonengineers the frequency domain is more obstacle than enabler, at least initially.

## II. EXPLAINING HETEROGENEITY

A. Performance heterogeneity

Human sensorimotor control has very heterogeneous performance, which can be observed using a simple experiment (though better demos are available on our website). First, place your hand in front of your face, close enough to clearly see the fine details of the lines in your palm. Then, perform the following movements:

(i) Oscillate your hand horizontally back and forth, increasing the frequency until the lines blur.
(ii) Hold your hand still and shake your head (in a "no" pattern) at increasing frequencies until blurring occurs.

The blur is either caused by delays in tracking or limitations in visual resolution. The blurring in experiment (i) typically occurs at around 1-2 Hertz whereas in experiment (ii) at a much higher frequency, illustrating the heterogeneous performance (in the level of tolerable disturbance) for the control loops that are responsible during (i) or (ii).

It is essential that the visual system be able to coordinate control of the eye muscles when tracking a moving object in the presence of head motion, and in humans (particularly as hunter-gatherers or athletes) the latter generally creates much higher rates, and thus needs much faster control. There are
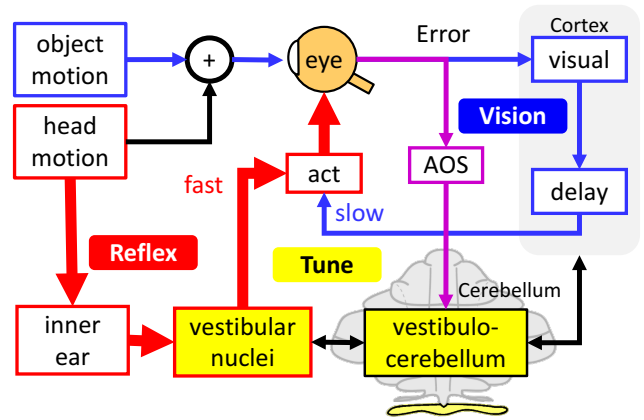
(a) Cartoon diagram of the vision and vestibular-ocular reflex (VOR) system. Objects are tracked with vision, while the much faster VOR system compensates for head motion. The fast and slow path gains sharing the eye muscle actuator must match, and the fast VOR gain is tuned via the vestibular nuclei.
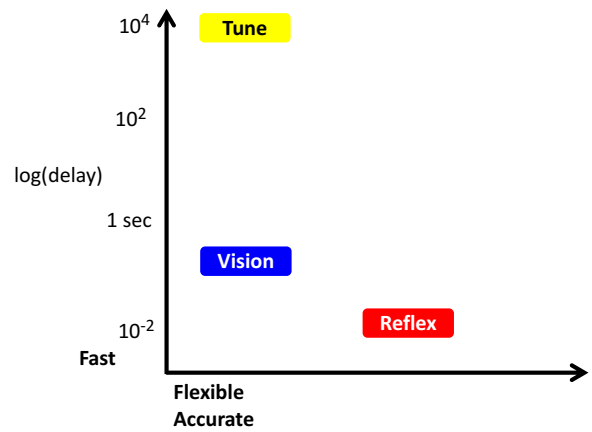
(b) Reflexes are fast but inaccurate and inflexible, whereas vision is slower but accurate and flexible. Tuning is even slower but is essential to have integrated layered control.

Fig. 2. Heterogeneity in the vision/VOR system.

thus two separate but coordinated feedback control loops that implement this process: compensation of the head motion and tracking of the object motion. Head motion is compensated for via the fast vestibulo-ocular reflex (VOR), whereas the object is tracked via slow cortical visual feedback. These are two separate processes, and their gains are tuned to match with each other using neural mechanisms involving the cerebellum, vestibular nuclei, and the accessory optical system (AOS).

The tradeoffs and mechanisms involved in the vision/VOR system are sketched in the cartoons in Fig. 2. Head motion disturbances are controlled with the VOR, which connects quite directly to the eye muscles and has a delay of approximately $10ms$, making it one of the fastest reflexes in the human body. This VOR does not depend on vision, and eyes will track despite head motion even when the eyes are closed, though this is harder to verify. Object motion is tracked using vision with the same eye muscle actuators as VOR but is slow ($> 200ms$ delay), due mostly to long nerve paths and
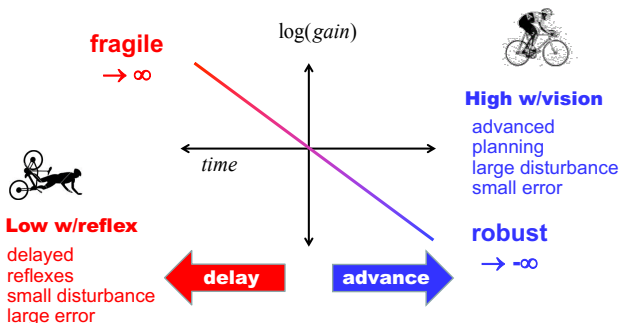
Fig. 3. Extremes in sensorimotor control between delayed reflexes and advanced planning.

cortical computation. Vision however is both accurate and flexible (e.g. you are reading these words) whereas VOR is neither.

Control experts will immediately recognize that the gains on these two systems (the arrows labeled fast and slow entering the common eye muscle actuators) must precisely match, despite huge uncertainties in all the components, and thus not be surprised that there is a another loop to finely tune these gains. This is implemented in the vestibular nuclei and tuned by the vestibulo-cerebellum using the auxiliary optical system (AOS), a specialized pathway in the vision system. The time scales of tuning can be from hours to days depending on the extent of the detuning perturbation (e.g. by wearing prism glasses that distort the visual system).

This example illustrates universal features of the nervous systems that we will explore in more detail. Fast *and* accurate (and flexible) visual performance is achieved with parts that are either fast or accurate but not both. (As we will see, neural hardware that is both fast and accurate would be prohibitively expensive.) Yet most people remain unaware of this unless they perform the above experiment or have some condition that degrades the VOR (e.g. alcohol). The architecture of this system provides fast accurate behavior in normal conditions with mechanisms that are largely unconscious and automatic. The control system is *layered* in having largely independent vision and reflex layer controllers, though learning and tuning is needed to coordinate them, and it is crucial that the right functionality is implemented in the right layer.

Again imagine riding a mountain bike down a twisting rocky trail. As shown in the cartoon in Fig. 3, here too there are vision and reflex layers. Vision is used to see the trail ahead, and this advanced warning is used to plan a route that stays on the trail. Even large variations (disturbances) in the trail can be navigated with small error provided there is enough advanced warning relative to speed. But the bumps in the trail are handled by a reflex layer that has necessarily delayed reactions, which can easily result in catastrophically large errors and crashes to even small disturbances. These reflexes are in addition to the VOR which is stabilizing vision despite the bumps.

Thus with enough advanced warning and resources we can be almost infinitely robust, and at the opposite extreme,

almost infinitely fragile. Our nervous system architecture allows us to typically exploit the former and avoid the latter. These tradeoffs are not unique to sensorimotor control or even the nervous system. As an extreme example of the RHS of Fig.3, we can easily thread a needle on a different continent given enough time and money, but can die due to exposure to tiny amounts of toxins. Due to a vaccine, which also puts the immune system on the RHS of Fig.3, there have been no smallpox deaths since 1980. In contrast, before the vaccine, smallpox infections are on the LHS and caused roughly 400 million deaths from 1900-1980. For these universal tradeoffs and architectures, we need a correspondingly universal theoretical framework.

B. Axon heterogeneity

The VOR and cortical feedback have heterogeneous performances and delays, due to the signaling delay of the vestibular nerve (in VOR feedback) or optic nerve (plus cortical delays in vision). The propagation speed of an action potential in a myelinated axon is approximately proportional to its radius, so the difference in mean diameter of vestibular and optic nerve axons ($2.88\mu m$ and $0.64\mu m$) means the vestibular nerve is on average faster than the optic nerve. This delay and size heterogeneity can be observed also in other types of nerves, and additionally, the heterogeneity can be observed also in the number of axons.

The heterogeneity in size and number of axons are summarized in Fig. 4a for several well-studied nerves. These nerves have axons with extreme variance in size and number: peripheral nerves can have a single large axon (greater than 20μ in diameter); the olfactory nerve has 6 million small axons (approximately $\approx .1\mu$ in diameter and unmyelinated). Other cranial and spinal nerves, including the optic and vestibular vary in between. Later we will translate Fig. 4a into a tradeoff between speed and accuracy in axonal signaling (Fig. 4c) because propagation speed is proportional to axon diameter, signaling bandwidth is proportional to the number of axons, and the spike rate is proportional to axon diameter. These can be combined to derive a simple model of bandwidth/delay tradeoffs that can be plugged into the theory developed in the next section.

In the next section, we develop a control theory specifically to model the sensorimotor system. The model then explains why the observed performance heterogeneity is necessary given the hardware constraints, and why the observed axon heterogeneities are beneficial.

C. Relevant theory

Consider the human sensorimotor system composed of sensors (eyes, inner ear), communication (cranial and spinal nerves of the CNS), controllers (also in the CNS), and actuators (muscle/limb). The sensors use nerve fibers to transmit information to the CNS. The CNS utilizes sensory information to decide the control action and send to the muscle. The muscle executes the control action upon receiving the CNS's command. Of course, in reality this is almost bewilderingly complex and distributed, and a cartoon like
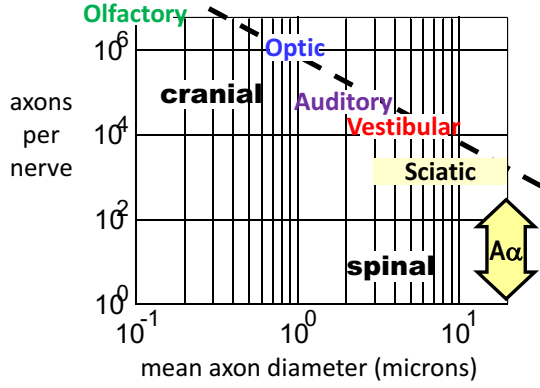
(a) Size and number of axons for selected nerves. A line of equal cross sectional area is indicated. The 4 middle nerves shown have similar areas but wildly different compositions of axon size and number.
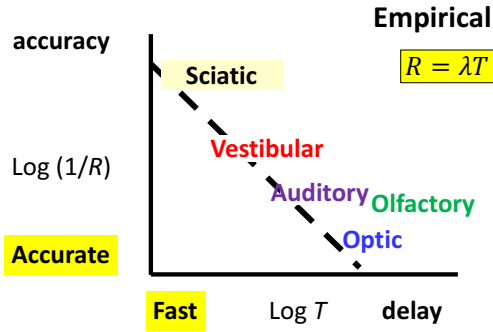
(b) The empirical speed/accuracy or bandwidth/delay tradeoff for different sensorimotor nerves.
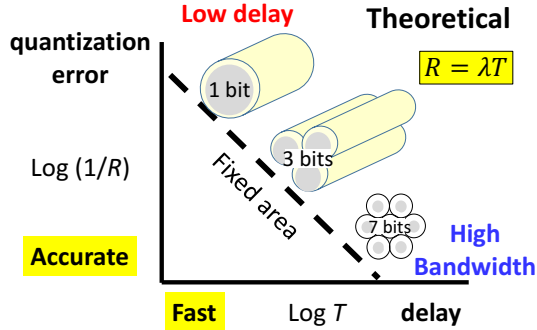
(c) Stylized depiction of the tradeoff between bandwidth R (in bits per unit time) vs delay. The accuracy in transmission scales like 2^{-R}.

Fig. 4.   Nerve heterogeneity and speed accuracy tradeoffs.

Fig. 2a would have delays and bandwidth limits in every box and connection. The simplest possible centralized version with just one of everything is

$$\mathcal{P}: \quad x(t+1) = ax(t) + w(t - T_w) + \mathcal{R}\left[u(t - T_u)\right]$$
$$\mathcal{C}: \quad u(t) = \mathcal{K}(x(0:t), w(0:t), u(0:t-1)) \qquad (1)$$

where $x(t) \in \mathbb{R}$ is a scalar state measuring deviations between desired and actual limb or muscle (or bike) state, $w(t) \in \mathbb{R}$ is a disturbance, and $u(t) \in \mathbb{R}$ is a scalar control action aimed to modify the actual limb/muscle state. We use $\mathcal{P}$ to denote the "plant," which describes the dynamic interaction between muscle/limb state, disturbances and control

action, and $\mathcal{C}$ to denote the "controller," which computes control actions to be applied given the state, disturbance and control action histories.

There are four other parameters of note in dynamics (1).

- a delay $T_u \geq 0$ in the control loop that determines how quickly computed control actions affect the state of the system via actuators;
- a delay $T_w \geq 0$ on the disturbance that allows $w(t)$ to be used in computing control actions prior to its effects being felt on the plant indicating that the system has advanced warning or preview control.
- the full information controller map $\mathcal{K}$, which uses the state, disturbance and control input histories to compute a suitable control action to be applied;
- a quantizer $\mathcal{R}: \mathbb{R} \to \{0,1\}^R$ that restricts the control action to be drawn from a finite set of values of cardinality 2^R per unit time (we denote R as bandwidth).

It is the interplay between these parameters that allows us to connect this simple control theoretic model with the neurophysiology and sensorimotor behavior described in the previous section. Roughly speaking, the disturbance delay T_w is determined by environmental and sensory factors (e.g., am I using vision to detect a moving target, or touch to detect a bump on a trail), and the control delay T_u and the number of bits R of the quantizer $\mathcal{R}$ are determined by neuronal hardware (and as we show below, are constrained to satisfy a speed/resolution tradeoff).

We temporarily set the neuro-interpretation of this problem aside and consider the following robust control problem

$$\min_{K, \mathcal{R}} \quad \max_{\|w\|_\infty \leq 1} \|x\|_\infty$$
$$\text{s.t.} \quad \text{dynamics (1)}.$$

We define the net warning (delay) from the disturbance to the control action as

$$\tau := T_w - T_u.$$

We can use this net warning(delay) to distinguish between two qualitatively different settings:

1) *delayed reaction:* when $\tau < 0$ control actions taken based on the disturbance $w(t)$ only affect the plant after the effects of $w(t)$ have been felt;
2) *advanced warning:* when $\tau \geq 0$ control actions taken based on the disturbance $w(t)$ can reach the plant before the effects of $w(t)$ have been felt.

These two regimes lead to qualitatively different optimal costs:

$$\begin{cases} \sum_{i=1}^{|\tau|} |a^{i-1}| + |a^{|\tau|}| \left(2^R - |a|\right)^{-1} & \text{if } \tau < 0 \\ \left(2^R - |a|\right)^{-1} & \text{if } \tau \geq 0. \end{cases} \qquad (2)$$

We can interpret these costs as follows. When in the advanced warning regime ($\tau \geq 0$), the cost incurred is due solely to quantization error, and reduces to the rate-distortion like expression

$$\left(2^R - |a|\right)^{-1} \qquad (3)$$

For the delayed reaction setting, we see that the cost is the sum of two (essentially independent) components. The first term,

$$\sum_{i=1}^{|\tau|} |a^{i-1}|, \tag{4}$$

is simply the L_1-norm of open loop impulse response of the system over the horizon of the net delay $|\tau|$ – in particular, if the net-delay is set to 0, then this term vanishes. The second term is given by the same quantization error (3) as the advanced warning case, but scaled by $|a^{|\tau|}|$ to capture the effects of the delay $\tau < 0$. Notice here that if we let the number of bits $R \to \infty$, then the total cost incurred in this regime reduces to (4).

To study sensorimotor control in detail, this model will need to be vastly more complex, including multiple disturbances and actuators, and layered, distributed communications and control, but we will add features incrementally and seek insight from the simplest possible models. Next, we consider how implementation in spiking neurons puts constraints on the optimal performance of the system.

D. Designing optimal nerves

Here, we develop a connection between a component level model of nerves in the CNS and overall sensorimotor behavior. In order to do so, we assume that the control action delay T_u can be expressed as the sum

$$T_u = T + T_c,$$

where the signaling delay T is the focus of our nerve design (see below), and there is some additional delay T_c (which can model communication or computation delay in other parts of the system) which we will assume is fixed. Slightly overloading (but simplifying) notation, we define the *fixed net warning* of the system to be

$$\tau := T_w - T_c,$$

as this is an inherent feature of the system, regardless of the chosen signaling delay T.

Our remaining task is thus to quantify the speed/accuracy tradeoff that arises when building a nerve of fixed metabolic resources so that we can select appropriate values for the signaling delay T and number of bits R used by our sensorimotor controller in (1). The metabolic cost to build and maintain a nerve is roughly proportional to its volume [1], and we assume that the overall physical layout of white and grey matter is fixed. Hence we'll assume the lengths of nerves are also fixed, and thus the resources (space and energy) devoted to a nerve are proportional to the cross-sectional area α of it (Fig. 1).

If each nerve is composed of m axons with average radius ρ, then the space used by the nerve is proportional to $\pi m \rho^2$ — the sum of cross-sectional area of m axons. The propagation speed ($\propto \frac{1}{T}$) and firing rate ϕ of an action potential in a myelinated axon is known to be approximately proportional to its radius ρ, *i.e.*, $\phi \propto \rho$ and $T \propto \rho^{-1}$ [1]. Thus, the resource cost satisfies the relation $\alpha \propto \pi m \rho^2$. If we

assume that the nerve is a lossless communication channel (e.g.codeword '1' is the presence of an action potential, and 0' its absence and no bit flips occur) then it can transmit $R = m\phi$ bits per unit time, where ϕ is the maximum firing rate. The assumption of lossless signaling is legitimate for myelanated axons with large diameter (the ones on the right half of Fig. 4a). Combining these dependencies we see that $R \propto \alpha T$, which defines the following tradeoff between signaling speed and accuracy:

$$R = \lambda T, \tag{5}$$

where $\lambda \propto \alpha$ is a proportionality constant determined by the resource consumption of the nerve (Fig. 1). Going forward we fix the axon diameter α, and hence write $R = \lambda T$. As we now show, this tradeoff is the key needed to connect component level constraints with system level behavior via our control theoretic results (2).

In particular, if we substitute $R = \lambda T$ into (2), we obtain an expression that quantifies the impact of the nerve speed/accuracy tradeoff on system performance. For the special case $a = 1$ the optimal cost for both the delayed reaction and advanced warning settings can be succinctly written as

$$\max [0, T - \tau] + (2^{\lambda T} - 1)^{-1}. \tag{6}$$

Once again, this cost can be decomposed into a delay term $\max [0, T - \tau]$ (which only contributes to the cost if the signaling delay T is larger than the fixed net warning τ), and a quantization term $(2^{\lambda T} - 1)^{-1}$.

Fig. 5 shows how the optimal cost varies as we sweep across different values of the signaling delay T (and thus bandwidth R), for the fixed net warning $\tau = 0$ and a fixed proportionality constant λ. As expression (6) makes clear, increasing the signaling delay T incurs a corresponding linear penalty in the delay term, but leads to an exponential decrease in the quantization component of the cost. This quantifies the familiar notion that *bits are powerful*. Consequently, the optimal choice of axon size (as parameterized by the signaling delay T) is achieved at a sweet spot of intermediate levels of delay and bandwidth. This contrasts dramatically with what is suggested by information theory, which has dominated theoretical neuroscience [4], and emphasizes maximizing bandwidth without consideration of delay.
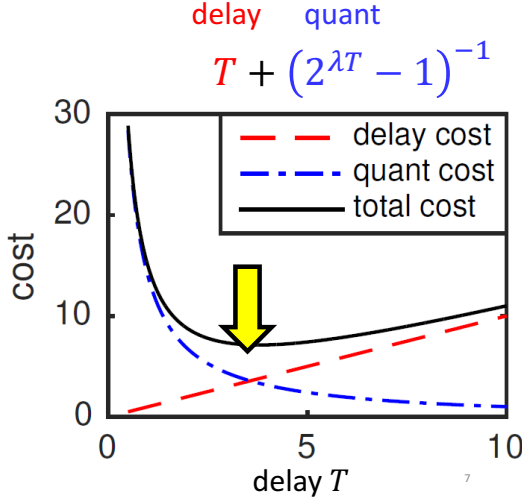
The remaining results do not depend crucially on any of the details of this model, and certainly not that $R = \lambda T$, but only that there is some monotonic tradeoff between R and T. This is the simplest possible starting point but more complex models including, say, bit losses and complex coding schemes, can be included at the cost of greater control complexity.

E. Delayed reaction versus advanced warning

We now revisit the delayed reaction and advanced warning regimes, previously considered solely from a control theoretic perspective, in light of the previous discussion and the derived speed/accuracy tradeoff. Fig. 6a shows the signaling
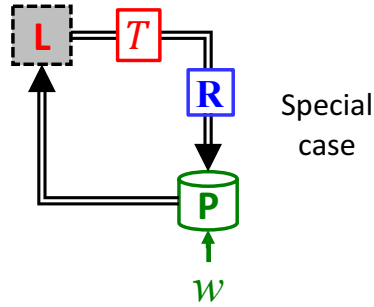
(a) The diagram of a simplified system to illustrate delay versus bandwidth tradeoffs.

$$\min_{T} \left[\max(0, T - \tau) + \left(2^{\lambda T} - 1\right)^{-1} \right]$$

(a) The delay, quantization, total cost when varying advanced warning $\tau > 0$, shown in the RHS of Figure 6b, or additional delay $\tau < 0$, shown in the LHS of Figure 6b.
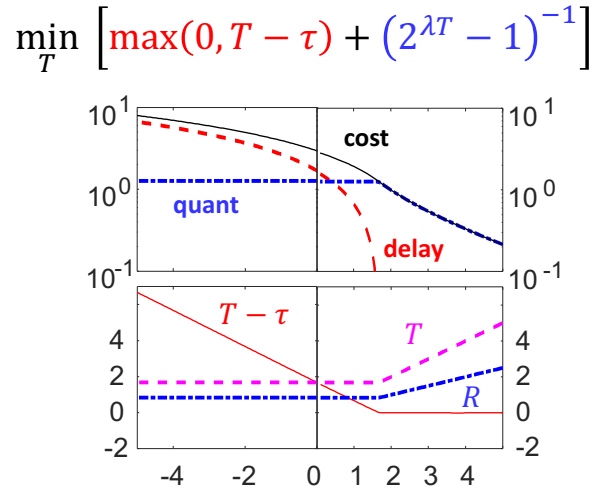
$$T + \left(2^{\lambda T} - 1\right)^{-1}$$

(b) The impact of speed/accuracy or delay/bandwidth tradeoffs on the simplified system in Fig 5a. The delay cost, quantization cost, the total cost is shown with varying delay T. Other system parameters are set to $\lambda = .1$, $a = 1$.

Fig. 5. The speed/accuracy tradeoff on system performance with net fixed warning $\tau = 0$ and varying signaling delay T.
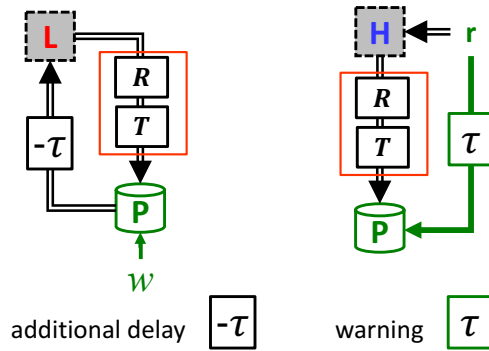

(b) The left hand side.

Fig. 6. Delayed vs warned system

delay T, the resulting total delay $T - \tau$, and the corresponding bandwidth $R = \lambda T$ that achieve the minimum total error as we vary the net delay τ. What results are clearly two distinct regimes with distinct physiology – these are conceptually consistent with the two qualitative regimes observed for expression (2).

(i) Delayed reaction: Here we assume that the fixed net warning satisfies $\tau < 0$. In this regime, Fig. 6a shows that the delay term increasingly dominates the total cost. Further this total cost can be much larger than the disturbance, and goes to infinity as delay increases or instability of the system increases (e.g., a small bump on a trail can cause my bike to flip over). Since the contributions due to the delay term dominate the overall cost, the optimal nerve is one that sacrifices bandwidth in order to minimize delay.

(ii) Advanced warning: As we increase the delay of the disturbance T_w, we eventually enter the fixed net warning $\tau \geq 0$ regime. In this regime, Fig. 6a shows that the quantization term increasingly dominates the total cost. Further this total cost can be made to approach zero if we allow R

to grow large – this further incurs no delay term penalty if we assume that our fixed net advanced warning τ grows correspondingly large (e.g., I can thread a needle in a different country given enough time). Since the contributions due to the quantization term dominate the overall cost, the optimal nerve is one that maximizes bandwidth by (approximately) matching the signaling delay T to the fixed net advanced warning τ of the system.

This model at least qualitatively matches both the cartoon in Fig. 3 and the empirical data on nerves as shown in Fig. 4b. Delayed reaction control will be implemented in nerves with low delay and bandwidth (e.g. sciatic and vestibular), and advanced warning uses high bandwidth and delay (e.g. optic and olfactory). In humans, less resources are devoted to olfaction than vision.

F. A simple layered model

Sensorimotor control must deal with a huge range of physical and temporal scales, and does so by combining extremely heterogeneous components in layered and distributed control architectures. We will next model a minimal combination of advance warning and delayed reactions to study their interaction. As a starting point in a task like

mountain bike riding, we abstract and simplify the layered sensorimotor system into two major feedback control loops shown in Fig. 7a: a planning loop that determines a movement path to follow a trail r using vision with advance warning τ, and a reflex loop that provides corrections against disturbances w from bumps for bike balance. (Multiple reflexes, including VOR for vision, will be present in reality but we'll just consider a single reflex system for simplicity.) Initially we'll assume that both loops act in common via a motor nerve with bandwidth R and delay $T < \tau$, and we apply the relevant theory from Section II-C, ultimately aiming to explain why the observed axon delay heterogeneity is beneficial.

To do so, we modify the dynamics (1) to reflect the diversity of the disturbances that affect bike (or eye movement, as illustrated in Fig. 7a). The resulting dynamics are given by

$$
\begin{aligned}
x(t+1) &= ax(t) + \mathcal{R}[u(t-T)] + v(t) \\
v(t) &= w(t) + r(t-\tau) \\
u(t) &= \mathcal{L}(x(0:t), w(0:t))) + \mathcal{H}(x(0:t), w(0:t))
\end{aligned}
\tag{7}
$$

where now the scalar state $x(t)$ is the error between the desired and actual trajectory. Further, the disturbance $v(t)$ is now composed of two terms: (i) a component $r(t-\tau)$ that is the trail observed with advance warning and (ii) a component $w(t)$ due to bumps in the trail or head motion. Notice that the object position $r(t-\tau)$ effects the error with a positive delay of τ, which is a simple model of our vision system's ability to see objects at a distance, thus providing us with advanced warning.

The control action $u(t)$ is generated by two nominally independent feedback loops, each having their own sensors, computing, and communication components, but for simplicity we will initially assume they act through a common motor nerve pathway with bandwidth R and delay T. We initially ignore other bandwidth and delay limits, which of course occur in every component. We assume that the change in object or trail position is bounded by $\|r\|_\infty \leq 1$ and bumps or head motion by $\|w\|_\infty \leq \delta$ (typically $\delta \ll 1$, but these constants are chosen out of convenience, and without loss of generality). The worst-case performance of the visual tracking system can then be quantified via the optimal cost to the following robust control problem:

$$
\underset{\mathcal{H}, \mathcal{L}}{\text{minimize}} \quad \underset{\|w\|_\infty \leq \delta, \|r\|_\infty \leq 1}{\sup} \|x\|_\infty \tag{8}
$$

A (shockingly) simple extension of the theory developed in Section II-C leads to the following optimal cost (Fig 7b):

$$
\left\{ \sum_{i=1}^{T} |a^{i-1}| + |a^T| \left(2^R - |a|\right)^{-1} \right\} \delta + \left(2^R - |a|\right)^{-1}.
$$

The constraint $R = \lambda T$ can then be added immediately. Intuitively, this result follows from the decomposition of problem (8) into two independent sub-problems, one for each of the feedback loops involving $\mathcal{H}$ and $\mathcal{L}$. A later paper in this collection discusses an experimental platform that
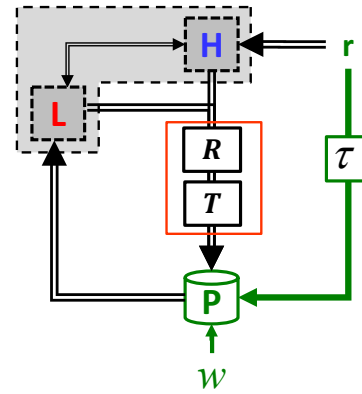
(a) Advanced planning $\mathcal{H}$ plus delayed reflex $\mathcal{L}$ control with shared motor nerve channel.

$$
\delta \left(\sum_{i=1}^{T} |a^{i-1}| + |a^T|(2^R - |a|)^{-1} \right) \quad +(2^R - |a|)^{-1}
$$

delay   **delay+quant**   **quant**

$$T < \tau$$
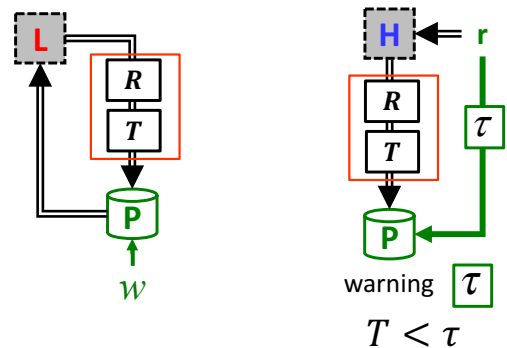
(b) The optimal cost of the system in Fig. 7a is almost the sum of the simpler delayed reflex system (left) and warned system (right).

Fig. 7. The layered model

implements planning and bumps in a video game with force feedback steering wheels, and confirms at least qualitatively that these formulas hold.

Fig 8 is similar but focuses on separate nerves connecting sensors of r to H versus w (and x) to L, and is arguably more directly related to the vision/VOR system. The resulting formulas still decompose as in Fig 7a with only slightly more complex expressions. Obviously a further combination of Figs 7a and 8 would be even more realistic but the formulas get fairly unreadable while giving little more insight.

G. Optimal resource allocation

The VOR system is a classic delayed reflex system whereas vision is human's preferred sensor for advanced planning.By comparing the two feedback loops in the visual system with the theory shown in Figs. 5 -8, we can observe why the their axon heterogeneity is beneficial. We have deliberately kept each of these Figures simple so that the consequences can be easily interpreted, but they must now be used in concert.

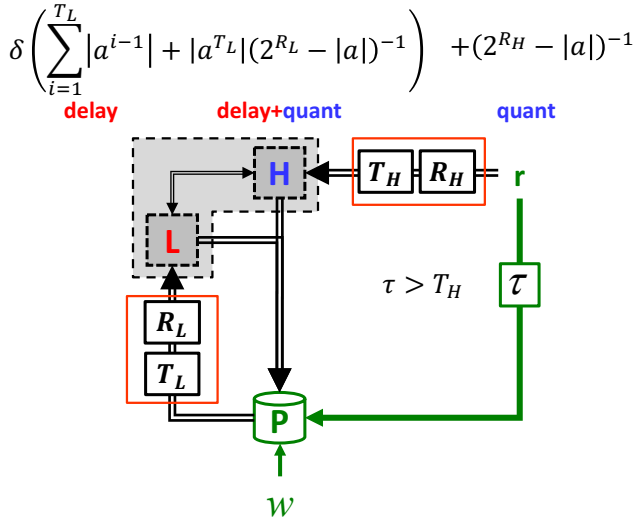*(i) The VOR system:* The feedback loop via vestibular

$$\delta\left(\sum_{i=1}^{T_L}|a^{i-1}| + |a^{T_L}|(2^{R_L} - |a|)^{-1}\right) + (2^{R_H} - |a|)^{-1}$$

Fig. 8. Planning and reflex control with separate nerves connecting sensors of r to H versus w (and x) to L, .

Fig. 9. The control of eye motion includes vision, the vestibular-ocular reflex (VOR), and the auxiliary optical system (AOS), but connects with several other systems. See text for explanation.

nerve and inner ear sensors attempt to cancel the head motion. Because our head motion can only be detected only after the motion occurs, we can assume $T_w \approx 0$ and consider it as a delayed system. From Section II-E, the delayed system achieves better performance for small signaling delay T, which is obtained by a few large and fast axons. This is also true of the sciatic nerve involved in reflexes in balance of the body.

(ii) The visual system: The feedback loop via optic nerve and the brain attempt to steer the eye toward relatively slow moving and distant objects, or trails and obstacles in the biking example, whose remote location allows for advanced planning and we can assume $\tau > 0$ and consider it as a warned system. From Section II-E, the warned system achieves better performance for large bandwidth, which is obtained by many small and slow axons.

This phenomena can be indeed observed in the real visual systems (Fig.4a). Specifically, the optic nerve has approximately 1M axons of mean diameter $0.64\mu m$ with CV $0.46\mu m$, while the 20K vestibular axons have mean diameter $2.88\mu m$ with CV 0.41, significantly larger and less numerous and slightly less variable. Note also that vision and VOR work so well together that we are normally unaware of the complex control system involved. Thus every aspect we have considered about VOR/vision fits our theory, at least qualitatively. To get a more quantitative test we will describe in a later paper a computer game aimed to mimic bumpy trail riding, and preliminary data taken from a prototype has a strikingly good fit to the theory.

H. Summary

The sensorimotor control are implemented via the nervous system which is limited by signaling speed and accuracy. In order to mitigate the deleterious effect of the delay and quantization, the sensorimotor control is implemented using layers of feedback loops, each with different optimal delay
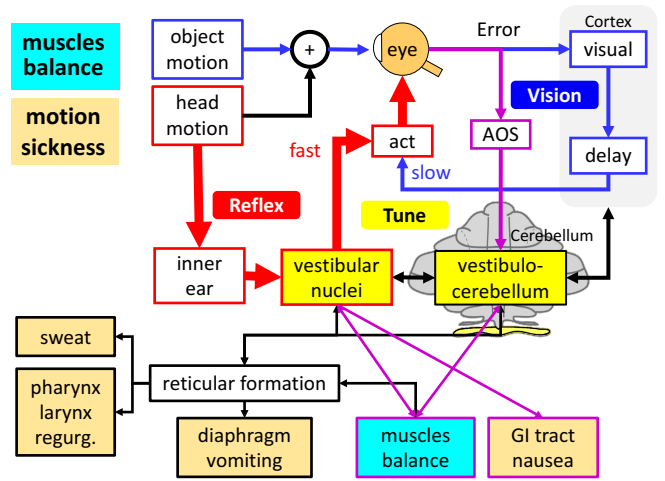
and bandwidth given the resource constraint. This explains why we observe different performance for similar task using different feedback loops, and why the sensory nerves are built to have size and number heterogeneity.

The vision/VOR system has some fascinating connections with the rest of the nervous system, as illustrated in Figure 9. While many details are unresolved, the connection with proprioception and balance is not surprising. An interesting experiment is to limit vision and proprioception by closing both eyes and balancing on one leg, which highlights the relative inaccuracy of the VOR system. Less obvious is the connection of the VOR tuning to mechanisms involved in motion sickness. The conventional wisdom is that the VOR tuning system is an exquisite sensor for poisons, and thus serves as an early warning system, and is connected with mechanisms to induce vomiting to expel poisons. This system is confused by modern technologies that create VOR tuning confusion, leading to motion sickness.

III. Stick Balancing

Balancing a stick on a person's hand is a popular case study in the sensorimotor control literature [5], [6], [7], but it also offers an attractive experiment to illustrate the theory we've shown here, and motivate some additional results, particularly regarding unstable zeros. Stick balancing can also be modeled as an inverted pendulum on a moving cart, which is familiar to engineers and scientists [8], [9]. Here we would like to use this model to illustrate a few interesting concepts in robust control theory that will shed some lights on the origin of oscillations in unstable system. A more complete discussions on robust control theory using the stick balancing as a case study is available at [10], which focuses on frequency domain analysis.

In this paper, we will consider an even more simplified stick balancing model. We understand that ordinary differential equations is not taught in high school, but this is the simplest model we have to explain important concepts related

to robust control theory and most importantly, the waterbed effect. In fact, because stick balancing is very easy to do using a standard extendable pointer, a lot of the theoretical results can be easily verified by the readers at home or office.

A. A Very Simple Balancing Model

Consider a very simple model of human stick balancing (see Fig. 10) where we assume that the stick mass is a lot smaller than the mass of the human balancing the stick and acceleration input. Then, the dynamics of the stick is characterized by

$$\ddot{x}(t) = u(t)$$
$$\ddot{x}(t)\cos\theta(t) + l\ddot{\theta}(t) + g\sin\theta(t) = 0$$
$$z(t) = x(t) + l_0\sin\theta(t) \qquad y(t) = z(t) + n(t) \qquad (9)$$

where x is the horizontal displacement of the hand, θ is the pendulum tilt angle from the vertical, u is the control input assumed to be the hand *acceleration*, y is the position measurement using the eye of z, the position of interest, n is sensor noise, g is the gravitational acceleration, l_0 is the fixation point, and l is the effective stick length (to the center of mass). Sensorimotor control delay due to signal transmissions and processing is represented as τ whereby visual processing is the major contributor [11], [12], [13]. (Note that vision and muscle motions are more quantized than noise, so a blend of this model with methods in this paper would be ideal, and will be explored in detail in future work.)

This model can be linearized and simplified into a linear dynamical system

$$\ddot{x}(t) = u(t) \qquad \ddot{x}(t) + l\ddot{\theta}(t) + g\theta(t) = 0$$
$$z(t) = x(t) + l_0\theta(t) \qquad y(t) = z(t) + n(t) \qquad (10)$$

where the pole and zero of the system is given in Table I.

The pole quantifies the (in)stability of the system, and it is related to the length of the stick. On the other hand, the zero quantifies the stability of the system when the output z is held at zero, and it is related to the fixation location l_0.
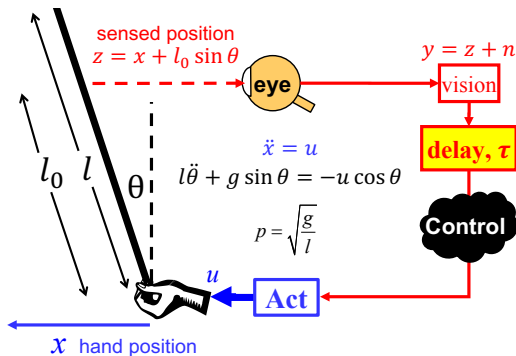

Fig. 10.   A schematic of balancing an inverted pendulum on one's palm.

B. What We Learn From This

This case study illustrates a remarkably rich set of ideas in robust control theory that can easily be demoed with a standard extendable pointer and also easily related to the theory developed here:

1) Balancing a short stick is hard: A shorter stick has larger unstable pole, and large unstable pole makes a system less robust and hard to control, especially with delays.

2) Low fixation point makes it hard to balance: A fixation point below the center of mass of a stick introduces an unstable zero into the system reducing the system robustness and causing the stick to become harder to balance. Peripheral vision must be occluded (e.g. by a cap or the other hand) to be sure that the measurement point is where the eye is looking.

3) Delay hurts system performance: Delay prevents a system from responding to disturbances quickly. As a result, the system becomes less robust to disturbances. This effect is exacerbated further if the system is unstable.

4) Waterbed effects are fundamental to system design, and oscillations are unavoidable as a result: As the closed loop system goes unstable, particularly for short sticks, the angle oscillates. This is a classic example of a waterbed effect and the oscillations can be reduced somewhat if the hand is allowed to slowly drift rather than have small steady state error. But as the length is shortened the oscillations become unavoidably larger until the stick crashes. The waterbed effect will be discussed more in detail in a later paper in this series.

5) The oscillations are stronger and crashes more frequent in the sagittal or lateral plane: The simple model above has one spatial direction whereas reality has three, but 3d models approximately decouple into 1d models, but the two main horizontal directions (vertical is less interesting) have different errors. The direction that the eye looks in has much larger errors because stereopsis is needed to determine range to the pointer, and this adds delay and additional quantization error.

6) It is difficult to balance a stick with one eye closed: With the loss of stereopsis, balancing becomes nearly impossible, unless the stick is held up so that the one eye can see underneath the tip, and see both directions of horizontal motion.

7) Unlike planning and reflex, stick balancing multiplexes poorly: Most biological processes are unstable, but are stabi-

Positions	Poles	Zeros
Upright		$\pm i\sqrt{\frac{g}{l_0-l}}$ if $l_0 > l$
	$0, \pm\sqrt{\frac{g}{l}}$	none if $l_0 = l$
		$\pm\sqrt{\frac{g}{l-l_0}}$ if $l_0 < l$
Downward		$\pm\sqrt{\frac{g}{l_0-l}}$ if $l_0 > l$
	$0, \pm i\sqrt{\frac{g}{l}}$	none if $l_0 = l$
		$\pm i\sqrt{\frac{g}{l-l_0}}$ if $l_0 < l$

TABLE I

POLES AND ZEROS OF 9.

lized with automatic and unconscious reflex and homeostatic control that multiplex perfectly. Here we are using vision to stabilize, which is convenient, and highlights the above features of the theory, but is not the natural use of vision. So balancing two sticks simultaneously in different hands is possible, but it is very fragile and crashes at relatively long stick lengths. Running while balancing a short stick is nearly impossible.

REFERENCES

[1] P. Sterling and S. Laughlin, *Principles of neural design*. MIT Press, 2015.

[2] A. J. Nagengast, D. A. Braun, and D. M. Wolpert, "Risk sensitivity in a motor task with speed-accuracy trade-off," *Journal of neurophysiology*, vol. 105, no. 6, pp. 2668–2674, 2011.

[3] M. A. Dahleh and I. J. Diaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

[4] W. Bialek, F. Rieke, R. van Steveninck, and D. Warland, "Spikes: Exploring the neural code (computational neuroscience)," 1999.

[5] N. P. Reeves, P. Pathak, J. M. Popovich, and V. Vijayanagar, "Limits in motor control bandwidth during stick balancing," *Journal of Neurophysiology*, vol. 109, no. 10, pp. 2523–2527, 2013.

[6] H. S. Harrison, D. G. Kelty-Stephen, D. V. Vaz, and C. F. Michaels, "Multiplicative-cascade dynamics in pole balancing," *Phys. Rev. E*, vol. 89, p. 060903, Jun 2014.

[7] J. G. Milton, "The delayed and noisy nervous system: implications for neural control," *Journal of Neural Engineering*, vol. 8, no. 6, p. 065005, Dec. 2011.

[8] A. D. Kuo, "An optimal state estimation model of sensory integration in human postural balance," *Journal of Neural Engineering*, vol. 2, no. 3, p. S235, Sep. 2005. [Online]. Available: http://iopscience.iop.org/1741-2552/2/3/S07

[9] O. Boubaker, "The inverted pendulum: A fundamental benchmark in control theory and robotics," in *Int. Conf. on Education and e-Learning Innovations (ICEELI)*, July 2012, pp. 1–6.

[10] Y. P. Leong and J. C. Doyle, "Understanding robust control theory via stick balancing," 2016.

[11] P. Lennie, "The physiological basis of variations in visual latency," *Vision Research*, vol. 21, no. 6, pp. 815–824, 1981.

[12] R. Nijhawan, "Visual prediction: Psychophysics and neurophysiology of compensation for time delays," *Behavioral and Brain Sciences*, vol. 31, no. 02, pp. 179–198, 2008.

[13] P. Cavanagh and P. Komi, "Electromechanical delay in human skeletal muscle under concentric and eccentric contractions," *European Journal of Applied Physiology and Occupational Physiology*, vol. 42, no. 3, pp. 159–163, 1979.

Balance, Oscillations, and Waterbed Effects

Yoke Peng Leong, John C. Doyle

Abstract— **Various waterbed effects are well known in the control community as a phenomena due to the conservation laws such as Bode's integral formula. This effect prevents a closed-loop system from having minimal oscillations at all frequencies without fundamentally changing the system design. This paper introduce a new formulation of the waterbed effect in the time domain that shows that good steady state disturbance rejection aggravates oscillations at higher frequencies. The result applies to all forms of single-input-single-output linear time invariant systems regardless of stability, state dimensions, and time continuity. Nonetheless, the derivation is simple and intuitive, and therefore, accessible to audiences who are not experts in control. Most importantly, we argue that oscillations commonly observed in many natural systems is the manifestation of the waterbed effect.**

I. INTRODUCTION

Traditionally, the *waterbed effect* refers to the inevitable effect on closed-loop system performance due to the conservation of Bode's integral [1]. Because the Bode's integral is conserved, the waterbed effect prevents a closed-loop system from having small or no oscillations at all frequencies without fundamentally changing the system design. In other words, if the closed-loop system performance is enhanced for a specific frequency range, then the closed-loop system performances at other frequency range will necessarily degrade.

In this paper, a more general form of waterbed effect is introduced via a time domain derivation and L_1 norm. We show that an improved closed-loop system performance for a type of disturbances will degrade the closed-loop system performance for other disturbances. This result is interesting for a few reasons. First, this result applies to any single-input-single-output linear time invariant system independent of stability, state dimensions, and time continuity. The only assumption is that the closed-loop impulse response of the system is positive for as long as the delay. This assumption simplifies the derivation, and we believe that it could potentially be removed in future iterations of this work. Second, in contrast to frequency domain analysis, derivations in time domain is more intuitive to the general audiences, and the math involved is a lot more accessible. Therefore, the derivation and the result are more friendly to people who do not have any advanced math backgrounds such as complex analysis. Lastly, this result also provides a simple explanation to the observation of oscillations in many natural systems [2], [3].

Many biological systems and "low level" physiological processes such as temperature regulation, blood oxygenation, and others, are unstable [2], [3], [4], and controlled automatically and unconsciously. Under specific circumstances,

they exhibit oscillation that, we argue, is an inevitable side effect of degrading system performance in dynamical systems with delay for enhancing system performance under another circumstances. In other words, to achieve efficiency under a certain case, we pay for oscillations in another case. In the past, the origin of oscillations in unstable systems have been poorly explained by experts in control, and it is poorly understood by researchers outside of the control community. In this paper, we hope to close this gap by providing a simple and intuitive explanation to the existence of seemingly cryptic oscillations in many natural systems, and shed some lights on the origin of oscillations in natural systems.

To illustrate the waterbed effect and the resulting oscillation, we will consider a case study in sensorimotor control – human stick balancing. Balancing a stick on a person's hand is a popular case study in the sensorimotor control literature [5], [6], [7], which can be modeled as an inverted pendulum on a moving cart that is familiar to engineers and scientists [8], [9]. A more complete discussions on robust control theory using the stick balancing as a case study is available at [10]. In this paper, we will consider an even more simplified stick balancing model. We understand that ordinary differential equation is not taught in high school, but this is the simplest model we have to explain important concepts in this paper.

Despite its simplicity, this case study is interesting because it highlights a few important points in system design [10]:

- *Balancing a short stick is hard.* A shorter stick has larger unstable pole, and large unstable pole makes a system less robust and harder to control.
- *Low fixation point makes it hard to balance.* Fixation point below the center of mass of a stick introduces an unstable zero into the system reducing the system robustness and causing the stick to become harder to balance.
- *Delay hurts system performance.* Delay prevents a system from responding to disturbances quickly. As a result, the system becomes less robust to disturbances. This effect is exacerbated further if the system is unstable.
- *Waterbed effects are fundamental to system design, and oscillations are unavoidable as a result.* An enhanced system performance for a particular disturbance will inevitably cause a degraded performance for other disturbances.

The last bullet point will be the focus of this paper. More detailed discussions of all the statements above are available in [10], which uses the frequency domain analysis.

A. Organization

The rest of this paper is organized as follows. Section II explains the waterbed effects and oscillations using a simple pictorial example. Section III formally defines the concepts in Section II. In Section IV, a simple stick balancing model is used to illustrate the main results. Lastly, Section V summarizes the main points of this paper.

B. Notation

The set of real numbers is represented as $\mathbb{R}$. The set of real $m \times n$ matrices is represented as $\mathbb{R}^{m \times n}$. A state x at time t is written as $x(t)$, and a trajectory x from time 0 to T is written as $x(0:T)$. The infinity norm of a trajectory x is given by $\|x\|_\infty = \sup_t x(t)$.

II. WATERBED EFFECTS AND OSCILLATIONS

The waterbed effect of bode integral suggests that an enhanced system performance for some input frequency will inevitably produce a degraded system performance for other frequencies. The proof of waterbed effect uses frequency domain analysis [1]. In this paper, we attempt to explain this phenomenon using only time domain. Before we proceed with the math in the next section, we would like to conceptually explain the waterbed effect and the resulting system oscillations using a few illustrative plots in Fig. 1.

In this work, the waterbed effect refers to an enhanced system performance for a specific disturbance model will produce a degraded system performance for other disturbance models. We quantifies system performance using the L_1 norm of the system output y

$$\max_{\|w\|_\infty \leq 1} \|y\|_\infty$$

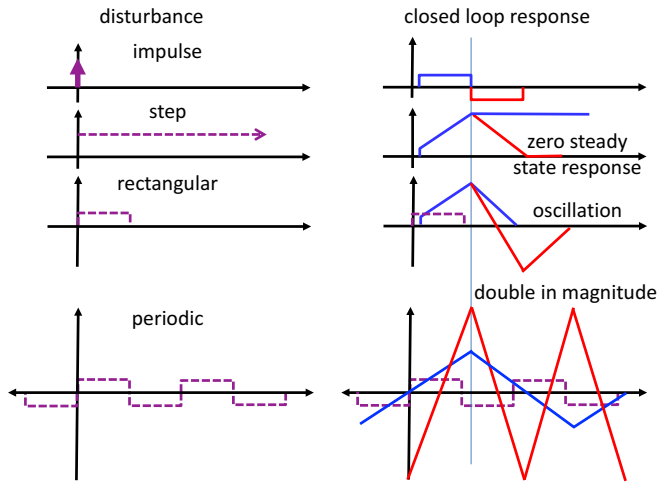where w represents the disturbance.

Fig. 1. A system's closed loop optimal L_1 responses. Blue curves represent the case when only optimal L_1 norm is required. Red curves represent the case when optimal L_1 norm and a zero steady state step response are required. The vertical light blue line denotes the delay. Oscillation occurs and maximum state magnitude doubles as a result of requiring zero steady state step response.

Suppose a system achieves the closed loop responses given by blue curves in Fig. 1 with respect to minimizing the L_1 norm. Using this controller, the step response achieves a non-zero steady state. However, in many applications, the step response is required to be zero at steady state.

Now, consider another case where we not only require for a minimal L_1 norm, but also zero steady state step response. The closed loop responses are shown in Fig. 1 as the red curves. Note that oscillation occurs under the rectangular disturbance model and the peak of the state's magnitude doubles for the periodic disturbance model. Therefore, oscillations and a degraded system performance at some disturbance models are consequences that one has to pay for to achieve a desired system performance at another disturbance model.

This simple explanation using Fig. 1 is very powerful because it is simple, it does not require the system dynamics to be either discrete or continuous in time, and it also does not require the system to be stable. In fact, the mathematical proof is exactly what is given by this figure, explained next.

III. FORMAL DISCUSSION

This section discusses the mathematical formulation that gives rise to the results in Section II. The formulation here is independent of the system stability or time continuity. Consider a single-input-single-output linear time invariant system S where in continuous time, S is (with some abuse of notation)

$$\dot{x}(t) = Ax(t) + B(u(t-T) + w(t))$$
$$y(t) = Cx(t)$$
$$u(t) = \mathcal{K}(x(0,t), w(0,t), u(0,t-1)),$$

and in discrete time, S is

$$x(t+1) = Ax(t) + B(u(t-T) + w(t))$$
$$y(t) = Cx(t)$$
$$u(t) = \mathcal{K}(x(0:t), w(0:t), u(0:t-1))$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}$ is the control action, $w(t) \in \mathbb{R}^n$ is the disturbance, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$, and T is the delay. Let $x(0) = 0$, and S has a positive open loop impulse response up to time T. The control action $u(t)$ is a function of the past states, and the goal of the controller is to reject disturbances $w(t)$.

Let $y_i(t)$ be the open loop impulse response for output $y(t)$ and $y_i(t) \geq 0$ for all $t \leq T$. Then, the step response of the state is given by

$$z_i(t) = \int_0^t y_i(\tau)d\tau$$

in continuous time and

$$z_i(t) = \sum_{\tau=0}^t y_i(\tau)$$

in discrete time.

A. Case 1: Optimal L_1 norm

First, let us consider the case when the system performance of S is optimal with respect to the following cost function

$$\min_{\mathcal{K}} \max_{\|w\|_\infty \leq 1} \|y\|_\infty \qquad (1)$$

which minimizes the L_1 norm of S.

The optimal impulse response $y_1(t)$ of output $y(t)$ in Case 1 is given by

$$y_1(t) = \begin{cases} y_i(t) & 0 < t \leq T \\ 0 & otherwise \end{cases}$$

where $y_i(t)$ is the open loop impulse response for output $y(t)$. This response corresponds to the blue curves in the first row of Fig. 1. Then, the optimal step response of output $y(t)$ is

$$z_1(t) = \begin{cases} 0 & t \leq 0 \\ z_i(t) & 0 < t \leq T \\ z_i(T) & otherwise \end{cases}$$

where $z_i(t)$ is the open loop step response for output $y(t)$. This response corresponds to the blue curve in the second row of Fig. 1.

Consider the following worst case disturbance model

$$w(t) = \begin{cases} 1 & t \in [0, \tau] \\ 0 & t \notin [0, \tau] \end{cases}. \qquad (2)$$

Then, the closed loop response would be

$$q_1(t) = \begin{cases} z_i(t) & 0 < t \leq T \\ z_i(T) - z_i(t - T) & T < t \leq 2T \\ 0 & otherwise \end{cases} \qquad (3)$$

whereby this response corresponds to the blue curve in the third row of Fig. 1.

More generally, the worst case disturbance model given by a periodic function

$$w(t) = (-1)^{\gamma(t)} \qquad (4)$$

where $\gamma(t) = \lfloor t/T \rfloor$ results in the closed loop response

$$p_1(t) = (-1)^{\gamma(t)}(z_i(t - \gamma(t)T) - z_i((\gamma(t)+1)T - t))$$

and

$$\min_{\mathcal{K}} \max_{\|w\|_\infty \leq 1} \|y\|_\infty = \|p_1\|_\infty = \sup_{t \in [0,T]} |z_i(t)| \qquad (5)$$

gives the optimal L_1 norm of the system S in Case 1. This response corresponds to the blue curve in the last row of Fig. 1.

If S is one dimensional system (i.e. $n = 1$) and $A = 1$ is marginally stable, the closed loop responses for Case 1 are exactly the blue curves in Fig. 1.

B. Case 2: Optimal L_1 norm + zero steady state step response

Now, consider the case the system performance is optimal with respect to (1) and also satisfies the zero steady state step response requirement.

Then, the optimal impulse response for output $y(t)$ in Case 2 needs to be

$$y_2(t) = \begin{cases} y_i(t) & 0 < t \leq T \\ -y_i(2T - t) & T < t \leq 2T \\ 0 & otherwise \end{cases}$$

and the optimal step response would then be

$$z_2(t) = \begin{cases} z_i(t) & 0 < t \leq T \\ z_i(2T - t) & T < t \leq 2T \\ 0 & otherwise \end{cases}$$

where zero steady state is achieved. These responses correspond to the red curves in the first and second rows of Fig. 1 respectively.

In this case, with the disturbance model (2), the closed loop response would be

$$q_2(t) = \begin{cases} z_i(t) & 0 < t \leq T \\ z_i(2T - t) - z_i(t - T) & T < t \leq 2T \\ -z_i(3T - t) & 2T < t \leq 3T \\ 0 & otherwise \end{cases} \qquad (6)$$

where by this response corresponds to the red curves in the third row of Fig. 1 respectively. Notice that (6) oscillates while (3) does not.

Furthermore, the closed loop response with the periodic disturbance (4) would be

$$p_2(t) = 2(-1)^{\gamma(t)}(z_i(t - \gamma(t)T) - z_i((\gamma(t)+1)T - t))$$

and

$$\min_{\mathcal{K}} \max_{\|w\|_\infty \leq 1} \|y\|_\infty = \|p_2\|_\infty = \sup_{t \in [0,T]} 2|z_i(t)| \qquad (7)$$

gives the optimal L_1 norm of the system S in Case 2. This response corresponds to the red curves in the last row of Fig. 1 respectively. Note that (7) is two times (5).

Similar to Case 1, if S is one dimensional (i.e. $n = 1$) and $A = 1$ is marginally stable, the closed loop responses for Case 2 are the red curves in Fig. 1.

C. Summary

Based on simple calculations above, we can observe two consequences of designing for a system to have zero steady state step response and to also achieve L_1 optimal system performance. First, the system performance degrades. The optimal L_1 norm doubles in this case. Hence, optimizing for one type of disturbance will inevitable results in degraded performance under other types of disturbance – the essential messages of waterbed effect. Second, oscillation occurs. This example shows a very important origin of oscillations in many systems. The presence of delay makes oscillation an

unavoidable side effect when the step response of the systems needs to be zero in steady state.

Furthermore, the formulation presented in this section is independent of whether the system dynamics is stable or unstable, discrete or continuous in time, and single state or more. This methodology is thus applicable to many applications unconstrained by the type of dynamical model available.

IV. STICK BALANCING

This section will apply the formulations from the previous section in the stick balancing case study discussed earlier. Consider a very simple model of human stick balancing (see Fig. 2) where we assume that the stick mass is a lot smaller than the mass of the human balancing the stick and acceleration input. Then, the linearized dynamics of the stick is characterized by

$$\ddot{x}(t) = u(t - \tau) + w(t)$$
$$\ddot{x}(t) + l\ddot{\theta}(t) + g\theta(t) = 0$$
$$y(t) = x(t) + l_0\theta(t) \tag{8}$$

where x is the horizontal displacement of the arm, θ is the pendulum tilt angle from the vertical, u is the control force, y is the position measurement using the eye of z, the position of interest, w is the actuation noise, g is the gravitational acceleration, l_0 is the fixation point, and l is the effective stick length. Sensorimotor control delay due to signal transmissions and processing is represented as τ whereby visual processing is the major contributor [11], [12], [13].

Simple algebraic manipulation of (8) gives

$$\dot{\bar{x}}(t) = A\bar{x}(t) + B(u(t - T) + w(t)) \tag{9}$$
$$y(t) = C\bar{x}(t)$$
$$\bar{x}(t) = \begin{bmatrix} x(t) & v(t) & \theta(t) & \omega(t) \end{bmatrix}^T$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{g}{l} & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -\frac{1}{l} \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 & l_0 & 0 \end{bmatrix}$$

where $v(t) = \dot{x}(t)$ and $\omega(t) = \dot{\theta}(t)$.

Fig. 2. A schematic of balancing an inverted pendulum on one's palm.

(a) Impulse
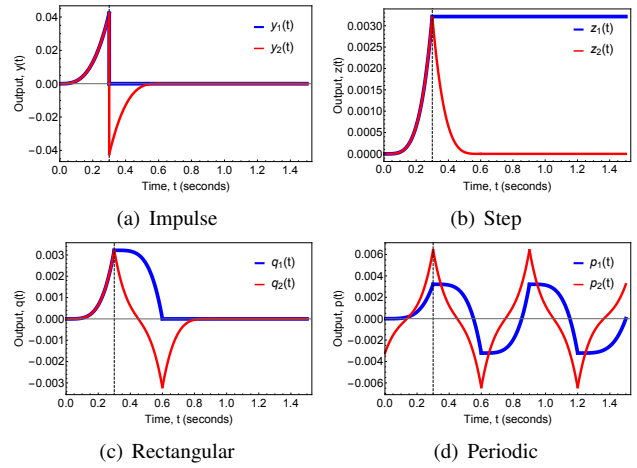
(b) Step

(c) Rectangular

(d) Periodic
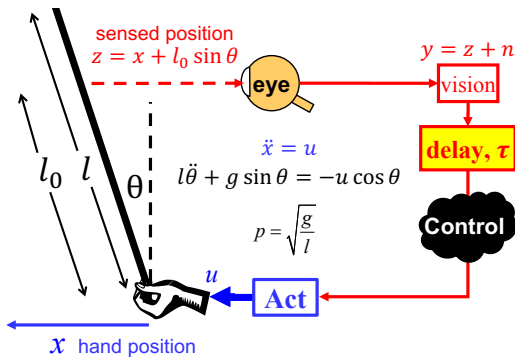
Fig. 3. Closed loop responses for system (8) under different disturbance model. Blue curves represent the case when only optimal L_1 norm is required. Red curves represent the case when optimal L_1 norm and a zero steady state step response are required. The vertical black dashed line denotes the delay at 300ms. Oscillation occurs in (c) and maximum state magnitude doubles from 0.003 to 0.006 in (d) as a result of requiring zero steady state step response.

Given (9) and let $g = 9.8 m/s^2$, $l = l_0 = 1m$ and $\tau = 0.3s$, the open loop impulse response is

$$y_i(t) = Ce^{At}B = t - 0.32\sin(3.13t)$$

and the open loop step response is

$$z_i(t) = \int_0^t Ce^{A(t-s)}B \, ds$$
$$= 0.5t^2 - 0.20\sin(1.57t)^2.$$

We can then compute the optimal closed loop responses given in Table I and Fig. 3. As expected, from Table I and Fig. 3, observe that the closed loop response in Case 2 oscillates under the rectangular disturbance model, and the infinity norm of the closed loop response in Case 2 is twice as much as the infinity norm of the closed loop response in Case 1 under the periodic disturbance model.

V. CONCLUSION

Using L_1 norm as the metric for system performance allows us to analyze system in the time domain which is more intuitive and realistic. This methodology inspires new ways of understanding waterbed effects, which is the "conservation law" in control, and the presence of oscillations in many natural systems. The result states that by improving closed-loop system performance for a specific disturbance, the closed-loop system performance for other disturbances will have to degrade and oscillation will occur as a result.

In the future, we plan to do the derivation more carefully, and generalize it to systems that are not necessarily single-input-single-output or have positive closed-loop impulse response for as long as the delay. We also encourage readers of this paper to further explore this new way of analysis in various other areas of control theory.

REFERENCES

[1] J. Doyle, B. Francis, and A. Tannenbaum, *Feedback Control Theory*. Macmillan Publishing Co., 1990.

[2] F. A. Chandra, G. Buzi, and J. C. Doyle, "Glycolytic oscillations and limits on robust efficiency," *Science*, vol. 333, no. 6039, pp. 187–192, 2011.

[3] A. Mitchell, P. Wei, and W. A. Lim, "Oscillatory stress stimulation uncovers an achilles heel of the yeast mapk signaling network," *Science*, vol. 350, no. 6266, pp. 1379–1383, 2015.

[4] N. Li, J. Cruz, C. S. Chien, S. Sojoudi, B. Recht, D. Stone, M. Csete, D. Bahmiller, and J. C. Doyle, "Robust efficiency and actuator saturation explain healthy heart rate control and variability," *Proc. Natl. Acad. Sci.*, vol. 111, no. 33, pp. E3476–E3485, 2014.

[5] N. P. Reeves, P. Pathak, J. M. Popovich, and V. Vijayanagar, "Limits in motor control bandwidth during stick balancing," *Journal of Neurophysiology*, vol. 109, no. 10, pp. 2523–2527, 2013.

[6] H. S. Harrison, D. G. Kelty-Stephen, D. V. Vaz, and C. F. Michaels, "Multiplicative-cascade dynamics in pole balancing," *Phys. Rev. E*, vol. 89, p. 060903, Jun 2014.

[7] J. G. Milton, "The delayed and noisy nervous system: implications for neural control," *Journal of Neural Engineering*, vol. 8, no. 6, p. 065005, Dec. 2011.

[8] A. D. Kuo, "An optimal state estimation model of sensory integration in human postural balance," *Journal of Neural Engineering*, vol. 2, no. 3, p. S235, Sep. 2005. [Online]. Available: http://iopscience.iop.org/1741-2552/2/3/S07

[9] O. Boubaker, "The inverted pendulum: A fundamental benchmark in control theory and robotics," in *Int. Conf. on Education and e-Learning Innovations (ICEELI)*, July 2012, pp. 1–6.

[10] Y. P. Leong and J. C. Doyle, "Understanding robust control theory via stick balancing," 2016.

[11] P. Lennie, "The physiological basis of variations in visual latency," *Vision Research*, vol. 21, no. 6, pp. 815–824, 1981.

[12] R. Nijhawan, "Visual prediction: Psychophysics and neurophysiology of compensation for time delays," *Behavioral and Brain Sciences*, vol. 31, no. 02, pp. 179–198, 2008.

[13] P. Cavanagh and P. Komi, "Electromechanical delay in human skeletal muscle under concentric and eccentric contractions," *European Journal of Applied Physiology and Occupational Physiology*, vol. 42, no. 3, pp. 159–163, 1979.

Disturbance	Closed Loop Response	
	Case 1	Case 2
Imp $\quad w(t) = \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases}$	$y_1(t) = \begin{cases} t - 0.32\sin(3.13t) & t \in [0, \tau] \\ 0 & t \notin [0, \tau] \end{cases}$	$y_2(t) = \begin{cases} t - 0.32\sin(3.13t) & t \in [0, \tau] \\ -0.6 + t + 0.32\sin(1.88 \\ \quad -3.13t) & t \in (\tau, 2\tau] \\ 0 & otherwise \end{cases}$
Step $\quad w(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$	$z_1(t) = \begin{cases} 0 & t \leq 0 \\ 0.5t^2 - 0.20\sin^2(1.57t) & t \in (0, \tau] \\ 0.0032 & t > \tau \end{cases}$	$z_2(t) = \begin{cases} 0.5t^2 - 0.20\sin^2(1.57t) & t \in [0, \tau] \\ 0.08 + (-0.6 + 0.5t)t \\ \quad +0.10\cos(1.88 - 3.13t) & t \in [\tau, 2\tau] \\ 0 & otherwise \end{cases}$
Rect $\quad w(t) = \begin{cases} 1 & t \in [0, \tau] \\ 0 & t \notin [0, \tau] \end{cases}$	$q_1(t) = \begin{cases} 0.5t^2 - 0.20\sin^2(1.57t) & t \in [0, \tau] \\ 0.06 + (0.3 - 0.5t)t \\ \quad -0.10\cos(0.94 - 3.13t) & t \in (\tau, 2\tau] \\ 0 & otherwise \end{cases}$	$q_2(t) = \begin{cases} 0.5t^2 - 0.20\sin^2(1.57t) & t \in [0, \tau] \\ 0.14 - 0.3t \\ \quad -0.10\cos(0.94 - 3.13t) \\ \quad +0.10\cos(1.88 - 3.13t) & t \in [\tau, 2\tau] \\ -0.30 + (0.9 - 0.5t)t \\ \quad -0.10\cos(2.82 - 3.13t) & t \in [2\tau, 3\tau] \\ 0 & otherwise \end{cases}$
Per $\quad w(t) = (-1)^{\gamma(t)}$	$p_1(t) = (-1)^{\gamma(t)}(z_i(t - \gamma(t)\tau)$ $\quad -z_i((\gamma(t) + 1)\tau - t))$	$p_2(t) = 2(-1)^{\gamma(t)}(z_i(t - \gamma(t)\tau)$ $\quad -z_i((\gamma(t) + 1)\tau - t))$

TABLE I

CLOSED LOOP RESPONSE FOR SYSTEM (8)

Complexity of ℓ_1 Optimal Controllers for Simple Systems

Emily Jensen, John Doyle

Abstract— The ℓ_1 optimal control framework is particularly convenient for introducing basic control concepts in a maximally accessible way. However, it is also known that even simple linear systems can have ℓ_1 optimal controllers that are of arbitrarily high dimension or are nonlinear, neither of which happens in comparable h_2 or h_∞ problems. Although this behavior in the ℓ_1 framework is known, the causes and consequences have yet to be resolved. Since these oddities are an obstacle to accessibility, this paper reconsiders the ℓ_1 "complex controller" with very simple and clear examples using stable and well behaved systems. Analytic formulas in this paper help to provide insight and clarity to this topic which has been mostly dormant for the past 20 years. Although we do not fully resolve the issues of existence of complex controllers and how to avoid them, the insight provided through simple examples is a first step resolving this annoying aspect of ℓ_1 optimal control.

I. INTRODUCTION

In this paper, we focus on ℓ_1 optimal control and its possible complexities, and show that in some cases the ℓ_1 optimal controller may be of arbitrarily high dimension. In particular, we will demonstrate through a simple example that this issue can occur even in stable systems, and for a range of objective functions and (stable) zero locations. Although zeros are not the only factor in causing high dimensionality, they do play a crucial role which is worth further examination. Papers in the past have highlighted this issue, but without clear explanations as to why. For instance, [1] has an example of arbitrarily high dimensional controllers, but this appeared to be caused by exotic conditions involving unstable poles and zeros, and there was no clear intuition to the behavior or its cause.

In our example, we minimize a convex combination of the output and control action for a stable system which is seemingly benign, having finite (in time) open loop impulse response (FIR). For certain weights in the objective function and a range of zero locations, we show that the optimal controller for this system can become arbitrarily complex. This behavior is important to understand, but hopefully it is ultimately not detrimental in most situations, as controller memory is relatively cheap and becomes another variable to be traded off. We provide analytic formulas of the optimal output, control action, and dimension of the optimal controller for this simple example. These formulas provide clarification to the cause of this strange high dimensionality and clearly demonstrate the trade off between the cost of control and cost of output.

Lastly, we use a well-known example from optimization, to construct a very simple linear problem in which the

E. Jensen, and J. C. Doyle are with the Control and Dynamical Systems, California Institute of Technology, Pasadena, CA, USA eljensen@caltech.edu, doyle@caltech.edu

optimal controller is nonlinear. Again, the possibility of this behavior is already known [2], but the exact conditions for when or why it occurs is lacking. We hope that studying this and other simple examples will help clarify where and why nonlinear controllers arise. This paper does not provide the final answer to the question of controller complexity, but it helps to demystify some of the unexplained behavior while addressing and clarifying key issues.

II. NOTATION AND PRELIMINARIES

Notation: For a vector $v \in \mathbb{R}^n$, let v_i denote the ith element of v. Then the ℓ_2 norm of v is defined as $||v||_2 = \sqrt{\sum_{i=1}^n v_i^2}$ and the ℓ_1 norm of a v is defined as $||v||_1 = \sum_{i=1}^n |v_i|$.
The ℓ_1 norm of a sequence of real numbers x is defined as $||x||_1 = \sum_{i=1}^\infty |x_i|$ where x_i is the ith entry in the sequence x.

Definition 1: For a discrete time single input single output system with input u and output y, the zero dynamics of the system are defined to be the dynamics the input u is required to follow in order to maintain an output of $y(t) = 0$ for all times t. $z \in \mathbb{R}$ is a zero of a system if it is a pole of the zero dynamics of that system. A zero at z is an unstable zero if it has magnitude $|z| > 1$.

III. A SIMPLE SYSTEM WITH A HIGH DIMENSIONAL CONTROLLER

A new example shows that even simple, well-behaved systems can have ℓ_1 optimal controllers which are dynamic and of arbitrarily high dimension. This example will motivate the definitions of a zero and zero dynamics and will provide insight as to how these affect the behavior of the system. Let

$$x(t+1) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t)$$
$$y(t) = \begin{pmatrix} \zeta^2 & -2\zeta \end{pmatrix} x(t) + u(t) \tag{1}$$

where $x(t) \in \mathbb{R}$ is the state at time t, $y(t) \in \mathbb{R}$ is the output at time t, and $u(t) \in \mathbb{R}$ is the control input at time t. This system is stable, with two poles at 0. We wish to find the optimal stabilizing control strategy which minimizes a convex combination of the output, $||y||_1$, and control action, $||u||_1$, of system (1), given the initial condition $x(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We assume that the control action at time t is full information, having access to the current and previous states and previous control actions. The optimization problem

can be written as:

$$\underset{u}{\text{minimize}} \quad (1-\gamma)||y||_1 + \gamma||u||_1$$

$$\text{subject to} \quad (u,y) \text{ satisfy dynamics (1)}$$

$$x(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2}$$

The special case when $\zeta = 1$ allows us to easily solve for the optimal controller analytically and gain insight to the form of the solution. For this choice of ζ, it can be seen that the open loop response of system (1) is: $y(0) = -2$, $y(1) = 1$, $y(t) = 0$ for all $t \geq 2$. Therefore, this system is stable and has an open loop response which remains zero for $t \geq 2$. Strangely, even with such a seemingly benign system, we will show that the optimal controller solving optimization problem (2) may be become arbitrarily complex for small enough values of γ. With $\zeta = 1$, explicitly iterating 1 gives:

$$y(0) = u(0) - 2$$
$$y(1) = 1 - 2u(0) + u(1) \tag{3}$$
$$y(t) = u(t-2) - 2u(t-1) + u(t), t \geq 2$$

and the zero dynamics for system (1) are given by:

$$u(t) = 2u(t-1) - u(t-2) \tag{4}$$

so the system has 2 zeros at 1. The optimization problem (2), with $\zeta = 1$, can be solved numerically using a linear program (LP). If we constrain the output to be zero after some finite time t_f we can rewrite $||y||_1$ as a finite sum:

$$||y||_1 = |u(0) - 2| + |1 - 2u(0) + u(1)| +$$
$$\sum_{t=2}^{t_f} |u(t-2) - 2u(t-1) + u(t)| \tag{5}$$

If $y(t) = 0$ for all $t \geq t_f$, then it must be that $u(t) = 0$ for all $t \geq t_f - 1$, in order to guarantee internal stability. Then optimization problem (2) with $\zeta = 1$ can be written as a linear program:

$$\min_u \Bigg((1-\gamma)\Big(|u(0) - 2| + |1 - 2u(0) + u(1)| +$$
$$\sum_{t=2}^{t_f} |u(t-2) - 2u(t-1) + u(t)| \Big) + \gamma \sum_{t=0}^{t_f} |u(t)| \Bigg)$$
$$\text{subject to} \quad y(t) = 0, t \geq t_f$$
$$u(t) = 0, t \geq t_f - 1 \tag{6}$$

Numerically solving the linear program (6) for various values of γ always gives a sparse output which is 0 except at two times: $t = 0$ and $t = T$, where T varies with γ. Ideally, we'd prove this directly rather than relying on numerical experiments, but very few problems can be solved analytically, so fixing this is not a priority. The main motivation for analytic results here is to get insight into the problem.

If we do add the additional constraint to (2) that the output is zero except at times $t = 0$ and $t = T$, then it is easy to solve this problem analytically. Indeed, a simple formula can be derived which describes the final time T in terms of the

weight γ while providing clear insights into the form of the solutions. The solution to the optimization problem (2) with this added constraint corresponds to an optimal controller which is $(T - 2)$ dimensional, where T is the time of the final nonzero output. Then, as the following theorem states, the weight on the output in the objective function can be chosen so that the optimal controller is of arbitrarily large dimension.

Theorem 2: Fix $\zeta = 1$ and additionally constrain (2) so that the output y must be 0 at all times except $t = 0$ and $t = T$. For any such T, we can choose a weight γ so that the optimal controller is of dimension $(T - 2)$.

To prove theorem (2), we will first prove a lemma which shows that the norms of the output and control action of the system can be written only in terms of one parameter - the final time T.

Lemma 3: The norms of the optimal control action and output solving (2) with the additional constraints as stated above are:

$$||u||_1 = \frac{T - 1}{2} \quad \text{and} \quad ||y||_1 = 1 + \frac{2}{T}.$$

Proof of lemma (3): If the output y is zero for all times except 0 and T and the control action is stabilizing, then $u(t) = 0$ for all $t > (T - 2)$, and using equation (3) the norms of the output and control action can be written as:

$$||y||_1 = |y(0)| + |y(T)| = |u(0) - 2| + |u(T-2)| \tag{7}$$

$$||u||_1 = \sum_{t=0}^{T-2} |u(t)| \tag{8}$$

Using the fact that the control action u is following the zero dynamics (4) whenever the output is zero, the general form of the control action at time t can be written as:

$$u(t) = (t+1)u(0) - t \tag{9}$$

To see why formula (9) is valid, note that if $y(1) = 0$, then $u(1) = 2u(0) - 1$, and if $y(2) = 0$ also, then $u(2) = 3u(0) - 2$. Using induction, it can be seen that in general, if $y(t) = 0$, then $u(t) = (t+1)u(0) - (t+2)$. Since $u(T-1) = 0$, equation (9) can be evaluated at time $t = (T-1)$ to provide an analytic formula for $u(0)$:

$$u(0) = \frac{T - 1}{T} \tag{10}$$

Combining equation (10) with the formula for $u(t)$ given by (9), we see that the control action at time $t \leq (T-1)$ can be written as:

$$u(t) = 1 - \frac{t+1}{T} \tag{11}$$

Using this general formula for $u(t)$ along with equation (3), we can write the general form of the output at times $t = 0$ and $t = T$:

$$y(0) = -1 - \frac{1}{T}$$
$$y(T) = \frac{1}{T} \tag{12}$$

Combining equation (12) with equation (7), the norm of the output can be written in a simple intuitive form: $||y||_1 = |y(0)| + |y(T)| = \left|-1 - \frac{1}{T}\right| + \left|\frac{1}{T}\right|$, which can be simplified to:

$$||y||_1 = 1 + \frac{2}{T} \tag{13}$$

Using equations (11) and (8), the norm of the control action can also be written in a simple form. Since $u(t) = 0$ for all $t > (T-2)$, $||u||_1 = \sum_{t=0}^{T-1} |u(t)| = \sum_{t=0}^{T-1} \left|1 - \frac{t+1}{T}\right|$. This can be simplified to:

$$||u||_1 = \frac{T-1}{2} \tag{14}$$

proving lemma (3) □.

Proof of theorem (2): Using lemma (3), the optimization problem (2), with $\zeta = 1$ and the additional constraint that the output is nonzero only at times $t = 0$ and $t = T$, can be rewritten so that the decision variable is the final time T:

$$\underset{T \in N}{\text{minimize}}(1 - \gamma)\left(1 + \frac{2}{T}\right) + \gamma\left(\frac{T-1}{2}\right) \tag{15}$$

The objective function in (15) is convex in the domain $T > 0$ for $T \in \mathbb{R}$ and γ fixed in $(0, 1)$, and global minimum is given by $T_\gamma = \sqrt{\frac{4(1-\gamma)}{\gamma}}$. Because of convexity, the optimal integer value of T must be one of the two closest integers to T_γ. Note that $T_\gamma \to \infty$ as $\gamma \to 0$; choosing $\gamma = \frac{4}{T^2+4}$ results in an optimal controller which is $(T-2)$ dimensional, proving theorem (2) □.

In proving theorem (2), we showed that there is a correspondence between the weight γ and the time of the final nonzero output T. The dimension of the optimal controller, as well as the norms of the optimal output and control action are easily parameterized in terms of the parameter T. Given any T, we can choose a corresponding weight $\gamma = \frac{4}{T^2+4}$. We will therefore parameterize the optimization problem in terms of the time T from now on, as it simplifies the analysis of the system.

This example clarifies the behavior of the system, and also shows the trade off between the cost of control, $||u||_1$, and the cost of output, $||y||_1$. When $T = 1$, the optimal control action is to apply no control, and the optimal output is the open loop response. This corresponds to an objective which only weights the control action. Conversely, when $T = \infty$, only the output is being weighted. This corresponds to a controller which zeros the output immediately and maintains a zero output at all times, so that the cost of control is infinite and the system loses internal stability. A value of $1 < T < \infty$ weights both the control action and the output. The corresponding optimal control action zeros the output at time $t = 1$, and then ramps downward following the zero dynamics until reaching zero at time $t = (T-1)$, after which it is held at zero. Zeroing the control action at time $t = (T-1)$ causes the output to be nonzero at time $t = T$.

The following figure demonstrates the output and control action for $T = 1$, $1 < T < \infty$, and $T = \infty$:
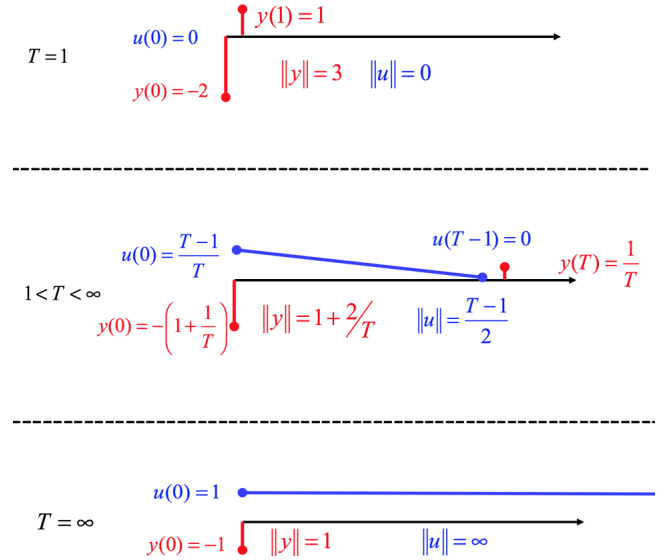

Fig. 1. Trajectory of optimal output and control action for various values of T. For $2 \leq T < \infty$ the optimal controller is of order $T - 2$ and thus the dimension of the controller can be made arbitrarily large.

The above figure gives a clear picture of the trade off that occurs as the parameter T is varied, allowing us to gain intuition as to why the dimension of the controller varies based on how much the cost of output and control action are weighted.

This problem can be generalized by allowing $\zeta \neq 1$. In this case, the zero dynamics are given by:

$$u(t) = 2\zeta u(t-1) - \zeta^2 u(t-2) \tag{16}$$

and the system has two zeros at ζ. With $\zeta \neq 1$, there does not exist an obvious analytic solution, and for most problems in the ℓ_1 optimal control framework this is indeed the case. The analytic solution for the simple case $\zeta = 1$ provided insight, but for problems even slightly more complicated we must rely on numerical solutions. Numerical solutions show that for a fixed zero location $\zeta \neq 1$, the dimension of the optimal controller increases as the weight on the output is increased. High dimensional controllers occur even for a range of stable zero locations, $|\zeta| < 1$, further demonstrating that although zero location has an impact, it is not this condition alone which causes controller complexity. Numerical solutions do not provide a specific condition for when scenarios of controller complexity occur, but they do provide evidence that complicated control laws arise in a variety of cases, and are not solely the cause of instability or unstable zero dynamics.

It is worth noting that although the dimension of the optimal controller can become arbitrarily large, for our exam-

ple, we can easily restrict the controller to be of any finite dimension we choose, and this is equivalent to increasing the weight on the control. This restriction to a dimension $(T-2)$ controller in optimization problem (2) gives a sparse output which zeros out in finite time T, and is exactly the optimal solution to this optimization problem with different appropriate choice of weight γ.

IV. A SIMPLE SYSTEM WITH A NONLINEAR CONTROLLER

As a final aside, in this section we recall a well-known example from optimization that shows that, even in simple cases, an ℓ_1 optimal controller may be nonlinear. Although this almost trivial example does not provide an explanation regarding exact conditions of when or why this behavior occurs, it does motivate further study of this topic. That ℓ_1 optimal controllers may be nonlinear is already known [2]. However, this behavior may seem even more odd than the case of dynamic, high-dimensional controllers, and it is unclear as to when this situation arises.

To start, consider the simple ℓ_2 optimization problem:

$$\min_x \left\| \begin{pmatrix} y - Ax \\ x \end{pmatrix} \right\|_2 \tag{17}$$

where $y \in \mathbb{R}^n$, $x \in \mathbb{R}^m$, and $A \in \mathbb{R}^{n \times m}$. It is well-known that there exists a simple closed-form solution to the above optimization problem of the form $x = My$ for some matrix M. In this simple ℓ_2 optimization problem it is clear that the optimal solution is simple and linear. However, even if we add more complexity to this problem- adding dynamics in continuous or discrete time, optimizing over the worst case or average error, or adding ℓ_2 regularization terms- the optimal solution remains minimally complex and linear as long as we restrict our choice of norm to ℓ_2. An almost trivial version of (17):

$$\min_x \left\{ (y - x)^2 + \lambda (x)^2 \right\} \tag{18}$$

has the familiar solution:

$$x = \frac{y}{1 + \lambda} \tag{19}$$

In contrast, even minimal departures from this ℓ_2 problem can have optimal solutions which are nonlinear, when use of the ℓ_1 norm is included. The simplest case is called ℓ_1 regularized denoising:

$$\min_x \left\{ (y - x)^2 + 2 |x| \right\} \tag{20}$$

This is a well-studied problem, known to have a nonlinear soft thresholding solution:

$$x = soft(y, 1) \tag{21}$$

where $soft(y, \lambda) = sgn(y) \max(0, |y| - \lambda)$.

Since such a small deviation from a pure ℓ_2 problem can lead to nonlinear solution, it should not be surprising that complex or nonlinear optimal controllers arise in the optimal ℓ_1 framework. This was a static example but it is easy to turn it into a very short FIR dynamic example.
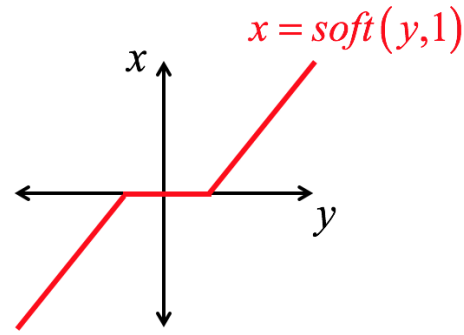

Fig. 2. Nonlinear solution to linear ℓ_1 regularized denoising.

V. CONCLUSIONS

It has been known for decades that ℓ_1 optimal controllers may be high dimensional and/or nonlinear; examples in the previous two sections can hopefully help to clarify the understanding of these odd behaviors. One reason these complex behaviors of ℓ_1 optimal control seem surprising, is the contrast from the well known h_2 and h_∞ problems, for which the optimal full information controllers are static and linear, and output feedback controllers are the same state dimension as the plant. It remains unclear, however, exactly when these issues arise, and what if any are their practical significance. In the context of this series of tutorial papers, seemingly gratuitously complex controllers are primarily a potential obstacle to accessibility. If we are to "revive" the ℓ_1 framework then these issues will need broader and deeper study as part of a more general treatment of scalability to more complex problems.

REFERENCES

[1] I. J. Diaz-Bobillo and M. A. Dahleh, "State feedback ℓ_1 optimal controllers can be dynamic," in *Systems & Control Letters, Volume 19, Issue 2*, 1992, pp. 87–93.

[2] A. A. Stoorvogel, "Nonlinear l_1 optimal controllers for linear systems," in *IEEE Transactions on Automatic Control, Vol. 40 No. 4*, 1995, pp. 694–696.

An experimental video game platform for sensorimotor control

Adam Dai, Emily Jensen, Dimitar Ho, John Doyle

Abstract— The aim of this paper is to describe a low-cost experimental video-game platform that can potentially be used to explore the theory presented in the earlier papers in this tutorial collection. We want to quantify if the theory correctly predicts planning performance as a function of advanced warning, and then how close to optimal are humans in multiplexing planning and reflex tasks. Both are motivated by examples such as riding a mountain bike down a steep twisting rocky trail. To that end we've created a game that uses a standard force feedback steering wheel to generate "bumps" together with a screen image that shows a "trail" to follow as well as the deviation from the trail both to bumps and poor planning. The game can have bumps and/or a trail individually or together and the advanced warning of the trail can be adjusted to any value. Preliminary α-testing of the platform has generated data that is strikingly consistent with the theory, which suggests it is suitable for more serious experimental studies. Everything is low cost and open source so should be easily replicated.

I. INTRODUCTION

This paper describes a computer game aimed to illustrate and explore the theory from [1], and figures and equations from [1] will be denoted by a ′. In this section we will describe the setup of our "driving" simulation, which is intended to capture some essential features of riding a mountain bike down a steep, turning, rocky trail as in Fig. 1′, but with no danger of real crashes. We will then explain the experiments and compare them with the theory [1].

The main goal of the driving game is to test how well a player can drive along a given trail in the virtual environment with possibly limited advanced warning while being exposed to force disturbances. As shown in the screen capture of the game (1), the player sees a red vertical line which is following a white trail. As the white trail is moving down the screen, the player has the task to move the blue vertical line such that it tracks the red vertical line along the large horizontal line. The blue line is the wheel angle of a Logitech G27 Racing Wheel connected via USB to a computer.

The G27 steering wheel has force feedback capability, i.e., can be programmed to apply force, thus causing the wheel to move independent of user input, and thus the user must supply a counter force to have a small error. This allows us to simulate disturbances, analogous to the ones one might experience from bumps in the trail as felt in handlebars. This is to add an element of reflexive control to multiplex with the visual task of following the game's path.

For every game run we can program different force disturbances, change the amount of advance warning the player receives of the white trail and it's difficulty. Overall, the main functionality of our driving game setup is to enable us to design experiments/game runs which allow us to analyze how the tracking error (difference between red and player's vertical line) changes when we expose the player to different disturbances, look-ahead windows and difficulty of the trail. In the next section we will discuss the experiments performed in this paper and show how they are aimed to illustrate the main results of [1].
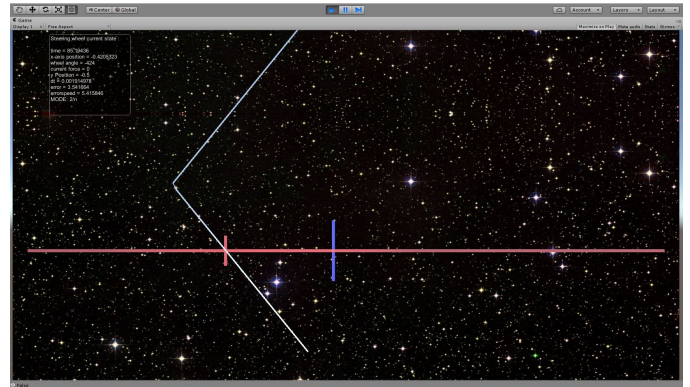

Fig. 1: Screen capture of running experiment

II. EXPERIMENTS

A. Bumps, Trails, Bump and Trails

This experiment aims to give us insight into how the player's performance changes during three scenarios:

1) **"Bumps":** Track a constant trail (the red trail line is not moving) subject to force disturbances on the wheel:
2) **"Trails":** Track moving trail while having look-ahead but no forces:
3) **"Trail with Bumps":** Track a moving trail while there are force disturbances on the wheel:

The player spends 60 seconds in each of the above scenarios in the order (Bumps, Trail, Trails with Bumps) with 5 seconds rest preceding each scenario. Furthermore the game is setup to give the player a standard look-ahead window/advanced warning of 1 second, which has been chosen such that the player doesn't experience an additional handicap from lacking advanced Warning. The disturbance and the trail the player experiences during the isolated phases "Bumps" and "Trails" are the same ones as in the combined "Trails with Bumps" phase. This is to cleanly compare the players performance when facing the separate tasks versus when he has to multiplex both of them.

As a disturbance we initially chose a binary random signal, whose amplitude is the maximum possible torque the motor of the steering wheel can exert. Furthermore, the switch decisions between max positive and negative torque is made every 100ms. A similar random binary switching is

L_1-Norm Error	E_b	E_t	$E_b + E_t$	E_{bt}
Random Binary Dist.	0.182	0.155	0.337	0.288
Adversarial Feedback Dist.	0.178	0.152	0.330	0.306

TABLE I: Comparison of L_1-Tracking Error Norm: Different Disturbance Models vs. Different Experiment Phases.

L_∞-Norm Error	E_b	E_t	$E_b + E_t$	E_{bt}
Random Binary Dist.	0.39	0.86	1.25	1.23
Adversarial Feedback Dist.	0.34	0.61	0.95	1.17

TABLE II: Comparison of L_∞-Tracking error norm: Different Disturbance Models vs. Different Experiment Phases

of the trail derivative. More specifically, the trail travels at a constant velocity but randomly switches its directions such that it always stays in the screen range comfortably visible by the player. We adjusted the velocity of the trail on the screen such that the required steering wheel turning rate is approximately $75°/\text{sec}$.

Figure 2 illustrates 5 second snapshots of our experiment design and corresponding tracking errors of a typical experiment run by one of the authors (Ho). The trail presented to the player is painted in black and is depicted in Sub-Figures (2a),(2b) and (2c), where each of the plots considers the "Bumps", "Trails" and "Trails with Bumps" phase of the experiment, respectively. The players trajectory is presented in the same plots with the colors blue, red and green, indicating the three different phases of the experiment in each of the Sub-Figures. In Addition, Sub-Figures (2d)-(2f) correspond to the same snapshots of the experiment as (2a)- (2c), but plotting the tracking error of the player, i.e. difference between trail and player's. The colors red, blue and green indicate in the same way the tracking error corresponding to the particular phase of the experiment. Finally Sub-Figure (2g) overlays Sub-Figures (2d)- (2f) to display the comparison of the player's tracking performance in the different experiment phases.
The purple-empty and orange-filled stem plots in Figures (2a) - (2g) indicate the timing and direction of the disturbances and trail curve changes, respectively. Note that both the wheel forces and the trail rates are square waves, and the stems indicate where these square waves switch (so are derivatives of the forces and rates).

Comparing the stem plots of (2a), (2b) and (2d), (2e) to (2c) and (2f), notice that the experimental setup of the "Trails with Bumps" phase is a superposition of the disturbance and trail characteristic of the individual "Bumps" and "Trails" phases. Furthermore notice that although the player's tracking error of the "Trails with Bumps" phase appears bigger than the error of the separate "Bumps" and "Trails" it still is not very much larger than the individual error signals. This hints that humans seem to be capable of multiplexing the tasks of rejecting a force disturbance and visually tracking a trail fairly well. This phenomenon can be explained with the recent results of [1] and will be discussed in the next section. In Addition, we implemented an adversarial feedback model of the force disturbance and running experiments showed that we received error signals for the "Trail with Bumps" scenario which demonstrated still good multiplexing performance of the player.

To quantify this observation, table (I) and (II) present the l_1 and l_∞-error-norm of a 50 second window, taken from the middle of each experimental phase corresponding to the two different disturbance models. While the columns E_b, E_t, E_{tb} denote the phases "Bumps", "Trails" and "Trails with Bumps" of the experiment, the column $E_b + E_t$ simply adds up the error-norms of the first two. Inspecting this table, we observe that for the binary random disturbance, the error-norms of the combined phase "Trail with Bumps" is close if not even smaller then the sum $E_b + E_t$. We also notice from this data, that the error-norm of the "Trails with Bumps" phase and their relation to the other errors depends on the disturbance model. More specifically, the experiments with our naive adversarial disturbance show that the errors of the combined task show a bigger ratio of $E_{bt}T$ to $E_b + E_t$ with respect to the binary random model. This also opens up potentially interesting follow-up question of what is the worst-possible disturbance to maximize that ratio.

B. Varying Advanced Warning

The purpose of this experiment is to evaluate how the length of the look-ahead window, i.e advanced warning affects the player's performance to track a trail without being exposed to disturbances. In the words of the previous section, this experiment requires to track the randomly generated trail from the "Trail" phase without "Bumps" but instead with different advanced Warning intervals.

The experiment is 400 seconds long and consists of one continuous "Trail" Phase which reduces every 30 seconds the look-ahead window. Starting off with the first two look-ahead windows being 1sec and 0.75 secs, the game then starts to decrease the window from 0.5sec in 0.1 sec steps down to -0.5 sec, where negative look-ahead means that we are displaying only past information of the trail.

An evolution of the absolute tracking error of the player as the game progresses is depicted in Figure (3). The plot displays only the mid 20 seconds of each of the 30 second intervals to neglect the effect of the player adjusting to the new look-ahead window. The progression of the error in the blocks of constant look-ahead demonstrate the intuitive effect, that the player looses performance with a decrease of advanced Warning.

To quantify that effect more in detail, we evaluate the l_1, l_2 and l_∞ error norm for every 20 second long window corresponding to an advanced warning level. Summarizing these calculations in a plot gives us Figure (4), which demonstrates how the players error-norm does not change
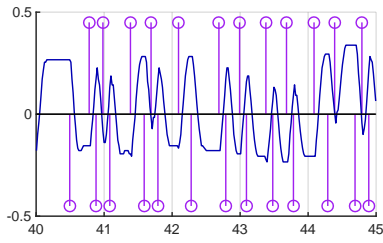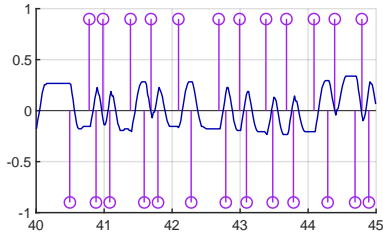
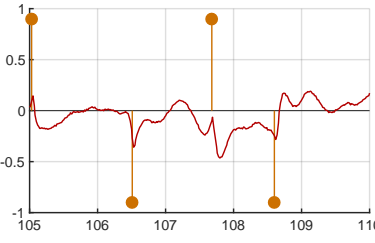(a) "Bump" Phase: Trail Pos. (black), Player Pos. (blue) vs. Game Time in sec
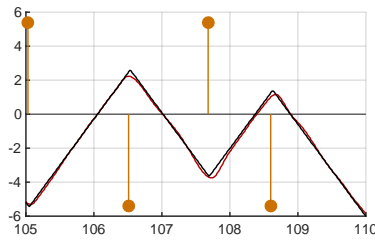
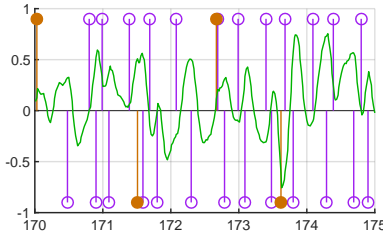(b) "Trail" Phase: Trail Pos. (black) vs. Player Pos. (red) vs. Game Time in sec
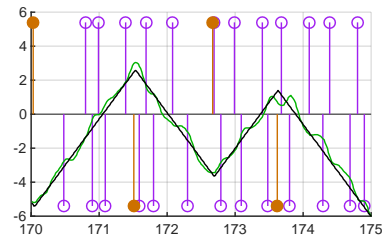
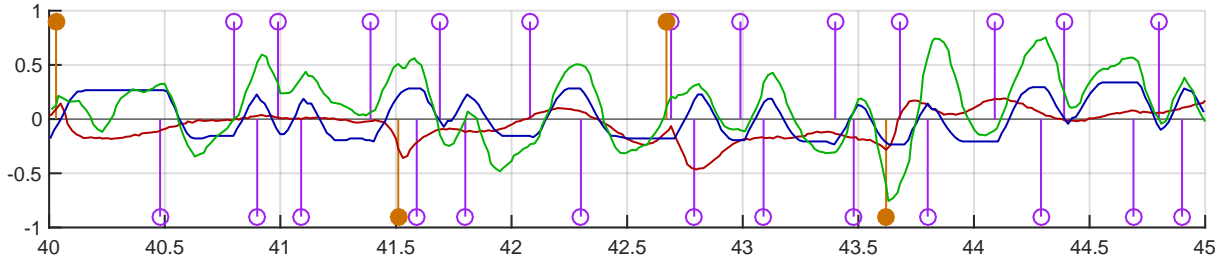(c) "Trail with Bumps" Phase: Trail Pos. (black), Player Pos. (gree) vs. Game Time in sec

(d) "Bump" Phase: Tracking Error (blue) vs. Game Time in sec

(e) "Trail" Phase: Tracking Error (red) vs. Game Time in sec

(f) "Trail with Bumps" Phase: Tracking Error (green) vs. Game Time in sec

(g) Overlayed Tracking Errors of Phases "Bump" (blue), "Trail" (red) and "Trails with Bumps" (green)

Fig. 2: Data-Snapshot of combined "Bumps" (blue) , "Trails" (red) and "Trails with Bumps" (green) Experiment. Orange/Purple Stem: Disturbance/Trail Curve Occurrence and Direction.
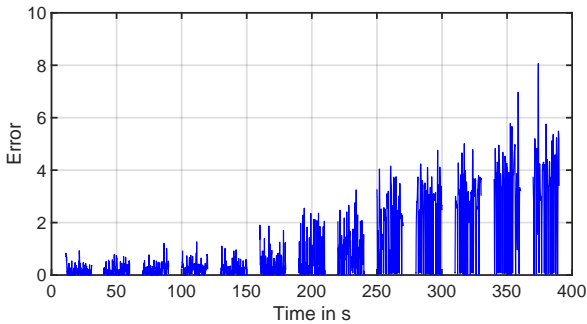
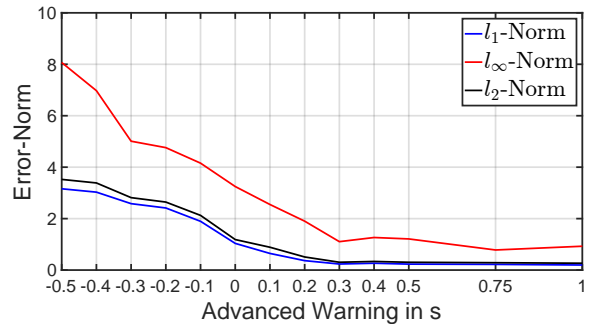Fig. 3: Absolute Error Evolution during Game

Fig. 4: Error Norms vs. Given Advanced Warning Interval

until the advanced warning reaches 0.3 seconds and then increases in approximately linear fashion.

III. RELATION TO THEORY

The first comparison is between the data in Fig 2, and Fig 7′ from [1]. The experiment has very high damping and minimal inertia, so is roughly first order and neutrally

stable, so is roughly like the model in Fig 7′ with $a = 1$ though further modeling and system ID is needed to make this more precise. The disturbance r would correspond to the trail changes, and w the wheel forces. As in Fig 7′, the player must control errors to both via a common actuator of arms and hands on the wheel. The theory predicts that these planning and reflex tasks have combined error roughly equal

to the sum of the individual errors, which is the case for the data in Fig 2. This is encouraging that the platform has potential, but obviously more modeling and experimentation is needed to draw further conclusions.

The data in Fig 4 can be compared more directly to Fig 6′ but for a fixed nerve:

$$\max\left[0, T - \tau\right] + (2^{\lambda T} - 1)^{-1}. \tag{1}$$

where τ is the advanced warning, which is the x axis in Fig 4. The theory predicts that when $\tau < T$ the error will increase linearly with τ, where T is the time delay of the visual feedback system. This fits the data in Fig 6′ strikingly well with $T \approx 250ms$, which is typical for humans. Note that this is one run and not average over multiple runs. Again, this is encourage that the platform could be used for more serious studies. This game will be demoed in a booth at CDC, so come try it yourself.

REFERENCES

[1] Y. Nakahira, Y. P. Leong, N. Matni, and J. C. Doyle, "Teaching controls: Neuroscience motivation," in *Conference on Decision and Control*, 2016.