Universal Laws and Architectures (普遍法则与结构)

DOYLE C John (道耀)[1], NAKAHIRA Yorie (依惠)[2]

1. California Institute of Technology
E-mail: doyle@caltech.edu

2. California Institute of Technology
E-mail: ynakahir@caltech.edu

Abstract: This talk, the opening plenary, aims to accessibly describe progress on a theory of network architecture relevant to neuroscience, biology, medicine, and technology, particularly Software Defined Networks and Network Function Virtualization (SDN/NFV) and cyberphysical systems. Key ideas are motivated by familiar examples from neuroscience, including live demos using audience brains, and compared with examples from technology and biology. A major challenge is to both broaden the applications and impact of control theory and make it more accessible to larger audiences. This paper briefly summarizes some background with links to additional material online, and particularly a series of videos for which this talk can be viewed as a short trailer.

Key Words: universal laws, theory of architecture

1 Introduction

This talk will review recent research aimed at developing a more "unified" theory for complex networks motivated by and drawing lessons from neuroscience [1] [2], cell biology [3], medical physiology [4] [5], and multiscale physics [6] [7] [8]. Most of these case studies mixed "off the shelf" theory and algorithms with somewhat hand-crafted upgrades, and were published in domain journals, but in parallel we have been developing new theory specifically aimed at complex network architecture, control, and evolution. This theory involves several elements: hard limits, tradeoffs, and constraints on achievable robust, efficient performance ("laws"), the organizing principles that succeed or fail in achieving them ("architectures" and protocols), the resulting high variability data and "robust yet fragile" behavior observed in real systems and case studies (behavior, data, statistics), the processes by which systems adapt and evolve (variation, selection, design, layering), and their unavoidable fragilities (hijacking, parasites, predation, zombies). A final crucial element is scalable algorithms to allow study and design of complex networks using this theory.

A ubiquitous tradeoff is between robustness and efficiency as illustrated in the cartoon in Fig. 1. Humans are roughly 4 times more efficient at long distance running than chimps, our closest primate relatives, but are much slower, weaker, and more fragile. Adding a bicycle improves efficiency but further increases fragilities to even small disturbances. Extending control theory to make robust efficiency tradeoffs precise and rigorous have been crucial to resolving long-standing mysteries regarding the origins and nature of glycolytic oscillations [3], heart rate variability [5], and turbulence [7]. Of course, robustness and efficiency are just two "principal components" in a high dimensional cyber-physical design space, and each has many additional dimensions. For example, within control systems, robustness often contains further speed accuracy tradeoffs (SAT), which can be nicely illustrated with examples from vision, cognition, and sensorimotor control (e.g. Fitts' Law). This will be the starting point in the talk and in this background paper.

This work is supported by NSF and AFOSR.

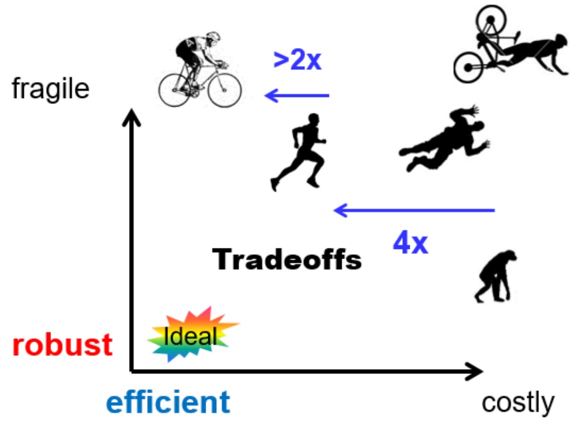


Fig. 1: Robustness efficiency tradeoffs in primate locomotion.

2 Case study in sensorimotor control

Human sensorimotor control systems have particularly interesting tradeoffs between achieving low delay versus low signaling error in nervous system communication. This phenomenon can be illustrated using the experiments described in Fig. 2. First, place your hand in front of your face, then move your hand horizontally back and forth at increasing frequency until the lines in your palm start to blur. Second, hold your hand still and shake your head (in a "no" pattern) at increasing frequencies until blurring occurs. The blurring for hand motion occurs at just a few Hz., whereas the blurring for head motion normally occurs at much higher frequency than for the hand motion. While the actuators are the eye muscles in both cases, there is a large delay heterogeneity in the systems that control the eye muscles depending on the source of the disturbance and sensors involved [9].

Fig. 3 is a minimal cartoon block diagram of the physiology underlying this experiment. For both hand and head motion, the eye muscles are the actuators that move the eye to compensate, and keep the image clear and in the high resolution central fovea of the eye. Object motion is compensated for by the visual system, which has a relatively slow but high resolution optic nerve and large amounts of cortical compu-

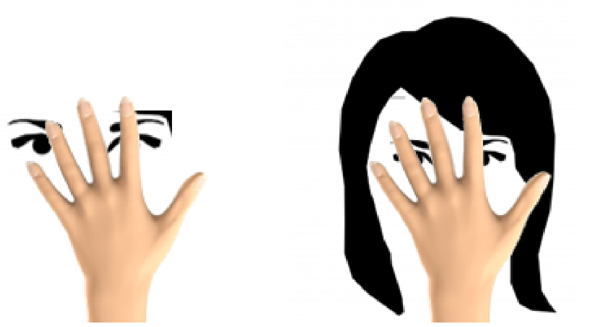


Fig. 2: Experiments that demonstrate delay heterogeneity in different feedback loops in the visual system.

tation before actuating the eye muscles, resulting in large delays (greater than $200ms$) and relatively poor performance. In constrast, the vestibular-ocular reflex (VOR) system directly measures head motion using sensors in the inner ear, and is connected directly to the eye muscles via the fast but low resolution vestibular nerve, with much less delay (approximately $10ms$) and thus much better performance. The gains on these two loops must match, and there is a gain module on the VOR system that is tuned by the cerebellum using the auxillary optical system (AOS) on timescales of hours to days.

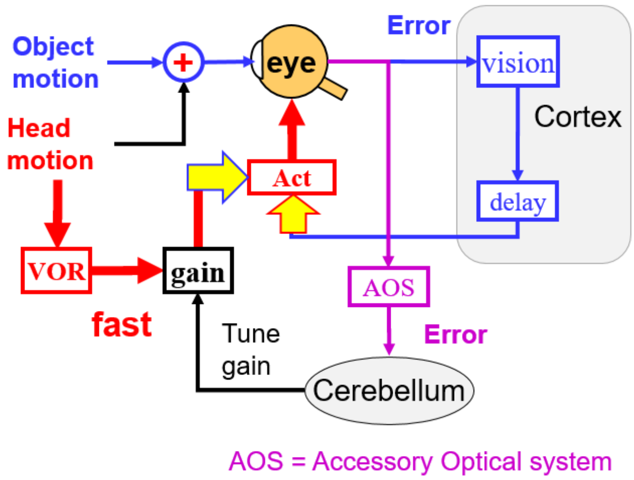
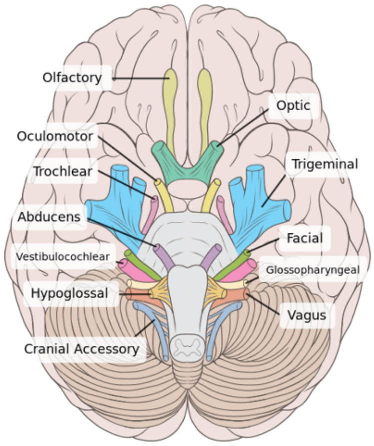


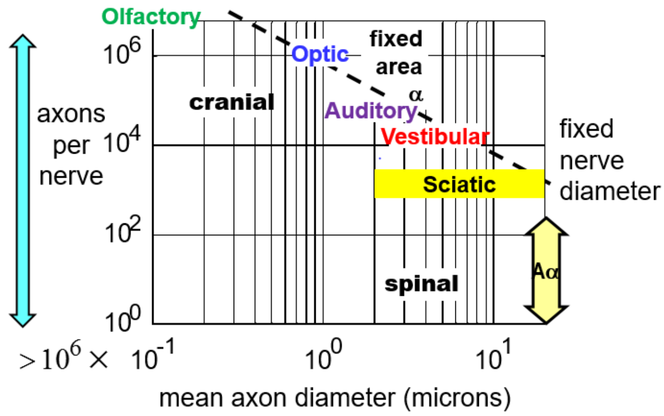
Fig. 3: Control of eye motion including vision, the vestibular-ocular reflex (VOR), and the auxilary optical system (AOS)

Thus the delays in just Fig. 3 vary by at least six orders of magnitude, with even greater heterogeneity in other parts of the neuroendocrine control system. Fig. 4 shows comparable heterogeneity in the composition of several cranial and spinal (sciatic) nerves. Nerves are discrete bundles of axons that connect collections of neurons over long distances. While the nerves in Fig. 4 all have roughly the same cross sectional area, they differ widely in the size and number of axons [10, 11].

This heterogeneity in delay and composition in performance and physiology motivates a minimal model to explain the connection between them and explore the design principles involved, as well as other issues suggest by Fig. 3,



(a) Illustration of existing nerves.



(b) Size and number for different types nerves [10].

Fig. 4: Composition of cranial and spinal nerves.

and particularly layered, distributed, and localized control. We will also briefly consider scalability of design, computation, and implementation to the large scale and extreme complexity of the full nervous system, and related scalability in technological systems. In Section 2.1 we introduce the simplest possible theory that studies the impact of delay and signaling error, and the insights it gives into the visual system in Section 2.2, which summarizes the work [2]. Its content is extended in the forthcoming papers [12, 13] from theoretical point of view and [14] from experimental point of view. Specifically, the work [12] shows a parallel result for linear quadratic cost function, and [13] studies the effect of zeros on system performance. The effect of zeros are also demonstrated in stick balancing, a popular case study in both sensorimotor control and engineering literature [14]).

2.1 Impact of delay and quantization on robust performance

The impact of delay and quantization on achievable robust performance have been widely studied (e.g. [15][2]). As a minimal toy model, let the scalar $x(t)$ be the difference between the true state and the desired state with dynamics

$$x(t+1) = ax(t) + u(t-T) + w(t). \tag{1}$$

The sequence $u(t-T)$ is the control action with delay T and $w(t)$ is the disturbance. The achievable robust performance using a controller with delay $T \geq 0$ and bandwidth R (sampling per unit time) is given by

$$\max_{\|w\|_\infty \leq 1} \|x\|_\infty = \sum_{i=0}^{T} |a^{i-1}| + |a^T|(2^R - |a|)^{-1} \qquad (2)$$

The optimal performance using a controller with just delay T is $\sum_{i=0}^{T} |a^{i-1}|$ (denote as the delay cost). The optimal performance using a controller with bandwidth R (and $T = 0$) is $(2^R - |a|)^{-1}$ (denote as the quantization cost). Surprisingly, the achievable performance using a controller with delay T and bandwidth R is a simple combination of these two terms. While the proofs are a bit tedious, they can be done entirely with high school algebra.

Assuming the lengths of nerves are given and constrained by other physiology, the metabolic overhead to build and maintain a nerve is roughly proportional to its cross sectional area [11]. For a fixed area (and thus cost) there is then a tradeoff between having high bandwidth with many small axons, or high speed (and low delay) with fewer larger axons. This tradeoff can be approximately quantified in [2] as:

$$R = \lambda T_s. \qquad (3)$$

Fast but inaccurate signaling is achieved by nerves with a few large axons, whereas slower but more accurate signaling occurs in nerves with many small axons. The constant λ is proportional to the resource (space and energy) consumption of the signaling nerve fibers. The tradeoffs in Fig. 4 can be redrawn assuming $R = \lambda T_s$ to give the speed-accuracy tradeoff in Fig. 5. What follows does not depend crucially on the specific function $R = \lambda T_s$ but simply that there is some tradeoff between bandwidth and delay, and thus this framework would readily incorporate alternative assumptions.

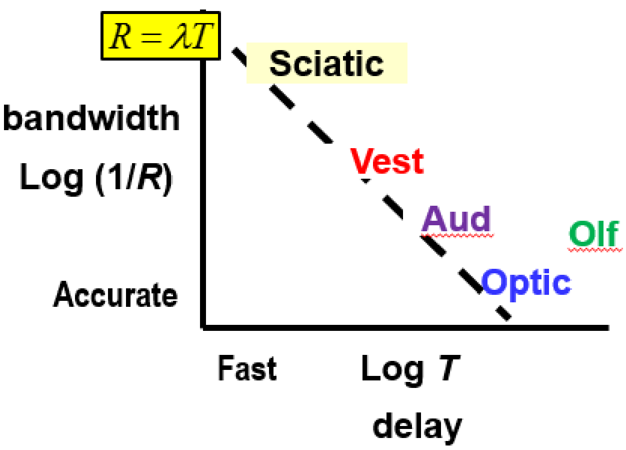


Fig. 5: Speed and accuracy for different nerves.

If we initially assume that the signaling delay in nerve fibers are the only cause of delay in control, *i.e.*, $T = T_s$ then combining with (1), the achievable robust performance can be written as

$$\max_{\|w\|_\infty \leq 1} \|x\|_\infty = \sum_{i=0}^{T} |a^{i-1}| + |a^T|(2^{\lambda T} - |a|)^{-1} \qquad (4)$$

Figure 6 shows the system performance when we vary delay $T = \lambda^{-1} R$ for fixed resource λ and $a = 1$, which simplifies

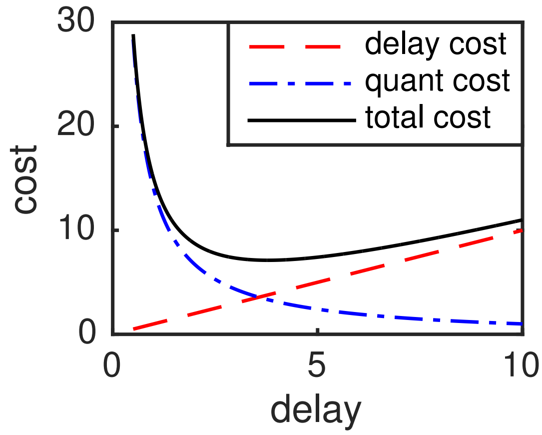


Fig. 6: Impact of speed versus accuracy. The cost of delay, the cost of quantization, and the total cost is shown with varying delay $T = \lambda^{-1} R$ when $\lambda = .1$ and $a = 1$. This plot assumes that the signaling delay in nerve fibers are the only cause of delay in control, *i.e.*, $T = T_s$.

the cost to

$$\max_{\|w\|_\infty \leq 1} \|x\|_\infty = T + (2^{\lambda T} - 1)^{-1} \qquad (5)$$

Increased delay increases the delay cost but reduces the quantization cost. Consequently, the optimal system level performance is achieved at intermediate levels of delay and bandwidth. This contrasts dramatically with what is suggested by information theory, which has dominated theoretical neuroscience [16], and emphasizes maximizing mutual information, and thus bandwidth, and entirely neglects delay.

Next we add additional delay or advance warning in other parts the feedback control loop and discuss its impact on robust performance and optimal nerve physiology. Assume now that the total delay or warning in control T is determined by nerve signaling delay T_s, some additional computational delay T_c, and also advanced warning T_w due to remote sensing such as vision, as

$$T = T_s + T_c - T_w = T_s - \tau \qquad (6)$$

where $\tau = T_w - T_c$. Then we focus on optimal design of one nerve with all other components fixed, and assume this nerve has the bandwith/delay tradeoff $R = \lambda T_s$.

Figure 7 shows the optimal signaling delay T_s and bandwidth $R = \lambda T_s$ that achieve the minimum total error when varying $T_w \geq 0$ and $T_c \geq 0$ separately in the two special cases: (i) $\tau = T_w - T_c > 0$ (denote as the warned system); (ii) $\tau = T_w - T_c < 0$ (denote as the delayed system). As a consequence, there are two qualitatively distinct regimes with strikingly different optimal physiology:

(i) The delayed system: When the computation delay T_c is greater than the advanced warning T_w, then $\tau < 0$ and the controller unavoidably has a net positive delay $T > 0$. Since the delay cost dominates the total cost in this regime, relatively small bandwidth R and signaling delay T_s is optimal. Furthermore, the total cost grows as T_c increases (and τ decreases).

(i) The warned system: When the advanced warning T_w is greater than the computation delay T_c, then $\tau > 0$, and

the controller can potentially exploit this advanced warning of future disturbance. Since the quantization cost now dominates the total cost, large bandwidth R and signaling delay T_s is optimal. Furthermore, the total cost goes to zero as T_w increases, exactly the opposite of the delayed system.

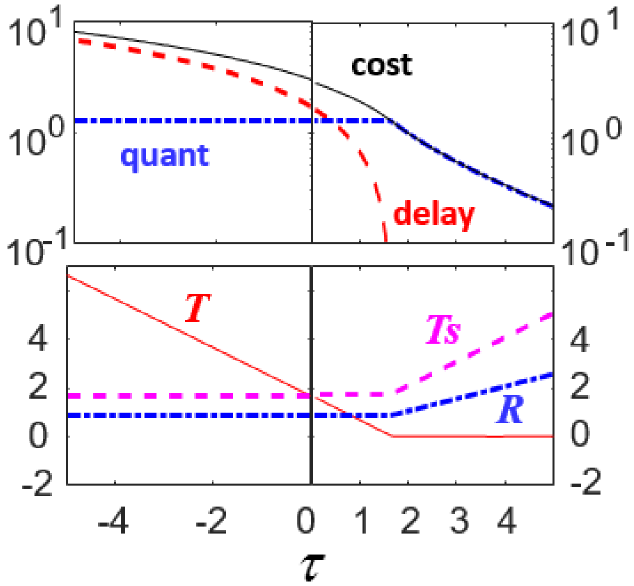


Fig. 7: The behavior of the delayed and warned system.

2.2 Comparing theory and experiment

This section briefly compares qualitatively the experiments from Fig. 2, the minimal block diagram cartoon in Fig. 3, the nerve compositions in Fig. 4, and the toy model summarized in Figs. 5, 6, and 7. The most obvious connections is in comparing the VOR and vision systems. Any top predators that stalk and chase their prey using vision, like humans, must be able to maintain sharp visual tracking despite large, fast head motions, and thus the sensors, communication, and control via the VOR and vestibular nerve clearly fit the "delayed" left side of Fig. 7, with relatively few but large axons for relatively low delay and accuracy. The sciatic nerve is even more extreme, and is involved in even faster and lower accuracy communications in lower limb reflexes. In general, the left side of Fig. 7 applies to the fast, automatic, unconscious, and reflexive components that prevent the crashes as depicted as cartoons in Fig. 1.

The vision system as used in the hand motion experiment is extremely suboptimal compared to VOR for tracking fast moving nearby objects, but this is fortunately not a task it needs to do well for survival. Instead, central vision is used to provide high resolution images of remote objects and their motion, giving advanced warning for control, and thus is on the "warned" right side of Fig. 7, with relatively many but smaller axons and thus with high delay and accuracy. With the advanced warning and planning that vision enables, humans are able to take huge initial condition errors and with time, drive them nearly to zero, consistent with the far right side of Fig. 7. Peripheral vision in contrast is fast and low resolution, and not used for tracking but instead to provide rapid alerts to fast moving objects in the periphery.

Thus we have a simple and direct connection between the neuron-level physiology in Figs. 4 and 5 and behavioral plausible system level tradeoffs as in Fig. 7.

3 Generalizations

The story so far is remarkably relevant to sensorimotor control given how extremely simplified it is, and for the first time rigorously connects neuron-level tradeoffs such as in Fig. 5 with system level robust performance as in Fig. 7. But real control systems are not only delayed and quantized but also layered, distributed, and localized, and all are necessary to make controller synthesis and implementation scalable. For large-scale systems such as the smart grid, software defined networks, automated highway systems, and biological systems, centralized controllers cannot be implemented due to the delay constraints imposed by the communication network interconnecting different subsystem controllers. Specifically, the underlying assumption of the centralized scheme is that the information from *all* the subsystems in the network must be collected *instantaneously* and transmitted *perfectly*. Similar to the SISO model described in the previous section, the practical communication delays degrade the performance of the centralized scheme substantially and make it unappealing to large-scale systems.

3.1 Distributed and Localized Control

The delay constraints between sensors, actuators, and subsystem controllers in a network system lead to an asymmetry in the information available to each of the local decision makers. This makes the MIMO control problems with delays fundamentally different from the SISO control problems with delays. In particular, even in a system with 2 states, the asymmetry in information can make the distributed optimal control problem computationally intractable (i.e., non-convex, non-linear, and NP-hard) [17, 18]. Due to the fundamental difficulties of solving general distributed optimal control problems, recent approaches have shifted toward the identification of tractable subclass of distributed optimal control problem.

It has been shown that the tractability of the distributed optimal control problems is determined by the relation between the delay constraints and the physical plant. Specifically, when the delay constraints satisfy a technical condition known as quadratic invariance [19–21], then the distributed optimal control problem can be solved via convex optimization. Roughly speaking, quadratic invariance holds when local controllers can communicate with each other faster than their control actions propagate through the physical system. This assumption is valid for many of the applications we are interested in (power grid, software-defined networking, automated highway systems, and neuroscience applications). Following the idea of quadratic invariance, we have made various progresses in finding finite dimensional reformulations of the optimal control problem or deriving explicit formulas for specific structured systems [22–24].

As promising as all of these results have been, they all suffer from a limitation that is inherent to quadratic invariance based distributed optimal control: they are not scalable to large systems. In particular, for a system with strongly connected dynamics, a controller architecture is quadratically invariant if and only if each sub-controller eventually collects and estimates the *global* system state. In particular, this implies that although the distributed optimal controller respects the communication constraints of the system, its control law is actualy more difficult to compute and/or implement than that of a corresponding centralized controller.

In order to circumvent the issues of *scalability*, we have introduced the *localized optimal control and estimation* framework [25–27]. By relying on an alternative parameterization of the controller in terms of system closed loop responses, we are able to impose arbitrary convex constraints on the controller architecture. In particular, this means that we can formulate distributed optimal control problems in which each local sub-controller only needs to collect a local subset of state measurements, and only requires a local sub-model of the full system. In this way, the complexity of the synthesis and implementation of a distributed controller depends only on the size of these local sub-problems, potentially allowing for arbitrary scalability. The caveat is that we ask for a slightly stronger condition on the communication delays: they must be such that local sub-controllers can coordinate to localize the effect of disturbances, effectively isolating disparate but overlapping sections of the global plant from each other. Using these techniques, we have been able to compute near globally optimal controllers for heterogeneous systems with over 50,000 states [26]. We note that the computational bottleneck in this case was the need to simulate such a system using a single laptop – if such a system were actually implemented in parallel the scalability of our approach would essentially be infinite.

3.2 From theory and practice

These theoretical results, while important, do not lend themselves to direct practical application due to the potentially unrealistic assumptions that need to be made to be able to answer them. These include the use of communication channels with infinite bandwidth and constant delay, the availability of good state-space models, and fixed controller architecture. Other lines of our research program aim to systematically address these unrealistic assumptions, with the ultimate goal of closing the gap between theory and practice.

In [28], we study the stability and optimal performance for distributed discrete time controllers with time-varying delay, quantization, saturation, sampling, and external disturbances. We propose method that designs distributed/localized controllers for such sytems with linear programing complexity. In [29], new identification results based on recursive Lasso have been proposed, which enables identification accuracy using a small number of noisy samples with theoretical guarantees.

On the other hand, the controller architecture, i.e., the actuators, sensors and communication network between them, can also be designed rather than as something given. Leveraging tools from Quadratic Invariance and Structured Linear Inverse Problems, we have developed a unifying computationally tractable framework built around convex optimization for the co-design of a controller architecture and its corresponding optimal control law [30]. We also extended the regularization for design (RFD) framework to work in conjunction with localized optimal control. Specifically, we have shown how the Localized LQR Optimal control problem can be suitably modified to simultaneously design a sparse actuation architecture and the corresponding optimal localized control law, despite pre-specified locality constraints [31].

3.3 Theory of Architecture

Network protocols in layered architectures have historically been obtained on an ad hoc basis, and many of the recent cross-layer designs are also conducted through piecemeal approaches. Network protocol stacks may instead be holistically analyzed and systematically designed as distributed solutions to some global optimization problems. On the other hand, the controller can also be built in separate layers. The work [32] proposed the Layering as Optimization (LAO) framework that modeled the overall communication network as a generalized network utility maximization problem, where each layer corresponds to a decomposed subproblem. The interfaces among layers are quantified as functions of the optimization variables coordinating the subproblems.

Recently, we generalized the Layering as Optimization (LAO) framework to incorporate not only optimization, but dynamics and control as well [33]. We show that by suitably relaxing an optimal control problem that jointly addresses determining and following an optimal trajectory, one can naturally recover a layered architecture composed of a low-level tracking layer and a top-level planning layer. The tracking layer consists of a distributed optimal controller that takes as an input a reference trajectory generated by the top-level layer, where this top-level layer consists of a trajectory planning problem that optimizes the weighted sum of a utility function and a "tracking penalty" regularizer. This latter term can be viewed as the planning layer's "virtual model" of the underlying physics of the system, and serves as a balance between the two by ensuring that the planned trajectory can indeed be efficiently followed by the tracking layer.

4 Additional material

Additional materials (slides, videos, and references) can be found in the following websites:

John Doyle's website:
www.cds.caltech.edu/ doyle
Nikolai Matni's website:
http://www.cds.caltech.edu/ nmatni/home/Welcome.html
Yorie Nakahira's website:
http://users.cms.caltech.edu/ ynakahir/

References

[1] John C Doyle and Marie Csete. Architecture, constraints, and behavior. *Proceedings of the National Academy of Sciences*, 108(Supplement 3):15624–15630, 2011.

[2] Yorie Nakahira, Nikolai Matni, and John C Doyle. Hard limits on robust control over delayed and quantized communication channels with applications to sensorimotor control. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 7522–7529, Dec 2015.

[3] Fiona A Chandra, Gentian Buzi, and John C Doyle. Glycolytic oscillations and limits on robust efficiency. *science*, 333(6039):187–192, 2011.

[4] Rami Namas, Ruben Zamora, Rajaie Namas, Gary An, John Doyle, Thomas E Dick, Frank J Jacono, Ioannis P Androulakis, Gary F Nieman, Steve Chang, et al. Sepsis: Some-

thing old, something new, and a systems view. *Journal of critical care*, 27(3):314–e1, 2012.

[5] Na Li, Jerry Cruz, Chenghao Simon Chien, Somayeh Sojoudi, Benjamin Recht, David Stone, Marie Csete, Daniel Bahmiller, and John C Doyle. Robust efficiency and actuator saturation explain healthy heart rate control and variability. *Proceedings of the National Academy of Sciences*, 111(33):E3476–E3485, 2014.

[6] Henrik Sandberg, Jean-Charles Delvenne, and John C Doyle. On lossless approximations, the fluctuation-dissipation theorem, and limitations of measurements. *Automatic Control, IEEE Transactions on*, 56(2):293–308, 2011.

[7] Dennice F Gayme, Beverley J McKeon, Bassam Bamieh, Antonis Papachristodoulou, and John C Doyle. Amplification and nonlinear mechanisms in plane couette flow. *Physics of Fluids (1994-present)*, 23(6):065108, 2011.

[8] Morgan T Page, David Alderson, and John Doyle. The magnitude distribution of earthquakes near southern california faults. *Journal of Geophysical Research: Solid Earth*, 116(B12), 2011.

[9] Zhaoping Li. *Understanding vision: Theory, models, and data*. Oxford University Press, USA, 2014.

[10] János A Perge, Jeremy E Niven, Enrico Mugnaini, Vijay Balasubramanian, and Peter Sterling. Why do axons differ in caliber? *The Journal of Neuroscience*, 32(2):626–638, 2012.

[11] Peter Sterling and Simon Laughlin. *Principles of neural design*. MIT Press, 2015.

[12] Yorie Nakahira and John C Doyle. Lq vs. l_∞ in controller design for systems with delay and quantization, in preparation.

[13] John C Doyle. Teaching control theory in high school. *To appear, Conference on Decision and Control, Las Vegas*, 2016.

[14] Leong Yoke Peng and John C Doyle. Understanding robust control theory via stick balancing, in preparation.

[15] Girish N Nair, Fabio Fagnani, Sandro Zampieri, and Robin J Evans. Feedback control under data rate constraints: An overview. *Proceedings of the IEEE*, 95(1):108–137, 2007.

[16] Fred Rieke. *Spikes: exploring the neural code*. MIT press, 1999.

[17] Hans S Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal of Control*, 6(1), 1968.

[18] John N. Tsitsiklis and Michael Athans. On the complexity of decentralized decision making and detection problems. In *IEEE Conference on Decision and Control (CDC)*, 1984.

[19] Michael Rotkowitz and Sanjay Lall. A characterization of convex problems in decentralized control. *Automatic Control, IEEE Transactions on*, 51(2):274–286, 2006.

[20] Şurban Sabău and Nuno C. Martins. Youla-like parametrizations subject to QI subspace constraints. *Automatic Control, IEEE Transactions on*, 59(6):1411–1422, 2014.

[21] Laurent Lessard and Sanjay Lall. Convexity of decentralized controller synthesis. *IEEE Transactions on Automatic Control, To appear*, 2016.

[22] Andrew Lamperski and John C. Doyle. The $\mathcal{H}_2$ control problem for quadratically invariant systems with delays. *Automatic Control, IEEE Transactions on*, 60(7):1945–1950, 2015.

[23] Andrew Lamperski and Laurent Lessard. Optimal decentralized state-feedback control with sparsity and delays. *Automatica*, 58:143–151, 2015.

[24] Nikolai Matni. Distributed control subject to delays satisfying an $\mathcal{H}_\infty$ norm bound. In *2014 53rd IEEE Conference on Decision and Control (CDC)*, 2014.

[25] Yuh-Shyang Wang, Nikolai Matni, and John C. Doyle. Localized LQR optimal control. In *2014 53rd IEEE Conference on Decision and Control (CDC)*, 2014.

[26] Yuh-Shyang Wang, Seungil You, and Nikolai Matni. Localized distributed Kalman filters for large-scale systems. In *5th IFAC Workshop on Distributed Estimation and Control in Networked Systems*, 2015.

[27] Yuh-Shyang Wang and Nikolai Matni. Localized LQG optimal control for large-scale systems. In *2016 IEEE American Control Conference (ACC), To appear*, 2016.

[28] Yorie Nakahira and Seungil You. A linear programming framework for networked control system design. *IFAC-PapersOnLine*, 48(22):34–39, 2015.

[29] Ye Yuan, Yorie Nakahira, and Claire Tomlin. On identification of parameterized switched linear systems. In *35th Chinese Control Conference*, Jul 2016.

[30] Nikolai Matni and Venkat Chandrasekaran. Regularization for design. *Accepeted, IEEE Transactions on Automatic Control*.

[31] Yuh-Shyang Wang, Nikolai Matni, and John C. Doyle. Localized LQR control with actuator regularization. In *2016 IEEE American Control Conference (ACC), To appear*, 2016.

[32] Mung Chiang, Steven H Low, A Robert Calderbank, and John C Doyle. Layering as optimization decomposition: A mathematical theory of network architectures. *Proceedings of the IEEE*, 95(1):255–312, 2007.

[33] Nikolai Matni and John C Doyle. A theory of dynamics, control and optimization in layered architectures. In *the 2016 American Control Conference*, 2016.