Mathematisches Forschungsinstitut Oberwolfach

Report No. 10/2018

DOI: 10.4171/OWR/2018/10

Mini-Workshop: Entropy, Information and Control

Organised by
Fritz Colonius, Augsburg
Tomasz Downarowicz, Wroclaw
Christoph Kawan, Passau
Girish Nair, Melbourne

4 March – 10 March 2018

Abstract. This mini-workshop was motivated by the emerging field of networked control, which combines concepts from the disciplines of control theory, information theory and dynamical systems. Many current approaches to networked control simplify one or more of these three aspects, for instance by assuming no dynamical disturbances, or noiseless communication channels, or linear dynamics. The aim of this meeting was to approach a common understanding of the relevant results and techniques from each discipline in order to study the major, multi-disciplinary problems in networked control.

Mathematics Subject Classification (2010): 93B52, 93C41, 37A35, 94A17.

Introduction by the Organisers

Sixteen researchers from Europe, North and South America and the Asia-Pacific area participated in this mini-workshop, including the organisers Fritz Colonius, Tomasz Downarowicz, Christoph Kawan, and Girish Nair. In total, twenty-two talks were featured over 5 days. Due to the varying backgrounds of the participants, these included several tutorial-style and overview talks on relevant material from control theory (Colonius), dynamical systems entropy (Downarowicz), information theory (Sahai), dimension and entropy (Gelfert), data rate theorems (Franceschetti), invariance entropy (Kawan) and nonstochastic information (Nair). These talks provided important background to the presentations of new and recent research results, and exposed participants from each discipline to the main techniques and terminology of the others. In addition, there was an evening "Open

Problems" session, a closing discussion, and an impromptu joint session with the mini-workshop on "Deep Learning and Inverse Problems".

In broad terms, the research presentations and Open Problems followed two themes. The first theme focused on entropy-related concepts for controlled or uncontrolled dynamical systems without channel or system noise. Liberzon discussed the notion of *estimation entropy* for such uncontrolled systems, and showed that it characterises the minimum bit rate needed to be able to estimate the system states with a specified exponential speed. In a similar setting, Pogromsky proposed the concept of *restoration entropy*, and proved that this characterised the minimum errorless bit rates needed to estimate system states in *regular* and *fine* senses. Gelfert discussed the box dimension of skew-product systems with partially hyperbolic and hyperbolic structure, and showed that it can be determined in terms of *pressure*, a notion closely related to entropy and inspired by thermodynamics. Santana proposed extensions of the pressure concept to control systems, via the notions of *invariance pressure* and *outer invariance pressure.* Serafin discussed a notion of *universal* topological systems, and showed that the class of dynamical systems with measure-theoretic entropies lying in any given nonnegative interval always admits a universal topological system. Downarowicz described the notion of *symbolic extension entropy*, which characterises the minimum bit rate needed to estimate the initial state of a topological dynamical system with error approaching zero with time, and showed how it is related to the *entropy structure* of the topological system. Colonius proposed an extension of invariance entropy to measure-theoretic systems governed by quasi-stationary measures, and showed that this measure-theoretic invariance entropy is invariant under measurable transformations and determined by the control sets of the system. The question of whether there is a variational principle linking this new concept to invariance entropy remains open.

The second theme focused on the information-theoretic aspects of systems with noise or uncertainty. Yüksel spoke about the control of nonlinear stochastic systems via noisy channels, under stability notions such as asymptotic mean stationarity, ergodicity and Harris recurrence, and gave characterizations of the largest class of channels for which there exist coding and control policies so that the closed-loop system can be made stochastically stable. In a similar setting, Kawan discussed the relationship between the channel capacity needed to achieve a given objective, and a measure-theoretic entropy-like quantity defined by representing the control loop as a random dynamical system. Franceschetti considered the problem of asymptotically stabilizing a linear system via an errorless channel with independent random delays. He showed that it is necessary for the entropy of the linear system to be no greater than the ordinary Shannon capacity of the delay viewed as a timing channel, a condition which is also sufficient when the delays are exponentially distributed. A related problem was discussed by Linsenmayer, who considered the containability of linear systems controlled via a digital channel with uncertain delay, and derived sufficient conditions in terms of the delay bound, the number of bits per packet, and the system dynamical parameters. Ranade considered the stabilization in probability of linear systems with multiplicative random noise, and showed that the stabilizability condition could be viewed in terms of a new concept called *control capacity*. She also discussed the open problem of determining tight conditions for the stabilizability and optimal nonlinear control of scalar linear plants with multiplicative measurement noise. Kostina spoke about the control of noisy linear systems via noisy channels and presented fundamental trade-offs between the Marko-Massey's directed information, equivalent to the expected bit-rate, sent across a noisy channel, and control performance as measured by the mean square state. Ishii considered a similar class of systems and showed that if the additional requirement of asymptotic second-order (i.e. wide-sense) stationarity was also imposed, then fundamental trade-offs between directed information and the disturbance rejection performance could be derived in frequency domain in terms of generalised Bode integrals. Sahai described the classical problem of the optimal decentralized control of Gaussian linear systems under a quadratic performance cost, and discussed connections to rate-distortion theory, dirty paper coding and sphere-packing bounds in information theory.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of a junior researcher in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows".

Mini-Workshop: Entropy, Information and Control

Table of Contents

Abstracts

On Stationary and Ergodic Properties of Stochastic Non-Linear Systems Controlled over Communication Channels

Serdar Yüksel

Consider a controlled non-linear system described by

$$x_{t+1} = f(x_t, u_t, w_t), \tag{1}$$

where x_t is the $\mathbb{R}^N$-valued state, u_t is the control action variable, and $\{w_t\}$ an i.i.d. noise process. This system is controlled over a noisy channel as shown in Figure 1. A *coding policy* Π is a sequence of functions $\{\gamma_t^e, t \in \mathbb{Z}_+\}$ such that $\gamma_t^e(x_{[0,t]}, q'_{[0,t-1]}) = q_t \in M$, where M is the finite channel input alphabet. The channel maps $q_t \in M$ to $q'_t \in M'$ in a stochastic fashion so that $P(q'_t \in \cdot | q_t, q_{[0,t-1]}, q'_{[0,t-1]})$ is a conditional probability measure for $t \in \mathbb{Z}_+$. A *controller policy* γ is a sequence of functions $\{\gamma_t, t \in \mathbb{Z}_+\}$ such that $u_t = \gamma_t(q'_{[0,t]})$.

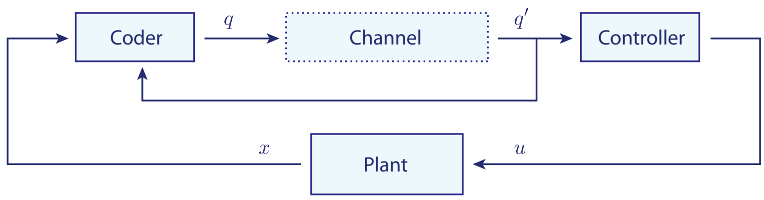

Figure 1. Control over a noisy channel with feedback.

This talk is concerned with the following: Given such a nonlinear system controlled over a channel, what is the largest class of channels for which there exist coding and control policies so that the closed-loop system can be made stochastically stable? Stability notions considered are asymptotic mean stationarity, ergodicity, and positive Harris recurrence. We do not restrict the state space to be compact, for example, systems considered can be driven by unbounded noise. Necessary and sufficient conditions are obtained for a large class of systems and channels. It is shown that the conditions obtained are tight for the case of linear systems driven by unbounded noise, and for such systems the stability criteria can be strengthened to positive Harris recurrence for noiseless and erasure channels [2, Chp. 6-8]. Connections with the metric entropy and Lyapunov exponents for random dynamical systems, prior work on deterministic nonlinear systems, and the state estimation problem will be presented, and some open problems involving the interaction of dynamical systems, information theory and networked control will be discussed. Some results related to this talk are given in [1].

References

[1] S. Yüksel. Stationary and ergodic properties of stochastic non-linear systems controlled over communication channels. *SIAM J. on Control and Optimization*, 54:2844–2871, 2016.
[2] S. Yüksel and T. Başar. *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints.* Springer, New York, 2013.

Stabilizing a linear system with phone calls

Massimo Franceschetti
(joint work with Mohammad Khostajesh, Gireeja Ranade)

A first-order abstraction of a cyber-physical system is that of a networked control system where the feedback loop is closed over a communication channel [1][2]. To quantify the effect of the communication channel on the ability to stabilize the system, several data-rate theorems have emerged in the literature, see [3, 4] for a survey. Roughly speaking, they state that to achieve stabilization the communication rate available in the feedback loop should be at least as large as the *intrinsic entropy rate* of the system, given by the sum of its unstable modes. Within this framework, a portion of the literature studied stabilization over "bit pipe channels," where a rate-limited, possibly time-varying, noiseless communication channel is present in the feedback loop. In the case of noisy channels, Tatikonda and Mitter [5] showed that for almost sure (a.s.) stabilization of undisturbed linear plants the Shannon capacity of the channel should be larger than the entropy rate of the plant. Matveev and Savkin [6] showed that this condition is also sufficient for discrete memoryless channels, but a stronger condition is required in the presence of disturbances, namely the zero-error capacity should be larger than the entropy rate of the plant [7]. Sahai and Mitter [8] considered the less stringent requirement of moment-stabilization over noisy channels and in the presence of system disturbances, and provided a data-rate theorem in terms of the anytime capacity of the channel. For a comprehensive treatment of data-rate theorems one can refer to the books [9, 10] and to the surveys [3, 4].

Another important aspect of CPS is event-triggering control.In this case, communication occurs in an opportunistic manner and the primary focus is on minimizing the number of transmissions while simultaneously ensuring the control or objective. Since the timing of the state-dependent triggering events carries information that can be used for stabilization. The amount of timing information, however, is sensitive to the delay in the communication channel. While for small delay stabilization can be achieved with data-rate arbitrarily close to zero, for large values of the delay this is not the case [11, 12].

In this work, we point out that state-dependent triggering is only one possible strategy to encode information in time. In order to characterize the fundamental limitations of using timing information for stabilization, we consider a timing channel where information is encoded by adjusting the transmission time of a single symbol [13]. We show that in order to drive the state to zero, the timing capacity of this channel should be at least as large as the entropy rate of the system.

In addition, in the case transmission is subject to exponentially distributed delay, we provide a tight sufficient condition. Finally, we point out that while in our analysis we restrict to transmitting a symbol from a unitary alphabet, it would be of practical interest to develop "mixed" strategies, using both timing information and physical data transmitted over a larger alphabet, and this is left for future work.

References

[1] K.-D. Kim and P. R. Kumar, "Cyber–physical systems: A perspective at the centennial," *Proceedings of the IEEE*, vol. 100 (Special Centennial Issue), pp. 1287–1308, 2012.
[2] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, 2007.
[3] M. Franceschetti and P. Minero, "Elements of information theory for networked control systems," in *Information and Control in Networks*. Springer, 2014, pp. 3–37.
[4] B. G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.
[5] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE transactions on Automatic Control*, vol. 49, no. 7, pp. 1196–1201, 2004.
[6] A. S. Matveev and A. V. Savkin, "An analogue of shannon information theory for detection and stabilization via noisy discrete communication channels," *SIAM journal on control and optimization*, vol. 46, no. 4, pp. 1323–1367, 2007.
[7] ——, "Shannon zero error capacity in the problems of state estimation and stabilization via noisy communication channels," *International Journal of Control*, vol. 80, no. 2, pp. 241–255, 2007.
[8] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link. Part I: Scalar systems," *IEEE transactions on Information Theory*, vol. 52, no. 8, pp. 3369–3395, 2006.
[9] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Springer Science & Business Media, 2013.
[10] A. S. Matveev and A. V. Savkin, *Estimation and control over communication networks*. Springer Science & Business Media, 2009.
[11] M. J. Khojasteh, P. Tallapragada, J. Cortés, and M. Franceschetti, "The value of timing information in event-triggered control," *arXiv preprint arXiv:1609.09594*, 2016.
[12] M. J. Khojasteh, M. Hedayatpour, J. Cortes, and M. Franceschetti, "Event-triggered stabilization of disturbed linear systems over digital channels," *arXiv preprint arXiv:1801.08704*, 2018.
[13] V. Anantharam and S. Verdu, "Bits through queues," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 4–18, 1996.

Entropy and minimal bit rates for state estimation and model detection

Daniel Liberzon
(joint work with Sayan Mitra)

We discuss the notion of *estimation entropy*, formulated in terms of the number of system trajectories that approximate all other trajectories up to an exponentially decaying error. We also consider an alternative definition of entropy which uses approximating functions that are not necessarily trajectories of the system.

We show that the two entropy notions turn out to be equivalent. We proceed to establish an upper bound of $(M + \alpha)n/\ln 2$ for the estimation entropy of an n-dimensional nonlinear dynamical system whose Jacobian matrix f_x has matrix measure bounded by M, when the desired exponential convergence rate of the estimate is α . We also develop a lower bound of $(\inf \mathrm{tr} f_x + \alpha n)/\ln 2$ on the estimation entropy, where the infimum is taken over the reachable states of the system. For linear systems, the upper and lower bounds can be refined so that they coincide and give an exact expression for the estimation entropy in terms of the eigenvalues of the system matrix.

Next, we propose an iterative procedure that uses quantized and sampled state measurements to generate state estimates that converge to the true state at the desired exponential rate. The main idea in the algorithm is to exponentially increase the resolution of the quantizer while keeping the number of bits sent in each round constant. This is achieved by using the quantized state measurement at each round to compute a bounding box for the state of the system for the next round. Then, at the beginning of the next round, this bounding box is partitioned to make a new and more precise quantized measurement of the state. We show that the bounding box is exponentially shrinking in time at a rate α when the average bit rate utilized by this procedure matches the upper bound $(M + \alpha)n/\ln 2$ on the estimation entropy. We also show that no other algorithm of this type can perform the same estimation task with bit rates lower than the estimation entropy. In other words, the "efficiency gap" of our estimation procedure is at most as large as the gap between the estimation entropy of the dynamical system and the above upper bound on it.

Moreover, we present an application of the estimation procedure in solving a model detection problem. Suppose we are given two competing candidate models of a dynamical system, and from the quantized and sampled state measurements we would like to determine which one is the true model. For example, the different models may arise from different parameter values or they could model "nominal" and "failure" operating modes of the system. We demonstrate that under a mild assumption of *exponential separation* of the candidate models' trajectories, a modified version of our estimation procedure can always definitively detect the true model in finite time.

Entropy for switched and hybrid systems and its role in state estimation and model detection as well as control of such systems is a subject of ongoing work. At the end of the talk, we discuss some preliminary results on computing entropy of switched linear systems with diagonal and triangular structure.

The talk is largely based on the work reported in [1]. The results on entropy for switched systems will appear in [2].

References

[1] D. Liberzon and S. Mitra. Entropy and minimal bit rates for state estimation and model detection. *IEEE Trans. Automat. Control*, 2018. To appear.

[2] G. Yang, A. J. Schmidt, and D. Liberzon. Topological entropy of switched linear systems with diagonal, triangular, and general matrices. 2018. Submitted.

State estimation of deterministic systems via information channels with finite capacity

Alexander Pogromsky
(joint work with A.S. Matveev)

The multidisciplinary area of networked control systems lies at the crossroad of control, communication, and computer sciences and integrates their classical topics into a whole, while responding to new special challenges born out of their union. One of them is caused by bottlenecks in the process of information transmission among the network nodes that may be due to constraints on the data transmission bit-rate allocated to every particular transmitter/receiver pair within a shared fieldbus device.

It is this observability problem that is addressed in the talk: What is the minimal bit-rate of data transfer that is needed to build a reliable and efficient state estimate in the remote location. To be more precise, we consider a nonlinear time-invariant plant

$$\dot{x} = f(x), \; x \in \mathbb{R}^n, \; t \in [0, \infty), \; x_0 \in K, \tag{1}$$

where x is the state, the map $f : \mathbb{R}^n \to \mathbb{R}^n$ is C^1-smooth, and $K \neq \emptyset$ is a positively invariant compact set of the feasible initial states. As time t runs from 0, a valid estimate $\widehat{x}(t, \widehat{x}_0)$ of the solution $x(t, x_0)$ of (1) should be continuously generated in real time t at a certain faraway site, S_{est}, where direct observation of x is impossible. Meanwhile, $x(t, x_0)$ is fully accessible at time t in another location $S_{\text{sen}} \neq S_{\text{est}}$ (typically that of the plant) on the basis of sensor data. The problem to be treated stems from the bottleneck caused by a finite bit-rate of data communication from S_{sen} to S_{est}.

Similarly to our previous results on discrete-time observation problem, [1], the communication channel is characterized by its bit-rate capacity c. The lower thresholds on this bit-rate which guarantee solvability of the regular and fine observation problem (denoted by $\mathfrak{R}_{\text{ro}}$ and $\mathfrak{R}_{\text{fo}}$ respectively) coincide due to the positive invariance of the set K [1].

The *restoration entropy* $H_{\text{res}}(f, K)$ of the system around the set K is defined as follows. Let B_a^δ be the ball of radius δ centered at the point a. Let $p(T, a, \delta)$ be the minimal number of δ-balls required to cover the image $x(T, (B_a^\delta \cap K))$. We define

$$H_{\text{res}}(f, K) := \lim_{T \to \infty} \frac{1}{T} \overline{\lim_{\delta \to 0}} \sup_{a \in K} \log_2 p(T, a, \delta). \tag{2}$$

The outer limit exists due to positive invariance of K. The relations between the observability bit-rate thresholde and the restoration entropy are established in a form of the Data Rate Theorem [2]: $\mathfrak{R}_{\text{ro}} = \mathfrak{R}_{\text{fo}} = H_{\text{res}}(f, K)$. The main results of the talk are focused on a constructive way to estimate the restoration entropy (lower and upper bounds). While the lower estimate is given in the spirit of the first Lyapunov method, the upper estimate goes along the lines of second Lyapunov method [2].

References

[1] A. Matveev, A. Pogromsky *Observation of nonlinear systems via finite capacity channels: constructive data rate limits*, Automatica, 70:217-229, 2016.

[2] A. Matveev, A. Pogromsky, *Observation of nonlinear systems via finite capacity channels. Part II: Restoration entropy and its estimates*, submitted, a preliminary version is available upon request.

Dimensions and critical regularity of hyperbolic graphs

Katrin Gelfert

(joint work with Lorenzo J. Díaz, Maik Gröger, Tobias Jäger)

The investigation of fractal attractors, repellers, horseshoes, and other types of hyperbolic sets has been a major driving force for many important developments in ergodic theory and its interfaces with mathematical physics and fractal geometry. The situation is fairly well understood essentially in a conformal setting, only (see [3, 4]), comparable to the study of conformal repellers. Extending the theory to higher-dimensional and genuinely nonconformal situations is well known to be difficult, and there exist only few and specific results in this direction (see, for example, [2]). Different phenomena here complicate matters:

- The possible loss of equality between Hausdorff and box dimensions,
- both dimensions may not vary continuously with the dynamics.

We proceed by studying the graphs in three-dimensional skew product systems

$$T\colon \Xi\times\mathbb{R}\to\Xi\times\mathbb{R},\quad T(\xi,x)=(\tau(\xi),T_\xi(x)),$$

with simultaneously a partially-hyperbolic and a hyperbolic structure. Here $\tau\colon\Xi\to\Xi$ are hyperbolic surface diffeomorphisms, or their restrictions to basic pieces. Dimension is intimately related to thermodynamic quantities such as entropy and Lyapunov exponents, where pressure often encodes such relations. Building on [2, 1], we show that, except in a nongeneric case when the graph is Lipschitz, its box dimension is given by $d^{\rm s}+d$, where $d^{\rm s}$ is the dimension of stable slices of Ξ and where d is determined as the unique solution of the pressure equation

$$P_{\tau|_\Xi}(\varphi^{\rm cu}+(d-1)\varphi^{\rm u})=0.$$

Here $\varphi^{\rm cu},\varphi^{\rm u}$ are geometric potentials taking into account the expansion rates in the fiber center unstable and the strong unstable directions, respectively.

A key ingredient to establish a lower bound for the dimension are ideas from multifractal analysis. Here we study the entropy of sets of points which simultaneously have given prescribed Birkhoff averages for the potentials $\varphi^{\rm cu}$ and $\varphi^{\rm u}$, respectively. These entropies govern the exponential growth of the number of Markov rectangles with an approximate prescribed size and enable an effective estimate for dimension. The presence of a so-called blender-like horseshoe will be essential in the case when the basic set in the base is a Cantor set.

References

[1] T. Bedford, *On the box dimension of self-affine graphs and repellers*, Nonlinearity **2** (1989), 53–71.
[2] J. L. Kaplan, J. Mallet-Paret, and J. A. Yorke, *The Lyapunov dimension of a nowhere differentiable attracting torus*, Ergodic Theory Dynam. Systems **4** (1984), 261–281.
[3] H. McCluskey and A. Manning, *Hausdorff dimension for horseshoes*, Ergodic Theory Dynam. Systems **3** (1983), 251–260.
[4] J. Palis and M. Viana, *On the continuity of Hausdorff dimension and limit capacity for horseshoes*, Dynamical Systems (Valparaiso, 1986) (Lecture Notes in Mathematics, 1331). Eds. R. Bamón, R. Labarca and J. Palis. Springer, Berlin, 1988, pp. 150–160.

Universal Systems, Constructions and Obstacles

Jacek Serafin
(joint work with Tomasz Downarowicz)

The main part of the talk is based on [1]. We call a topological system (Z, F) *universal* for a class $\mathcal{H}$ of ergodic measure-theoretic systems if its simplex of invariant measures contains, up to an isomorphism, all elements of the class $\mathcal{H}$ and no elements from outside this class (to this end, Z is a compact metric space and F is a homeomorphism of Z).

Note that this differs from a classical notion of universality which has been studied in various contexts, in particular for toral automorphisms or systems with specification. So, classically, (Z, F) is universal if for every ergodic system $(\mathcal{X}, \mathcal{F}, \mu, T)$, where $T : \mathcal{X} \to \mathcal{X}$ is an automorphism of a standard nonatomic probability space $(\mathcal{X}, \mathcal{F}, \mu)$ whose Kolmogorov–Sinai entropy $h_\mu(T)$ is strictly smaller than the topological entropy $h_{top}(Z, F)$, there exists an F-invariant measure ν on Z such that the systems $(\mathcal{X}, \mathcal{F}, \mu, T)$ and $(Z, \mathcal{B}_\nu, \nu, F)$, with $\mathcal{B}_\nu$ denoting the Borel sigma-algebra in Z completed with respect to ν, are measure-theoretically isomorphic. Note that the Krieger Finite Generator Theorem is equivalent to saying that the full shift on n symbols is universal in the class of systems with entropy smaller than $\log n$.

We make the definition of universality for a class $\mathcal{H}$ more precise and subtle by adding the requirement that for every ergodic F-invariant probability measure ν on Z, the system $(Z, \mathcal{B}_\nu, \nu, F)$ belongs to $\mathcal{H}$. Note now that the above mentioned full shift is no longer universal in the class $\mathcal{H} = \{h < \log n\}$ as the uniform Bernoulli measure has entropy equal $\log n$; on the other hand the class $\mathcal{H} = \{h \le \log n\}$ contains products of uniform Bernoulli measure with zero-entropy systems, such products are not supported by the simplex of the full n-shift.

Our main result states that if $I = [c, C]$, with $0 \le c < C \le \infty$, then there exists a universal system for the class $\mathcal{H}_I$ of systems with measure-theoretic entropies belonging to I. We construct the universal system in the form of an inverse limit of subshifts (we do not know whether the universality can be achieved in the form of a subshift); the construction depends on a delicate argument that all systems with entropies in I have a Jewett-Krieger representation as a limit of strictly ergodic subshifts projecting on the fixed strictly ergodic model for the Bernoulli shift with entropy equal c.

On the other, we recall the result in [2], which treats the case $c = C = 0$. It turns out that there does not exist a universal model for the class of zero-entropy systems. Here is the outline of the argument: any topological system of zero entropy is a topological factor of a zero-entropy subshift. This subshift has its symbolic complexity $(U(n))$ (the sequence of the number of blocks of increasing lengths, occurring in the system); if the subshift supports an invariant measure and we partition the space into finitely many elements, then the measure-theoretic complexity (as defined by Ferenczi) with respect to that partition must is always bounded above by the symbolic complexity. Finally we show that there is no universal bound on the measure-theoretic complexity among the zero-entropy systems: given a sequence $(U(n))$ such that $\frac{1}{n}\log(U(n)) \to 0$ we construct a measure-theoretic zero-entropy system and a partition such that the measure-theoretic complexity is not bounded above by the sequence $(U(n))$.

In the final part (based on [3]) of the talk we recall an elementary construction of a finite generating partition for an ergodic finite-entropy automorphism of a probability space (one could call it a weak form of the Krieger Generator Theorem, as we do not produce a generator with the minimal possible number of elements). We are able to reduce the cardinality of the generator from countable to finite by applying the Kraft inequality (a well-known tool in information theory) and producing a prefix-free set of codewords, with the length function satisfying the requirements needed to run the coding procedure. All necessary calculations rely on a simple application of the ergodic theorem, and on the Kac recurrence lemma.

References

[1] T. Downarowicz, J.Serafin, *Universal Systems for Entropy Intervals*, Journal of Dynamics and Differential Equations **29** (2017), 1411–1422.

[2] J. Serafin, *Non-Existence of a Universal Zero-Entropy System*, Israel Journal of Mathematics **194** (2013), 349–358.

[3] M. Keane, J. Serafin, *Generators*, Periodica Mathematica Hungarica **44** (2002), 187–195.

A general connection between channel capacity and dynamical entropy

Christoph Kawan

We derive a very general lower bound on the smallest channel capacity of a noise-free channel above which an unspecified control objective can be achieved via a periodic coding scheme. This lower bound establishes a general connection between some type of dynamical entropy of the closed-loop system and the capacity of the channel. The system considered here is of the general form

$$x_{t+1} = f(x_t, u_t, w_t),$$

where $x_t \in X$, $u_t \in U$ and $w_t \in W$, X and U being measurable spaces and W a probability space with probability measure ν. We assume that $(w_t)_{t\in\mathbb{Z}_+}$ is an i.i.d. sequence of random variables with $w_t \sim \nu$ and that the initial state x_0 is a random variable with distribution $x_0 \sim \pi_0$, independent of (w_t). The coder is assumed to encode the state in a periodic fashion, applying a fixed quantizer at every time instant $k\tau$, $k \in \mathbb{Z}_+$, where τ is the sampling period. The encoded state is then sent through the channel during the forthcoming time interval of length τ, and we write $q_t \in \mathcal{M}$ (the coding alphabet) for the signal sent at time t. Assuming that an unspecified control objective is achieved under these hypotheses, a lower bound on the channel capacity $C = \log_2 |\mathcal{M}|$ is obtained. Looking at the discrete random variable $q_{[0,k\tau-1]} = (q_0, q_1, \dots, q_{k\tau-1}) \in \mathcal{M}^{k\tau}$, we find that

$$C \geq \limsup_{k\to\infty} \frac{1}{k\tau} H(q_{[0,k\tau-1]}).$$

Conditioning on the noise (w_t), this leads to

$$C \geq \limsup_{k\to\infty} \frac{1}{k\tau} H(q_{[0,k\tau-1]}|(w_t)_{t\in\mathbb{Z}_+}) + \liminf_{k\to\infty} \frac{1}{k\tau} I\left[q_{[0,k\tau-1]}; (w_t)_{t\in\mathbb{Z}_+}\right].$$

Writing $\mathcal{W} = W^{\mathbb{Z}_+}$ and $\mu := \nu^{\mathbb{Z}_+}$, the first term on the right-hand side can be transformed into

$$\limsup_{k\to\infty} \frac{1}{k\tau} \int_{\mathcal{W}} H_{\pi_0}(\mathcal{C}^k(\bar{w}))\mathrm{d}\mu(\bar{w}), \tag{1}$$

where we integrate over the Shannon entropy of the dynamical partition $\mathcal{C}^k(\bar{w})$ obtained as the common refinement of the pullbacks of the coding partition under the closed-loop dynamics. The argument $\bar{w}$ indicates the noise realization. The key point of this derivation is that the expression (1) resembles very closely the measure-theoretic entropy of a random dynamical system (with respect to a partition). To be more precise, this random dynamical system is defined by equipping $\mathcal{W}$ with the standard left shift operator θ and regarding the closed-loop dynamics as a cocycle over the ergodic base $(\mathcal{W}, \mu, \theta)$.

The only, but crucial difference to the entropy of a random dynamical system consists in the fact that π_0 is usually not a stationary measure for the state process. In fact, even when f is assumed to be continuous and X to be compact, the existence of a stationary measure is not guaranteed, since the closed-loop dynamics is discontinuous.

However, in the situation when there is no control, i.e., $f(x,u,w) = g(x,w)$, and the control task reduces to a computational task (estimation, for instance), this problem does not appear and we may assume that the initial distribution π_0 is a stationary measure of the state process (or something slightly weaker). In this case, the $\limsup$ in (1) is a limit and the expression equals the measure-theoretic entropy with respect to the coding partition (modulo the factor τ). In this case, results obtained previously for state estimation objectives can be recovered as special cases of our estimate, see [1, 2, 3].

References

[1] C. Kawan, S. Yüksel, *Entropy bounds on state estimation for stochastic non-linear systems under information constraints.* Submitted, 2016. arXiv:1612.00564

[2] A. Matveev, A. Pogromsky, *Observation of nonlinear systems via finite capacity channels: constructive data rate limits.* Automatica J. IFAC 70 (2016), 217–229.

[3] A. V. Savkin, *Analysis and synthesis of networked control systems: Topological entropy, observability, robustness and optimal control.* Automatica J. IFAC 42 (2006), no. 1, 51–62.

An informational perspective on uncertainty in control

Gireeja Ranade
(joint work with Jian Ding, Victoria Kostina, Yuval Peres, Govind Ramnarayan, Anant Sahai, Mark Sellke, and Alex Zhai.)

High-performance cyber-physical systems rely on many sensors, actuators and hardware components for successful operation. Control strategies for these devices require an understanding of how unpredictability in these components might impair performance. Our aim is to quantify the informational bottlenecks imposed by unpredictable system models in a manner that is compatible with standard information-theoretic tools for understanding communication and uncertainty limits in systems.

We will see examples of systems where it is useful to consider the weaker notion of "stability-in-probability," since the standard second-moment notion of stability can be too conservative. This motivates a notion of "control capacity," as provides a fundamental limit on a controller's ability to stabilize a system with random time-varying parameters (modeled as multiplicative noise). We can use this to quantify the value of side-information regarding the uncertainty in the system (in bits), in order to answer questions such as: "what is the value of adding an extra sensor to the system?" We will show in a simple case of vector control with dropped packets, non-causal information about uncertainty in the control channel can improve the performance of the controller.

I will also contrast systems with noisy actuation (e.g., when motors on a drone cannot precisely execute control actions) to noisy sensing (e.g., miscalibrated cameras). We use techniques from information-theory and probability-theory to show that these systems exhibit surprisingly different behavior — a fact that the linear control perspective does not reveal. If time permits, I will also consider the case where the system gain itself is unpredictable.

Open Problem

The curious difference between the optimality of linear strategies for multiplicative actuation noise and their suboptimality in the case of multiplicative observation noise is only partially understood. Consider the second-moment stability of the scalar system:

$$X_{t+1} = aX_n + U_n$$
$$Y_n = C_n X_n.$$

Here let the variables C_n be drawn i.i.d. from some continuous distribution, such as $\mathcal{N}(1,1)$ or $Unif[1,2]$. U_n can be any causal function of the observation $Y_0 \cdots Y_n$. In this case, we do not know the optimal control strategy for the system, nor do we know that maximum system gain a is, such that the system can be stabilized. We know from [3] that non-linear strategies can outperform linear strategies. However, there is a large gap between the achievable strategy and the converse bound provided in [3]. We also know that among linear strategies, memoryless ones are optimal, but that non-linear strategies with memory can outperform these. The open question remains: what is the optimal control strategy for the system? Can we provide a tight impossibility bound?

References

[1] G. Ranade, A. Sahai, *Control capacity*, arXiv:1701.04187.

[2] G. Ramanarayan, G. Ranade, A. Sahai, *Side-information in control and estimation*, ISIT 2014.

[3] J. Ding, Y. Peres, G. Ranade, A. Zhai, *When multiplicative noise stymies control*, arXiv:1612.03239.

Coding near Shannon-theoretic limits in control

Victoria Kostina

(joint work with B. Hassibi, A. Khina, A. Khisti, E. R. Gårding, G. M. Pettersson, Y. Nakahira, F. Xiao, J. C. Doyle)

In remote control of stochastic processes, the observer encodes its observations of a stochastic process under a rate constraint and a stringent delay constraint. The controller aims to bring the process to the target while only having access to these encoded observations.

A simple model of dynamical systems is that of a linear stochastic system, that evolves according to

$$X_{t+1} = aX_t + U_t + V_t, \tag{1}$$

where X_t is the system state, U_t is the control signal, chosen based on the entire history of the data received by the controller, V_t is the system noise, which we assume to be independent and identically distributed; $a > 1$ is a constant known as the system gain. The goal is to minimize, under communication rate constraints, the average mean-square deviation of system state from the target state, 0.

We adopt a Shannon-theoretic view of remote stochastic linear control. We introduce an information-theoretic measure for this scenario, a certain rate-cost function, $\mathbb{R}(b)$, which is equal to the minimum communication rate (directed information rate through the feedback loop of the dynamical system) required to attain mean-square cost b. For example, for the system in (1) we prove that $\mathbb{R}(b)$ is bounded as [1]:

$$\mathbb{R}(b) \geq \log a + \frac{1}{2} \log \left(1 + \frac{N(V)}{b - \mathrm{Var}[V]} \right), \tag{2}$$

where $N(V)$ and $\mathrm{Var}[V]$ is the entropy power and the variance of the common distribution of $V_1, V_2, \ldots$, respectively.

We show impossibility (converse) theorems linking this informational quantity to operational meaning (that is, the number of bits sent) over both noiseless and noisy rate-limited channels.

For control over several channels of interest, namely, variable-length rate-limited noiseless channels [1], Gaussian channels [2], rate-limited packet drop channels [3], and biomolecular channels [4], we propose coding strategies that can approach the bound in (2) from above.

While our techniques generalize to vector linear systems with general quadratic cost, many open questions remain. These include practical coding schemes for control over general noisy channels, rate-limited control of nonlinear systems, and joint optimal sampling / coding / control strategies.

References

[1] V. Kostina and B. Hassibi, "Rate-cost tradeoffs in control," in *Proceedings 54th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, Oct. 2016.

[2] A. Khina, G. M. Pettersson, V. Kostina, and B. Hassibi, "Multi-rate control over AWGN channels: An analog joint source-channel coding perspective," in *Proceedings 2016 IEEE Conference on Decision and Control*, Las Vegas, NV, Dec. 2016.

[3] A. Khina, V. Kostina, A. Khisti, and B. Hassibi, "Sequential coding of Gauss–Markov sources with packet erasures and feedback," in *Proceedings 2017 IEEE Information Theory Workshop*, Kaohsiung, Taiwan, Nov. 2017.

[4] Y. Nakahira, F. Xiao, V. Kostina, and J. Doyle, "Fundamental limits and achievable performance in biomolecular control," in *Proceedings 2018 American Control Conference*, Milwaukee, WI, June. 2018, to appear.

Tradeoffs in Networked Control Systems: An Information Theoretic Approach

Hideaki Ishii

(joint work with Song Fang and Jie Chen)

Today, for remotely controlling physical systems in real time, communication networks play an indispensable role to connect numerous sensors and actuators with controllers. In such networked control systems, the quality in communication can have significant impacts on the achievable control performances. The analysis of such limitations and tradeoffs arising in networked control lies at the intersection of both control theory and information theory.

In this talk, we focus on the so-called Bode integral through an information theoretic approach. Bode integral is a classical result, representing one of the most fundamental limitations in feedback control systems [1, 8]. It shows that performance in terms of disturbance attenuation cannot be improved over all frequency ranges by any stabilizing controllers. We present our recent works on novel Bode-type integral inequalities, developed for networked control systems [3, 4, 5].

Our approach explicitly takes account of the noises introduced at communication channels by employing tools such as entropy rates [2] of signals [9, 7, 6]. This enables us to derive general results for a broad class of systems consisting of linear time-invariant plants and causal, possibly nonlinear and time-varying controllers communicating over general noisy channels through encoders and decoders. We introduce novel measures for the quality of communication such as blurredness and negentropy rates and characterize their effects on Bode-type integrals. We further discuss performance bounds for general networked systems in control as well as estimation problems.

References

[1] H. W. Bode, *Network Analysis and Feedback Amplifier Design*, D. Van Nostrand, 1945.
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, 1991.
[3] S. Fang, J. Chen, and H. Ishii, *Towards Integrating Control and Information Theories: From Information-Theoretic Measures to Control Performance Limitations*, Lecture Notes in Control and Information Sciences, vol. 465, Springer, 2017.
[4] S. Fang, J. Chen, and H. Ishii, Design constraints and limits of networked feedback in disturbance attenuation: An information theoretic analysis, *Automatica*, 79: 65–77, 2017.
[5] S. Fang, H. Ishii, and J. Chen, Trade-offs in networked feedback systems: From information-theoretic measures to Bode-type integrals, *IEEE Trans. Automatic Control*, 62: 1046–1061, 2017.
[6] H. Ishii, K. Okano, and S. Hara, Achievable sensitivity bounds for MIMO control systems via an information theoretic approach, *Systems & Control Letters*, 60: 111–118, 2011.
[7] N. C. Martins, M. A. Dahleh, and J. C. Doyle, Fundamental limitations of disturbance attenuation in the presence of side information. *IEEE Trans. Automatic Control*, 52: 56–66, 2007.
[8] M. M. Seron, J. H. Braslavsky, G. C. Goodwin, *Fundamental Limitations in Filtering and Control*, Springer, 1997.
[9] G. Zang and P. A. Iglesias, Nonlinear extension of Bode's integral based on an information-theoretic interpretation, *Systems & Control Letters*, 50: 11–19, 2003.

Implicit Communication in Decentralized Control

Anant Sahai
(joint work with Pulkit Grover, Se Yong Park)

The real-world problem of decentralized control involves a group of agents trying to make a physical system behave in a desired manner by working together. The main contrast with centralized control is that each agent must act on the basis of only the local information available to it. The following discrete-time example illustrates the issues here:

Let $X(t)$ be the scalar state at time t, where t ranges over the natural numbers. The evolution of the state is given by:

$$X(t+1) = \lambda X(t) + U_a(t) + U_b(t) + W(t) \tag{1}$$

where the $W(t)$ are independent and identically distributed Gaussian random variables with zero mean and unit variance. The initial condition $X(0)$ can be taken to be a random Gaussian random variable (independent of the $W(t)$ random variables) with zero mean and variance σ_0^2. The constant λ represents the open-loop system gain. The $U_a(t)$ and $U_b(t)$ are the control inputs applied by controllers a and b respectively at time t, where each of these are allowed to depend only on the sequence of past and present observations $Y_a(0,1,\ldots,t)$ for $U_a(t)$ and similarly $Y_a(0,1,\ldots,t)$ for $U_b(t)$. These observations are noisy versions of the state:

$$Y_a(t) = X(t) + V_a(t) \tag{2}$$

$$Y_b(t) = X(t) + V_b(t) \tag{3}$$

where the $V_a(t)$ are also independent and identically distributed Gaussian random variables with zero mean and variance σ_a^2; and the $V_b(t)$ are independent and identically distributed Gaussian random variables with zero mean and variance σ_b^2. The $\{V_a(t)\}$ and $\{V_b(t)\}$ are independent of each other as well as the $\{W(t)\}$ and initial condition $X(0)$.

The goal is to minimize the long-term average expected cost

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} E[X^2(t)] + Q_a E[U_a^2(t)] + Q_b E[U_b^2(t)] \tag{4}$$

by suitably choosing the control functions that give rise to the U_a and U_b sequences from observations.

Although this problem is still very much open, it is possible to give nonlinear control strategies that depend on the system gain λ, observation noise variances σ_a^2, σ_b^2 and the control costs Q_a, Q_b, that achieve performance no more than a constant factor worse than the optimal cost no matter what values these problem parameters take.

It turns out that the conceptual heart of the problem comes from cases where Q_a is high (it is painful for the first controller to act) but σ_a^2 is small (the first controller can see what the state is quite accurately; while the situation is reversed for the second controller. In these cases, intuitively, the first controller would like to somehow count on the second controller to do the work of reducing the state while using its own limited control actions to somehow implicitly communicate its better quality observation to the second controller. This is precisely the context that is distilled into the famous Witsenhausen counterexample in a minimalist form.

Consider only three time steps: 0, 1, 2. After $X(0)$ is realized, only the first controller gets to act. For simplicity, there is no observation noise for this controller and so $Y_a(0) = X(0)$. The state evolves to $X(1) = X(0) + U_a(X(0))$. At the next time step, only the second controller gets to act, based on observation noise with variance σ_b^2. The state evolves to $X(2) = X(1) + U_b(Y_b) = X(0) + U_a + U_b$. The goal is to minimize $E[X^2(2)] + Q_a E[U_a^2(X(0))]$. There is no penalty on either the state $X(1)$ or the control U_b.

It turns out that this problem can be reinterpreted as a problem of "assisted interference suppression" in communication with the first controller playing the role of an encoder and the second controller playing the role of a decoder. Information-theoretic bounds can be derived leveraging rate-distortion theory (also called $\epsilon-$entropy by Kolmogorov). Furthermore, by looking at the infinite-dimensional limit of the problem (i.e. replacing all scalar random variables with high-dimensional vectors with i.i.d. distributions across the different dimensions), nonlinear control strategies can be found based on the idea of dirty-paper-coding (also called the Gaussian Gelfand-Pinsker channel) in communication. The basic information-theoretic bounds can then be tightened in the case of the scalar problem (and finite-dimensions more generally) by exploiting the ideas behind sphere-packing bounds derived through change-of-measure arguments. These can be used to establish constant-factor optimality for the Witsenhausen counterexample.

All of these results were developed with inspiration provided by simple bubble-diagrams that attempt to distill the informational interaction betwen real-valued random variables by viewing them as idealized bit strings written out in binary notation.

References

[1] P. Grover, A. Wagner, and A. Sahai *Information Embedding and the Triple Role of Control*, IEEE Transactions on Information Theory, **61** (April 2015), 1539–1549.

[2] P. Grover, S.Y. Park, and A. Sahai *Approximately Optimal Solutions to the Finite-Dimensional Witsenhausen Counterexample*, IEEE Transactions on Automatic Control **58** (Sep 2013), 2189–2204.

Invariance pressure

Alexandre J. Santana

(joint work with Fritz Colonius, Joao A.N. Cossich)

In this work we present the concepts of invariance pressure and outer invariance pressure for continuous time control systems based on weights for the control values. These concepts are an extension of the notion of invariance entropy (see e.g. [1] and [2]) and are inspired by the classical concept of topological pressure in the theory of dynamical systems (see [3] and [4]). The considered control systems are given by ordinary differential equations on a smooth manifold M. For a compact subset K of a set $Q \subset M$ and $\tau \in \mathbb{N}$ a set $\mathcal{S}$ of control functions is called (τ, K, Q)-spanning, if for each initial $x \in K$ there is a control ω in $\mathcal{S}$ such that the corresponding trajectory remains in Q up to time τ. Then, for a continuous real valued function f on the control range U and $(S_\tau f)(\omega) = \frac{1}{\tau}\int_0^\tau f(\omega(t))dt$ let

$$a_\tau(f, K, Q) = \inf\left\{\sum_{\omega\in\mathcal{S}} e^{(S_\tau f)(\omega)} \mid \mathcal{S} \text{ is } (\tau, K, Q)\text{-spanning}\right\}$$

and define the invariance pressure of f as

$$P_{inv}(f, K, Q) = \limsup_{\tau\to\infty} \frac{1}{\tau} \log a_\tau(f, K, Q).$$

The outer invariance pressure $P_{out}(f, K, Q)$ is obtained by replacing Q by an ε-neighborhood of Q and taking the limit for $\varepsilon \to 0$. It can be shown that many properties of topological pressure of dynamical systems have analogues for invariance pressure. Our main result characterizes the outer invariance pressure for linear control systems with compact control range U. Given a continuous function

$f : U \to \mathbb{R}$ that achieves its minimal value in an equilibrium which can be reached from every point of Q one finds for the outer invariance pressure

$$P_{out}(f, Q, Q) = \min_{u \in U} f(u) + h_{inv,out}(Q).$$

The outer invariance entropy $h_{inv,out}(Q)$ in this linear case is given by the unstable determinant (cf. [2]). Open questions in this area include upper and lower bounds for the invariance pressure analogous to known bounds for invariance entropy. Furthermore, the assumptions on the function f in the result presented above should be weakened.

References

[1] F. Colonius and C. Kawan, *Invariance entropy for control systems*, SIAM J. Control Optim, (2009), 1701-1721.

[2] C. Kawan, *Invariance Entropy for Deterministic Control Systems*, An Introduction, vol. 2089 of Lecture Notes in Mathematics, Springer-Verlag, 2013.

[3] M. Viana and K. Oliveira, *Foundations of Ergodic Theory*, Cambridge University Press, 2016.

[4] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, 1982.

Bit rate bounds for containability of scalar systems with event-based sampling

Steffen Linsenmayer

(joint work with Rainer Blind, Hideaki Ishii, Frank Allgöwer)

When studying stabilization problems for continuous-time control system, the state is usually sampled periodically before being coded and sent over a channel. In such a scenario, fundamental bounds on the necessary bit rate for stabilization are known. The study by Kofman and Braslavsky in [1] on the other hand, showed that using a sampling mechanism that employs state information, the necessary bit rate for stabilizing an unstable SISO control system can be made arbitrarily small. Recently, this initiated research on the influence of such state-dependent, often referred to as event-based, sampling strategies on the necessary bit rates for given control tasks.

Here we consider a setup, where the controller is assumed to be static and the coder and decoder are assumed to be memoryless. Such a setup is studied using a conventional sampling scheme in [2]. Therein the notion of containability is defined. According to that definition, a finite communication control system on $\mathbb{R}^n$ is containable if for any sphere N centered at the origin there exists an open neighborhood of the origin M and coding and feedback control laws such that any trajectory started in M remains in N for all time.

We discuss the influence of event-based sampling on bit rate bounds for containability of scalar, unstable, linear control systems. A first investigation confirms the findings of Kofman and Braslavsky for this scenario in the sense that, if no transmission delay is being present, the asymptotic average bit rate for containability can be made arbitrarily small. If delays are present, and one assumes that the only knowledge about the delays is a uniform lower and upper bound, this is not possible anymore. The required bit rate for this case can be quantified as a function of the delay bound [3].

After noticing that uncertainty in time affects the bit rate bound, a further study is conducted that analyzes the effect of uncertain system parameters. In this scenario, an interesting finding is given by the fact that if no delays are present it is again possible to make the necessary bit rate arbitrarily small, but only if one uses an alphabet which contains at least three symbols, while an alphabet consisting of two symbols was sufficient without uncertainty. We characterize the bit rate bounds with those two alphabets and draw conclusions regarding the optimal choice.

An issue that is discussed is the extension to higher order systems. It is possible to derive a procedure that preserves the general possibilities for a system with two unstable eigenvalues, but a generalization towards higher dimensions in the given setup remains an open problem.

References

[1] E. Kofman and J. H. Braslavsky, *Level crossing sampling in feedback stabilization under data-rate constraints*, in Proc. 45th IEEE Conf. Decision and Control (CDC), 2006, pp. 4423–4428.

[2] W. S. Wong and R. W. Brockett, *Systems with finite communication bandwidth constraints. II. Stabilization with limited information feedback*, IEEE Trans. Automat. Contr., vol. 42, no. 9, pp. 1294–1299, 1997.

[3] S. Linsenmayer, R. Blind, and F. Allgöwer, *Delay-dependent data rate bounds for containability of scalar systems*, in Proc. of the 20th IFAC World Congress, 2017, pp. 7875–7880.

Entropy structure

Tomasz Downarowicz

Suppose we observe a discrete time topological dynamical system (a homeomorphism T acting on a compact metric space X) and we need to send through a digital channel enough information, so that the receiver can identify the initial state within some ϵ-ball, where ϵ tends to zero with time. Such an encoding could be called a *lossless digitalization* of our system. How much capacity must the channel have (or what bitrate of the signal must be applied) to achieve this goal?

The answer to this very natural question relies on the *theory of symbolic extensions and entropy structure* developed in [2] and [3] and summarized in [4]. In the measure-theoretic setup, the famous Krieger Generator Theorem settles the question: an ergodic automorphism T with finite Kolmogorov–Sinai entropy $h_\mu(T)$ can be digitalized using $|\Lambda| = \lfloor e^{h_\mu(T)} \rfloor + 1$ symbols and the minimal capacity of the channel in question is precisely the Kolmogorov–Sinai entropy. By analogy, one might hope that in the topological setup the topological entropy $h_{top}(T)$ of the dynamical system (X, T) plays a similar role. However, an obstacle occurs: there is no hope to find (in general) a subshift which is topologically conjugate to (X, T). But one may still find a *symbolic extension* of (X, T), i.e., a subshift (Y, S)

and a (surjective) topological factor map $\pi : (Y,S) \to (X,T)$. Such an extension could play the role of a lossless digitalization of (X,T). Then the minimal capacity of a channel capable of sending the digital signal encoded in Y would be equal to the topological entropy $h_{top}(S)$ of (Y,S) and the number of symbols needed would be $\lfloor e^{h_{top}(S)} \rfloor + 1$. So the key question becomes this: what is the infimum of the topological entropies of symbolic extensions of (X,T)? Let us denote this infimum by $h_{sex}(T)$ and call it the *symbolic extension entropy* of (X,T). The most pending questions are: is $h_{sex}(X,T)$ always equal to $h_{top}(T)$ (or can it be larger)? Is it at least always finite when $h_{top}(T)$ is? (Infinite $h_{sex}(X,T)$ is equivalent to nonexistence of symbolic extensions.)

It turns out that the answers to both questions are negative. The first example of a systems with $h_{sex} > h_{top}$ is due to Mike Boyle [1], and can be described as follows: Let X be the set of 0-1 arrays $x = [x_{k,n}]_{k\in\mathbb{N},n\in\mathbb{Z}}$ satisfying the following rules: the first row (corresponding to $k = 1$) is an arbitrary 0-1 (bi-infinite) sequence. All other rows are filled with zeros, except when the first row is periodic with minimal period p. Then, in row number $1+p$ all 0-1 sequences are allowed. It is elementary to check that this set of arrays is closed and invariant under the horizontal shift $\sigma([x_{k,n}]) = [x_{k,n+1}]$. It is also easy to see, using the variational principle, that $h_{top}(X,T) = \log 2$. It takes a bit more effort to verify that any symbolic extension of this system must have entropy at least $\log 4$. A slightly modified version of this example produces a system with finite topological entropy and infinite symbolic extension entropy.

The main question in the theory of symbolic extensions is how to compute the value of $h_{sex}(X,T)$ in terms of internal properties of (X,T) (i.e., without constructing its symbolic extensions). The answer is surprisingly complicated and needs the notion of an "entropy structure", which we describe below. For simplicity, let us assume that (X,T) is zero-dimensional. Then it is an inverse limit of subshifts over finite alphabets: $(X,T) = \overleftarrow{\lim}_k(X_k,T_k)$. Each (X_k,T_k) is a subshift and a topological factor of (X_{k+1},T_{k+1}), as well as of (X,T). Every invariant measure μ on X (we will write $\mu \in \mathcal{M}_T(X)$) projects to a shift invariant measure μ_k on X_T (i.e., $\mu_k \in \mathcal{M}_{T_k}(X_k)$). On $\mathcal{M}_T(X)$ we can thus define a sequence of nonnegative entropy functions $h_k(\mu) = h_{\mu_k}(T_k)$. These functions are obviously affine, upper semicontinuous (this a well-known property of subshifts) and converge nondecreasingly to the entropy function $h(\mu) = h_\mu(T)$. By the *entropy structure* $\mathcal{H}$ of (X,T) we will mean the entire sequence of functions $\mathcal{H} = (h_k)_{k\geq 1}$. A function E on $\mathcal{M}_T(X)$ is called a *superenvelope* of $\mathcal{H}$ if $E \geq h$ and $E - h_k$ is upper semicontinuous for each k. It is not hard to see that the pointwise infimum of all superenvelopes of $\mathcal{H}$ is again a superenvelope of $\mathcal{H}$. We call it the *minimal superenvelope* of $\mathcal{H}$ and denote by $E_{\mathcal{H}}$. We are in a position to formulate the main theorem in the theory of symbolic extensions and entropy structure.

Symbolic Extension Entropy Theorem [2, 4]:

(1) If $\pi : (Y,S) \to (X,T)$ is a symbolic extension of (X,T), then the function $h^\pi(\mu) = \sup\{h_\nu(S) : \pi^*(\nu) = \mu\}$ (where π^* is the natural projection of

$\mathcal{M}_S(Y)$ onto $\mathcal{M}_T(X)$ determined by the factor map π) is a superenvelope of the entropy structure $\mathcal{H}$ on (X, T). Note that this function is also affine.

(2) Conversely, any affine superenvelope the entropy structure $\mathcal{H}$ equals the function $h^\pi(\mu)$ for some symbolic extension $\pi : (Y, S) \to (X, T)$.

(3) $h_{sex}(X, T) = \sup\{E_{\mathcal{H}}(\mu) : \ \mu \in \mathcal{M}_T(X)\}$.

We conclude with the remark that every nondecreasing sequence of affine upper semicontinuous functions defined on any Choquet simplex represents an entropy structure of some zero-dimensional dynamical system (X, T) (see [5]). Thus in order to give examples of phenomena associated with entropy structure and, in particular, with the symbolic extension entropy, it suffices to build examples of Choquet simplices and sequences of affine upper semicontinuous functions on them. Most of the interesting phenomena can be observed already on simplices with countably many extreme points, and this is done in the book [4].

References

[1] M. Boyle, Transparencies for a lecture given in a conference in Sapporo (1992) and private communiation.

[2] M. Boyle and T. Downarowicz, *The entropy theory of symbolic extensions, Inventiones Mathematicae* **156** (2004), 119–161.

[3] T. Downarowicz, *Entropy structure*, J. dAnalyse **96** (2005), 57–116.

[4] T. Downarowicz, *Entropy in Dynamical Systems*, New Mathematical Monographs 18, Cambridge University Press, Cambridge, 2011.

[5] T. Downarowicz and J. Serafin, *Possible entropy functions*, Israel J. Math. **135** (2003), 221–250.

Metric invariance entropy, quasi-stationary measures and control sets

Fritz Colonius

The topological notion of invariance entropy has been studied since some time and a number of results are available (cf. Kawan [6]). In particular, under hyperbolicity assumptions da Silva and Kawan [5] could show that the invariance entropy of control sets is determined by the exponential growth rate of the unstable determinant.

In view of the fruitful interplay between topological and measure theoretic versions of entropy of dynamical systems it seems desirable to develop also a measure-theoretic version of invariance entropy. The research reported undertakes some steps in that direction. Since here the behavior of trajectories within a non-invariant subset of the state space is of interest, I use a generalization of invariant measures given by quasi-stationary measures with respect to a given probability measure on the control range. Quasi–stationary measures are frequently employed in the theory of absorbed Markov processes, where they occur as Yaglom limits and describe the behavior under the condition that the trajectory remains in the considered subset. General references to quasi-stationary measures are the monograph [2] and the survey [1].

Compared to entropy for dynamical systems, in the construction of invariance entropy one replaces partitions and open covers by invariant partitions and invariant open covers, respectively, which use feedbacks keeping the system in the given subset of the state space up to a finite time. Due to quasi-stationarity, the relevant probability measure for the associated Shannon-entropy has to be weighted according to the considered time. Since the minimal required bit rate is of relevance for control theoretic purposes, the infimum of the associated bit rates over all invariant partitions is taken.

The main results show that this entropy, which is always bounded above by the topological invariance entropy, is invariant under measurable transformations and that it is already determined by control sets, which are certain subsets of the state space which are characterized by controllability properties.

Open problems in this field include the question if measure theoretic invariance entropy can be arbitrarily close to the topological version. Furthermore, there is a lack of explicit examples where the measure theoretic invariance entropy can be computed.

References

[1] N. Champagnat and D. Villemonais, *General criteria for the study of quasi-stationarity*, arXiv:1712.08092v2 [math.PR] 26 Jan 2018.

[2] P. Collett, S. Martinez, and J. San Martin, *Quasi-Stationary Distributions: Markov Chains, Diffusions, and Dynamical Systems*, Springer-Verlag, Berlin, 2013.

[3] F. Colonius, *Metric invariance entropy and conditionally invariant measures,* Ergodic Theory and Dynamical Systems. First published online: 20 October 2016. doi: 10.1017/etds.2016.72.

[4] F. Colonius, *Invariance entropy, quasi-stationary measures and control sets*, Discrete Contin. Dyn. Syst. A, 38 (2018), 2093-2123.

[5] A. da Silva and C. Kawan, *Invariance entropy of hyperbolic control sets*, Discrete Contin. Dyn. Syst. A, 36 (2016), 97–136.

[6] C. Kawan, *Invariance Entropy for Deterministic Control Systems*, An Introduction, vol. 2089 of Lecture Notes in Mathematics, Springer-Verlag, 2013.

Reporter: Steffen Linsenmayer