Estimating a linear process using phone calls

Mohammad Javad Khojasteh, Massimo Franceschetti, Gireeja Ranade

***Abstract*— We consider the problem of estimating an undisturbed, scalar, linear process over a "timing" channel, namely a channel where information is communicated through the timestamps of the transmitted symbols. Each transmitted symbol is received at the decoder subject to a random delay. The encoder can encode messages in the holding times between successive transmissions and the decoder must decode the message from the inter-reception times of successive symbols. This set-up is analogous to a telephone system where a transmitter signals a phone call to the receiver through a "ring" and, after the random time required to establish the connection, is aware of the "ring" being received. We show that for the estimation error to converge to zero in probability, the *timing capacity* of the channel should be at least as large as the entropy rate of the process. In the case the symbol delays are exponentially distributed, we show a tight sufficient condition using a random-coding strategy.**

I. INTRODUCTION

A networked control system with a feedback loop over a communication channel provides a first-order approximation of a cyber-physical system (CPS) [1], [2]. In this setting, data-rate theorems quantify the impact of the communication channel on the ability to stabilize (and estimate) the system state. Roughly speaking, these theorems state that to achieve estimation with arbitrary small error (and stabilization) the communication rate available in the feedback loop should be at least as large as the intrinsic entropy rate of the system, expressed by the sum of its unstable modes [3]–[6].

We consider a specific communication channel — a *timing channel*. Here, information is communicated through the timestamps of the symbols transmitted over the channel; the "time" is carrying the message. This formulation is motivated by recent works in event-triggering estimation and control, showing that the timing of the triggering events carries information that can be used for stabilization [7]–[12]. However, while in these works the timing information is not explicitly quantified, our goal is to precisely determine what is the value of a timestamp from an information-theoretic perspective, when this is used for estimation and control.

We consider estimation of a scalar, undisturbed, continuous-time process over a timing channel and rely on work in information theory that defines the *timing capacity* of the channel, namely the amount of information that can be encoded using time stamps [13]–[16]. In this setting, the encoder can communicate with the decoder by choosing the timestamps at which symbols from a unitary alphabet are transmitted. The decoder receives each transmitted symbol after a *random delay* is added to the timestamp. We show that in order to achieve arbitrary small estimation error the timing capacity should be proportional to the entropy-rate of the system, with a proportionality factor of at least one, that accounts for the difference in time scales. In the case the random delays are exponentially distributed, we show that a random coding strategy can be used to achieve this bound. While our analysis is restricted to transmitting symbols from a unitary alphabet, it is natural to extend this and develop "mixed" strategies that use both timing information and data payload, as in event-triggered estimation and control. Finally, our results show that state-dependent triggering is only one of many possible strategies to encode information in time, thus opening several new venues for future investigation.

A. Background

The books [3], [4], [17] and the surveys [5], [6] provide detailed discussions on data-rate theorems and related results. A portion of the literature studied estimation and stabilization over "bit pipe channels," where a rate-limited, possibly time-varying, noiseless communication channel is present [18]–[22]. In the case of noisy channels, Tatikonda and Mitter [23] showed that for almost sure (a.s.) estimation and stabilization of undisturbed linear systems the Shannon capacity of the channel should be larger than the entropy rate of the system. Matveev and Savkin [24] 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 of the channel must be larger than the entropy rate of the system [25]. Nair [26] derived a similar information-theoretic result in a deterministic setting. Sahai and Mitter [27] considered the less stringent requirement of moment-estimation (and 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 [28]–[32].

Another important aspect of CPS is event-triggered estimation and control [33], [34]. One primary focus of event-triggered control is on minimizing the number of transmissions while simultaneously ensuring the control objective [35], [36]. In this context, the works in [7]–[12] show that the timing of the state-dependent triggering events carries information that can be used for stabilization. The amount of timing information that is transmitted 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 and the data-rate must be increased [8], [12]. In this

M. J. Khojasteh and M. Franceschetti are with the Department of Electrical and Computer Engineering, University of California, San Diego. {mkhojasteh,massimo}@ucsd.edu

G. Ranade is with Microsoft Research and the EECS department at the University of California, Berkeley. Some of this work was performed while at the Simons Institute for the Theory of Computing. gireeja@eecs.berkeley.edu

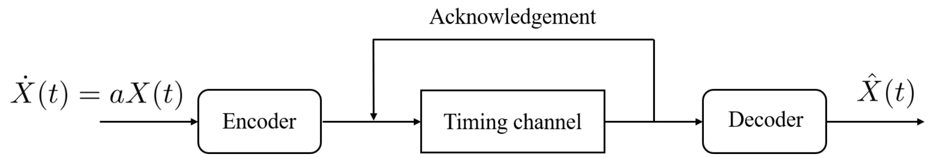

Fig. 1. The estimation problem.

context, our work explicitly quantifies the value of the timing information, independent of any transmission strategy, and also shows its dependence on the random delay, which plays the role of the channel noise in an information-theoretic setting.

In the remainder of the paper, Section II formulates the problem, Section III describes our results. Section IV concludes the paper with some open problems. Proofs of the results are omitted and can be found in [37].

B. Notation

Let $X^n = (X_1, \cdots, X_n)$ denote a vector of random variables and let $x^n = (x_1, \cdots, x_n)$ denote its realization. If the $X_1, \cdots, X_n$ are independent and identically distributed (i.i.d), then we refer to a generic $X_i \in X^n$ by X and skip the subscript i. We use $\log$ and $\ln$ to denote the logarithms to base 2 and e respectively. We use $H(X)$ to denote the Shannon entropy of a discrete random variable X and $h(X)$ to denote the differential entropy of a continuous random variable X. Further, we use $I(X,Y)$ for the mutual information between random variables X and Y. We will write $X_n \xrightarrow{P} X$ if X_n converges in probability to X. Similarly, we will write $X_n \xrightarrow{a.s.} X$ if X_n converges almost surely to X.

II. Problem formulation

We consider the estimation problem depicted in Fig. 1. The process dynamics are described by a scalar, continuous-time, noiseless process

$$\dot{X}(t) = aX(t), \tag{1}$$

where $X(t) \in \mathbb{R}$ is the process state. The constants $a \in \mathbb{R}$ such that $a > 0$. The initial state, $X(0)$, is random and is drawn from a distribution with bounded support, such that $|X(0)| < L$ and $h(X(0)) < \infty$. Conditioned on the realization of $X(0)$, the system evolves deterministically. The decoder has knowledge of system dynamics in (1).

We assume the encoder can measure the state of the system with infinite precision. The encoder is connected to the decoder through a timing channel (or the telephone signaling channel in [13]) as follows.

The encoder can choose to transmit the symbol ♠ at fixed times to the decoder. This symbol is delivered to the decoder after a random delay. The encoder receives an instantaneous but causal acknowledgement when the symbol is delivered, similar to [12], [27], [38].

The encoder uses a "waiting time" to encode information, i.e., after the i^{th} ♠ has been received by the decoder, the encoder waits for W_{i+1} seconds to transmit the next symbol. We assume that the channel is initialized with a symbol received at $t = 0$.

This is similar to a telephone system where a transmitter signals a phone call to the receiver through a "ring". After the random time required to establish the connection, the receiver is aware of the "ring". Thus, communication between the transmitter and receiver can occur without any vocal exchange, by encoding messages in the waiting times between consecutive calls.

Let D_i be the inter-reception time between two consecutive symbols, i.e.,

$$D_i = W_i + S_i, \tag{2}$$

where $\{S_i\}$ are random delays that are assumed to be i.i.d. Fig. 2 provides an example of the timing channel in action.

We assume the use of a random codebook, namely the holding times $\{W_i\}$ used to encode any given message are i.i.d. and also independent of the random delays $\{S_i\}$. This assumption is made to simplify our analysis, and does not change the capacity of the communication channel.

We assume the blocklength of a codeword is n, i.e. the decoder will use a set of n timestamps to decode the message. The reception time of the last symbol is given by $\mathcal{T}_n = \sum_{i=1}^{n} D_i$. We are interested in estimating the process (1) at a sequence of times $\{t_n\}$ such that

$$1 < \lim_{n\to\infty} \frac{t_n}{\mathbb{E}[\mathcal{T}_n]} \le \Gamma, \tag{3}$$

i.e. we want $|X(t_n) - \hat{X}(t_n)| \xrightarrow{P} 0$ as $n \to \infty$.

The following definitions are derived from [13], incorporating our random coding assumption.

Definition 1: A (n, M, T, δ)-i.i.d.-feedback-timing code for the telephone signaling channel consists of a codebook of M codewords $\{(w_i^{(m)}, i = 1, \ldots, n),\ m = 1 \ldots M\}$, where the symbols in each codeword are picked i.i.d. from a common distribution as well as a decoder, which upon observation of $(D_1, \ldots, D_n)$ selects the correct transmitted codeword with probability at least $1 - \delta$. Moreover, the codebook is such that the expected random arrival time of the n^{th} symbol, given by $\mathcal{T}_n = \sum_{i=1}^{n} D_i$, is not larger than T,

$$\mathbb{E}\left[\mathcal{T}_n\right] \le T. \tag{4}$$

Definition 2: The rate of an (n, M, T, δ)-i.i.d.-feedback-timing code is

$$R = (\log M)/T. \tag{5}$$

Definition 3: The timing capacity C of the telephone signaling channel is the supremum of the achievable rates, namely the largest R such that for every $\gamma > 0$ there exists

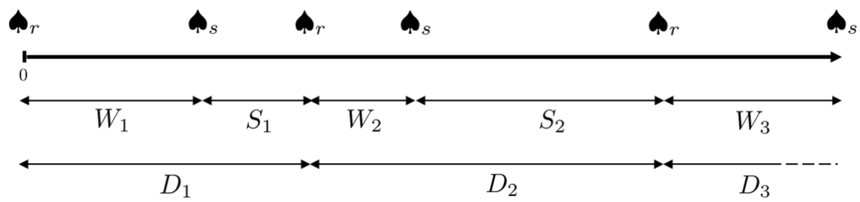

Fig. 2. The timing channel. Subscripts s and r are used to denote sent and received symbols, respectively.

a sequence of $(n, M_n, T_n, \delta_{T_n})$-i.i.d.-feedback-timing codes that satisfy

$$\frac{\log M_n}{T_n} > R - \gamma, \tag{6}$$

and $\delta_{T_n} \to 0$ as $n \to \infty$.

The capacity definition in [13] is slightly more general and does not include a random coding assumption. However, the following result [13, Theorem 8] applies to our random coding set-up, since the capacity in [13] is achieved by random codes.

Theorem 1 (Anantharam and Verdú): The timing capacity of the telephone signaling channel is given by

$$C = \sup_{\chi > 0} \sup_{\substack{W \geq 0 \\ \mathbb{E}[W] \leq \chi}} \frac{I(W; W+S)}{\mathbb{E}[S] + \chi}, \tag{7}$$

and if S is exponentially distributed then

$$C = \frac{1}{e\mathbb{E}[S]} \quad \text{[nats/sec]}. \tag{8}$$

III. Statement of the Results

The first theorem provides a necessary rate for the state estimation problem.

Theorem 2: Consider the estimation problem depicted in Fig. 1 with process dynamics (1). Consider transmitting n symbols over the telephone signaling channel (2), and a sequence of estimation times satisfying (3). If $|X(t_n) - \hat{X}(t_n)| \xrightarrow{P} 0$, then

$$I(W; W+S) \geq a\ \Gamma\ \mathbb{E}[W+S] \quad \text{[nats]}. \tag{9}$$

The entropy-rate of the process (1) is given by a nats/time [39]. Setting $\Gamma = 1$ our result recovers a scenario that parallels the data-rate theorems, stating that the mutual information between an encoding symbol W and its received noisy version $W + S$ should be larger than the average “information growth” of the state during the inter-reception interval D, which is given by

$$\mathbb{E}[aD] = a\ \mathbb{E}[S + W]. \tag{10}$$

In this case, using (7) we also obtain from (9) that

$$C \geq a \quad \text{[nats/sec]}. \tag{11}$$

On the other hand, for $\Gamma > 1$ our result shows that we must pay a penalty of Γ in the case that there is a time lag between the reception time $\mathcal{T}_n$ of the last symbol in the codeword and the observation time t_n, see Fig. 3. Using (7), in this case we obtain from (9) that

$$C \geq a\Gamma \quad \text{[nats/sec]}. \tag{12}$$

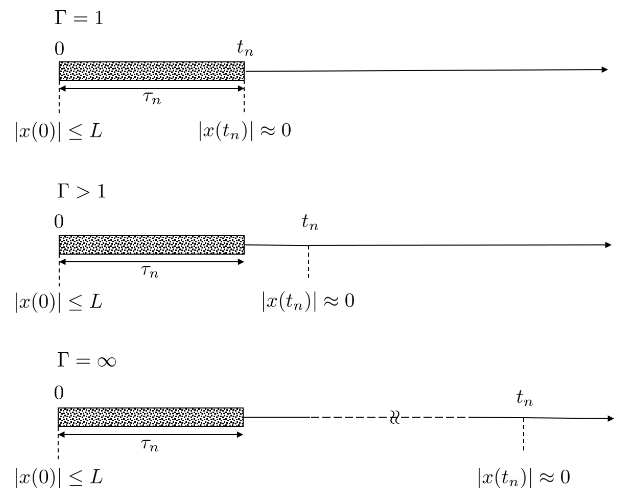

Fig. 3. Typical realization of codeword transmission for large n.

The case $\Gamma = \infty$ requires transmission of a codeword carrying an infinite amount of information over a channel of infinite capacity, thus revealing the initial state of the system with infinite precision. Once this state is known to the decoder, the estimation error can be kept small, even at observation times that are arbitrarily far in the future. This case is equivalent to transmitting a single real number over a channel without error.

The second theorem provides a sufficient condition for convergence of the estimation error to zero in probability for any sequence of times $\{t_n\}$ that satisfies (3) in the case of exponentially distributed delays.

Theorem 3: Consider the estimation problem depicted in Fig. 1 with process dynamics (1). Consider transmitting n symbols over the telephone signaling channel (2). Assume $\{S_i\}$ are drawn i.i.d. from an exponential distribution with mean $\mathbb{E}[S]$. If the capacity of the timing channel, C, is at least

$$C > a\Gamma \quad \text{[nats/sec]}, \tag{13}$$

then for any sequence of times $\{t_n\}$ that satisfies (3), we can compute an $\hat{X}(t_n)$ such that:

$$|X(t_n) - \hat{X}(t_n)| \xrightarrow{P} 0. \tag{14}$$

Remark 1: Tatikonda and Mitter in [23] considered the problem of estimation of the discrete-time version of the system in (1) over an erasure channel. In their model, at each time step of the system's evolution the encoder transmits a packet of I bits to the decoder and this is delivered with probability $1 - \mu$, or it is dropped with probability μ. It is shown that a necessary condition for $|X(k) - \hat{X}(k)| \xrightarrow{a.s} 0$ where $k \in \mathbb{N}$ is

$$(1-\mu)I \geq \log a \quad [\text{bits/sec}]. \tag{15}$$

For $\Gamma = 1$ in Theorem 2 we obtain the following necessary condition for $|X(t_n) - \hat{X}(t_n)| \xrightarrow{P} 0$:

$$\frac{I(W; W+S)}{\mathbb{E}[W+S]} \geq a \qquad [\text{nats/sec}], \tag{16}$$

for any sequence of times $\{t_n\}$ that satisfies (3). We now compare (15) and (16). The rate of expansion of the state space of the continuous time process is a nats per unit time, while for the discrete system is $\log a$ bits per unit time. Accordingly, in the case of (16) the decoder should receive at least $a\mathbb{E}[W+S]$ nats representing the initial state during a time interval of average length $\mathbb{E}[W+S]$. Similarly, in the case of (15) the decoder should receive at least $\log a/(1-\mu)$ bits representing the initial state over a time interval whose average length corresponds to the average number of trials before the first successful reception

$$(1-\mu)\sum_{k=0}^{\infty}(k+1)\mu^k = \frac{1}{1-\mu}. \tag{17}$$

•

IV. Conclusion

Event triggering policies for control exploit timing information related to the trigger-event to stabilize a system over a digital communication channel. This paper studied the fundamental limits of using timing information for estimation, as a precursor to control. We showed that for estimation of an undisturbed scalar linear system over a channel with a unitary alphabet, the timing capacity should be at least as large as the entropy rate of the system. In addition, in the case of exponentially distributed delay, we provided a tight sufficient condition. Important open problems for future research include understanding control of systems using timing information. In addition, it is important to consider the effect of system disturbances as well as considering symbols that carry a payload in addition to communicating a timestamp.

Acknowledgements

This research was partially supported by NSF award CNS-1446891. A part of this work was done while M. J. Khojasteh was visiting Microsoft Research. The authors would like to thank Yuval Peres, Serdar Yuksel and Girish Nair for helpful discussions.

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] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Springer Science & Business Media, 2013.

[4] A. S. Matveev and A. V. Savkin, *Estimation and control over communication networks*. Springer Science & Business Media, 2009.

[5] M. Franceschetti and P. Minero, "Elements of information theory for networked control systems," in *Information and Control in Networks*. Springer, 2014, pp. 3–37.

[6] 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.

[7] E. Kofman and J. H. Braslavsky, "Level crossing sampling in feedback stabilization under data-rate constraints," in *45st IEEE Conference on Decision and Control (CDC)*. IEEE, 2006, pp. 4423–4428.

[8] 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.

[9] Q. Ling, "Bit rate conditions to stabilize a continuous-time scalar linear system based on event triggering," *IEEE Transactions on Automatic Control*, 2016.

[10] M. J. Khojasteh, P. Tallapragada, J. Cortés, and M. Franceschetti, "Time-triggering versus event-triggering control over communication channels," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 5432–5437.

[11] S. Linsenmayer, R. Blind, and F. Allgöwer, "Delay-dependent data rate bounds for containability of scalar systems," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 7875–7880, 2017.

[12] M. J. Khojasteh, M. Hedayatpour, J. Cortés, and M. Franceschetti, "Event-triggered stabilization of disturbed linear systems over digital channels," in *Information Sciences and Systems (CISS), 2018 52nd Annual Conference on*. IEEE, 2018, pp. 1–6.

[13] V. Anantharam and S. Verdu, "Bits through queues," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 4–18, 1996.

[14] A. S. Bedekar and M. Azizoglu, "The information-theoretic capacity of discrete-time queues," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 446–461, 1998.

[15] T. J. Riedl, T. P. Coleman, and A. C. Singer, "Finite block-length achievable rates for queuing timing channels," in *Information Theory Workshop (ITW), 2011 IEEE*. IEEE, 2011, pp. 200–204.

[16] A. B. Wagner and V. Anantharam, "Zero-rate reliability of the exponential-server timing channel," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 447–465, 2005.

[17] H. Ishii and B. A. Francis, *Limited data rate in control systems with networks*. Springer Science & Business Media, 2002, vol. 275.

[18] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, 2004.

[19] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413–436, 2004.

[20] J. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," in *Proc. 15th Int. Symp. on Mathematical Theory of Networks and Systems (MTNS)*, 2002.

[21] P. Minero, M. Franceschetti, S. Dey, and G. N. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, p. 243, 2009.

[22] P. Minero, L. Coviello, and M. Franceschetti, "Stabilization over Markov feedback channels: the general case," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 349–362, 2013.

[23] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE transactions on Automatic Control*, vol. 49, no. 7, pp. 1196–1201, 2004.

[24] 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.

[25] ——, "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.

[26] G. Nair, "A non-stochastic information theory for communication and state estimation," *IEEE Transactions on Automatic Control*, vol. 58, pp. 1497–1510, 2013.

[27] 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.

[28] G. Como, F. Fagnani, and S. Zampieri, "Anytime reliable transmission of real-valued information through digital noisy channels," *SIAM Journal on Control and Optimization*, vol. 48, no. 6, pp. 3903–3924, 2010.

[29] R. Ostrovsky, Y. Rabani, and L. J. Schulman, "Error-correcting codes for automatic control," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 2931–2941, 2009.

[30] R. T. Sukhavasi and B. Hassibi, "Linear time-invariant anytime codes for control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3826–3841, 2016.

[31] A. Khina, W. Halbawi, and B. Hassibi, "(Almost) practical tree codes," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 2404–2408.

[32] P. Minero and M. Franceschetti, "Anytime capacity of a class of markov channels," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1356–1367, 2017.

[33] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1680–1685, 2007.

[34] W. P. M. H. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 3270–3285.

[35] P. Tallapragada and J. Cortés, "Event-triggered stabilization of linear systems under bounded bit rates," *IEEE Transactions on Automatic Control*, vol. 61, no. 6, pp. 1575–1589, 2016.

[36] J. Pearson, J. P. Hespanha, and D. Liberzon, "Control with minimal cost-per-symbol encoding and quasi-optimality of event-based encoders," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2286–2301, 2017.

[37] M. J. Khojasteh, M. Franceschetti, and G. Ranade, "Stabilizing a linear system using phone calls," *arXiv preprint arXiv:1804.00351*, 2018.

[38] Q. Ling, "Bit rate conditions to stabilize a continuous-time linear system with feedback dropouts," *IEEE Transactions on Automatic Control*, 2017.

[39] G. N. Nair, R. J. Evans, I. M. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1585–1597, 2004.