METRIC INVARIANCE ENTROPY AND CONDITIONALLY INVARIANT MEASURES

FRITZ COLONIUS*
INSTITUT FÜR MATHEMATIK, UNIVERSITÄT AUGSBURG, AUGSBURG, GERMANY

Abstract. Two notions of metric invariance entropy are constructed with respect to conditionally invariant measures for control systems in discrete time and it is shown that they are invariant under conjugacies.

April 26, 2016

Key words. invariance entropy, conditionally invariant measures, quasi-stationary measures, control systems

MSC 2010: 37A35, 37B40, 93C10, 94A17

1. Introduction. This paper proposes a notion of metric invariance entropy in analogy to the topological notion of invariance entropy of deterministic control systems, cf. Nair, Evans, Mareels, and Moran [15] and the monograph Kawan [11]. We consider control systems in discrete time of the form

$$x_{n+1} = f(x_n, u_n), n \in \mathbb{N}_0 = \{0, 1, \ldots\}, \tag{1.1}$$

where $f : M \times \Omega \to M$ is continuous and M and Ω are metric spaces. Abbreviate $f_\omega := f(\cdot, \omega) : M \to M$ for $\omega \in \Omega$, and for $u = (u_n)_{n\in\mathbb{N}_0} \in \mathcal{U} := \Omega^{\mathbb{N}_0}$ write the solutions as

$$\varphi : \mathbb{N}_0 \times M \times \mathcal{U} \to M, \quad \varphi(0, x, u) := x, \varphi(n, x, u) := f_{u_{n-1}} \circ \cdots \circ f_{u_0}(x) \text{ for } n \geq 1.$$

The system should be kept in a given subset Q of M. (In the literature there are colorful terms to describe this situation: One may think of $M \setminus Q$ as a "trap" or as a "hole" in the state space or leaving Q means "killing" the system). The notion of invariance entropy $h_{inv}(Q)$ of a compact subset $Q \subset M$ describes the average data rate needed to keep the system in Q (forward in time). It is constructed with some analogy to topological entropy of dynamical systems. This is done in Nair, Evans, Mareels, and Moran [15] via the version of Adler, Konheim and McAndrews [1], and in Kawan [11, 10] via the version due to Bowen and Dinaburg based on spanning sets. In the presence of hyperbolicity conditions, the invariance entropy has been discussed in da Silva and Kawan [4]. We refer, e.g., to Walters [18] and Downarowicz [7] for the entropy theory of dynamical systems. A major difference of entropy in a control context to entropy for dynamical systems is that the minimal required entropy for the considered control task is of interest, instead of the total entropy generated by the dynamical system.

Presumably, the earliest connection between control for deterministic systems and ergodic theory has been established by Delchamps [5] who studied when quantized feedbacks for stabilization lead to a nontrivial invariant measure. The contribution [14] by Mehta, Vaidya and Banaszuk considers measure-theoretic notions of entropy in the context of fundamental limitations in control. Here, however, the uncertainty arises due to disturbances, called conditional dither.

*Supported by DFG grant Co 124/19-1 and Brazilian - European Partnership in Dynamical Systems FP7-PEOPLE-2012-IRSES 318999 BREUDS (EU)

If one wants to construct a metric entropy, the choice of an appropriate probability measure is crucial. The present paper proposes to use conditionally invariant measures for this purpose. In the dynamical systems literature, conditionally invariant (also called relatively invariant) measures have been introduced by Pianigiani and Yorke [17]; cf. the survey Demers and Young [6] and also Keller and Liverani [12]. For random systems, the related notion of quasi-stationary distributions (or measures) is a classical topic; cf. the recent monograph Collett, Martinez, and San Martin [2] and also Zmarrou and Homburg [19]. Intuitively, quasi-stationary measures describe "the distribution of trajectories which are on the verge of falling in the trap" [2, p. 15].

Control system (1.1) may be viewed as a skew product dynamical system by considering the left shift θ on $\mathcal{U}$ given by

$$(\theta u)_n := u_{n+1}, n \in \mathbb{N}_0, \text{ for } u = (u_n) \in \mathcal{U}. \tag{1.2}$$

Then $S : (u, x) \mapsto (\theta u, f(x, u_0))$ is a skew product map on $\mathcal{U} \times M$ and its iterations define a skew product dynamical system. Note that the product topology makes $\mathcal{U} = \Omega^{\mathbb{N}_0}$ into a metrizable space which is compact if Ω is compact. If one wants to keep system (1.1) in a closed subset $Q \subset M$ it appears appropriate to look at conditionally invariant measures μ for S on $\mathcal{U} \times M$ with respect to $\mathcal{U} \times Q$. We construct a metric invariance entropy with respect to such a conditionally invariant measure. This is done using feedbacks and alternatively (open-loop) time-dependent control functions and results in two versions of metric entropy, a feedback invariance μ-entropy and a controlled invariance μ-entropy. It will be shown that each is invariant under appropriately defined conjugacies.

The main contribution of the present paper is the construction of metric invariance entropy. Also the existence of quasi-stationary measures is briefly discussed; they yield special conditionally invariant measures. The constructions for the metric invariance entropy are conveniently done for general conditionally invariant measures. In the monograph by Collett, Martinez, and San Martin other sufficient conditions for the existence of quasi-stationary measures are derived; cf. [2, Proposition 2.10 and Theorem 2.11]. Demers and Young [6] show that conditionally invariant measures always exist and discuss their properties mainly for deterministic maps and with regard to absolutely continuous conditionally invariant measures and their escape rates.

The contents of this paper is as follows: Section 2 discusses conditionally invariant measures for maps and for control systems; here also quasi-stationary measures are considered and notation is fixed. Section 3 constructs the metric invariance entropy and proves invariance under conjugacies.

2. Conditionally invariant measures. In this section we collect some basic information on conditionally invariant measures and fix some notation.

For a map $S : X \to X$ on a metric space (with metric d) and $A \subset X$ we let $S^{-1}A := \{x \in X \mid S(x) \in A\}$.

Definition 2.1. *Let $S : X \to X$ be a continuous map on a metric space X and consider a closed subset $Y \subset X$. A probability measure μ on X endowed with the Borel σ-algebra $\mathcal{B}(X)$ is called conditionally invariant with respect to Y with constant ρ if $0 < \rho = \rho(\mu) := \mu(S^{-1}Y \cap Y) \leq 1$ and*

$$\mu(A) = \frac{\mu(S^{-1}A \cap Y)}{\mu(S^{-1}Y \cap Y)} \text{ for all } A \in \mathcal{B}(X).$$

Often the number $\rho(\mu)$ is considered as an escape rate from Y. Putting $A = Y$

in Definition 2.1 one sees that the topological support of μ given by

$$\text{supp}\mu := \{x \in X \mid \mu(N) > 0 \text{ for each open set with } x \in N\}$$

is contained in Y. Hence, if we identify the probability measures on $\mathcal{B}(X)$ which have support in Y with the set $\mathcal{P}(Y)$ of probability measures on the Borel σ-algebra $\mathcal{B}(Y)$, a measure $\mu \in \mathcal{P}(Y)$ is conditionally invariant if and only if $\mu(S^{-1}Y) > 0$ and

$$\mu(A) = \frac{\mu(S^{-1}A)}{\mu(S^{-1}Y)} \text{ for all } A \in \mathcal{B}(Y).$$

Observe that we allow $\mu(S^{-1}Y \cap Y) = 1$ for a conditionally invariant measure μ on $\mathcal{B}(X)$, hence this includes invariant measures on Y.

The following proposition gives equivalent characterizations of conditionally invariant measures.

PROPOSITION 2.2. *Let S be a continuous map on X. Fix a closed subset Y of X and let $\mu \in \mathcal{P}(X)$. Suppose that $\rho = \mu(S^{-1}Y \cap Y) > 0$. Then the following are equivalent:*

(i) The measure μ is conditionally invariant for S with respect to Y.

(ii) For every bounded continuous real function $f \in C_b(X)$ with $f(x) = 0$ for $x \in X \setminus Y$

$$\mu(S^{-1}Y \cap Y) \int_Y f d\mu = \int_Y f \circ S \; d\mu.$$

(iii) The measure μ satisfies for every $k \in \mathbb{N}$ and every $A \in \mathcal{B}(Y)$

$$\mu\left(S^{-k}(Y) \cap ... \cap Y\right) \mu(A) = \mu\left(S^{-k}(A) \cap S^{-(k-1)}Y \cap ... \cap Y\right). \tag{2.1}$$

If (i) (or (ii), (iii)) holds, it follows that for all $k \in \mathbb{N}$

$$\mu\left(S^{-k}(Y) \cap ... \cap Y\right) = \mu(S^{-1}Y \cap Y)^k = \rho^k. \tag{2.2}$$

Proof. Let μ be conditionally invariant for Y. Then for every $A \in \mathcal{B}(Y)$ the characteristic function $\mathbf{1}_A$ satisfies

$$\int \mathbf{1}_A d\mu = \mu(A) = \frac{\mu(S^{-1}A)}{\mu(S^{-1}Y)} = \frac{\int (\mathbf{1}_A \circ S) \cdot \mathbf{1}_Y d\mu}{\mu(S^{-1}Y \cap Y)}.$$

The same is true for all simple functions, and then also for all integrable, hence for all bounded continuous functions. This shows that (i) implies (ii). The converse follows by approximating characteristic functions by continuous functions.

For $k = 1$, assertion (2.1) reduces to (i). In order to see that conversely (i) implies (iii), we proceed by induction. Suppose that (iii) holds for k. One finds for every $A \in \mathcal{B}(Y)$

$$\begin{aligned}&\mu\left(S^{-1}\left(S^{-k}(A) \cap S^{-(k-1)}Y \cap ... \cap Y\right) \cap Y\right)\\ &= \mu(S^{-(k+1)}A \cap S^{-k}Y \cap ... \cap S^{-1}(Y) \cap Y).\end{aligned}$$

Applying this also to $A = Y$ and using (i) one finds, as claimed,

$$\mu\left(S^{-(k+1)}Y \cap S^{-k}Y \cap ... \cap Y\right) \mu(A) = \mu\left(S^{-(k+1)}A \cap S^{-k}Y \cap ... \cap Y\right).$$

If (i) holds, assertion (2.2) is valid for $k = 1$, and if it holds for k it follows by (i) that

$$\begin{aligned}\mu\left(S^{-(k+1)}Y \cap S^{-k}Y \cap ... \cap Y\right) &= \mu(S^{-1}Y \cap Y)\mu\left(S^{-k}Y \cap ... \cap Y\right) \\ &= \mu(S^{-1}Y \cap Y)^{k+1}.\end{aligned}$$

□

REMARK 2.3. *If we introduce the restriction $S_Y := S_{|Y} : Y \to X$, we can rewrite the requirement for conditional invariance of μ as $\mu(S_Y^{-1}Y)\mu(A) = \mu\left(S_Y^{-1}(A)\right), A \in \mathcal{B}(Y)$, with $\rho = \mu(S_Y^{-1}Y) > 0$. Furthermore, with a slight abuse of notation, formula (2.1) can be written as*

$$\mu(S_Y^{-k}Y)\mu(A) = \mu\left(S_Y^{-k}(A)\right).$$

Thus μ is a conditionally invariant measure for the map S_Y^k with constant $\rho^k = \mu(S_Y^{-k}(Y))$. It is a trivial observation that $\mu(S_Y^{-k}Y) \to 0$ for $k \to \infty$ if $\rho < 1$. Furthermore, $\rho^{-k}\mu$ defines a probability measure on $S_Y^{-k}Y$.

The following proposition gives some information on the support of conditionally invariant measures (cf. Demers and Young [6, p. 380]).

PROPOSITION 2.4. *For a conditionally invariant measure $\mu \in \mathcal{P}(X)$ with respect to a compact set Y one has*

$$\text{supp}\mu \subset \{x \in Y \mid S^{-n}(x) \cap Y \neq \emptyset \text{ for all } n \in \mathbb{N}\}.$$

Proof. We first show that for every $n \in \mathbb{N}$

$$\text{supp}\mu \subset \{x \in Y \mid S^{-n}(x) \cap Y \neq \emptyset\}. \tag{2.3}$$

The set $E_n := \{x \in Y \mid S^{-n}(x) \cap Y = \emptyset\}$ satisfies $\mu(S_Y^{-n}E_n) = 0$, and hence

$$\mu(E_n) = \frac{\mu(S_Y^{-n}E_n)}{\left(\mu(S_Y^{-1}Y)\right)^n} = 0.$$

The complement of E_n in Y is closed: Consider $x_k \to x$ in Y such that there are $y_k \in S^{-n}(x_k) \cap Y$, hence $S^n(y_k) = x_k$. Then a subsequence (y_{k_i}) converges to some $y \in Y$ and hence $S^n(y_{k_i}) \to S^n(y) = x$ for $i \to \infty$. Thus $S^{-n}(x) \cap Y \neq \emptyset$. This shows that E_n is open and assertion (2.3) follows. Furthermore, σ-additivity implies $\mu(E) = 0$ for the open set

$$E := \{x \in Y \mid \text{ there is } n \in \mathbb{N} \text{ with } S^{-n}(x) \cap Y = \emptyset\} = \bigcup_{n \geq 1} E_n.$$

□

In the rest of this section we consider control systems. Control system (1.1) can be described by the continuous skew product map S defined by

$$S : \mathcal{U} \times M \to \mathcal{U} \times M, (u, x) \mapsto (\theta u, f(x, u_0)), \tag{2.4}$$

where the shift θ is given by (1.2). Similarly to Remark 2.3, we also write $S_Q := S_{|\mathcal{U}\times Q} : \mathcal{U} \times Q \to \mathcal{U} \times M$ for the restriction. Thus $S^{-1}(\mathcal{U} \times Q) \cap (\mathcal{U} \times Q) = S_Q^{-1}(\mathcal{U} \times Q)$.

DEFINITION 2.5. *A conditionally invariant measure μ for the map S defined in (2.4) with respect to a closed subset Q of M is a probability measure on the Borel σ-algebra of $\mathcal{U} \times M$ such that $0 < \rho := \mu(S^{-1}(\mathcal{U} \times Q) \cap (\mathcal{U} \times Q)) \leq 1$ and*

$$\rho\mu(B) = \mu(S^{-1}B \cap (\mathcal{U} \times Q)) \text{ for all } B \in \mathcal{B}(\mathcal{U} \times M).$$

Thus μ is a special case of the measures specified in Definition 2.1 with $Y = \mathcal{U} \times Q \subset X = \mathcal{U} \times M$. Any probability measure μ with marginal ν on $\mathcal{U}$ can be disintegrated in the form

$$\mu(B) = \int_{\mathcal{U}} \int_{M} \mathbf{1}_B(u,x)\eta_u(dx)\nu(du), B \in \mathcal{B}(\mathcal{U} \times M), \tag{2.5}$$

where η_u are probability measures on M and for all $A \in \mathcal{B}(M)$ the real map $u \mapsto \eta_u(A)$ is measurable with respect to $\mathcal{B}(\mathcal{U})$; the measures η_u are uniquely determined ν-almost everywhere; cf. Gänssler and Stute [9, p. 196].

As an example we consider random maps of the form

$$x_{n+1} = f(x_n, u_n), \tag{2.6}$$

where $f : M \times \Omega \to M$ is as above and a probability measure ν_0 on the Borel σ-algebra $\mathcal{B}(\Omega)$ of Ω is fixed.

Let $p(x, A) := \nu_0 \{\omega \in \Omega \mid f(x, \omega) \in A\}, x \in M$, $A \in \mathcal{B}(M)$, be the associated Markov transition probability. Recall that a stationary measure $\eta \in \mathcal{P}(M)$ is a probability measure such that

$$\eta(A) = \int_M p(x, A)\eta(dx) \text{ for all } A \in \mathcal{B}(M). \tag{2.7}$$

It is well known that the stationary measures uniquely correspond to the invariant Borel measures $\nu \times \eta, \nu := \nu_0^{\mathbb{N}_0}$, of the skew product map S defined in (2.4); cf. Kifer [13].

Definition 2.6. *Let Q be a closed subset of M. A quasi-stationary measure with respect to Q for the random map (2.6) is a probability measure η on $\mathcal{B}(M)$ such that $0 < \int_Q p(x, Q)d\eta \leq 1$ and*

$$\eta(A) = \frac{\int_Q p(x, A)\eta(dx)}{\int_Q p(x, Q)\eta(dx)} \text{ for all } A \in \mathcal{B}(Q).$$

Putting $A = Q$, one sees that the support of η is contained in Q. Observe that $1 - \int_Q p(x, Q)\eta(dx)$ is the average probability to exit in one step from Q. The measure η is stationary if and only if $\int_Q p(x, Q)\eta(dx) = 1$.

The following proposition shows that quasi-stationary measures correspond to conditionally invariant measures for S. The proof is included for the reader's convenience, although it essentially coincides with the one for Zmarrou and Homburg [19, Lemma 5.2].

Proposition 2.7. *A probability measure $\eta \in \mathcal{P}(Q)$ is quasi-stationary with respect to Q for the random map (2.6) if and only if the probability measure $\mu := \nu \times \eta, \nu := \nu_0^{\mathbb{N}_0}$, is conditionally invariant with respect to Q for the skew product map S in (2.4). In this case, it follows that $\rho = \mu(S_Q^{-1}(\mathcal{U} \times Q)) = \int_Q p(x, Q)\eta(dx)$.*

Proof. For any probability measure η on M and $A \in \mathcal{B}(M)$ one has

$$\begin{aligned}
\int_M p(x, A)\eta(dx) &= \int_M \int_\Omega \mathbf{1}_{\{\omega \in \Omega \mid f(x,\omega) \in A\}} \nu_0(d\omega)\eta(dx) \\
&= \int_M \int_{\mathcal{U}} \mathbf{1}_{\{u \in \mathcal{U} \mid f(x,u_0) \in A\}} \nu(du)\eta(dx) \\
&= \int_M \int_{\mathcal{U}} \mathbf{1}_{S^{-1}(\mathcal{U} \times A)} \nu(du)\eta(dx) \\
&= (\nu \times \eta)\left(S^{-1}(\mathcal{U} \times A)\right).
\end{aligned} \tag{2.8}$$

Suppose that $\nu \times \eta$ is conditionally invariant for S with respect to Q. Then η is quasi-stationary, since formula (2.8) implies that for all $A \in \mathcal{B}(Q)$

$$\eta(A) = (\nu \times \eta)(\mathcal{U} \times A) = \frac{(\nu \times \eta)(S^{-1}(\mathcal{U} \times A))}{(\nu \times \eta)(S^{-1}(\mathcal{U} \times Q))} = \frac{\int_Q p(x, A)\eta(dx)}{\int_Q p(x, Q)\eta(dx)}.$$

Conversely, suppose that $\eta \in \mathcal{P}(Q)$ is quasi-stationary, hence

$$\eta(A) = \frac{\int_Q \int_\Omega \mathbf{1}_A(f(x, \omega))\nu_0(d\omega)\eta(dx)}{\int_Q \int_\Omega \mathbf{1}_Q(f(x, \omega))\nu_0(d\omega)\eta(dx)} \text{ for all } A \in \mathcal{B}(Q). \tag{2.9}$$

Take a Borel set $V \times A$ with $V \subset \mathcal{U}$ and $A \subset Q$ and compute

$$\begin{aligned}(\nu \times \eta)\left(S^{-1}(V \times A)\right) &= (\nu \times \eta)\left(\bigcup_{u_0 \in \Omega} \left((\{u_0\} \times V) \times f^{-1}(A, u_0)\right)\right) \\ &= \nu(V) \cdot (\nu \times \eta)\left(\bigcup_{u_0 \in \Omega} \left(\{u_0\} \times f^{-1}(A, u_0)\right)\right) \\ &= \nu(V) \cdot \int_Q \int_\Omega \mathbf{1}_A(f(x, u_0))\nu_0(du_0)\eta(dx).\end{aligned}$$

Together with (2.9) this implies

$$\begin{aligned}(\nu \times \eta)(V \times A) = \nu(V)\eta(A) &= \frac{\nu(V)}{\nu(\mathcal{U})} \frac{\int_Q \int_\Omega \mathbf{1}_A(f(x, \omega))\nu_0(d\omega)\eta(dx)}{\int_Q \int_\Omega \mathbf{1}_Q(f(x, \omega))\nu_0(d\omega)\eta d(x)} \\ &= \frac{(\nu \times \eta)\left(S^{-1}(V \times A)\right)}{(\nu \times \eta)\left(S^{-1}(\mathcal{U} \times Q)\right)}.\end{aligned}$$

This is the assertion for Borel sets of the form $B = V \times A$ and hence it follows for all $B \in \mathcal{B}(\mathcal{U} \times Q)$. These arguments also prove the second assertion in the proposition. □

REMARK 2.8. *For the random map (2.6) we obtain the following result for the iterates of S_Q. Define for $x \in Q$ and $A \in \mathcal{B}(Q)$ the n-step transition probability $p_Q^{(n)}(x, A)$ iteratively by $p^{(0)}(x, A) = p(x, A)$ and*

$$p_Q^{(n+1)}(x, A) := \int_Q p_Q^{(n)}(y, A)p(x, dy), n \geq 0.$$

Then for a quasi-stationary measure η and for all $A \in \mathcal{B}(Q)$ and $n \geq 1$

$$\int_Q p_Q^{(n)}(x, A)\eta(dx) = (\nu \times \eta)(S_Q^{-n}(\mathcal{U} \times A)) \text{ and } \eta(A) = \frac{\int_Q p_Q^{(n)}(x, A)\eta(dx)}{\int_Q p_Q^{(n)}(x, Q)\eta(dx).} \tag{2.10}$$

Next we will briefly discuss the existence of quasi-stationary measures.

THEOREM 2.9. *Let $Q \subset M$ be a compact set and consider the random map (2.6) for a probability measure $\nu_0 \in \mathcal{P}(\Omega)$ on the Borel σ-algebra of Ω.*

(i) Assume that there is $\rho_0 > 0$ such that for every $x \in Q$ one has $p(x, Q) \geq \rho_0$ and $p(x, \partial Q) = 0$. Then there exists a quasi-stationary measure η with respect to Q.

(ii) If $p(x, M \setminus Q) > 0$ for every $x \in Q$, then there is no stationary measure η with support contained in Q.

(iii) Let $x \in Q$ and suppose that $p(x,\cdot)$ has a density with respect to a probability measure λ with $\lambda(\partial Q) = 0$. Then $p(x,\partial Q) = 0$ follows.

Proof. (i) Consider the map $\hat{S} : \mathcal{P}(Q) \to \mathcal{P}(Q)$ given by

$$(\hat{S}\eta)(A) := \frac{(\nu \times \eta)\left(S^{-1}(\mathcal{U} \times A)\right)}{(\nu \times \eta)\left(S^{-1}(\mathcal{U} \times Q)\right)}, A \in \mathcal{B}(Q).$$

In fact, $\hat{S}\eta$ is a probability measure on Q, since it is a nonnegative measure on Q and $(\hat{S}\eta)(Q) = 1$.

The assumption $p(x,Q) \geq \rho_0$ for all $x \in Q$ implies that

$$\begin{aligned}(\nu \times \hat{S}\eta)(S^{-1}(\mathcal{U} \times Q)) &= (\nu \times \hat{S}\eta)\{(u,x) \mid f(x,u_0)) \in Q\} \\ &= \int_Q \int_{\mathcal{U}} \mathbf{1}_{\{(u,x)|\, f(x,u_0))\in Q\}} \nu(du)(\hat{S}\eta)(dx) \\ &= \int_Q \int_{\Omega} \mathbf{1}_{\{(\omega,x)|\, f(x,\omega))\in Q\}} \nu_0(d\omega)(\hat{S}\eta)(dx) \\ &= \int_Q p(x,Q)(\hat{S}\eta)(dx) \geq \rho_0,\end{aligned}$$

since $\hat{S}\eta$ is a probability measure. It follows that the map $\hat{S}$ can be restricted to a map on the compact and convex set

$$\mathcal{P}(Q,\rho_0) := \{\eta \in \mathcal{P}(Q) \mid (\nu \times \eta)\, S^{-1}(\mathcal{U} \times Q) \geq \rho_0\}.$$

The map associating to $\eta \in \mathcal{P}(Q) \subset C(Q)^*$ the measure

$$(\nu \times \eta)\left(S^{-1}(\mathcal{U} \times A)\right), A \in \mathcal{B}(Q),$$

is weak* continuous, cf. Walters [18, Theorem 6.7]. Furthermore, [18, Remark 3(iv) on p. 149] shows that the map associating to $\eta \in \mathcal{P}(Q)$ the real number $(\nu \times \eta)\, S^{-1}(\mathcal{U} \times Q)$ is continuous, if for every $\eta \in \mathcal{P}(Q)$

$$(\nu \times \eta)\, S^{-1}(\partial(\mathcal{U} \times Q)) = 0. \tag{2.11}$$

Hence the map $\hat{S}$ on $\mathcal{P}(Q)$ is weak* continuous if (2.11) holds. By (2.4)

$$\begin{aligned}(\nu \times \eta)\, S^{-1}(\partial(\mathcal{U} \times Q)) &= (\nu \times \eta)\,\{(u,x) \mid f(x,u_0)) \in \partial Q\} \\ &= (\nu_0 \times \eta)\,\{(\omega,x) \mid f(x,\omega)) \in \partial Q\} \\ &= \int_Q p(x,\partial Q)\eta(dx) = 0,\end{aligned}$$

since by assumption $0 = p(x,\partial Q)$ for all $x \in Q$. Thus (2.11) holds and $\hat{S}$ is continuous on the compact convex subset $\mathcal{P}(Q,\rho_0)$ of a locally convex topological vector space. Then the Schauder-Tychonov fixed point theorem (cf. Dunford/Schwartz [8, p. 456]) shows that $\hat{S}$ has a fixed point $\eta \in \mathcal{P}(Q,\rho_0)$. This fixed point is quasi-stationary.

(ii) If η is a stationary measure with support contained in Q, then $\int_Q p(x,Q)\eta(dx) = 1$, hence $p(x,Q) = 1$ for η-almost all $x \in Q$. Clearly, $1 = p(x,M) = p(x,Q) + p(x,M \setminus Q)$ for every $x \in M$. If $p(x,M \setminus Q) > 0$ for all $x \in Q$ it follows that $p(x,Q) < 1$ for all $x \in Q$ and hence, in particular for all $x \in \text{supp}\eta \subset Q$. This contradiction proves (ii).

(iii) If $p(x,\cdot)$ has a density $k(x,\cdot) \in L^1(Q,\lambda)$, then $\lambda(\partial Q) = 0$ implies $p(x,\partial Q) = \int_{\partial Q} k(x,y)\lambda(dy) = 0$. □

REMARK 2.10. *There are alternative constructions for conditionally invariant measures in a variety of situations. Demers and Young [6] discuss this mainly for maps, without taking into account a skew product structure, and with regard to absolutely continuous conditionally invariant measures, cf. in particular [6, Section 5]. For the case of quasi-stationary measures, Collett, Martinez, and San Martin [2, Proposition 2.10 and Theorem 2.11] prove a general result based on the analysis of an associated operator P_1 and require that P_1 maps the space $C(Q), Q$ compact, of continuous real valued functions into itself (i.e., it is Feller). For a random map of the form (2.6), this amounts to the following: Define for bounded measurable functions $h : Q \to \mathbb{R}$*

$$[P_1(h)](x) := \int_{\{\omega\in\Omega | f(x,\omega)\in Q\}} h(f(x,\omega))\nu_0(d\omega).$$

If P_1 maps continuous functions h into continuous functions, a variant of a theorem due to Krein (see Oikhberg and Troitsky [16, Theorem 4]) yields the existence of a positive eigenvector for the dual operator which determines a quasi-stationary measure η. (The theorem can be applied, since $C(Q)$ is an ordered Banach space which has an element h_0 with $\|h_0\| = 1$ such that $h_0 \geq h$ for all h with $\|h\| = 1$.)

More explicitly, the theorem implies that there is a positive measure η on Q such that $P_1^*\eta = \beta\eta$. Thus for all $h \in C(Q)$

$$(P_1^*\eta)(h) = \int_Q (P_1 h)d\eta = \int_Q \int_{\{\omega\in\Omega | f(x,\omega)\in Q\}} h(f(x,\omega))\nu_0(d\omega)\eta(dx) = \beta\int_Q h\eta(dx).$$

We may assume that η is a probability measure. Then for every Borel measurable set $A \subset Q$

$$\int_Q P(x,A)\eta(dx) = \int_Q \int_{\{\omega\in\Omega | f(x,\omega)\in A\}} \nu_0(d\omega)\eta(dx) = \beta\eta(A).$$

It follows that $\beta = \int_Q P(x,Q)\eta(dx)$ and hence η is quasi-stationary.

The map P_1 leaves the space of continuous functions invariant in the following situation: for $h \in C(Q)$, continuity of f guarantees that the map $(x,\omega) \mapsto h(f(x,\omega))$ is continuous. If $x_n \to x$ in Q, compactness of Ω implies that the Hausdorff distance $d_H(A_n, A) \to 0$ for the compact sets $A_n := \{\omega \in \Omega \,|\, f(x_n,\omega) \in Q\}$ and $A := \{\omega \in \Omega \,|\, f(x,\omega) \in Q\}$. If, for example, $\Omega \subset \mathbb{R}^m$ and ν_0 has a density $g \in L^1(\Omega)$ with respect to the Lebesgue measure, it follows that $\nu_0(A_n \triangle A) = \int_{A_n\triangle A} g(\omega)d\omega \to 0$. Thus $[P_1(h)](x_n) \to [P_1(h)](x)$ and the desired continuity of $P_1 h$ follows.

3. Metric invariance entropy. In this section, we construct two versions of metric invariance entropy in analogy to metric entropy for dynamical systems: a feedback version and a version based on controllability properties.

We consider control system (1.1) on M and suppose that Q is a closed subset of M. Throughout this section, we fix a conditionally invariant measure μ with constant $\rho = \rho(\mu) \in (0,1]$ for Q and the skew product map S on $\mathcal{U} \times M$; cf. Definition 2.5. Recall that we write $S_Q := S_{|\mathcal{U}\times Q} : \mathcal{U}\times Q \to \mathcal{U}\times M$ for the restriction. Thus for $n \in \mathbb{N}$ the measure μ is conditionally invariant for S_Q^n with constant ρ^n. Since μ lives in $\mathcal{U}\times Q$ we construct certain partitions for $\mathcal{U}\times Q$ whose entropy with respect to μ will be used

to define the metric invariance entropy. While this seems fairly straightforward for the component in Q, more work will be needed for the $\mathcal{U}$-component.

For motivational purposes consider a sequence of measurable partitions satisfying, for some fixed $\tau \in \mathbb{N}$, for all $n, m \in \mathbb{N}$

$$\mathfrak{A}_{n+m} = \mathfrak{A}_n \vee S^{-n\tau}\mathfrak{A}_m. \tag{3.1}$$

It follows that $\mathfrak{A}_n = \bigvee_{i=0}^{n-1} S^{-i\tau}\mathfrak{A}_1$ for all $n \geq 1$. Recall that for a probability measure μ on a space X and measurable partitions $\mathfrak{A}$, $\mathfrak{C}$ and $\mathfrak{D}$ the conditional entropy is, with $\phi(x) = x \log x, x \in [0,1]$, defined by

$$H_\mu(\mathfrak{A}\,|\mathfrak{C}\,) = -\sum_{C\in\mathfrak{C}} \mu(C) \sum_{A\in\mathfrak{A}} \phi\left(\frac{\mu(A\cap C)}{\mu(C)}\right) = -\sum_{A\in\mathfrak{A}, C\in\mathfrak{C}} \mu(A\cap C)\log\frac{\mu(A\cap C)}{\mu(C)},$$

and $H_\mu(\mathfrak{A}\vee\mathfrak{C}\,|\mathfrak{D}\,) = H_\mu(\mathfrak{A}\,|\mathfrak{D}\,) + H_\mu(\mathfrak{C}\,|\mathfrak{A}\vee\mathfrak{D}\,)$ (cf. Walters [18, Theorem 4.3(i)]). Suppose that μ is an invariant measure for a map S. Then, applying repeatedly this formula, one finds for $n \in \mathbb{N}$

$$H_\mu(n,\mathfrak{A}_1) := H_\mu(\mathfrak{A}_n) = \sum_{i=0}^{n-1} H_\mu\left(S^{-i\tau}\mathfrak{A}_1 \,\middle|\, \bigvee_{\ell=0}^{i-1} S^{-\ell\tau}\mathfrak{A}_1\right) = \sum_{i=0}^{n-1} H_\mu\left(\mathfrak{A}_{i+1}\,|\mathfrak{A}_i\,\right); \tag{3.2}$$

here, and in the following, conditioning for $\bigvee_{\ell=0}^{-1}$ is omitted and $\mathfrak{A}_0 := \{X\}$. The last equation in (3.2) follows, since for all $i \geq 0$

$$H_\mu\left(\mathfrak{A}_{i+1}\,|\mathfrak{A}_i\,\right) = H_\mu\left(S^{-i\tau}\mathfrak{A}_1 \vee \mathfrak{A}_i\,|\mathfrak{A}_i\right) = H_\mu\left(S^{-i\tau}\mathfrak{A}_1\,|\mathfrak{A}_i\right).$$

This will guide our definition of invariance entropy for a conditionally invariant measure μ.

The following definition is taken from Kawan [11, Definition 2.8].

DEFINITION 3.1. *An invariant partition is a triple $\mathcal{C} = (\mathcal{P}, \tau, F)$ where $\mathcal{P}$ is a finite partition of Q into Borel measurable sets, $\tau \in \mathbb{N}$, and $F : \mathcal{P} \to \Omega^\tau$ is a map assigning to each set P in $\mathcal{P}$ a control function such that $\varphi(k, P, F(P)) \subset Q$ for all $k \in \{1, ..., \tau\}$.*

Note that F should be considered as a (piecewise constant) feedback.

REMARK 3.2. *Analogously, if $\mathcal{P}$ is an open cover of Q, the triple $(\mathcal{P}, \tau, F)$ is called an invariant open cover of Q. In the definition of topological feedback entropy by Nair, Evans, Mareels and Moran [15] invariant open covers replace the open covers used for topological entropy of dynamical systems.*

For an invariant partition $\mathcal{C} = (\mathcal{P}, \tau, F)$ with $\mathcal{P} = \{P_1, ..., P_q\}$ we abbreviate $F_i := F(P_i) \in \Omega^\tau, i = 1, ..., q$, and define for every word $a := [a_0, a_1, ..., a_{N-1}], N \in \mathbb{N}$, with $a_j \in \{1, ..., q\}$ a control function u_a on $\{0, ..., N\tau - 1\}$ by applying these feedback maps one after the other:

$$(u_a)_{i\tau+k} := (F_{a_i})_k \text{ for } i = 0, ..., N-1 \text{ and } k = 0, ..., \tau - 1. \tag{3.3}$$

We also write $u_a = (F_{a_0}, F_{a_1}, ..., F_{a_{N-1}})$. The word a is called admissible for $\mathcal{C}$ if there exists a point $x \in Q$ with

$$\varphi(i\tau, x, u_a) \in P_{a_i} \text{ for } i = 0, 1, ..., N-1.$$

The admissible words for $\mathcal{C}$ describe the possible sequences of partition elements under the feedbacks associated with $\mathcal{C}$. For $P \in \mathcal{P}$ we define

$$A(P) := \{(u, x) \in \mathcal{U} \times Q \,|\, x \in P \text{ and } \varphi(k, P, u) \subset Q \text{ for } k = 1, ..., \tau\} \tag{3.4}$$

and

$$\mathfrak{A}(\mathcal{C}) := \{A(P)\,|P \in \mathcal{P}\} \text{ with union } \mathcal{A}(\mathcal{C}) = \bigcup_{P\in\mathcal{P}} A(P).$$

Here and in the following, we will omit the reference to $\mathcal{C}$ if the considered invariant partition is clear from the context. This union is disjoint, and, in general, $\mathcal{A}$ will be a proper subset of $\mathcal{U} \times Q$. For $(u, x) \in A(P)$, the controls u (on the relevant interval $\{0, ..., \tau-1\}$) should be considered as feedbacks since they are applied to all elements of P. Observe that $S_Q^{-n\tau}\mathfrak{A}$ is a measurable partition of the set $S_Q^{-n\tau}\mathcal{A} \subset S_Q^{-n\tau}(\mathcal{U}\times Q)$. A sequence $(A_0, ..., A_{n-1})$ is called admissible if $A_i = A(P_{a_i}) \in \mathfrak{A}$ and $a = [a_0, ..., a_{n-1}]$ is an admissible word, and we define

$$D_a = A_0 \cap ... \cap S_Q^{-(n-1)\tau}A_{n-1} \in \bigvee_{i=0}^{n-1} S_Q^{-i\tau}\mathfrak{A}. \tag{3.5}$$

We also call intersections of the form D_a admissible. Let

$$\mathfrak{A}_n := \left\{ D_a \in \bigvee_{i=0}^{n-1} S_Q^{-i\tau}\mathfrak{A}\,|a \text{ admissible} \right\}, \mathcal{A}_n := \bigcup\nolimits_{D\in\mathfrak{A}_n(\mathcal{C})} D. \tag{3.6}$$

Observe that $\mathfrak{A}_1 = \mathfrak{A}$ and, for convenience, we set $\mathfrak{A}_0 = \mathcal{U} \times Q$. Furthermore, for all sets $D \in \mathfrak{A}_n$ let

$$\mathfrak{A}_n^+(D) := \{A \in \mathfrak{A} \,\Big|\, D \cap S_Q^{-n\tau}A \text{ is admissible}\},\ \mathcal{A}_n^+(D) := \bigcup\nolimits_{A\in\mathfrak{A}_n^+(D)} A,$$
$$\mathfrak{A}_{n+1}(D) := D \cap S_Q^{-n\tau}\mathfrak{A}_n^+(D),\ \mathcal{A}_{n+1}(D) := D \cap S_Q^{-n\tau}\mathcal{A}_n^+(D).$$

Thus for $A \in \mathfrak{A}_n^+(D_a)$ the intersection

$$A_0 \cap ... \cap S_Q^{-(n-1)\tau}A_{n-1} \cap S_Q^{-n\tau}A$$

corresponds to an admissible word of length $n + 1$, hence is admissible. It is immediately clear that $\mathfrak{A}_{n+1} \subset \mathfrak{A}_n \vee S_Q^{-n\tau}\mathfrak{A}_1$ and that, in general, this inclusion is proper. Furthermore, $\mathcal{A}_{n+1}$ is the disjoint union of the sets $\mathcal{A}_{n+1}(D), D \in \mathfrak{A}_n$.

The following lemma shows that the sequence $\mathfrak{A}_n, n \in \mathbb{N}$, satisfies a condition similar to (3.1).

LEMMA 3.3. *(i) For all $n, m \in \mathbb{N}$*

$$\mathfrak{A}_{n+m} \subset \mathfrak{A}_n \vee S_Q^{-n\tau}\mathfrak{A}_m. \tag{3.7}$$

More precisely, the collection $\mathfrak{A}_{n+m}$ consists of all intersections $D_{a^{[n]}} \cap S_Q^{-n\tau}D_{a^{[m]}}$, where $D_{a^{[m]}} \in \mathfrak{A}_m$ corresponds to an admissible word $a^{[m]}$ of length m such that the word $a = a^{[n]}a^{[m]}$ is admissible.

(ii) Assume that the invariant partition $\mathcal{C} = (\mathcal{P}, \tau, F)$ satisfies the following property: for every $x \in Q$ there are $P \in \mathcal{P}$ and $y \in P$ such that $\varphi(\tau, y, F(P)) = x$. Then for every admissible word $a^{[m]}$ of length m there exist an admissible word $a^{[n]}$ such that the word $a = a^{[n]}a^{[m]}$ is admissible.

Proof. (i) Consider for an admissible word a

$$A_0 \cap ... \cap S_Q^{-(n+m-1)\tau}A_{n+m-1} \in \mathfrak{A}_{n+m}$$

with $A_i = A(P_{a_i})$. Then, clearly, $A_0 \cap ... \cap S_Q^{-(n-1)\tau} A_{n-1}$ is an element of $\mathfrak{A}_n$, since $[a_0, ..., a_{n-1}]$ is an admissible word. Also $[a_n, ..., a_{n+m-1}]$ is admissible, since

$$\varphi(j\tau, \varphi(n\tau, x, u_a), \theta^{n\tau} u_a) = \varphi((n+j)\tau, x, u_a) \in P_{a_{n+j}} \text{ for } j = 0, 1, ..., m-1.$$

It follows that

$$\begin{aligned} &S_Q^{-n\tau} A_n \cap ... \cap S_Q^{-(n+m-1)\tau} A_{n+m-1} \\ &= S_Q^{-n\tau} \left(A_n \cap ... \cap S_Q^{-(m-1)\tau} A_{n+m-1} \right) \in S_Q^{-n\tau} \mathfrak{A}_m. \end{aligned}$$

(ii) Let $a^{[m]}$ be an admissible word of length m, i.e., there is $x \in Q$ such that

$$\varphi(j\tau, x, u_{a^{[m]}}) \in P_{a_j^{[m]}} \text{ for } j = 0, 1, ..., m-1.$$

Then there are $P = P_{a_{n-1}}$, a point $x_{-1} \in P_{a_{n-1}} \subset Q$ and $u_{-1} = F(P_{a_{n-1}}) \in \Omega^\tau$ such that

$$\varphi(\tau, x_{-1}, F(P_{a_{n-1}})) = x.$$

Proceeding inductively, one finds $x_0 \in Q$ and an admissible word $a^{[n]}$ such that

$$\varphi(n\tau, x_0, u_{a^{[n]}}) = x \text{ and } \varphi(i\tau, x_0, u_{a^{[n]}}) \in P_{a_i^{[n]}} \text{ for } i = 0, ..., n-1.$$

Hence the word $a = a^{[n]} a^{[m]}$ is admissible. □

The converse inclusion in (3.7) is not valid, since a nonvoid intersection $D \cap S_Q^{-n\tau} D'$ with $D \in \mathfrak{A}_n$ and $D' \in \mathfrak{A}_m$ may not correspond to an admissible word of length $n+m$. Furthermore, the collection $\mathfrak{A}_n$ is not a partition of $\mathcal{U} \times Q$ and μ is not invariant for S_Q.

We will use the following version of the notion of conditional entropy.

DEFINITION 3.4. *Let $\mathfrak{D}$ and $\mathfrak{E}(D), D \in \mathfrak{D}$, be finite families of measurable, pairwise disjoint sets in a space X with probability measure $\mathbf{m}$. Let $\mathcal{E}(D) := \bigcup_{E \in \mathfrak{E}(D)} E$ for $D \in \mathfrak{D}$, and suppose that for every $D \in \mathfrak{D}$ the collection $\{D \cap E \,|\, E \in \mathfrak{E}(D)\}$ is a partition of $D \cap \mathcal{E}(D)$. Then the conditional entropy of $\mathfrak{E}(\cdot)$ given $\mathfrak{D}$ is*

$$H_{\mathbf{m}}(\mathfrak{E}(\cdot) \,|\, \mathfrak{D}) = -\sum_{D \in \mathfrak{D}} \mathbf{m}(D \cap \mathcal{E}(D)) \sum_{E \in \mathfrak{E}(D)} \phi\left(\frac{\mathbf{m}(D \cap E)}{\mathbf{m}(D \cap \mathcal{E}(D))} \right).$$

This differs from the usual definition of conditional entropy (it might be called a conditional pseudo-entropy), since it refers to measurable families, not necessarily partitions, and in the considered collections $\{D \cap E \,|\, E \in \mathfrak{E}(D)\}$ the collections of allowed sets E may depend on D. For every $D \in \mathfrak{D}$, the conditional measure $\frac{\mathbf{m}(D \cap \cdot)}{\mathbf{m}(D \cap \mathcal{E}(D))}$ is a probability measure on $D \cap \mathcal{E}(D)$, and we sum up the corresponding entropies of the partitions induced by $\mathfrak{E}(D)$ on $D \cap \mathcal{E}(D)$ with weights $\mathbf{m}(D \cap \mathcal{E}(D))$. Thus the entropy induced on the complement of the union of all elements of $\mathfrak{D}$ is disregarded; and the weights $\mathbf{m}(D \cap \mathcal{E}(D))$ only take into account the probability of the intersection of D with the union $\mathcal{E}(D)$ of the elements of $\mathfrak{E}(D)$. Naturally, this conditional entropy can also be written as

$$H_{\mathbf{m}}(\mathfrak{E}(\cdot) \,|\, \mathfrak{D}) = -\sum_{D \in \mathfrak{D}} \sum_{E \in \mathfrak{E}(D)} \mathbf{m}(D \cap E) \log \frac{\mathbf{m}(D \cap E)}{\mathbf{m}(D \cap \mathcal{E}(D))}. \tag{3.8}$$

For the definition of invariance entropy we will consider the conditional entropy $H_{\rho^{-i\tau}\mu}\left(S_Q^{-i\tau}\mathfrak{A}_i^+(\cdot)\,|\mathfrak{A}_i\right)$ of the family $\left\{S_Q^{-i\tau}\mathfrak{A}_i^+(D), D\in\mathfrak{A}_i\right\}$ given the collection $\mathfrak{A}_i$ defined in (3.6). Note that for every $D\in\mathfrak{A}_i$ the collection

$$\{D\cap S_Q^{-i\tau}A\,\big|A\in\mathfrak{A}_i^+(D)\}$$

is a partition of $D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D)$, and every $D\in\mathfrak{A}_i$ can be extended to an admissible intersection of length $i+1$ by some element $S_Q^{-i\tau}A$. Furthermore, by Remark 2.3, $\rho^{-i\tau}\mu$ is a probability measure on $S_Q^{-i\tau}(\mathcal{U}\times Q)$, and

$$\frac{\rho^{-i\tau}\mu(D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D)\cap\cdot)}{\rho^{-i\tau}\mu(D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D))}$$

defines a probability measure on $D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D)$. Thus we sum up the entropies induced on $D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D), D\in\mathfrak{A}_i$, with weights given by $\rho^{-i\tau}\mu(\cdot)$,

$$\begin{aligned}
&H_{\rho^{-i\tau}\mu}\left(S_Q^{-i\tau}\mathfrak{A}_i^+(\cdot)\,|\mathfrak{A}_i\right)\\
&=-\sum_{D\in\mathfrak{A}_i}\sum_{A\in\mathfrak{A}_i^+(D)}\rho^{-i\tau}\mu(D\cap S_Q^{-i\tau}A)\log\frac{\mu(D\cap S_Q^{-i\tau}A)}{\mu(D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D))}\\
&=-\sum_{D\in\mathfrak{A}_i}\rho^{-i\tau}\mu(D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D))\sum_{A\in\mathfrak{A}_i^+(D)}\phi\left(\frac{\rho^{-i\tau}\mu(D\cap S_Q^{-i\tau}A)}{\rho^{-i\tau}\mu(D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D))}\right).
\end{aligned}$$

Next we define the metric invariance entropy of an invariant partition in analogy to condition (3.2).

DEFINITION 3.5. *For an invariant partition $\mathcal{C}=(\mathcal{P},\tau,F)$ of Q the invariance μ-entropy of $\mathfrak{A}(\mathcal{C})$ up to time $n\tau$ is defined as the following sum of conditional entropies,*

$$H_\mu(n,\mathfrak{A}(\mathcal{C}))=\sum_{i=0}^{n-1}H_{\rho^{-i\tau}\mu}\left(S_Q^{-i\tau}\mathfrak{A}_i^+(\cdot)\,|\mathfrak{A}_i\right),\tag{3.9}$$

and the feedback invariance μ-entropy of $\mathcal{C}$ is

$$h_\mu^{fb}(\mathcal{C})=\liminf_{n\to\infty}\frac{1}{n\tau}H_\mu(n,\mathfrak{A}(\mathcal{C})).\tag{3.10}$$

Formula (3.9) measures the average increase of information. Observe that only the elements of the families $S_Q^{-i\tau}\mathfrak{A}_i^+(D), D\in\mathfrak{A}_i$, determine the additional information in every time step.

The following definition introduces the central concept of this paper.

DEFINITION 3.6. *For the skew product map S from (2.4) associated with control system (1.1) the feedback invariance μ-entropy of Q with respect to a conditionally invariant measure μ is*

$$h_\mu^{fb}(Q,S):=\inf_{\mathcal{C}}h_\mu^{fb}(\mathcal{C}),$$

where the infimum is taken over all invariant partitions $\mathcal{C}=(\mathcal{P},\tau,F)$ of Q. If no invariant partition exists, we set $h_\mu^{fb}(Q,S):=\infty$.

REMARK 3.7. *For an invariant partition $\mathcal{C}$, the entropy $H_{\rho^{-i\tau}\mu}\left(S_Q^{-i\tau}\mathfrak{A}_i^+(\cdot)\,|\mathfrak{A}_i\right)$ can also be written in a somewhat more concise form. Define*

$$H_{\rho^{-i\tau}\mu}(\mathfrak{A}_{i+1}\,|\mathfrak{A}_i\,) := -\rho^{-i\tau}\sum_{D\in\mathfrak{A}_i}\sum_{E\in\mathfrak{A}_{i+1}}\mu(D\cap E)\log\frac{\mu(D\cap E)}{\mu(\mathcal{A}_{i+1}(D))}.$$

Then the entropy satisfies

$$H_{\rho^{-i\tau}\mu}(\mathfrak{A}_{i+1}\,|\mathfrak{A}_i\,) = H_{\rho^{-i\tau}\mu}\left(S_Q^{-i\tau}\mathfrak{A}_i^+(\cdot)\,|\mathfrak{A}_i\right). \tag{3.11}$$

In fact, for $D\cap S_Q^{-i\tau}A$ with $D\in\mathfrak{A}_i$ and $A\in\mathfrak{A}_i^+(D)$ one finds with

$$E := D\cap S_Q^{-i\tau}A = A_0\cap\ldots\cap S_Q^{-(i-1)\tau}A_{i-1}\cap S_Q^{-i\tau}A\in\mathfrak{A}_{i+1}$$

that $D\cap E = D\cap S_Q^{-i\tau}A$ and $\mathcal{A}_{i+1}(D) = D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D)$. Conversely, let $E\in\mathfrak{A}_{i+1}$ and $D\in\mathfrak{A}_i$ be given by

$$E = A_0'\cap\ldots\cap S_Q^{-(i-1)\tau}A_{i-1}'\cap S_Q^{-i\tau}A_i',\ D = A_0\cap\ldots\cap S_Q^{-(i-1)\tau}A_{i-1}.$$

Since $\mathcal{A}_{i+1}$ is the disjoint union of the sets $\mathcal{A}_{i+1}(D), D\in\mathfrak{A}_i$, it follows that there is a unique $D\in\mathfrak{A}_i$ with $E = D\cap S_Q^{-i\tau}A, A\in\mathfrak{A}_i^+(D)$. The sets A, A' in $\mathfrak{A}_i^+(D)$ are pairwise disjoint, hence $D\cap S_Q^{-i\tau}A = D\cap S_Q^{-i\tau}A'\neq\emptyset$ implies $A = A'$. Thus there is a unique $A\in\mathfrak{A}_i^+(D)$ with $D\cap E = D\cap S_Q^{-i\tau}A, A\in\mathfrak{A}_i^+(D)$, and $\mathcal{A}_{i+1}(D) = D\cap S_Q^{-i\tau}\mathcal{A}_i^+(D)$. Then (3.11) follows.

REMARK 3.8. *If μ is an invariant measure for S, the formula for $H_\mu(n,\mathfrak{A}(\mathcal{C}))$ reduces to (3.2). One may wonder why a general time step $\tau\in\mathbb{N}$ is used above, while $\tau = 1$ might appear as the natural choice. We look at itineraries of trajectories in the elements of the partition $\mathcal{P}$ at integer multiples $i\tau, i\in\mathbb{N}$. Then on some fixed time interval from 0 to T the choice $\tau > 1$ yields less itineraries than $\tau = 1$.*

In Definition 3.6 the metric invariance entropy $h_\mu^{fb}(Q,S)$ is defined using the infimum over all invariant partitions $\mathcal{C} = (\mathcal{P},\tau,F)$, hence also over all $\tau\in\mathbb{N}$. All constructions are equally valid if we would take the fixed time step $\tau = 1$ throughout. The same is true, if we require that the limit for $\tau\to\infty$ is considered, i.e., we could take the infimum over all invariant partitions $\mathcal{C} = (\mathcal{P},\tau,F)$ for a fixed $\tau\in\mathbb{N}$ and then take the limit inferior for $\tau\to\infty$. The definitions above are given for general $\tau\in\mathbb{N}$ in analogy to feedback entropy as defined in Nair, Evans, Mareels and Moran [15]. Also Kawan's proof of Theorem 3.14 cited below uses in an essential way general time steps τ. His proof also shows that for the topological invariance entropy it suffices to take arbitrarily large time steps τ.

In the following, we argue why a straightforward generalization of the definition of entropy to conditionally invariant measures leads to a trivial notion. Recall the definition of $\mathfrak{A}_n$ from (3.6) and define for all $n\in\mathbb{N}$ partitions of $\mathcal{U}\times Q$ in the following way:

$$\tilde{\mathfrak{A}}_n := \mathfrak{A}_n\cup\{Z_n\} \text{ where } Z_n := (\mathcal{U}\times Q)\setminus\bigcup\nolimits_{D\in\mathfrak{A}_n}D.$$

We will show that $\lim_{n\to\infty}\frac{1}{n}H_\mu(\tilde{\mathfrak{A}}_n) = 0$ if $\rho < 1$. In view of $\mathfrak{A}_{n+m}\subset\mathfrak{A}_n\vee S_Q^{-n\tau}\mathfrak{A}_m$

consider

$$-H_\mu(\tilde{\mathfrak{A}}_{n+m}) - \phi(\mu(Z_{n+m})) \geq \sum_{A_0\in\mathfrak{A}_n, A_1\in\mathfrak{A}_m} \phi\left(\mu(A_0\cap S_Q^{-n\tau}A_1)\right)$$
$$= \sum_{A_0\in\mathfrak{A}_n, A_1\in\mathfrak{A}_m} \mu(A_0\cap S_Q^{-n\tau}A_1)\log\mu(A_0)$$
$$+ \sum_{A_0\in\mathfrak{A}_n, A_1\in\mathfrak{A}_m} \mu(A_0\cap S_Q^{-n\tau}A_1)\log\frac{\mu(A_0\cap S_Q^{-n\tau}A_1)}{\mu(A_0)}.$$

Observe that $\mathfrak{A}_n$ is complemented to the partition $\tilde{\mathfrak{A}}_n$ of $\mathcal{U}\times Q$. The corresponding summands will be negative, hence the first sum is bounded below by

$$\sum_{A_0\in\tilde{\mathfrak{A}}_n} \mu(A_0\cap S_Q^{-n\tau}(\mathcal{U}\times Q))\log\mu(A_0) \geq -H_\mu(\tilde{\mathfrak{A}}_n).$$

For the second summand denote the elements of $\mathfrak{D} = S_Q^{-n\tau}\tilde{\mathfrak{A}}_m$ by D. We obtain, using convexity of ϕ and conditional invariance, that it is bounded below by

$$\sum_{\substack{A_0\in\tilde{\mathfrak{A}}_n,\\ A_1\in\tilde{\mathfrak{A}}_m}} \mu(A_0\cap S_Q^{-n\tau}A_1)\log\frac{\mu(A_0\cap S_Q^{-n\tau}A_1)}{\mu(A_0)} = \sum_{\substack{A_0\in\tilde{\mathfrak{A}}_n,\\ A_1\in\tilde{\mathfrak{A}}_m}} \mu(A_0)\phi\left(\frac{\mu(A_0\cap S_Q^{-n\tau}A_1)}{\mu(A_0)}\right)$$
$$\geq \sum_{A_1\in\tilde{\mathfrak{A}}_m} \phi\left(\sum_{A_0\in\tilde{\mathfrak{A}}_n} \mu(A_0\cap S_Q^{-n\tau}A_1)\right) = \sum_{A_1\in\tilde{\mathfrak{A}}_m} \phi(\mu(S_Q^{-n\tau}A_1))$$
$$= \rho^{n\tau}\sum_{A\in\tilde{\mathfrak{A}}_m} \mu(A)\log\mu(A) + \rho^{n\tau}\log\rho^{n\tau}.$$

Then, with $b_n := \rho^{n\tau}\log\rho^{n\tau} - \phi(\mu(Z_{n+m}))$, it follows that

$$H_\mu(\tilde{\mathfrak{A}}_{n+m}) \leq H_\mu(\tilde{\mathfrak{A}}_n) + \rho^{n\tau}H_\mu(\tilde{\mathfrak{A}}_m) - b_n.$$

Note that $\mu(Z_{n+m}) \to 1$ for $n\to\infty$, since the sets in $\mathfrak{A}_{n+m}$ are contained in $S_Q^{-(n+m)}(\mathcal{U}\times Q)$. Thus the term b_n tends to 0 for $n\to\infty$. A modification of the subadditivity lemma in Walters [18, Theorem 4.9] shows that this implies, as claimed, $\lim_{n\to\infty}\frac{1}{n}H_\mu(\tilde{\mathfrak{A}}_n) = 0$.

Next we introduce a variant of Definition 3.6 where we omit the feedback property required in $A(P)$. For an invariant partition $\mathcal{C} = (\mathcal{P}, \tau, F)$ and $P\in\mathcal{P}$ define

$$B(P) := \{(u,x)\in\mathcal{U}\times Q \,|\, x\in P \text{ and } \varphi(k,x,u)\in Q \text{ for } k=1,...,\tau\}, \tag{3.12}$$

and let

$$\mathfrak{B}(\mathcal{C}) := \{B(P)\,|\,P\in\mathcal{P}\} \text{ with union } \mathcal{B}(\mathcal{C}) = \bigcup_{P\in\mathcal{P}} B(P).$$

This union is disjoint, and, in general, $\mathcal{B}$ will be a proper subset of $\mathcal{U}\times Q$. Observe that $S_Q^{-n\tau}\mathfrak{B}$ is a measurable partition of the set $\bigcup_{P\in\mathcal{P}} S_Q^{-n\tau}B(P) \subset S_Q^{-n\tau}(\mathcal{U}\times Q)$ for

all $n \in \mathbb{N}$. A sequence $(B_0, ..., B_{n-1})$ is called admissible for $\mathcal{C}$ if $B_i = B(P_{a_i}) \in \mathfrak{B}$ and $a = [a_0, ..., a_{n-1}]$ is an admissible word, and we define

$$D_a = B_0 \cap ... \cap S_Q^{-(n-1)\tau} B_{n-1} \in \bigvee_{i=0}^{n-1} S_Q^{-i\tau} \mathfrak{B}, \tag{3.13}$$

and let

$$\mathfrak{B}_n := \left\{ D_a \in \bigvee_{i=0}^{n-1} S_Q^{-i\tau} \mathfrak{B} \,|a \text{ admissible} \right\}.$$

Observe that $\mathfrak{B}_1 = \mathfrak{B}$. Furthermore, for all sets $D \in \mathfrak{B}_n$ let

$$\mathfrak{B}_n^+(D) := \{B \in \mathfrak{B} \,\Big|\, D \cap S_Q^{-n\tau} B \text{ is admissible}\}.$$

As in Lemma 3.3 one shows that

$$\mathfrak{B}_{n+m} \subset \mathfrak{B}_n \vee S_Q^{-n\tau} \mathfrak{B}_m.$$

DEFINITION 3.9. *For an invariant partition $\mathcal{C} = (\mathcal{P}, \tau, F)$ of Q the invariance μ-entropy of $\mathfrak{B}(\mathcal{C})$ up to time $n\tau$ is defined as the following sum of conditional entropies,*

$$H_\mu(n, \mathfrak{B}(\mathcal{C})) = \sum_{i=0}^{n-1} H_{\rho^{-i\tau}\mu} \left(S_Q^{-i\tau} \mathfrak{B}_i^+(\cdot) \,|\mathfrak{B}_i \right), \tag{3.14}$$

and the controlled invariance μ-entropy of $\mathcal{C}$ is

$$h_\mu^{co}(\mathcal{C}) = \liminf_{n\to\infty} \frac{1}{n\tau} H_\mu(n, \mathfrak{B}(\mathcal{C})). \tag{3.15}$$

This leads to the following notion of metric invariance entropy.

DEFINITION 3.10. *For the skew product map S from (2.4) associated with control system (1.1) the controlled invariance μ-entropy of Q with respect to a conditionally invariant measure μ is*

$$h_\mu^{co}(Q, S) := \inf_{\mathcal{C}} h_\mu^{co}(\mathcal{C}),$$

where the infimum is taken over all invariant partitions $\mathcal{C}$ of Q. If no such invariant partition exists, we set $h_\mu^{co}(Q, S) := \infty$.

Next we consider the behavior of the metric invariance entropies under appropriate conjugacies. For notational simplicity, we suppose that the sets of control values coincide.

DEFINITION 3.11. *Consider two control systems of the form (1.1) on M_1 and M_2, respectively, given by*

$$x_{n+1} = f_1(x_n, u_n) \text{ and } y_{n+1} = f_2(y_n, u_n) \text{ with } (u_n) \in \mathcal{U}.$$

Let μ_1 and μ_2 be conditionally invariant measures with respect to closed subsets $Q_1 \subset M_1$ and $Q_2 \subset M_2$, respectively. A bimeasurable bijection $\pi : Q_1 \to Q_2$ is called a conjugacy of these systems, if

$$\pi(f_1(x, \omega)) = f_2(\pi x, \omega) \text{ for all } \omega \in \Omega \text{ and } x \in Q_1 \tag{3.16}$$

and $\mathrm{id}_{\mathcal{U}} \times \pi : \mathcal{U} \times Q_1 \to \mathcal{U} \times Q_2$ *maps* μ_1 *onto* μ_2*, i.e.,*

$$\mu_1 \left((\mathrm{id}_{\mathcal{U}} \times \pi)^{-1} (B) \right) = \mu_2(B) \text{ for all } B \in \mathcal{B}(\mathcal{U} \times Q_2). \tag{3.17}$$

In terms of the solutions, conjugacy condition (3.16) means

$$\pi\varphi_1(n, x_0, u) = \varphi_2(n, \pi x_0, u) \text{ for all } n \geq 0.$$

With the associated skew product maps $S_i(u, x) = (\theta u, f_i(x, u_0)), i = 1, 2$, one obtains the skew conjugacy condition for $(u, x) \in \mathcal{U} \times Q_1$

$$\begin{aligned} (\mathrm{id}_{\mathcal{U}} \times \pi) \circ S_1(u, x) &= (\mathrm{id}_{\mathcal{U}} \times \pi) (\theta u, f_1(x, u_0)) = (\theta u, f_2(\pi x, u_0)) \\ &= S_2 \circ (\mathrm{id}_{\mathcal{U}} \times \pi) (u, x). \end{aligned} \tag{3.18}$$

Since we may interchange the roles of S_1, μ_1 and S_2, μ_2, respectively, conjugacy is an equivalence relation. The metric invariance entropies turn out to be invariant under this conjugacy relation.

THEOREM 3.12. *Suppose that there is a conjugacy* π *of two control systems of the form (1.1) on* M_1 *and* M_2 *with associated skew product maps* S_1 *and* S_2 *and conditionally invariant measures* μ_1 *and* μ_2 *for closed sets* $Q_1 \subset M_1$ *and* $Q_2 \subset M_2$*, respectively. Then the feedback invariance* μ*-entropy and the controlled invariance* μ*-entropy satisfy*

$$h^{fb}_{\mu_1}(Q_1, S_1) = h^{fb}_{\mu_2}(Q_2, S_2) \text{ and } h^{co}_{\mu_1}(Q_1, S_1) = h^{co}_{\mu_2}(Q_2, S_2).$$

Proof. First observe that conjugacy properties (3.17) and (3.18) imply $Q_2 = \pi Q_1$ and

$$\rho_1 = \mu_1(S_1^{-1}(\mathcal{U} \times Q_1)) = \mu_2(S_2^{-1}(\mathcal{U} \times Q_2)) = \rho_2.$$

Consider an invariant partition $\mathcal{C}_1 = (\mathcal{P}_1, \tau, F_1)$ of Q_1. Then $\{\pi(P) \,|\, P \in \mathcal{P}_1\}$ is a measurable partition of $Q_2 = \pi Q_1$ and every measurable partition of Q_2 can be written in this form. For $P \in \mathcal{P}_1$ and $u \in \mathcal{U}$ with $\varphi_1(k, P, u) \subset Q_1$ for $k = 1, ..., \tau$ one finds

$$\varphi_2(k, \pi P, u) = \pi\varphi_1(k, P, u)) \subset Q_2 \text{ for all } k \in \{1, \ldots, \tau\}.$$

Thus $\mathrm{id}_{\mathcal{U}} \times \pi$ maps the invariant partition $\mathcal{C}_1 = (\mathcal{P}_1, \tau, F_1)$ to the invariant partition $\mathcal{C}_2 = (\mathcal{P}_2, \tau, F_1)$ and the collection $\mathfrak{A}(\mathcal{C}_1)$ of $\mathcal{U} \times Q_1$ is mapped to the corresponding collection $\mathfrak{A}(\mathcal{C}_2)$ of $\mathcal{U} \times Q_2$,

$$\mathfrak{A}(\mathcal{C}_2) = \{(\mathrm{id}_{\mathcal{U}} \times \pi) A \,|\, A \in \mathfrak{A}(\mathcal{C}_1)\}.$$

One finds for the entropy

$$H_{\mu_1}(n, \mathfrak{A}(\mathcal{C}_1)) = H_{\mu_2}(n, \mathfrak{A}(\mathcal{C}_2)), n \in \mathbb{N}, \text{ and } h^{fb}_{\mu_1}(\mathcal{C}_1) = h^{fb}_{\mu_2}(\mathcal{C}_2).$$

It follows that $h^{fb}_{\mu_1}(Q_1, S_1) = h^{fb}_{\mu_2}(Q_2, S_2)$.

These arguments also show that $h^{fb}_{\mu_1}(Q_1, S_1) = \infty$ if and only if $h^{fb}_{\mu_2}(Q_2, S_2) = \infty$. For the controlled invariance μ-entropy one argues analogously. □

Finally, we briefly discuss the relation of metric invariance entropy to other notions of invariance entropy. In Kawan [11, Definition 2.2 and Proposition 2.3(ii)] the following notion is considered.

DEFINITION 3.13. *Consider system (1.1) with skew product map S and let $Q \subset M$ be a compact controlled invariant set, i.e., for every $x \in Q$ there is $\omega_x \in \Omega$ with $f(x, \omega_x) \in Q$. For $\tau \in \mathbb{N}$ a set $\mathcal{R} \subset \mathcal{U}$ is called (τ, Q)-spanning if for all $x \in Q$ there is $u \in \mathcal{R}$ with $\varphi(n, x, u) \in Q$ for all $n = 1, ..., \tau$. By $r_{inv}(\tau, Q)$ we denote the minimal number of elements such a set can have (if no finite (τ, Q)-spanning set exists, $r_{inv}(\tau, Q) := \infty$). The invariance entropy is defined by*

$$h_{inv}(Q, S) := \lim_{\tau \to \infty} \frac{1}{\tau} \log r_{inv}(\tau, Q).$$

Write $\mathcal{W}_N(\mathcal{C})$ for the set of all admissible words of length N of an invariant partition $\mathcal{C}$ and let

$$h(\mathcal{C}) := \lim_{N \to \infty} \frac{\log \#\mathcal{W}_N(\mathcal{C})}{N\tau} = \inf_{N \in \mathbb{N}} \frac{\log \#\mathcal{W}_N(\mathcal{C})}{N\tau}.$$

The following characterization of invariance entropy is given in [11, Theorem 2.3].

THEOREM 3.14. *For a compact and controlled invariant set Q it holds that*

$$h_{inv}(Q, S) = \inf_{\mathcal{C}} h(\mathcal{C}),$$

where the infimum is taken over all invariant partitions $\mathcal{C}$ of Q.

The metric entropies $h_\mu^{fb}(Q, S)$ and $h_\mu^{co}(Q, S)$ in the present paper have been constructed in analogy to this result. Instead of counting the sequences corresponding to admissible words, the conditionally invariant measure has been used to associate a probability to them, and then a corresponding notion of entropy is considered. One might expect that the additional information provided by a conditionally invariant measure would reduce the entropy, i.e., that $h_\mu^{fb}(Q, S)$ and $h_\mu^{co}(Q, S)$ are bounded above by $h_{inv}(Q, S)$. This would require to estimate the terms in (3.9) by the number of admissible words. At present it is not clear to me what the relation is between metric invariance entropies and $h_{inv}(Q, S)$. The relation of $h_{inv}(Q, S)$ to feedback invariance entropy (cf. Nair, Evans, Mareels, Moran [15]) has been clarified in Colonius, Kawan, Nair [3].

For given invariant partition $\mathcal{C}$ the feedback invariance μ-entropy describes the information associated with itineraries in $\mathcal{P}$ of the corresponding trajectories under feedbacks (leaving the partition elements in Q). Controlled invariance μ-entropy describes the information associated with itineraries in $\mathcal{P}$ of trajectories for arbitrary controls. Concerning the relation between these two notions of metric invariance entropy one immediately sees that for an invariant partition $\mathcal{C}$ every set $A(P)$ defined in (3.4) is contained in the corresponding set $B(P)$ defined in (3.12). It is not clear if this induces a corresponding inequality for the conditional entropies up to time $n\tau$ (cf. (3.9) and (3.14)).

Finally, it would be important to compute the metric invariance entropy for special classes of systems and for specific classes of conditionally invariant measures (e.g., for those generated by quasi-stationary measures) and to determine relations to the escape rate $\rho(\mu)$ and to a corresponding Perron-Frobenius operator.

Acknowledgement. I am grateful to an anonymous reviewer whose comments and queries helped to improve the paper considerably.

REFERENCES

[1] R. Adler, A. Konheim, and M. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc., 114 (1965), pp. 61–85.

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

[3] F. Colonius, C. Kawan, and G. Nair, *A note on topological feedback entropy and invariance entropy*, Systems and Control Letters, 62 (2013), pp. 377–381.

[4] A. da Silva and C. Kawan, *Invariance entropy of hyperbolic control sets*, Discrete and Continuous Dynamical Systems A, 36 (2016), pp. 97–136.

[5] D. Delchamps, *Stabilizing a linear system with quantized state feedback*, IEEE Trans. Aut. Control, 35 (1990), pp. 916–924.

[6] M. F. Demers and L.-S. Young, *Escape rates and conditionally invariant measures*, Nonlinearity, 19 (2006), pp. 377–397.

[7] T. Downarowicz, *Entropy in Dynamical Systems*, Cambridge University Press, 2011.

[8] N. Dunford and J. T. Schwartz, *Linear Operators, Part I: General Theory*, Wiley-Interscience, 1977.

[9] P. Gänssler and W. Stute, *Wahrscheinlichkeitstheorie*, Springer-Verlag, 1977.

[10] C. Kawan, *Lower bounds for the strict invariance entropy*, Nonlinearity, 24 (2011), pp. 1909–1935.

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

[12] G. Keller and C. Liverani, *Rare events, escape rates, quasistationarity: Some exact formulae*, J. Stat. Phys., 135 (2009), pp. 519–534.

[13] Y. Kifer, *Ergodic Theory of Random Transformations*, Birkhäuser, 1986.

[14] P. Mehta, U. Vaidya, and A. Banaszuk, *Markov chains, entropy, and fundamental limitations in nonlinear stabilization*, IEEE Trans. Aut. Control, 53 (2008), pp. 784–791.

[15] G. Nair, R. J. Evans, I. Mareels, and W. Moran, *Topological feedback entropy and nonlinear stabilization*, IEEE Trans. Aut. Control, 49 (2004), pp. 1585–1597.

[16] T. Oikhberg and V. Troitsky, *A theorem of Krein revisited*, Rocky Mountain Journal of Mathematics, 35 (2005), pp. 195–210.

[17] G. Pianigiani and J. A. Yorke, *Expanding maps on sets which are almost invariant: Decay and chaos*, Trans. Amer. Math. Soc., 252 (1979), pp. 351–366.

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

[19] H. Zmarrou and A. J. Homburg, *Bifurcations of stationary measures of random diffeomorphisms*, Ergodic Theory and Dynamical Systems, 27 (2007), pp. 1651–1692.