Towards an information-theoretic model of the Allison mixture

Lachlan Gunn,[1] François Chapeau-Blondeau,[2] Andrew Allison,[3] and Derek Abbott[4]

[1] *School of Electrical and Electronic Engineering, The University of Adelaide, SA 5005, Australia*
e-mail address: lachlan.gunn@adelaide.edu.au

[2] *Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), University of Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France.*
e-mail address: chapeau@univ-angers.fr

[3] *School of Electrical and Electronic Engineering, The University of Adelaide, SA 5005, Australia*
e-mail address: andrew.allison@adelaide.edu.au

[4] *School of Electrical and Electronic Engineering, The University of Adelaide, SA 5005, Australia*
e-mail address: derek.abbott@adelaide.edu.au

I. INTRODUCTION

The Allison Mixture[1] is a process formed by random sampling of two parent processes, and which can have the unintuitive property of being autocorrelated, despite all its values being drawn from uncorrelated processes. However, this correlation vanishes if the parent processes are of equal mean, suggesting the use of autoinformation[2,3] as an alternative to correlation, providing a canonical measure of the strength of the memory of the Allison mixture. We apply this measure to the Allison mixture, producing analytic expressions for the k-step autoinformation of its sampling process.

II. THE ALLISON MIXTURE

The Allison mixture[1] is a process in which samples are drawn from one of two distributions, the choice determined by the state of a Markov chain, shown in Fig. (1). The marginal distribution of this process is a mixture of the two source distributions, the mixing constant determined by the stationary distribution of the Markov chain.

Definition II.1 (Allison mixture[1])**.** *An Allison mixture X_t of two processes U_t and V_t is given by*

$$X_t = S_t U_t + (1 - S_t) V_t \tag{1}$$

where the sampling process S_t is a Markov chain, shown in Fig. (1), having states $\{0, 1\}$ and transition probabilities α_0 and α_1 when in states 0 and 1 respectively.

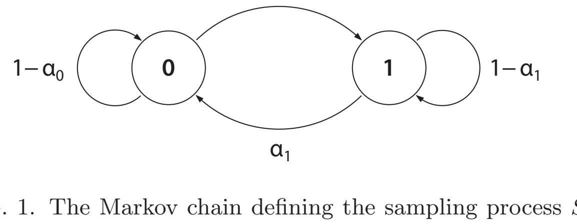

FIG. 1. The Markov chain defining the sampling process S_t of the Allison mixture. It is parametrised by the probabilities α_0 and α_1 of leaving states 0 and 1 respectively.

The stationary distribution of S_t is given by

$$\pi_0 = \frac{\alpha_1}{\alpha_0 + \alpha_1} \tag{2}$$

$$\pi_1 = \frac{\alpha_0}{\alpha_0 + \alpha_1}. \tag{3}$$

We use a spectral decomposition of the transition matrix P in order to compute the k-step probability matrix P^k and so the k-step transition probabilities $\alpha_{0,k}$ and $\alpha_{1,k}$.

Theorem II.1. *The sampling process S_t has k-step transition probabilities*

$$\alpha_{0,k} = \pi_0 \left[1 - (1 - \alpha_0 - \alpha_1)^k\right] \tag{4}$$

$$\alpha_{1,k} = \pi_1 \left[1 - (1 - \alpha_0 - \alpha_1)^k\right]. \tag{5}$$

III. AUTOINFORMATION OF THE ALLISON MIXTURE SAMPLING PROCESS

The autoinformation function is an alternative to the autocovariance function as a measure of dependence, defined as follows:

Definition III.1 (Autoinformation function[3])**.** *The autoinformation function of a stochastic process S_t is the mutual information*

$$I_{xx}[t, k] = I(S_t, S_{t-k}) \tag{6}$$

$$= H(S_t, S_{t-k}) - H(S_t) - H(S_{t-k}). \tag{7}$$

If S_t is stationary, then we may omit t as a parameter, leaving us with

$$I_{xx}[k] = I(S_t, S_{t-k}) \tag{8}$$

$$= H(S_t, S_{t-k}) - 2H(S_t). \tag{9}$$

The autoinformation improves on the autocovariance function as a measure of dependence by providing a condition both sufficient and necessary—whereas a lack of correlation does not necessarily indicate independence, two variables will have zero mutual information only if they are statistically independent; this is vital when the processes U_t and V_t of the system being modelled have identical means but differing variances, such as particle velocities in statistical mechanics.

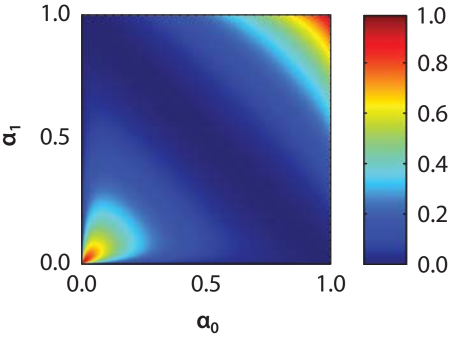

FIG. 2. Single-step autoinformation of the Allison mixture sampling process as a function of α_0 and α_1. Note the lines of zero autoinformation along $\alpha_0 = 0$, $\alpha_1 = 0$, and $\alpha_0 + \alpha_1 = 1$.

Lemma III.1. *Let S_t be a binary-valued random process with transition probabilities and a stationary distribution equal to that of the Markov chain in Definition II.1. Then, in the fully-mixed regime the single-step autoinformation is given by*

$$\begin{aligned} I_{xx}[1] = \; & \frac{\alpha_1(1-\alpha_0)\log_2 \frac{1-\alpha_0}{\alpha_1}}{\alpha_0+\alpha_1} \\ & + \frac{\alpha_0(1-\alpha_1)\log_2 \frac{1-\alpha_1}{\alpha_0}}{\alpha_0+\alpha_1} \\ & + \log_2(\alpha_0+\alpha_1), \end{aligned} \tag{10}$$

where both α_0 and α_1 are nonzero, zero if exactly one of α_0 and α_1 is equal to zero, and undefined if both are equal to zero.

Thus the autoinformation is equal to zero when $\alpha_0 = 0$, $\alpha_1 = 0$, or $\alpha_0+\alpha_1 = 1$, and so these previously-described[1] conditions for decorrelation of the sampling process imply zero mutual information and therefore genuine independence.

Importantly, we have not assumed the Markov property, instead directly demanding that the formulae for the stationary probabilities hold. This weakening is intended to allow us later to generalise to the Allison mixture proper.

The mutual information as a function of (α_0, α_1) is shown in Fig. (2). As one would expect, we see a peak near $(\alpha_0, \alpha_1) = (0, 0)$, where consecutive states are highly dependent. Similarly, we see a large autoinformation near $(1, 1)$, where the strong anticorrelation makes consecutive states highly predictable. Between these two extremes lies a valley, its nadir falling along the line $\alpha_0 + \alpha_1 = 1$; along this line, consecutive states of the sampling process are completely independent.

Theorem III.1. *The k-step autoinformation of a fully mixed two-state Markov chain with exit probabilities α_0 and α_1, as in Fig. (1), is given by Lemma III.1 under the substitution*

$$\alpha_0 \longrightarrow \pi_0 \left[1 - (1-\alpha_0-\alpha_1)^k\right] \tag{11}$$

$$\alpha_1 \longrightarrow \pi_1 \left[1 - (1-\alpha_0-\alpha_1)^k\right]. \tag{12}$$

This final theorem allows us to extend our single-step results to arbitrary time-lags, completing our characterisation of the Allison mixture sampling process.

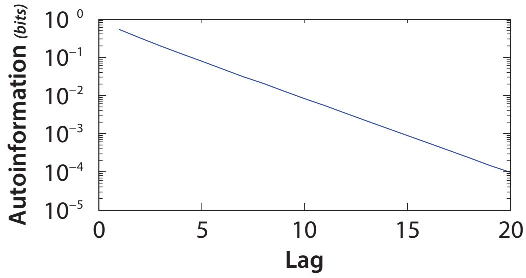

FIG. 3. Exponentially-decaying autoinformation of an Allison mixture sampling process with $\alpha_0 = 0.1, \alpha_1 = 0.1$.

We show the autoinformation in Fig. (3) as a function of lag; it can be seen to decay at a roughly exponential rate.

IV. OPEN QUESTIONS

The theorems that we have presented allow computation of the autoinformation function of the Allison mixture sampling process, and can be readily extended to binary-valued Allison mixtures. However, many physical systems are described by continuous-valued processes, and their autoinformation cannot be calculated by Lemma III.1. It remains to be seen whether the autoinformation can be computed by transformation of the sampling process autoinformation in a similar fashion to that of the autocovariance function[1].

Furthermore, the information-theoretic approach that we have presented provides the starting point for an investigation of the transfer entropy[4] between the sampling process and the Allison mixture; previous works on transfer entropy have focussed on complex systems, leaving room for the analysis of simpler and analytically tractable models in order to better probe its properties.

[1] L. J. Gunn, A. Allison, and D. Abbott, International Journal of Modern Physics **33**, 1460360 (2014).
[2] G. A. Darbellay, Physica A **287**, pp. 429–439 (2000).
[3] F. Chapeau-Blondeau, Physica A **380**, pp. 1–18 (2007).
[4] T. Schreiber, Physical Review Letters **85**, pp. 461–464 (2000).