ESAM 495 - Data-Driven Methods for Dynamical Systems
Project 3: Transfer entropy

Zihao Wang, Thomas Stiadle, April Zhi Zhou

1 Introduction

The rates at which time series from stochastic or deterministically chaotic systems generate information are quantified by the entropy. Given a time series that contains multiple components, it is important to learn each component's contributions to the information production and the rate at which they exchange information. One way to quantify the exchange of information among systems is to use mutual information, which is defined to be the excess amount of code produced by erroneously assuming that the two systems are independent. Consider two processes I and J, with joint probabilities $p_{IJ}(i,j)$, the mutual information is given by:

$$M_{IJ} = \sum p(i,j) \log \frac{p(i,j)}{p(i)p(j)} = H_I + H_J - H_{IJ}.$$

To further capture the dynamics and directionality of the information exchange, Schreiber proposes an alternatively measure called transfer entropy, which is defined to be

$$T_{J\to I} = \sum p(i_{n+1}, i_n^{(k)}, j_n^{(l)}) \log \left(\frac{p(i_{n+1}|i_n^{(n)}, j_n^{(n)})}{p(i_{n+1}|i_n^{(k)})} \right).$$

It can be interpreted as the how much information is lost in the probability in the distribution of state I when we don't consider the information flow from state J to state I. In this report, we construct one dimensional lattices of uni-directionally coupled maps and study the information transport in the lattices using transfer entropy.

2 Transfer entropy tests

We first construct 1D lattices of uni-directionally coupled maps defined by

$$x_{n+1}^m = f(\epsilon x_n^{m-1} + (1-\epsilon)x_n^m),$$

where ϵ is the coupling strength. The information in the lattices flows in the direction of increasing m. Tent map

$$f(x) = \begin{cases} 2x & x < 0.5 \\ 2 - 2x & x > 0.5 \end{cases}$$

and Ulam map $f(x) = 2 - x^2$ are used in the implementation. At zero coupling ($\epsilon = 0$), $T_{I^{m-1}\to I^m} = T_{I^m \to I^{m-1}} = 0$; for nonzero coupling, $T_{I^m\to I^{m-1}} = 0$ and $T_{I^{m-1}\to I^m} > 0$ due to the uni-directionality. We use a periodic lattice of 100 tent maps and compute the averages of transfer entropies from 10 runs of $20,000$ iterative steps after $20,000$ transient steps. In Figures 2(a) and 2(b), the averaged transfer entropy is plotted against the coupling strength ϵ (Figure 2(a) excludes the zero coupling case). As the coupling strength increases, we observe that the transfer entropy increases as well. We repeat the same process for the lattice of Ulam maps using the same parameters, and the numerical result is plotted in Figure 1 (including the zero coupling case), which is similar to the results in Figure 2 in Schreiber's paper.

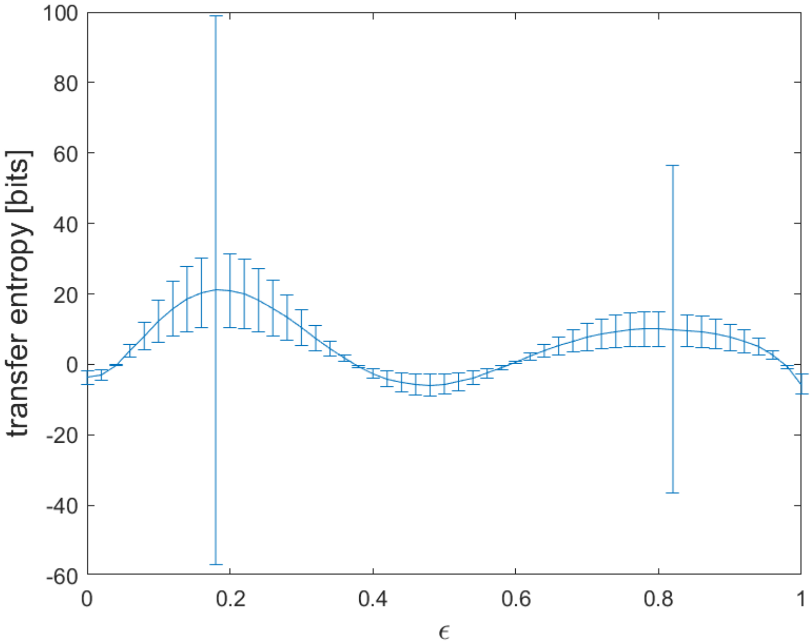
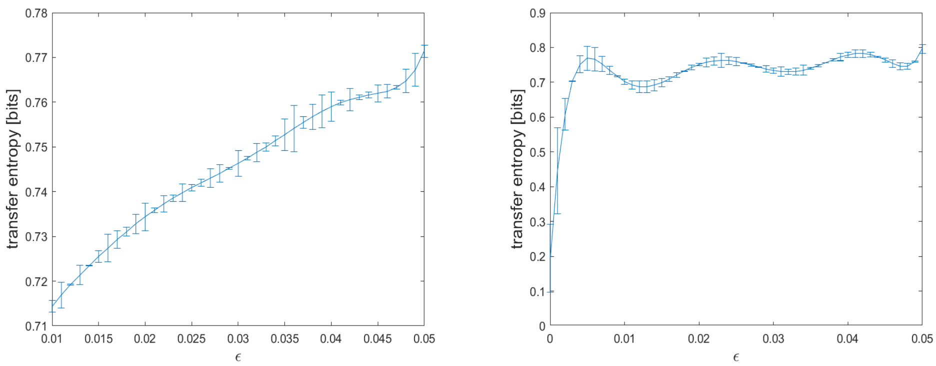


Figure 1: Transfer entropies $T_{X^{m-1}} \to T_{X^m}$ as functions of the coupling strength ϵ for a unidirectionally coupled Ulam lattice (resolution $r = 10^{-5}$).



(a) (b)

Figure 2: Transfer entropy $T_{X^{m-1}} \to T_{X^m}$ for the coupling direction as a function of the coupling strength ϵ in a tent map lattice (resolution $r = 10^{-5}$).

3 Transfer entropy test failures

3.1 Example 1: inadequate sample size

In this part, we alter the above model so that the transfer entropy test is expected to fail. First, we reduce the sample size in the 1D lattice of tent maps by reducing the total number of iterations from $N = 40,000$ to 2000 while holding everything else fixed. The numerical results for various values of ϵ (excluding the zero coupling case) are shown in Figure 3, from which, we fail to observe positive correlation between transfer entropy and coupling strength, contrary to Figure 2(a) where the sample size is large.

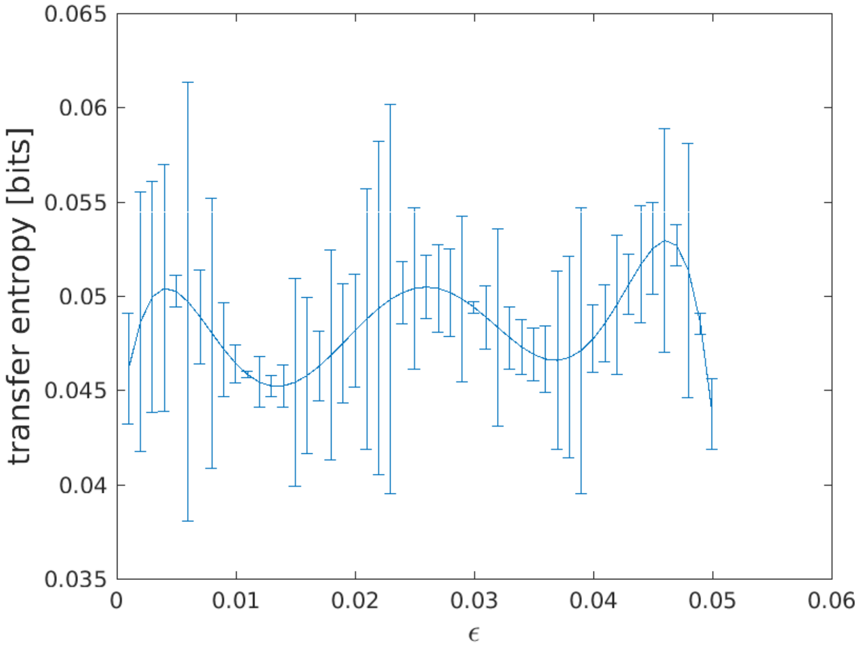


Figure 3: The fail test for tent map, with small iterations $N = 2000$.

3.2 Example 2: large resolution

Second, we increase the resolution r used for binning in the probability estimations. In the lattice of 100 uni-directional tent maps, we increase the resolution from $r = 10^{-5}$ to $r = 10^{-1}$ while holding everything else fixed. The numerical results for different values of ϵ (excluding the zero coupling case) are presented in Figure 4, from which we observe decrease in transfer entropy as the coupling strength increases, which is false.

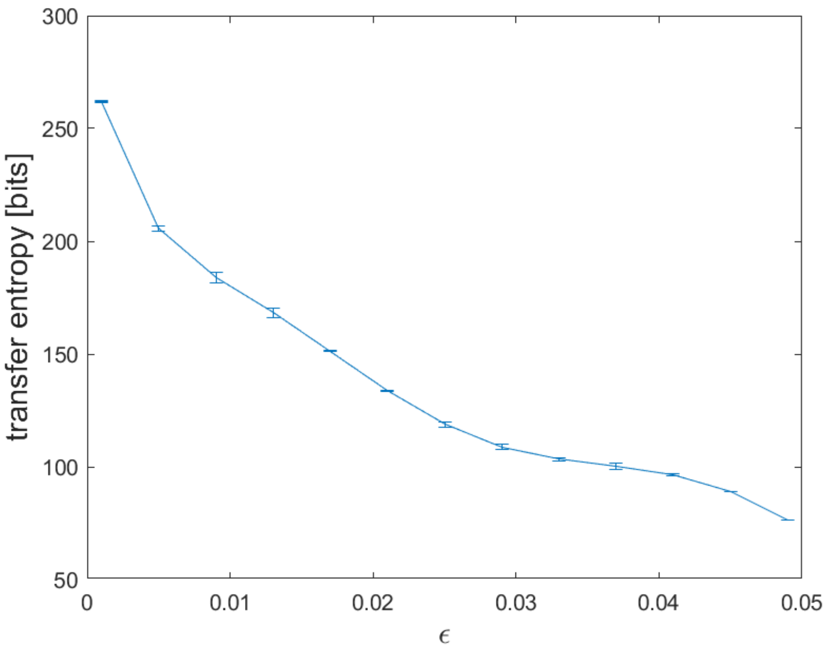


Figure 4: Transfer entropies $T_{X^{m-1}} \rightarrow T_{X^m}$ as functions of the coupling strength ϵ for a uni-directionally coupled Ulam lattice (resolution $r = 0.3$).

4 Conclusion

In conclusion, the transfer entropy tests work well with deterministic maps. In our implementations, the tests successfully detect directionality and coupling strength in uni-directionally coupled maps. Although, the tests are subject to failures if the sample size is too small or the binning resolution is too large. Note that since the tent maps and Ulam maps are continuous, the transfer entropy will diverge and the test will fail if the resolution is too small.

A Pseudocodes

procedure COMPUTE TRANSFER ENTROPY
  initialize number of points - M , number of iterations N, coupling strength ϵ, and resolution r.
  for $j = 1 : 10$ **do**
    construct a $M \times N$ lattice `X` with

$$x_{n+1}^m = f(\epsilon x_n^{m-1} + (1-\epsilon)x_n^m)$$

using the tent map or the Ulam map;
    remove the first $N/2$ iterations of the transient steps: `X = X(:,N/2:N)`
    compute transfer entropy between the first points x^1 and x^2

$$T_j = \sum_{n=1}^{N/2-1} \hat{p}(x_{n+1}^2, x_n^2, x_n^1) \log \left| \frac{\hat{p}(x_{n+1}^2 | x_n^2, x_n^1)}{\hat{p}(x_{n+1}^2 | x_n^2)} \right|$$

  end for
  compute the average of the transfer entropy $\{T_j\}_{j=1}^{10}$
end procedure

Algorithm 1: Transfer entropy implementation

procedure COMPUTE CONDITIONAL AND JOINT PROBABILITIES
  Receive vectors $y = x^1$, $x = x^2$, and index of point of interest, n
  Set resolution r
  Construct matrix $Z = [z_1, z_2, z_3]$, where $z_1 = x(2:N) - x(n+1)$, $z_2 = x(1:N-1) - x(n)$, and $z_3 = y(1:N-1) - y(n)$
  Find the maximum absolute value in each row of Z across columns 1 and 2. Count how many of these values are less than r (call this count a)
  Define $p(x_{n+1}, x_n) = \frac{a}{N-1}$
  Count how many of the absolute values of column 1 of Z are less than r (call this count k)
  Define $p(x_n) = \frac{k}{N-1}$
  Define $p(x_{n+1}|x_n) = \frac{p(x_{n+1}, x_n)}{p(x_n)}$
  Find the maximum absolute value in each row of Z. Count how many of these values are less than r (call this count q)
  Define $p(x_{n+1}, x_n, y_n) = \frac{q}{(N-1)}$
end procedure

Algorithm 2: Estimates of Required Probabilities