A SECOND LAW FOR OPEN MARKOV PROCESSES

BLAKE S. POLLARD

Department of Physics and Astronomy
University of California
Riverside, CA 92521

Abstract. In this paper we define the notion of an open Markov process. An open Markov process is a generalization of an ordinary Markov process in which populations are allowed to flow in and out of the system at certain boundary states. We show that the rate of change of relative entropy in an open Markov process is less than or equal to the flow of relative entropy through its boundary states. This can be viewed as a generalization of the Second Law for open Markov processes.

1. Introduction

Markov processes are special cases of random walks or stochastic processes. Their utility stems from the fact that many otherwise intractable questions and concepts can be answered and explored using the framework of Markov processes. A Markov process can be viewed as a collection of states on which populations live. The ‘master equation’ describes how populations hop from state to state. In this paper we define an open Markov process as one in which there are internal states, where the populations obey the master equation, and boundary states where populations do not obey the master equation because they interact with the external world.

One can visualize an open Markov process as a graph where the edges are labelled by positive real numbers. Each vertex is a ‘state’ and the numbers attached to the edges are transition rates. In the image below, internal states are shaded and boundary states are white.

Figure 1. An open Markov process can be represented by a labelled graph. The numbers on each edge are transition rates. The shaded circles are internal states and the white circles are boundary states.

E-mail address: bpoll1002@ucr.edu.
More precisely, an open Markov process is a triple \((V, H, B)\) where \(V\) is a finite set of states, \(H : \mathbb{R}^V \to \mathbb{R}^V\) is the Hamiltonian which is an infinitesimal stochastic linear operator:

\[
\sum_{i \in V} H_{ij} = 0 \quad \text{and} \quad H_{ij} \geq 0, \quad i \neq j.
\]

The finite set \(B \subseteq V\) is a subset of states called the boundary states. This also defines a set of internal states, \(V - B\). Our dynamical variables are the populations of each state \(p_i \in [0, \infty), \ i \in V\). The vector whose entries are the populations of each state at time \(t\) we call the population distribution, \(p(t) \in [0, \infty)^V\). The Hamiltonian generates the time evolution of the populations at internal states via the master equation

\[
\frac{dp_i(t)}{dt} = \sum_{j \in V} H_{ij}p_j(t), \quad i \in V - B.
\]

In an open Markov process, populations at the boundary states do not obey the master equation, whereas populations at the internal states obey the master equation. An ordinary Markov process is an open Markov process, \((V, H, B)\), with all states being internal, i.e. \(B = \emptyset\). We simply write the pair \((V, H)\) for an ordinary Markov process. Note that in an ordinary Markov process the populations of all states satisfy the master equation.

Given two population distributions, the entropy of \(p\) relative to \(q\) or the relative entropy is given by:

\[
S(p, q) = \sum_{i \in V} p_i \ln \left( \frac{p_i}{q_i} \right).
\]

The relative entropy is sometimes referred to as the Kullback–Leibler divergence \([7]\). Moran, Morimoto, and Csiszar proved that, in an ordinary Markov process, the entropy of any distribution relative to the equilibrium distribution is non-increasing \([3, 9, 10]\). Dupuis and Fischer proved that the relative entropy between any two distributions satisfying the master equation is non-increasing \([4]\). Merhav argues that the Second Law of thermodynamics can be viewed as a special case of the monotonicity in time of the relative entropy in Markov processes \([8]\).

The reason for using relative entropy instead of the usual entropy \(-\sum p_i \ln p_i\) is that the usual entropy is not necessarily a monotonic function of time in Markov processes. If a Markov process has the uniform distribution as its equilibrium distribution, then the usual entropy will increase \([9]\). A Markov process has the uniform distribution as its equilibrium distribution if and only if its Hamiltonian is infinitesimal doubly stochastic, meaning that both the columns and the rows sum to zero. Relative entropy is non-increasing even for Markov processes whose equilibrium distribution is not uniform \([2]\). This suggests the importance of a deeper underlying idea, that of the Markov ordering on the population distributions themselves; see \([5]\) for details. For more information on reversibility and stochastic processes see \([1, 6]\).

The goal of this paper is to study relative entropy in open Markov processes. We show that in an open Markov process \((V, H, B)\), if \(p(t)\) and \(q(t)\) obey the master equation at internal states then the rate of change of relative entropy satisfies the following inequality involving the behavior of the populations at the boundary.
In this expression, $\frac{Dp_i}{Dt}$ is the inflow at the $i$th state, which is given by:

$$\frac{Dp_i}{Dt} = \frac{dp_i}{dt} - \sum_j H_{ij} p_j.$$ 

The inflow measures the amount by which the evolution of the population differs from that given by the master equation. The above inequality is our Second Law for open Markov processes. This inequality tells us that the rate of change of relative entropy in an open Markov process is less than or equal to the rate of change of relative entropy at the boundary. In Section 3 we derive this inequality.

2. Composition of Open Markov Processes

Part of the motivation for considering open Markov processes is to make precise the notion of composition of open Markov processes. One should be able to take two open Markov processes and combine them to get a new Markov process, where probability or population can now flow between the two original processes. This composition is accomplished by gluing two open Markov processes together along some set of boundary states.

Consider the open Markov processes associated to each of the graphs below:

![Figure 2](image2.png)

**Figure 2.** This is a graphical representation of two open Markov processes.

We can compose these two open Markov processes by identifying state 3 with 3':

![Figure 3](image3.png)

**Figure 3.** Composition of the two open Markov processes depicted in Figure 2 results in an ordinary Markov process with no boundary states.

Notice that the state labelled 3 is now an internal state. For this particular example the resulting graph is the special case of an open Markov process with no
boundary states, which is an ordinary Markov process. In general, composition of open Markov processes may result in another open Markov process, i.e. after composition there may be a non-empty set of boundary states.

Suppose we have an ordinary Markov process \((V, H)\), which is the composite of two open Markov processes \((V_1, H_1, B_1)\) and \((V_2, H_2, B_2)\). We consider the case, as in Figures 2 and 3, when \(V = V_1 \cup V_2\) and \(B_2 = B_1 = V_1 \cap V_2\). Given two population distributions \(p\) and \(q\) on \(V\), let us define the following notation for relative entropy:

\[
S_V(p, q) = \sum_{i \in V} \ln \left( \frac{p_i}{q_i} \right).
\]

Using this notation, we write the relative entropy of the composite as,

\[
S_{V_1 \cup V_2}(p, q) = S_{V_1}(p, q) + S_{V_2}(p, q) - S_{V_1 \cap V_2}(p, q).
\]

The third term comes from the fact that the contributions to the relative entropy from the boundary states are counted in both the sum over \(V_1\) and the sum over \(V_2\).

3. The Second Law for Open Markov Processes

In this section we show that the rate of change of relative entropy in an open Markov process is less than or equal to the relative entropy flowing through its boundary states. We use the fact that relative entropy is non-increasing in an ordinary Markov process.

Given a Markov process \((V, H)\) and two population distributions \(p(t), q(t) \in \mathbb{R}^V\), each of which are solutions to the master equation, the entropy of \(p(t)\) relative to \(q(t)\) is

\[
S(p(t), q(t)) = \sum_i p_i(t) \ln \left( \frac{p_i(t)}{q_i(t)} \right).
\]

Following Dupuis and Fischer [4], we can see that relative entropy is non-increasing for Markov processes:

\[
\frac{dS(p(t), q(t))}{dt} = \frac{d}{dt} \sum_i p_i \ln \left( \frac{p_i}{q_i} \right)
\]

\[
= \sum_i \frac{dp_i}{dt} \ln \left( \frac{p_i}{q_i} \right) + \sum_i \frac{q_i}{dt} \ln \left( \frac{p_i}{q_i} \right)
\]

\[
= \sum_i \left[ \sum_j H_{ij} p_j \ln \left( \frac{p_i}{q_i} \right) + \sum_j H_{ij} p_j - \sum_j \frac{p_i}{q_i} H_{ij} q_j \right]
\]

\[
= \sum_i \left[ \sum_{j \neq i} H_{ij} p_j \left[ \ln \left( \frac{p_i}{q_i} \right) - \frac{p_i q_j}{p_j q_i} + 1 \right] + H_{ii} p_i \ln \left( \frac{p_i}{q_i} \right) \right]
\]

\[
= \sum_i \sum_{j \neq i} H_{ij} p_j \left[ \ln \left( \frac{p_i}{q_i} \right) - \frac{p_i q_j}{p_j q_i} + 1 \right] + \sum_j H_{jj} p_j \ln \left( \frac{p_j}{q_j} \right)
\]
\[= \sum_i \sum_{j \neq i} H_{ij} p_j \left[ \ln \left( \frac{p_i}{q_i} \right) - \frac{p_i q_j}{p_j q_i} + 1 \right] - \sum_j \sum_{i \neq j} H_{ij} p_j \ln \left( \frac{p_j}{q_j} \right)\]

\[= \sum_i \sum_{j \neq i} H_{ij} p_j \left[ \ln \left( \frac{p_i q_j}{q_i p_j} \right) - \frac{p_i q_j}{q_i p_j} + 1 \right] \leq 0.\]

The last line follows from the fact that \(H_{ij} \geq 0\) for \(i \neq j\) along with the fact that the term in the brackets \(\ln(x) - x + 1\) is everywhere negative except at \(x = 1\) where it is zero. As \(q_i \to 0\) for some \(i \in V\), the rate of change of relative entropy tends towards negative infinity. One has to allow infinity as a possible value for relative entropy and its first time derivative, in which case the above inequality still holds. Thus, we conclude that for any ordinary Markov process,

\[\frac{d}{dt} S(p(t), q(t)) \leq 0. \] (2)

This gives a version of the Second Law that holds for ordinary Markov processes. Relative entropy is more like free energy, in that it decreases with time. If the reference distribution \(q\) is taken to be the uniform distribution \(q_i = c\) for all \(i\) and for some constant \(c\), then the relative entropy becomes

\[S(p, q) = \sum_i p_i \ln(p_i) - \sum_i p_i \ln(c).\]

If \(\sum_i p_i\) is constant, then for \(q\) uniform, the relative entropy equals the negative of the usual entropy minus a constant. Thus the above calculation for \(\frac{dS(p,q)}{dt}\) gives the usual Second Law.

Now we calculate the rate of change of relative entropy in an open Markov process \((V, H, B)\). Recall that the inflow at the \(i^{th}\) vertex is given by

\[\frac{dp_i}{dt} = \frac{dp_i}{dt} - \sum_{j \in V} H_{ij} p_j.\]

Note that the inflow is zero for internal states as the master equation holds at internal states. Also note the following relations:

\[\frac{\partial S(p, q)}{\partial p_i} = \sum_i \left( \ln \left( \frac{p_i}{q_i} \right) + 1 \right)\]

and

\[\frac{\partial S(p, q)}{\partial q_i} = -\sum_i \frac{p_i}{q_i}.\]
Taking the time derivative of the relative entropy we obtain
\[
\frac{d}{dt} S(p(t), q(t)) = \sum_{i \in V} \frac{dp_i}{dt} \ln \left( \frac{p_i}{q_i} \right) + 1 - \sum_{i \in V} \frac{p_i}{q_i} \frac{dq_i}{dt} \\
= \sum_{i \in V} \sum_{j \in V} H_{ij} p_j \left[ \ln \left( \frac{p_i}{q_i} \right) + 1 - \frac{p_i q_j}{q_i p_j} \right] \\
+ \sum_{i \in B} \left[ \frac{dp_i}{dt} \ln \left( \frac{p_i}{q_i} \right) + 1 - \frac{p_i}{q_i} \frac{dq_i}{dt} \right].
\]

In the last step we separated the contributions from internal and boundary states and used the master equation for the internal states. Now let us add and subtract terms so that the first term corresponds to the rate of change of relative entropy for a Markov process with no boundary states:
\[
\frac{d}{dt} S(p(t), q(t)) = \sum_{i \in V} \sum_{j \in V} H_{ij} p_j \left[ \ln \left( \frac{p_i}{q_i} \right) + 1 - \frac{p_i q_j}{q_i p_j} \right] \\
+ \sum_{i \in B} \sum_{j \in V} \left( \frac{dp_i}{dt} - H_{ij} p_j \right) \left( \ln \left( \frac{p_i}{q_i} \right) + 1 \right) \\
- \sum_{i \in B} \sum_{j \in V} \left( \frac{dq_i}{dt} - H_{ij} q_j \right) \frac{p_i}{q_i}.
\]

The first term is the rate of change of relative entropy for an ordinary Markov process, which is less than or equal to zero by inequality (2). Therefore, we have
\[
\frac{d}{dt} S(p(t), q(t)) \leq \sum_{i \in B} \sum_{j \in V} \left( \frac{dp_i}{dt} - H_{ij} p_j \right) \left( \ln \left( \frac{p_i}{q_i} \right) + 1 \right) \\
- \sum_{i \in B} \sum_{j \in V} \left( \frac{dq_i}{dt} - H_{ij} q_j \right) \frac{p_i}{q_i}.
\]

We can write this more compactly as
\[
\frac{d}{dt} S(p(t), q(t)) \leq \sum_{i \in B} \frac{Dp_i}{Dt} \frac{\partial S}{\partial p_i} + \frac{Dq_i}{Dt} \frac{\partial S}{\partial q_i}.
\]

This gives a version of the Second Law that holds for open Markov processes. One can see that this result reduces to the usual Second Law for an ordinary Markov process, where all states are internal and there are no boundary states.

**Acknowledgements.** I would like to thank John C. Baez for his constant guidance as well as for numerous conversations which helped shape this paper. I thank the Centre for Quantum Technologies and everyone there for their hospitality during my visit. I also thank the FQXi for funding my visit.
OPEN MARKOV PROCESSES

References

[1] P. M. Alberti and A. Uhlmann, Stochasticity and Partial Order: Doubly Stochastic Maps and Unitary Mixing, D. Reidel, Dordrecht, 1982. ↑
[2] T. M. Cover, Which processes satisfy the Second Law?, in Physical Origins of Time Asymmetry, eds. J. J. Halliwell, J. Perez-Mercader and W. H. Zurek, Cambridge University Press, New York, 1994, pp. 98–107. ↑
[3] I. Csiszár, Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten, Publ. Math. Inst. Hungar. Acad. Sci., 8 (1963), 85–108. ↑
[4] P. Dupuis and M. Fischer, On the construction of Lyapunov functions for nonlinear Markov processes via relative entropy, preprint 2012. ↑
[5] A. N. Gorban, P. A. Gorban and G. Judge, Entropy: The Markov ordering approach, Entropy, 12 (2010), 1145–1193. ↑
[6] F. P. Kelly, Reversibility and Stochastic Networks, Wiley, Chichester, 1979. ↑
[7] S. Kullback, R. A. Leibler, On information and sufficiency, Annals of Mathematical Statistics, 22 (1951), 79–86. ↑
[8] N. Merhav, Data processing theorems and the Second Law of Thermodynamics, IEEE Transactions on Information Theory, 58 (2011), 4926–4939. ↑
[9] P. A. P. Moran, Entropy, Markov processes and Boltzmann’s H-Theorem, Proceedings of the Cambridge Philosophical Society, 57 1961, 833–842. ↑
[10] T. Morimoto, Markov processes and the H-Theorem, J. Phys. Soc. Jap, 12 1963, 328–331. ↑