Metastability for the Contact Process on $\mathbb{Z}$

Owen Lynch   Kacper Urbansky

December 4, 2020

1 Overview

The essential idea of metastability is that, as some parameter goes to infinity, we can "coarse-grain" a Markov process ξ_t into something that looks like the Markov process in figure **??**.

Figure 1: Metastability in a Nutshell
??

The system begins in the so-called "metastable state", and has a very small chance to move to the stable state at any time. The stable state is either absorbing, or close-to absorbing; in the case that we will talk about in this paper, the stable state is absorbing.

In order to make this "coarse-graining" happen, we need two things to be true asymptotically.

1. The hitting time of the stable state is exponentially distributed.

2. Up until the hitting time, temporal means of measurements made to the process converge to a stationary distribution.

Together, these two properties intuitively allow us to approximate the entire process by sampling from the stationary distribution up until the hitting time, and then putting the system in the absorbing state after the hitting time.

The difficulty comes in stating these two properties precisely. To do this, suppose that N is the parameter of the system that goes to ∞, (i.e. we have a collection of processes $\xi_N(t)$), T_N is the hitting time of the absorbing state, μ is the stationary distribution, and R_N is a "time scale" parameter that satisfies $R_N / \mathbb{E} \, T_N \to 0$. Then we rewrite the two properties more formally as

1. $T_N / \mathbb{E} \, T_N \to \text{Exp}(1)$ in distribution as $N \to \infty$.

2. For any f,
$$\int_S^{S+R_N} f(\xi_N(t)) \to \mu(f)$$

as $N \to \infty$, for any $S + R_N < T_N$.

This second statement is still very imprecise, and actually mathematically meaningless as currently posed. Also, it turns out that we want a much stronger statement than that. However, we hope that this first statement should "innoculate" the reader to the precise statement, which is fairly dense on its own.

2 Review of Contact Process

2.1 Basic Definitions

The contact process on $\mathbb{Z}$ can be defined by the Markov pregenerator L with domain cylindrical functions $2^{\mathbb{Z}} \to \mathbb{R}$ given by

$$Lf(\eta) = \sum_{x \in \mathbb{Z}} c(x, \eta)(f(\eta^x) - f(\eta)) \tag{1}$$

where

$$c(x, \eta) = \begin{cases} 1 & \text{if} \quad \eta(x) = 1 \\ (\eta(x-1) + \eta(x+1)) & \text{if} \quad \eta(x) = 0 \end{cases}$$

λ is the single parameter for the contact process; we will discuss this more later.

However, there is an alternative definition that lends itself better to certain constructions that are useful in proofs.

First we construct a "percolation structure", which consists of for each $x \in \mathbb{Z}$

1. A poisson process P_x with rate 1, which we call the "death" process at x.

2. A poisson process $P_{x \to x+1}$ with rate λ, which we call the "right infection" process at x.

3. A poisson process $P_{x \to x-1}$ with rate λ, which we call the "left infection" process at x.

We consider P_x to be a random element of $\mathcal{P}(\mathbb{R})$, i.e. $t \in P_x$ if and only if the poisson process "ticks" at time t.

Figure 2: Example Percolation Structure

We define a "path" between $(x, s), (y, t) \in \mathbb{Z} \times \mathbb{R}$ with $s \leq t$ to be a sequence $(z_0, r_0), \ldots, (z_n, r_n)$ with $r_i \leq r_{i+1}$ such that for all $(z_i, r_i), (z_{i+1}, r_{i+1})$, either

1. $r_i = r_{i+1}$, $|z_i - z_{i+1}| = 1$, and $r_i \in P_{z_i \to z_{i+1}}$. In this case, we are jumping laterally by one line at a time of infection.

2. $z_i = z_{i+1}$, and $[r_i, r_{i+1}] \cap P_{z_i} = \emptyset$. In this case, we are moving along a vertical line that has no deaths in it.

Then $\xi^A(t)$ is the set of y such that there is a path from $(x, 0)$ to (y, t) for some $x \in A$. If the superscript is omitted, then we assume $A = \mathbb{Z}$, i.e. $\xi(t) = \xi^{\mathbb{Z}}(t)$.

For any $A \subseteq B$, we define $\xi_B^A(t)$ to be the set of y such that there is a path from $(x, 0)$ to (y, t) for some $x \in A$ that stays entirely within B. As a special case, we let $\xi_N^A(t) = \xi_{[-N,N]}^A$, for $N \in \mathbb{N}$. Note that ξ_B takes values exclusively in $\mathcal{P}(B)$.

2.2 Useful Properties

One of the most important facts about the contact process is that there is a critical value of λ, λ_c. For $\lambda < \lambda_c$, $\xi(t)$ has unique extremal invariant measure $\delta_\emptyset$, but for $\lambda > \lambda_c$, $\xi(t)$ has another extremal invariant measure μ defined by

$$\mu(f) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathbb{E}\, f(\xi(t))\, \mathrm{d}t$$

and this measure is not concentrated at $\emptyset$.

For $\xi_N(t)$, the only invariant measure is $\delta_\emptyset$, because $\emptyset$ is a trap and ξ_N takes values on a finite state space. However, as mentioned before, we will see later on that $\xi_N(t)$ *approximately* has invariant measure μ all the way up until it reaches $\emptyset$.

TODO: give brief explanations for each of these.

Proposition 1. $\mathbb{P}(\xi(t) \cap A \neq \emptyset)$ *monotonically decreases to* $\mu(\eta \mid \eta \cap A \neq \emptyset)$.

Proposition 2. *The contact process is* self-dual.

Corollary 2.1. $\mathbb{P}(\xi^A(s) \neq \emptyset, \forall s \geq 0) = \mu(\eta \mid \eta \cap A \neq \emptyset)$

Proposition 3. *Let* $\rho = \mathbb{P}(\xi^{\{0\}}(s) \neq \emptyset) = \mu(\eta \mid \eta(0) = 1)$, *if* λ *is the parameter for* ξ. *Then if* $\lambda > \lambda_c$ *then* $\rho > 0$. *This will later be an important quantity for proofs.*

Proposition 4. μ *is translation-invariant, and even better, ergodic under translations. This means that one can approximate* $\mu(f)$ *by averaging* $\mathbb{E}\,\tau_i f(\xi(t))$ *over a large number of* i, *where* $\tau_i f(\eta)$ *is* f *applied to* η *shifted by* i. *Finally,* μ *has exponentially decaying correlations; that is if* $\{X_i\}_{i \in \mathbb{Z}}$ *is distributed according to* μ, *then the correlation between* X_i *and* X_j *decays exponentially with* $|i - j|$.

3 Hitting Time of Stable State

Theorem 5. *If* $T_N = \inf\{t > 0 \mid \xi_N(t) \neq \emptyset\}$, *then*

$$\frac{T_N}{\mathbb{E}\, T_N} \to \mathrm{Exp}(1)$$

in distribution.

Without further ado, we get down to the proof. First of all, we replace $\mathbb{E}\,T_N$ by the unique (by monotonicity) β_N such that $\mathbb{P}(T_N > \beta_N) = \mathrm{e}^{-1}$. At the end, we will show that $\frac{\mathbb{E}\,T_N}{\beta_N} \to 1$.

Now let $G_N(t) = \mathbb{P}[\frac{T_N}{\beta_N} > t]$, the CDF of T_N/β_N. Similarly, let $G_N^A(t) = \mathbb{P}[\frac{T_N^A}{\beta_N} > t]$. Note that $G_N^A(t) \le G_N(t)$ To show that T_N/β_N converges in distribution to $\mathrm{Exp}(1)$, we must show that $G_N(t) \to \mathrm{e}^{-t}$. This can be accomplished by showing that

$$\lim_{N \to \infty} |G_N(t+s) - G_N(t)G_N(s)| = 0$$

for all $t, s > 0$.

Note that $G_N(t) = \mathbb{P}[\xi_N(t) \ne \emptyset]$, as ξ_N is "alive" at time t if and only if $T_N > t$. Thus,

$$
\begin{aligned}
G_N(t+s) &= \mathbb{P}[\xi_N(t+s) \ne \emptyset] \\
&= \sum_{A \ne \emptyset} \mathbb{P}[\xi_N(t+s) \ne \emptyset | \xi_N(t) = A]\,\mathbb{P}[\xi_N(t) = A] \\
&= \sum_{A \ne \emptyset} G_N^A(s)\,\mathbb{P}[\xi_N(t) = A] \\
&\le G_N(s) \sum_{A \ne \emptyset} \mathbb{P}[\xi_N(t) = A] \\
&= G_N(s)G_N(t)
\end{aligned}
$$

Therefore, we are looking to show that $G_N(s)G_N(t) - G_N(t+s) \to 0$ (we can forget the absolute value signs).

It is at this point that we introduce a curious little construction, which seems to not make much sense at first but turns out to be the key to the entire proof. Define F_b for $b > 0$ by

$$F_b = \{A \subseteq \mathbb{Z} \mid \frac{|A \cap [-b, -1]|}{b} > frac\rho 2, \frac{|A \cap [1, b]|}{b} > frac\rho 2\}$$

Recall that ρ was defined in **??** to be the probability that $\xi^{\{0\}}$ never dies. The intuition that the reader should hold for F_b is that it is the set of A that are "sufficiently dense" on $[-b, -1]$ and $[1, b]$ so that ξ_N^A behaves just like ξ_N. The reason we have this "two-sided" condition is that we will use the processes $\xi_{(-\infty, N)}$ and $\xi_{(-N, \infty)}$, which have invariant measure μ as well, to talk about ξ_N.

Now, by similar reasoning as before,

$$
\begin{aligned}
G_N(t+s) &= \sum_{A \ne \emptyset} G_N^A(t)\,\mathbb{P}[\xi_N(\beta_N s) = A] \\
&\ge \sum_{A \in F_b} G_N^A(t)\,\mathbb{P}[\xi_N(\beta_N s) = A] \\
&\ge \min_{A \in F_b} G_N^A(t)\,\mathbb{P}[\xi_N(\beta_N s) \in F_b]
\end{aligned}
$$

4

(We assume that $A \subseteq [-N, N]$ when such an assumption is necessary).

Therefore,

$$
\begin{aligned}
G_N(t)G_N(s) - G_N(t+s) &\leq G_N(t)G_N(s) - \min_{A \in F_b} G_N^A(t)\, \mathbb{P}[\xi_N(\beta_N s) \in F_b] \\
&= G_N(s)(G_N(t) - \min_{A \in F_b} G_N^A(t)) \\
&\quad + \min_{A \in F_b} G_N^A(t)(G_N(s) - \mathbb{P}[\xi(\beta_N s) \in F_b]) \\
&\leq (G_N(t) - \min_{A \in F_b} \mathbb{P}[\xi_N^A(\beta_N t) \neq \emptyset]) \\
&\quad + \mathbb{P}[\xi_N(\beta_N s) \neq \emptyset, \xi_N(\beta_N s) \notin F_b]
\end{aligned}
$$

We will have finished if for any $\varepsilon > 0$, we can find $b(\varepsilon)$ and $N(\varepsilon) > b(\varepsilon)$ such that for $N \geq N(\varepsilon)$ and $A \in F_b$, we have both

$$G_N(t) - G_N^A(t) < \varepsilon \tag{2}$$

$$\mathbb{P}[\xi_N(\beta_N s) \neq \emptyset, \xi_N(\beta_N s) \notin F_b] < \varepsilon \tag{3}$$

We tackle (2) first. Remember that $\xi_N(t)$ and ξ_N^A are defined on the same percolation structure. Therefore, $\xi_N(t) \supset \xi_N^A(t)$, so we have

$$\mathbb{P}[\xi_N(\beta_N t) \neq \emptyset] - \mathbb{P}[\xi_N^A(\beta_N t) \neq \emptyset] = \mathbb{P}[\xi_N(\beta_N t) \neq \emptyset, \xi_N^A(\beta_N t) = \emptyset] \leq \mathbb{P}[T_N \neq T_N^A]$$

The intuition for why the right hand side of this is small is that once $\xi_N(t_0) = \xi_N^A(t_0)$, then for all $t > t_0$, $\xi_N(t) = \xi_N^A(t)$. Therefore, as long as the event $\xi_N(t_0) = \xi_N^A(t_0)$ happens before T_N^A, we will have $T_N^A = T_N$. Let E be this event; we must show that $\mathbb{P}(E) > 1 - \varepsilon$.

As $\lambda > \lambda_c$, we can find $n(\varepsilon)$ such that

$$\mu(B \mid B \cap [1, n(\varepsilon)] = \emptyset) \leq \frac{\varepsilon}{2}$$

Then take $b = b(\varepsilon)$ such that $n(\varepsilon) \leq b \cdot \rho/2$. For $A \in F_b$, $|A \cap [-b, -1]| \geq b \cdot \rho/2 \geq n(\varepsilon)$, so

$$\mathbb{P}[T_{[-N,\infty)}^{A \cap [-b,-1]} = \infty] \geq \mathbb{P}[T_{[-N,\infty]}^{[-N,\ldots,-N+n(\varepsilon)]} = \infty] \geq 1 - \frac{\varepsilon}{2}$$

TODO give argument for the first inequality. The second inequality follows from duality and monotone convergence.

Similarly, $\mathbb{P}[T_{(-\infty,N]}^{A \cap [1,b]} = \infty] \geq 1 - \frac{\varepsilon}{2}$. Let E' be the event that both of these stopping times are infinite; clearly $P(E') \geq 1 - \varepsilon$. It remains to show that $E \supset E'$.

Define stopping times

$$
\begin{aligned}
U &= \inf\{t > 0 \mid N \in \xi_{[-N,\infty)}^{A \cap [-b,-1]}(t)\} \\
V &= \inf\{t > 0 \mid -N \in \xi_{(-\infty,N]}^{A \cap [1,b]}(t)\}
\end{aligned}
$$

Until U, $\xi_N^{A \cap [-b,-1]}(t) = \xi_{[-N,\infty)}^{A \cap [-b,-1]}(t)$. Therefore, on E', $\xi_N^{A \cap [-b,-1]}(t)$ is alive up until U, whence ξ_N^A is alive up until U. Similarly, ξ_N^A is alive up until V. Therefore, on E', $T_N \geq T_N^A > \max(U, V)$.

However, for $t > \max(U, V)$, I claim that $\xi_N(t) = \xi_N^A(t)$. To see this, any path to $x \in \xi_N(t)$ must either cross the path that goes from $A \cap [-b, -1]$ at time 0 to N at time U or the path that goes from $A \cap [1, b]$ at time 0 to $-N$ at time V. Therefore, there is a path from A at time 0 to x at time t, whence $x \in \xi_N^A(t)$. Therefore, on E', $T_N = T_N^A$, and we have shown (2).

Equation 3 also relies crucially on the percolation structure. Let $D_b = F_b^C \setminus \{\emptyset\}$. Then we must show that $\mathbb{P}(\xi_N(\beta_N s) \in D_b) < \varepsilon$. To do this, note that as long as $\xi_N(t)$ is alive, we have

$$\xi_N(t) = \xi(t) \cap [\min \xi_N(t), \max \xi_N(t)]$$

This is because any path to $x \in \xi(t) \cap [\min \xi_N(t), \max_N(t)]$ must either stay inside $[\min \xi_N(t), \max_N(t)]$, in which case $x \in \xi_N(t)$, or pass through the path to $\min \xi_N(t)$ or $\max \xi_N(t)$, in which case we also have $x \in \xi_N(t)$.

Therefore, as long as $\min \xi_N(\beta_N s) < -N + L$, $\max \xi_N(\beta_N s) > N - L$ and $\xi_N(\beta_N s) \neq \emptyset$, we have $\xi(\beta_N s) \in D_b$ iff $\xi_N(\beta_N s) \in D_b$.

We can minimize $\xi(\beta_N s) \in D_b$ by picking $b > b''(\varepsilon)$ sufficiently large. In order to use this, we use the following decomposition

$$\begin{aligned}
\mathbb{P}[\xi_N(\beta_N s) \in D_b] \leq\ &\mathbb{P}[\xi_N(\beta_N s) \in D_b, \min \xi_N(\beta_N s) < -N + L, \max \xi_N(\beta_N s) > N - L] \\
&+ \mathbb{P}[\min \xi_N(\beta_N s) \geq -N + L, \xi_N(\beta_N s) \neq \emptyset] \\
&+ \mathbb{P}[\max \xi_N(\beta_N s) \leq N - L, \xi_N(\beta_N s) \neq \emptyset]
\end{aligned}$$

4 Almost-Ergodicity up to Hitting Time