

Metastability for the Contact Process on \mathbb{Z}

Owen Lynch

Kacper Urbański

December 14, 2020

Metastability is the coexistence of an equilibrium and a “quasi-equilibrium” (metastable) state in a system. If we look at a metastable system in time domain, it initially appears to have stabilized in the “quasi-equilibrium” state, until it suddenly relaxes to the true equilibrium. Metastability is a common phenomenon in nature, with examples ranging from physics to economics and social sciences. This work is a summary of a paper by R. Schonmann [Sch85], which frames this phenomenon as a rigorous property of an abstract interacting particle system - the Contact Process.

1 Overview

1.1 Metastability

Let $\{\xi_N(t)\}_{N \in \mathbb{N}}$ be a family of Markov processes. The essential idea behind metastability is that, if we were to discretize (“coarse-grain”) time, increasing N would make the process under consideration look more and more like the Markov process in figure 1.

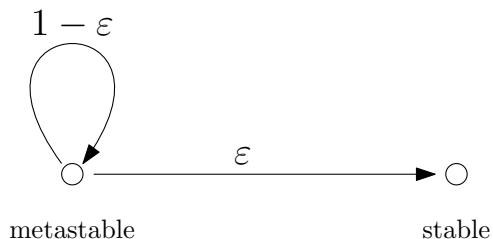


Figure 1: Metastability in a Nutshell

In other words, the system begins in a 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 an expectation of those measurements with respect to some stationary distribution.

Informally, these two properties 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.

The difficulty comes in stating these two properties precisely. To do this, suppose that 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 $R_N/\mathbb{E} T_N \rightarrow 0$. Then we rewrite the two properties more formally as

1. $T_N/\mathbb{E} T_N \rightarrow \text{Exp}(1)$ in distribution as $N \rightarrow \infty$.
2. For any f cylindrical,

$$\int_S^{S+R_N} f(\xi_N(t)) dt \rightarrow \mu(f)$$

as $N \rightarrow \infty$, 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.

1.2 The Contact Process on \mathbb{Z}

We will use the “percolation structure” definition of the contact process, which lends itself better to certain useful constructions.

This “percolation structure”, 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 $P_{x \rightarrow x+1}$ with rate λ , which we call the “right infection” process at x .
3. A Poisson process $P_{x \rightarrow x-1}$ 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 .

We define a “path” between $(x, s), (y, t) \in \mathbb{Z} \times \mathbb{R}$ with $s \leq t$ to be a sequence $(z_0, r_0), \dots, (z_n, r_n)$ 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 $r_i \in P_{z_i \rightarrow z_{i+1}}$. 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, uninterrupted by any deaths.

Define $\xi^A(t)$ to be the set of y such that there is a path from $(x, 0)$ to (y, t) for some $x \in A$. 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 $\mathcal{P}(B)$.

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 a unique extremal invariant measure δ_\emptyset . At $\lambda = \lambda_c$ the system undergoes a phase transition - for $\lambda > \lambda_c$, another extremal invariant measure μ appears, with the property that

$$\mu(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbb{E} f(\xi(t)) dt$$

This measure is not concentrated at \emptyset .

For $\xi_N(t)$, the only invariant measure is δ_\emptyset , because \emptyset is a trap and ξ_N takes values on a finite state space. However, as mentioned before, for $\lambda > \lambda_c$ an analogue of a phase transition takes place. The system starts being metastable, and stays distributed *approximately* as μ , before a sudden fluctuation takes it to \emptyset .

We assume that the reader has a general familiarity with the contact process, and thus we will use some properties later on in the proofs without explicitly discussing them here.

1.3 Summary

The object of this paper is to give an overview of the two conditions for metastability for the family of contact processes ξ_N , in the supercritical regime $\lambda > \lambda_c$.

2 Main Theorems

2.1 Exponential Distribution of Hitting Time

Theorem 1 (from [Sch85]). *If $T_N = \inf\{t > 0 \mid \xi_N(t) \neq \emptyset\}$, then*

$$\frac{T_N}{\mathbb{E} T_N} \rightarrow \text{Exp}(1)$$

in distribution.

To prove this, we first replace $\mathbb{E} T_N$ by the unique (by monotonicity) β_N such that $\mathbb{P}(T_N > \beta_N) = e^{-1}$. At the end, we will show that $\frac{\mathbb{E} T_N}{\beta_N} \rightarrow 1$.

Let $G_N(t) = \mathbb{P}[\frac{T_N}{\beta_N} > t]$, the cumulative distribution function of T_N/β_N . Similarly, let $G_N^A(t) = \mathbb{P}[\frac{T_N^A}{\beta_N} > t]$. Note that $G_N^A(t) \leq G_N(t)$. To show that T_N/β_N converges in

distribution to $\text{Exp}(1)$, we must show that $G_N(t) \rightarrow e^{-t}$. This can be accomplished by showing that

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

for all $t, s > 0$.

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

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

where $\rho = \mu(\{\eta \mid \eta(0) = 1\})$.

It is not trivial to show that

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

However, the proof is not terribly interesting, and so we refer the reader to Schonman for the details.

The intuition for what this equation is claiming is that we can show that starting in $A \in F_b$ is not too different from starting in $[-N, N]$, and we can show that ending *anywhere* non-empty is not too different from ending up somewhere in F_b . Then a typical process will have probability $G_N(t)$ to end up in $A \in F_b$ at time $\beta_N t$, and then probability $G_N(s)$ to end up anywhere non-empty at time $\beta_N(t+s)$, starting in some $A \in F_b$, so total probability to end up non-empty at time $\beta_N(t+s)$ is $G_N(t)G_N(s)$, as required.

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

$$\mathbb{P}[\xi_N(\beta_N t) \neq \emptyset] - \mathbb{P}[\xi_N^A(\beta_N t) \neq \emptyset] = G_N(t) - G_N^A(t) < \varepsilon \quad (1)$$

$$\mathbb{P}[\xi_N(\beta_N s) \neq \emptyset, \xi_N(\beta_N s) \notin F_b] < \varepsilon \quad (2)$$

We tackle (1) 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

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

We discussed this proof in the presentation. In short, at a certain time t , $\xi_N(t) = \xi_N^A(t) \neq \emptyset$, and by picking b large enough, we can ensure that the process is most likely still alive at this time. Then because of the percolation structure definition, for all $s > t$, $\xi_N(s) = \xi_N^A(s)$, whence $T_N = T_N^A$ with probability greater than $1 - \varepsilon$.

Equation 2 also relies crucially on the percolation structure. Because of the percolation structure, as long as $\xi_N(t) \neq \emptyset$

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

To make use of this, we say that $\xi_N(t)$ is “wide” if $\min \xi_N(t) < -N + L$, $\max \xi_N(t) > N - L$, for some fixed L . Then as long as $[-b, b] \subseteq [-N + L, N - L]$, for wide $\xi_N(t)$, $\xi_N(t) \in F_b$ if and only if $\xi(t) \in F_b$. Now, let $D_b = F_b^C \setminus \{\emptyset\}$.

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

By what we noted earlier, the first term is less than $\mathbb{P}(\xi(\beta_N s) \in D_b)$, and we can pick b such that this is less than $\frac{\varepsilon}{3}$.

It remains to minimize the last two terms; by symmetry we only show how to minimize the first. Using a percolation structure argument, it is easy to show that as long as $\xi_N(\beta_N s) \neq \emptyset$, $\min \xi_N(t) = \min \xi_{[-N, \infty)}$. Therefore,

$$\begin{aligned} \mathbb{P}[\min \xi_N(\beta_N s) \geq -N + L, \xi_N(\beta_N s) \neq \emptyset] &\geq \mathbb{P}[\min \xi_{[-N, \infty)} \geq -N + L] \\ &\geq \mu_{[-N, \infty)}\{A \in [-N, \infty) \mid A \cap [-N, -N + L - 1] = \emptyset\} \end{aligned}$$

We can pick L large enough to make that last term less than $\frac{\varepsilon}{3}$, and we have shown that everything can be made as small as we like, so we are done.

The last thing to do is to show that $\beta_N / \mathbb{E} T_N \rightarrow 1$; this follows from the above proof, as the unique β such that $\mathbb{P}[(1) > \beta] = e^{-1}$ is $1 = \mathbb{E}[(1)]$.

2.2 Quasi-stationarity up to the Hitting Time

Recall our initial description of Quasi-stationarity in Section 1. To make it more precise, we will define

the intermediate timescale to be a choice of $R_N \in \mathbb{R}_+$ with $R_N / \mathbb{E}[T_N] \rightarrow 0$

an observable quantity to be a cylindrical (local) $f : \{0, 1\}^{\mathbb{Z}} \rightarrow \mathbb{R}$

the temporal mean of observable quantity f to be

$$A_{R_N}^N(s, f) := R_N^{-1} \int_s^{s+R_N} f(\xi_N(t)) dt$$

the fixed probability distribution to be μ (the upper invariant measure of $\xi(t)$)

Note that $A_{R_N}^N$ is a random variable. We will say that this quantity is *close* to $\mu(f)$ if the two converge in probability when $N \rightarrow \infty$.

It seems like the definitions given above would be sufficient to make our natural language definition of metastability formal. However, it turns out that an additional technical condition is needed. Define

$$\Lambda(f) := \text{smallest } B \subset \mathbb{Z} \text{ s.t. } f(A) = f(A \cap B) \forall A \subset \mathbb{Z}$$

Although this is not strictly true, we can think of $\Lambda(f)$ as the “support” of f .

Our theorem will hold for a specific $L(\varepsilon, f) \in \mathbb{N}$, and we need to have that

$$L < N \quad \text{and} \quad \Lambda(f) \subset [-N + L, N - L]$$

Notice that L does not depend on N . Thus, since we chose to grow $N \rightarrow \infty$, having to choose this L doesn't restrict our choice of f - it merely sets the minimum N we can consider.

We're now ready to give a simplified formulation of what we mean by "quasi-stationarity"

Theorem 2 (from [Sch85]). *If $\lambda > \lambda^*$ there is a sequence $\{R_N\}_{N \in \mathbb{N}} \subset \mathbb{R}_+$ such that:*

(a) $R_N/\beta_N \rightarrow 0$ as $N \rightarrow \infty$

(b) For all $\varepsilon > 0$ and $f : \{0, 1\}^{\mathbb{Z}} \rightarrow \mathbb{R}$ cylindrical, $\exists L(\varepsilon, f) \in \mathbb{N}$ such that

$$\mathbb{P} \left[\max_{\mathbb{N}_0 \ni k < K_N} |A_{R_N}(kR_N, f) - \mu(f)| > \varepsilon \right] \rightarrow 0$$

as $N \rightarrow \infty$, where $K_N = \max\{k \in \mathbb{N}_0 : kR_N < T_N\}$ and $\Lambda(f) \subset [-N + L, N - L] \cap \mathbb{Z}$

A full proof of this statement is beyond the scope of this article—we will merely give the reader some insight into its key points.

A central claim is that we can make the exceedance probability (given below) arbitrarily small.

$$\mathbb{P} [|A_{R_N}(kR_N, f) - \mu(f)| > \varepsilon]$$

We do so in a way which is uniform across all values of K_N and $\mathbb{N}_0 \ni k < K_N$. We estimate this probability with a triangle inequality.

$$\mathbb{P} \left[\left| R_N^{-1} \int_{kR_N}^{(k+1)R_N} f(\xi(t)) dt - \mu(f) \right| > \varepsilon/2 \right] + \mathbb{P} \left[\left| R_N^{-1} \int_{kR_N}^{(k+1)R_N} f(\xi_N(t)) - f(\xi(t)) dt \right| > \varepsilon/2 \right] \quad (3)$$

The first term can be made arbitrarily small, as a consequence of $\xi(t) \rightarrow \mu(f)$ and exponentially decaying temporal correlations of $\xi(t)$.

To shrink the second term in (3) we make extensive use of the graphical model construction of the Contact Process. As before, we say $\xi_N(t)$ is *wide* at t if $\min \xi_N(t) < -N + L \wedge \max \xi_N(t) > N - L$. By the fact that we construct all processes on the same percolation structure, and that interactions are between nearest neighbours, we have

Lemma 3 (Shielding by a wide process). *If $\xi_N(t)$ is wide at t , then*

$$\xi_N(t) = \xi(t) \quad \text{on} \quad [-N + L, N - L] \cap \mathbb{Z}$$

In particular, since $\Lambda(f) \subset [-N + L, N - L]$, we have $f(\xi_N(t)) = f(\xi(t))$

Thus, the second term of (3) will go to 0 if we're able to make it arbitrarily likely for ξ_N to be wide. For any given N , it is clear that we can find such an L , but we must show that we can find a L that works for all N .

To show this, we again make use of the percolation structure. Note that near $-N$, ξ_N will “look similar” to $\xi_{[-N, \infty)}$ (this is true as long as ξ_N is alive). By the same reasoning, ξ_N will “look similar” to $\xi_{(-\infty, N]}$ near N . Hence we set L such that

$$\mu_{[-N, \infty)}(\{A : A \cap [-N, -N + L] = \emptyset\}) \leq \varepsilon/(16\|f\|)$$

This makes all possible states of ξ_N that do not intersect $[-N, -N + L]$ very unlikely (by “looking similar” to $\xi_{[-N, \infty)}$). By symmetry, this also makes states that do not intersect $[N - L, N]$ very unlikely. As a final stroke, we use translation invariance to notice that we will get exactly the same effect if we set L

$$\mu_{[0, \infty)}(\{A : A \cap [0, L] = \emptyset\}) \leq \varepsilon/(16\|f\|)$$

This concludes our overview of the proof of Theorem 2.

References

- [Sch85] Roberto H. Schonmann. “Metastability for the contact process.” In: *Journal of Statistical Physics* Vol. 41, Issue 3 (1985).