Metastability for the Contact Process on Z: Part 2

Kacper Urbański

Universiteit van Amsterdam

kacper.urbanski@protonmail.com

December 4, 2020

Overview

A bit of terminology

- ξ - 'xi'
- ξ_N - 'xi n'
- $\xi_{[-N,\infty)}$ - 'xi plus inf'
- $\xi_{(-\infty,N]}$ - 'xi minus inf'
- $[-N, N]$ - 'main interval'
- $\xi_N(t) \neq \varnothing$ - 'process is still alive'

Natural language definition of metastability

Recall that a system is metastable if:

1. It stays out of its equilibrium during a memoryless random time
2. During this time in which the system is out of equilibrium it stabilizes

Let's elaborate some more on point 2.

1. Assume that for a given N we have some intermediate timescale such R_N that $R_N << \beta_N$
2. Say we measure a temporal mean of some observable quantity of a system (e.g. particle density) over this timescale
3. We say system has stabilized if this mean is close to the expectation of this observable quantity w.r.t. some fixed probability distribution on $\{0, 1\}^{\mathbb{Z}}$

To ensure $R_N << \beta_N$, let's require $R_N/\beta_N \to 0$ as $N \to \infty$.

For our purposes, define **observable quantity of a system** as f, such that:

- $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$
- f is local

Define

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

We can think of $\Lambda(f)$ as the "support" of f.

Towards the rigorous definition

Define **temporal mean** of observable quantity $f(\xi_N(t))$ as:

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

Where:

- s is the time in which we start our measurement
- R is the duration over which we calculate the temporal mean

- Take our **fixed probability distribution** to be μ (i.e. the non-zero invariant measure of the contact process in the supercritical regime).
- Then, we can define **expectation of observable quantity** w.r.t to this probability distribtuion as $\mu(f) := \int f d\mu$
- Take convergence in probability as how we understand **closeness**.

Let's try to work out a 1st draft of the theorem.

We want to have a sequence R_N, such that:

- $R_N/\beta_N \to 0$ as $N \to \infty$
- for all $\varepsilon > 0$ and all observable quantities f

$$\mathbb{P}\left[|A_{R_N}(s, f) - \mu(f)|\langle \varepsilon\right] \to 0$$

As $N \to \infty$.

Are we done now?

Let's try to work out a 1st draft of the theorem.

We want to have a sequence R_N, such that:

- $R_N/\beta_N \to 0$ as $N \to \infty$
- for all $\varepsilon > 0$ and all observable quantities f

$$\mathbb{P}\left[|A_{R_N}(s, f) - \mu(f)|\langle\, \varepsilon\right] \to 0$$

As $N \to \infty$.

Are we done now? **No!** How do we choose s (the starting point of temporal mean measurement)?

Define

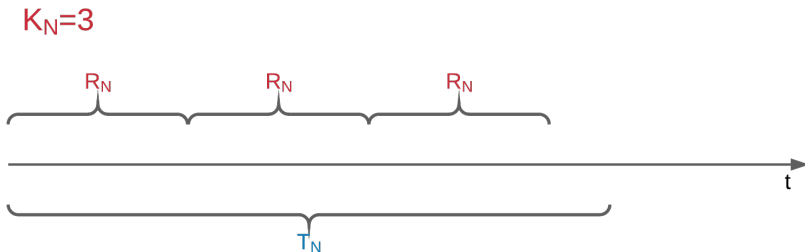$$K_N = \max\{k \in \mathbb{N}_0 : kR_N < T_N\}$$

$K_N = 3$

Figure: Example of K_N

Turns out that even if we choose the worst-case interval for averaging, we can still make it highly probable that our temporal mean is as close to the $\mu(f)$ as we want, given that we're free to increase N.

In other words, we want to have a sequence R_N, such that:

- $R_N/\beta_N \to 0$ as $N \to \infty$
- for all $\varepsilon > 0$ and all observable quantities f

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

As $N \to \infty$.

This draft is almost ready - we now only have one more technicality left.

We need to choose $L(\varepsilon, f) \in \mathbb{N}$, and we need to have that

- $L < N$
- $\Lambda(f) \subset [-N + L, N - L]$ (call $[-N + L, N - L]$ **inside region**)

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

We need to choose $L(\varepsilon, f) \in \mathbb{N}$, and we need to have that

- $L < N$
- $\Lambda(f) \subset [-N + L, N - L]$ (call $[-N + L, N - L]$ **inside region**)

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

Additionally, define

- $[-N, -N + L)$ - **left boundary region**
- $(N - L, N]$ - **right boundary region**

Figure: Partition of the main interval

Theorem 2

We're finally ready to formulate the theorem.

Theorem (Thermalization)

If $\lambda > \lambda^*$ there is a sequence $\{R_N\}_{N \in \mathbb{N}} \subset \mathbb{R}_+$ such that:

- $R_N/\beta_N \to 0$ as $N \to \infty$
- For all $\varepsilon > 0$ and observable quantities f $\exists L(\varepsilon, f) \in \mathbb{N}$ such that

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

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

Proof

Set
$$B_k^N = \{|A_{R_N}(kR_N, f) - \mu(f)| > \varepsilon\}$$

On B_k^N the difference between temporal mean of observable measured over the k-th interval and its expectation is greater than we'd like to. B_k^N means **failure** (in k-th interval).

We want the probability of having no failures to go to 1

$$\mathbb{P}\left[\max_{\mathbb{N}_0 \ni k < K_N} |A_{R_N}(kR_N, f) - \mu(f)| \leq \varepsilon\right] = \mathbb{P}\left[\bigcap_{\mathbb{N}_0 \ni k < K_N} \left(B_k^N\right)^C\right] \to 1$$

as $N \to \infty$

Proof

It is easy to show that $\mathbb{P}(K_N = 0) \to 0$. Thus, we can safely focus only on the subset of Ω where $K_N \geq 1$.

Combining this with the event from the previous slide and applying some simple algebra, we arrive at the following bound

$$\mathbb{P}\left[K_N \geq 1, \bigcap_{\mathbb{N}_0 \ni k < K_N} \left(B_k^N\right)^C\right] \geq \mathbb{P}[1 \leq K_N \leq m] - m^2 \max_{1 \leq j} \max_{0 \leq k < j} \mathbb{P}\left[B_k^N, K_N = j\right]$$

Our objective will be to find a sequence $\{m_N\}_{N \in \mathbb{N}}$ such that:

- $\mathbb{P}[1 \leq K_N \leq m_N] \to 1$
- $m_N^2 \max_{1 \leq j} \max_{0 \leq k < j} \mathbb{P}\left[B_k^N, K_N = j\right] \to 0$

As $N \to \infty$

Proof

We will start with the second term. However, first we need some intermediate results. We will say $\xi_N(t)$ is **wide** at t if it intersects both boundary regions. In other words:

$$\min \xi_N(t) < -N + L \wedge \max \xi_N(t) > N - L$$

We will call the process **narrow** otherwise.

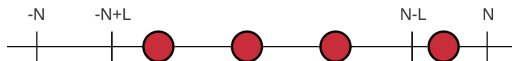Figure: Snapshot at t of a process wide at t

Figure: Snapshot at t of a process narrow at t

Proof

Lemma (Shielding by a wide process)

If $\xi_N(t)$ is wide at t, then

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

In particular we have $f(\xi_N(t)) = f(\xi(t))$

Why would this be true?

Essentially, rightmost and leftmost infected individuals "shield" the entire space between them from outside influence. GIF SHOWING WHAT I MEAN

Proof

Define

$$h_L(\eta) = I_{\{\xi : \xi \cup [-N, -N+L] = \varnothing\}}$$

Notice that h_L can tell us whether a process is wide or not. If we take S to be a horizontal flip operator, then $h_L(\xi_N(t))h_L(S\xi_N(t))$ is the desired indicator.
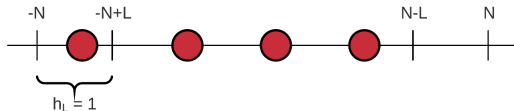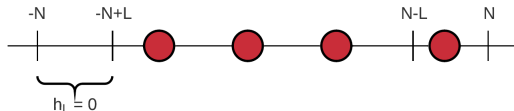

Figure: $h_L(\xi_N(t)) = 1$

Figure: $h_L(S\xi_N(t)) = 0$

Lemma (Shielding of the left boundary region)

$$\{T_N > t\} \subset \{h_L(\xi_N(t)) = h_L(\xi_{[-N,\infty)}(t))\}$$

In other words, if $\xi_N(t)$ is still alive, it fully determines whether $\xi_{[-N,\infty)}(t)$ intersects the left boundary region.

Why would this be true?

- If $\xi_N(t)$ intersects left boundary region, $\xi_{[-N,\infty)}$ intersects it too
- If $\xi_N(t)$ does not intersect the left boundary region, but has at least one node still alive, this node shields the left boundary region from outside influence. Moreover, no influence can propagate from $-N$. Hence, they need to agree on the left boundary region.

GIF

Proof

Now we're ready to proceed. Recall, we want

- $m_N^2 \max_{1 \le j} \max_{0 \le k < j} \mathbb{P}\left[B_k^N, K_N = j\right] \to 0$

Let's take a closer look at $\mathbb{P}\left[B_k^N, K_N = j\right]$. We will estimate the difference between the temporal average of $f(\xi_N)$ and it's expectation with a triangle inequality. For $k < j$ we have

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

Proof

Let's take the first term in the alternative and try to find a bound for

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

It looks like we could almost apply weak convergence here, but we there's something missing. . .

By weak convergence, as $R \to \infty$, we have

$$R_N^{-1}\int_{kR_N}^{(k+1)R_N} \mathbb{E}\left[f(\xi(t))\right]dt \to \mu(f)$$

Proof

Let's take the first term in the alternative and try to find a bound for

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

It looks like we could almost apply weak convergence here, but we there's something missing. . .

By weak convergence, as $R \to \infty$, we have

$$R_N^{-1}\int_{kR_N}^{(k+1)R_N} \mathbb{E}\left[f(\xi(t))\right]dt \to \mu(f)$$

If we can now find a way to bind what's below, we're in business! Notice that this is an exceedance probability for a random variable. . . Could Chebyshev be of help here?

$$\mathbb{P}\left[\left|R_N^{-1}\int_{kR_N}^{(k+1)R_N} f(\xi(t)) - \mathbb{E}\left[f(\xi(t))\right]dt\right| > \varepsilon/4\right]$$

Proof

Indeed, we can find a bound on variance, and thus, by Chebyshev, on exceedance probability. This has to do with the fact that time correlations of $\xi(t)$ decay exponentially fast:

Theorem (Exponentially decaying correlations)

For any $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ local, there are constants C, γ such that

$$|\mathbb{C}ov\left[f(\xi(t)), f(\xi(s))\right]| \le Ce^{-\gamma|s-t|}$$

Proof

Indeed, we can find a bound on variance, and thus, by Chebyshev, on exceedance probability. This has to do with the fact that time correlations of $\xi(t)$ decay exponentially fast:

Theorem (Exponentially decaying correlations)

For any $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ local, there are constants C, γ such that

$$|\mathbb{C}ov\left[f(\xi(t)), f(\xi(s))\right]| \leq Ce^{-\gamma|s-t|}$$

Now, consider the variance of the expression from the previous slide:

$$\mathbb{E}\left[\left(R_N^{-1} \int_{kR_N}^{(k+1)R_N} f(\xi(t)) - \mathbb{E}\left[f(\xi(t))\right] dt\right)^2\right]$$

Proof

Indeed, we can find a bound on variance, and thus, by Chebyshev, on exceedance probability. This has to do with the fact that time correlations of $\xi(t)$ decay exponentially fast:

Theorem (Exponentially decaying correlations)

For any $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ local, there are constants C, γ such that

$$|\mathbb{C}ov\left[f(\xi(t)), f(\xi(s))\right]| \leq Ce^{-\gamma|s-t|}$$

Now, consider the variance of the expression from the previous slide:

$$\mathbb{E}\left[\left(R_N^{-1} \int_{kR_N}^{(k+1)R_N} f(\xi(t)) - \mathbb{E}\left[f(\xi(t))\right] dt\right)^2\right]$$

- Looks very much like we could get an expression for a covariance from this!

Proof

Indeed, we can find a bound on variance, and thus, by Chebyshev, on exceedance probability. This has to do with the fact that time correlations of $\xi(t)$ decay exponentially fast:

Theorem (Exponentially decaying correlations)

For any $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ local, there are constants C, γ such that

$$|\mathbb{C}ov\left[f(\xi(t)), f(\xi(s))\right]| \leq Ce^{-\gamma|s-t|}$$

Now, consider the variance of the expression from the previous slide:

$$\mathbb{E}\left[\left(R_N^{-1}\int_{kR_N}^{(k+1)R_N} f(\xi(t)) - \mathbb{E}\left[f(\xi(t))\right] dt\right)^2\right]$$

- Looks very much like we could get an expression for a covariance from this!
- Then, Chebyshev will finish out our job for us.

Proof

Recall, we've had:

$$\mathbb{P}\left[B_k^N, K_N = j\right] \leq \mathbb{P}\left[\left|R_N^{-1}\int_{kR_N}^{(k+1)R_N} f(\xi(t))dt - \mu(f)\right| > \varepsilon/2\right] + $$

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

We managed to bind the 1st term using:

- Weak convergence
- Exponentially decaying correlations + Chebyshev

Will this go through for the 2nd term?

- Weak convergence is a no-go. $\xi_N(t)$ converges weakly to δ_0, but $\xi(t)$ converges to μ...
- However, we know that in our intermediate timescale $\xi_N(t)$ is distributed close to μ, so we should be able, in principle, to find a bound for the 2nd term

Proof

Wait, didn't we talk about a situation where $f(\xi_N(t))$ is actually equal to $f(\xi(t))$ before?

$$\left| R_N^{-1} \int_{kR_N}^{(k+1)R_N} f(\xi_N(t)) - f(\xi(t)) dt \right| > \varepsilon/2$$

Recall that

- If ξ_N is wide at t, we have $f(\xi_N(t)) = f(\xi(t))$ (shielding by a wide process).
- We can use h_L to tell whether a process is wide or not (flipping trick).
- ξ_N is alive for all t we're considering. Hence $h_L(\xi_N(t)) = h_L(\xi_{[-N,\infty)}(t))$ (shielding of the left boundary region).

We get that the condition above implies

$$2\|f\| R_N^{-1} \int_{kR_N}^{(k+1)R_N} h_L(\xi_{[-N,\infty)}(t)) + h_L(S\xi_{(-\infty,N]}(t)) dt > \varepsilon/2$$

Proof

By symmetry,

$$\mathbb{P}\left[2\|f\|R_N^{-1}\int_{kR_N}^{(k+1)R_N} h_L(\xi_{[-N,\infty)}(t)) + h_L(S\xi_{(-\infty,N]}(t))dt > \varepsilon/2\right] \leq$$
$$2\mathbb{P}\left[2\|f\|R_N^{-1}\int_{kR_N}^{(k+1)R_N} h_L(\xi_{[-N,\infty)}(t)) > \varepsilon/4\right]$$

This form is something we could use the previous "machinery" (weak convergence + covariance + chebyshev) on, but:

- That method would us to make the time average arbitrarily close to $\mu_{[-N,\infty)}(h_L)$, here we need to make it close to 0...
- Solution: make $\mu_{[-N,\infty)}(h_L)$ close to 0! We can still vary L!

Proof

Set L to be such that

$$\mu_{[-N,\infty)}\left(\{A : A \cap [0,L] = \varnothing\}\right) = \mu_{[0,\infty)}\left(\{A : A \cap [0,L] = \varnothing\}\right) \le \varepsilon/(16\|f\|)$$

Notice how we have used that $\mu_{[-N,\infty)}$ is translation invariant. Defining L in this way makes it independent of N, as we wanted. Surprising, eh?

Proof

We've arrived at

$$\mathbb{P}\left[B_k^N, K_N = j\right] \le \mathbb{P}\left[\Gamma_k^{R_N}\right] + 2\mathbb{P}\left[\tilde{\Gamma}_k^{R_N, L}\right]$$

Where

$$\Gamma_k^{R_N} := \left\{\left|R_N^{-1}\int_{kR_N}^{(k+1)R_N} f(\xi(t))dt - \mu(f)\right| > \varepsilon/2\right\}$$

$$\tilde{\Gamma}_k^{R_N} := \left\{2\|f\|R_N^{-1}\int_{kR_N}^{(k+1)R_N} h_L(\xi_{[-N,\infty)}(t)) > \varepsilon/4\right\}$$

Formally, our previous considerations give us (for appropriate R_N, L)

$$\mathbb{P}\left[B_k^N, K_N = j\right] \le \mathbb{P}\left[\Gamma_k^{R_N}\right] + 2\mathbb{P}\left[\tilde{\Gamma}_k^{R_N, L}\right] \le C/R_N$$

Proof

Recall we needed to find m_N, R_N such that

- $\mathbb{P}[1 \leq K_N \leq m_N] \to 1$
- $m_N^2 \max_{1 \leq j} \max_{0 \leq k < j} \mathbb{P}\left[B_k^N, K_N = j\right] \to 0$

We can bind

- $\max_{1 \leq j} \max_{0 \leq k < j} \mathbb{P}\left[B_k^N, K_N = j\right] \leq C/R_N$

Another lemma shows that $N/\beta_N \to 0$. We can then set

$$m_N = \beta_N^{1/5}/N^{1/5}, \quad R_N = \beta_N^{9/10} N^{1/10}$$

As our solution.

Proof

Originally, the paper also considers shifted versions of f, but I thought this would only introduce more confusing indices. If τ_i, $i \in \mathbb{Z}$ is the shift operator we can also prove (using exactly the same techniques) that

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

We must restrict i to be such that $\Lambda(\tau_i f) \subset [-N + L, N - L]$, though.

Proof

Originally, the paper also considers shifted versions of f, but I thought this would only introduce more confusing indices. If τ_i, $i \in \mathbb{Z}$ is the shift operator we can also prove (using exactly the same techniques) that

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

We must restrict i to be such that $\Lambda(\tau_i f) \subset [-N + L, N - L]$, though.

As a corollary, same statements also holds for spatial means of observable quantities

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

Where $\bar{f} = \frac{1}{\#i} \sum_i \tau_i f$ and $\#i$ is the total number of all is allowed.