involve

a journal of mathematics

Bounds on the artificial phase transition for perfect simulation of hard core Gibbs processes

Mark L. Huber, Elise Villella, Daniel Rozenfeld and Jason Xu

msp

2012 vol. 5, no. 3

INVOLVE 5:3 (2012)
dx.doi.org/10.2140/involve.2012.5.247

# Bounds on the artificial phase transition for perfect simulation of hard core Gibbs processes

Mark L. Huber, Elise Villella, Daniel Rozenfeld and Jason Xu

(Communicated by John C. Wierman)

Repulsive point processes arise in models where competition forces entities to be more spread apart than if placed independently. Simulation of these types of processes can be accomplished using dominated coupling from the past with a running time that depends on the intensity of the number of points. These algorithms usually exhibit what is called an artificial phase transition, where below a critical intensity the algorithm runs in finite expected time, but above the critical intensity the expected number of steps is infinite. Here the artificial phase transition is examined. In particular, an earlier lower bound on this artificial phase transition is improved by including a new type of term in the analysis. In addition, the results of computer experiments to locate the transition are presented.

## 1. Introduction

A spatial point process is a random collection of points in a set $S$. In most applications, $S$ is a continuous space and all of the points are distinct. For instance, the locations of trees in a forest [Møller and Waagepetersen 2007] and the locations of cities in a country [Glass and Tobler 1971] can be modeled using spatial point processes.

One simple spatial point process is the Poisson point process. Suppose that $S$ is a bounded Borel set with positive and finite Lebesgue measure. The basic Poisson point process is the outcome of the following algorithm. First choose a random number of points $N$ according to a Poisson distribution with parameter $\lambda\mu(S)$ (so $\mathbb{P}(N=i)=\exp(-\lambda\mu(S))(\lambda\mu(S))^i/i!$ for nonnegative integers $i$.) Here $\mu$ is Lebesgue measure and $\lambda > 0$ is a parameter of the model. Next, choose points $X_1, \ldots, X_n$ independently and uniformly from the set $S$. The resulting set $\{X_1, \ldots, X_N\}$ is a Poisson point process.

---

*MSC2010:* primary 62M30; secondary 60G55, 60K35.
*Keywords:* spatial point process, dominated coupling from the past, birth-death chain.
This work was conducted as part of a summer research for undergraduates program funded through NSF grant DMS-05-48153, in cooperation with the Claremont Colleges REU.

Since the points are drawn independently, this model fails to capture situations where the locations of points are not independent. In both the forest and cities examples mentioned earlier, the points tend to be farther apart than in the independent situation since the entities involved are competing for space and resources. The points appear to act as particles with the same charge, and so they exhibit repulsion.

There are several ways to account for this repulsion. The *hard core Gibbs process* [Mase et al. 2001] is a Poisson point process conditioned on the event that none of the points lie within distance $R$ of each other. In other words, each point is surrounded by a hard core of radius $R/2$. The cores are "hard" in the sense that the cores are not allowed to overlap. Here $R$ is a parameter of the model.

In frequentist approaches, this model can be used to construct maximum likelihood estimators for $R$ and $\lambda$. The values of these estimators can be approximated by methods which use random draws of the point process from the model. See, for example, [Geyer and Møller 1994; Geyer 1999; Møller and Waagepetersen 2004] for details.

In Bayesian approaches, this model (together with a prior on $\lambda$ and $R$) can be used to build a posterior for the parameters. This posterior is quite complex, and depends on a normalizing constant (also known as partition function) that is difficult to compute exactly. The auxiliary variable method of Møller et al. [2006] can be used to create a Markov chain for these problems: this Markov chain also requires the ability to draw random variates from the model in question.

***Spatial birth and death chains.*** Preston [1975] created a coupled pair of jump processes $(X_t, Y_t)$ where the stationary distribution of $Y_t$ is a Poisson point process, and the stationary distribution of $X_t$ is the target process. In a jump process, the state stays the same until abruptly jumping to a new state (these jumps are called *events*). The time until the next jump is an exponential random variable whose rate depends only on the current state. Conditioned on this rate, the exponential is independent of all prior history of the process. For $Y_t$, a *birth* is an addition of a point to the process, and occurs at rate equal to $\lambda\mu(S)$. If a birth event occurs, the point added is chosen uniformly from $S$ (again this choice is independent of the prior history of the process.) Each point when born is given a time of death that is the current time plus an exponential random variable with mean 1. This exponential is once more independent of the prior history of the process. At time of death, the point is removed from the process.

For a jump process $A_t$, let

$$A_{t^-} = \bigcap_{\epsilon>0} \bigcap_{t-\epsilon<t'<t} A_{t'}$$

be the state of the process immediately before time $t$. To use Preston's method for the hard core Gibbs process, suppose that the point $v$ is born at time $t$ in the $Y$

process. Then $v$ is added to the $X_t$ state if and only if it is not within distance $R$ of a point in $X_{t^-}$. So births are always added to the $Y$ process, but only sometimes to the $X$ process in order to maintain the hard core property.

If a point $w \in Y_{t-}$ dies, at time $t$ it is removed from the $Y$ process. If $w$ is also in $X_{t-}$ it is also removed from the $X$ process at time $t$. With this coupling,

$$X_{t'} \subseteq Y_{t'} \implies X_t \subseteq Y_t$$

for all $t' < t$, so the $Y$ process is referred to as the *dominating process*.

Preston's approach yields a jump process whose limiting distribution of $X_t$ is the hard core Gibbs process, but $X_t$ will never exactly be in the correct distribution. Ferrari, Fernández, and Garcia [2002] developed a method for drawing samples exactly from the desired distribution using a clan of ancestors approach. In turn, Kendall and Møller [2000] developed a much faster algorithm, *dominated coupling from the past* (DCFTP), which can be used to sample from a variety of distributions that include the hard core Gibbs process.

Previous analysis showed that when using the standard Euclidean distance, the DCFTP method was provably fast when $\lambda < 1/(\pi R^2)$ [Huber 2012]. In this work we build upon this analysis, providing a wider set of conditions on $\lambda$ and $R$ for the DCFTP method to run quickly. The original argument used a term depending on the number of points in the configuration, while the new method uses the number of points as well as the area spanned by these points. This extra area term is what leads to the stronger proof. For ease of exposition we use the Euclidean metric to measure the distance between points and only operate in $\mathbb{R}^2$ throughout this work; we simply note that the same argument can easily be applied to any metric and to problems in higher dimensions.

The remainder of the work is organized as follows. Section 2 gives our new result: improved sufficient conditions on the parameters of the model for dominated coupling from the past to operate quickly. Section 3 gives computer results to complement the theoretical results of the previous section, and we close with our conclusions.

## 2. Bounding the running time of DCFTP

The time necessary to run DCFTP is related to the *clan of descendants* (cod) of a point $v$, defined as follows. For any point $v \in Y_0$, couple another point process $C_t(v)$ to $Y_t$ as follows. Let $C_0(v) = \{v\}$. If a point $w$ is born to $Y_{t^-}$ at time $t$, add $w$ to $C_t$ if and only if $w$ is within distance $R$ of a point in $C_{t^-}$. If a point $w'$ dies in the $Y$ process at time $t$, and is also in the C process, remove it from $C_t$ as well.

Then the cod of $v$ is

$$C(v) = \bigcup_{t \geq 0} C_t(v).$$

The clan of ancestors in [Ferrari et al. 2002] is the time reversal of the cod, so they have the same size. In addition, the expected running time of DCFTP is bounded by a constant times the expected size of the cod. If there is a chance that the cod grows indefinitely, DCFTP has the same chance of taking forever to generate a sample, so the algorithm is only useful when the cod is finite with probability 1.

To bound the size of the cod, we wish to show that $\#C_t$ converges to 0 (so that $C_t = \varnothing$) with probability 1 after a finite number of births and deaths that affect the cod. In particular,

**Theorem.** *For $\lambda < [8/(3\sqrt{3}+4\pi)]/R^2$, the expected number of births and deaths that affect the cod is bounded above by*

$$\left(\frac{8/(3\sqrt{3}+4\pi)}{R^2} - \lambda\right)^{-1}.$$

As noted in Section 1, a similar previous result in [Huber 2012] had a constant of $1/\pi \approx 0.3183$ in front of the $R^{-2}$ factor, while the new result has $8/(3\sqrt{3}+4\pi) \approx 0.4503$. Hence this result proves the efficacy of the DCFTP method (and mixing time of the chain) over values of $\lambda$ that are 41% larger than previously known.

***Avoiding boundary effects.*** In order to avoid having to worry about boundary effects arising from finite $S$, we first build another point process that dominates $C_t(v)$. As with the regular process, start with $C_0^+(v) = \{v\}$. Let $S(C_t^+(v), R)$ be all points within distance $R$ of a point in $C_t^+(v)$. Then births in $S(C_t^+(v), R)$ will occur at rate $\lambda \cdot \mu(S(C_t^+(v), R))$. Points in $C_t^+$ die at rate 1. Births and deaths in $S$ can be coupled to the births and deaths in $Y_t$, but there might be extra points in $C_t^+$ that were born outside of $S$. Therefore, $C_t(v) \subseteq C_t^+(v)$, and to show that $\#C_t(v)$ converges to zero, it suffices to show $\#C_t^+(v)$ converges to zero.

***Useful facts.*** Before proving the Theorem, we show some facts that will be useful. We are only interested in how $C_t^+$ changes with births and deaths. Hence let $t_i$ denote the time of the $i$-th event that is either a death of a point in the cod, or the proposed birth of a point within distance $R$ of the cod. Let $D_i = C_{t_i}^+$, so $D_i$ represents a superset of the cod after $i$ such events have occurred. Let $\#D_i$ denote the number of points in this set.

For a configuration $x$, let $A(x)$ denote the Lebesgue measure of the region within distance $R$ of at least one point in $x$. In particular, $A(D_i)$ is the measure of the area of the region within distance $R$ of points in the cod. So $A(D_i)$ is proportional to the rate at which births occur that increase $\#D_i$ by 1. Our first lemma limits the average area that is added when such a birth occurs.

**Lemma 1.** $\mathbb{E}[A(D_{i+1}) - A(D_i) \mid$ *a birth occurs at time* $t_{i+1}] \le R^2 3\sqrt{3}/4$.

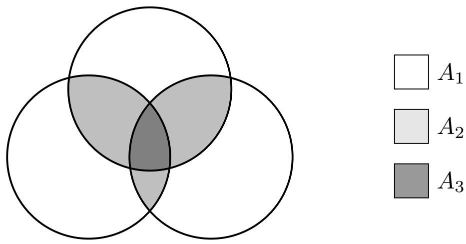


**Figure 1.** For circles of radius $R$, $3\pi R^2 = A_1 + 2A_2 + 3A_3$.

*Proof.* Let $w$ be a proposed birth point. Then in order to add to the clan of descendants, $w$ must be within distance $R$ of a point $v$ of $D_i$. The area of the new setup does not increase by $\pi R^2$, however, since only the region within $R$ of $w$ and not within $R$ of $v$ can be added area. Because $w$ is conditioned to lie within distance $R$ of $v$, the distance between centers is a random variable with density $f_r(a) = (2a/R^2) \cdot \mathbf{1}(0 \le a \le R)$.[1] Hence, the expected area added can be written as

$$\begin{aligned}\mathbb{E}[A(D_{i+1}) - A(D_i) \,|\, \text{birth}] &\le \int_0^R \frac{2a}{R^2}\left[\pi R^2 - 4\int_{a/2}^R \sqrt{R^2 - x^2}\, dx\right] da \\ &= R^2 3\sqrt{3}/4.\end{aligned}$$

This is an upper bound on the expected value of $A(D_{i+1}) - A(D_i)$ because $w$ might be within distance $R$ of other points in $D_i$ as well, which would reduce the added area. □

The last lemma gives an upper bound on the area added when a birth occurs. The next lemma gives a lower bound on the area removed when a death occurs.

**Lemma 2.**

$$\mathbb{E}[A(D_{i+1}) - A(D_i) \,|\, \textit{ a death occurs at time } t_{i+1}] \ge [2A(D_i)/\#D_i] - \pi R^2.$$

*Proof.* Let $A_k$ denote the area of the region that is within distance $R$ of exactly $k$ points of $D_i$. Then (see Figure 1)

$$\pi R^2 \# D_i = A_1 + 2A_2 + 3A_3 + \cdots + (\#D_i)A_{\#D_i},$$

and $A(D_i) = A_1 + A_2 + A_3 + \cdots + A_{\#D_i}$. Therefore

$$2A(D_i) - \pi R^2 \# D_i = A_1 - A_3 - 2A_4 - \cdots - (\#D_i - 2)A_{\#D_i} \le A_1.$$

If the points in $D_i$ are labeled $1, 2, \ldots, \#D_i$, then $A_1 = a_1 + a_2 + \cdots + a_{\#D_1}$, where $a_k$ is the area of the region within distance $R$ of point $i$ and no other points.

[1] We use $\mathbf{1}(P(a))$ for the indicator function of $P(a)$, defined as 1 if $P(a)$ is true and as 0 otherwise.

When a death occurs, every point in $\#D_i$ is equally likely to be chosen to be removed, so the average area removed is

$$\frac{1}{\#D_i}a_1 + \cdots + \frac{1}{\#D_i}a_{\#D_i} = \frac{1}{\#D_i}A_1 \geq \frac{2A(D_i)}{\#D_i} - \pi R^2. \qquad \square$$

*Proof of the Theorem.* For a configuration $x$, let $\phi(x) = A(x) + c \cdot \#x$, where $c > 0$ is a constant to be chosen later. Note that $\phi(x)$ is positive unless $x$ is the empty configuration, in which case it equals 0. Let $\tau = \inf\{i : D_i = \varnothing\}$. Using $a \wedge b$ to denote the minimum of $a$ and $b$, we shall show that $\phi(D_{i\wedge\tau}) + (i \wedge \tau)\delta$ is a supermartingale with

$$\delta = \frac{2 - \lambda R^2(3\sqrt{3}/4)}{1+\lambda}.$$

The rest of the result then follows as a consequence of the optional sampling theorem (OST). See Chapter 5 of [Durrett 2010] for a description of supermartingales and the OST.

When $i \geq \tau$, $\phi(D_{i\wedge\tau}) + (i \wedge \tau)\delta$ is a constant, and so trivially is a supermartingale.

When $i < \tau$, $\phi(D_{i+1})$ either grows when a birth occurs in the cod, or shrinks when a death occurs. First consider how $\#D_i$ changes. Births occur at rate $\lambda A(D_i)$, and deaths at rate $\#D_i$. Hence the probability that an event that changes $\#D_i$ is a birth is $A(D_i)/(A(D_i) + \#D_i)$, with the rest of the probability going towards deaths. So

$$\begin{aligned}\mathbb{E}[\#D_{i+1} - \#D_i \mid \phi(D_i)] &= \mathbb{E}\big[\mathbb{E}[\#D_{i+1} - \#D_i \mid D_i] \bigm| \phi(D_i)\big] \\ &\leq \mathbb{E}\left[\mathbf{1}(i<\tau)\left(\frac{\lambda A(D_i)}{A(D_i)+\#D_i} - \frac{\#D_i}{A(D_i)+\#D_i}\right) \middle| \phi(D_i)\right].\end{aligned}$$

(The analysis in [Huber 2012] only considered this term in $\phi$, which is why the result is weaker than what is given here.)

From our first lemma, a birth increases (on average) the area covered by the cod by at most $R^2 3\sqrt{3}/4$. Our second lemma provides a lower bound on the average area removed when a death occurs. Combining these results yields

$$\begin{aligned}&\mathbb{E}[A(D_{i+1}) - A(D_i) \mid \phi(D_i)] = \mathbb{E}\big[\mathbb{E}[A(D_{i+1}) - A(D_i) \mid D_i] \bigm| \phi(D_i)\big] \\ &\leq \mathbb{E}\left[\mathbf{1}(i<\tau)\left(\frac{\lambda A(D_i)}{A(D_i)+\#D_i}R^2\frac{3\sqrt{3}}{4} - \frac{\#D_i}{A(D_i)+\#D_i}\left(\frac{2A(D_i)}{\#D_i} - \pi R^2\right)\right) \middle| \phi(D_i)\right].\end{aligned}$$

Note that $\mathbf{1}(i < \tau)$ is measurable with respect to $\phi(D_i)$, so bringing that out front and adding the inequalities gives

$$\begin{aligned}&\mathbb{E}[\phi(D_{i+1}) - \phi(D_i) \mid \phi(D_i)] \\ &\qquad \leq \mathbf{1}(i<\tau)\,\mathbb{E}\left[\frac{A(D_i)(\lambda((R^2 3\sqrt{3}/4) + c) - 2) + \#D_i(\pi R^2 - c)}{A(D_i)+\#D_i} \middle| \phi(D_i)\right].\end{aligned}$$

Now $c$ can be set to

$$c = \frac{\pi R^2 + 2 - \lambda R^2(3\sqrt{3}/4)}{1+\lambda},$$

so that

$$\mathbb{E}\left[\phi(D_{i+1}) - \phi(D_i) \mid \phi(D_i)\right] \le \mathbf{1}(i < \tau)\mathbb{E}\left[\frac{A(D_i)(-\delta) + \#D_i(-\delta)}{A(D_i) + \#D_i} \,\middle|\, \phi(D_i)\right]$$
$$= -\delta\mathbf{1}(i < \tau).$$

Hence $\phi(D_{i\wedge\tau}) + (i \wedge \tau)\delta$ is a supermartingale. $\square$

## 3. Experimental results

This theoretical result increases the known lower bound for the value of $\lambda$ where the clan of descendants is finite, but this is still just a lower bound. Computer experiments can estimate this critical value of $\lambda$ more precisely.

For the estimates in this section, the following protocol was used. We began a clan of descendants superset $C^+(v)$ from a single point, and recorded whether the clan died out or reached a size of 750. This was repeated 200 times, and used to estimate the probability that the clan dies out for a given value of $\lambda$. The results indicate that somewhere in $[0.625, 0.626]$, the probability begins to drop from 1 down towards 0 (see Figure 2 for how the extinction probability changes with $\lambda$).

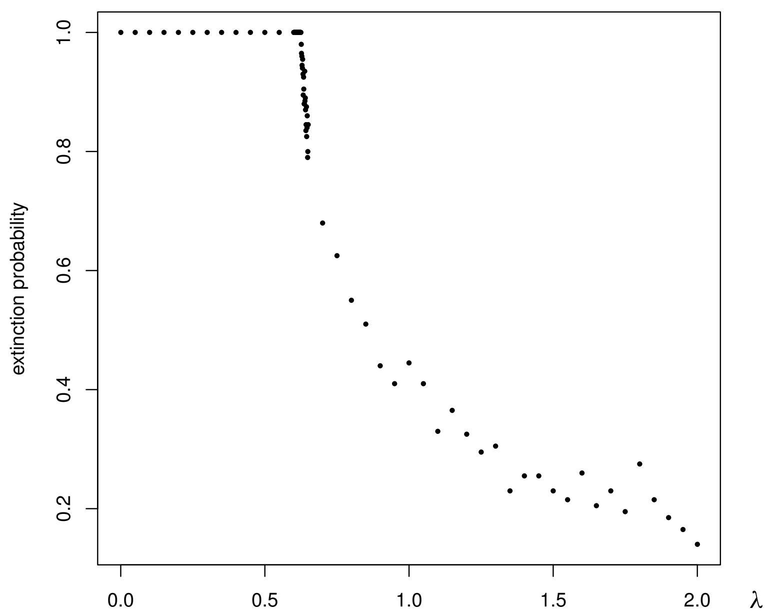


**Figure 2.** Estimate of extinction probability using 200 trials. The maximum cod size is 750 points.

This indicates that while the new 0.4503 theoretical result is an improvement over the old result of 0.3183, there is still work to be done to reach the true value. Increasing the ceiling size from 750 to 1500 did not alter the results within experimental error.

In short, by including a term for the area covered by the points in the potential function, a stronger theoretical lower bound on the artificial phase transition for dominated coupling from the past applied to the hard core gas model has been found. This method appears to be very general and should apply to a wide variety of repulsive processes.

## References


[Durrett 2010] R. Durrett, *Probability: theory and examples*, 4th ed., Cambridge University Press, 2010. MR 2011e:60001 Zbl 1202.60001

[Ferrari et al. 2002] P. A. Ferrari, R. Fernández, and N. L. Garcia, "Perfect simulation for interacting point processes, loss networks and Ising models", *Stochastic Process. Appl.* **102**:1 (2002), 63–88. MR 2003j:60140

[Geyer 1999] C. Geyer, "Likelihood inference for spatial point processes", pp. 79–140 in *Stochastic geometry* (Toulouse, 1996), edited by O. E. Barndorff-Nielsen et al., Monogr. Statist. Appl. Probab. **80**, Chapman & Hall/CRC, Boca Raton, FL, 1999. MR 1673118

[Geyer and Møller 1994] C. J. Geyer and J. Møller, "Simulation procedures and likelihood inference for spatial point processes", *Scand. J. Statist.* **21**:4 (1994), 359–373. MR 95i:62082

[Glass and Tobler 1971] L. Glass and W. R. Tobler, "Uniform distribution of objects in a homogeneous field: Cities on a plain", *Nature* **233** (1971), 67–68.

[Huber 2012] M. L. Huber, "Spatial birth-death-swap chains", *Bernoulli* **18**:3 (2012), 1031–1041. Zbl 06064472

[Kendall and Møller 2000] W. S. Kendall and J. Møller, "Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes", *Adv. in Appl. Probab.* **32**:3 (2000), 844–865. MR 2001h:62176

[Mase et al. 2001] S. Mase, J. Møller, D. Stoyan, R. P. Waagepetersen, and G. Döge, "Packing, densities and simulated tempering for hard core Gibbs point processes", *Ann. Inst. Statist. Math.* **53**:4 (2001), 661–680. MR 2003b:60067 Zbl 1086.60512

[Møller and Waagepetersen 2004] J. Møller and R. P. Waagepetersen, *Statistical inference and simulation for spatial point processes*, Monographs on Statistics and Applied Probability **100**, Chapman & Hall/CRC, Boca Raton, FL, 2004. MR 2004h:62003 Zbl 1044.62101

[Møller and Waagepetersen 2007] J. Møller and R. P. Waagepetersen, "Modern statistics for spatial point processes", *Scand. J. Statist.* **34**:4 (2007), 643–684. MR 2009h:60091 Zbl 1157.62067

[Møller et al. 2006] J. Møller, A. N. Pettitt, R. Reeves, and K. K. Berthelsen, "An efficient Markov chain Monte Carlo method for distributions with intractable normalising constants", *Biometrika* **93**:2 (2006), 451–458. MR 2278096

[Preston 1975] C. Preston, "Spatial birth-and-death processes", *Bull. Inst. Internat. Statist.* **46**:2 (1975), 371–391. MR 57 #14170 Zbl 0379.60082


Received: 2010-10-15 Revised: 2012-01-04 Accepted: 2012-01-13

mhuber@cmc.edu
*Department of Mathematical Sciences, Claremont McKenna College, 850 Columbia Avenue, Claremont, CA 91711, United States*

elisemccall@gmail.com
*Massachusetts Institute of Technology, 320 Memorial Drive, Cambridge, MA 02139, United States*

Daniel_J_Rozenfeld@hmc.edu
*Harvey Mudd College, 340 East Foothill Boulevard, Claremont, CA 91711, United States*

qxu@email.arizona.edu
*University of Arizona, 1021 W. Green Pebble Drive, Tucson, AZ 85755, United States*

# involve

msp.org/involve





See inside back cover or msp.org/involve for submission instructions. The subscription price for 2012 is US $105/year for the electronic version, and $145/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve
2012 vol. 5 no. 3



1944-4176(2012)5:3;1-A