

University of Groningen

## Rare Events in Random Geometric Graphs

Hirsch, Christian; Moka, Sarat B.; Taimre, Thomas; Kroese, Dirk P.

*Published in:*  
Methodology and Computing in Applied Probability

*DOI:*  
[10.1007/s11009-021-09857-7](https://doi.org/10.1007/s11009-021-09857-7)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2022

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Hirsch, C., Moka, S. B., Taimre, T., & Kroese, D. P. (2022). Rare Events in Random Geometric Graphs. *Methodology and Computing in Applied Probability*, 24, 1367–1383. <https://doi.org/10.1007/s11009-021-09857-7>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



# Rare Events in Random Geometric Graphs

Christian Hirsch<sup>1</sup> · Sarat B. Moka<sup>2</sup> · Thomas Taimre<sup>2</sup> · Dirk P. Kroese<sup>2</sup>

Received: 9 July 2020 / Revised: 3 February 2021 / Accepted: 9 March 2021 /  
Published online: 4 May 2021  
© The Author(s) 2021

## Abstract

This work introduces and compares approaches for estimating rare-event probabilities related to the number of edges in the random geometric graph on a Poisson point process. In the one-dimensional setting, we derive closed-form expressions for a variety of conditional probabilities related to the number of edges in the random geometric graph and develop conditional Monte Carlo algorithms for estimating rare-event probabilities on this basis. We prove rigorously a reduction in variance when compared to the crude Monte Carlo estimators and illustrate the magnitude of the improvements in a simulation study. In higher dimensions, we use conditional Monte Carlo to remove the fluctuations in the estimator coming from the randomness in the Poisson number of nodes. Finally, building on conceptual insights from large-deviations theory, we illustrate that importance sampling using a Gibbsian point process can further substantially reduce the estimation variance.

**Keywords** Rare event · Random geometric graph · Conditional Monte Carlo · Strauss process

**Mathematics Subject Classification (2010)** 60K35 · 60F10 · 82C22

## 1 Introduction

In this paper, we focus on rare events associated with the number of edges in the *Gilbert graph*  $G(X)$  on a homogeneous Poisson point process  $X = \{X_i\}_{i \geq 1}$  with intensity  $\lambda > 0$  in

---

✉ Christian Hirsch  
c.p.hirsch@rug.nl

Sarat B. Moka  
s.babumoka@uq.edu.au

Thomas Taimre  
t.taimre@uq.edu.au

Dirk P. Kroese  
kroese@maths.uq.edu.au

<sup>1</sup> Bernoulli Institute, University of Groningen, Nijenborgh 9, 9747 AG Groningen, The Netherlands

<sup>2</sup> School of Mathematics and Physics, The University of Queensland, Brisbane, Australia

$\mathbb{R}^d$ , for  $d \geq 1$ . Consequently, the nodes of  $G(X)$  are the points of  $X$  and there is an edge between  $X_i, X_j \in X$  if  $\|X_i - X_j\| \leq 1$ , where the upper bound 1 is the *threshold* of the Gilbert graph and  $\|\cdot\|$  denotes the Euclidean norm. Our goal is to analyze the probability of the rare event that the number of edges in a bounded sampling window  $W \subset \mathbb{R}^d$  deviates considerably from its expected value. More succinctly, we ask:

*What is the probability that the Gilbert graph has at least twice its expected number of edges?*

*What is the probability that the Gilbert graph has at most half its expected number of edges?*

These seemingly innocuous questions have an intriguing connection to large deviations of heavy-tailed sums. Indeed, suppose that  $\{W_i\}$  is an equal-volume partition of  $W$  such that the diameter of each  $W_i$  is at most 1. Then, letting  $Z_i := |X \cap W_i|$  denote the number of Poisson points in  $W_i$ , the edges entirely inside  $W_i$  contribute  $Z_i(Z_i - 1)/2$  to the total edge count. Since  $Z_i$  follows a Poisson distribution, the tail probability  $\mathbb{P}(Z_i \geq n)$  is of the order  $\exp(-cn \log n)$  for some constant  $c > 0$ . Hence, the tails of  $Z_i^2$  are of the order  $\exp(-c\sqrt{n} \log n/2)$  and therefore  $Z_i^2$  does not have exponential moments. This is critical to note, because it means that the problem at hand is tightly related to large deviations of heavy-tailed sums, where large deviations typically come from extreme realizations of the largest summand (Asmussen and Kroese 2006; Blanchet and Glynn 2008; Rojas-Nandayapa 2013).

The analysis and simulations in this paper rely on two methods: *conditional Monte Carlo (MC)* and *importance sampling*; see Chapters 9.4 and 9.7 of Kroese et al. (2013), respectively. Importance sampling techniques have been in use for the analysis of rare events in the setting of the Erdős–Rényi graph (Bhamidi et al. 2015), which is different from the Gilbert graph considered here. Other than that, there has been very little literature on the topic. Note that the same analysis can be extended to Gilbert graphs with the threshold not equal to 1 by modifying the size of the window  $W$  and the intensity  $\lambda$  of the Poisson point process.

Seen in a larger context, our work is the first to provide specific simulation-based tools for the estimation of rare events in the Gilbert graph, which is the simplest example of a spatial random network. This is of particular relevance when taking into account the central place that such networks have in models in materials science and wireless networks (Baccelli and Błaszczyszyn 2009; Stenzel et al. 2014). When these models form the basis for applications in security-sensitive contexts, it becomes essential to provide estimates for rare-event probabilities. Indeed, when introducing a new technology, such as for instance device-to-device networks, it is not enough to know that the system works well on average. At the same time, the probabilities for catastrophic events should be estimated sufficiently precisely such that they can be deemed negligible. A prototypical application scenario based on measurement data is outlined by Keeler et al. (2018).

We also mention that it would be possible to derive analytical approximations for rare-event probabilities, for instance by assuming that the quantity of interest is approximately Poisson or normally distributed. The appeal of such an approach would lie in its simplicity, as it only requires first- and second-moment information. However, in the analytical approximations, it is difficult to quantify the estimation error, whereas the simulation-based methods outlined in the present paper are unbiased, and therefore do not suffer from this weakness, provided the number of simulations is large enough.

The rest of the presentation is organized in three parts. First, in Section 2, we explore the potential of conditional Monte Carlo in a one-dimensional setting, where we can frequently derive explicit closed-form expressions. Surprisingly, when moving to the far end

of the upper tails, it is possible to avoid simulation altogether, as we derive a fully analytic representation. Then, in Section 3, we move to higher dimensions. Here, we apply conditional MC to remove the randomness coming from the random number of nodes. Finally, in Section 4, we present a further refinement of the conditional MC estimator, by combining it with importance sampling using a Strauss-type Gibbs process. To ease notation, we henceforth identify the Gilbert graph with its edge set, so that  $|G(X \cap W)|$  yields the number of edges in the Gilbert graph on  $X \cap W$ .

## 2 Conditional MC in Dimension 1

In this section, we consider the one-dimensional setting. More precisely, we consider a line segment  $W = [0, w]$  as sampling window. In Section 2.1, we describe a specific conditional MC scheme that leads to estimators for the rare-event probability of the number of edges being small. In a simulation study in Section 2.2, we show that these new estimators are substantially better than a crude MC approach.

Then, Section 2.3 discusses rare events corresponding to the number of missing edges, i.e., the number of point-pairs that are not connected by an edge. The analysis is motivated from the observation that the Erdős–Rényi graph with edge probability  $p \in [0, 1]$  exhibits a striking duality with its complement. Specifically, the missing edges of this graph again form an Erdős–Rényi graph but now with probability  $1 - p$ . In Section 2.3, we point out that in the Gilbert graph such a duality is much more involved. We still elucidate how to compute the probability of observing no missing edges or precisely one missing edge.

In this section, we assume that the points  $\{X_i\}_{i \geq 1}$  of the Poisson point process on  $[0, \infty)$  are ordered according to their occurrence; that is,  $X_i \leq X_j$ , whenever  $i \leq j$ .

### 2.1 Few Edges

Henceforth, let

$$E_{\leq k} := \{|G(X \cap [0, w])| \leq k\}$$

denote the event that the number of edges in the Gilbert graph on  $X \cap [0, w]$  is at most  $k \geq 0$ . For fixed  $k$  and large  $w$ , the probability

$$p_{\leq k} := \mathbb{P}(E_{\leq k})$$

becomes small, and we discuss how to use both the natural ordering on the real half-line and the independence property of the Poisson point process to derive a refined estimator.

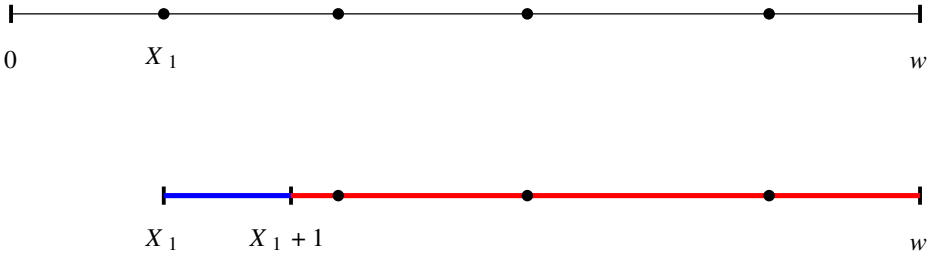
We focus only on the cases  $k = 0, 1$ . In principle, the methods could be extended to cover estimation of probabilities of the form  $p_{\leq k}$  for  $k \geq 2$ . However, for large values of  $k$  the combinatorial analysis becomes quickly highly involved; see Remark 2 for more details.

#### 2.1.1 No Edges

To begin with, let  $k = 0$ . That is, we analyze the probability that all vertices in  $X \cap [0, w]$  are isolated in the sense that their vertex degree is 0. The key idea for approaching this probability is to note that  $E_0 := E_{\leq 0}$  occurs if and only if  $X \cap [X_1, (X_1 + 1) \wedge w] = \emptyset$  and the Gilbert graph restricted to  $X \cap [X_1 + 1, w]$  does not contain edges; see Fig. 1. Here, we adhere to the convention that  $[a, b] = \emptyset$  if  $a > b$ .

According to the Palm theory for one-dimensional Poisson point processes, the process

$$X^{(1)} := (X - X_1) \cap [1, \infty)$$



**Fig. 1** For  $G(X \cap [0, w]) = \emptyset$ , the blue interval may not contain points of  $X$  and the Gilbert graph restricted to the red interval may not contain edges

again forms a homogeneous Poisson point process, which is independent of  $X_1$ ; see Last and Penrose (2017, Theorem 7.2). In particular, writing

$$\mathcal{F}_1 := \sigma(X_1, X^{(1)})$$

for the  $\sigma$ -algebra generated by  $X_1$  and  $X^{(1)}$  allows for a partial computation of  $p_0 := p_{\leq 0}$  via conditional MC (Kroese et al. 2013, Chapter 9.4). More precisely, when computing  $\mathbb{P}(E_0 | \mathcal{F}_1)$  we explicitly throw away the information from the configuration of  $X$  inside the interval  $[X_1, X_1 + 1]$ .

**Theorem 1** (No edges) *Suppose that  $w \geq 1$ . Then,*

$$\mathbb{P}(E_0 | \mathcal{F}_1) = e^{-\lambda((w-X_1) \wedge 1)} \mathbb{1}\{G(X^{(1)} \cap [1, w - X_1]) = \emptyset\} \quad \text{almost surely.} \tag{1}$$

*Proof* First,  $X_1$  is isolated if there are no further vertices in the interval  $[X_1, (X_1 + 1) \wedge w]$ . Moreover, after conditioning on  $X_1$ , the remaining vertices form a Poisson point process in  $[(X_1 + 1) \wedge w, w]$ ; see Last and Penrose (2017, Theorem 7.2). Therefore,

$$\begin{aligned} \mathbb{P}(E_0 | \mathcal{F}_1) &= \mathbb{P}(X \cap (X_1, (X_1 + 1) \wedge w) = \emptyset | \mathcal{F}_1) \mathbb{1}\{G(X^{(1)} \cap [1, w - X_1]) = \emptyset\} \\ &= e^{-\lambda(((X_1+1) \wedge w) - X_1)} \mathbb{1}\{G(X^{(1)} \cap [1, w - X_1]) = \emptyset\} \end{aligned}$$

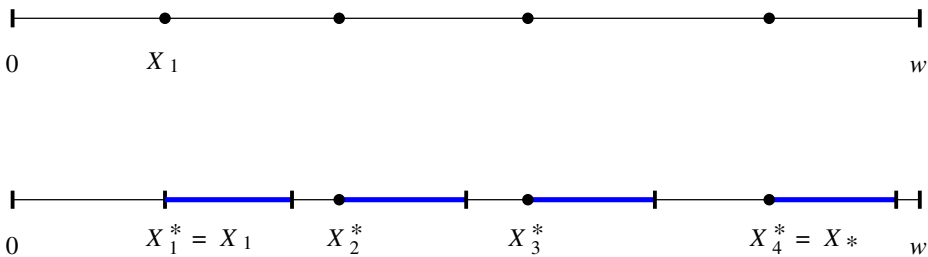
as asserted. □

In other words, invoking the Rao–Blackwell theorem (Billingsley 1995), Theorem 1 indicates that the right-hand side of identity (1) is an attractive candidate for estimating  $p_0$  via conditional Monte Carlo. The Rao–Blackwell theorem is a powerful tool in situations where  $p_0$  is not available in closed form. Note that elementary properties of the conditional expectation imply that the conditional MC estimator  $\mathbb{P}(E_0 | \mathcal{F}_1)$  is unbiased and exhibits smaller variance than the crude MC estimator  $\mathbb{1}\{E_0\}$ .

Moreover, the right-hand side of identity (1) features another indicator of an isolation event. Hence, it becomes highly attractive to refine the estimator further by proceeding iteratively. To make this precise, we define an increasing sequence  $X_1^* \leq X_2^* \leq \dots$  of points of  $X$  recursively as follows. First,  $X_1^* = X_1$  denotes the left-most point of  $X$ . Next, once  $X_m^*$  is available,

$$X_{m+1}^* := \inf\{X_i \in X : X_i \geq X_m^* + 1\}$$

denotes the first point of  $X$  to the right of  $X_m^* + 1$ . Then, the event  $E_0$  occurs if none of the intervals  $[X_i^*, X_i^* + 1]$  contains points from  $X$ ; see Fig. 2.



**Fig. 2** For  $G(X \cap [0, w]) = \emptyset$ , the blue intervals may not contain any points

Of particular interest is the last index

$$I_* := \sup\{i \geq 1 : X_i^* \leq w\},$$

where  $X_{i_*}^*$  remains inside  $[0, w]$ , together with the associated point

$$X_* := X_{i_*}^*.$$

If  $X_1 > w$ , we set  $I_* = 0$  and  $X_* = w$ . Let

$$\mathcal{F}^* := \sigma(X_1^*, X_2^*, \dots),$$

be the  $\sigma$ -algebra generated by  $\{X_i^* : i \geq 1\}$ .

**Theorem 2** (No edges – iterated) *Suppose that  $w \geq 1$ . Then,*

$$\mathbb{P}(E_0 | \mathcal{F}^*) = e^{-\lambda((I_*-1)_+ + (w-X_*) \wedge 1)} \text{ almost surely.}$$

*Proof* To prove the claim, we first define the shifted process

$$X^{(m)} := (X - X_m^*) \cap [1, \infty)$$

and write

$$\mathcal{F}_m := \sigma(X_1^*, \dots, X_m^*, X^{(m)})$$

for the  $\sigma$ -algebra generated by  $X_1^*, \dots, X_m^*$  and  $X^{(m)}$ . Observe that  $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \dots \supseteq \mathcal{F}^*$ . In particular, by the tower property of conditional expectation,

$$\begin{aligned} \mathbb{P}(E_0 | \mathcal{F}^*) &= \sum_{m \geq 0} \mathbb{E}[\mathbb{1}\{I_* = m\} \mathbb{1}\{E_0\} | \mathcal{F}^*] = \sum_{m \geq 0} \mathbb{1}\{I_* = m\} \mathbb{P}(E_0 | \mathcal{F}^*) \\ &= \sum_{m \geq 0} \mathbb{1}\{I_* = m\} \mathbb{E}[\mathbb{P}(E_0 | \mathcal{F}_m) | \mathcal{F}^*] \mathbb{1}\{I_* \geq m\}. \end{aligned}$$

Hence, it suffices to show that for every  $m \geq 0$ ,

$$\mathbb{P}(E_0 | \mathcal{F}_m) \mathbb{1}\{I_* \geq m\} = e^{-\lambda((m-1)_+ + (w-X_m^*) \wedge 1)} \mathbb{1}\{G(X^{(m)} \cap [1, w - X_m^*]) = \emptyset\} \mathbb{1}\{I_* \geq m\}, \tag{2}$$

because  $X^{(I_*)} \cap [1, w - X_*] = \emptyset$ . To achieve this goal, we proceed by induction on  $m$ . For  $m = 0$  and  $m = 1$  we are in the setting of Theorem 1. To pass from  $m$  to  $m + 1$ , the induction hypothesis yields that

$$\begin{aligned} \mathbb{P}(E_0 | \mathcal{F}_{m+1}) \mathbb{1}\{I_* \geq m + 1\} &= \mathbb{E}[\mathbb{P}(E_0 | \mathcal{F}_m) | \mathcal{F}_{m+1}] \mathbb{1}\{I_* \geq m + 1\} \\ &= e^{-\lambda m} \mathbb{P}(G(X^{(m)} \cap [1, w - X_m^*]) = \emptyset | \mathcal{F}_{m+1}) \mathbb{1}\{I_* \geq m + 1\}. \end{aligned} \tag{3}$$

Since  $X_{m+1}^*$  is the first point of  $X$  after  $X_m^* + 1$ , by applying Theorem 1 to the Poisson point process  $X^{(m)}$ , we obtain

$$\mathbb{P}(G(X^{(m)} \cap [1, w - X_m^*]) = \emptyset | \mathcal{F}_{m+1}) = e^{-\lambda((w-X_m^*) \wedge 1)} \mathbb{1}\{G(X^{(m+1)} \cap [1, w - X_{m+1}^*]) = \emptyset\} \tag{4}$$

Combining (3) and (4) yields the assertion. □

The conditional MC estimator from Theorem 2 leads to Algorithm 1. Here,  $\text{Exp}(\lambda)$  is an exponential random variable with parameter  $\lambda$  that is independent of everything else.

---

**Algorithm 1** Conditional MC estimator for  $p_0$ .

---

```

input: Number  $N \geq 1$  of MC runs and Poisson intensity  $\lambda > 0$ .
output: Conditional MC estimator of  $p_0$ .
1 sum  $\leftarrow$  0
2 for  $i \leq N$  do
3   Draw  $Z$  from  $\text{Exp}(\lambda)$ 
4    $p \leftarrow 1$ 
5   while  $Z \leq w$  do
6      $p \leftarrow pe^{-\lambda((w-Z) \wedge 1)}$ 
7     Draw  $Y$  from  $\text{Exp}(\lambda)$ 
8      $Z \leftarrow Z + 1 + Y$ 
9   sum  $\leftarrow$  sum +  $p$ 
10 return sum/ $N$ 

```

---

**2.1.2 At Most One Edge**

Here, let  $k = 1$ ; i.e., we propose an estimator for the probability  $p_1 := p_{\leq 1}$  that the Gilbert graph on  $X \cap [0, w]$  has at most one edge. Let

$$I^+ := \inf\{i \geq 2 : X_i - X_{i-1} \leq 1\}$$

be the index of the first point of  $X$  whose predecessor is at distance at most 1. Putting  $X^+ := (X - X_{I^+}) \cap [1, \infty)$ , Fig. 3 illustrates that the event  $E_{\leq 1}$  is equal to the intersection of the events  $\{X_{I^++1} \geq X_{I^+} + 1\}$  and  $\{G(X^+ \cap [1, w - X_{I^+}]) = \emptyset\}$ .

Moreover, we write

$$\mathcal{F}^+ := \sigma(X_1, X_2, \dots, X_{I^+}, X^+)$$

for the  $\sigma$ -algebra generated by  $X_1, X_2, \dots, X_{I^+}$  and  $X^+$ .

**Theorem 3** (At most one edge) *Suppose that  $w \geq 1$ . Then,*

$$\mathbb{P}(E_{\leq 1} | \mathcal{F}^+) = e^{-\lambda((w-X_{I^+}) \wedge 1)} \mathbb{1}\{G(X^+ \cap [1, w - X_{I^+}]) = \emptyset\}, \quad \text{almost surely.} \tag{5}$$



**Fig. 3** For  $|G(X \cap [0, w])| \leq 1$ , the blue interval may not contain any points and the Gilbert graph restricted to the red interval may not contain edges

*Proof* Since the proof is very similar to that of Theorem 1, we only point to the most important ideas. Equation 5 obviously holds for  $X_{I^+} > w$ . For the case  $X_{I^+} \leq w$ , relying again on the Palm theory of the one-dimensional Poisson process, we have

$$\begin{aligned} \mathbb{P}(E_{\leq 1} | \mathcal{F}^+) &= \mathbb{P}(X \cap (X_{I^+}, (X_{I^+} + 1) \wedge w] = \emptyset | \mathcal{F}^+) \mathbb{1}\{G(X^+ \cap [1, w - X_{I^+}]) = \emptyset\} \\ &= e^{-\lambda((X_{I^+} + 1) \wedge w - X_{I^+})_+} \mathbb{1}\{G(X^+ \cap [1, w - X_{I^+}]) = \emptyset\}, \end{aligned}$$

as asserted. □

Similarly to Section 2.1.2, Theorem 3 yields a conditional MC estimator, which we describe in Algorithm 2.

---

**Algorithm 2** Conditional MC estimator for  $p_1$ .

---

```

input: Number  $N \geq 1$  of MC runs and Poisson intensity  $\lambda > 0$ .
output: Conditional MC estimator of  $p_1$ .
1 sum  $\leftarrow$  0
2 for  $i \leq N$  do
3   Draw  $Z$  from  $\text{Exp}(\lambda)$ 
4    $p \leftarrow 1$ 
5   repeat
6     Draw  $Y$  from  $\text{Exp}(\lambda)$ 
7      $Z \leftarrow Z + Y$ 
8   until  $Y \leq 1$ 
9   while  $Z \leq w$  do
10     $p \leftarrow p e^{-\lambda((w-Z) \wedge 1)}$ 
11    Draw  $Y$  from  $\text{Exp}(\lambda)$ 
12     $Z \leftarrow Z + 1 + Y$ 
13  sum  $\leftarrow$  sum +  $p$ 
14 return sum/ $N$ 
    
```

---

*Remark 1* (Symmetric window) The methods described above could also be applied for a Poisson point process in an interval of the form  $[-w/2, w/2]$ . Then, in addition to  $I_*$  and  $I^+$ , we would need to take into account the corresponding quantities located to the left of the origin.

*Remark 2* ( $k \geq 2$ ) The method to handle  $k = 0, 1$  could certainly be extended to larger  $k \geq 2$ , but the configurational analysis would quickly become very involved. For instance, for  $k = 2$  we would first need to require that the interval  $[X_{I_*-1}, X_{I_*-1} + 1]$  contains only the point  $X_{I_*}$ . Next, if  $X_{I_*+1} \leq X_{I_*} + 1$ , then there may be no more edges to the right of  $X_{I_*+1}$ . On the other hand, the analysis will be different if  $X_{I_*+1} > X_{I_*} + 1$ , because then there can still be one edge in the graph to the right of  $X_{I_*+1}$ .

**2.2 Simulations**

In this section, we illustrate how to estimate the rare-event probabilities  $p_0$  and  $p_1$  via MC. After presenting the crude MC estimator, we illustrate how the conditional MC estimators described in Section 2.1 improve the efficiency drastically. In the simulation study, we estimate  $p_0$  and  $p_1$  for sampling windows of size  $w \in \{5, 7.5, 10\}$  and Poisson intensity  $\lambda = 2$ .



**Table 1** Crude MC estimates for  $p_0 = p_{\leq 0}$  and  $p_{\leq 1}$  for sampling intervals of size  $w \in \{5, 7.5, 10\}$  and Poisson intensity  $\lambda = 2$

$w$	$p_0^{\text{CMC}}$	$p_{\leq 1}^{\text{CMC}}$
5	$4.056 \times 10^{-3} \pm 6.36 \times 10^{-5}$	$1.676 \times 10^{-2} \pm 1.28 \times 10^{-4}$
7.5	$2.410 \times 10^{-4} \pm 1.55 \times 10^{-5}$	$1.354 \times 10^{-4} \pm 3.68 \times 10^{-5}$
10	$1.100 \times 10^{-5} \pm 3.32 \times 10^{-6}$	$8.500 \times 10^{-5} \pm 9.23 \times 10^{-6}$

Standard deviations reported as  $\pm$

Both the crude MC and the conditional MC estimator are computed based on  $N = 10^6$  samples.

To estimate the rare-event probabilities  $p_{\leq k}$  using crude MC, we draw iid samples  $X^{(1)}, \dots, X^{(N)}$  of the Poisson point process on  $[0, w]$  and set

$$p_{\leq k}^{\text{CMC}} := \frac{1}{N} \sum_{i \leq N} \mathbb{1}\{|G(X^{(i)} \cap [0, w])| \leq k\}$$

for the proportion of samples leading to a Gilbert graph with at most  $k$  edges.

The estimates reported in Table 1 reveal that as the size of the sampling window grows, the rare-event probabilities decrease rapidly and that the estimators exhibit a high relative error.

Next, we estimate  $p_{\leq k}$  for  $k = 0, 1$  with the conditional MC methods described in Algorithms 1 and 2. If  $P^{(1)}, \dots, P^{(N)}$  denote the simulation outputs, then we set

$$p_{\leq k}^{\text{Cond}} := \frac{1}{N} \sum_{i \leq N} P^{(i)}.$$

The corresponding estimates shown in Table 2 highlight that the theoretical improvements over crude MC predicted from Theorems 2 and 3 also manifest themselves in the simulation study. This is particularly striking in the setting  $k = 0$ , where the variance can be reduced by several orders of magnitude.

### 2.3 Few Missing Edges

We now focus on computing the rare-event probabilities of having few missing edges. More precisely, we write

$$M_w := \binom{|X \cap [0, w]|}{2} - |G(X \cap [0, w])|$$

**Table 2** Conditional MC estimates for  $p_0 = p_{\leq 0}$  and  $p_{\leq 1}$  for sampling intervals of size  $w \in \{5, 7.5, 10\}$  and Poisson intensity  $\lambda = 2$

$w$	$p_0^{\text{Cond}}$	$p_{\leq 1}^{\text{Cond}}$
5	$4.148 \times 10^{-3} \pm 1.41 \times 10^{-5}$ (20.30)	$1.670 \times 10^{-2} \pm 7.63 \times 10^{-4}$ (2.8)
7.5	$2.244 \times 10^{-4} \pm 1.06 \times 10^{-6}$ (216.3)	$1.304 \times 10^{-4} \pm 1.87 \times 10^{-5}$ (3.9)
10	$1.296 \times 10^{-5} \pm 1.06 \times 10^{-7}$ (984.9)	$9.637 \times 10^{-5} \pm 4.56 \times 10^{-6}$ (4.1)

Variance improvements in comparison to the crude estimator in parentheses



**Fig. 4** For  $M_w = 0$ , the red interval may not contain any points

for the number of edges of  $G(X \cap [0, w])$  that are missing from the complete graph on the vertices  $X \cap [0, w]$ . We write

$$p'_{\leq k} := \mathbb{P}(M_w \leq k)$$

for the probability that at most  $k \geq 0$  edges are missing. Surprisingly, this seemingly more complicated task is more accessible than the probability of seeing few edges considered in Section 2.1, as both  $p'_0 = p'_{\leq 0}$  and  $p'_{\leq 1}$  are amenable to closed-form expressions.

For  $p'_0$ , the key insight is to note that  $M_w = 0$  if and only if  $X \cap [X_1 + 1, w] = \emptyset$ ; see Fig. 4.

**Theorem 4** (No missing edges) *Suppose that  $w \geq 1$ . Then,*

$$p'_0 = e^{-\lambda(w-1)} + (w - 1)\lambda e^{-\lambda(w-1)}.$$

*Proof* As observed in the above remark, we need to compute  $\mathbb{P}(X \cap [X_1 + 1, w] = \emptyset)$ . Hence, invoking the void probability for a Poisson point process,

$$p'_0 = \int_0^\infty \lambda e^{-\lambda x_1} \mathbb{P}(X \cap [x_1 + 1, w] = \emptyset) dx_1 = e^{-\lambda(w-1)} + \int_0^{w-1} \lambda e^{-\lambda x_1} e^{-\lambda(w-1-x_1)} dx_1,$$

which equals  $e^{-\lambda(w-1)} + (w - 1)\lambda e^{-\lambda(w-1)}$ , as asserted. □

Next, we compute the probability of observing at most one missing edge.

**Theorem 5** (At most one missing edge) *Suppose that  $w \geq 2$ . Then,*

$$p'_{\leq 1} = p'_0 + \frac{\lambda^2(w - 2)^2}{2} e^{-\lambda w} + (w - 3/2)\lambda^2 e^{-\lambda(w-1)}.$$

*Proof* We decompose  $p'_{\leq 1}$  as

$$p'_{\leq 1} = p'_0 + \mathbb{P}(M_w = 1, X \cap [X_1 + 2, w] \neq \emptyset) + \mathbb{P}(M_w = 1, X \cap [X_1 + 2, w] = \emptyset)$$

and compute the second and third probability separately. This corresponds to Case 1 and Case 2 below. Note that under the event  $\{X \cap [X_1 + 2, w] \neq \emptyset\}$ , we have  $|X \cap [0, w]| = 2$ , since more points would imply at least two missing edges.

**Case 1.**  $(X \cap (X_1, X_1 + 2] = \emptyset$  and  $|X \cap [X_1 + 2, w]| = 1)$

Conditioning on  $X_1$ , the probability of this event becomes

$$\int_0^{w-2} \lambda e^{-\lambda x_1} \mathbb{P}(X \cap (x_1, x_1 + 2] = \emptyset, |X \cap [x_1 + 2, w]| = 1) dx_1.$$

Inserting the void probability for the Poisson process, we arrive at

$$\int_0^{w-2} \lambda e^{-\lambda x_1} e^{-2\lambda(w-2-x_1)} e^{-\lambda(w-2-x_1)} dx_1 = \lambda^2 e^{-\lambda w} \int_0^{w-2} (w-2-x_1) dx_1 = \frac{\lambda^2(w-2)^2}{2} e^{-\lambda w}.$$



**Fig. 5** Configuration where  $M_w = 1$  and  $X \cap [X_1 + 2, w] = \emptyset$ . Here,  $[X_1, X'_1 - 1]$  is in red and  $[X_1 + 1, X_1 + 2]$  in blue

It remains to treat the case  $X \cap [X_1 + 2, w] = \emptyset$ . First, note that conditioned on  $X_1 = x_1$ , the point process  $X \cap [x_1, w]$  is Poisson. Thinking of this Poisson point process to be formally extended to  $-\infty$  to the left, we write  $X'_1$  for the first Poisson point to the left of  $w$ . Then, Fig. 5 illustrates that the event  $\{M_w = 1\} \cap \{X \cap [X_1 + 2, w] = \emptyset\}$  is equivalent to  $X'_1 \in [X_1 + 1, X_1 + 2]$  and  $X \cap (X_1, X'_1 - 1) = \emptyset$ .

For Case 2, we distinguish between the cases  $X_1 \leq w - 2$  and  $X_1 \in [w - 2, w - 1]$ .

**Case 2a.** ( $X_1 \leq w - 2, X'_1 \in [X_1 + 1, X_1 + 2]$  and  $X \cap (X_1, X'_1 - 1) = \emptyset$ ) Then, we compute the desired probability as

$$\int_0^{w-2} \lambda e^{-\lambda x_1} \int_{x_1+1}^{x_1+2} \lambda e^{-\lambda(w-x'_1)} \mathbb{P}(X \cap (x_1, x'_1 - 1) = \emptyset) dx'_1 dx_1 = \int_0^{w-2} \int_{x_1+1}^{x_1+2} \lambda^2 e^{-\lambda(w-1)} dx'_1 dx_1,$$

which equals  $(w - 2)\lambda^2 e^{-\lambda(w-1)}$ .

**Case 2b.** ( $X_1 \in [w - 2, w - 1], X'_1 \in [X_1 + 1, w]$  and  $X \cap (X_1, X'_1 - 1) = \emptyset$ ) Finally,

$$\int_{w-2}^{w-1} \lambda e^{-\lambda x_1} \int_{x_1+1}^w \lambda e^{-\lambda(w-x'_1)} \mathbb{P}(X \cap (x_1, x'_1 - 1) = \emptyset) dx'_1 dx_1 = \int_{w-2}^{w-1} \int_{x_1+1}^w \lambda^2 e^{-\lambda(w-1)} dx'_1 dx_1,$$

which equals  $\frac{\lambda^2}{2} e^{-\lambda(w-1)}$ . Assembling the different cases together concludes the proof.  $\square$

### 3 Conditional MC in Higher Dimensions

In Section 2, we analyzed rare events related to few edges or few missing edges in a one-dimensional setting. There, closed-form expressions for conditional probabilities were derived from the natural ordering of the Poisson points. Now, we proceed to higher dimensions and also consider more general deviations from the mean number of edges. First, we again illustrate that substantial variance reductions are possible through a surprisingly simple conditional MC method. Loosely speaking, the Poisson point process consists of 1) an infinite sequence of random points in the window determining the locations of points and 2) a Poisson random variable determining the number of points in the sampling window. We use that after conditioning on the spatial locations, the rare-event probability is available in closed form. This type of Poisson conditioning is novel in a spatial rare-event estimation context, but has strong ties to approaches appearing earlier in seemingly unrelated problems. More precisely, for instance in reliability theory, related conditional MC schemes lead to spectacular variance reductions (Lomonosov and Shprungin 1999; Vaisman et al. 2015).

The rest of this section is organized as follows. First, Section 3.1 describes how to estimate the rare event probabilities related to too few and too many edges relative to their expected number through conditional MC. Then, Section 3.2 presents a simulation

study illustrating that this estimator can reduce the estimation variance by several orders of magnitude.

### 3.1 Conditioning on a Spatial Component

We consider a full-dimensional sampling window  $W \subset \mathbb{R}^d$  and rare events of the form

$$F_{<a} := \{|G(X \cap W)| < (1 - a)\mu\} \quad \text{and} \quad F_{>a} := \{|G(X \cap W)| > (1 + a)\mu\},$$

where  $\mu := \mathbb{E}[|G(X \cap W)|]$  denotes the mean (i.e., expected) number of edges, which can be estimated through simulations. Alternatively, if  $W$  is so large that edge effects can be neglected, then the Slivnyak–Mecke formula (Last and Penrose 2017, Theorem 4.4) gives the approximation  $\mu \approx \frac{1}{2}|W|\lambda^2\kappa_d$ , where  $\kappa_d$  denotes the volume of the unit ball in  $\mathbb{R}^d$ .

The key idea for developing a conditional MC scheme is to use the explicit construction of a Poisson point process in a bounded sampling window. More precisely, let  $X_\infty = \{X_n\}_{n \geq 1}$  be an iid family of uniform random points in  $W$  and  $K$  be an independent Poisson random variable with parameter  $\lambda|W|$ . Then,  $\{X_n\}_{n \leq K}$  is a Poisson point process in  $W$  with intensity  $\lambda$ , Last and Penrose (2017, Theorem 3.6).

In conditional MC, we use the fact that the rare-event probabilities can be computed in closed form after we condition on the locations  $X_\infty$ . More precisely, we let

$$K_{<a}(X_\infty) := \inf\{k \geq 1 : |G(\{X_1, \dots, X_k\} \cap W)| \geq (1 - a)\mu\}$$

denote the first time where the Gilbert graph on the nodes  $\{X_1, \dots, X_k\}$  contains at least  $(1 - a)\mu$  edges and refer to Fig. 6 for an illustration.

Similarly, let

$$K_{>a}(X_\infty) := \sup\{k \geq 1 : |G(\{X_1, \dots, X_k\} \cap W)| \leq (1 + a)\mu\}$$

denote the largest  $k$  for which the Gilbert graph on the nodes  $\{X_1, \dots, X_k\}$  contains at most  $(1 + a)\mu$  edges. Then,

$$\begin{aligned} \mathbb{P}(F_{<a}) &= \mathbb{E}[\mathbb{P}(F_{<a} | X_\infty)] = \mathbb{E}[F_{\text{Poi}}(K_{<a}(X_\infty) - 1)], \\ \mathbb{P}(F_{>a}) &= \mathbb{E}[\mathbb{P}(F_{>a} | X_\infty)] = \mathbb{E}[1 - F_{\text{Poi}}(K_{>a}(X_\infty))], \end{aligned} \tag{6}$$

where  $F_{\text{Poi}} : \mathbb{Z}_{\geq 0} \rightarrow [0, 1]$  denotes the cumulative distribution function of a Poisson random variable with parameter  $\lambda|W|$ .

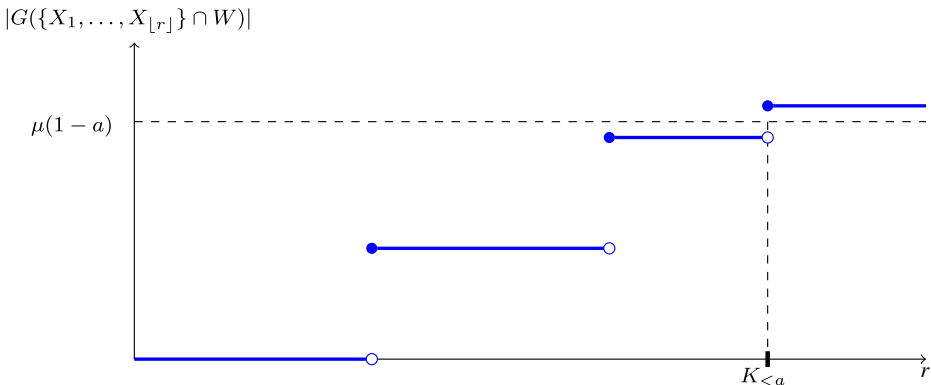


Fig. 6 Conceptual illustration of  $K_{<a}$

### 3.2 Numerical Results

We sample planar homogeneous Poisson point processes  $X^S = \{X_i^S\}_{i \geq 1}$ ,  $X^M = \{X_i^M\}_{i \geq 1}$  and  $X^L = \{X_i^L\}_{i \geq 1}$  with intensity  $\lambda = 2$  in windows of size  $20 \times 20$ ,  $25 \times 25$ , and  $30 \times 30$ , respectively. Here, S, M, and L stand for small, medium, and large, respectively.

In Section 1, we mentioned that a major challenge in devising efficient estimators for rare-event probabilities related to the edge count comes from a qualitatively different tail behavior: light on the left, heavy on the right. In other words, we expect that in the left tail, we see changes throughout the sampling window, whereas in the right tail, a singular particular configuration in a small part of the window is sufficient to induce the rare event. We recall that the left tail of a random variable  $Z$  refers to the probabilities  $\mathbb{P}(Z \leq r)$  for small  $r$  and the right tail refers to the probabilities  $\mathbb{P}(Z \geq r)$  for large  $r$ .

Although on a bounded sampling window, the theoretical difference between the left and the right tail is subtle, we illustrate in Table 3 that it does become visible when considering the quantiles  $Q_\alpha$  and  $Q_{1-\alpha}$  for the number of edges if  $\alpha$  is small. Here, the empirical quantiles for  $10^6$  samples of the edge counts in a  $(20 \times 20)$ -window are shown. For instance the 1%-quantile is 16.4% lower than the mean, which is a similar deviation as the 18% exceedance of the 99%-quantile. However, when moving to the 0.01%-quantile, then it is 25.2% lower than the mean, whereas the corresponding 99.99% quantile exceeds the mean by 30.0%. These figures are an early indication of the difference in the tails that will reappear far more pronouncedly in the numerical results concerning the estimation of the rare-event probabilities  $\mathbb{P}(F_{<0.2})$  and  $\mathbb{P}(F_{>0.2})$ .

In the rest of the section, we estimate the rare-event probabilities

$$q_{<0.2} := \mathbb{P}(F_{<0.2}) \quad \text{and} \quad q_{>0.2} := \mathbb{P}(F_{>0.2})$$

corresponding to 20% deviations from the mean. Here, we draw  $N = 10^5$  samples of  $X_\infty$ . Then, taking into account the representation (6), we set

$$p_{<0.2}^{\text{Cond}} := \frac{1}{N} \sum_{i \leq N} F_{\text{Poi}}(K_{<a}(X_\infty(i)) - 1),$$

$$p_{>0.2}^{\text{Cond}} := \frac{1}{N} \sum_{i \leq N} (1 - F_{\text{Poi}}(K_{>a}(X_\infty(i)))).$$

The variances of the crude MC estimators are equal to  $q_{<0.2}(1 - q_{<0.2})$ .

We report the estimates  $p_{<0.2}^{\text{Cond}}$  and  $p_{>0.2}^{\text{Cond}}$  in Table 4. First, we see that the exceedance probabilities  $p_{>0.2}^{\text{Cond}}$  are always smaller than the corresponding undershoot probabilities  $p_{<0.2}^{\text{Cond}}$ . This supports the preliminary impression of the difference in the tail behavior hinted at in Table 3. Moreover, in all examples conditional MC reduces the estimation variance massively and the efficiency gains become more pronounced the rarer the event.

**Table 3** Empirical quantiles for the number of edges in a  $(20 \times 20)$ -window: absolute values (upper two rows) and relative deviation from the mean  $\mu$  (lower two rows)

$\alpha$	$10^{-2}$	$10^{-3}$	$10^{-4}$
$Q_\alpha$	2012	1892	1800
$Q_{1-\alpha}$	2841	2999	3130
$\frac{Q_\alpha - \mu}{\mu}$	-0.164	-0.214	-0.252
$\frac{Q_{1-\alpha} - \mu}{\mu}$	0.180	0.246	0.300

**Table 4** Estimates for  $q_{<0.2}$  and  $q_{>0.2}$  based on conditional MC for different window sizes based on  $N = 10^5$  samples

	$P_{<0.2}^{\text{Cond}}$	$P_{>0.2}^{\text{Cond}}$
$X^S$	$2.023 \times 10^{-3} \pm 6.98 \times 10^{-6}$ (414.8)	$5.118 \times 10^{-3} \pm 1.63 \times 10^{-5}$ (193.7)
$X^M$	$1.542 \times 10^{-4} \pm 7.05 \times 10^{-7}$ (3106.4)	$6.764 \times 10^{-4} \pm 2.77 \times 10^{-6}$ (878.6)
$X^L$	$6.912 \times 10^{-6} \pm 4.19 \times 10^{-8}$ (39, 415.8)	$6.242 \times 10^{-5} \pm 3.24 \times 10^{-7}$ (5, 911.1)

Variance improvements in comparison to the crude estimator in parentheses

## 4 Importance Sampling

The conditional MC estimators constructed in Section 3 take into account that an atypically large number of Poisson points leads to a Gilbert graph exhibiting considerably more edges than expected. However, not only the number but also the location of points play a pivotal role. For instance if the points tend to repel each other, then we would typically observe fewer edges. Similarly, clustered point patterns should induce more edges.

In order to implement these insights, we resort to the technique of *importance sampling* (Kroese et al. 2013, Chapter 9.7). That is, we draw samples from a point process with distribution  $\mathbb{Q}$  for which the rare event becomes typical and then correct the estimation bias by weighting with likelihood ratios. In Sections 4.1 and 4.2 below, we explain how to implement these steps through configuration-dependent birth-death processes reminiscent of the Strauss process from spatial statistics (Møller and Waagepetersen 2004). Finally, in Section 4.3, we illustrate in a simulation study how importance sampling of the spatial locations reduces the estimation variance further.

### 4.1 Lower Tails

The key observation is that under the rare event of seeing exceptionally few edges, we expect a repulsion between points. More precisely, the most likely reason for the rare event are changes to the configuration of the underlying Poisson point process throughout the entire sampling window. The large-deviation analysis of Seppäläinen and Yukich (2001) suggests to perform importance sampling where, instead of considering the distribution  $\mathbb{P}$  of the a priori Poisson point process, we draw samples according to a different stationary point process with distribution  $\mathbb{Q}$  such that under  $\mathbb{Q}$ , the original rare event becomes typical and whose deviation from  $\mathbb{P}$ , as measured through the Kullback–Leibler divergence  $h(\mathbb{Q} | \mathbb{P})$ , is minimized. We implement this repulsion by a dependent thinning inspired from the Strauss process.

Here, we start from a realization of the Gilbert graph on  $n_0 = \lfloor \lambda |W| \rfloor$  iid points  $\{X_1, \dots, X_{n_0}\}$ , and then, we thin out points successively. An independent thinning of points would give rise to uniformly distributed locations without interactions. In the importance sampling, we thin instead via a configuration-dependent birth-death process; see e.g., Chapter 9.7 of Kroese et al. (2013).

To describe the death mechanism more precisely, we draw inspiration from the Strauss process and choose the probability  $p_i$  to remove point  $X_i$  proportional to  $\gamma^{\deg(X_i)}$ , where  $\deg(X_i)$  denotes the degree of  $X_i$  in the Gilbert graph and  $\gamma > 1$  is a parameter of the algorithm. Algorithm 3 shows the pseudo-code leading to the importance sampling estimator  $q_{<a}^{\text{IS}}$  for  $q_{<a}$ ; in the algorithm we write  $|X|$  for the number of points of  $X$ . To understand

the principle behind Algorithm 3, we briefly expound on the general approach in importance sampling, and refer the reader to Chapter 9.7 of Kroese et al. (2013) for an in-depth discussion.

As mentioned above, when thinning out points independently until the number of edges in the Gilbert graph falls below  $\mu(1 - a)$ , we would arrive at the random variable  $K_{<a}(X_\infty)$  from Section 3. However, thinning out according to a configuration-dependent probability distorts its distribution towards a probability measure  $\mathbb{Q}$  that is in general different from the true distribution  $\mathbb{P}$ . Still, by construction,  $\mathbb{Q}$  is absolutely continuous with respect to  $\mathbb{P}$ , and we let  $\rho := d\mathbb{P}/d\mathbb{Q}$  be the likelihood ratio. Then, from a conceptual point of view, Algorithm 3 first draws  $N \geq 1$  iid samples  $X_\infty(1), \dots, X_\infty(N)$  from the distorted distribution  $\mathbb{Q}$  with associated likelihood ratios  $\rho_1, \dots, \rho_N$ , and then computes

$$q_{<a}^{\text{IS}} := \frac{1}{N} \sum_{i \leq N} \rho_i F_{\text{Poi}}(K_{<a}(X_\infty(i)) - 1),$$

where we recall that  $F_{\text{Poi}}$  denotes the cumulative distribution function of a Poisson random variable with parameter  $\lambda|W|$ .

---

**Algorithm 3** Importance sampling estimator  $q_{<a}^{\text{IS}}$  for  $q_{<a}$ .

---

```

input: Number  $N \geq 1$  of MC runs; parameter  $\gamma$ ; Poisson intensity  $\lambda$ ; sampling window  $W$ .
output: Importance sampling estimator  $q_{<a}^{\text{IS}}$  for  $q_{<a}$ .
1  $n_0 \leftarrow \lfloor \lambda|W| \rfloor$ 
2  $\text{sum} \leftarrow 0$ 
3 for  $i \leq N$  do
4    $X \leftarrow \{X_1, \dots, X_{n_0}\}$  iid uniform in  $W$ 
5    $\rho \leftarrow 1$ 
6   while  $|G(X \cap W)| \geq \mu(1 - a)$  do
7     draw  $X_i \in X$  with probability  $r_i = \gamma^{\deg(X_i)} / \sum_{X_j \in X} \gamma^{\deg(X_j)}$ 
8      $\rho \leftarrow \rho / (|X|r_i)$ 
9      $X \leftarrow X \setminus \{X_i\}$ 
10   $\text{sum} \leftarrow \text{sum} + \rho F_{\text{Poi}}(|X|)$ 
11 return  $\text{sum}/N$ 

```

---

We now explain why the  $q_{<a}^{\text{IS}}$  returned by the algorithm is a good approximation for the probability  $q_{<a}$ . The general philosophy is that when performing importance sampling with a proposal distribution that is close to the conditional distribution under the rare event, we can be optimistic that importance sampling decreases the estimation variance. For instance, in Kroese et al. (2013, Section 13.2), it is shown that an unbiased zero-variance estimator could be obtained if the proposal distribution coincided with the conditional distribution under the rare event. In our setting, this means that we can expect  $q_{<a}^{\text{IS}}$  to be a good approximation for the probability  $q_{<a}$  if the proposal density in the importance sampling is close to the conditional distribution under the event  $F_{<a}$ .

To achieve this goal, intuitively, we would like to choose the thinning parameter  $\gamma > 1$  such that the number of edges in the rare event  $F_{<a}$  should match the expected number of edges under the thinning. To develop a heuristic for this choice, we consider a Strauss process, where we restrict to the two-dimensional setting to allow for an accessible presentation. To make the article self-contained, we briefly recall the definition of a Strauss process from Møller and Waagepetersen (2004). The *Strauss process* in the observation window  $W$

is a Gibbs process whose density  $f(\varphi)$  with respect to a unit-intensity Poisson point process on  $W$  is given by

$$f(\varphi) = \alpha \beta^{|\varphi|} \gamma^{s_R(\varphi)},$$

with parameters  $R > 0$  (interaction radius),  $\beta > 0$  (interaction parameter) and  $\gamma \in [0, 1)$  (interaction strength) where  $\alpha > 0$  is a normalizing constant and  $s_R(\varphi) = \#\{\{x, y\} \subseteq \varphi : 0 < |x - y| \leq R\}$  denotes the number of unordered point pairs in  $\varphi$  within a distance at most  $R$ .

Hence, matching the number of edges in the rare event to the expected number of edges leads to the equation

$$(1 - a)\mu = (\lambda^{\text{Str}})^2 \int_0^1 \pi r \rho^{\text{Str}}(r) dr, \tag{7}$$

where  $\lambda^{\text{Str}} > 0$  and  $\rho^{\text{Str}} : [0, \infty) \rightarrow [0, \infty)$  denote the intensity and pair-correlation function of the Strauss process, respectively (Møller and Waagepetersen 2004). In contrast to the Poisson setting, neither  $\lambda^{\text{Str}}$  nor  $\rho^{\text{Str}}$  are available in closed form. Still, both quantities admit accurate saddle-point approximations  $\lambda^{\text{PS}} > 0$  and  $\rho^{\text{PS}} : [0, \infty) \rightarrow [0, \infty)$  that are ready to implement in the Strauss case (Baddeley and Nair 2012).

More precisely,  $\lambda^{\text{PS}}$  is the unique positive solution of the equation

$$\lambda^{\text{PS}} G e^{\lambda^{\text{PS}} G} = \lambda G,$$

where  $G = (1 - \beta)\pi$  and  $\beta = \log(\gamma)$ , see Baddeley and Nair (2012). Then, recalling that we work in a planar setting, we put

$$\rho^{\text{PS}}(r) := \beta \exp((1 - \beta)^2 b(r/2) \lambda^{\text{PS}}),$$

where

$$b(r/2) := 2 \cos^{-1}(r/2) - r \sqrt{1 - (r/2)^2}$$

denotes the intersection area of two unit disks at distance  $r$ . Inserting these approximations into (7) and solving the resulting implicit equation yields  $\beta \approx \log(1.018)$  and a value of  $\gamma \approx 1.018$ .

### 4.2 Upper Tails

Similar to the lower tails, we can strengthen the estimator by combining conditional MC with importance sampling on the spatial locations. For the upper tails, it is natural to devise an importance sampling scheme favoring clustering rather than repulsion. From the point of view of large deviations, this phenomenon was considered in Chatterjee and Harel (2020). There, it is shown that all excess edges come from a ball of size 1 containing a highly dense configuration of Poisson points. More precisely, the method is particularly powerful in settings where the rare event is the epitome of a condensation phenomenon. That is, a peculiar behavior of the point process in a small part in the window becomes the most likely explanation of the rare event. As laid out in Chatterjee and Harel (2020), at least in the asymptotic regime of large deviations, the rare event of observing too many edges is governed by the above-described condensation phenomenon.

We propose to take the clustering into account via a birth mechanism favoring the generation of points in areas that would lead to a large number of additional edges. To ease implementation, the density of the birth mechanism is discretized and remains constant in bins of a suitably chosen grid. The density in a bin at position  $x \in W$  is proportional to  $\gamma^{n(x)}$ , where  $\gamma > 1$  is a parameter governing the strength of the clustering and  $n(x)$  denotes the



**Table 5** Estimates for  $q_{<0.2}$  and  $q_{>0.2}$  based on importance sampling with Strauss-type processes for different window sizes based on  $N = 10^5$  samples

	$q_{<0.2}^{\text{IS}}$	$q_{>0.2}^{\text{IS}}$
$X^S$	$2.025 \times 10^{-3} \pm 6.22 \times 10^{-6}$ (523.3)	$5.125 \times 10^{-3} \pm 1.57 \times 10^{-5}$ (207.9)
$X^M$	$1.544 \times 10^{-4} \pm 6.16 \times 10^{-7}$ (4071.0)	$6.744 \times 10^{-4} \pm 2.66 \times 10^{-6}$ (951.8)
$X^L$	$6.935 \times 10^{-6} \pm 3.63 \times 10^{-8}$ (52, 665.8)	$6.240 \times 10^{-5} \pm 3.09 \times 10^{-7}$ (6, 537.22)

Variance improvements in comparison to the crude estimator in parentheses

number of Poisson points in a suitable neighborhood around  $x$ , such as the bin containing  $x$  together with all adjacent bins.

Similar to the lower tails case, a subtle point in this approach pertains choosing the parameter  $\gamma > 1$ . Unfortunately, an attractive Strauss process is ill-defined in the entire Euclidean space, so that the saddle-point approximation from Section 4.1 does not apply. In Section 4.3 below, we therefore rely on a pilot run suggesting  $\gamma = 1.01$  as a good choice for further variance reduction. Although in this pilot run, the number of simulations is small, and estimates of the variance are still volatile, we found that it provides a good indication for simulations on a larger scale.

Section 3.2 revealed that even with a sample size of  $N = 10^5$  the conditional MC estimators still exhibit a considerable relative error. Now, we illustrate that importance sampling may be an appealing option to reduce this error.

Table 5 reports the estimates  $q_{<0.2}^{\text{IS}}$  and  $q_{>0.2}^{\text{IS}}$  for different sizes of the sampling window as in Section 3.2. In the left tail, we see massive variance improvements when compared to the crude MC estimator. In the right tail, the gains are substantial, but a little less pronounced. This matches the intuition that the root of the rare events in the right tails should be a condensation phenomenon, which goes against the heuristic of changing the point process homogeneously throughout the window.

**Acknowledgments** The authors thank the anonymous referee for reading through the entire manuscript thoroughly and for providing us with high-quality and constructive feedback. Major parts of this research were carried out while Christian Hirsch was on a research visit at the University of Queensland and while Sarat B. Moka was on a research visit at Ulm University. We thank both hosts for their hospitality and the Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS) for the support under grant number CE140100049.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Asmussen S, Kroese DP (2006) Improved algorithms for rare event simulation with heavy tails. *Adv Appl Probab* 38(2):545–558
- Baccelli F, Błaszczyszyn B (2009) *Stochastic Geometry and Wireless Networks*. Now Publishers Inc

- Baddeley A, Nair G (2012) Approximating the moments of a spatial point process. *Stat* 1(1):18–30
- Bhamidi S, Hannig J, Lee CY, Nolen J (2015) The importance sampling technique for understanding rare events in Erdős-Rényi random graphs. *Electron J Probab*. 20(107):30
- Billingsley P (1995) *Probability and measure*, 3rd edn. Wiley, New York
- Blanchet J, Glynn P (2008) Efficient rare-event simulation for the maximum of heavy-tailed random walks. *Ann Appl Probab* 18(4):1351–1378
- Chatterjee S, Harel M (2020) Localization in random geometric graphs with too many edges. *Ann Probab* 48(2):574–621
- Keeler HP, Jahnelt B, Maye O, Aschenbach D, Brzozowski M (2018) Disruptive events in high-density cellular networks. 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)
- Kroese DP, Taimre T, Botev ZI (2013) *Handbook of monte carlo methods*. Wiley, New York
- Last G, Penrose MD (2017) *Lectures on the poisson process*. Cambridge University Press, Cambridge
- Lomonosov M, Shpungin Y (1999) Combinatorics of reliability Monte Carlo. *Random Struct Algorithm* 14(4):329–343
- Møller J, Waagepetersen RP (2004) *Statistical inference and simulation for spatial point processes*. CRC, Boca Raton
- Rojas-Nandayapa L (2013) A review of conditional rare event simulation for tail probabilities of heavy tailed random variables. *Bol Soc Mat Mex* (3) 19(2):159–182
- Seppäläinen T, Yukich JE (2001) Large deviation principles for Euclidean functionals and other nearly additive processes. *Probab Theory Rel Fields* 120(3):309–345
- Stenzel O, Hirsch C, Brereton T, Baumeier T, Andrienko D, Kroese DP, Schmidt V (2014) A general framework for consistent estimation of charge transport properties via random walks in random environments. *Multiscale Model Simul* 12:1108–1134
- Vaisman R, Kroese DP, Gertsbakh IB (2015) Improved sampling plans for combinatorial invariants of coherent systems. *IEEE Trans Rel* 65(1):410–424

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.