

Interference-Aware Molecular Detector Design for Clustered Bio-Nanonetworks

Seyed Mohammad Azimi-Abarghouyi*, Harpreet S. Dhillon†, and Leandros Tassiulas‡

*KTH Royal Institute of Technology, Stockholm, Sweden, seyaa@kth.se

†Virginia Tech, Blacksburg, VA USA, hddhillon@vt.edu

‡Yale University, New Haven, CT USA, leandros.tassiulas@yale.edu

Abstract—We present a comprehensive approach to the modeling and design of clustered molecular bio-nanonetworks in which nano-machines of different clusters release an appropriate number of molecules to transmit their sensed information to their respective fusion centers. The fusion centers decode this information by counting the number of molecules received in the given time slot. Owing to the propagation properties of the biological media, this setup suffers from both inter- and intra-cluster interference that needs to be carefully modeled. We first develop a novel spatial model for this setup by modeling nano-machines as a Poisson cluster process with the fusion centers forming its parent point process. For this setup, we then derive a new set of distance distributions in the three-dimensional space, resulting in a remarkably simple result for the special case of the Thomas cluster process. Accordingly, total interference from previous symbols and different clusters is characterized and its expected value is obtained. Then, using the expected value, a simple detector suitable for biological applications is proposed. The impact of different parameters on the performance of the detector is also investigated.

Index Terms—Molecular communications, clustered nanonetworks, stochastic geometry, Thomas cluster process.

I. INTRODUCTION

Molecular communications (MC) has attracted significant research interest due to its numerous applications in biological and communication engineering such as nanoscale sensing, drug delivery, and body area networks [2]. In MC, biological nano-machines (NMs) release molecules into the environment to realize nano-scale information transfer. Information can be encoded in the concentration or type of the molecules or at the time instants at which they are released. A detector decodes the information transmitted by a NM based on the number of captured molecules. Some examples of modulation and inter-symbol interference (ISI) mitigation techniques can be found in [3], [4]. These works focus on a single pair of transmitting and receiving NMs. On the other hand, nanonetworking and internet of bio-nano things paradigms require information collection from multiple arbitrarily-located NMs sensing potentially different phenomena [5]. Hence, interaction and interference of released molecules of multiple NMs over time play a crucial role in the performance characterization and design of bio-nanonetworks.

Several prior works consider single source interference [6] and fixed-location configurations with a limited number of interfering NMs [7], [8]. Due to the randomness and irregularity in the locations of NMs, stochastic geometry approaches based on the Poisson point process (PPP) assumption have also been considered for the modeling and analysis of nanonetworks [9]–[14]. However, these early stochastic geometry-based works lack in two important aspects that inspired this paper. First, these works did not explore the dependence of detectors on interference even though interference causes significant increase in the molecule count. Second, despite the usefulness of the PPP for modeling the uniform deployments of NMs, it cannot model deployments where fusion centers (FCs) are located at places with high NM density to sense and process their information. The latter is particularly important for capturing the fact that many biological phenomena, such as abnormalities, are localized and multiple cooperative NMs, e.g., each monitoring a different feature of a target, might be required for fast, accurate, and diverse sensing [15]. For instance, one can think of multiple tumors in the body as targets, which might need to be monitored for smart drug delivery. In such cases, it is important to take into account non-uniformity as well as the correlation that exists between the locations of the NMs and FCs. Models based on Poisson cluster processes (PCPs) [16] have recently been well studied for wireless networks over two-dimensional (2D) regions [17]–[19]. In these works, the network usually follows a Thomas cluster process (TCP) [16] or Matérn cluster process (MCP) [16]. Further, key distance distributions for both TCP and MCPs have been derived in the recent years [20], [21].

In this paper, we consider a bio-nanonetwork setup consisting of multiple NMs forming clusters around their respective FCs in a three-dimensional (3D) biological media. In each time slot, one NM is assumed to transmit its sensed information to its respective FC. This not only results in inter-cluster interference from the simultaneous transmissions in the current time slot but also both inter- and intra-cluster interference from the transmissions corresponding to the previous time slots. A rigorous analysis of these interferences requires a spatial model that is rich enough to capture salient features of the setup while being tractable enough for mathematical derivations. Inspired by this, we develop a novel approach

based on PCPs to facilitate tractable modeling and analysis of 3D clustered bio-nanonetworks. We first characterize relevant distance distributions of 3D PCPs that are inspired by this setup and provide closed-form results for 3D TCPs. In particular, we identify a specific structure for 3D TCPs that provides a remarkably simple distribution. Both ISI and interference from interfering NMs in different clusters are characterized, and the expected value of intra- and inter-cluster interferences are derived. Then accordingly, a low-complexity detector for FCs is designed that can efficiently adjust to the interference. We also investigate the impact of different parameters of the system model on the performance in terms of the error probability and the expected value of interference. As expected, our analysis reveals that a higher intensity of cluster centers, distance of a reference NM to the center of its cluster, or number of time slots has a degrading effect on the performance. Also, increasing the time slot duration decreases the error probability.

II. SYSTEM MODEL

Important aspects of the system model are discussed next.

A. Spatial Model

We consider a 3D clustered bio-nanonetwork as shown in Fig. 1, where the locations of NMs are modeled as a PCP in $\mathbb{R}^3$. A 3D PCP Φ can be defined as a union of offspring points in $\mathbb{R}^3$ that are located around parent points, i.e., cluster centers. The parent point process is a 3D PPP Φ_p with intensity λ_p , and the offspring point processes are conditionally independent. The set of offspring points of $\mathbf{x} \in \Phi_p$ is denoted by $\mathcal{N}^{\mathbf{x}}$, such that $\Phi = \cup_{\mathbf{x} \in \Phi_p} \mathcal{N}^{\mathbf{x}}$, and the conditional probability density function (PDF) of each element being at a location $\mathbf{y} + \mathbf{x} \in \mathbb{R}^3$ is $f_{\mathbf{Y}}(\mathbf{y})$. After characterizing theoretical results in terms of general distribution $f_{\mathbf{Y}}(\mathbf{y})$, we specialize the results to the 3D TCP where the points are distributed around cluster centers according to an independent Gaussian distribution

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{3}{2}}\sigma^3} \exp\left(-\frac{\|\mathbf{y}\|^2}{2\sigma^2}\right). \quad (1)$$

In the center of each cluster, there is an FC that detects, gathers and processes the transmitted information of the NMs of the same cluster. In our analysis, we will model each FC as a ball of non-zero radius r_0 that will place some restrictions on the placement of FCs and NMs. For instance, NMs cannot lie inside the FCs. While the exact analysis of this modified setting is complicated and will lead to significant loss in tractability, we will include the non-zero radius r_0 in many components of our analysis, thus capturing these additional restrictions while maintaining tractability.

B. Propagation Model

We consider a time-slotted transmission with time slots of duration T . This implicitly assumes perfect synchronization of all NMs, which is a common assumption in the MC literature. However, it is not required for the accuracy of our results since the received molecule distribution is primarily determined by the NM locations.

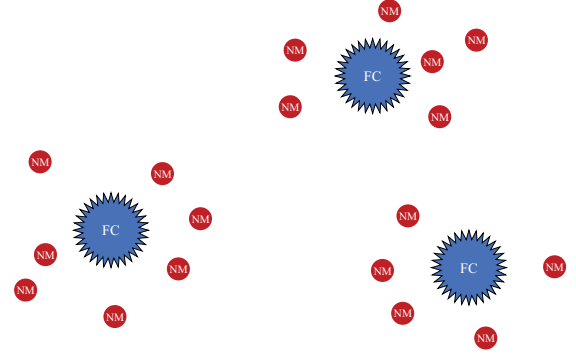


Fig. 1. An illustration of clustered bio-nanonetworks projected on a 2D space.

An NM at a location $\mathbf{z}$ can release $X_{\mathbf{z}} \in \{x_0, x_1\}$ molecules at the beginning of a time slot, x_0 for bit 0 and x_1 for bit 1, each with equal probability [22].¹ All NMs are assumed to release the same type of molecule. The molecules released by the NMs propagate through the biological medium and are observed at the FCs. We assume that a molecule is absorbed at an FC and then counted when it hits the surface of the FC. Let $p_{iL}(d)$ denote *hitting probability*, i.e., the probability that the molecule with distance d to the FC and released at the i -th time slot hits the FC during the L -th time slot. The hitting probability depends on the medium and type of detector [3]. For instance, for a general 3D environment with a point-source NM and a spherical absorbing FC with radius r_0 , the hitting probabilities are given by [23]

$$p_{iL}(d) = g((L-i+1)T, d) - g((L-i)T, d), \quad \forall i \in \{1, \dots, L-1\}, \quad (2)$$

$$p_{LL}(d) = g(T, d), \quad (3)$$

where

$$g(t, d) = \frac{r_0}{2d} \left[\exp\left(-\sqrt{\frac{\mu}{D}}(d-r_0)\right) \operatorname{erfc}\left(\frac{d-r_0}{\sqrt{4Dt}} - \sqrt{\mu t}\right) + \exp\left(\sqrt{\frac{\mu}{D}}(d-r_0)\right) \operatorname{erfc}\left(\frac{d-r_0}{\sqrt{4Dt}} + \sqrt{\mu t}\right) \right], \quad (4)$$

where D is the diffusion coefficient, μ denotes the reaction rate constant of molecular degradation, and $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$.

C. Transmission Scheme

In order to mitigate interference in each cluster of our clustered MC setup, in each time slot, only one of the NMs of the cluster is scheduled to release molecules according to its information availability status. Let the NM at $\mathbf{y}_{\mathbf{x}}^i$ in the cluster with $\mathbf{x} \in \Phi_p$ be selected to release $X_{\mathbf{x}+\mathbf{y}_{\mathbf{x}}^i}$ molecules in the i -th time slot. According to the propagation model and due to the fact that the input-output relationships between the NMs and the FCs can be modeled as independent Poisson channels with additive Poisson noise [3], at the end of L -th time slot,

¹Inspired by the current experimental implementations of molecular communications [10]–[13], [22], we consider two-level modulation, which will also allow us to cleanly describe the effect of different parameters.

we can model the number of received molecules at the FC at $\mathbf{z} \in \Phi_p$ as

$$Y_L^{\mathbf{z}} \sim \text{Poisson} \left(\lambda_0 T + \sum_{\mathbf{x} \in \Phi_p} \sum_{i=1}^L p_{iL}(\|\mathbf{x} + \mathbf{y}_x^i - \mathbf{z}\|) X_{\mathbf{x} + \mathbf{y}_x^i} \right), \quad (5)$$

where $\lambda_0 T$ is the mean of the number of noise molecules during the time slot. In (5), the effect of ISI due to the transmissions in the previous $i \in \{1, \dots, L-1\}$ -th time slots is also included.

III. DISTANCE DISTRIBUTIONS

In this section, we present the following theorem on the conditional PDF of the distance of any (arbitrary) element in the set $\mathcal{N}^x$ of the cluster centered at $\mathbf{x} \in \Phi_p$ to the origin $\mathbf{o}$ and the subsequent corollary as a special case of the theorem for $\mathcal{N}^o$. These results will be used later in derivations of the expected value of interference in Section IV.

Theorem 1: Conditioned on $\|\mathbf{x}\|$, i.e., the distance of the parent point $\mathbf{x}$ from the origin, the PDF of the distances $d = \|\mathbf{y} + \mathbf{x}\|$, $\forall \mathbf{y} \in \mathcal{N}^x$, is

$$\begin{aligned} f_d(y|\|\mathbf{x}\|) &= \int_{z_1=-y}^y \int_{z_2=-\sqrt{y^2-z_1^2}}^{\sqrt{y^2-z_1^2}} \frac{y}{\sqrt{y^2-z_1^2-z_2^2}} \\ &\times \left[f_{\mathbf{Y}} \left(z_1 - \|\mathbf{x}\|, z_2, \sqrt{y^2-z_1^2-z_2^2} \right) \right. \\ &\left. + f_{\mathbf{Y}} \left(z_1 - \|\mathbf{x}\|, z_2, -\sqrt{y^2-z_1^2-z_2^2} \right) \right] dz_2 dz_1; \quad y \geq 0, \end{aligned} \quad (6)$$

which is specialized for TCP as

$$\begin{aligned} f_d(y|\|\mathbf{x}\|) &= \frac{y}{\sqrt{2\pi}\sigma\|\mathbf{x}\|} \left[\exp \left(-\frac{(y-\|\mathbf{x}\|)^2}{2\sigma^2} \right) \right. \\ &\left. - \exp \left(-\frac{(y+\|\mathbf{x}\|)^2}{2\sigma^2} \right) \right]; \quad y \geq 0. \end{aligned} \quad (7)$$

Proof: See Appendix A. ■

Corollary 1: The PDF of the distances $d = \|\mathbf{y}\|$, $\forall \mathbf{y} \in \mathcal{N}^o$, is

$$\begin{aligned} f_d(y|0) &= \int_{z_1=-y}^y \int_{z_2=-\sqrt{y^2-z_1^2}}^{\sqrt{y^2-z_1^2}} \frac{y}{\sqrt{y^2-z_1^2-z_2^2}} \\ &\times \left[f_{\mathbf{Y}} \left(z_1, z_2, \sqrt{y^2-z_1^2-z_2^2} \right) \right. \\ &\left. + f_{\mathbf{Y}} \left(z_1, z_2, -\sqrt{y^2-z_1^2-z_2^2} \right) \right] dz_2 dz_1; \quad y \geq 0, \end{aligned} \quad (8)$$

which is specialized for TCP as

$$f_d(y|0) = \sqrt{\frac{2}{\pi}} \frac{y^2}{\sigma^3} \exp \left(-\frac{y^2}{2\sigma^2} \right); \quad y \geq 0. \quad (9)$$

Proof: By putting $\|\mathbf{x}\| = 0$ in Theorem 1, we get the results. However, for the TCP result, we achieve $\frac{0}{0}$ and need to use L'Hopital's rule as

$$\lim_{\|\mathbf{x}\| \rightarrow 0} \frac{y}{\sqrt{2\pi}\sigma\|\mathbf{x}\|} \left[\exp \left(-\frac{(y-\|\mathbf{x}\|)^2}{2\sigma^2} \right) \right.$$

$$\begin{aligned} &\left. - \exp \left(-\frac{(y+\|\mathbf{x}\|)^2}{2\sigma^2} \right) \right] = \lim_{\|\mathbf{x}\| \rightarrow 0} y \times \\ &\frac{\frac{y-\|\mathbf{x}\|}{\sigma^2} \exp \left(-\frac{(y-\|\mathbf{x}\|)^2}{2\sigma^2} \right) + \frac{y+\|\mathbf{x}\|}{\sigma^2} \exp \left(-\frac{(y+\|\mathbf{x}\|)^2}{2\sigma^2} \right)}{\sqrt{2\pi}\sigma} \\ &= \sqrt{\frac{2}{\pi}} \frac{y^2}{\sigma^3} \exp \left(-\frac{y^2}{2\sigma^2} \right). \end{aligned} \quad (10)$$

This completes the proof. ■

IV. FUSION CENTER DETECTOR DESIGN

Our analysis will focus on the performance of the typical cluster/FC, which we term as the reference cluster/FC. Thanks to Slivnyak's theorem [16], we simply add the reference FC to the origin along with its cluster $\mathcal{N}^o$. Also, we consider the L -th time slot, where the reference FC desires to detect the information of a reference NM at location $\mathbf{y}_0$ in the reference cluster. Then, from (5), the number of received molecules at the reference FC can be rewritten as

$$Y_L^o \sim \text{Poisson} \left(p_{LL}(\|\mathbf{y}_0\|) X_{\mathbf{y}_0} + \mathcal{I}_L + \lambda_0 T \right), \quad (11)$$

where $\mathcal{I}_L = \mathcal{I}_L^{\text{intra}} + \mathcal{I}_L^{\text{inter}}$ is the total interference. The intra-cluster interference $\mathcal{I}_L^{\text{intra}}$ from NMs inside the reference cluster and inter-cluster interference $\mathcal{I}_L^{\text{inter}}$ from NMs of other clusters are given by

$$\mathcal{I}_L^{\text{intra}} = \sum_{i=1}^{L-1} p_{iL}(\|\mathbf{y}_o^i\|) X_{\mathbf{y}_o^i}, \quad (12)$$

and

$$\mathcal{I}_L^{\text{inter}} = \sum_{\mathbf{x} \in \Phi_p} \sum_{i=1}^L p_{iL}(\|\mathbf{x} + \mathbf{y}_x^i\|) X_{\mathbf{x} + \mathbf{y}_x^i}. \quad (13)$$

Due to the unknown transmitted symbols and locations of NMs, the value of $\mathcal{I}_L$ is unknown to the FCs. However, using the statistical characteristics of the interference $\mathcal{I}_L$, we propose an approximate maximum likelihood (ML) decision rule as

$$\begin{aligned} \hat{X}_{\mathbf{y}_0} &= \arg \max_{j=\{0,1\}} \mathbb{P}(Y_L^o = y \mid X_{\mathbf{y}_0} = x_j) \approx \arg \max_{j=\{0,1\}} \\ &\frac{e^{-(p_{LL}(\|\mathbf{y}_0\|)x_j + \mathbb{E}\{\mathcal{I}_L\} + \lambda_0 T)} (p_{LL}(\|\mathbf{y}_0\|)x_j + \mathbb{E}\{\mathcal{I}_L\} + \lambda_0 T)^y}{y!}, \end{aligned} \quad (14)$$

where the interference $\mathcal{I}_L$ is approximated by its expected value $\mathbb{E}\{\mathcal{I}_L\}$. It can be obtained as

$$\mathbb{E}\{\mathcal{I}_L\} = \mathbb{E}\{\mathcal{I}_L^{\text{intra}}\} + \mathbb{E}\{\mathcal{I}_L^{\text{inter}}\}, \quad (15)$$

where for $\mathbb{E}\{\mathcal{I}_L^{\text{intra}}\}$, we have

$$\begin{aligned} \mathbb{E}\{\mathcal{I}_L^{\text{intra}}\} &= \mathbb{E} \left\{ \sum_{i=1}^{L-1} p_{iL}(\|\mathbf{y}_o^i\|) X_{\mathbf{y}_o^i} \right\} \\ &= \mathbb{E}\{X_{\mathbf{y}_o^1}\} \mathbb{E} \left\{ \sum_{i=1}^{L-1} p_{iL}(\|\mathbf{y}\|) \mid \|\mathbf{y}\| > r_0 \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{x_0 + x_1}{2} \sum_{i=1}^{L-1} \frac{\int_{r_0}^{\infty} p_{iL}(y) f_d(y|0) dy}{\int_{r_0}^{\infty} f_d(y|0) dy} \\
&= \frac{x_0 + x_1}{2 \int_{r_0}^{\infty} f_d(y|0) dy} \int_{r_0}^{\infty} f_d(y|0) \sum_{i=1}^{L-1} p_{iL}(y) dy, \quad (16)
\end{aligned}$$

where the term $\int_{r_0}^{\infty} f_d(y|0) dy = \mathbb{P}(\|\mathbf{y}\| > r_0)$ is for the condition that NMs are outside the reference FC, i.e., the ball with radius r_0 centered at the origin. From (2), $\sum_{i=1}^{L-1} p_{iL}(y) = g(LT, y) - g(T, y)$ in (16). Then, $\mathbb{E}\{\mathcal{I}_L^{\text{inter}}\}$ can be obtained as

$$\begin{aligned}
\mathbb{E}\{\mathcal{I}_L^{\text{inter}}\} &= \mathbb{E}\left\{\sum_{\mathbf{x} \in \Phi_p} \sum_{i=1}^L p_{iL}(\|\mathbf{x} + \mathbf{y}_x^i\|) X_{\mathbf{x} + \mathbf{y}_x^i}\right\} \\
&= \mathbb{E}\{X_{\mathbf{x}_1 + \mathbf{y}_{x_1}^1}\} \mathbb{E}\left\{\sum_{\mathbf{x} \in \Phi_p} \sum_{i=1}^L p_{iL}(\|\mathbf{x} + \mathbf{y}\|) \middle| \|\mathbf{y}\| > r_0\right\} \\
&\stackrel{(a)}{=} \frac{x_0 + x_1}{2} \mathbb{E}\left\{\sum_{\mathbf{x} \in \Phi_p} \sum_{i=1}^L \left(\int_{r_0}^{\infty} p_{iL}(y) \times \frac{f_d(y|\|\mathbf{x}\|)}{\int_{r_0}^{\infty} f_d(y|\|\mathbf{x}\|) dy} dy\right)\right\} \stackrel{(b)}{=} 4\pi\lambda_p \times \frac{x_0 + x_1}{2} \times \\
&\int_{2r_0}^{\infty} \frac{1}{\int_{r_0}^{\infty} f_d(y|x) dy} \sum_{i=1}^L \left(\int_{r_0}^{\infty} p_{iL}(y) f_d(y|x) dy\right) x^2 dx \\
&= 2\pi\lambda_p(x_0 + x_1) \int_{2r_0}^{\infty} \frac{x^2 \int_{r_0}^{\infty} f_d(y|x) \sum_{i=1}^L p_{iL}(y) dy}{\int_{r_0}^{\infty} f_d(y|x) dy} dx, \quad (17)
\end{aligned}$$

where (a) follows from the fact that the distance of each NM in the cluster $\mathbf{x} \in \Phi_p$ to the origin is i.i.d. with distribution $f_d(\cdot|\|\mathbf{x}\|)$ in (6) [17], [19] and the term $\int_{r_0}^{\infty} f_d(y|\|\mathbf{x}\|) dy = \mathbb{P}(\|\mathbf{y} + \mathbf{x}\| > r_0)$ is for the condition that NMs are outside the ball with radius r_0 centered at the origin. Then, (b) follows from the Campbell's theorem for PPPs [16] and the fact that FCs have at least $2r_0$ distance from each other. From (2)-(3), $\sum_{i=1}^L p_{iL}(y) = g(LT, y)$ in (17). For relatively large time slot duration when ISI can be ignored, the expected value of interference is given as in the following lemma.

Lemma 1: For the TCP model and $T \rightarrow \infty$, the expected value of $\mathcal{I}_L$ is simplified to

$$\mathbb{E}\{\mathcal{I}_L\} = 2\pi\lambda_p r_0(x_0 + x_1) e^{\frac{\sigma^2}{2} \frac{\mu}{B} + r_0 \sqrt{\frac{\mu}{B}}} \int_{2r_0}^{\infty} \frac{x A(x)}{B(x)} dx, \quad (18)$$

where

$$\begin{aligned}
A(x) &= e^{-x\sqrt{\frac{\mu}{B}}} \left(1 - \frac{1}{2} \text{erfc}\left(\frac{x - \sigma^2 \sqrt{\frac{\mu}{B}} - r_0}{\sqrt{2}\sigma}\right)\right) \\
&\quad - \frac{1}{2} e^{x\sqrt{\frac{\mu}{B}}} \text{erfc}\left(\frac{x + \sigma^2 \sqrt{\frac{\mu}{B}} + r_0}{\sqrt{2}\sigma}\right), \quad (19)
\end{aligned}$$

and

$$\begin{aligned}
B(x) &= \frac{\sigma}{\sqrt{2\pi}x} \left(\exp\left(-\frac{(x - r_0)^2}{2\sigma^2}\right) - \exp\left(-\frac{(x + r_0)^2}{2\sigma^2}\right)\right) \\
&\quad + 1 - \frac{1}{2} \text{erfc}\left(\frac{x - r_0}{\sqrt{2}\sigma}\right) + \frac{1}{2} \text{erfc}\left(\frac{x + r_0}{\sqrt{2}\sigma}\right). \quad (20)
\end{aligned}$$

Proof: The result is obtained from the following facts:

$$\lim_{t \rightarrow \infty} g(t, y) = \frac{r_0}{y} \exp\left(-\sqrt{\frac{\mu}{D}}(y - r_0)\right), \quad (21)$$

where $g(t, y)$ is the channel function in (4) and

$$\begin{aligned}
&\int_{r_0}^{\infty} x^2 f_d(y|x) \sum_{i=1}^L p_{iL}(y) dy = \frac{r_0 x}{\sqrt{2\pi}\sigma} \times \\
&\int_{r_0}^{\infty} e^{-\frac{(y-x)^2}{2\sigma^2} - \sqrt{\frac{\mu}{D}}(y-r_0)} - e^{-\frac{(y+x)^2}{2\sigma^2} - \sqrt{\frac{\mu}{D}}(y-r_0)} dy = \\
&\frac{r_0 x}{\sqrt{2\pi}\sigma} \int_{r_0}^{\infty} e^{-\frac{(y - (x - \sigma^2 \sqrt{\frac{\mu}{D}}))^2}{2\sigma^2} + \frac{\sigma^2}{2} \frac{\mu}{B} - (x - r_0) \sqrt{\frac{\mu}{D}}} \\
&- e^{-\frac{(y + (x + \sigma^2 \sqrt{\frac{\mu}{D}}))^2}{2\sigma^2} + \frac{\sigma^2}{2} \frac{\mu}{B} + (x + r_0) \sqrt{\frac{\mu}{D}}} dy = \\
&\stackrel{(c)}{=} \frac{r_0 x}{2} e^{\frac{\sigma^2}{2} \frac{\mu}{B} - (x - r_0) \sqrt{\frac{\mu}{D}}} \frac{2}{\sqrt{\pi}} \int_{-\frac{x - \sigma^2 \sqrt{\frac{\mu}{D}} - r_0}{\sqrt{2}\sigma}}^{\infty} e^{-t^2} dt \\
&- \frac{r_0 x}{2} e^{\frac{\sigma^2}{2} \frac{\mu}{B} + (x + r_0) \sqrt{\frac{\mu}{D}}} \frac{2}{\sqrt{\pi}} \int_{\frac{x + \sigma^2 \sqrt{\frac{\mu}{D}} + r_0}{\sqrt{2}\sigma}}^{\infty} e^{-t^2} dt \\
&= \frac{r_0 x}{2} e^{\frac{\sigma^2}{2} \frac{\mu}{B} - (x - r_0) \sqrt{\frac{\mu}{D}}} \left(2 - \text{erfc}\left(\frac{x - \sigma^2 \sqrt{\frac{\mu}{D}} - r_0}{\sqrt{2}\sigma}\right)\right) \\
&- \frac{r_0 x}{2} e^{\frac{\sigma^2}{2} \frac{\mu}{B} + (x + r_0) \sqrt{\frac{\mu}{D}}} \text{erfc}\left(\frac{x + \sigma^2 \sqrt{\frac{\mu}{D}} + r_0}{\sqrt{2}\sigma}\right), \quad (22)
\end{aligned}$$

where (c) comes from the variable changes $t = \frac{y - x + \sigma^2 \sqrt{\frac{\mu}{D}}}{\sqrt{2}\sigma}$ and $t = \frac{y + x + \sigma^2 \sqrt{\frac{\mu}{D}}}{\sqrt{2}\sigma}$. Also, we have the fact

$$\begin{aligned}
&\int_{r_0}^{\infty} f_d(y|x) dy = \frac{1}{\sqrt{2\pi}\sigma x} \int_{r_0}^{\infty} y e^{-\frac{(y-x)^2}{2\sigma^2}} - y e^{-\frac{(y+x)^2}{2\sigma^2}} dy \\
&= \frac{\sigma}{\sqrt{2\pi}x} \int_{r_0}^{\infty} \frac{y - x}{\sigma^2} e^{-\frac{(y-x)^2}{2\sigma^2}} - \frac{y + x}{\sigma^2} e^{-\frac{(y+x)^2}{2\sigma^2}} dy \\
&+ \frac{1}{\sqrt{2\pi}\sigma} \int_{r_0}^{\infty} e^{-\frac{(y-x)^2}{2\sigma^2}} + e^{-\frac{(y+x)^2}{2\sigma^2}} dy \\
&= \frac{\sigma}{\sqrt{2\pi}x} \left(e^{-\frac{(x-r_0)^2}{2\sigma^2}} - e^{-\frac{(x+r_0)^2}{2\sigma^2}}\right) \\
&+ \frac{1}{2} \left(\frac{2}{\sqrt{\pi}} \int_{-\frac{x-r_0}{\sqrt{2}\sigma}}^{\infty} e^{-t^2} dt + \frac{2}{\sqrt{\pi}} \int_{\frac{x+r_0}{\sqrt{2}\sigma}}^{\infty} e^{-t^2} dt\right) \\
&= \frac{\sigma}{\sqrt{2\pi}x} \left(e^{-\frac{(x-r_0)^2}{2\sigma^2}} - e^{-\frac{(x+r_0)^2}{2\sigma^2}}\right) \\
&+ \frac{1}{2} \left(2 - \text{erfc}\left(\frac{x - r_0}{\sqrt{2}\sigma}\right) + \text{erfc}\left(\frac{x + r_0}{\sqrt{2}\sigma}\right)\right). \quad (23)
\end{aligned}$$

From (18), it can be seen that the expected interference linearly increases with λ_p and $\frac{x_0 + x_1}{2}$. ■

TABLE I
PARAMETER VALUES

λ_p	r_0	$\ y_0\ $	D	μ	σ	T	L	$\{x_0, x_1\}$
$2 \times 10^{-6} \mu\text{m}^3$	$5 \mu\text{m}$	$2r_0$	$40 \times 10^{-12} \frac{\text{m}^2}{\text{s}}$	0.1 s^{-1}	$20 \mu\text{m}$	0.5 s	5	$\{0, 40\}$

Assuming the order $x_0 < x_1$, it can be shown that the rule in (14) can be further simplified to

$$\hat{X}_{y_0} = \begin{cases} x_0 & \text{if } y < \text{th}, \\ x_1 & \text{if } y \geq \text{th}, \end{cases} \quad (24)$$

where the threshold is as

$$\text{th} = \frac{p_L(\|y_0\|)x_1 - p_L(\|y_0\|)x_0}{\ln \left(\frac{p_L(\|y_0\|)x_1 + \mathbb{E}\{\mathcal{I}_L\} + \lambda_0 T}{p_L(\|y_0\|)x_0 + \mathbb{E}\{\mathcal{I}_L\} + \lambda_0 T} \right)}. \quad (25)$$

This simple detector has low complexity to be implemented on an NM.

V. NUMERICAL RESULTS

In this section, we provide numerical results for specific scenarios of clustered bio-nanonetworks with the parameter values given in Table 1, unless otherwise stated. We use Monte Carlo simulations to evaluate the error probability of the proposed detector in (24)-(25). For completeness, note that the analytical results for the error probability are available in the expanded version of this article [1].

In Fig. 2, the error probability is shown as a function of the number of time slot L for $D = 10 \times 10^{-12}$ and 40×10^{-12} . It is observed that the error probability increases as the number of time slots increases. It is because the interference from ISI increases. Other observation is related to the increase of the diffusion coefficient D that can significantly decrease the error probability. It is due to the direct effect of increase in D on the increase in the hitting probability p_{LL} in (3) and accordingly the improvement of the link from the reference NM.

The error probability as a function of time slot duration T for $\|y_0\| = 2.5r_0$ and $3r_0$ is studied in Fig. 3. It is observed that as T increases, the error probability decreases. It is due to the fact that the hitting probabilities p_{iL} in (2) decrease, which decreases the ISI. Also, when the reference NM is located closer to the reference FC, the performance improves since p_{LL} increases. In Fig. 4, the expected number of interference molecules in (15) and its Monte Carlo simulation, and the threshold given in (25) are shown as a function of T for $\|y_0\| = 2.5r_0$. It reveals that ignoring exclusion zones of the interfering FCs in the analysis was without any loss of accuracy. Also, we can observe that as T increases, the expected number of interference molecules increases and the threshold follows the increasing trend by unit step function increments. Also, the duration of constant thresholds increases.

In Fig. 5, the error probability as a function of the parent intensity λ_p is plotted for $\sigma = 3r_0$ and $5r_0$. The error probability increases with the increase in λ_p . It is because the number of interfering clusters increases. Also, increasing σ improves the performance. This is intuitive because the probability of the event that the reference FC is located farther from the NMs of the reference cluster increases with σ and this leads to decrease in p_{iL} in (2).

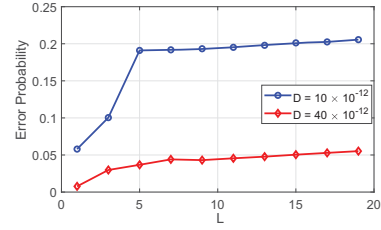


Fig. 2. Error probability as a function of L .

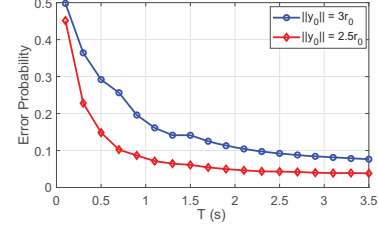


Fig. 3. Error probability as a function of T .

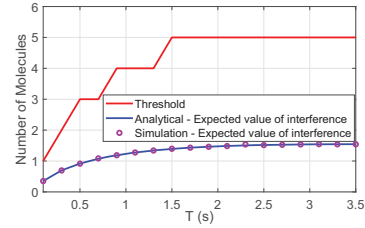


Fig. 4. Number of molecules as a function of T .

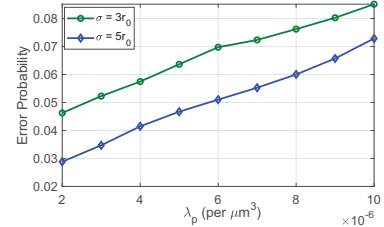


Fig. 5. Error probability as a function of λ_p .

VI. CONCLUSIONS

In this paper, we used tools from stochastic geometry to develop the first comprehensive framework for the modeling and design of MC in bio-nanonetworks whose NMs form clusters around their respective FCs. In order to capture the coupling in the locations of NMs and their respective FCs, we modeled NMs as a PCP with the FCs forming the parent point process. This departs significantly from the known approaches that rely on PPP-based models and often ignore the complex dependence of detectors on interference. For the proposed model, we first characterized the distributions of the distances from a reference FC to various intra- and inter-cluster NMs. We also identified a specific structure for 3D TCPs that provided a remarkably simple expression for the distance distribution. Then, based on the expected value of intra- and inter-cluster interferences, we proposed a simple detector for FCs that is suitable for biological applications. Our analysis revealed that decreasing the intensity of cluster centers or the distance of the NM from the center of its cluster and also increasing the time slot duration improve the error probability.

APPENDIX A
PROOF OF THEOREM 1

Defining $\mathbf{z} = \mathbf{x} + \mathbf{y} \in \mathbb{R}^3$, where $\mathbf{z} = (z_1, z_2, z_3)$ and $\mathbf{x} = (x_1, x_2, x_3)$, the conditional cumulative distribution function (CDF) of the distance $d = \|\mathbf{z}\|$ with realization $y = \sqrt{z_1^2 + z_2^2 + z_3^2}$ is

$$\begin{aligned} \mathbb{P}(d < y \mid \mathbf{x}) &= \int_{z_1=-y}^y \int_{z_2=-\sqrt{y^2-z_1^2}}^{\sqrt{y^2-z_1^2}} \int_{z_3=-\sqrt{y^2-z_1^2-z_2^2}}^{\sqrt{y^2-z_1^2-z_2^2}} f_{\mathbf{Y}}(z_1 - x_1, z_2 - x_2, z_3 - x_3) dz_3 dz_2 dz_1 \\ &\stackrel{(a)}{=} \int_{z_1=-y}^y \int_{z_2=-\sqrt{y^2-z_1^2}}^{\sqrt{y^2-z_1^2}} \int_{z_3=-\sqrt{y^2-z_1^2-z_2^2}}^{\sqrt{y^2-z_1^2-z_2^2}} f_{\mathbf{Y}}(z_1 - \|\mathbf{x}\|, z_2, z_3) dz_3 dz_2 dz_1 = \mathbb{P}(d < y \mid \|\mathbf{x}\|) \\ &= F_d(y \mid \|\mathbf{x}\|), \end{aligned} \quad (26)$$

where (a) is due to the fact that $f_{\mathbf{Y}}(\mathbf{y})$ is rotationally invariant. It is notable that the result of (26) is dependent on the norm of $\mathbf{x}$. Then, by taking a derivative of the CDF $F_d(y \mid \|\mathbf{x}\|)$, the conditional PDF $f_d(y \mid \|\mathbf{x}\|)$ is obtained with the help of the Leibniz integral rule [24] and simplifications.

For the special case of TCP, by substituting the following $f_{\mathbf{Y}}$ from (1) as

$$f_{\mathbf{Y}}(z_1 - \|\mathbf{x}\|, z_2, z_3) = \frac{1}{(2\pi)^{\frac{3}{2}} \sigma^3} \exp\left(-\frac{(z_1 - \|\mathbf{x}\|)^2 + z_2^2 + z_3^2}{2\sigma^2}\right), \quad (27)$$

into (6), we obtain

$$\begin{aligned} f_d(y \mid \|\mathbf{x}\|) &= \frac{2y}{(2\pi)^{\frac{3}{2}} \sigma^3} \exp\left(-\frac{y^2 + \|\mathbf{x}\|^2}{2\sigma^2}\right) \int_{z_1=-y}^y \\ &\exp\left(\frac{\|\mathbf{x}\|z_1}{\sigma^2}\right) \int_{z_2=-\sqrt{y^2-z_1^2}}^{\sqrt{y^2-z_1^2}} \frac{1}{\sqrt{y^2-z_1^2-z_2^2}} dz_2 dz_1, \end{aligned} \quad (28)$$

where $\int_{z_2=-\sqrt{y^2-z_1^2}}^{\sqrt{y^2-z_1^2}} \frac{1}{\sqrt{y^2-z_1^2-z_2^2}} dz_2 = \pi$. Then, we obtain

$$\begin{aligned} f_d(y \mid \|\mathbf{x}\|) &= \frac{y}{\sqrt{2\pi}\sigma\|\mathbf{x}\|} \exp\left(-\frac{y^2 + \|\mathbf{x}\|^2}{2\sigma^2}\right) \times \\ &\left[\exp\left(\frac{y\|\mathbf{x}\|}{\sigma^2}\right) - \exp\left(-\frac{y\|\mathbf{x}\|}{\sigma^2}\right) \right], \end{aligned} \quad (29)$$

which leads to the final result.

ACKNOWLEDGMENT

H. S. Dhillon and L. Tassiulas acknowledge the support of the US National Science Foundation Grant ECCS-2030215 and Grant 2128530 RINGS, respectively.

REFERENCES

- [1] S. M. Azimi-Abarghouyi, H. S. Dhillon, and L. Tassiulas, "Fundamentals of clustered molecular nanonetworks," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, under review, available online: arxiv.org/abs/2208.14213.
- [2] I. F. Akyildiz, J. M. Jornet, and M. Pierobon, "Nanonetworks: A new frontier in communications," *Commun. of the ACM*, vol. 54, no. 11, pp. 84-89, Nov. 2011.
- [3] H. Arjmandi, A. Gohari, M. Nasiri Kenari, and F. Bateni, "Diffusion-based nanonetworking: A new modulation technique and performance analysis," *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 645-648, Apr. 2013.
- [4] H. Arjmandi, M. Movahednasab, A. Gohari, M. Mirmohseni, M. Nasiri-Kenari, and F. Fekri, "ISI-avoiding modulation for diffusion-based molecular communication," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 3, no. 1, pp. 48-59, Mar. 2017.
- [5] I. F. Akyildiz, M. Pierobon, S. Balasubramaniam, and Y. Koucheryavy, "The internet of bio-nano things," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 32-40, Mar. 2015.
- [6] T. N. Cao, N. Zlatanov, P. L. Yeoh, and J. S. Evans, "Optimal detection interval for absorbing receivers in molecular communication systems with interference," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 6, no. 3, pp. 184-198, Dec. 2020.
- [7] A. Noel, K. C. Cheung, and R. Schober, "A unifying model for external noise sources and ISI in diffusive molecular communication," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 12, pp. 2330-2343, Dec. 2014.
- [8] Z. Cheng, Y. Zhang, and M. Xia, "Performance analysis of diffusive mobile multiuser molecular communication with drift," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 4, no. 4, pp. 237-247, Dec. 2018.
- [9] M. Pierobon and I. F. Akyildiz, "A Statistical-physical model of interference in diffusion-based molecular nanonetworks," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 2085-2095, Jun. 2014.
- [10] E. Dinc and O. B. Akan, "Theoretical limits on multiuser molecular communication in internet of nano-bio things," *IEEE Trans. Nanobiosci.*, vol. 16, no. 4, pp. 266-270, Jun. 2017.
- [11] Y. Deng, A. Noel, W. Guo, A. Nallanathan, and M. ElKashlan, "Analyzing large-scale multiuser molecular communication via 3-D stochastic geometry," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 3, no. 2, pp. 118-133, Jun. 2017.
- [12] M. B. Dissanayake, Y. Deng, A. Nallanathan, M. ElKashlan, and U. Mitra, "Interference mitigation in large-scale multiuser molecular communication," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4088-4103, Jun. 2019.
- [13] N. V. Sabu and A. K. Gupta, "Analysis of diffusion based molecular communication with multiple transmitters having individual random information bits," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 5, no. 3, pp. 176-188, Dec. 2019.
- [14] F. Zabini, "Spatially distributed molecular communications via diffusion: Second-order analysis," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 5, no. 2, pp. 112-138, Nov. 2019.
- [15] R. Mosayebi, V. Jamali, N. Ghoroghchian, R. Schober, M. Nasiri-Kenari, and M. Mehrabi, "Cooperative abnormality detection via diffusive molecular communications," *IEEE Trans. Nanobiosci.*, vol. 16, no. 8, pp. 828-842, Dec. 2017.
- [16] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.
- [17] S. M. Azimi-Abarghouyi, B. Makki, M. Haenggi, M. Nasiri-Kenari, and T. Svensson, "Stochastic geometry modeling and analysis of single- and multi-cluster wireless networks," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4981-4996, Oct. 2018.
- [18] S. M. Azimi-Abarghouyi, M. Nasiri-Kenari, and M. Debbah, "Stochastic design and analysis of user-centric wireless cloud caching networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 7, pp. 4978-4993, Jul. 2020.
- [19] M. Afshang and H. S. Dhillon, "Poisson cluster process based analysis of HetNets with correlated user and base station locations," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2417-2431, Apr. 2018.
- [20] M. Afshang, C. Saha, and H. S. Dhillon, "Nearest-neighbor and contact distance distributions for Thomas cluster process," *IEEE Wireless Commun. Lett.*, vol. 6, no. 1, pp. 130-133, Feb. 2017.
- [21] M. Afshang, C. Saha, and H. S. Dhillon, "Nearest-neighbor and contact distance distributions for Matrn cluster process," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2686-2689, Dec. 2017.
- [22] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Modulation techniques for communication via diffusion in nanonetworks," *IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, Jun. 2011.
- [23] A. C. Heren, H. B. Yilmaz, C. B. Chae, and T. Tugcu, "Effect of degradation in molecular communication: Impairment or enhancement?," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 1, no. 2, pp. 217-229, Jun. 2015.
- [24] P. J. Olver, *Applications of Lie Groups to Differential Equations*. New York, NY, USA: Springer, 1986.