

Reception Modeling of Sphere-to-Sphere Molecular Communication via Diffusion

Gaye Genc^{a,*}, Yunus Emre Kara^b, Tuna Tugcu^a, Ali Emre Pusane^c

^a*NETLAB, Bogazici University Department of Computer Engineering, Istanbul, Turkey*

^b*PILAB, Bogazici University Department of Computer Engineering, Istanbul, Turkey*

^c*WCL, Bogazici University Department of Electrical and Electronics Engineering, Istanbul, Turkey*

Abstract

As a widely accepted information transfer method in the nanonetworking domain, molecular communication via diffusion (MCvD) presents many advantages as well as challenges. In order to assess the capabilities and restrictions of MCvD, a thorough understanding of the reception process through the first passage time distribution holds utmost importance. As the network setup becomes more realistic, analytical derivations become increasingly difficult. Using statistical methods on empirical data is a remedy to this challenge. In this paper, we propose two novel heavy-tail distributions, which are well-equipped to model the first passage time distribution for a reflective sphere transmitter and a fully absorbing sphere receiver pair. We present their modeling power using the Kolmogorov-Smirnov goodness of fit test and how the modeling performance behaves under diverse deployment parameters. We also discuss the probability of molecule absorption, signal-to-interference ratio, and the advantages of using a reflective sphere transmitter.

Keywords: Molecular communications, nanonetworks, sphere transmitter, reception modeling, absorption probability, first passage time, heavy-tail distributions

*Corresponding author

Email addresses: gaye.genc@boun.edu.tr (Gaye Genc), yunus.kara@boun.edu.tr (Yunus Emre Kara), tugcu@boun.edu.tr (Tuna Tugcu), ali.pusane@boun.edu.tr (Ali Emre Pusane)

1. Introduction

Nanotechnology, as a key technology with a variety of current and potential future applications, deals with matter in the atomic and the molecular scale [1]. Nanomachines are used to describe devices ranging in size from 0.1 μm to 10 μm and constructed of nano or molecular scale components [2]. Apart from the man-made nanomachines, bioengineered cells that are programmed for a specific task and have basic signaling capabilities [3] are also regarded as nanomachines from this study's point of view.

Operating at the nano-scale is expected to require high cooperation among multiple devices to make an impact on the macro scale. Clusters composed of such machines or cells cooperating with each other enable the realization of complex applications such as health monitoring, tissue engineering, biomedicine, nanomedicine, and environment monitoring [4, 5]. Therefore, communication between these nodes is of high importance and the communication in that scale has different characteristics. Nanonetworking is a rapidly growing area of research due to the high level of collaboration needed by these devices. Specifically, it deals with the communication between nano and/or micro scale machines that has at least one component in the nano scale (up to 100nm according to the definition of IEEE P1906.1 [6]), and controlled or engineered by humans.

Having ties to nanotechnology, biotechnology, and communication technology, molecular communication is an ever-growing interdisciplinary research area in the nanonetworking domain [4]. Molecular communication via diffusion (MCvD) focuses on micro- and nanomachines communicating through molecules emitted into a viscous communication medium. The infrastructureless nature and the propagation of molecules by free diffusion makes MCvD a very effective and energy efficient method of communication. The emitted molecules, called messenger molecules (MM), are the main instrument of information transmission through the environment. After emission, the MMs roam the communication medium according to the laws of free diffusion and the physical characteristics of the channel. Some of the MMs hit the receiver and the properties (type, amount,

concentration, etc.) of the received molecules define the received signal.

From the communications perspective, one of the most critical components of assessing the capabilities and restrictions of MCvD is understanding the reception process. Many works in the literature define the reception process as the absorption and removal of the MMs from the communication environment
35 the absorption and removal of the MMs from the communication environment [7, 8, 9, 10, 11]. In this case, MMs can only contribute to the received signal once. The duration from which an MM is released into the environment up to its removal from the environment upon contact with the receiver is called the first passage time. In the literature, examination of the reception process considering the first passage starts with the 1-D case [7, 8]. In [9], the first passage
40 time is investigated in a 1-D environment with drift. In both [7] and [9], the first passage time in the 1-D medium with drift is shown to follow an inverse Gaussian distribution. Using the inverse Gaussian and Lévy distributions for the first passage time is also a common practice in studying the capacity of
45 molecular timing channels [12, 13].

The first passage process in 3-D is more complex than the 1-D case; thus, studying it in a tractable manner requires making several assumptions on transmitter or channel properties. Two common assumptions are considering point transmitters and spherical receivers. In [10], the authors derive the expected
50 number of absorbed MMs in an interval for the point transmitter and fully absorbing spherical receiver case. Another work under the same geometrical assumptions considers the case where the receiver surface has receptors [14]. The cases where the MMs degrade or the communication medium contains enzymes that neutralize the MMs are also investigated [15, 16, 17].

It is also possible to consider passive receivers in 3-D where the MMs diffuse
55 freely in and out of the receiver body, thus contributing to the received signal more than once. The point transmitter and passive spherical setting is investigated in [18] both with and without the drift component. Additionally, the same environment with enzymes is also investigated [19].

In this paper, we consider a reflecting spherical transmitter and fully absorb-
60 ing spherical receiver. This setting is more realistic than the point transmitter

case in terms of molecular communication since the MMs will not appear out of a singularity in the communication medium, but rather be emitted from a transmitting body that is capable of MM production. Furthermore, using a reflective transmitter has an advantage of providing directivity gain [20]. Recent studies
65 about the first passage time in the literature have been carried out under the spherical transmitter assumption [20, 21, 22, 23, 24]. In [24], a very simplistic method is proposed to model the reception process for the reflective spherical transmitter case using machine learning. Reflecting spherical transmitters are
70 considered together with inter-symbol interference issues in [21, 22]. An analytical study for the first passage time is carried out in [23] for a passive spherical transmitter (i.e. the MMs diffuse freely in and out of the transmitter.)

As of yet, no analytical solution exists for the first passage time distribution in sphere-to-sphere MCvD for a reflective transmitter and an absorbing receiver.
75 The methodology presented in [10] for the point transmitter and absorbing spherical receiver is not generalizable to the sphere-to-sphere setting. This is due to the fact that the authors in [10] use the radial symmetry about the receiver body in their derivations, which cannot be adapted to our setting due to the lack of symmetry originating from the reflective transmitter.

80 In this paper, we aim to overcome the difficulty of the analytical derivation approach by modeling the first passage time distribution using a statistical approach. The contributions of this paper can be summarized as follows:

- We investigate suitable parametric distribution alternatives to represent the first passage time distribution for sphere-to-sphere MCvD.
- 85 • We derive and introduce two new heavy-tail distributions, namely generalized beta-generated inverse of generalized gamma (GBIGG) and Kummer beta-generated inverse of generalized gamma (KBIGG), to the communications domain.
- The distributions that we introduce serve as efficient tools for analyzing
90 the first passage time probability in sphere-to-sphere MCvD. Once the

parameters have been evaluated for the given environment, the first passage time probability can be represented with a few parameters, instead of storing empirical first passage time distributions or running simulations again.

- 95 • Using classical error metrics, we test the modeling performance of these distributions under diverse and challenging scenarios by comparing them against empirical densities obtained from extensive simulations.
- In addition, we rigorously test and affirm the modeling success of our proposed distributions using the powerful Kolmogorov-Smirnov goodness
100 of fit test.
- The approach that we describe in this paper enables us to investigate the probability of absorption and signal-to-interference ratio, both of which could not be calculated using simulations otherwise.
- We emphasize the advantages of using a reflective sphere transmitter over
105 the point transmitter.
- The distributions that we propose in this work make it easier to conduct further research such as inter-symbol interference, modulation, and channel capacity for sphere-to-sphere MCvD.

The remainder of this paper is organized as follows: In Section 2, we present
110 the sphere-to-sphere MCvD system model, and explain our motivation for choosing to work with a reflective sphere transmitter. In Section 3, we present the candidate distributions for modeling the first passage time distribution of sphere-to-sphere MCvD, along with two novel heavy-tail distributions. We present and validate the modeling performance of the candidate distributions in Section 4,
115 followed by a preliminary network performance analysis. We conclude with a summary of key observations and future directions in Section 5.

2. Sphere-to-Sphere Molecular Communication via Diffusion System Model

We model a communication system composed of a fluid environment and a pair of devices, each called Nanonetworking-enabled Node (NeN); one as the transmitter and the other as the receiver. The NeNs are inspired by the cells in the sense that they are able to produce energy by converting raw materials in their surroundings. In this work, we assume that the NeNs have the basic functionalities necessary for communication and are able to use part of the produced energy for communication purposes.

In MCvD, the information transfer is achieved between the transmitter and the receiver via the diffusion-based propagation of specific MMs [25]. The MMs can be chosen as a specific type of protein, peptide, DNA sequence, or other molecular structure.

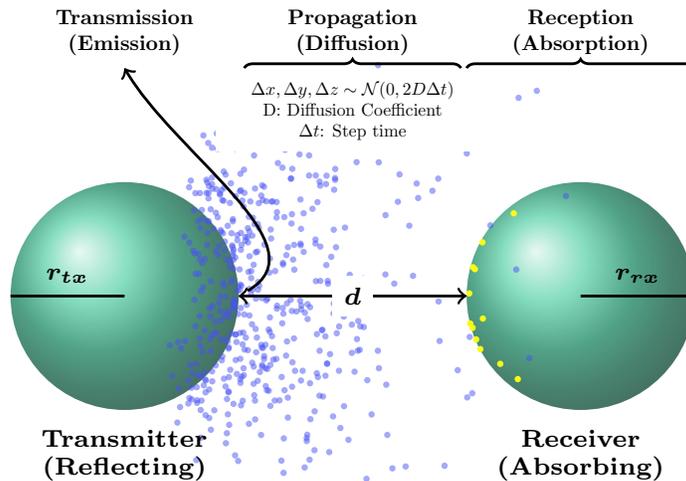


Figure 1: MCvD system process model. The visual is constructed based on real diffusion simulation data from our particle tracking based simulation tool. Received molecules are colored yellow and molecules in transit are colored blue. In this setting, diffusion coefficient $D=75 \mu\text{m}^2/\text{s}$, transmitter NeN radius $r_{tx}=10 \mu\text{m}$, receiver NeN radius $r_{rx}=10 \mu\text{m}$, distance $d=20 \mu\text{m}$, and the simulation snapshot is taken at $t=0.6 \text{ s}$.

The MCvD system is composed of five main processes: encoding, emission,

propagation, reception, and decoding. As we show in Figure 1, emission, propagation, and the reception processes are in the focus of this paper.

Both the transmitter and the receiver are assumed to have spherical bodies, with radii denoted as r_{tx} and r_{rx} , respectively. In our work, we consider a
 135 reflective transmitter which upon emission of MMs, does not absorb them back. The MMs are emitted from the region closest to the receiver. In contrast, the receiver is capable of absorbing MMs such that whenever an MM contacts the body of the receiver, the molecule is received by being absorbed from the environment. The transmitter and the receiver NeNs are deployed d μm apart
 140 in a 3-D no-drift liquid medium that has viscosity and temperature similar to blood.

2.1. Messenger Molecule Propagation via Brownian Motion

Propagation process consists of the free diffusion of MMs in the molecular scale when the environment does not have any drift. In this scale, the movement of particles inside a fluid is modeled by the Brownian motion. We do not consider the collisions between MMs for the sake of simplicity. In a 1-D space, the displacement of a single MM in a unit time is a random variable Δx , which follows a normal distribution with zero mean and σ^2 variance

$$\Delta x \sim \mathcal{N}(\Delta x; 0, \sigma^2) \quad (1)$$

where $\sigma = \sqrt{2D\Delta t}$ and D is the diffusion coefficient that describes the tendency of the propagating molecules to diffuse through the fluid [26].

In our model, the particles propagate through an unbounded 3-D environment without drift. This movement can be modeled as three independent displacements (one for each dimension) [27] and the total displacement, \vec{r} , in one time step can be found as

$$\vec{r} = (\Delta x, \Delta y, \Delta z). \quad (2)$$

145 *2.2. Motivation: The Advantage of the Reflecting Sphere Transmitter Over the*
Point Transmitter

Researchers often operate under simplifying assumptions, and in many cases these assumptions are crucial to initiate the in-depth analysis of an emerging field. However, as the nanonetworking and molecular communication domains start to gain significant popularity, it is also necessary to target more realistic scenarios. In this aspect, evolving the point transmitter assumption into a sphere transmitter setting is an appropriate step. The additional reflectivity feature for the sphere transmitter complicates the work to be done, but our motivation is that it also brings a significant advantage from a communications point-of-view.

When a point transmitter is used, the MMs are dispersed in all directions. Many of these MMs go in the opposite direction from the receiver. Some MMs never reach the receiver or they reach too late and cause inter-symbol interference. In our setting, the reflecting transmitter emits the MMs from its closest point to the receiver, which is shown to work best in terms of mean first passage time and probability of absorption [28]. Most of the reflections occur just after transmission, when the MMs are still close to the transmitter surface facing the receiver. The reflective nature of the transmitter biases the total movement of the MMs towards the receiver.

In Figure 2, we present the signal boost advantage of using a reflective sphere transmitter over a point transmitter. The number of received MMs for the reflecting sphere case attains a sharper peak and has a slightly lighter tail compared to the point transmitter case. This phenomenon is the direct result of the increased MM absorption probability. As a consequence, it is possible to emit a lower number of MMs and still attain a similar number of absorbed MMs as the point transmitter case. This results in a more budget-friendly consumption of MMs, especially if there are energy constraints on the transmitter NeN regarding the production of MMs.

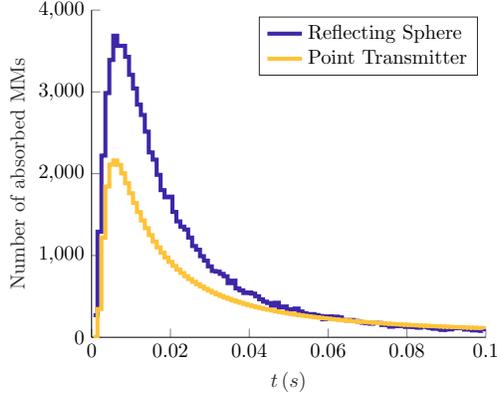


Figure 2: The advantage of using a reflecting sphere transmitter over a point transmitter is evident in the number of absorbed MMs ($r_{tx}=10\ \mu\text{m}$, $r_{rx}=10\ \mu\text{m}$, $d=2\ \mu\text{m}$, $D=125\ \mu\text{m}^2/\text{s}$, 100 000 MMs are sent.)

3. Modeling the First Passage Time Distribution and Absorption Probabilities for Sphere-to-Sphere MCvD

175

Arising from the probabilistic dynamics of Brownian motion, the MMs move randomly. During this random movement, it is possible to miss and never coincide with the receiver NeN since the first passage process is not recurrent in 3-D environment [29]. This means that even if we wait infinitely, some MMs will never reach the receiver NeN. From our point of view, successful communication 180 relies on the absorption of MMs. Therefore, we signify an absorbed MM as “successful” and denote its absorption probability as $p_a < 1$. Unsuccessful MMs are those that evade the absorption process with probability $1 - p_a$. Due to its simple nature, MM absorption is a Bernoulli process. The parameter of 185 the Bernoulli distribution can be easily estimated using the sample mean, i.e. by calculating the ratio of received MMs to the emitted MMs.

In the context of this work, the *first passage time* refers to the time elapsed from the release of an MM from the transmitter NeN until the time the MM is absorbed by the receiver NeN. First passage time is an improper random variable in the sense that its probability density function integrates to $p_a < 1$. However, the first passage time of the **received** MMs is a well-defined random

variable with the conditional density

$$f(t) = \frac{f^*(t)}{p_a} \tag{3}$$

where $f^*(t)$ is the (improper) first passage time distribution of all MMs and p_a is the probability of absorption. In the remainder of this paper, whenever we mention the “first passage time distribution”, we refer to the **conditional** density introduced above. First passage time probability is affected mainly by the diffusion coefficient, transmitter-receiver properties, and the distance between the transmitter and the receiver.

In this section, our main focus is on finding a suitable parametric probability distribution that best describes the first passage time distribution ($f(t)$ in (3)) of received MMs for sphere-to-sphere MCvD. We also need to find the absorption probability (p_a in (3)) of an MM since some MMs never arrive at the receiver and the first passage time is defined only for absorbed MMs. In the remainder of this section, we start by elaborating on the point-to-sphere special case of the first passage distribution. Then, we consider the generalized sphere-to-sphere communication by discussing several candidate distributions and introducing two new distributions. We conclude the section by presenting an estimator for the absorption probability.

3.1. Special Case: Point-to-Sphere MCvD

For a point transmitter ($r_{tx} = 0$) and a sphere receiver ($r_{rx} > 0$) in 3-D environment, the hitting rate $n_{hit}(t)$ of the MMs to the receiver is formulated in [10] as

$$n_{hit}(t) = \frac{r_{rx}}{d + r_{rx}} \frac{d}{\sqrt{4\pi Dt^3}} e^{-\frac{d^2}{4Dt}}. \tag{4}$$

In this context, the hitting rate $n_{hit}(t)$ refers to the improper first passage time distribution of all MMs. When inspected closely, (4) is actually a scaled version of the inverse gamma (IG^1) distribution. Inverse gamma is a two-parameter

¹The abbreviation IG is sometimes used to indicate the inverse Gaussian distribution. The reader should note that throughout this paper the abbreviation IG only refers to the inverse gamma distribution.

distribution with support $t > 0$, shape parameter $\alpha > 0$, and scale parameter $\theta > 0$. The probability density function of the inverse gamma distribution is

$$f_{IG}(t; \alpha, \theta) = \frac{1}{\theta \Gamma(\alpha)} (\theta/t)^{\alpha+1} e^{-\theta/t}, \quad (5)$$

where $\Gamma(\cdot)$ is the gamma function. The cumulative distribution function of the inverse gamma distribution is also referred to as the regularized gamma function and has the form

$$F_{IG}(t; \alpha, \theta) = \frac{\Gamma(\alpha, \theta/t)}{\Gamma(\alpha)}, \quad (6)$$

where the numerator $\Gamma(\cdot, \cdot)$ is the upper incomplete gamma function. Thus, we can represent (4) as

$$\underbrace{n_{hit}(t)}_{f^*(t)} = \underbrace{\frac{r_{rx}}{d + r_{rx}}}_{p_a} \underbrace{f_{IG}\left(t; \frac{1}{2}, \frac{d^2}{4D}\right)}_{f(t)}, \quad (7)$$

where $f^*(t)$ is the improper first passage time distribution, p_a is the absorption probability of an MM, and $f(t)$ is the conditional first passage time distribution of the received MMs as introduced in (3) ².

Our aim is to find both p_a and $f(t)$ that work not only for the point transmitter ($r_{tx} = 0$), but also for sphere transmitters as well (i.e. $r_{tx} \geq 0$). Thus, as the baseline distribution for $f(t)$, we start by relaxing both shape ($\alpha = 1/2$) and scale ($\theta = d^2/4D$) parameter values defined in (7) and fit the inverse gamma distribution described in (5) and (6). However, in Section 4, we show that even this relaxed inverse gamma distribution is not adequate for modeling the first passage time distribution $f(t)$ for sphere-to-sphere MCvD. In the upcoming subsections, we introduce candidate distribution families which are generalizations of the inverse gamma distribution.

²Note that the inverse gamma distribution with shape parameter $\alpha = 1/2$ and scale parameter θ is equivalent to the Lévy distribution with location parameter 0 and scale parameter 2θ , i.e. $f_{IG}(t; 1/2, \theta) \sim f_{Levy}(t; 0, 2\theta)$. The Lévy distribution is also considered for the first passage time probability in the literature [13].

3.2. IGG: Inverse of Generalized Gamma

In 1962, the Generalized Gamma distribution was introduced by Stacy [30]. Rewriting the generalized gamma distribution for the random variable $1/t$ results in the inverse of generalized gamma distribution (IGG). The probability density function of the IGG distribution is

$$f_{IGG}(t; \alpha, \theta, \beta) = \frac{\beta}{\theta \Gamma(\alpha)} (\theta/t)^{\alpha\beta+1} e^{-(\theta/t)^\beta} \quad (8)$$

where α and θ are same as in IG distribution, and $\beta > 0$. The cumulative distribution function of the inverse of the Generalized Gamma distribution is

$$F_{IGG}(t; \alpha, \theta, \beta) = \frac{\Gamma\left(\alpha, (\theta/t)^\beta\right)}{\Gamma(\alpha)}. \quad (9)$$

The inverse of generalized gamma distribution becomes the IG distribution when $\beta = 1$:

$$f_{IGG}(t; \alpha, \theta, 1) = f_{IG}(t; \alpha, \theta).$$

3.3. GBIGG: Generalized beta-generated inverse of generalized gamma

Similar to the IG distribution, we find the IGG distribution not satisfactory enough for modeling first passage time probability of sphere-to-sphere MCvD. We base this claim on the results shown in Section 4. To remedy this problem, we propose a new six parameter distribution named generalized beta-generated inverse of generalized gamma distribution. Before presenting the PDF and CDF of this distribution, we discuss the generalization method that we use in detail.

In the statistics literature, there are several methods for generalizing probability distributions by introducing additional parameters that control tail weights, skewness, kurtosis, etc. [31, 32, 33, 34]. The family of generalized beta-generated distributions introduced by Alexander *et al.* [31] is one of them.

In 1984, McDonald introduced the generalized beta distribution of the first kind [35], which can be characterized by its density as

$$f_{GB}(u; a, b, c) = cB(a, b)^{-1} u^{ac-1} [1 - u^c]^{b-1} \quad (10)$$

where $0 < u < 1$, $a > 0$, $b > 0$, $c > 0$, and $B(a, b)$ denotes the beta function.

Alexander *et al.* use the generalized beta distribution for introducing a new family of distributions, namely generalized beta-generated distributions [31]. Given a probability density function $g(x; \tau)$ with support S (i.e. $x \in S$), parameter vector τ , and its cumulative distribution function $G(x; \tau)$, the *generalized beta-generated g distribution* has the probability density function of the following form

$$f_{GBg}(x; \tau, a, b, c) = cB(a, b)^{-1}g(x; \tau)G(x; \tau)^{ac-1}[1 - G(x; \tau)^c]^{b-1} \quad (11)$$

where $x \in S$, a , b and c are shape parameters. Its cumulative distribution function is

$$F_{GBg}(x; \tau, a, b, c) = I(G(x; \tau)^c; a, b) = B(a, b)^{-1} \int_0^{G(x; \tau)^c} w^{a-1}(1-w)^{b-1} dw \quad (12)$$

where $I(x; a, b) = B(a, b)^{-1} \int_0^x w^{a-1}(1-w)^{b-1} dw$ denotes the regularized incomplete beta function.

230

We propose a new six-parameter distribution by incorporating inverse of generalized gamma distribution (IGG) into the generalized beta-generator, utilizing (11) and (12). We call this new distribution the *generalized beta-generated inverse of generalized gamma*, abbreviated as GBIGG. The probability density function of GBIGG is

$$f_{GBIGG}(t; \alpha, \theta, \beta, a, b, c) = \frac{\beta c B(a, b)^{-1} (\theta/t)^{\alpha\beta+1} e^{-(\theta/t)^\beta} \Gamma(\alpha, (\theta/t)^\beta)^{ac-1}}{\theta \Gamma(\alpha)^{c(a+b-1)} \left(\Gamma(\alpha)^c - \Gamma(\alpha, (\theta/t)^\beta)^c \right)^{1-b}} \quad (13)$$

where $t \in (0, \infty)$.

The cumulative distribution function of GBIGG is

$$F_{GBIGG}(t; \alpha, \theta, \beta, a, b, c) = I\left(\frac{\Gamma(\alpha, (\theta/t)^\beta)^c}{\Gamma(\alpha)^c}; a, b\right) \quad (14)$$

where $I(x; a, b)$ denotes the regularized incomplete beta function.

Note that, generalized beta-generated distributions reduce to beta-generated distributions [32] when $c = 1$ and to Kumaraswamy-generated distributions [33]

235 when $a = 1$. In addition, both beta-generated and Kumaraswamy-generated
distributions reduce to exponentiated distributions when $b = 1$. In Table 1, we
summarize these special cases of the GBIGG distribution. Note that GBIGG is
a rather complex distribution. It is always a good practice to check whether the
modeling performance of simpler distributions are satisfactory. Thus, in addi-
240 tion to the GBIGG distribution, we also investigate the modeling performance
of its special cases.

GBIGG	Generalized beta-generated IGG	α	θ	β	a	b	c
BIGG	Beta-generated IGG	α	θ	β	a	b	1
EIGG	Exponentiated IGG	α	θ	β	a	1	1
KwIGG	Kumaraswamy IGG	α	θ	β	1	b	c
EIGG	Exponentiated IGG (alt. expression)	α	θ	β	1	1	c
IGG	Inverse of generalized gamma	α	θ	β	1	1	1

Table 1: Some special cases of the GBIGG distribution.

3.4. KBIGG: Kummer beta-generated inverse of generalized gamma

Arising from the dynamics of Brownian motion, the first passage time dis-
tribution in MCvD has a very heavy tail [11, 21]. Additionally, in the set-
245 tings where the communicating pair of NeNs are closer, MMs are absorbed very
quickly. This results in a first passage time distribution skewed to the left ex-
treme. Due to this nature of Brownian motion and MCvD, it is beneficial to
employ a generalization method suitable to remedy these issues. The Kummer
beta-generator introduced in [34] has the advantage of modeling heavy-tailed
250 distributions and it offers flexibility to the left and right extremes. Thus, we
use this generator to propose another new six-parameter distribution named
Kummer beta-generated inverse of generalized gamma. Before presenting the
PDF and the CDF of this distribution, we discuss the generalization method
that we use in detail.

In 1995, Ng and Kotz proposed the Kummer beta distribution [36] with

probability density and cumulative distribution functions

$$f_{KB}(u; a, b, c) = Ku^{a-1}(1-u)^{b-1}e^{-cu}, \quad (15)$$

$$F_{KB}(u; a, b, c) = K \int_0^u w^{a-1}(1-w)^{b-1}e^{-cw} dw, \quad (16)$$

respectively, where $0 < u < 1$, $a > 0$, $b > 0$, and $-\infty < c < \infty$. The regularization factor K is defined as

$$K = \left(\int_0^1 w^{a-1}(1-w)^{b-1}e^{-cw} dw \right)^{-1} \quad (17)$$

$$K = {}_1F_1(a; a+b; -c)^{-1} B(a, b)^{-1} \quad (18)$$

where $B(\cdot, \cdot)$ is the beta function and

$${}_1F_1(a; a+b; -c) = B(a, b)^{-1} \int_0^1 w^{a-1}(1-w)^{b-1}e^{-cw} dw \quad (19)$$

255 is the confluent hypergeometric function [37].

Based on the Kummer beta distribution, Pescim *et al.* propose the *Kummer beta-generated g distribution* [34] with probability density and cumulative distribution functions

$$f_{KBg}(x; \tau, a, b, c) = Kg(x; \tau)G(x; \tau)^{a-1}(1-G(x; \tau))^{b-1}e^{-cG(x; \tau)}, \quad (20)$$

$$F_{KBg}(x; \tau, a, b, c) = K \int_0^{G(x; \tau)} w^{a-1}(1-w)^{b-1}e^{-cw} dw, \quad (21)$$

respectively, where $g(x; \tau)$ and $G(x; \tau)$ are the probability density and cumulative distribution functions of the input g distribution with parameters τ .

Using the Kummer beta-generator, we introduce a new six-parameter distribution based on inverse of generalized gamma distribution. We call this new distribution the *Kummer beta-generated inverse of generalized gamma*, abbreviated as KBIGG. The probability density function of KBIGG is

$$f_{KBIGG}(t; \alpha, \theta, \beta, a, b, c) = \frac{K\beta\Gamma\left(\alpha, (\theta/t)^\beta\right)^{a-1} (\theta/t)^{\alpha\beta+1} e^{-(\theta/t)^\beta - c \frac{\Gamma(\alpha, (\theta/t)^\beta)}{\Gamma(\alpha)}}}{\theta\Gamma(\alpha)^{a+b-1} \left(\Gamma(\alpha) - \Gamma\left(\alpha, (\theta/t)^\beta\right)\right)^{1-b}} \quad (22)$$

The cumulative distribution function of KBIGG is

$$F_{KBIGG}(t; \alpha, \theta, \beta, a, b, c) = K \int_0^{\frac{\Gamma(\alpha, (\theta/t)^\beta)}{\Gamma(\alpha)}} w^{\alpha-1} (1-w)^{b-1} e^{-cw} dw \quad (23)$$

Although the Kummer beta generated generalized gamma distribution introduced in [38] is somewhat similar, KBIGG distribution is more suitable for our purposes because it encompasses the inverse gamma distribution as we mention in Section 3.1. In Table 2, we present some special cases of the KBIGG distribution and how they relate to each other.

KBIGG	Kummer beta-generated IGG	α	θ	β	a	b	c
BIGG	Beta-generated IGG	α	θ	β	a	b	0
EIGG	Exponentiated IGG	α	θ	β	a	1	0
IGG	Inverse of generalized gamma	α	θ	β	1	1	0

Table 2: Some special cases of the KBIGG distribution

In Figure 3 we present a graphical summary of all candidate distributions that we consider for modeling the first passage time distribution. Both GBIGG and KBIGG are introduced in this paper for the first time, encompassing five other suitable models already existing in the literature. In Section 4, we provide an in-depth analysis of these distributions' performances.

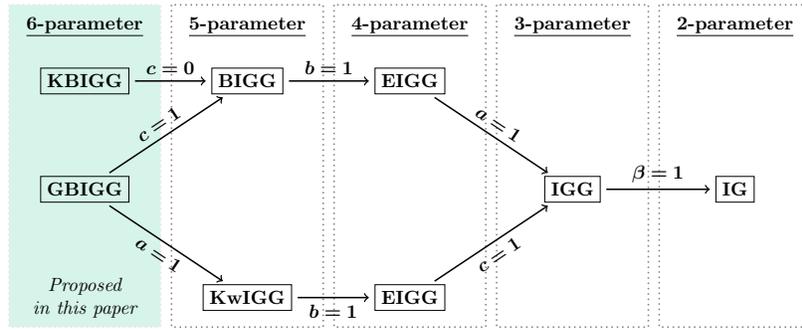


Figure 3: Hierarchy of candidate distributions. We start out with the inverse gamma distribution and introduce some of its generalizations currently existing in the literature. The 6-parameter GBIGG and KBIGG distributions are introduced in this paper for the first time.

3.5. Intricacies of fitting distributions to empirical data

Although the candidate distributions for the first passage time distribution have their support on $[0, \infty)$, it is practically impossible to observe the absorption of MMs in a simulation environment indefinitely. Since simulations provide us with finite-time observations, we need to take this fact into account when finding the suitable first passage time probability distribution. The absorption of an MM is analogous to the death event in lifetime data analysis. Researchers in this domain often come across finite-time observations, which are specified as right-truncated [39, 40]. That is, the researchers are unable to observe the event of interest after some time T . In our case, we are unable to observe the absorption of some MMs after the simulation end time T . Thus, we need to fit a right-truncated distribution to the first passage time data.

For a given distribution with probability density function $f(t)$ and cumulative distribution function $F(t)$, the distribution right-truncated at T is expressed with the probability density function $\frac{f(t)}{F(T)}$ and the cumulative distribution function $\frac{F(t)}{F(T)}$. The right-truncated versions of the candidate distributions introduced earlier in this section should be used for fitting the simulation data.

As we mention earlier in this section, $f^*(t)$ is the improper first passage time distribution of all MMs. By definition, its cumulative distribution function $F^*(t)$ defines the ratio of MMs absorbed until time t to all emitted MMs. Let N_e be the total number of molecules emitted at time $t = 0$ and $\hat{N}_a(t)$ be the expected number of molecules absorbed until time t . Then, the estimator of $F^*(t)$ can simply be expressed as

$$\hat{F}^*(t) = \frac{\hat{N}_a(t)}{N_e}. \quad (24)$$

For estimating the probability of absorption, we refer to the definition of the improper first passage time distribution. From (3), it follows that

$$f^*(t) = p_a f(t). \quad (25)$$

We integrate both sides with respect to t from 0 to T :

$$\int_0^T f^*(t)dt = \int_0^T p_a f(t)dt. \quad (26)$$

Since the probability of absorption p_a is independent of time, we have

$$F^*(T) = p_a F(T) \quad (27)$$

$$p_a = \frac{F^*(T)}{F(T)}. \quad (28)$$

Substituting the estimator for $F^*(T)$ from (24) yields the estimator for p_a as

$$\hat{p}_a = \frac{\hat{N}_a(T)}{N_e F(T)}. \quad (29)$$

285 4. Results

We test the modeling performance of the candidate distributions introduced in Section 3 under 1900 scenarios dictated by four factors: diffusion coefficient, transmitter/receiver size, and distance. To this end, we implement a particle tracking based MCvD environment and monitor the movement of each MM
 290 governed by Brownian motion. We operate in an environment as realistic as possible by mimicking the pancreatic islets. The NeN sizes are similar to the pancreatic beta cells and the distance between them are typical of such a setting. The diffusion coefficients reflect the speed of the insulin hormone in a no-drift environment with temperature and viscosity similar to those of blood. For each
 295 scenario, the probability of absorption, p_a , differs depending on the aforementioned four factors. In order to keep the number of received MMs similar, we emit more molecules for impaired scenarios, which is the reason N_e are drastically different across all scenarios. Further details of the simulation parameters are presented in Table 3.

300 For each MM, we record whether or not it was received and if reception occurs, the time of absorption is recorded. For each scenario, absorption times are used to construct an empirical CDF of the first passage time probability. Then, least squares fitting is performed to match truncated candidate CDFs to the empirical CDFs for each scenario. The implementation for fitting is

Parameter	Variable	Value
Diffusion coefficient	D	50, 75, 100, and 125 $\mu\text{m}^2/\text{s}$
Transmitter radius	r_{tx}	6, 8, 10, 12, and 14 μm
Receiver radius	r_{rx}	6, 8, 10, 12, and 14 μm
Tx-Rx distance	d	2 μm to 20 μm
Simulation time step	Δt	10^{-5} s
Simulation duration	T	1000 s
Number of emitted MMs per scenario	N_e	800 000 to 3 200 000

Table 3: Simulation Parameters

305 done using MATLAB and GNU Scientific Library (GSL). Specifically, we use
the `lsqcurvefit` function in MATLAB for curve fitting while performing in-
tegration operations using GSL due to its superior numerical stability. We
investigate the modeling performance of seven candidate distributions, namely
IG, IGG, EIGG, KwIGG, BIGG, GBIGG, and KBIGG, which have 2, 3, 4, 5,
310 5, 6, and 6 parameters, respectively.

4.1. Mean Squared Error Based Performance

In order to test the performance of the candidate distributions, we perform
several goodness of fit tests for each simulation scenario. We start by observ-
ing three metrics, namely root mean squared error (RMSE), normalized mean
squared error (NMSE), and normalized root mean squared error (NRMSE) de-

defined as

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^M (x_i - y_i)^2}, \quad (30)$$

$$\text{NMSE} = 1 - \frac{\sum_{i=1}^M (x_i - y_i)^2}{\sum_{i=1}^M (y_i - \bar{y})^2}, \quad (31)$$

$$\text{NRMSE} = 1 - \sqrt{\frac{\sum_{i=1}^M (x_i - y_i)^2}{\sum_{i=1}^M (y_i - \bar{y})^2}} \quad (32)$$

where M is the sample size, x_i are the test data, y_i are the reference data, and \bar{y} is the mean value of the reference data. We use these three metrics for measuring the overlap between the empirical CDF (reference data) and the candidate CDF (test data). For RMSE, values closer to zero indicate a better fit. For NMSE and NRMSE, $-\infty$ indicates a bad fit and 1 indicates a perfect fit. For the latter two metrics, the outcome of zero indicates that the test data models the reference data only as good as its mean value.

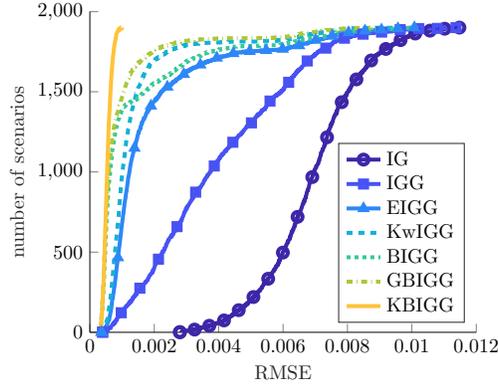


Figure 4: Cumulative match characteristic (CMC) curves for the RMSE of each candidate distribution. A good candidate performs well across all scenarios, which is indicated by a high number of scenarios for a target RMSE value.

In Figure 4, we present the cumulative match characteristic (CMC) curves

320 for the RMSE for each candidate distribution. The horizontal axis represents the
target RMSE and the vertical axis represents the number of scenarios for which
the RMSE is less than the target RMSE. A good candidate distribution performs
well across all scenarios, which is indicated by a high number of scenarios for
a small target RMSE value. KBIGG achieves best fit performance since in
325 all scenarios RMSE is lower than the competitor distributions. Recall that,
the working model for the point-to-sphere first passage time distribution given
in (7) is a special case of IG. Even with its parameters relaxed as in (5), it
performs worst among all distributions and fails to model the sphere-to-sphere
first passage time distribution.

		IG	IGG	EIGG	KwIGG	BIGG	GBIGG	KBIGG
RMSE	Min	0.0028	0.0004	0.0004	0.0003	0.0003	0.0002	0.0003
	Max	0.0115	0.0114	0.0101	0.0096	0.0101	0.0087	0.0010
	Mean	0.0069 ± 0.0015	0.0039 ± 0.0021	0.0019 ± 0.0018	0.0014 ± 0.0013	0.0013 ± 0.0018	0.0010 ± 0.0012	0.0005 ± 0.0001
NMSE	Min	0.9975	0.9975	0.9976	0.9980	0.9976	0.9984	~ 1
	Max	0.9999	~ 1	~ 1				
	Mean	0.9993 ± 0.0004	0.9997 ± 0.0004	0.9999 ± 0.0003	0.9999 ± 0.0002	0.9999 ± 0.0003	~ 1 ± 0.0002	~ 1
NRMSE	Min	0.9502	0.9502	0.9514	0.9548	0.9514	0.9605	0.9961
	Max	0.9903	0.9986	0.9987	0.9988	0.9990	0.9991	0.9990
	Mean	0.9749 ± 0.0067	0.9854 ± 0.0091	0.9928 ± 0.0078	0.9948 ± 0.0058	0.9948 ± 0.0076	0.9963 ± 0.0054	0.9980 ± 0.0005

Table 4: Goodness of fit results for the candidate distributions according to RMSE, NMSE, and NRMSE criteria. A good fit is indicated by a low RMSE value, and NMSE or NRMSE values close to 1.

330 For each candidate distribution, the minimum, maximum, and average val-
ues for RMSE, NMSE, and NRMSE are presented in Table 4. A good fit is
indicated by a low RMSE, high NMSE, and high NRMSE values. The best
fitting distribution across all criteria is the KBIGG distribution, followed by the
GBIGG distribution. NMSE is the least discriminative metric among the three

335 criteria and gives very close results. (Note that, the value 1 indicates a perfect fit; however, the values presented in the table read as 1 due to rounding.) Although these metrics clarify the performance difference of the 7 candidate distributions with respect to each other, they lack the necessary statistical rigor to prove that KBIGG is indeed a well chosen candidate.

340 *4.2. Statistically Validating the Equality of Candidate Distribution to Empirical Data*

In order to carry out a more statistically significant comparison, we employ the Kolmogorov-Smirnov (KS) test. Kolmogorov-Smirnov test is a powerful statistical tool for assessing the equality of a candidate distribution to an empirical distribution. For an empirical CDF $G(t)$ and a candidate CDF $F(t)$, the Kolmogorov-Smirnov statistic D^* is calculated as the supremum of the distance between two distributions:

$$D^* = \sup_t (|G(t) - F(t)|). \quad (33)$$

The Kolmogorov-Smirnov statistic is used for testing the null hypothesis that the candidate distribution and the underlying distribution of the empirical data are equal. The null hypothesis is rejected if

$$\alpha > \underbrace{1 - F_{Kol}(\sqrt{M}D^*)}_{\text{p-value}} \quad (34)$$

where $\alpha \in (0, 1)$ is the significance level, M is the number of samples, and $F_{Kol}(\cdot)$ is the cumulative distribution function of Kolmogorov distribution. Rejection of the null hypothesis means that the candidate distribution does not model the simulation data well. A higher significance level stands for a stricter test and a higher p-value indicates a better fit.

345

Figure 5 shows the CMC curves of each candidate distribution according to their KS test performances. The horizontal axis represents the significance level α of the KS test and the vertical axis represents the number of successful scenarios passing the test at the given α significance level (i.e. scenarios with p-values higher than α .) KBIGG is by far the most successful distribution

350

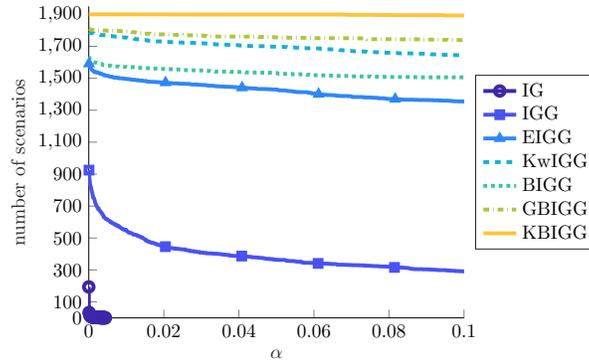


Figure 5: CMC curves of each candidate distribution according to their KS test performances. A successful candidate passes the KS test across most of the scenarios even at high significance level α .

to model the first passage time probability. All but one of the 1900 different scenarios pass the test even when a high significance level is targeted.

	p-value			Good fits	
	Min	Max	Mean	$\alpha = 0.01$	$\alpha = 0.05$
IG	1.979e-129	0.0044	1.257e-05 \pm 0.0001	0%	0%
IGG	6.205e-62	0.9978	0.0733 \pm 0.1945	28.89%	19.16%
EIGG	1.679e-54	~ 1	0.409 \pm 0.3357	78.74%	75.26%
KwIGG	1.679e-54	~ 1	0.5371 \pm 0.3123	92.32%	89.37%
BIGG	1.679e-54	~ 1	0.7005 \pm 0.3936	82.68%	80.68%
GBIGG	1.679e-54	~ 1	0.779 \pm 0.3163	94.37%	92.42%
KBIGG	0.04317	~ 1	0.9116 \pm 0.1521	100%	99.95%

Table 5: Minimum, maximum, and mean p-values for each candidate distribution along with the percentage of good fits at significance level α , out of 1900 scenarios. A higher p-value and percentage indicate a statistically better fit.

Table 5 presents a summary of the CMC curves in Figure 5. At both significance levels 0.01 and 0.05, KBIGG is the best candidate distribution to model the first passage time probability by passing the KS test, followed by GBIGG and KwIGG, respectively. At significance level 0.01, KBIGG passes the KS test

across all 1900 scenarios. It is also worth mentioning that the worst p-value achieved by KBIGG is by far greater than its counterparts for other distribu-
 360 tions. (Some of the maximum p-values are very close to 1 and they appear as 1 due to rounding.)

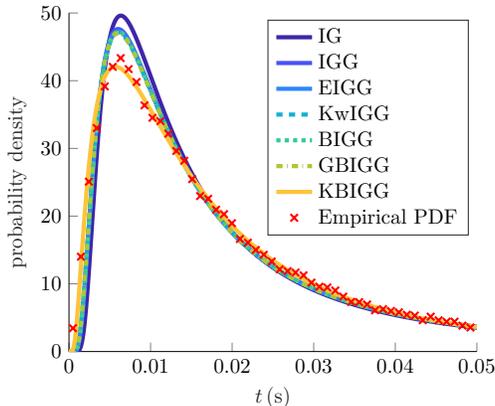


Figure 6: Observing the candidate distributions closely: The PDF of the “worst KBIGG fit” where it achieves the minimum p-value in comparison to its contenders. All candidate distributions are displayed in their truncated forms ($f(t)/F(T)$) since the empirical observations stop at time T . In this scenario, $D=125 \mu\text{m}^2/\text{s}$, $r_{tx}=14 \mu\text{m}$, $r_{rx}=10 \mu\text{m}$, and $d=2 \mu\text{m}$. Even the worst fit of KBIGG outperforms its counterparts.

Modeling the first passage time probability turns out to be more challenging for specific scenarios where the diffusion coefficient is high, both transmitting and receiving NeNs are large and the distance between them is small. In Fig-
 365 ure 6, we observe this challenge closely. This figure shows all of the candidate PDFs for the scenario where the p-value achieved by KBIGG is minimum among all scenarios. This specific scenario results in a rush of fast moving MMs across a small gap, in addition to being reflected and absorbed by large NeN bodies. Most of the candidate distributions struggle with modeling the first passage time
 370 distribution of such a scenario due to the empirical CDF’s significant skewness towards the left extreme (notice the short duration in the vertical axis.) Recall that our reason for choosing the Kummer-beta generator for deriving the KBIGG distribution is its power at offering flexibility to extreme values and its

ability to support heavy tails. The outcome of this scenario shows that KBIGG
 375 is indeed a very suitable choice for modeling the first passage time probability
 of sphere-to-sphere MCvD. In the remainder of this paper, we focus our analysis
 on the KBIGG distribution.

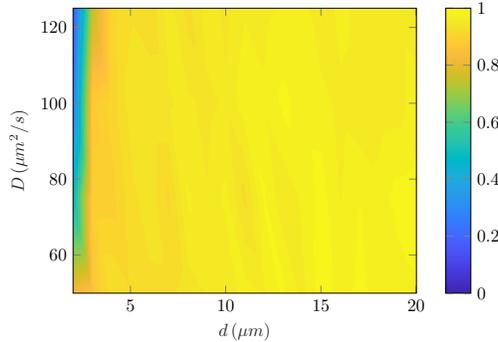


Figure 7: The effect of diffusion coefficient D and transmitter-receiver distance d on the p-values for KBIGG. Higher intensity areas indicate a higher p-value. First passage probability is harder to model under fast diffusion and small distances. ($r_{tx}=10 \mu\text{m}$, $r_{rx}=10 \mu\text{m}$)

The difficulty of modeling the first passage probability under the fast diffusion and small distance combination can be further observed in Figure 7. In
 380 this figure, p-values for KBIGG is given as a heatmap encompassing all possible diffusion coefficient and distance pairs. Higher intensity areas indicate a higher p-value where KBIGG models the empirical data well. The difficulty presented by the diffusion coefficient is only observed for small distances; when the distance is larger, the diffusion coefficient does not affect the fitting performance.
 385 No significant change is observed in p-value with respect to transmitter/receiver size and, therefore, its graph is omitted.

Figure 8 shows the change in the PDFs of KBIGG distribution under different transmitter/receiver distances. The separation between the communicating bodies is a very prominent factor in MCvD. The shape of the first passage time
 390 distribution changes drastically with different d values. Increasing the separation results in a great impairment of communication where the first passage process becomes much more uniform in time, and the distribution's mode shifts

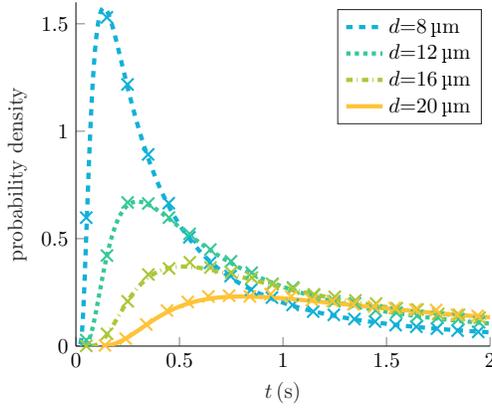


Figure 8: The change in the PDFs of KBIGG distribution under different transmitter/receiver distances. Corresponding empirical PDF values are marked in ‘x’. ($D=75 \mu\text{m}^2/\text{s}$, $r_{tx}=10 \mu\text{m}$, $r_{rx}=10 \mu\text{m}$)

towards right. KBIGG’s high modeling performance can be observed once more by the significant overlap between the curves and the empirical PDF data.

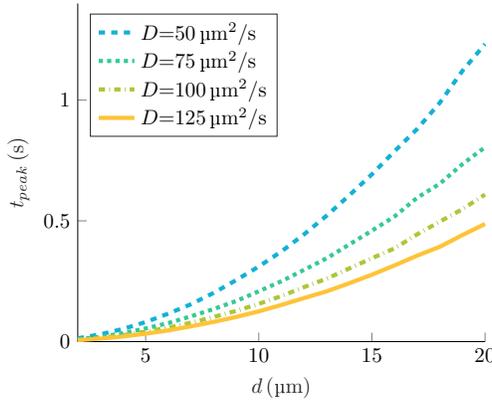


Figure 9: The change in peak signal time with respect to distance and diffusion coefficient. ($r_{tx}=10 \mu\text{m}$, $r_{rx}=10 \mu\text{m}$)

395 In Figure 9, we observe how the mode of the fitted KBIGG distributions change with respect to distance and diffusion coefficient. The vertical axis named ‘ t_{peak} ’ denotes the mode of the distribution, i.e. the peak time of the signal. Similar to the point source setting [10], t_{peak} is quadratically proportional

to d and inversely proportional to D .

400 Up until this point, we have established the statistical significance of KBIGG
being the most suitable choice for modeling the first passage time probability of
sphere-to-sphere MCvD. In all candor, we would like to point out that there is
still the need for a mechanism that maps the four network scenario parameters
 D , r_{tx} , r_{rx} , and d to the six distribution parameters α , θ , β , a , b , and c . Note
405 that, finding this mapping is out of the scope of this paper and our focus is on
introducing the suitable distribution family for which this mapping should be
investigated.

4.3. Preliminary Network Performance Analysis

Although the focus of this work is on modeling the first passage time distri-
410 bution of sphere-to-sphere MCvD, we also dwell on a couple of network perfor-
mance criteria based on the KBIGG distribution. We observe the effect of the
four scenario parameters D , r_{tx} , r_{rx} , and d .

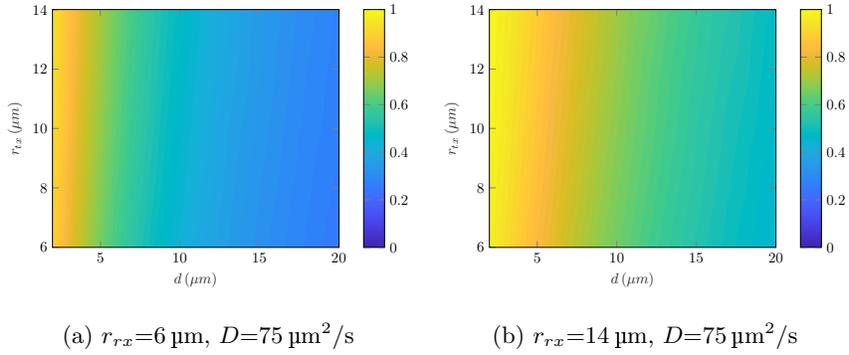


Figure 10: The change in probability of absorption with respect to transmitter/receiver size and distance. Higher intensity areas indicate a higher probability of absorption.

We first investigate the probability of absorption p_a under different scenarios
in Figure 10. In this heatmap, higher intensity areas indicate a higher probabili-
415 ty of absorption. In Figures 10a and 10b, we show the change in probability of
absorption with respect to changing distance and transmitter size for $r_{rx}=6 \mu\text{m}$
and $r_{rx}=14 \mu\text{m}$, respectively. We observe that increased distance affects ab-

sorption negatively, since the MMs are scattered away before they can reach the receiver. A large receiver body is capable of absorbing more MMs, which is shown by the larger bright intensity area in Figure 10b. An interesting trend to observe is the effect of the transmitter size on absorption. When the receiver body is large, increasing the transmitter's size affects the absorption process favorably since the MMs reflect off of the transmitter surface towards the receiver. This trend is more prominent for small distances; reflected molecules are absorbed by the large receiver body before they have a chance to dissipate. The behavior is almost non-existent for the case with a smaller receiver. Even when the MMs are reflected by the transmitter, they are less likely to coincide with a small body as they dissipate.

Like in any other communication application, MCvD also requires a continuous transmission of information in the communication medium. One way to provide this continuity is to divide the time into symbol durations and periodically release batches of MMs to convey information. In such a setting, the heavy tailed nature of the first passage time probability creates a risk. Lagging MMs which are absorbed after their own symbol duration cause inter-symbol interference and make MCvD a channel with memory [21]. Signal to interference ratio (SIR) is a simple metric to assess the effect of the heavy tail on the intended signal. For a symbol duration of t , SIR is defined as

$$\text{SIR}(t) = \frac{F(t)}{1 - F(t)}. \quad (35)$$

In Figure 11, we present the effect of the transmitter and receiver size on SIR, as well as the reflecting sphere transmitter's advantage over the point transmitter. In order to observe the effect of transmitter reflectivity better, we investigate a scenario where the distance between the communicating NeNs is small and the diffusion coefficient is high. Increasing the transmitter size has a distinctly positive effect on the SIR, especially when the receiver is larger, as shown in Figure 11b. This trend is still observable but has a slightly diminished effect when the receiver is small as in Figure 11a.

Another important point worth mentioning is that increasing the receiver

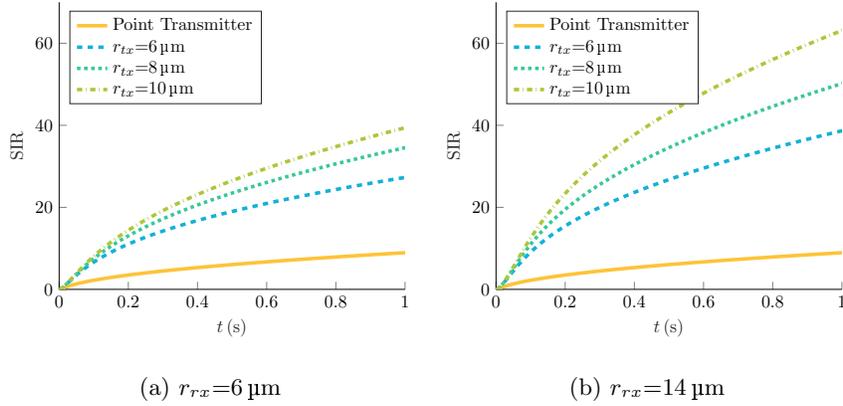


Figure 11: The effect of the transmitter and receiver size and the sphere transmitter’s advantage over the point transmitter is portrayed in terms of signal-to-interference ratio ($d=2\ \mu\text{m}$, $D=125\ \mu\text{m}^2/\text{s}$.)

size has no effect on SIR when the point transmitter is used since the first passage time distribution in that case depends only on d and D (as shown in
 440 (7)). In contrast, using a reflective sphere transmitter not only boosts SIR, but the effect is also amplified when a larger receiver is used. Therefore, in addition to being more realistic, the sphere transmitter also has a positive effect on the communication quality when it is of reflective nature, thus confirming once again our motivation regarding this paper.

445 Note that, for calculating the SIR using simulation results, two values would be needed: the number of absorbed molecules until time t , $N_a(t)$, and the total number of absorbable molecules, $N_a(\infty)$. The estimator of the SIR would then be $\frac{N_a(t)}{N_a(\infty) - N_a(t)}$. Here, $N_a(t)$ can be easily counted, however it is impossible to find the value of $N_a(\infty)$ using simulations. Only the total number of absorbed
 450 molecules until the end of simulation, $N_a(T)$, can be calculated. Without any closed-form representation of the first passage time distribution, the SIR can only be estimated using $N_a(T)$ instead of $N_a(\infty)$. This would result in an inaccurate and optimistic estimation of SIR since $N_a(\infty) > N_a(T)$. By using the KBIGG distribution to represent the first passage time distribution, we are
 455 able to provide an accurate analysis of the SIR.

While we strive to work in a communication setting with a high SIR value, it follows from the definition in (35) that the SIR approaches infinity as $t \rightarrow \infty$. A trade-off is needed to continue communication in a timely fashion with a short symbol duration, while keeping the SIR as high as possible. Finding the optimal
460 symbol duration under these considerations is an open issue that is yet to be investigated.

5. Conclusion

In this paper, we investigate the first passage time distribution of sphere-to-sphere molecular communication via diffusion in the 3-D environment. In
465 the literature, this issue is mostly examined for the 1-D case; for 3-D, the first passage time probability is shown to follow an inverse gamma distribution for point-to-sphere communication. However, realistic channel modeling requires a transmitting body, which makes a full analytical approach a challenge. To this end, we investigate several candidate distributions to represent the first
470 passage time probability by using empirical densities obtained from extensive simulations. In addition to investigating several existing distributions in the literature, we introduce two new distributions: the generalized beta-generated inverse of generalized gamma (GBIGG) and Kummer beta-generated inverse of generalized gamma (KBIGG). We fit these candidate distributions to the
475 empirical distributions we construct from simulation data. We meticulously evaluate the quality of these fits with several methods, including the powerful Kolmogorov-Smirnoff test for goodness of fit. We show that KBIGG models sphere-to-sphere MCvD almost perfectly, and it is the best performing distribution across a variety of simulation scenarios, followed by GBIGG. We also show
480 that point and sphere transmitters affect the behavior of MCvD significantly since the inverse gamma distribution (which is suitable the point-to-sphere approach) falls short of statistically acceptable modeling for the sphere-to-sphere MCvD. Additionally, we conduct a preliminary network performance analysis focusing on the probability of molecule absorption and signal-to-interference

485 ratio. The network performance analysis results emphasize the importance of
using a reflective spherical transmitter instead of the point transmitter. We also
shed light on several open issues in the area which need to be addressed in the
future, such as finding a mapping of network parameters (distance, diffusion
coefficient, etc.) to the distribution parameters of KBIGG and symbol duration
490 optimization targeting inter-symbol interference.

Acknowledgments

This work has been partially supported by the State Planning Organization
(DPT) of Republic of Turkey under the project TAM with the Project Number
2007K120610.

495 References

- [1] T. Nakano, A. W. Eckford, T. Haraguchi, *Molecular communication*, Cambridge University Press, 2013.
- [2] A. O. Tarakanov, L. B. Goncharova, Y. A. Tarakanov, *Carbon nanotubes towards medicinal biochips*, *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* 2 (1) (2010) 1–10.
500
- [3] J. Beal, T. Lu, R. Weiss, *Automatic compilation from high-level biologically-oriented programming language to genetic regulatory networks*, *PLoS One* 6 (8) (2011) e22490.
- [4] S. Hiyama, Y. Moritani, T. Suda, R. Egashira, A. Enomoto, M. Moore,
505 T. Nakano, *Molecular communication*, *Journal-Institute of Electronics Information and Communication Engineers* 89 (2) (2006) 162.
- [5] I. F. Akyildiz, J. M. Jornet, M. Pierobon, *Nanonetworks: A new frontier in communications*, *Communications of the ACM* 54 (11) (2011) 84–89.
- [6] *IEEE Approved Draft Recommended Practice for Nanoscale and Molecular Communication Framework*, IEEE P1906.1/D2.1.
510

- [7] K. V. Srinivas, A. W. Eckford, R. S. Adve, Molecular Communication in Fluid Media: The Additive Inverse Gaussian Noise Channel, *IEEE Transactions on Information Theory* 58 (7) (2012) 4678–4692.
- [8] T. Nakano, Y. Okaie, J.-Q. Liu, Channel Model and Capacity Analysis of Molecular Communication with Brownian Motion, *IEEE Communications Letters* 16 (6) (2012) 797–800.
- [9] S. Kadloor, R. S. Adve, A. W. Eckford, Molecular communication using Brownian motion with drift, *IEEE Transactions on NanoBioscience* 11 (2) (2012) 89–99.
- [10] H. B. Yilmaz, A. C. Heren, T. Tugcu, C.-B. Chae, Three-dimensional channel characteristics for molecular communications with an absorbing receiver, *IEEE Communications Letters* 18 (6) (2014) 929–932.
- [11] M. Ş. Kuran, H. B. Yilmaz, T. Tugcu, B. Özerman, Energy model for communication via diffusion in nanonetworks, *Elsevier Nano Communication Networks* 1 (2) (2010) 86–95.
- [12] H. Li, S. M. Moser, D. Guo, Capacity of the Memoryless Additive Inverse Gaussian Noise Channel, *IEEE Journal on Selected Areas in Communications* 32 (12) (2014) 2315–2329. doi:10.1109/JSAC.2014.2367673.
- [13] N. Farsad, Y. Murin, M. Rao, A. Goldsmith, On the capacity of diffusion-based molecular timing channels with diversity, in: 2016 50th Asilomar Conference on Signals, Systems and Computers, 2016, pp. 1117–1121. doi:10.1109/ACSSC.2016.7869544.
- [14] A. Akkaya, H. B. Yilmaz, C.-B. Chae, T. Tugcu, Effect of receptor density and size on signal reception in molecular communication via diffusion with an absorbing receiver, *IEEE Communications Letters* 19 (2) (2015) 155–158.
- [15] A. C. Heren, F. N. Kilicli, G. Genc, T. Tugcu, Effect of messenger molecule decomposition in communication via diffusion, in: *Proceedings of ACM*

- The First Annual International Conference on Nanoscale Computing and
Communication, ACM, 2014, p. 12.
- 540
- [16] A. C. Heren, H. B. Yilmaz, C.-B. Chae, T. Tugcu, Effect of degradation
in molecular communication: Impairment or enhancement?, *IEEE Trans-
actions on Molecular, Biological and Multi-Scale Communications* 1 (2)
(2015) 217–229.
- 545 [17] H. B. Yilmaz, Y.-J. Cho, W. Guo, C.-B. Chae, Interference reduction
via enzyme deployment for molecular communication, *Electronics Letters*
52 (13) (2016) 1094–1096.
- [18] A. Noel, K. C. Cheung, R. Schober, Optimal receiver design for diffusive
molecular communication with flow and additive noise, *IEEE Transactions*
550 *on NanoBioscience* 13 (3) (2014) 350–362.
- [19] A. Noel, K. C. Cheung, R. Schober, Improving receiver performance of
diffusive molecular communication with enzymes, *IEEE Transactions on*
NanoBioscience 13 (1) (2014) 31–43.
- [20] H. B. Yilmaz, G.-Y. Suk, C.-B. Chae, Chemical propagation pattern for
555 molecular communications, *IEEE Wireless Communications Letters* 6 (2)
(2017) 226–229.
- [21] G. Genc, Y. E. Kara, H. B. Yilmaz, T. Tugcu, ISI-Aware Modeling and
Achievable Rate Analysis of the Diffusion Channel, *IEEE Communications*
Letters 20 (9) (2016) 1729–1732.
- 560 [22] G. Genc, H. B. Yilmaz, T. Tugcu, Reception enhancement with protrusions
in communication via diffusion, in: *Proc. of IEEE International Black Sea*
Conference on Communications and Networking (BlackSeaCom), 2013, pp.
89–93.
- [23] A. Noel, D. Makrakis, A. Hafid, Channel impulse responses in diffusive
565 molecular communication with spherical transmitters, in: *Proc. CSIT Bi-
ennial Symposium on Communications*, 2016.

- [24] H. B. Yilmaz, C. Lee, Y. J. Cho, C.-B. Chae, A machine learning approach to model the received signal in molecular communications, arXiv preprint arXiv:1611.06079.
- 570 [25] M. Pierobon, I. F. Akyildiz, A physical end-to-end model for molecular communication in nanonetworks, *IEEE Journal on Selected Areas in Communications* 28 (4) (2010) 602–611.
- [26] H. C. Berg, *Random walks in biology*, Princeton University Press, 1993.
- [27] M. J. Saxton, Modeling 2D and 3D diffusion, in: *Methods in Membrane Lipids*, Springer, 2007, pp. 295–321.
- 575 [28] M. Ş. Kuran, H. B. Yilmaz, T. Tugcu, Effects of routing for communication via diffusion system in the multi-node environment, in: *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, 2011, pp. 461–466.
- 580 [29] S. Redner, *A guide to first-passage processes*, Cambridge University Press, 2001.
- [30] E. W. Stacy, A generalization of the gamma distribution, *The Annals of Mathematical Statistics* (1962) 1187–1192.
- [31] C. Alexander, G. M. Cordeiro, E. M. Ortega, J. M. Sarabia, Generalized beta-generated distributions, *Computational Statistics & Data Analysis* 56 (6) (2012) 1880–1897.
- 585 [32] M. Jones, Families of distributions arising from distributions of order statistics, *Test* 13 (1) (2004) 1–43.
- [33] G. M. Cordeiro, M. de Castro, A new family of generalized distributions, *Journal of Statistical Computation and Simulation* 81 (7) (2011) 883–898.
- 590 [34] R. R. Pescim, G. M. Cordeiro, C. G. Demétrio, E. M. Ortega, S. Nadarajah, The new class of Kummer beta generalized distributions, *SORT-Statistics and Operations Research Transactions* 36 (2) (2012) 153–180.

- [35] J. B. McDonald, Some generalized functions for the size distribution of
595 income, *Econometrica: Journal of the Econometric Society* (1984) 647–
663.
- [36] K. Ng, S. Kotz, Kummer-Gamma and Kummer-Beta univariate and mul-
tivariate distributions, Tech. Rep. 84, Department of Statistics, The Uni-
versity of Hong Kong, Hong Kong (1995).
- 600 [37] M. Abramowitz, I. A. Stegun, et al., Handbook of mathematical functions,
Applied Mathematics Series 55 (62) (1966) 39.
- [38] G. M. Cordeiro, R. R. Pescim, C. G. Demétrio, E. M. Ortega, The Kummer
Beta Generalized Gamma Distribution, *Journal of Data Science* 12 (2014)
661–698.
- 605 [39] J. P. Klein, M. L. Moeschberger, Survival analysis: techniques for censored
and truncated data, Springer Science & Business Media, 2005.
- [40] J. F. Lawless, Statistical models and methods for lifetime data, Vol. 362,
John Wiley & Sons, 2011.