Aalborg Universitet

AALBORG UNIVERSITY

Analysis of Stochastic Radio Channels with Temporal Birth-Death Dynamics

A Marked Spatial Point Process Perspective

Jakobsen, Morten Lomholt; Pedersen, Troels; Fleury, Bernard Henri

Link to publication from Aalborg University

Analysis of Stochastic Radio Channels with Temporal Birth-Death Dynamics: A Marked Spatial Point Process Perspective

Morten Lomholt Jakobsen, Troels Pedersen and Bernard Henri Fleury

Abstract—We employ the theory of spatial point processes to revisit and reinterpret a particular class of time-variant stochastic radio channel models. Common for all models in this class is that individual multipath components are emerging and vanishing in a temporal birth-death like manner, with the underlying stochastic birth-death mechanism governed by two facilitating assumptions. Well-known analytical properties of this class of channel models are reestablished by simple arguments and several new results are derived. The primary tool used to obtain these results is Campbell's Theorem which enables novel assessment of the autocorrelation functions of random processes used in the general channel model description. Under simplifying assumptions the channel transfer function is shown to be wide-sense stationary in both time and frequency (despite the birth-death behavior of the overall channel). The proof of this result is a consequence of the point process perspective, in particular by circumventing enumeration issues arising from the use of integer-indexed path components in traditional channel modeling approaches. The practical importance of being able to analytically characterize the birth-death channel models is clearly evidenced, e.g. by the fact that key parameters enter explicitly in measurable quantities such as the power-delay profile.

Index Terms—Stochastic radio channel modeling, time-variant channel transfer function, time-frequency correlation function, marked spatial point processes, Campbell's Theorem.

I. INTRODUCTION

IN the historical development of *time-invariant* stochastic radio channel models still being favored nowadays, the first use of *point processes* can be traced back to the seminal work by Turin [1], [2]. Specifically, Turin suggested modeling the occurrences of multipath delay components via a one-dimensional (possibly *inhomogeneous*) Poisson point process. One-dimensional point processes were similarly involved in the later developments by Suzuki [3] and Hashemi [4] as a convenient tool for modeling and simulation. Despite a pronounced use for the *modeling* of stochastic radio channels, neither point processes nor their underlying theoretical framework have dominated as tools for *analysis*. This trend persists in the popular contribution by Saleh and Valenzuela [5] as well as in the more recent extension by Spencer et al. [6]. Essentially, point processes are employed only in the channel model specifications whereas tools from the underlying theory have not been used for analytical characterizations. In fact, this trend exists also for certain *time-variant* channel models, see [7]–[10]. This trend can most likely be attributed in part

All three authors are with Aalborg University, Department of Electronic Systems, Fredrik Bajers Vej 7, DK-9220 Aalborg East, Denmark.

to scientific tradition and to the scarce proportion of readily accessible, engineering-targeted literature on point processes around the time of Turin's initial work. The textbook by Snyder [11] is one of the earliest of its kind, targeting engineers, and covering numerous examples and application areas. However, the engineering-targeted exposition in [11] (or its successor [12]) does not appear to have fully convinced the channel modeling community to also start analyzing their models using the variety of well-established tools from the theory of point processes.

Recently, the theoretical framework of *spatial*[1] *point processes* has facilitated the analytical characterization of various stochastic radio channel models, see e.g. [13]–[17]. In [13], a point process approach has been employed to derive and analyze a non-stationary *geometric-stochastic* propagation model applicable within satellite-to-vehicle communications. The impulse response model by Saleh and Valenzuela [5] and more recent variations of it [6] have been analyzed in [14]–[16] with new and detailed insight gained. *Campbell's Theorem*, a standard tool from the theory of spatial point processes, has proven itself instrumental for deriving both well-known and new results via concise and rigorous arguments, e.g. as in [13] and [15].

Time- and space-varying multipath propagation phenomena, such as path components which emerge and vanish, occur partially due to movements of the communicating entities and the surrounding scatterers [18], [19]. To imitate such *birth-death dynamics* in the radio channel, the following two tractable assumptions have been invoked several times in the channel modeling literature:

- *i*) **Stationary emergences:** The collection of time instances where new path components emerge forms a stationary point process on the real line.
  - [†] (Facilitating special-case: Poisson point process).
- *ii*) **i.i.d. lifetimes:** The lifetimes of individual path components are drawn independently and identically distributed.
  - [†] (Facilitating special-case: Exponential distribution).

The special-cases $i)^\dagger$ and $ii)^\dagger$ entered first in [7] while later also in several other contributions, e.g. [8]–[10]. Under $i)^\dagger$ and $ii)^\dagger$ the instantaneous number of path components in the channel is a Poisson distributed random variable (with known mean parameter). This property is justified in [7] by

---

[1]When prepending the term "spatial" we refer to point processes in two dimensions or more. The essential distinction compared to one-dimensional point processes is the inherent absence of a natural ordering of the points.

reference to standard results from *queuing theory*. Basically, the random process governing the time-varying number of path components in the channel can be identified as an $M/M/\infty$ queue [20, Sec. 16-2]. This observation also appears elsewhere in the literature, see e.g. [10, Sec. III-C], where the construction via $i)^\dagger$ and $ii)^\dagger$ is identified as a *birth-death process* or, in the alternative terminology of [10], as a *marked Poisson process*. Analytical benefits and additional insight can often be gained from such structural identifications, e.g. as in [15] and [16] via their individual point process perspectives. Surprisingly, the promising potential of the available theory of marked spatial point processes does not seem to have been exploited so far in the literature on time-variant radio channel models. Yet, spatial point processes have already been successfully utilized in highly related areas of research, namely for the design and analysis of *wireless networks*, see [21] and the extensive list of references therein.

In the present contribution we analyze the class of time-variant stochastic radio channel models which fulfill the more general assumptions $i)$ and $ii)$ in contrast to the special-cases $i)^\dagger$ and $ii)^\dagger$. Random processes essential for radio channel characterization are constructed from underlying spatial point processes. Subsequently, the use of Campbell's Theorem facilitates novel analytical assessment of the associated autocorrelation functions. A key feature of the point process perspective is its ability to elegantly handle enumeration issues arising in traditional channel modeling approaches. Specifically, the traditional use of integer-indexed path components is naturally replaced by a tractable point process-based indexing method. Our first main contribution is an extensive analysis of the *temporal birth-death process* induced from the underlying assumptions $i)$ and $ii)$. We relate our analysis to the results already obtained in the literature regarding the special-cases $i)^\dagger$ and $ii)^\dagger$. Secondly, we derive a novel and general expression for the *time-frequency correlation function* associated with the class of temporal birth-death channel models. With example cases we illustrate how a particular channel model can be readily extended by including random marks in the point process construction. We show as well how the point process perspective offers crucial insight for the task of simulating channel realizations with temporal birth-death dynamics. Overall, our contributions are comprised by the point process formulations and the analytical characterizations induced thereof. Our approach relies on simple arguments and standard results from the theory of spatial point processes. In particular, Campbell's Theorem repeatedly reveals itself a highly useful and easily applicable tool.

The paper is organized as follows. Sec. II provides the necessary background information on the class of temporal birth-death channel models with an overview of previous approaches and contributions. We then present a concise but self-contained introduction to the basics of spatial point processes in Sec. III. Subsequently, we initiate our investigations of the class of time-variant channel models by analyzing and characterizing the temporal birth-death process in Sec. IV. We continue in Sec. V with an exhaustive analysis of the channel's time-frequency correlation function. Sec. VI provides a selection of examples illustrating how the channel model can be ex-

tended and we supplement with details and aspects related to computer simulation. Finally, we draw our conclusions in Sec. VII.

II. STOCHASTIC RADIO CHANNEL MODELS WITH TEMPORAL BIRTH-DEATH DYNAMICS

The birth-death behavior of the class of channel models under consideration is governed by the assumptions $i)$ and $ii)$ in Sec. I. Common to the approaches [7]–[10] is that they (with minor individual variations) essentially all propose a time-varying multipath *channel impulse response* of the form[2]

$$h(t,\tau) = \sum_{\ell=1}^{L(t)} \alpha_\ell(t)\delta\big(\tau - \tau_\ell(t)\big), \qquad (1)$$

where $L(t)$ is the number of path components at time t, $\alpha_\ell(t)$ is the complex-valued gain of the ℓ'th path component, $\tau_\ell(t)$ is the associated propagation delay, and $\delta(\cdot)$ is the Dirac delta. The *channel transfer function* corresponding to (1) reads

$$H(t,f) = \sum_{\ell=1}^{L(t)} \alpha_\ell(t)\exp\big(-j2\pi f\tau_\ell(t)\big). \qquad (2)$$

The *integer-indexed* multipath representations in (1) and (2), and slight variations thereof, are standard in the channel modeling literature [18], [19]. The special-case which has attracted most attention is when $L(t) = L$ is constant (deterministic or random) within each realization of the channel. That is, numerous traditional channel modeling approaches are disregarding the case when the number of path components is allowed to vary *within* each channel realization. However, the channel models in [7]–[10] all incorporate a temporal birth-death feature of the individual path components, either directly via the integer-valued random process[3] $L(\cdot)$ or via equivalent representations. Here, the transition times of $L(\cdot)$ are to be generated according to $i)$ and $ii)$ in Sec. I, but common to the contributions [7]–[10] is that they all specialize exclusively (by default) to the cases $i)^\dagger$ and $ii)^\dagger$.

In [7] each amplitude $|\alpha_\ell|$ is constant with time and the amplitudes of emerging components are drawn i.i.d. according to a log-normal distribution. Hence, the conditional second moment of $|\alpha_\ell|$ does *not* depend on the associated propagation delay τ_ℓ. From narrowband considerations the propagation delays are also modeled to be constant with time and they are drawn i.i.d. from a uniform distribution.

The model in [9] is inspired by [7]. However, the amplitudes $|\alpha_\ell(t)|$ are now varying with time. To ensure smooth transitions when birth and death events occur, a sequence of root-raised cosine functions are incorporated in a multiplicative manner. Each propagation delay $\tau_\ell(t)$ is a function of a random *initial delay* (the propagation delay when a path emerges) and it varies with time as a function of the receiver's position in

---

[2]In fact [7]–[10] rely on a *space-varying* approach which subsequently can be converted into a time-varying equivalent by appropriately assuming a receiver trajectory (most often a straight line in space). Furthermore, [7] and [10] are both modeling the *line-of-sight* component via a separate stochastic mechanism which alternates between being active and inactive.

[3]By $L(\cdot)$ we refer to the entire random process while with $L(t)$ we indicate that the time instance t is fixed. Hence, $L(t)$ is a random variable.

space. The *power-delay profile* is targeted to exhibit an overall exponential decay. See [22] for further details.

In [8] each time-varying amplitude $|\alpha_\ell(t)|$ is smoothed using the non-negative part of an ordinary sine function. The sine is stretched in time to match the individual lifetimes and $|\alpha_\ell(t)|$ is generated such that its conditional second moment depends on $\tau_\ell(t)$ in an exponentially decaying manner. Each propagation delay $\tau_\ell(t)$ varies with time as a function of the receiver's position together with an initial delay drawn from an exponential distribution. From this construction the power-delay profile appears to exhibit an overall exponential decay, but the sine smoothing and the average number of path components are both disregarded in the calculations.

The model in [10] is inspired by [8] and each path component is characterized by a transfer matrix (polarization, antenna characteristics, etc.) together with a propagation delay and directions of departure and arrival. Hence, the channel model in [10] accounts for small-scale fading, large-scale fading and polarization aspects. A key aspect is the modeling of path directions which are found to be Laplacian distributed for smaller propagation delays followed by migration into a uniform distribution for larger delays. Furthermore, a number of heuristic guidelines are proposed for standard technicalities related to computer simulation, e.g. initialization and time-discretization.

A. Consequences of Temporal Birth-Death Dynamics

The individual contributions [7]–[10] rely on different assumptions on the path gains, propagation delays, incidence directions, and so on. The approaches, notations, methodologies, and techniques in use are rather diverse in general. But common to all modeling approaches seems to be the fact that a thorough analytical characterization and the assessment of the resulting channel properties have been very difficult to carry out. Computer simulation does not adequately provide the desired insight and is often not tractable either – especially not when the number of overall model parameters grows as large as in [10].

The essential part of the analytical challenge comes with the birth-death behavior of the random process $L(\cdot)$ describing the time-varying number of path components. Conceptually, the birth-death process $L(\cdot)$ may appear straightforward to handle but difficulties quickly arise in the attempt to compute a correlation quantity such as $\mathbb{E}[H^*(t, f)H(t', f')]$, where $\mathbb{E}[\cdot]$ denotes expectation and $(\cdot)^*$ denotes complex conjugation. The difficulties emerge since for distinct time instances t and t', the integers $L(t)$ and $L(t')$ are not necessarily referring to the same *integer-indexed* path components anymore (glimpse at Fig. 4 on page 8). This "*enumeration issue*" repeatedly affects (1) and (2) every single time the birth-death process $L(\cdot)$ experiences a transition. Consequently, re-enumeration via the integer index ℓ or other means of non-trivial bookkeeping is recurrently needed. Surprisingly, the enumeration issue (which is a consequence of the birth-death behavior and the chosen indexing) is barely mentioned in [7]–[10]. Specifically, [7] is the only contribution explicitly acknowledging that a "time-influenced" integer-indexed representation equivalent to (2)

brings along an inexpedient enumeration issue. An approach is proposed based on *set-valued* random variables, see [7, Sec. 2.2.1], which appears to be the only "mitigation" of the enumeration issue published so far.

Fortunately, as we shall see in Sec. IV and Sec. V, the enumeration-technical and analytical difficulties encountered in [7]–[10] are swiftly circumvented by use of marked spatial point processes. Specifically, we demonstrate how thorough analytical insight can be obtained by virtue of this well-established mathematical framework. The approach is fully general in the sense that our results apply to any temporal birth-death channel model from the class governed by $i)$ and $ii)$, in particular to the models in [7]–[10], which all rely on $i)^\dagger$ and $ii)^\dagger$. However, as knowledge of the basics of spatial point processes will be essential for transparency in the later derivations, we provide in the following a concise, self-contained and engineering-targeted introduction to this mathematical framework.

III. SPATIAL POINT PROCESSES

A *spatial point process* [12], [23]–[26] is a random *countable* collection Y of points sitting in a d-dimensional Euclidian space S (either $\mathbb{R}^d$ or a subset of it). The term "spatial" is used here to stress the fact that d is larger than one, but the term is often skipped again for the sake of brevity. Since the peculiar ordering property of the real line obscures the overall theory, the reader is encouraged [24] to always think of the two-dimensional ($d = 2$) case, see Fig. 1.

For reasons of practical applicability and simplicity it is convenient to restrict attention to the class of *locally finite* and *simple* point processes. Here, locally finite means that only a *finite* number of points are falling in every *bounded* region $B \subset \mathbb{R}^d$. Furthermore, the term "simple" indicates that no two points of the process coincide. Both conditions are to be satisfied with probability one. As no two points from Y coincide, each individual realization of Y can be identified as a countable set of points $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \ldots\}$, $\boldsymbol{y}_i \in S$. The counting index i on $\boldsymbol{y}_i$ is used here only to distinguish individual points and to indicate countability. This index does not imply any ordering of the points (recall Fig. 1) and for this reason we deliberately skip it again.

A. Region Counts and the Intensity Function

A natural way of exploring the properties of a point process is to count the number of points falling in different regions [26]. Accordingly, for any set $B \subseteq S$ consider the *region count*

$$N_Y(B) := |Y \cap B| = \sum_{\boldsymbol{y} \in Y} \mathbb{1}[\boldsymbol{y} \in B], \qquad (3)$$

where $|\cdot|$ denotes set cardinality and $\mathbb{1}[\cdot]$ is a generic *indicator function* taking the value one if the logical statement in brackets is fulfilled and zero otherwise. As suggested by its name, the region count $N_Y(B)$ in (3) gives the random number of points from Y falling within the region B. For *fixed* and *bounded* B, the region count $N_Y(B)$ is an ordinary random variable with range $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$. An example illustration was already provided in Fig. 1. For a general point process Y,
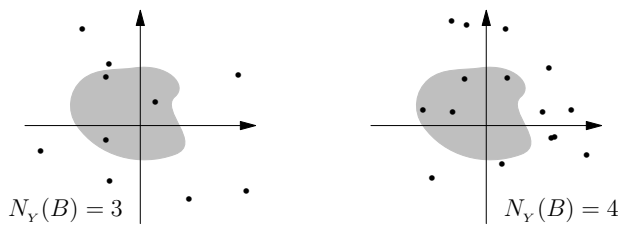
Fig. 1. A pair of realizations of a two-dimensional point process Y. Notice that the points are not enumerated/ordered and observe that each realization has a different number of points $N_Y(B)$ falling in the gray-shaded region B.

the probability distribution of $N_Y(B)$ depends on the region B via its d-dimensional *Lebesgue measure*, shape, location, orientation, and so on. The region B can be very complicated but certain properties of the region counts are easily established. In particular we have $N_Y(\emptyset) = 0$ and for disjoint regions $A, B \subset S$ we also have $N_Y(A \cup B) = N_Y(A) + N_Y(B)$. Indeed, various complicated regions can be build up from simpler ones by use of set operations for which the behavior of $N_Y(\cdot)$ is obvious.

The expected value of the region count $\mu_Y(B) := \mathbb{E}[N_Y(B)]$ induces a *measure* on S, the so-called *intensity measure* of Y. If the intensity measure can be expressed as an integral

$$\mu_Y(B) = \int_B \varrho_Y(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}, \tag{4}$$

for some *locally integrable* function $\varrho_Y : S \to [0, \infty)$, then $\varrho_Y(\cdot)$ is called the *intensity function* of Y. The intensity function does exist for virtually all applications of practical relevance. If $\varrho_Y(\cdot)$ is constant on S then Y is called a *homogeneous* point process and otherwise Y is said to be *inhomogeneous*. In general, the shape of $\varrho_Y(\cdot)$ indicates where points are more likely to occur. Besides this intuitive feature the intensity function plays an important role for the use of *Campbell's Theorem* [24, Chap. 3], [26, Thm. 2.2]. This powerful and widely applicable theorem states that the identity

$$\mathbb{E}\Big[\sum_{\boldsymbol{y} \in Y} g(\boldsymbol{y}) \Big] = \int_S g(\boldsymbol{y}) \varrho_Y(\boldsymbol{y}) \mathrm{d}\boldsymbol{y} \tag{5}$$

holds whenever the integral on the right is well-defined, where $g(\cdot)$ is any real- or complex-valued function defined on S. Campbell's Theorem is particularly useful as it enables straightforward calculation of expected values of scalar random variables of the form

$$\sum_{\boldsymbol{y} \in Y} g(\boldsymbol{y}). \tag{6}$$

Random variables of the type in (6) are frequently encountered when dealing with point processes. One example already appeared in (3) where $g(\cdot)$ is the indicator function for a region B.

B. Marked Point Processes

Let Y be a point process on $S \subseteq \mathbb{R}^d$ and consider the procedure of attaching a random label or a *mark $\boldsymbol{m_y}$* to each point $\boldsymbol{y} \in Y$. These marks can be of very general type but they must all belong to the same space M.

Definition 1. Let Y be a simple and locally finite point process on $S \subseteq \mathbb{R}^d$ and let M be some space. If a random mark $\boldsymbol{m_y} \in M$ is attached to each point $\boldsymbol{y} \in Y$, then

$$X := \{ (\boldsymbol{y}, \boldsymbol{m_y}) : \boldsymbol{y} \in Y \} \tag{7}$$

is called a *marked spatial point process* with points in S and marks in M.

By construction, X is a simple and locally finite random subset of $S \times M$. Accordingly, a marked point process X with points in S and marks in M can always be viewed as an *unmarked* point process on $S \times M$. However, the converse is not true since an arbitrary point process cannot always be projected onto a lower dimensional space and viewed as a marked point process. Such a collection of projected points is not necessarily simple nor locally finite.

C. Poisson Point Processes

The most fundamental types of spatial point processes are members of the *Poisson*[4] class:

Definition 2. [25, Def. 3.2] A point process Y on $S \subseteq \mathbb{R}^d$ is called a *Poisson point process* with intensity function $\varrho_Y(\cdot)$ if it fulfills the two conditions:

1) For any region $B \subseteq S$ with $\mu_Y(B) = \int_B \varrho_Y(\boldsymbol{s}) \mathrm{d}\boldsymbol{s} < \infty$ the region count $N_Y(B)$ is Poisson distributed with mean $\mu_Y(B)$.
2) Conditioned on the event $N_Y(B) = n \in \mathbb{N}$ such that $0 < \mu_Y(B) < \infty$, the joint distribution of $Y \cap B$ is the same as that of n points drawn i.i.d. according to the probability density function

$$f_B(\boldsymbol{s}) := \frac{\varrho_Y(\boldsymbol{s})}{\mu_Y(B)} \mathbb{1}[\boldsymbol{s} \in B], \quad \boldsymbol{s} \in S.$$

We write $Y \sim \texttt{PoissonPP}(S, \varrho_Y)$.

For a Poisson point process the individual region counts are Poisson distributed random variables. Hence the name of the process. A particularly useful property of Poisson point processes arises via the *Marking Theorem* [24, Sec. 5.2], [25, Prop. 3.9]. This theorem states that if mutually independent marks are attached to a Poisson point process, then the extended collection (of points and marks) is again a Poisson point process (in a higher dimensional space) and the associated intensity function is known. Loosely speaking, the Poisson property of the region counts is sustained when expanding the dimensionality. Notice that mutual independence is the key requirement of the Marking Theorem. For instance, the marks need not be identically distributed. Notice also that the Marking Theorem does not apply to point processes in general, only to Poisson point processes. This partly explains the facilitating aspect of the special-case assumption $i)^\dagger$ in Sec. I.

---

[4] A highly accessible introduction to Poisson point processes can be found in the book [24] by Kingman. Although not crucial, his approach and choice of presentation is better perceived with some minor knowledge on abstract *measure theory*. In contrast, several other books on spatial point processes require a solid background in measure theory just to get started.

IV. CHARACTERIZATION OF THE TEMPORAL BIRTH-DEATH PROCESS $L(\cdot)$

In this section we assess the fundamental properties of the temporal birth-death process associated with the class of time-variant stochastic radio channel models fulfilling the assumptions $i)$ and $ii)$ from Sec. I. Our approach is to reformulate and incorporate $i)$ and $ii)$ using a marked point process. This perspective facilitates a convenient definition of the temporal birth-death process $L(\cdot)$ which we briefly introduced in Sec. II. We show that $L(\cdot)$ is strict-sense stationary and we calculate its mean by use of Campbell's Theorem. Subsequently we extract a partial but crucial share of the second-order properties of $L(\cdot)$, again by application of Campbell's Theorem. Furthermore, we illustrate for the facilitating special-case $i)^\dagger$ how additional insight can be gained from changing perspective. Specifically, via the Marking Theorem for Poisson point processes we swap to a higher dimensional representation which allows us to readily identify that $L(t)$ is Poisson distributed for any fixed $t \in \mathbb{R}$. Finally, we indicate that $i)$ and $ii)$ are essential for preserving analytical tractability of $L(\cdot)$ as even minor relaxation attempts turn $L(\cdot)$ into a non-stationary random process.

A. Definition of $L(\cdot)$ Using a Marked Point Process

We begin by setting up a useful notation to incorporate assumption $i)$ from Sec. I. We introduce Y to represent the one-dimensional *stationary*[5] point process describing the time instances when path components emerge. For instance, Y could be a Poisson point process, a *renewal* point process [12, Chap. 6], a *Cox* point process [25, Chap. 5], and so on. A stationary point process is necessarily homogeneous [26, Cor. 1] and so the intensity function $\varrho_Y(\cdot)$ of Y is constant such that $\varrho_Y(y) = \lambda_e$ for all $y \in \mathbb{R}$. The subscript on the positive constant λ_e is used to abbreviate the term "emerge".

Secondly, we need to incorporate assumption $ii)$ from Sec. I. To account for the *period* (or lifetime) of each path component, a random non-negative mark p_y is attached to each element $y \in Y$ (lifetimes are necessarily non-negative). The letter p is used to emphasize the interpretation of each mark as a "period" while the subscript y on p_y serves as a *unique* identifier for its underlying point (with probability one). Thus, the countable collection of periods $\{p_y : y \in Y\}$ is *conveniently indexed* using the points from Y. The periods are drawn i.i.d. according to some probability density function $f_{\text{period}}(\cdot)$ with non-negative support and finite first-order moment

$$\mathbb{E}[p_y] = \mathbb{E}[p_\star] = \int_0^\infty p f_{\text{period}}(p)\mathrm{d}p < \infty. \tag{8}$$

In (8) we use the wildcard notation $\mathbb{E}[p_\star]$ to denote the mean of a "typical/arbitrary" mark $p_\star$ (as they are all drawn i.i.d.). The random collection

$$X := \{(y, p_y) : y \in Y\} \tag{9}$$

[5]In the sense of spatial point processes which similarly to ordinary random processes means that all distributional properties are preserved under arbitrary fixed translations [26, Def. 1.7].

is by construction a marked point process on $\mathbb{R} \times \mathbb{R}_+$ with points in $\mathbb{R}$ and marks in $\mathbb{R}_+$. This marked point process is analytically convenient as it allows directly for random variables of type (6) to be established.

By use of the random collection X in (9), the number of "active" path components at time t can now be formulated as

$$L(t) := \sum_{y \in Y} \mathbb{1}[y \leq t]\mathbb{1}[y + p_y > t], \quad t \in \mathbb{R}. \tag{10}$$

An arbitrary component (y, p_y) from (9) contributes to the value of the sum in (10) only if it *emerges before* and *vanishes after* time t (incorporated by the product of the two indicator functions). Obviously, $L(\cdot)$ as defined by (10) is a continuous-time random process with range $\mathbb{N}_0$ (once more, glimpse at Fig. 4 on page 8). Notice how the temporal birth-death process $L(\cdot)$ is readily generated as *a function* of the underlying random mechanism X. However, X cannot be reconstructed from a realization of $L(\cdot)$ in general, despite the fact that Y always can (from the time instances with upward jumps).

As a consequence of $i)$ and $ii)$, we show in Appendix A that the temporal birth-death process $L(\cdot)$ is *strict-sense stationary*. Intuitively, this is also to be expected since the underlying point process Y is stationary and since the marks/periods attached to it are mutually independent and identically distributed. By strict-sense stationarity, $L(\cdot)$ is also *wide-sense stationary* which means that $\mathbb{E}[L(t)]$ does not depend on t and $\mathbb{E}[L(t)L(t')]$ depends on $t' - t$ only.

B. First- And Second-Order Properties of $L(\cdot)$

Using the law of total expectation [27, Sec. 3.7] we obtain

$$\mathbb{E}[L(t)] = \mathbb{E}\big[\mathbb{E}[L(t) \,|\, Y]\big]$$
$$= \mathbb{E}\Big[\sum_{y \in Y} \mathbb{1}[y \leq t]\mathbb{E}\big[\mathbb{1}[y + p_y > t] \,\big|\, Y\big]\Big].$$

We emphasize that expectation and summation cannot be directly interchanged in (10) since Y is a random collection. Since each mark p_y does not depend on $Y \backslash \{y\}$, we now get

$$\mathbb{E}[L(t)] = \mathbb{E}\Big[\sum_{y \in Y} \underbrace{\mathbb{1}[y \leq t]\text{Pr}(p_y > t - y|y)}_{g(y;t)}\Big]$$
$$= \int_{\mathbb{R}} g(y;t)\varrho_Y(y)\mathrm{d}y,$$

where the last step follows by application of Campbell's Theorem. We have not (yet) used the fact that Y is stationary and neither have we used the assumption that the periods are identically distributed. Hence, the above integral formula is valid for more general settings where the mean $\mathbb{E}[L(t)]$ may depend on time t. We continue by exploiting $i)$ such that

$$\mathbb{E}[L(t)] = \int_{-\infty}^t \text{Pr}(p_y > t - y)\lambda_e \mathrm{d}y \tag{11}$$

and the substitution $\xi := t - y$ together with $ii)$ then yields

$$\mathbb{E}[L(t)] = \lambda_e \int_0^\infty \text{Pr}(p_{t-\xi} > \xi)\mathrm{d}\xi = \lambda_e \mathbb{E}[p_\star], \tag{12}$$

i.e. the mean $\mathbb{E}[L(t)]$ does not depend on t, in accordance with $L(\cdot)$ being strict-sense stationary. Readers familiar with

queueing theory will recognize the expression in (12) to be consistent with well-known results for the $M/G/\infty$ queue [20, Sec. 16-2], i.e. when specializing $i)$ to the case $i)^\dagger$.

The interpretation of (12) is quite intuitive: the average number of path components in the channel is governed by the *rate* (λ_e) at which new components emerge, together with the *inverse rate* $(\mathbb{E}[p_\star])$ at which they vanish again. The mean $\mathbb{E}[L(t)]$ is *not* affected by the exact shape of the probability density function $f_\mathrm{period}(\cdot)$ shared by all the marks. Only the first-order moment $\mathbb{E}[p_\star]$ matters.

By strict-sense stationarity we know that the autocorrelation function $\mathbb{E}[L(t)L(t')]$ depends only on the time difference $t' - t$. To obtain further insight we introduce and split the autocorrelation function as

$$R_L(t,t') := \mathbb{E}[L(t)L(t')] = \mathbb{E}[(\diamond_1)] + \mathbb{E}[(\diamond_2)], \qquad (13)$$

with the definitions

$$(\diamond_1) := \sum_{y \in Y} \mathbb{1}\big[y \le \min\{t,t'\},\ y + p_y > \max\{t,t'\}\big] \quad (14)$$

$$(\diamond_2) := \sum_{y,\widetilde{y} \in Y}^{\neq} \mathbb{1}\big[y \le t,\ y + p_y > t\big]\mathbb{1}\big[\widetilde{y} \le t',\ \widetilde{y} + p_{\widetilde{y}} > t'\big].$$

The splitting in (13) reflects a deliberate choice as the first quantity $(\diamond_1)$ gives the random number of terms (y, p_y) from (9) contributing *jointly* to both $L(t)$ and $L(t')$. The mean of $(\diamond_1)$ is readily assessed using the same manipulations as those leading to (11), in particular the law of total expectation and Campbell's Theorem:

$$\mathbb{E}[(\diamond_1)] = \mathbb{E}\big[\mathbb{E}[(\diamond_1)|Y]\big] = \lambda_\mathrm{e} \int_{|t'-t|}^{\infty} \Pr(p_\star > \xi)\mathrm{d}\xi \quad (15)$$

with $|t'-t| = \max\{t,t'\} - \min\{t,t'\}$.

The second quantity $(\diamond_2)$ is not as easily handled as $(\diamond_1)$ and its interpretation is also not as straightforward. The symbol $\neq$ above the summation in $(\diamond_2)$ is used to indicate that the sum is taken over pairwise distinct points y and $\tilde{y}$. Calculating the mean of $(\diamond_2)$ is indeed possible but more involved as it requires to know the statistical properties of joint occurrences of points from Y. In general, this knowledge is not available through the intensity function $\varrho_Y(\cdot)$ of Y. It is contained in a function called the *second-order product density* [25, Def. 4.3] or the *second moment density* [26, Def. 2.5]. In the sequel we calculate for illustration purposes the mean of $(\diamond_2)$ for the facilitating special-case $i)^\dagger$.

As expected, the quantity in (15) depends only on the time difference $\Delta t := t' - t$, and not on the specific time instances t and t' (by strict-sense stationarity). Obviously, the same conclusion holds for $\mathbb{E}[(\diamond_2)]$, having calculated this term explicitly or not. Opposite to the mean $\mathbb{E}[L(t)]$ in (12) which is affected only by the first-order moment of $f_\mathrm{period}(\cdot)$, the exact shape of this probability density function affects directly the autocorrelation function $\mathbb{E}[L(t)L(t')]$ in (13). As will be shown in Sec. V, the quantity in (15) plays a crucial role as it influences the *time-frequency correlation function* of the channel. To the contrary, the quantity $\mathbb{E}[(\diamond_2)]$ does not. Accordingly, we do not insist on calculating this particular term.

To the best of our knowledge, equally concise and rigorous derivations of the results (12) and (15) have not appeared elsewhere in the literature. The novelty here is that our conclusions are valid for the general assumptions $i)$ and $ii)$ and that our findings emerge as a concise and direct result of the point process perspective, in particular, as a result of using Campbell's Theorem.

C. The Poisson Special-Case $i)^\dagger$

By invoking $i)^\dagger$ we immediately have (recall Sec. III-C)

$$Y \sim \texttt{PoissonPP}(\mathbb{R}, \varrho_Y), \qquad \varrho_Y(y) = \lambda_\mathrm{e},\ y \in \mathbb{R}.$$

Due to $ii)$, the marks $\{p_y : y \in Y\}$ are mutually independent and each individual period p_y does not depend on $Y \backslash \{y\}$. Thus, by the Marking Theorem (recall again Sec. III-C), it follows that the marked collection X in (9) is itself a Poisson point process, namely

$$X \sim \texttt{PoissonPP}(\mathbb{R} \times \mathbb{R}_+, \varrho_X).$$

The associated intensity function $\varrho_X(\cdot)$ reads [25, Prop. 3.9]

$$\varrho_X(\boldsymbol{x}) = \varrho_Y(y) f_\mathrm{period}(p), \qquad \boldsymbol{x} = (y,p). \quad (16)$$

To ease the notation it is often convenient to drop the index y on p_y as we just did. We simply think of each $\boldsymbol{x} \in X$ as a two-dimensional point (y,p) and not as a one-dimensional point with a mark attached. Notice that neither $\varrho_Y(\cdot)$ nor $f_\mathrm{period}(\cdot)$ depends on y. Hence, the intensity function $\varrho_X(\cdot)$ varies only with its second argument p via the shape of $f_\mathrm{period}(\cdot)$.

In view of the two-dimensional Poisson point process X it is a straightforward exercise to see that $L(t)$ is Poisson distributed for any fixed $t \in \mathbb{R}$. In fact, $L(t)$ as defined in (10) can now be recognized as a region count associated with X. The region in question must necessarily be indexed by time t and from inspection of the two indicator functions in (10) we "mechanically" define

$$B_t := \{(y,p) : y \le t,\ y + p > t\} \subset \mathbb{R} \times \mathbb{R}_+. \quad (17)$$

With this definition of the region B_t it follows that $L(t)$ as defined in (10) coincides directly with the number of points from X falling within B_t, i.e. (see also Fig. 2)

$$L(t) = N_X(B_t) = \sum_{\boldsymbol{x} \in X} \mathbb{1}[\boldsymbol{x} \in B_t], \qquad t \in \mathbb{R}. \quad (18)$$

Since X is a Poisson point process the region count $N_X(B_t)$ is a Poisson distributed random variable (recall Sec. III-C). By the identity in (18) it follows immediately that $L(t)$ is Poisson distributed. We emphasize the simplicity of this argument and we highlight how our argument relates naturally to the point process framework. In particular, no results from queuing theory were needed as compared to the approach in [7].

Under the facilitating assumption $i)^\dagger$, we have now shown that the strict-sense stationary random process $L(\cdot)$ has Poisson distributed marginals, i.e. the random variable $L(t)$ is *Poisson distributed* for any fixed t. This is quite natural since under $i)^\dagger$ both Y and X are Poisson point processes. In the general case $i)$, however, no such exceptional circumstance holds jointly for Y and X. In the general case we have already
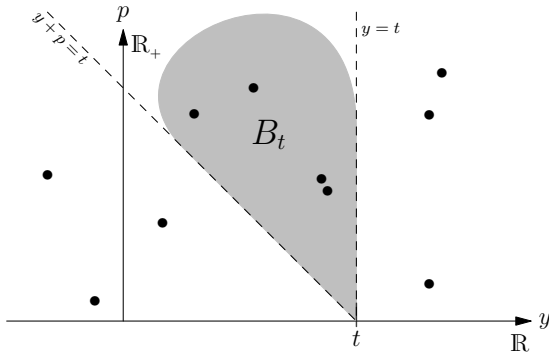
Fig. 2. The region $B_t \subset \mathbb{R} \times \mathbb{R}_+$ for arbitrary $t \in \mathbb{R}$ with black bullets representing points from X. Points falling within B_t indicate path components which are *active* in the radio channel at time t, i.e. $L(t) = N_X(B_t)$. Notice that the collection of points from X contributing jointly to both $L(t)$ and $L(t')$ are exactly those falling in the region $B_t \cap B_{t'}$.

calculated in (12) the mean $\mathbb{E}[L(t)]$ by use of the law of total expectation and Campbell's Theorem. It is nonetheless informative to reconsider the steps leading to (12) in the light of the two-dimensional perspective when X is of Poisson type. Specifically, the mean of the region count $N_x(B_t)$, and hence the mean of $L(t)$, is given by the value of the associated intensity measure at B_t, namely as (recall the general relationship in (4))

$$\mu_x(B_t) = \int_{B_t} \varrho_X(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \lambda_e \int_0^\infty p f_{\text{period}}(p) \mathrm{d}p = \lambda_e \mathbb{E}[p_\star].$$

The above integration step is natural in the sense that is it "dictated" by the point process framework: to get the intensity measure just integrate the intensity function. In particular, we did not even make use of Campbell's Theorem as opposed to the more involved steps leading to (12)[6].

Another analytical benefit of Y and X being Poisson point processes is that we can explicitly calculate the cumbersome autocorrelation term $\mathbb{E}[(\diamond_2)]$ in (13). Since Y is a Poisson point process, special-case use of Campbell's Theorem yields [25, Prop. 4.1]

$$\mathbb{E}[(\diamond_2)] = \mathbb{E}[L(t)]\mathbb{E}[L(t')] = (\lambda_e \mathbb{E}[p_\star])^2.$$

Hence, under the facilitating assumption $i)^\dagger$ we find the autocorrelation function of $L(\cdot)$ to be

$$R_L(t,t') = \mathbb{E}[(\diamond_1)] + \mathbb{E}[(\diamond_2)]$$
$$= \lambda_e \int_{|\Delta t|}^\infty \Pr(p_\star > \xi) \mathrm{d}\xi + (\lambda_e \mathbb{E}[p_\star])^2. \quad (19)$$

Notice that by evaluating (19) at $\Delta t = 0$, one immediately identifies the well-known property that the mean and the variance of a Poisson distributed random variable coincide.

Example 1. Assume that $ii)^\dagger$ holds as in [7]–[10]. We introduce $\mathbb{E}[p_\star] = 1/\lambda_v$ with the subscript "v" abbreviating "vanish". Then (19) specializes to

$$R_L(t,t') = \frac{\lambda_e}{\lambda_v} \exp\left(-\lambda_v |\Delta t|\right) + \left(\frac{\lambda_e}{\lambda_v}\right)^2. \quad \square$$

[6]This observation is well-aligned with the claim by Kingman in the preface of his book [24], namely that the theory of (Poisson) point processes is more natural and powerful in higher dimensions.

Example 2. For comparison we assume instead that all periods in the collection $\{p_y : y \in Y\}$ are uniformly distributed on the interval $[0, \frac{2}{\lambda_v}]$. The mean parameter is still $1/\lambda_v$ but in this case the autocorrelation function in (19) specializes to

$$R_L(t,t') =$$
$$\lambda_e \left(\frac{\lambda_v}{4} |\Delta t|^2 - |\Delta t| + \frac{1}{\lambda_v}\right) \mathbb{1}\left[|\Delta t| \leq \frac{2}{\lambda_v}\right] + \left(\frac{\lambda_e}{\lambda_v}\right)^2.$$

As intuitively expected, when the periods are uniformly distributed on the interval $[0, \frac{2}{\lambda_v}]$, we find that the Poisson distributed random variables $L(t)$ and $L(t')$ are uncorrelated whenever t and t' are displaced further than $2/\lambda_v$ apart. $\quad \square$

The result in (19) was derived in [23, Sec. 5.6 (iii)] under the facilitating special-case assumption $i)^\dagger$. However, the derivation is more involved since [23, Sec. 5.6 (iii)] does not rely on the straightforward application of Campbell's Theorem.

D. Relaxing The General Assumptions $i)$ and $ii)$

A compelling motivation for employing the special-case assumptions $i)^\dagger$ and $ii)^\dagger$ is that these together lead to (quoting [7]) *simple mathematics*. A practically relevant question is whether these general assumptions can be relaxed while preserving the ability to assess and comprehend the analytical properties inherited by $L(\cdot)$.

From the result in (12), one may tend to think that the relationship $\mathbb{E}[L(t)] = \lambda_e \mathbb{E}[p_\star]$ would stay unaffected if we relaxed $i)$ such that Y would just be homogenous and not stationary. Furthermore, $ii)$ could potentially be relaxed such that the marks were no longer identically distributed but still sharing the same mean parameter $\mathbb{E}[p_\star]$. Obviously, $L(\cdot)$ would then no longer be strict-sense stationary as all arguments from Appendix A would break down. However, $L(\cdot)$ could potentially remain wide-sense stationary. A simple construction inspired from Examples 1 and 2 in the previous subsection shows that this is not the case.

Example 3. Consider drawing the periods using a threshold procedure at the origin such that

$$p_y \sim f_{\text{period}}(p) = \begin{cases} \lambda_v e^{-\lambda_v p} & \text{if } y \geq 0, \\ \frac{\lambda_v}{2} \mathbb{1}[0 \leq p \leq \frac{2}{\lambda_v}] & \text{if } y < 0. \end{cases} \quad (20)$$

One can then readily check via (11) and (12) that the particular construction in (20) gives rise to a mean function depending on time t in such a way that $\mathbb{E}[L(t)] = \frac{\lambda_e}{\lambda_v}$ for $t < 0$ and

$$\mathbb{E}[L(t)] =$$
$$\frac{\lambda_e}{\lambda_v}\left(1 - e^{-\lambda_v t}\right) + \lambda_e \left(\frac{\lambda_v}{4} t^2 - t + \frac{1}{\lambda_v}\right) \mathbb{1}\left[0 \leq t \leq \frac{2}{\lambda_v}\right]$$

for $t \geq 0$ (see Fig. 3). Hence, the mean $\mathbb{E}[L(t)]$ is clearly affected by the sudden change of the mark distribution as given in (20). However, as time progresses the aftereffect of this change in $f_{\text{period}}(\cdot)$ becomes less and less noticeable (the birth-death mechanism stabilizes again). $\quad \square$

The example-construction in (20) shows that it affects the properties of the temporal birth-death process $L(\cdot)$ when the periods are not identically distributed. The influence of (20)
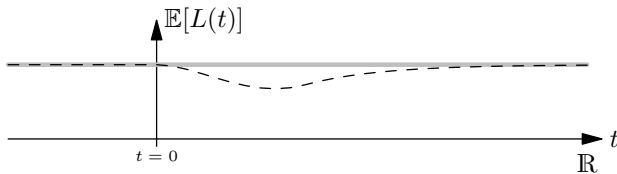
Fig. 3.   The effect of (20) upon the mean $\mathbb{E}[L(t)]$ from (12) as a result of an attempt to relax the assumptions $i)$ and $ii)$.
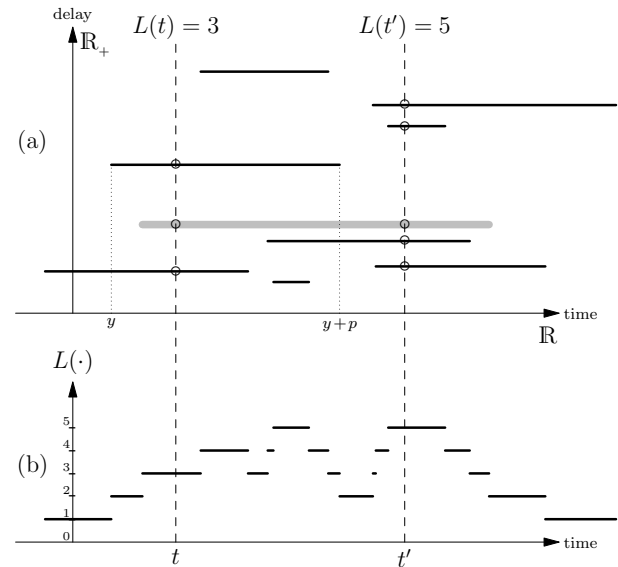

Fig. 4.   Line segments in (a) indicating the time instances of birth-death transition of the random process $L(\cdot)$ in (b). In this particular illustration, notice that only a single path component (thick gray line) contributes *jointly* to the count $L(t) = 3$ and the count $L(t') = 5$.

on the autocorrelation function $\mathbb{E}[L(t)L(t')]$ is even more complicated than for the mean $\mathbb{E}[L(t)]$. In conclusion, to preserve analytical tractability of the birth-death process $L(\cdot)$, the assumption $ii)$ cannot be relaxed. Assumption $i)$ can be relaxed in the sense of replacing the term "stationary" with "homogeneous", and the resulting birth-death process $L(\cdot)$ will *sometimes* remain wide-sense stationary but examples where this is not the case can be readily constructed.

V. THE TIME-FREQUENCY CORRELATION FUNCTION

In this section our goal is to calculate and analyze the structure of the *time-frequency correlation function*

$$R_H(t, t', f, f') := \mathbb{E}[H^*(t, f)H(t', f')]. \quad (21)$$

This function is often considered for radio channel characterization as the information it carries is useful for several reasons [28]. Among others, the time-frequency correlation function reveals if the time-variant channel transfer function $H(\cdot, \cdot)$ in (2) is wide-sense stationary (in time, in frequency or in both domains simultaneously). This knowledge is crucial since in practice we typically ask for stochastic models which indeed are wide-sense stationary, mainly to facilitate analytical insight and to simplify our system designs. A frequently recurring example application of the time-frequency correlation function emerges in *linear minimum mean-squared error* channel estimation (see e.g. [29] for an OFDM use case). Wide-sense stationarity of $H(\cdot, \cdot)$ in both time and frequency notably simplifies the design and implementation of such channel estimators. Furthermore, wide-sense stationarity of $H(\cdot, \cdot)$ in frequency allows for inferring on the rate of decay of received power versus propagation delay, namely, to infer on the channel's power-delay profile. Power-delay profiles are crucial to the design of modern positioning and communication systems and such profiles are often estimated in practice using channel sounding measurements.

Calculating the correlation function in (21) is conceptually straightforward when the number of path components (deterministic or random) remains fixed within each realization of the channel in (2). The traditional procedure is to introduce two separate integer-indices ℓ and k, one for each term in (21), and then to pair together those path components for which $\ell = k$. However, in the current time-variant setup we cannot compute (21) by traditional means since the temporal birth-death process $L(\cdot)$ in (2) is changing without explicit reference to the enumeration of the underlying random quantities (path gains and propagation delays). The overall situation and the enumeration issue are together illustrated in Fig. 4. Hence, with the considered class of temporal birth-death channel

models, we necessarily need to account for the enumeration issue.

A. Formulating a Generic Channel Transfer Function Using the Point Process X

To compute (21) we bypass the enumeration issue by adopting our point process perspective from Sec. IV. The basic idea is to replace the integer-indexed sum in the traditional expression (2) by a point process indexed sum like the one in (10). Motivated by the previous modeling approaches in [7]–[10] we consider here a generic multipath representation with

$$H(t, f) = \sum_{\boldsymbol{x} \in X} w(\boldsymbol{x}, t, f) H_{\boldsymbol{x}}(t, f), \quad (22)$$

where X is the two-dimensional point process defined in (9). Our definition in (22) displays a deliberate choice, namely that we assign $\{w(\boldsymbol{x}, \cdot, \cdot) : \boldsymbol{x} \in X\}$ as *window functions* to "implement" the birth-death process $L(\cdot)$ while $\{H_{\boldsymbol{x}}(\cdot, \cdot) : \boldsymbol{x} \in X\}$ are "per path" contributions to the overall channel transfer function. Conditioned on each $\boldsymbol{x} \in X$, the corresponding window $w(\boldsymbol{x}, \cdot, \cdot)|\boldsymbol{x}$ is considered a *deterministic function* (of t and f) whereas $H_{\boldsymbol{x}}(\cdot, \cdot)|\boldsymbol{x}$ is a *random process*. These choices are indeed arbitrary (i.e. other choices could have been made instead) but our definition in (22) intentionally allows to restore the models in [7]–[10] and allows further to readily assess (21).

Each window $w(\boldsymbol{x}, \cdot, \cdot)$ must have a temporal support corresponding to the indicator function $\mathbb{1}[\boldsymbol{x} \in B_t]$ since from (18) we recall how the birth-death process $L(\cdot)$ can be interpreted as a "time-sliding" region count. In general, the window functions (which could depend on frequency) serve to provide sufficiently smooth birth-death transitions as proposed in [8] and [9] (recall Sec. II).

Example 4. The simplest to consider is crude "on/off" windows which do not depend on frequency, namely (recall (17))

$$w(\boldsymbol{x}, t, f) = \mathbb{1}[\boldsymbol{x} \in B_t], \qquad \boldsymbol{x} = (y, p) \in X. \quad (23)$$

Example 5. As suggested in [8], the window functions can be defined such that

$$w((y, p), t, f) = \sin\left(\pi \frac{t-y}{p}\right) \mathbb{1}\left[(y, p) \in B_t\right],$$

leading to smooth birth-death transitions. As mentioned in [9], the windows can also be defined using other smooth functions, e.g. root-raised cosines with the same temporal support. □

For the "per path" component processes $\{H_{\boldsymbol{x}}(\cdot, \cdot)\}$ we assume a zero-mean condition together with mutual uncorrelatedness:

$$\mathbb{E}[H_{\boldsymbol{x}}(t, f) | \boldsymbol{x}] = 0, \qquad \boldsymbol{x} \in X, \quad (24)$$
$$\mathbb{E}[H_{\boldsymbol{x}}^*(t, f) H_{\widetilde{\boldsymbol{x}}}(t', f') | \boldsymbol{x}, \widetilde{\boldsymbol{x}}] = R_2(\boldsymbol{x}; t, t', f, f') \mathbb{1}[\boldsymbol{x} = \widetilde{\boldsymbol{x}}]. \quad (25)$$

The random processes $\{H_{\boldsymbol{x}}(\cdot, \cdot)\}$ can be defined in numerous different ways. The following constructions are illustrative examples which we shall also get back to later in this section.

Example 6. For the geometric-stochastic modeling approach in [9], the random processes $\{H_{\boldsymbol{x}}(\cdot, \cdot)\}$ can be defined as

$$H_{\boldsymbol{x}}(t, f) = \alpha_{\boldsymbol{x}}(t) \mathrm{e}^{-j2\pi f \tau_{\boldsymbol{x}}(t)}, \qquad \boldsymbol{x} \in X, \quad (26)$$

with the point $\boldsymbol{x}$ serving merely as an index. In fact, with the windows in (23) and the random processes in (26) we restore the traditional expression in (2), only now it is indexed differently. *Thus, we use the points from X to index individual random processes while simultaneously generating the birth-death process $L(\cdot)$ using the very same points from X.* □

Example 7. A frequently encountered special-case of (26) is

$$H_{\boldsymbol{x}}(t, f) = \alpha_{\boldsymbol{x}} \mathrm{e}^{j2\pi t\nu_{\boldsymbol{x}}} \mathrm{e}^{-j2\pi f \tau_{\boldsymbol{x}}}, \qquad \boldsymbol{x} \in X. \quad (27)$$

This model can be constructed from the marked point process

$$\{(\boldsymbol{x}, (\alpha_{\boldsymbol{x}}, \nu_{\boldsymbol{x}}, \tau_{\boldsymbol{x}})) : \boldsymbol{x} \in X\}, \quad (28)$$

using a mark collection $\{(\alpha_{\boldsymbol{x}}, \nu_{\boldsymbol{x}}, \tau_{\boldsymbol{x}}) : \boldsymbol{x} \in X\}$ of i.i.d. members indexed via the points from X. Each three-dimensional mark designates a triplet of a path gain, a Doppler shift, and a propagation delay, respectively. The construction in (28) is fundamentally similar to the one in (9), only now the marks are random vectors instead of scalar random variables. □

B. Assessing the Time-Frequency Correlation Function Using Campbell's Theorem

By (24) it follows readily that the generic channel transfer function in (22) has zero mean. To calculate the time-frequency correlation function we enter (22) in (21) and proceed at first via intermediate conditioning on X (the law of total expectation) such that

$$R_H(t, t', f, f') = \mathbb{E}\big[\mathbb{E}[H^*(t, f) H(t', f') | X]\big] \quad (29)$$
$$= \mathbb{E}\Big[\sum_{\boldsymbol{x} \in X} w^*(\boldsymbol{x}, t, f) w(\boldsymbol{x}, t', f') R_2(\boldsymbol{x}; t, t', f, f')\Big], \quad (30)$$

where the second equality is due to the assumption in (25). Then, by application of Campbell's Theorem we get

$$R_H(t, t', f, f') =$$
$$\int w^*(\boldsymbol{x}, t, f) w(\boldsymbol{x}, t', f') R_2(\boldsymbol{x}; t, t', f, f') \varrho_X(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \quad (31)$$

The integral expression in (31) is one of our main results which we shall use in the following to analyze a number of special-case scenarios where further assumptions are introduced. The recently obtained results in [13] are centered around a similar integral-type correlation function, but that particular channel modeling framework did not include temporal birth-death transitions of path components.

A simple but relevant restriction that can be invoked is when the correlation function $R_2(\boldsymbol{x}; t, t', f, f')$ from (25) does not depend on $\boldsymbol{x}$, which is the case when $\boldsymbol{x}$ serves only as an index (e.g. in the example in (27)). Then (31) factorizes as

$$R_H(t, t', f, f') = R_1(t, t', f, f') R_2(t, t', f, f') \quad (32)$$

with

$$R_1(t, t', f, f') := \int w^*(\boldsymbol{x}, t, f) w(\boldsymbol{x}, t', f') \varrho_X(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \quad (33)$$

The product form in (32) holds an attractive and intuitive interpretation: The first term R_1 reports the correlation properties of the temporal birth-death process $L(\cdot)$ via the associated window functions (a large-scale quantity). The second term R_2 reports the "per path" time-frequency correlation properties (a small-scale quantity). Hence, if changes are made to the window functions $w(\boldsymbol{x}, \cdot, \cdot)$ in (22) these affect the correlation term R_1 whereas changes made to the component processes $H_{\boldsymbol{x}}(\cdot, \cdot)$ affect the correlation term R_2.

C. A Doubly Wide-Sense Stationary Special-Case Model

The class of doubly wide-sense stationary stochastic channel models has been favored both in literature and practice ever since Bello introduced his seminal *WSSUS characterization* [28]. Bello did not assume a specific form of the time-variant channel transfer function as we do here. Instead he treated the overall channel transfer function $H(\cdot, \cdot)$ as a general stochastic process in two variables, but for consistency we restrict here our attention to the repeatedly favored multipath model in (22).

When the on/off windows from (23) are employed, the frequency-variable dependencies disappear and the correlation term in (33) simplifies to

$$R_1(t, t') = \int_{B_t \cap B_{t'}} \varrho_X(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \mu_X(B_t \cap B_{t'})$$
$$= \mathbb{E}\big[N_X(B_t \cap B_{t'})\big],$$

i.e. the first factor in (32) provides the expected number of points from X contributing *jointly* to both $H(t, \cdot)$ and $H(t', \cdot)$, recall Fig. 2 and Fig. 4. Notice that the region count $N_X(B_t \cap B_{t'})$ coincides with the random quantity $(\diamond_1)$ defined in (14) in Sec. IV-B and its expectation has already been calculated in (15). Since the expression in (15) is a function of Δt only, so is the correlation term $R_1(t, t') = R_1(\Delta t)$. If we specialize to traditional modeling approaches where $L(t) = L$

is constant within individual channel realizations, then all we have to modify is the term $R_1(\Delta t)$. If L is a fixed constant then $R_1(\Delta t)$ is to be replaced by L itself, whereas if L is a random variable, then $R_1(\Delta t)$ is to be replaced by $\mathbb{E}[L]$. However, for the temporal birth-death channels considered in this paper the (large-scale) correlation term $R_1(\Delta t)$ is given by $\mathbb{E}[N_x(B_t \cap B_{t'})]$, i.e. the expected number of path components contributing *jointly* to both $H(t, \cdot)$ and $H(t', \cdot)$.

An important additional restriction to consider is when the "per-path" correlation function in (25), apart from not depending on $\boldsymbol{x}$, also does not depend directly on (t, t', f, f') but only on the time and frequency differences $\Delta t = t' - t$ and $\Delta f = f' - f$. In this case the time-frequency correlation function in (32) specializes to

$$R_H(\Delta t, \Delta f) = R_1(\Delta t) R_2(\Delta t, \Delta f) \qquad (34)$$

which overall is a function of time and frequency differences only. Hence, the generic channel transfer function $H(\cdot, \cdot)$ in (22) can (via suitable assumptions) become *wide-sense stationary in both time and frequency* (despite the birth-death behavior of individual path components).

Example 8. For the simplest but frequently encountered special-case model in (27) the general conditions in (24) and (25) are easily satisfied. We can for instance employ the assumption that the i.i.d. vector marks are drawn from a joint distribution such that

$$\mathbb{E}[\alpha_{\boldsymbol{x}} \,|\, \nu_{\boldsymbol{x}}, \tau_{\boldsymbol{x}}] = \mathbb{E}[\alpha_\star \,|\, \nu_\star, \tau_\star] = 0, \qquad \boldsymbol{x} \in X, \qquad (35)$$

where we make use of a wildcard notation like in (8). The assumption in (35) is usually justified from a default argument of uniformly distributed *initial phases*. With i.i.d. marks and (35) the condition in (25) is satisfied exactly such that

$$R_2(\boldsymbol{x}; t, t', f, f') = \mathbb{E}\left[|\alpha_\star|^2 e^{j2\pi(\Delta t \nu_\star - \Delta f \tau_\star)}\right] = R_2(\Delta t, \Delta f)$$

where the expectation is with respect to the joint distribution of a typical/arbitrary vector mark $(\alpha_\star, \nu_\star, \tau_\star)$. $\qquad \square$

In contrast to the special-case construction in (27), the more general model in (26) does not straightforwardly hold the ability to induce (double) wide-sense stationarity in (22). At a first glance it may seem rather innocent to ask for how and when (26) will induce wide-sense stationarity. *Surprisingly, it seems that the question is not so innocent after all.* If we make the simplistic assumption of $\alpha_{\boldsymbol{x}}(\cdot)$ and $\tau_{\boldsymbol{x}}(\cdot)$ being independent random processes (for each $\boldsymbol{x} \in X$) it means that we need to draw our conclusions via

$$R_2(\boldsymbol{x}; t, t', f, f') = \mathbb{E}[\alpha_{\boldsymbol{x}}^*(t)\alpha_{\boldsymbol{x}}(t')]\mathbb{E}\left[e^{-j2\pi(f'\tau_{\boldsymbol{x}}(t') - f\tau_{\boldsymbol{x}}(t))}\right].$$

Obviously, the random process $\alpha_{\boldsymbol{x}}(\cdot)$ must be wide-sense stationary in order to make the first expectation a function of Δt only. Surprisingly, the only way for the second expectation to be a function of Δf is when the random process $\tau_{\boldsymbol{x}}(\cdot)$ has constant realizations (and hence it does not depend on t and t'). A simple argument for this claim can be found in Appendix B.

D. Principal Representation of $H(\cdot, \cdot)$ and Its Associated Time-Frequency Correlation Function

The insight gained so far motivates the assumption of window functions from (23), random but constant propagation delays, and wide-sense stationary path gain processes. Accordingly, we introduce a tractable special-case expression of the generic channel transfer function in (22) reading

$$H(t, f) = \sum_{\boldsymbol{x} \in X} \mathbb{1}[\boldsymbol{x} \in B_t] \alpha_{\boldsymbol{x}}(t) e^{-j2\pi f \tau_{\boldsymbol{x}}}. \qquad (36)$$

This choice of the "per path" processes can be seen as a natural intermediate between (26) and (27) such that the overall model in (36) has the ability to become doubly wide-sense stationary. We now make use of a collection of random *processes* (path gains) together with a collection of random *marks* (propagation delays) with both collections being indexed via the points from X. We assume that the marks $\{\tau_{\boldsymbol{x}} : \boldsymbol{x} \in X\}$ comprise an i.i.d. collection drawn from a probability density function $f_{\text{delay}}(\cdot)$. Conditioned on all marks, we assume that the random processes $\{\alpha_{\boldsymbol{x}}(t) : \boldsymbol{x} \in X\}$ all hold zero mean, are mutually uncorrelated, and individually exhibit an autocorrelation function parameterized via the corresponding mark:

$$\mathbb{E}[\alpha_{\boldsymbol{x}}(t) \,|\, \tau_{\boldsymbol{x}}] = 0, \qquad \boldsymbol{x} \in X, \qquad (37)$$

$$\mathbb{E}[\alpha_{\boldsymbol{x}}^*(t)\alpha_{\widetilde{\boldsymbol{x}}}(t') \,|\, \tau_{\boldsymbol{x}}, \tau_{\widetilde{\boldsymbol{x}}}] = R_\alpha(\Delta t; \tau_{\boldsymbol{x}}) \mathbb{1}[\boldsymbol{x} = \widetilde{\boldsymbol{x}}], \qquad (38)$$

Among other virtues, (38) allows to conveniently assign conditional average power to each random process $\alpha_{\boldsymbol{x}}(\cdot)$ as a function of its corresponding propagation delay $\tau_{\boldsymbol{x}}$. Notice that (37) ensures (24) to be satisfied and furthermore that (38) ensures (25) to be satisfied with $R_2(\boldsymbol{x}; t, t', f, f') = R_2(\Delta t, \Delta f)$.

By repeating the previous calculations we now obtain a time-frequency correlation function for the model in (36) with

$$R_H(\Delta t, \Delta f) = R_1(\Delta t)\mathbb{E}\left[R_\alpha(\Delta t; \tau_\star) e^{-j2\pi\Delta f \tau_\star}\right], \qquad (39)$$

where the expectation is with respect to $f_{\text{delay}}(\cdot)$, i.e. we utilize again a wildcard notation $\tau_\star$ to represent a typical mark. By (39), the time-variant channel transfer function $H(\cdot, \cdot)$ in (36) is wide-sense stationary in both time and frequency.

To summarize this section we emphasize the tools which enabled our novel assessment of the time-frequency correlation function with the most general form as given in (31), and specialized in (32), (34), and (39). In a nutshell, our key step was to reformulate and generalize the traditional expression (2) into the generic expression in (22). This we did in order to circumvent the enumeration issue induced by the traditional integer-indexing of path components. Specifically, we reformulated (2) using the point process X which is *the same* random collection we also used for generating the temporal birth-death process $L(\cdot)$ in Sec. IV. Our generic representation in (22) was then employed in direct substitute of the traditional and widely accepted model (2) of the channel transfer function. The inherent structure of the generic substitute representation is the reason for it being appropriate for analysis using the combined application of the law of total expectation and Campbell's Theorem.

VI. Selected Examples and Simulation Aspects

In this section we show via specific examples how key parameters of the temporal birth-death channel model enter explicitly in quantities which can be measured in practice, e.g. the power-delay profile. Furthermore, depending on its purpose, we show how the channel model can be readily modified to incorporate fewer or more stochastic features. Finally, we illustrate from a simulation technical point of view the facilitating and convenient aspects of the assumptions $i)^{\dagger}$ and $ii)^{\dagger}$.

A. Power-Delay Profile Induced From A Separation Property

In Sec. V-D we did not introduce a specific choice for the autocorrelation function $R_\alpha(\cdot; \cdot)$ in (38). By making such a choice and by making also explicit choices about the probability density functions $f_{\text{period}}(\cdot)$ and $f_{\text{delay}}(\cdot)$, we end up with a final and specific expression for the time-frequency correlation function in (39).

Example 9. A simple choice for the autocorrelation function $R_\alpha(\cdot; \cdot)$ in (38) is to assume that

$$\mathbb{E}[\alpha_{\boldsymbol{x}}^*(t)\alpha_{\boldsymbol{x}}(t') \,|\, \tau_{\boldsymbol{x}}] = R_\alpha(\Delta t; \tau_{\boldsymbol{x}}) = \sigma_\alpha^2(\tau_{\boldsymbol{x}})J_0(2\pi\eta\Delta t),$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, η is a positive parameter to be specified, and $\sigma_\alpha^2(\cdot)$ assigns conditional average power to each random process $\alpha_{\boldsymbol{x}}(\cdot)$ as a function of its corresponding propagation delay $\tau_{\boldsymbol{x}}$. Similar assumptions (or choices) are often found in the literature [8], [29]. The use of the Bessel function $J_0(\cdot)$ originates from Clarke's seminal work [30]. Notice that we in fact assume the correlation function to be the same for all individual path gain processes $\{\alpha_{\boldsymbol{x}}(t) : \boldsymbol{x} \in X\}$, except for an individual scaling by $\sigma_\alpha^2(\tau_{\boldsymbol{x}})$. With these choices the time-frequency correlation function in (39) is seen to factorize in a product of time and frequency separated terms, namely as

$$R_H(\Delta t, \Delta f) = R_{\text{time}}(\Delta t)R_{\text{freq}}(\Delta f), \tag{40}$$

$$R_{\text{time}}(\Delta t) := R_1(\Delta t)J_0(2\pi\eta\Delta t), \tag{41}$$

$$R_{\text{freq}}(\Delta f) := \mathbb{E}\big[\sigma_\alpha^2(\tau_\star)e^{-j2\pi\Delta f\tau_\star}\big]. \tag{42}$$

The product form in the right-hand side of (40) reflects the separation property mentioned in [29]. An immediate and practically convenient consequence of the separation property in (40) is that it notably simplifies the design of linear minimum mean-squared error estimators of the channel transfer function $H(\cdot, \cdot)$, e.g. for OFDM applications [29].

Notice how even (41) factorizes, namely as a product of *large-scale* and *small-scale* fading induced correlation properties (recall also the discussion below (32) in Sec. V-B). To further concretize (41) we could invoke assumption $ii)^{\dagger}$ as done in [7]–[10] and in Example 1 in Sec. IV-B. The large-scale correlation term $R_1(\Delta t)$ would then exhibit an exponential decay. Specifically, we would have

$$R_{\text{time}}(\Delta t) = \frac{\lambda_{\text{e}}}{\lambda_{\text{v}}}\exp(-\lambda_{\text{v}}|\Delta t|)J_0(2\pi\eta\Delta t).$$

Exponentially decaying correlation functions for large-scale fading processes (*shadowing*) have been suggested in the literature several times, see e.g. [31] for one of the earliest occurrences. We have selected the Bessel function $J_0(\cdot)$ as the small-scale correlation structure only to provide a specific and popular example. Obviously, we could have made any other selection for the normalized correlation function and substituted it directly into (41) instead of $J_0(\cdot)$.

The expression in (42) holds important information as well. Specifically, as we can immediately identify (42) as the Fourier transform

$$R_{\text{freq}}(\Delta f) = \mathcal{F}\{\sigma_\alpha^2(\cdot)f_{\text{delay}}(\cdot)\}(\Delta f)$$

we find that the *power-delay profile* of the channel is

$$P_{\text{delay}}(\tau) := R_{\text{time}}(0)\mathcal{F}\{R_{\text{freq}}(\cdot)\}(\tau) = \frac{\lambda_{\text{e}}}{\lambda_{\text{v}}}\sigma_\alpha^2(\tau)f_{\text{delay}}(\tau).$$

This function characterizes the dispersion and rate of decay of received power versus propagation delay. Knowledge of the power-delay profile (sometimes also called the *delay-power spectrum*) is crucial for localization aspects and wireless communications in general. Motivated by empirical observations, a standard assumption is that $P_{\text{delay}}(\cdot)$ exhibits an overall exponential decay. In [8], [9] such an exponential decay is maintained by appropriate selections of the conditional power assigning function $\sigma_\alpha^2(\cdot)$ and the probability density function $f_{\text{delay}}(\cdot)$.

To conclude this example we stress the importance of being able to calculate and analytically assess the functions $R_{\text{time}}(\cdot)$, $R_{\text{freq}}(\cdot)$, $P_{\text{delay}}(\cdot)$, etc. The importance lies in the fact that key parameters of the channel model, such as λ_{e} and λ_{v}, enter explicitly in these practically measurable quantities. This offers a potential avenue for new and rigorously motivated parameter estimators and enables the temporal birth-death channel model to be utilized as a tool for measurement prediction. Such capabilities stand in notable contrast to the limitations of "pure" simulation models, where we usually do not know if different model parameters interact and where questions regarding stationarity properties most often remain inconclusive. □

At a first glance it may seem quite restrictive that we, in the above example, assumed a shared normalized correlation function for all individual path gains $\{\alpha_{\boldsymbol{x}}(t) : \boldsymbol{x} \in X\}$. A straightforward way to generalize this is to introduce a second collection of i.i.d. marks. We illustrate this idea in the sequel.

B. Modeling Flexibility

The forthcoming example illustrates an important attribute of the point process perspective: Once the underlying point process X in (9) is in place, we can in a natural and straightforward manner change the modeling details of the channel transfer function. Specifically, we can swap in dimensionality by attaching fewer or more marks to each $\boldsymbol{x} \in X$ depending on the number of features we wish the random process $H(\cdot, \cdot)$ to exhibit. Yet, the notation is not noticeably affected by such changes and the crucial details in the derivation of the resulting time-frequency correlation function (Sec. V-B) stay virtually unaffected. This attractive fact stands in notable contrast to a variety of model extensions proposed in the literature.

Certain well-known model extensions appear essentially to lack structure or to suffer from the absence of a profound theoretical modeling framework. Examples include Spencer's extension [6] of Saleh and Valenzuela's model [5] together with the ultra-wideband model [32] which is fundamentally based on [5] as well.

To extend the temporal birth-death channel model in (36) we can for example utilize the marked point process

$$\{(\boldsymbol{x}, (\tau_{\boldsymbol{x}}, \theta_{\boldsymbol{x}})) : \boldsymbol{x} \in X\}. \tag{43}$$

We still make use of a separate collection of random processes (path gains) but now the collection of random marks consists of vectors instead of scalars. Conceptually we are attaching two-dimensional marks $(\tau_{\boldsymbol{x}}, \theta_{\boldsymbol{x}})$ and the collection in (43) can of course be seen as a point process in a four-dimensional space (which may sometimes be useful, but not always). The newly added collection $\{\theta_{\boldsymbol{x}} : \boldsymbol{x} \in X\}$ models azimuth (incidence) directions for the individual path components as also addressed in e.g. [6] and [10]. Each $\theta_{\boldsymbol{x}}$ is drawn independently from a probability density function $f_{\text{azimuth}}(\cdot)$ with support set $[-\pi, \pi)$. The assumption in (38) is then replaced by

$$\mathbb{E}[\alpha_{\boldsymbol{x}}^*(t)\alpha_{\boldsymbol{x}}(t') \,|\, \tau_{\boldsymbol{x}}, \theta_{\boldsymbol{x}}] = \sigma_\alpha^2(\tau_{\boldsymbol{x}})R_\alpha(\Delta t; \theta_{\boldsymbol{x}}) \tag{44}$$

such that each path gain $\alpha_{\boldsymbol{x}}(\cdot)$ holds an individual *normalized* autocorrelation function $R_\alpha(\cdot; \theta_{\boldsymbol{x}})$ parameterized by $\theta_{\boldsymbol{x}}$. Based on (44), we now find that the resulting time-frequency correlation function takes the form

$$R_H(\Delta t, \Delta f) =$$
$$R_1(\Delta t)\mathbb{E}[R_\alpha(\Delta t; \theta_\star)]\mathbb{E}[\sigma_\alpha^2(\tau_\star)\mathrm{e}^{-j2\pi\Delta f\tau_\star}], \tag{45}$$

where the first expectation is with respect to $f_{\text{azimuth}}(\cdot)$ and the second with respect to $f_{\text{delay}}(\cdot)$.

Example 10. Conditioned on the two marks $(\tau_{\boldsymbol{x}}, \theta_{\boldsymbol{x}})$, we could for instance generate each path gain process $\alpha_{\boldsymbol{x}}(\cdot)$ in terms of azimuth-dispersed sub-components such that

$$\alpha_{\boldsymbol{x}}(t) := \sqrt{\frac{\sigma_\alpha^2(\tau_{\boldsymbol{x}})}{M}} \sum_{m=1}^M A_m \exp\left(j2\pi\eta\cos(\varphi_m)t\right)$$

where $\{A_m\} \stackrel{\text{i.i.d.}}{\sim} \mathcal{CN}(0, 1)$ and $\{\varphi_m\} \stackrel{\text{i.i.d.}}{\sim} \mathcal{VM}(\theta_{\boldsymbol{x}}, \kappa)$. Here we use $\mathcal{CN}$ and $\mathcal{VM}$ to denote the *complex normal* distribution and the *von Mises* distribution, respectively. Furthermore, (M, η, κ) are fixed parameters to be set according to the particular context (environment, physical speeds, carrier wavelength, etc.) but they are not of our main concern at present.

It follows that $\alpha_{\boldsymbol{x}}(t)|(\tau_{\boldsymbol{x}}, \theta_{\boldsymbol{x}}) \sim \mathcal{CN}(0, \sigma_\alpha^2(\tau_{\boldsymbol{x}}))$, and with some minor calculations we get (recall (44))

$$R_\alpha(\Delta t; \theta_{\boldsymbol{x}}) = \mathbb{E}[\exp(j2\pi\eta\cos(\varphi_\star)\Delta t)\,|\,\theta_{\boldsymbol{x}}] \tag{46}$$
$$= \frac{I_0(\sqrt{\kappa^2 - (2\pi\eta\Delta t)^2 + j4\pi\kappa\eta\Delta t\cos(\theta_{\boldsymbol{x}})}\,)}{I_0(\kappa)},$$

where in (46) the expectation is with respect to an arbitrary $\varphi_\star$ drawn from the von Mises density

$$f_{\mathcal{VM}}(\varphi; \theta_{\boldsymbol{x}}, \kappa) = \frac{\exp(\kappa\cos(\varphi - \theta_{\boldsymbol{x}}))}{2\pi I_0(\kappa)}, \quad \varphi \in [-\pi, \pi),$$

and $I_0(\cdot)$ denotes the zeroth-order *modified* Bessel function of the first kind. We refer to [33] for more details on the von Mises distribution and the properties of the resulting correlation function. To calculate the time-frequency correlation function in (45) we need to average (46) with respect to the density $f_{\text{azimuth}}(\cdot)$, i.e.

$$\mathbb{E}[R_\alpha(\Delta t; \theta_{\boldsymbol{x}})] = \mathbb{E}\big[\mathbb{E}[\exp(j2\pi\eta\cos(\varphi_\star)\Delta t)\,|\,\theta_{\boldsymbol{x}}]\big]$$
$$= \iint f_{\text{azimuth}}(\theta)f_{\mathcal{VM}}(\varphi; \theta, \kappa)\mathrm{e}^{j2\pi\eta\cos(\varphi)\Delta t}\mathrm{d}\varphi\mathrm{d}\theta, \tag{47}$$

where the region of integration is $[-\pi, \pi) \times [-\pi, \pi)$. The expression in (47) is difficult to assess except for certain special cases of the density $f_{\text{azimuth}}(\cdot)$. However, if we employ the assumption that $f_{\text{azimuth}}(\cdot)$ is the uniform density on $[-\pi, \pi)$, then (47) simplifies to the standard expression $J_0(2\pi\eta\Delta t)$ from Example 9. To see this, simply swap the order of integration in (47) and notice that $\int f_{\mathcal{VM}}(\varphi; \theta, \kappa)\mathrm{d}\theta = 1$ since the individual roles of the "variables" φ and θ are interchangeable in the von Mises density. Hence, from the very beginning we could have made the default selection of the Bessel function $J_0(2\pi\eta\Delta t)$ for the shared correlation function in (38). However, we could also have employed individual correlation functions as just shown. Yet, with the particular choices made the time-frequency correlation function in (45) coincide with the one from Example 9. $\square$

The above example illustrates a particular choice and recipe for how each path gain process $\alpha_{\boldsymbol{x}}(\cdot)$ can be parameterized and generated in a computer simulation. In the following we highlight a few selected details on tractable simulation procedures for the underlying birth-death process $L(\cdot)$. In addition to its analytical advantages, the point process perspective proves itself also particularly valuable for simulation purposes.

C. Simulation Aspects Regarding the Point Process X and the Temporal Birth-Death Process $L(\cdot)$

The approaches in [7]–[10] rely exclusively on the assumptions $i)^\dagger$ and $ii)^\dagger$. The motivation for using these assumptions is that they together endow the resulting channel model with *simple mathematics* [7]. In fact, $i)^\dagger$ and $ii)^\dagger$ are also particularly convenient for computer simulation, especially due to our theoretical knowledge from Sec. IV regarding the point process X.

The contribution in [7] focuses on modeling and experimental investigations and does not cover aspects related to computer simulation. Both [8] and [9] mention a few simulation aspects with example visualizations of generated channel impulse responses. However, exact details or guidelines are not provided. In particular, the enumeration issue related to the birth-death transitions of $L(\cdot)$ is not mentioned. In contrast, [10] proposes a number of heuristic guidelines for controlling, initializing and time-discretizing the temporal birth-death mechanism. For instance, the random process $L(\cdot)$ is always initialized such that no path components are present. Hence, if the *origin* is selected to be the arbitrary starting time, then [10] systematically assigns $L(0) = 0$. The motivation for this is (quoting [10, Sec. III-C]): *To avoid the problem of defining a specific starting state*. It is then suggested to initially let

the process run long enough to yield the "correct" value for $\mathbb{E}[L(t)]$. Based on $ii)^\dagger$, a value is then given for the minimum forerun needed to yield an error of at most 1% (in a certain sense). The actual simulation should subsequently take place after this forerun[7].

With our theoretical knowledge from Sec. IV, approximate simulation guidelines like the example mentioned above can be entirely circumvented. In particular, rather than *defining* a starting state, $L(0)$ should be drawn from a Poisson distribution instead of being systematically assigned to zero. Conditioned on $L(0)$, the task is then to calculate the conditional joint distribution of *emergence times* and *lifetimes* of those $L(0)$ path components which necessarily are present in the channel at initialization time $t = 0$. Obviously, all $L(0)$ path components must have emerged before time $t = 0$.

Example 11. Our goal in this example is to show the facilitating aspects of $i)^\dagger$ by itself. Accordingly, we combine $i)^\dagger$ with ii). In the following we show how to initialize the birth-death process $L(\cdot)$ using emergence times and lifetimes drawn from the equilibrium distribution.

Initially, recall Fig. 2 and shift the region B_t to the origin. Since X as defined in (9) is a Poisson point process we have that $L(0) = N_x(B_0)$ is a Poisson distributed random variable with mean $\mu_x(B_0) = \lambda_e\mathbb{E}[p_\star]$. Hence, we start by drawing the non-negative integer $L(0)$. By the second item of Definition 2 in Sec. III, we should then draw the $L(0)$ points $X \cap B_0$ mutually independent and identically distributed according to

$$f_{B_0}(\boldsymbol{x}) = \frac{\varrho_x(\boldsymbol{x})}{\mu_x(B_0)}\mathbb{1}[\boldsymbol{x} \in B_0] = \frac{f_{\text{period}}(p)}{\mathbb{E}[p_\star]}\mathbb{1}[(y,p) \in B_0],$$

where the intensity function $\varrho_x(\cdot)$ is taken from (16). Due to the shape of the unbounded triangular region B_0, it is most convenient to draw each two-dimensional point $\boldsymbol{x} = (y,p)$ by a two-step procedure. Marginalizing $f_{B_0}(\cdot)$ with respect to each of y and p yields respectively

$$f_y(y) := \int f_{B_0}(y,p)\mathrm{d}p \tag{48}$$
$$= \frac{\mathbb{1}[y \le 0]}{\mathbb{E}[p_\star]}\int_{-y}^{\infty} f_{\text{period}}(p)\mathrm{d}p = \frac{\Pr(p_\star > -y)}{\mathbb{E}[p_\star]}\mathbb{1}[y \le 0],$$

and

$$f_p(p) := \int f_{B_0}(y,p)\mathrm{d}y \tag{49}$$
$$= \frac{f_{\text{period}}(p)}{\mathbb{E}[p_\star]}\int_{-p}^{0} 1\mathrm{d}y = \frac{p\,f_{\text{period}}(p)}{\mathbb{E}[p_\star]}.$$

Both $f_y(\cdot)$ and $f_p(\cdot)$ integrate to unity. The corresponding conditional distributions read

$$f_{y|p}(y|p) = \frac{f_{B_0}(y,p)}{f_p(p)} = \frac{1}{p}\mathbb{1}[y \in (-p,0)] \tag{50}$$

and

$$f_{p|y}(p|y) = \frac{f_{B_0}(y,p)}{f_y(y)} = \frac{f_{\text{period}}(p)}{\Pr(p_\star > -y)}\mathbb{1}[p > -y], \tag{51}$$

which also both integrate to unity. Now, to generate a point $\boldsymbol{x} = (y,p)$, we can first generate the emergence time y according to $f_y(\cdot)$ in (48) and then generate the corresponding[8] period p according to $f_{p|y}(\cdot|y)$ in (51). Alternatively, we can first generate the period via (49) and then generate the corresponding emergence time via (50). The approach most preferable for implementation depends on our choice of the density $f_{\text{period}}(\cdot)$. In essence, to initialize the temporal birth-death process $L(\cdot)$ in equilibrium at time $t = 0$ we do as follows:

1. Draw $L(0) = N_x(B_0)$ from a Poisson distribution with mean $\lambda_e\mathbb{E}[p_\star]$.
2. Draw the points $X \cap B_0$ i.i.d. according to (48) and (51) (alternatively, use (49) and (50)).

With these two steps the temporal birth-death process $L(\cdot)$ exhibits its exact theoretical properties at initialization time $t = 0$ without imitating a forerun from the infinite past.   □

The example above illustrates in a convincing manner the benefits and the potential of the point process perspective. In contrast to the approximate and heuristic initialization guideline in [10], the channel can now be initialized exactly (in equilibrium) by means of a "mechanic procedure" dictated by the properties of the Poisson point process X (Definition 2 in Sec. III). Simulation aspects deduced specifically from assumption $ii)^\dagger$ can be found in [34].

VII. Conclusion

The theoretical framework of spatial point processes and its powerful tools, like Campbell's Theorem, comprise a natural environment for the engineering treatment of various stochastic radio channel models. Our analysis of the class of temporal birth-death channel models, governed by the assumptions i) and ii) in Sec. I, supports this conclusion and the usefulness of Campbell's Theorem has been demonstrated repeatedly. Overall, the proposed point process perspective is analytically beneficial due to its flexibilities with respect to dimensionality swapping and its ability to circumvent enumeration issues arising from the use of integer-indexed path components in traditional channel modeling approaches. Specifically, the key technique we employed consisted in replacing certain integer-indexed sums by equivalent expressions indexed by points from spatial point processes. In essence, this allows for keeping track of individual path components by use of the same stochastic mechanism which also generates the temporal birth-death behavior of the channel. In addition to its analytical advantages, the point process perspective has proven itself particularly valuable for simulation purposes as well. A complete and categorized overview of our findings is given in Fig. 5.

---

[7]This procedure resembles the well-known *burn-in* periods often used in *Markov Chain Monte Carlo* (MCMC) simulations [25, Sec. 8.1.2]. Such a burn-in is employed to ensure that the marginal distribution of the Markov chain's current state is sufficiently close to its (unknown) *equilibrium distribution* for all practical purposes.

[8]In this construction y and p are obviously dependent since we have conditioned on the fact that exactly $L(0)$ path components are present at time $t = 0$. To the contrary, the periods of those path components to enter in the future are to be drawn independently of their emergence times.
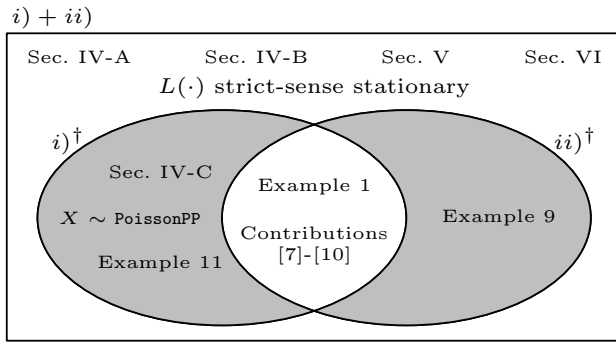
Fig. 5.   Overview of the present contribution in the form of a Venn-diagram. Notice that we have not indicated Sec. IV-D anywhere in the figure as this paragraph attempts to relax the assumptions $i)$ and $ii)$ displayed in terms of the rectangular box.

In Sec. IV we have shown that the temporal birth-death process $L(\cdot)$ is strict-sense stationary. The mean of $L(\cdot)$ does not depend on the exact shape of the probability density function induced via $ii)$, only the its first-order moment matters. However, the autocorrelation function of $L(\cdot)$ is directly affected via its shape. Finally, we indicated in Sec. IV the crucial roles of $i)$ and $ii)$ in that relaxation attempts in general turn $L(\cdot)$ into a non-stationary random process.

We derived in Sec. V an integral expression of the channel's time-frequency correlation function. To the best of our knowledge this general expression has not appeared elsewhere in the channel modeling literature. Under simplifying assumptions the time-frequency correlation function is comprised by the product of a large-scale and a small-scale term and we have shown that the channel transfer function can become wide-sense stationary in both time and frequency (despite the channel's temporal birth-death behavior).

We have exemplified in Sec. VI the paramount ability to explicitly calculate the time-frequency correlation function. Several key parameters of the birth-death channel model enter in practically measurable quantities such as temporal correlation functions and the power-delay profile. Immediate practical potentials are that new/novel parameter estimation procedures can be rigorously motivated and that the class of temporal birth-death channel models can as well be used as a tool for measurement prediction (as compared to a model class useful merely for simulation purposes).

APPENDIX A
STRICT-SENSE STATIONARITY OF THE TEMPORAL BIRTH-DEATH PROCESS $L(\cdot)$

Proposition. *As a consequence of* $i)$ *and* $ii)$, *the random process* $L(\cdot)$ *defined in* (10) *is strict-sense stationary.*

Proof: To see that $L(\cdot)$ is strict-sense stationary we have to show (for any fixed time shift $s \in \mathbb{R}$ and for any $k \in \mathbb{N}$) that

$$\Pr\big(L(t_1 + s) \le n_1, \ldots, L(t_k + s) \le n_k\big) \qquad (52)$$

does not depend on our choice of s. However, the common time shift in (52) corresponds in fact to nothing but a translation of the point process Y since by the definition of $L(t)$ in

(10) we have

$$L(t_i + s) = \sum_{y \in Y} \mathbb{1}[y \le t_i + s]\mathbb{1}[y + p_y > t_i + s]$$
$$= \sum_{\widetilde{y} \in \widetilde{Y}} \mathbb{1}\big[\widetilde{y} \le t_i,\ \widetilde{y} + p_{\widetilde{y}+s} > t_i\big], \quad i = 1, \ldots, k,$$

where $\widetilde{Y} := Y - s$ is the random collection $\{y - s : y \in Y\}$ of shifted points. Indeed, for any fixed $s \in \mathbb{R}$, the shifted collection $\widetilde{Y}$ is a stationary point process on the entire real line with the same statistical properties as Y. Additionally, the marks/periods are drawn i.i.d. irrespectively of the underlying point pattern. From this we now conclude that

$$\big(L(t_1 + s), \ldots, L(t_k + s)\big) \sim \big(L(t_1), \ldots, L(t_k)\big).$$

That is, (52) does not depend on s and so the random process $L(\cdot)$ defined in (10) is strict-sense stationary.   ∎

APPENDIX B
ON CONDITIONS FOR WIDE-SENSE STATIONARITY

Proposition. *Let* $\tau(\cdot)$ *be real-valued random process for which the product-moment* $\mathbb{E}[\tau(t)\tau(t')]$ *exists for all* $t, t' \in \mathbb{R}$. *Then the function*

$$g(t, t', f, f') := \mathbb{E}\big[\mathrm{e}^{-j2\pi(f'\tau(t') - f\tau(t))}\big] \qquad (53)$$

depends at most on $\Delta f = f' - f$ *only if* $\tau(\cdot)$ *is a random process with constant realizations.*

Proof: Observe that the expectation in (53) relates directly to the *characteristic function* (or the *moment generating function*) of the bivariate random variable $(X, Y) = \big(\tau(t), \tau(t')\big)$. Consider for simplicity the moment generating function

$$\mathcal{M}_{XY}(f_1, f_2) := \mathbb{E}\big[\mathrm{e}^{f_1 X + f_2 Y}\big], \qquad (54)$$

which relates to the characteristic function $\mathcal{C}_{XY}(\cdot, \cdot)$ by evaluating (54) at (jf_1, jf_2). Moreover, it is readily seen that we directly obtain (53) by evaluating (54) at $(j2\pi f, -j2\pi f')$. Requiring (53) to be a function of Δf essentially means that we require $\mathcal{M}_{XY}(f_1, f_2) = M(f_1 + f_2)$, for some function $M(\cdot)$. Then by using the fact that (54) is a moment generating function it follows for all $n, m \in \mathbb{N}_0$ that

$$\frac{\partial^n}{\partial f_1^n}\frac{\partial^m}{\partial f_2^m}\mathcal{M}_{XY}(f_1, f_2)\Big|_{(0,0)} = \mathbb{E}[X^n Y^m] = M^{(n+m)}(0),$$

where $M^{(k)}(\cdot)$ denotes the k'th order derivative of $M(\cdot)$. Hence, for first- and second-order properties of X and Y we find that $\mathbb{E}[X] = \mathbb{E}[Y]$ and $\mathbb{E}[X^2] = \mathbb{E}[XY] = \mathbb{E}[Y^2]$, which means that the *correlation coefficient* ρ_{XY} between X and Y is such that $|\rho_{XY}| = 1$. Accordingly, X and Y are related via some affine transform $Y = aX + b$ but only $a = 1$ and $b = 0$ together fulfills the first- and second-order requirements for X and Y. This completes the proof since these two random variables were arbitrary samples from the random process $\tau(\cdot)$.   ∎

REFERENCES

[1] G. L. Turin, "Communication Through Noisy, Random-Multipath Channels", Doctoral thesis, Massachusetts Institute of Technology, 1956.

[2] G. L. Turin, F. D. Clapp, T. L. Johnston, S. B. Fine, and D. Lavry, "A statistical model of urban multipath propagation", IEEE Trans. Vehic. Techn., Vol. 21, No. 1, pp. 1-9, 1972.

[3] H. Suzuki, "A statistical model for urban radio propagation", Doctoral thesis, University of California, Berkeley, 1975. See also IEEE Trans. Commun., Vol. 25, No. 7, pp. 673-680, 1977.

[4] H. Hashemi, "Simulation of the urban radio propagation channel", Doctoral thesis, University of California, Berkeley, 1977. See also IEEE Trans. Vehic. Techn., Vol. 28, No. 3, pp. 213-225, 1979.

[5] A. A. M. Saleh and R. A. Valenzuela, "A statistical model for indoor multipath propagation", IEEE J. Selec. Areas in Commun., Vol. 5, No. 2, pp. 128-137, 1987.

[6] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel", IEEE J. Selec. Areas in Commun., Vol. 18, No. 3, pp. 347-360, 2000.

[7] S. J. Papantoniou, "Modelling the Mobile Radio Channel", PhD thesis, ETHZ, Switzerland, 1990.

[8] H. Iwai and Y. Karasawa, "Wideband Propagation Model for the Analysis of the Effect of the Multipath Fading on the Near-Far Problem in CDMA Mobile Radio Systems", IEICE Trans. Commun., Vol. E76-B, No. 2, pp. 103-112, 1993.

[9] B. H. Fleury, U. P. Bernhard and R. Heddergott, "Advanced Radio Channel Model for Magic WAND", ACTS Mobile Telecommun. Summit, pp. 600-607, 1996.

[10] T. Zwick, C. Fischer and W. Wiesbeck, "A Stochastic Multipath Channel Model Including Path Directions for Indoor Environments", IEEE J. Selec. Areas in Commun., Vol. 20, No. 6, pp. 1178-1192, 2002.

[11] D. L. Snyder, "Random Point Processes", John Wiley & Sons, 1975.

[12] D. L. Snyder and M. I. Miller, "Random Point Processes in Time and Space", Springer, 1991.

[13] F. M. Schubert, M. L. Jakobsen, and B. H. Fleury, "Non-Stationary Propagation Model for Scattering Volumes with an Application to the Rural LMS Channel", IEEE Trans. on Anten. and Prop., Vol. 61, No. 5, pp. 2817-2828, 2013.

[14] J. A. Gubner, B. N. Bhaskar and K. Hao, "Multipath-Cluster Channel Models", IEEE Int. Conf. on UWB, pp. 292-296, 2012.

[15] M. L. Jakobsen, T. Pedersen and B. H. Fleury, "Analysis of the Stochastic Channel Model by Saleh & Valenzuela via the Theory of Point Processes", Proc. Int. Zurich Seminar on Commun., Switzerland, pp. 115-118, 2012.

[16] K. Hao, "Modeling and Statistical Analysis of Ultra-Wideband (UWB) Channels and Systems: A Point-Process Approach", PhD thesis, University of Wisconsin-Madison, USA, 2006.

[17] A. Ridolfi, "Power Spectra of Random Spikes and Related Complex Signals", PhD thesis, EPFL, Switzerland, 2004.

[18] H. Hashemi, "The Indoor Radio Propagation Channel", Proc. IEEE, Vol. 81, No. 7, pp. 943-968, 1993.

[19] H. Hashemi, "Impulse Response Modeling of Indoor Radio Propagation Channels", IEEE J. Selec. Areas in Commun., Vol. 11, No. 7, pp. 967-978, 1993.

[20] A. Papoulis, "Probability, Random Variables and Stochastic Processes", McGraw-Hill, 3rd ed., 1991.

[21] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse and M. Franceschetti, "Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks", IEEE J. Selec. Areas in Commun., Vol. 27, No. 7, pp. 1029-1046, 2009.

[22] R. Heddergott, U. P. Bernhard and B. H. Fleury, "Stochastic Radio Channel Model for Advanced Indoor Mobile Communication Systems", IEEE PIMRC, pp. 140-144, 1997.

[23] D. R. Cox and V. Isham, "Point Processes", Chapman & Hall, 1980.

[24] J. F. C. Kingman, "Poisson Processes", Oxford University Press, 1993.

[25] J. Møller and R. P. Waagepetersen, "Statistical Inference and Simulation for Spatial Point Processes", Chapman & Hall/CRC, 2004.

[26] A. J. Baddeley, "Spatial Point Processes and their Applications" (in "Stochastic Geometry - Lecture Notes in Mathematics"), Springer, 2007.

[27] P. Olofsson, "Probability, Statistics, and Stochastic Processes", John Wiley & Sons, 2005.

[28] P. A. Bello, "Characterization of Randomly Time-Variant Linear Channels", IEEE Trans. on Commun. Systems, pp. 360-393, 1963.

[29] Y. (G.) Li, L. J. Cimini, Jr. and N. R. Sollenberger, "Robust Channel Estimation for OFDM Systems with Rapid Dispersive Fading Channels", IEEE Trans. on Commun., Vol 46., No. 7, pp. 902-915, 1998.

[30] R. H. Clarke, "A Statistical Theory of Mobile-Radio Reception", Bell Syst. Tech. J., pp. 957-1000, 1968.

[31] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems", IEE Elect. Lett., Vol. 27, No. 23, pp. 2145-2146, 1991.

[32] A. F. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedal, J. Kunisch, H. G. Schantz, K. Siwiak and M. Z. Win, "A Comprehensive Standardized Model for Ultrawideband Propagation Channels", IEEE Trans. on Anten. and Prop., Vol. 54, No. 11, pp. 3151-3166, 2006.

[33] A. Abdi, J. A. Barger and M. Kaveh, "A Parametric Model for the Distribution of the Angle of Arrival and the Associated Correlation Function and Power Spectrum at the Mobile Station", IEEE Trans. on Vehic. Techn., Vol. 51, No. 3, pp. 425-434, 2002.

[34] M. L. Jakobsen, T. Pedersen and B. H. Fleury, "Simulation of Birth-Death Dynamics in Time-Variant Stochastic Radio Channels", Proc. Int. Zurich Seminar on Commun., Switzerland, pp. 124-127, 2014.

Morten Lomholt Jakobsen received his M.Sc. degree in (applied) mathematics in June 2009 from Department of Mathematical Sciences, Aalborg University, Denmark. In October 2013 he received his Ph.D. degree in wireless communications from Department of Electronic Systems, Aalborg University, Denmark. His research interests include stochastic radio channel modeling, radio channel characterization, signal and information processing for communications, and positioning/localization in wireless networks. He is currently a postdoctoral researcher at Department of Electronic Systems, Aalborg University, Denmark.

Troels Pedersen received the M.Sc.E. degree in digital communications and the Ph.D. degree in wireless communications from Aalborg University, Denmark, in 2004 and 2009, respectively. Since 2009 he has been with Section Navigation and Communications, at Department of Electronic Systems, Aalborg University, first as assistant professor and since 2012 as associate professor. He received the "Teacher of the Year 2011" award by the Study Board for Electronics and IT, Aalborg University. In Spring 2012 he was a visiting professor at IETR, University Rennes 1, France. His research interests lie in the area of statistical signal processing and communication theory, including sensor array signal processing, radio geolocation techniques, radio channel modeling, and radio channel sounding.

Bernard Henri Fleury (M'97-SM'99) received the Diploma in electrical engineering and in mathematics and the Ph.D. degree in electrical engineering from the Swiss Federal Institute of Technology Zurich (ETHZ), Zurich, Switzerland, in 1978, 1990, and 1990, respectively.

Since 1997, he has been with the Department of Electronic Systems, Aalborg University, Aalborg, Denmark, as a Professor of Communication Theory. He is the Head of the Section of Navigation and Communications, which is one of the 11 laboratories of this Department. From 2006 to 2009, he was a Key Researcher with the Telecommunications Research Center Vienna (FTW), Vienna, Austria. During 1978-1985 and 1992-1996, he was a Teaching Assistant and a Senior Research Associate, respectively, with the Communication Technology Laboratory, ETHZ. Between 1988 and 1992, he was a Research Assistant with the Statistical Seminar at ETHZ.

Prof. Fleury's research interests cover numerous aspects within communication theory, signal processing, and machine learning, mainly for wireless communications. His current scientific activities include stochastic modeling and estimation of the radio channel, especially for multiple-input-multiple-output (MIMO) systems operating in harsh conditions, iterative message-passing processing with focus on the design of efficient feasible architectures for wireless receivers, localization techniques in wireless terrestrial systems, and radar signal processing. Prof. Fleury has authored and coauthored more than 120 publications in these areas. He has developed, with his staff, a high-resolution method for the estimation of radio channel parameters that has found a wide application and has inspired similar estimation techniques both in academia and in industry.