The Annals of Applied Probability
2008, Vol. 18, No. 5, 1910–1943
DOI: 10.1214/07-AAP509

LINEAR AND QUADRATIC FUNCTIONALS OF RANDOM HAZARD RATES: AN ASYMPTOTIC ANALYSIS

BY GIOVANNI PECCATI[1] AND IGOR PRÜNSTER[2]

Université Paris VI and Università degli Studi di Torino

A popular Bayesian nonparametric approach to survival analysis consists in modeling hazard rates as kernel mixtures driven by a completely random measure. In this paper we derive asymptotic results for linear and quadratic functionals of such random hazard rates. In particular, we prove central limit theorems for the cumulative hazard function and for the path-second moment and path-variance of the hazard rate. Our techniques are based on recently established criteria for the weak convergence of single and double stochastic integrals with respect to Poisson random measures. The findings are illustrated by considering specific models involving kernels and random measures commonly exploited in practice. Our abstract results are of independent theoretical interest and can be applied to other areas dealing with Lévy moving average processes. The strictly Bayesian analysis is further explored in a companion paper, where our results are extended to accommodate posterior analysis.

1. Introduction. Survival analysis has been the focus of many contributions to Bayesian nonparametric theory and practice. Indeed, many statistical problems arising in the framework of survival analysis require function estimation and, hence, they are ideally suited for a nonparametric treatment. Essentially, two closely related lines of research have been pursued: the first is represented by the introduction of models for the random cumulative distribution function whereas the second deals with models for the random hazard rate and the random cumulative hazard. As for the former, most proposals fall within the class of neutral to the right processes due to Doksum [10]; among others, we mention [13–15, 23, 47, 48]. As for the latter, one can distinguish models leading to a cumulative hazard which is almost surely discrete and models for which it is almost surely absolutely continuous. The famous beta process derived in Hjort [17] belongs to the first class along the contributions of, for example, [6, 24–26]. The second class focuses on

Received October 2007; revised December 2007.

[1]Supported in part by ISI Foundation—Lagrange Project.

[2]Also affiliated with Collegio Carlo Alberto and ICER, Torino, Italy. Supported in part by MiUR Grant 2006/133449.

AMS 2000 subject classifications. 60G57, 62G20.

Key words and phrases. Asymptotics, Bayesian nonparametrics, central limit theorem, completely random measure, multiple Wiener–Itô integral, path-variance, random hazard rate, survival analysis.

the hazard rate which is modeled as a mixture and has recently received much attention due to a relatively simple implementation in applications. After the seminal papers [12, 32] important developments dealing also with more general multiplicative intensity models can be found in [18, 20–22, 33, 34], among others. Passing from a hazard rate function to the corresponding model for the cumulative distribution function is straightforward if the hazard rate is almost surely absolutely continuous, but quite subtle otherwise. See [17] and [23], which establishes a nice link via the notion of spatial neutral to the right process. It is also worth noting that all models share a common feature, namely, that their basic building block is represented by an increasing additive process (see [44]) or more generally by a completely random measure, a notion introduced in [27].

Let us focus attention on hazard rates that are modeled as mixtures. Denote by U a positive absolutely continuous random variable representing the lifetime and assume that its random hazard rate is of the form

$$\tilde{h}(t) = \int_{\mathbb{X}} k(t,x)\tilde{\mu}(dx), \tag{1}$$

where k is a kernel and $\tilde{\mu}$ a completely random measure on some space $\mathbb{X}$. The random measure $\tilde{\mu}$ will be often referred to as the *background driving* random measure, a terminology introduced in [2], which indicates that $\tilde{\mu}$ is acting on latent feature of the model. The cumulative hazard is then given by $\tilde{H}(t) = \int_0^t \tilde{h}(s)\,ds$. Note that, given $\tilde{\mu}$, $\tilde{h}$ represents the hazard rate of U, that is

$$\tilde{h}(t)\,dt = \mathbb{P}(t \le U \le t + dt | U \ge t, \tilde{\mu}).$$

From (1), provided $\tilde{H}(t) \to \infty$ for $t \to \infty$ almost surely, one can define a random density function f as

$$\tilde{f}(t) = \tilde{h}(t)\, \exp(-\tilde{H}(t))$$

where $\tilde{S}(t) := \exp(-\tilde{H}(t))$ is the so-called survival function providing the probability that $U > t$. Consequently the random cumulative distribution function of U is of the form $\tilde{F}(t) = 1 - \exp(-\tilde{H}(t))$. Such models, often referred to as life-testing models, have been considered in [12] and [32] with $\tilde{\mu}$ being an extended gamma process, also known as weighted gamma process. In [33], instead, a weighted version of a gamma compound Poisson process was used. Analysis beyond gamma-like choices of $\tilde{\mu}$ was not possible due to the lack of a suitable and implementable posterior characterization; however, in [22] this goal has been achieved and many choices for $\tilde{\mu}$ can now be explored. See also [18] for a posterior characterization via S-paths.

In this paper, we provide asymptotic results for random hazard rates constructed via a mixture approach as in (1). In particular, for $i = 1, 2, 3$, we will be interested in establishing the existence of two positive functions $\tau_i(T)$ and $\eta_i(T)$ such that the following central limit theorems (CLTs in the sequel) take place as $T \to +\infty$:

$$\eta_1(T) \times [\tilde{H}(T) - \tau_1(T)] \xrightarrow{\text{law}} X_1(\sigma_1), \tag{2}$$

$$\eta_2(T) \times \left[\frac{1}{T}\int_0^T \tilde{h}(t)^2\,dt - \tau_2(T)\right] \xrightarrow{\text{law}} X_2(\sigma_2), \tag{3}$$

$$\eta_3(T) \times \left[\frac{1}{T}\int_0^T [\tilde{h}(t) - \tilde{H}(T)/T]^2\,dt - \tau_3(T)\right] \xrightarrow{\text{law}} X_3(\sigma_3), \tag{4}$$

where, for $i = 1, 2, 3$, $X_i(\sigma_i)$ is a centered Gaussian random variable, with variance σ_i depending on the analytic structures of $\tilde{\mu}$ and k. For a fixed $T > 0$, the random objects $T^{-1}\int_0^T \tilde{h}(t)^2\,dt$ and $T^{-1}\int_0^T [\tilde{h}(t) - \tilde{H}(T)/T]^2\,dt$ are called, respectively, the (realized) *path-second moment* and the (realized) *path-variance* associated with $\tilde{h}$. As we will point out in the subsequent sections, weak convergence results such as (2), (3) and (4) give a description of the overall variability of the hazard rate $\tilde{h}(t)$, by providing a synthetic answer to the following crucial questions: (i) "How fast does the cumulative hazard rate diverge from its long-term trend $\tau_1(T)$?," (ii) "How fast increases the magnitude of the fluctuations of $\tilde{h}(t)$ above zero?," and (iii) "How big are the oscillations of $\tilde{h}(t)$ around its average value?." We stress that our choice of $+\infty$ as a limiting point is mainly conventional, and that one can easily modify our framework to deal with models that live within a finite window of time by using an appropriate deformation of the time scale. For instance, one can embed a hazard rate model defined on $[0, +\infty)$ into a finite time interval, by substituting the time parameter T in the previous discussion with an increasing function of the type $\log[T^*/(T^* - T)]$, where $T^* < +\infty$ and $0 \leq T < T^*$. Note that models defined on a finite time interval are more apt to represent hazard rates associated to the age of a given human population.

To the authors' knowledge, this represents a completely new line of research. Indeed, by now, many results have been obtained in terms of consistency of posterior distributions. See [16] for an exhaustive account. However, little is known about the distributional behavior of the prior ingredients of a Bayesian nonparametric model such as (1), in particular with reference to functionals of statistical relevance. In the more conventional setup of random probability measures, instead of the one concerning hazard rates considered here, the first results on linear functionals of the Dirichlet process were achieved in the pioneering paper of Cifarelli and Regazzini [5], whereas the variance functional is studied in [4] and [41]. One may try to adopt the approach of [42] based on Gurland's inversion formula to derive expressions for the distribution of linear functionals of general random hazards as in (1), but to tackle quadratic functionals seems impossible to date. In light of these considerations, it seems important to remark that, despite the theoretical relevance of our asymptotic results, they also turn out to be helpful in terms of prior specification: on one hand they can serve as a guide for deciding which particular completely random measure $\tilde{\mu}$ to use for defining the model (1) and on the other hand, once $\tilde{\mu}$ is chosen, provide hints for selecting the parameters of $\tilde{\mu}$. Indeed, up to now these two steps were carried out in a conventional way, leaving aside the problem of properly incorporating prior knowledge, in particular with respect to the choice of $\tilde{\mu}$. A first contribution highlighting the different clustering behaviors induced by alternative random measures in the context of mixtures for Bayesian density estimation is provided in [30]. See also [19].

It is important to remark that our results can be extended to the posterior case. More precisely, in [7] analogous CLTs for linear and quadratic functionals of random hazard rates are proved, after conditioning on an arbitrarily large (fixed) number of observations. One of the main findings in [7] is that, for the most widely used models, the CLTs associated with the posterior hazard rate are exactly the same as the prior ones. This is proved by combining the results of the present paper with the aforementioned general posterior characterization of random hazards of James [22]. In particular, James' theory allows to successfully tackle the fact that the models we are considering are *not conjugate*. More details are given in Section 5.

Before providing an outline of the paper, we shall also stress that our results can be of potential interest in other areas where moving averages of Lévy processes appear, such as finance (see, e.g., [2]) or telecommunications modeling (see, e.g., [50]).

The paper is structured as follows. In Section 2 we introduce some basic concepts and notation. In Section 3 we state the main results concerning linear and quadratic functionals of random hazard rates. In particular, we derive CLTs for the cumulative hazard function and for the path-second moment and path-variance of the hazard rate. Moreover, we provide a useful comparison theorem which allows to bypass the verification of the most delicate conditions thus leading to obtain CLTs for hazard rates based on complex kernels or random measures. Section 4 is devoted to applications: we consider specific models involving kernels and random measures commonly exploited in practice and analyze their asymptotic behavior in detail. Section 5 briefly describes the extension to the posterior case developed in [7]. In Section 6 the proofs of our results are provided and further useful techniques described. Section 7 contains some concluding remarks together with possible extensions and an outline of future work.

2. Basic concepts and notation. We start by introducing the main concepts and notation employed throughout the paper. Consider a measure space $(\mathbb{X}, \mathscr{X})$, where $\mathbb{X}$ is a complete and separable metric space and $\mathscr{X}$ is the usual Borel σ-field. Introduce a *Poisson random measure* $\tilde{N}$, defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and taking values in the set of nonnegative counting measures on $(\mathbb{R}^+ \times \mathbb{X}, \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X})$, with nonatomic *intensity measure* ν, that is,

$$\mathbb{E}[\tilde{N}(dv, dx)] = \nu(dv, dx)$$

and, for any $A \in \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X}$ such that $\nu(A) < \infty$, $\tilde{N}(A)$ is a Poisson random variable of parameter $\nu(A)$. Moreover, given any finite collection of pairwise disjoint sets, $A_1, \dots, A_k$, in $\mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X}$, the random variables $\tilde{N}(A_1), \dots, \tilde{N}(A_k)$

are mutually independent. Throughout the paper, $\mathbb{E}[\cdot]$ will denote expectation with respect to $\mathbb{P}$. Moreover, the intensity measure ν must satisfy

$$\int_{\mathbb{R}^+} (v \wedge 1)\nu(dv, \mathbb{X}) < \infty$$

where $a \wedge b = \min\{a, b\}$. See [8] for an exhaustive account on Poisson random measures.

Recall that, according to, for example, [8], a Borel measure μ on some Polish space endowed with the Borel σ-algebra is said to be *boundedly finite* if $\mu(A) < +\infty$ for every bounded measurable set A. Let now $(\mathbb{M}, \mathscr{B}(\mathbb{M}))$ be the space of boundedly finite measures on $(\mathbb{X}, \mathscr{B}(\mathbb{X}))$. We suppose that $\mathbb{M}$ is equipped with the topology of vague convergence and that $\mathscr{B}(\mathbb{M})$ is the corresponding Borel σ-field. Let $\tilde{\mu}$ be a random element, defined on $(\Omega, \mathscr{F}, \mathbb{P})$ and with values in $(\mathbb{M}, \mathscr{B}(\mathbb{M}))$, which can be represented as a linear functional of the Poisson random measure $\tilde{N}$ (with intensity ν) as follows:

$$\tilde{\mu}(B) = \int_{\mathbb{R}^+ \times B} s\tilde{N}(ds, dx) \qquad \forall B \in \mathscr{B}(\mathbb{X}).$$

It can be easily deduced from the properties of $\tilde{N}$ that $\tilde{\mu}$ is, in the terminology of [27], a *completely random measure* (CRM) on $\mathbb{X}$, that is:

(i) $\tilde{\mu}(\varnothing) = 0$ a.s.-$\mathbb{P}$;

(ii) for any collection of disjoint sets in $\mathscr{B}(\mathbb{X})$, denoted as $B_1, B_2, \ldots,$ the random variables $\tilde{\mu}(B_1), \tilde{\mu}(B_2), \ldots$ are mutually independent and $\tilde{\mu}(\bigcup_{j\geq 1} B_j) = \sum_{j\geq 1} \tilde{\mu}(B_j)$ holds true a.s.-$\mathbb{P}$.

Now let $\mathscr{G}_\nu$ be the space of functions $g : \mathbb{X} \to \mathbb{R}^+$ such that $\int_{\mathbb{R}^+ \times \mathbb{X}} [1 - e^{-sg(x)}]\nu(ds, dx) < \infty$. Then, the law of $\tilde{\mu}$ is uniquely characterized by its *Laplace functional* which, for any g in $\mathscr{G}_\nu$, is given by

$$\mathbb{E}[e^{-\int_{\mathbb{X}} g(x)\tilde{\mu}(dx)}] = \exp\left\{-\int_{\mathbb{R}^+ \times \mathbb{X}} [1 - e^{-sg(x)}]\nu(ds, dx)\right\}. \tag{5}$$

For details and further references on CRMs see [28]. From (5) it is apparent that the law of the CRM $\tilde{\mu}$ is completely determined by the corresponding intensity measure ν. This suggests a simple and useful distinction of the random measures we deal with, according to the decomposition of ν. Letting λ be a nonatomic and σ-finite measure on $\mathbb{X}$, we have:

(a) if $\nu(dv, dx) = \rho(dv)\lambda(dx)$, for some measure ρ on $\mathbb{R}^+$, we say that the corresponding $\tilde{N}$ and $\tilde{\mu}$ are *homogeneous*;

(b) if $\nu(dv, dx) = \rho(dv|x)\lambda(dx)$, where $\rho : \mathscr{B}(\mathbb{R}^+) \times \mathbb{X} \to \mathbb{R}^+$ is a kernel [i.e., $x \mapsto \rho(C|x)$ is $\mathscr{B}(\mathbb{X})$-measurable for any $C \in \mathscr{B}(\mathbb{R}^+)$ and $\rho(\cdot|x)$ is a σ-finite measure on $\mathscr{B}(\mathbb{R}^+)$ for any x in $\mathbb{X}$], we say that the corresponding random measures $\tilde{N}$ and $\tilde{\mu}$ are *nonhomogeneous*.

In the sequel we consider CRM $\tilde{\mu}$ whose intensity measures satisfy

$$\int_{\mathbb{R}^+ \times \mathbb{X}} \rho(dv|x)\lambda(dx) = +\infty. \tag{H1}$$

In the homogeneous case, (H1) reduces to $\max\{\rho(\mathbb{R}^+); \lambda(\mathbb{X})\} = +\infty$, which is tantamount to requiring either infinite activity of $\tilde{\mu}$, that is, $\tilde{\mu}$ jumping infinitely often on any bounded $A \in \mathscr{X}$, or to considering $\tilde{\mu}$ with unbounded support S such that $\lambda(S) = +\infty$. In the nonhomogeneous case, for (H1) to hold it is enough that $\tilde{\mu}$ jumps infinitely often on some bounded set of positive λ-measure. It is clear that (H1) is met by the CRM usually considered in the literature. In the subsequent sections, as illustrations of our general results, we will consider the following CRMs:

1. Generalized gamma CRM: its intensity measure is homogeneous and given by

$$\nu(dv, dx) = \frac{1}{\Gamma(1-\sigma)} \frac{e^{-\gamma v}}{v^{1+\sigma}} \, dv \, \lambda(dx) \tag{6}$$

   where $\sigma \in (0, 1)$ and $\gamma > 0$. This class, studied in [3], can be characterized as the tilted exponential family generated by the positive stable laws. It includes the inverse Gaussian CRM for $\sigma = 1/2$ and the gamma CRM as $\sigma \to 0$. The case $\sigma < 0$, corresponding to the so-called *compound gamma CRM*, is discussed in [33].
2. Extended gamma CRM: its nonhomogeneous intensity measure is of the form

$$\nu(dv, dx) = \frac{e^{-\beta(x)v}}{v} \, dv \, \lambda(dx) \tag{7}$$

   where β is a strictly positive function on $\mathbb{X}$. This class dates back to [12, 32]. See also [31]. The gamma CRM arises if β is constant.
3. Beta CRM: its nonhomogeneous intensity measure is given by

$$\nu(dv, dx) = \mathbb{I}_{(0,1)}(v) c(x) \frac{(1-v)^{c(x)-1}}{v} \, dv \, \lambda(dx) \tag{8}$$

   where c is some strictly positive function on $\mathbb{X}$ and $\mathbb{I}_A$ stands for the indicator function of set A. Note that the class of beta CRM, which is due to [17], has the particularity of allowing only jumps of sizes less than 1.

Having settled the basics regarding the background driving CRM in (1), we now have to define the kernel: k is a jointly measurable application from $\mathbb{R}^+ \times \mathbb{X}$ to $\mathbb{R}^+$, such that $\int_{\mathbb{X}} k(t, x)\lambda(dx) < +\infty$ and $\int_{\cdot} k(t, x) \, dt$ is a σ-finite measure on $\mathscr{B}(\mathbb{R}^+)$ for any x in $\mathbb{X}$. Given these two ingredients, the random hazard rate in (1) is properly defined.

A further technical assumption we will make throughout the paper is represented by the following conditions:

$$\int_{\mathbb{R}^+ \times \mathbb{X}} k(t, x)^j v^j \rho(dv|x)\lambda(dx) < +\infty \qquad \forall t, j = 1, 2, 4;$$

(H2) $$\int_{[0,T]}\int_{\mathbb{R}^+\times\mathbb{X}} k(t,x)^j v^j \rho(dv|x)\lambda(dx)\,dt < +\infty \qquad \forall T>0,\ j=1,2,4.$$

If, for $j=1,2,4$, the application $x \mapsto \int_{\mathbb{R}^+} v^j \rho(dv|x)$ is *bounded* by some finite constant (which is typically the case), then the first condition in (H2) reduces to requiring that the function $x \mapsto k(t,x)^j$ is integrable with respect to λ for every t, whereas the second line of (H2) boils down to the assumption that the application $(t,x) \mapsto k(t,x)$ is an element of $\bigcap_{j=1,2,4} L^j([0,T]\times\mathbb{X}, dt\,\lambda(dx))$ for every $T>0$. Hence, in the uniformly bounded case (H2) is a condition not involving the CRM, but just the kernel. Moreover, it is easy to see that the quantity $\int_{\mathbb{R}^+} v^j \rho(dv|x)$, $j=1,2,4$, is bounded in x whenever $\rho(dv|x)$ is associated to one of the three classes of CRMs defined above [see (6), (7) and (8)]. We shall also note that, in the homogeneous case, (H2) implies that $\int_{\mathbb{R}^+} v^j \rho(dv) < +\infty$, $j=1,2,4$. An example of a homogeneous CRM which does not meet (H2) is the stable CRM for which $\rho(dv)=v^{-1-\sigma}\,dv$ and $\sigma\in(0,1)$. Note that the stable CRM can be recovered from the generalized gamma class by allowing $\gamma=0$ in (6); we have excluded this possibility since it does not meet (H2).

REMARK 2.1. We conjecture that the results proved in this paper could be further generalized, in order to accommodate the analysis of the stable case [corresponding to the choice of $\gamma=0$ in (6)]. One possible direction is indicated in [22], page 1785. In such a reference it is indeed shown that every random hazard rate $\tilde{h}$ of the type (1) can be coupled with an ancillary random hazard h^*, verifying the relation: for every t,

$$\mathscr{L}(h^*(t)) = \mathscr{L}(\tilde{h}(t)|U\geq t),$$

where $\mathscr{L}$ indicates the law of a given random variable. A remarkable feature of the random hazard rate h^* is that it displays the same kernel k as $\tilde{h}$, integrated with respect to a CRM with *time-dependent* intensity, obtained via an "exponential correction" of the intensity associated with $\tilde{h}$. In the particular case of a stable background driving measure, the time-dependent intensity associated with $h^*(t)$ takes the form

$$\nu_t(dv,dx) = C\times \exp(-vk_t^*(x))v^{-\sigma-1}\,dv\,\lambda(dx),$$

where C is a suitable constant, and

$$k_t^*(x) = \int_0^t k(s,x)\,ds.$$

Observe that, for fixed t, this corresponds to a gamma-type intensity of the kind considered in (6), where $\gamma = k_t^*(x) > 0$. The idea would be therefore to study linear and quadratic functionals of $\tilde{h}$, by first resorting to those of the associated rate h^*. An additional difficulty would be of course the presence of time-dependency into the background driving CRM. This topic will be the object of a separate research.

2.1. *Further notation.* For $q, p \geq 1$, we note

$$L^p(\nu^q) = L^p((\mathbb{R}^+ \times \mathbb{X})^q, (\mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X})^q, \nu^q)$$

the Banach space of real-valued functions f on $(\mathbb{R}^+ \times \mathbb{X})^q$, such that $|f|^p$ is integrable with respect to $\nu^q := \nu^{\otimes q}$. We will systematically write $L^p(\nu^1) = L^p(\nu)$ for $p \geq 1$. The symbol $L_s^2(\nu^2)$ is used to denote the subspace of $L^2(\nu^2)$ composed of *symmetric functions* on $(\mathbb{R}^+ \times \mathbb{X})^2$. By symmetric, we mean that every $f \in L_s^2(\nu^2)$ is such that $f(s, x; t, y) = f(t, y; s, x)$ for every $(s, x), (t, y) \in \mathbb{R}^+ \times \mathbb{X}$. As an example of function in $L^2(\nu^2)$, one can take

$$f(s, x; t, y) = \mathbb{I}_A(s, x)\mathbb{I}_B(t, y),$$

where $A, B \in \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X}$ are such that $\nu(A) < \infty$ and $\nu(B) < \infty$, whereas a symmetric function in $L_s^2(\nu^2)$ is, for instance,

$$f(s, x; t, y) = \mathbb{I}_A(s, x)\mathbb{I}_B(t, y) + \mathbb{I}_A(t, y)\mathbb{I}_B(s, x),$$

where the sets A and B are as above. We also write $L_{s,0}^2(\nu^2)$ to indicate the subset of $L_s^2(\nu^2)$ composed of symmetric functions *vanishing on diagonals*, that is, such that their support is contained in the *purely nondiagonal set* $D_0^2 = \{(s, x; t, y) : (s, x) \neq (t, y)\}$.

We now turn to the definition of three basic auxiliary kernels which are associated to a given $f \in L_s(\nu^2)$:

(i) the kernel $f \star_1^0 f$ is defined on $(\mathbb{R}^+ \times \mathbb{X})^3$ and is given by

$$f \star_1^0 f(t_1, x_1; t_2, x_2; t_3, x_3) = f(t_1, x_1; t_2, x_2) f(t_2, x_2; t_3, x_3); \tag{9}$$

(ii) $f \star_1^1 f$ is defined on $(\mathbb{R}^+ \times \mathbb{X})^2$ and is actually a *contraction* equal to

$$f \star_1^1 f(t_1, x_1; t_2, x_2) = \int_{\mathbb{R}^+ \times \mathbb{X}} f(t_1, x_1; s, y) f(s, y; t_2, x_2) \nu(ds, dy); \tag{10}$$

(iii) $f \star_2^1 f$ is defined on $(\mathbb{R}^+ \times \mathbb{X})$ and is given by

$$f \star_2^1 f(t, x) = \int_{\mathbb{R}^+ \times \mathbb{X}} f(t, x; s, y)^2 \nu(ds, dy). \tag{11}$$

Note that, by the Cauchy–Schwarz inequality and by the symmetry and square-integrability of f, the kernel $f \star_1^1 f$ is necessarily an element of $L_s^2(\nu^2)$. The three kernels defined above are the fundamental building blocks to obtain explicit expressions for the moments and the cumulants of the linear and quadratic functionals associated with random hazard rates (when they exist). Such expressions enter implicitly in the statements of the subsequent results, and are mainly of a combinatorial nature. We refer the reader to [43] for an exhaustive analysis of the combinatorial machinery underlying the construction of stochastic integrals with respect to completely random measures.

In the subsequent sections it will be often convenient to work with the *compensated Poisson random measure* canonically associated to $\tilde{N}$. Such an object is indicated by

$$\tilde{N}^c = \{\tilde{N}^c(A) : A \in \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X}\}, \tag{12}$$

and is formally defined as the unique CRM on $(\mathbb{R}^+ \times \mathbb{X}, \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X})$ such that

$$\tilde{N}^c(A) = \tilde{N}(A) - \nu(A) \tag{13}$$

for every set A of finite ν-measure. For every $g \in L^2(\nu)$, we denote by

$$\tilde{N}^c(g) = \int_{\mathbb{R}^+ \times \mathbb{X}} g(s,x) \tilde{N}^c(ds, dx)$$

the Wiener–Itô integral of g with respect to $\tilde{N}^c$. We recall that, for every $g \in L^2(\nu)$, $\tilde{N}^c(g)$ is a centered and square-integrable random variable with an infinitely divisible law, such that, for every $\lambda \in \mathbb{R}$,

$$\mathbb{E}[e^{i\lambda \tilde{N}^c(g)}] = \exp\left\{\int_{\mathbb{R}^+ \times \mathbb{X}} [e^{i\lambda g(s,x)} - 1 - i\lambda g(s,x)] \nu(ds, dx)\right\} \tag{14}$$

[compare with (5)]. Moreover, for every $f, g \in L^2(\nu)$, one has the *isometric property*

$$\mathbb{E}[\tilde{N}^c(f)\tilde{N}^c(g)] = \int_{\mathbb{R}^+ \times \mathbb{X}} f(s,x) g(s,x) \nu(ds, dx) := (f, g)_{L^2(\nu)}. \tag{15}$$

Note that (5), (14) and the isometric property (15) imply that, for every $g \in L^2(\nu) \cap L^1(\nu)$,

$$\mathbb{E}[\tilde{N}(g)] = \int_{\mathbb{R}^+ \times \mathbb{X}} g(s,x) \nu(ds, dx), \tag{16}$$

$$\mathrm{Var}[\tilde{N}(g)] = \mathrm{Var}[\tilde{N}^c(g)] = \int_{\mathbb{R}^+ \times \mathbb{X}} g(s,x)^2 \nu(ds, dx). \tag{17}$$

3. Main results: CLTs for linear and quadratic functionals. In what follows, we shall develop several techniques, allowing to study the asymptotic behavior of linear and quadratic functionals associated to the random hazard rate $\tilde{h}(t)$ appearing in (1). Concerning quadratic functionals, we will be mainly interested in the path-variance and the path-second moment of $\tilde{h}(t)$. As will be clarified in Section 6, our approach exploits the fact that any quadratic functional of $\tilde{h}$ can be (uniquely) represented as a linear combination of its expectation and of the following two random elements: (i) the stochastic integral of a deterministic kernel with respect to $\tilde{N}^c$, and (ii) the double Wiener–Itô integral of a deterministic bivariate kernel with respect to the stochastic product measure associated to $\tilde{N}^c$. According to the results proved in [38] (see Section 6.1), the joint (weak) convergence of single and double Poisson integrals can be characterized in terms of the asymptotic negligibility of deterministic contraction kernels. We will show that such contractions are indeed explicit functionals of the kernel k defining $\tilde{h}$. We shall first state the main general results of the paper, and then describe in detail several applications. The proofs are deferred to Section 6.

Consider the random hazard rate $\tilde{h}$ defined in formula (1), and assume that the intensity of the underlying Poisson CRM $\tilde{N}$ verifies (H1), and that the positive kernel k satisfies (H2). Moreover, for every $T > 0$ define the kernel

$$k_T^{(0)}(s,x) = s\int_0^T k(t,x)\,dt, \qquad (s,x) \in \mathbb{R}^+ \times \mathbb{X}. \tag{18}$$

Our first result concerns the asymptotic behavior of the cumulative hazard rate $\tilde{H}(T) = \int_0^T \tilde{h}(t)\,dt$.

THEOREM 1. *Suppose that*: (i) $k_T^{(0)} \in L^3(\nu)$ *for every* T, *and* (ii) *there exists a strictly positive function* $T \mapsto C_0(k,T)$, *such that, as* $T \to +\infty$,

$$C_0^2(k,T) \times \int_{\mathbb{R}^+\times\mathbb{X}} [k_T^{(0)}(s,x)]^2 \nu(ds,dx) \to \sigma_0^2(k) > 0, \tag{19}$$

$$C_0^3(k,T) \times \int_{\mathbb{R}^+\times\mathbb{X}} [k_T^{(0)}(s,x)]^3 \nu(ds,dx) \to 0. \tag{20}$$

Then,

$$C_0(k,T) \times [\tilde{H}(T) - \mathbb{E}[\tilde{H}(T)]] \xrightarrow{\text{law}} X, \tag{21}$$

where $X \sim \mathscr{N}(0, \sigma_0^2(k))$.

Note that conditions (19)–(20) only involve the analytic form of the kernel k, and do not make any use of the asymptotic properties of the law of the process $\tilde{h}(t)$, such as, for example, mixing. We now focus on the limiting behavior of the quadratic functionals associated to the random hazard rate $\tilde{h}$. To this end, we associate to $k(\cdot,\cdot)$, and to each $T > 0$, the three auxiliary kernels:

$$k_T^{(1)}(s,x;t,y) = \frac{st}{T}\int_0^T k(u,x)k(u,y)\,du, \tag{22}$$

$$k_T^{(2)}(s,x) = \frac{s^2}{T}\int_0^T k(u,x)^2\,du, \tag{23}$$

$$k_T^{(3)}(s,x) = \int_{\mathbb{R}^+\times\mathbb{X}} k_T^{(1)}(s,x;u,w)\nu(du,dw). \tag{24}$$

The kernel $k_T^{(2)}$ can be obtained by restricting $k_T^{(1)}$ to the *diagonal set* $\{(s,x;t,y) : (s,x) = (t,y)\}$. We will see in Section 6 that the kernels $k_T^{(\cdot)}$ are intimately related to the objects defined in formulas (9)–(11). Note that, due to assumption (H2) and the Jensen and Cauchy–Schwarz inequalities, $k_T^{(1)} \in L_s^2(\nu^2) \cap L^4(\nu^2)$, and also

$k_T^{(2)} \in L^2(\nu)$. The following theorem provides a CLT for the path-second moment of random hazard rates.

THEOREM 2. *Suppose that* $k_T^{(3)} \in L^2(\nu) \cap L^1(\nu)$, $k_T^{(2)} \in L^3(\nu)$ *and that there exists a strictly positive function* $C_1(k, T)$ *such that the following asymptotic conditions are satisfied as* $T \to +\infty$:

1. $2C_1^2(k,T)\|k_T^{(1)}\|^2_{L^2(\nu^2)} \to \sigma_1^2(k) > 0$;
2. $C_1^4(k,T)\|k_T^{(1)}\|^4_{L^4(\nu^2)} \to 0$;
3. $C_1^4(k,T)\|k_T^{(1)} \star_1^1 k_T^{(1)}\|^2_{L^2(\nu^2)} \to 0$;
4. $C_1^4(k,T)\|k_T^{(1)} \star_2^1 k_T^{(1)}\|^2_{L^2(\nu)} \to 0$;
5. $C_1^2(k,T)\|k_T^{(2)} + 2k_T^{(3)}\|^2_{L^2(\nu)} \to \sigma_2^2(k) > 0$;
6. $C_1^3(k,T)\|k_T^{(2)} + 2k_T^{(3)}\|^3_{L^3(\nu)} \to 0$.

Then,

$$C_1(k,T) \times \left\{\frac{1}{T}\int_0^T \tilde{h}(t)^2\,dt - \frac{1}{T}\int_0^T \mathbb{E}[\tilde{h}(t)^2]\,dt\right\} \xrightarrow{\text{law}} X, \tag{25}$$

where $X \sim \mathscr{N}(0, \sigma_1^2(k) + \sigma_2^2(k))$.

Note that

$$\begin{aligned}\|k_T^{(3)}\|_{L^1(\nu)} &= \int_{\mathbb{R}^+\times\mathbb{X}}\int_{\mathbb{R}^+\times\mathbb{X}} k_T^{(1)}(s,x;u,w)\nu(du,dw)\nu(ds,dx)\\ &= \frac{1}{T}\int_0^T\left(\int_{\mathbb{R}^+\times\mathbb{X}} sk(t,x)\nu(ds,dx)\right)^2 dt.\end{aligned}$$

Also, by applying formulas (16) and (17) (for every $t > 0$) in the case $h(s,x) = sk(t,x)$, one obtains that

$$\begin{aligned}\frac{1}{T}\int_0^T \mathbb{E}[\tilde{h}(t)^2]\,dt &= \frac{1}{T}\int_0^T\left(\int_{\mathbb{R}^+\times\mathbb{X}} sk(t,x)\nu(ds,dx)\right)^2 dt\\ &\quad + \frac{1}{T}\int_0^T\int_{\mathbb{R}^+\times\mathbb{X}} s^2k(t,x)^2\nu(ds,dx)\,dt.\end{aligned} \tag{26}$$

The next theorem combines Theorems 1 and 2 to deal with path-variances of random hazard rates.

THEOREM 3. *Suppose that* $\tilde{h}$ *is such that assumptions* (19)–(20) *are verified, and conditions* 1–6 *of Theorem* 2 *are satisfied. If there exists a constant* $\delta(k) \geq 0$ *such that, as* $T \to +\infty$:

1. $C_1(k,T)/(TC_0(k,T))^2 \to 0$;

2. $2C_1(k,T)\mathbb{E}[\tilde{H}(T)]/(T^2C_0(k,T)) \to \delta(k)$;
3. $\|C_1(k,T)(k_T^{(2)}+2k_T^{(3)}) - \delta(k)C_0(k,T)k_T^{(0)}\|^2_{L^2(\nu)} \to \sigma_3^2(k) \geq 0$,

then,

$$C_1(k,T)\times\left\{\frac{1}{T}\int_0^T\left[\tilde{h}(t)-\frac{\tilde{H}(T)}{T}\right]^2dt - \frac{1}{T}\int_0^T\mathbb{E}[\tilde{h}(t)^2]\,dt + \frac{\mathbb{E}[\tilde{H}(T)]^2}{T^2}\right\}$$

$$\begin{aligned}(27)\qquad &= C_1(k,T)\times\left\{\frac{1}{T}\int_0^T\left[\tilde{h}(t)-\frac{\tilde{H}(T)}{T}\right]^2dt\right.\\ &\qquad\left. - \frac{1}{T}\int_0^T\mathbb{E}\left[\tilde{h}(t)-\frac{\mathbb{E}(\tilde{H}(T))}{T}\right]^2dt\right\} \xrightarrow{\text{law}} X,\end{aligned}$$

where $X \sim \mathscr{N}(0, \sigma_1^2(k)+\sigma_3^2(k))$.

In view of (17), one also has that

$$\frac{1}{T}\int_0^T \operatorname{Var}(\tilde{h}(t))\,dt = \frac{1}{T}\int_0^T\int_{\mathbb{R}^+\times\mathbb{X}} s^2k(t,x)^2\nu(ds,dx)\,dt.$$

To conclude this subsection, we state a useful *comparison theorem* for random hazard rates. To this end, consider two completely random Poisson measures (on $\mathbb{R}^+\times\mathbb{X}$) $\overline{N}$ and $\overline{\overline{N}}$, as well as positive kernels $\overline{k}$ and $\overline{\overline{k}}$. The σ-finite and nonatomic intensity measures of $\overline{N}$ and $\overline{\overline{N}}$ are denoted by $\overline{\nu}$ and $\overline{\overline{\nu}}$, respectively. We assume that $\overline{\nu}$ and $\overline{\overline{\nu}}$ both verify (H1), and that $\overline{k}$ and $\overline{\overline{k}}$ satisfy (H2). Finally, we suppose that, for every $B\in(\mathscr{B}(\mathbb{R}^+)\otimes\mathscr{X})$,

$$\overline{\nu}(B)\leq\overline{\overline{\nu}}(B),$$

and, for every $(t,x)\in\mathbb{R}^+\times\mathbb{X}$,

$$\overline{k}(t,x)\leq\overline{\overline{k}}(t,x).$$

Throughout the paper, for strictly positive sequences $\{a_n\}$ and $\{b_n\}$, we write $a_n \sim b_n$ if there exists $c\in(0,+\infty)$ such that $a_n/b_n\to c$, as $n\to\infty$.

THEOREM 4. *Suppose that the pair* (ν,k) *entering the definition of the random hazard* $\tilde{h}$ *in* (1) *is such that, for every* $B\in(\mathscr{B}(\mathbb{R}^+)\otimes\mathscr{X})$, $\overline{\nu}(B)\leq\nu(B)\leq\overline{\overline{\nu}}(B)$ *and, for every* $(t,x)\in\mathbb{R}^+\times\mathbb{X}$, $\overline{k}(t,x)\leq k(t,x)\leq\overline{\overline{k}}(t,x)$. *Then, the following three comparison criteria hold.*

(A) *Assume that the two kernels* $\overline{k}$ *and* $\overline{\overline{k}}$, *with* $\overline{\nu}$ *and* $\overline{\overline{\nu}}$ *substituting* ν, *satisfy the conditions* (19)–(20) *for some appropriate positive functions* $C_0(\overline{k},T)$ *and* $C_0(\overline{\overline{k}},T)$ *and constants* $\sigma_0^2(\overline{k})$ *and* $\sigma_0^2(\overline{\overline{k}})$. *Suppose also that* $C_0(\overline{k},T)\sim C_0(\overline{\overline{k}},T)$, *and consider a positive function* $C_0(k,T)$ *such that* $C_0(k,T)\sim C_0(\overline{k},T)$. *Then, for every diverging sequence* $T_n\to+\infty$, *there exists a subsequence* $T_{n'}$ *such that* *the CLT* (21) *holds as $n' \to +\infty$, with $T_{n'}$ substituting T, where X is a centered Gaussian random variable whose variance depends on the choice of $C_0(k,T)$ and on n'.*

(B) *Assume that $\overline{k}$ and $\overline{\overline{k}}$, with $\overline{\nu}$ and $\overline{\overline{\nu}}$ substituting ν, satisfy conditions* 1–6 *of Theorem* 2 *for some positive functions $C_1(\overline{k},T)$ and $C_1(\overline{\overline{k}},T)$ and constants $\sigma_j^2(\overline{k})$ and $\sigma_j^2(\overline{\overline{k}})$, $j=1,2$. Assume, moreover, that $C_1(\overline{k},T) \sim C_1(\overline{\overline{k}},T)$, and select a positive function $C_1(k,T)$ such that $C_1(k,T) \sim C_1(\overline{k},T)$. Then, for every sequence $T_n \to +\infty$, there exists a subsequence $T_{n'}$ such that the CLT* (25) *is verified* (*for $n' \to +\infty$ and with $T_{n'}$ substituting T*), *where X is a centered Gaussian random variable whose variance depends on $C_1(k,T)$ and n'.*

(C) *Suppose that $\overline{k}$, $\overline{\overline{k}}$, $C_j(\overline{k},T)$, $C_j(\overline{\overline{k}},T)$ and $C_j(k,T)$ $(j=0,1)$ satisfy the assumptions pinpointed in parts* (A) *and* (B), *and suppose that they also meet the conditions* 1–3 *of Theorem* 3. *Then, for every sequence $T_n \to +\infty$, there exists a subsequence $T_{n'}$ such that the CLT* (27) *holds, for $n' \to +\infty$ and with $T_{n'}$ substituting T.*

REMARK 3.1. The conclusions of Theorem 4 are less precise than those of Theorems 1–3, in the sense that they only apply to subsequences $T_{n'}$. Of course, this is due to the fact that, in the statement of Theorem 4, we do not make *any* assumption on the analytic properties of k and ν, besides the conditions $\overline{k} \le k \le \overline{\overline{k}}$ and $\overline{\nu} \le \nu \le \overline{\overline{\nu}}$. As will become clear in the subsequent sections, more exact information can be deduced by adding some specific requirements to the structure of k and ν.

4. Applications. We will now consider noteworthy examples of random hazard rates by specifying suitable kernels and the form of the background driving CRM. In the following we will always consider CRMs with λ being the Lebesgue measure on $\mathbb{R}^+$, which appears a natural choice in our context. This implies that assumption (H1) is met. Section 4.1 is devoted to the study of the asymptotic behavior of the cumulative hazard $\tilde{H}$, whereas in Section 4.2 we deal with quadratical functionals of the hazard rate.

4.1. *Asymptotics for the cumulative hazard.* As an illustration of Theorem 1, we consider different kernels and show how they are responsible for the rate of divergence of the cumulative hazard and how they influence the variance of the limiting Gaussian random variable in the CLT (21). We first consider general homogeneous CRM such that $\int_{[1,\infty)} v^4 \rho(dv) < \infty$, which is tantamount to requiring the part of condition (H2) involving the jump component of the Poisson intensity to be satisfied. Moreover, set, for notational convenience, $K_\rho^{(i)} = \int_0^\infty s^2 \rho(ds)$, $i=1,2$, and $I_i = I_i(T) = \int_{\mathbb{R}^+ \times \mathbb{X}} [k_T^{(0)}(s,x)]^i \nu(ds,dx)$ for $i=1,2,3$. Note that $I_1(T) = \mathbb{E}[\tilde{H}(T)]$.

(i) *Rectangular kernel.* The kernel $k(t,x) = \mathbb{I}_{(|t-x|\le\tau)}$ where $\tau > 0$ represents a bandwidth, is known as a uniform rectangular kernel. Such a kernel represents a sensible choice when no prior information on the shape of the hazard rate is available. See, for example, [20]. In this setup (H2) is clearly met,

$$k_T^{(0)}(s,x) = \begin{cases} s(x+\tau), & 0 < x < \tau, \\ s2\tau, & \tau \le x < T-\tau, \\ s[T+\tau-x], & T-\tau \le x < T+\tau, \\ 0, & \text{elsewhere}, \end{cases}$$

and $k_T^{(0)}(s,x) \in L^3(\nu)$ for all $T > 0$. We also have, as $T \to \infty$, $I_1(T) = K_\rho^{(1)}\{2T\tau - \frac{1}{2}\tau^2\} = 2\tau K_\rho^{(1)} T + o(T^{1/2})$, $I_2(T) \sim 4K_\rho^{(2)}\tau^2 T$ and $I_3(T) \sim cT$ for some $c > 0$. Hence, (19) and (20) are satisfied with $C_0(k,T) = T^{-1/2}$ and, by Theorem 1, we obtain

$$\frac{1}{\sqrt{T}}[\tilde{H}(T) - 2\tau K_\rho^{(1)} T] \xrightarrow{\text{law}} X, \tag{28}$$

where $X \sim \mathcal{N}(0, 4K_\rho^{(2)}\tau^2)$.

(ii) *Dykstra–Laud kernel.* If $k(t,x) = \mathbb{I}_{(0\le x\le t)}$, then the random hazard rate is monotone increasing. Such a kernel, which is widely exploited in practice, was first proposed in [12]. It is easy to see that (H2) is satisfied and that $k_T^{(0)}(s,x) = s(T-x)\mathbb{I}_{(0\le x\le T)} \in L^3(\nu)$ for all $T > 0$. Moreover, one obtains $I_1 = \frac{K_\rho^{(1)}}{2}T^2$, $I_2 = \frac{K_\rho^{(2)}}{3}T^3$ and $I_3 = \frac{K_\rho^{(2)}}{4}T^4$, so that (19) and (20) are met with $C_0(k,T) = T^{-3/2}$. Hence, by Theorem 1, we have

$$\frac{1}{T^{3/2}}\left[\tilde{H}(T) - \frac{K_\rho^{(1)}}{2}T^2\right] \xrightarrow{\text{law}} X, \tag{29}$$

where $X \sim \mathcal{N}(0, \frac{K_\rho^{(2)}}{3})$. Note that the Dykstra–Laud cumulative hazard has a quadratic asymptotic trend, whereas the trend obtained from a rectangular kernel is linear. Moreover, the speed at which the Dykstra–Laud cumulative hazard diverges from its trend is significantly faster than in the rectangular case. The reason may be that the former produces monotone increasing hazard rates whereas the latter does not. This phenomenon, well exemplified by our result, should be taken into account when deciding which kernel to adopt.

(iii) *Ornstein–Uhlenbeck kernel.* If $k(t,x) = \sqrt{2\kappa}\exp(-\kappa(t-x))\mathbb{I}_{(0\le x\le t)}$, then the random hazard rate is an Ornstein–Uhlenbeck-type process. Such models for the hazard rate are employed in [33, 34]. In this case, (H2) is met, $k_T^{(0)}(s,x) = s\sqrt{2/\kappa}(1 - e^{-\kappa(T-x)})\mathbb{I}_{(0\le x\le T)} \in L^3(\nu)$ for all $T > 0$, and we have that, as T diverges to infinity, $I_1(T) = K_\rho^{(1)}\sqrt{2/\kappa}\{T + e^{-T}/\kappa - \kappa^{-1}\} = K_\rho^{(1)}\sqrt{2/\kappa}T + o(T^{1/2})$, $I_2(T) \sim \frac{2K_\rho^{(2)}}{\kappa}T$ and $I_3(T) \sim cT$ for some constant $c > 0$. Hence,

(19) and (20) are satisfied with $C_0(k,T) = T^{-1/2}$. From Theorem 1 it follows that

$$\frac{1}{\sqrt{T}}\left[\tilde{H}(T) - K_\rho^{(1)}\sqrt{\frac{2}{\kappa}}T\right] \xrightarrow{\text{law}} X, \tag{30}$$

where $X \sim \mathcal{N}(0, \frac{2K_\rho^{(2)}}{\kappa})$. One may note that the trend and the rate of divergence from the trend associated with the Ornstein–Uhlenbeck kernel coincide with those arising from the rectangular kernel. Moreover, given the same background driving CRM, the variances of the limiting Gaussian random variables appearing in (28) and (30) coincide if the parameters are chosen in such a way that $\kappa = 1/(2\tau^2)$.

(iv) *U-shaped or bathtub kernel.* If $k(t,x) = \mathbb{I}_{(|t-\beta| \geq x)}$ with $\beta > 0$, then the corresponding hazard rates are U-shaped with minimum at β. Such a kernel is suggested by [32]. See also [20, 21]. It is easy to check that (H2) is met,

$$k_T^{(0)}(s,x) = \begin{cases} s(T-2x), & 0 < x < \beta, \\ s[T-(\beta+x)], & \beta \leq x < T-\beta, \\ 0, & \text{elsewhere}, \end{cases}$$

and $k_T^{(0)}(s,x) \in L^3(\nu)$ for all $T > 0$. Moreover, as $T \to +\infty$, $I_1(T) = \frac{1}{2}K_\rho^{(1)}T^2 + o(T^{3/2})$, $I_2 \sim \frac{K_\rho^{(2)}}{3}T^3$ and $I_3 \sim cT^4$ for some constant $c > 0$. Choosing $C_0(k,T) = T^{-3/2}$, (19) and (20) are satisfied and from Theorem 1 we deduce

$$\frac{1}{T^{3/2}}\left[\tilde{H}(T) - \frac{1}{2}K_\rho^{(1)}T^2\right] \xrightarrow{\text{law}} X, \tag{31}$$

where $X \sim \mathcal{N}(0, \frac{K_\rho^{(2)}}{3})$. Note that the bathtub kernel produces the same asymptotic behavior as the Dykstra–Laud kernel; this fact is not surprising since after reaching its minimum in β, also the bathtub kernel is monotone increasing. Of course, one can regard the Dykstra and Laud kernel as a degenerate bathtub kernel, corresponding to the case $\beta = 0$.

As apparent from the statement of Theorem 1 and from the discussion provided above, the variances of the limiting Gaussian random variables appearing in (21), (28), (29), (30) and (31) always depend on the jump part of the Poisson intensity. For instance, if $\tilde{\mu}$ is the generalized gamma CRM with intensity (6), then $K_\rho^{(2)} = \frac{(1-\sigma)}{\gamma^{2-\sigma}}$. This confirms the empirical finding, used in tuning the prior parameters, that a small γ induces a large variance. To avoid confusion, note that in the setting of, for example, [20] $\beta = 1/\gamma$ and, hence, their claim that a large β induces a noninformative prior is coherent with our result. As for σ, the variance is maximal in $\sigma = 0$ if $\gamma \leq e$, whereas it is maximized in $\sigma = (\log(\gamma) - 1)/\log(\gamma)$ if $\gamma \geq e$.

Let us now turn attention to hazards based on nonhomogeneous CRM, specifically the extended gamma and beta CRMs presented in Section 2. From (7) and (8) one can see that their nonhomogeneity is due to the strictly positive functions β and c, respectively. According to their structure we distinguish three cases:

Case a. If $\beta(x) = \bar{\beta}$ in (7) and $c(x) = \bar{c}$ in (8), the CRMs become homogeneous and the previous results hold with $K_\rho^{(2)}$ equal to $1/\bar{\beta}^2$ and $1/(1+\bar{c})$, respectively.

Case b. If β (or c) are bounded by some finite constant M, then one can apply Theorem 4 to conclude that $C_0(k, T)$ has the same order as in the examples above, thus depending on the choice of the kernel. Moreover, if β (or c) are eventually nondecreasing (nonincreasing), the convergence holds for any diverging sequence T_n with the variance of the limiting Gaussian random variable depending on the choice of β (or c) taking value in the range $[\sigma_0^2(\bar{k}), \sigma_0^2(\bar{\bar{k}})]$.

Case c. If β (or c) diverge to $+\infty$ as $x \to +\infty$, quite interesting phenomena appear, which shed some light on the possible use of the factor of nonhomogeneity represented by the functions β (or c). Set, for $i = 1, 2, 3$, $K_\rho^{(i)}(x) = \int_0^\infty s^i \rho(ds|x)$, so that I_i becomes $\int_{\mathbb{X}} K_\rho^{(i)}(x)[\int_0^T k(t,x)\,dt]^i\,dx$. For both CRMs, a diverging β (or c) implies that $K_\rho^{(2)}(x) \to 0$; this, indeed, affects the asymptotic behavior of the cumulative hazard $\tilde{H}$. We shall now present some more specific estimates.

(c1) Consider the Dykstra and Laud kernel combined with an extended gamma CRM such that $\beta(x) \sim \sqrt{x}$ as $x \to \infty$: it follows that $I_2 \sim \log(T)T^2$ and $I_3 \sim dT^3$ for some constant $d > 0$. Hence, (19) and (20) are satisfied with $C_0(k, T) = (\sqrt{\log(T)}T)^{-1}$ and, by Theorem 1, we have

$$\frac{1}{\sqrt{\log(T)}T}[\tilde{H}(T) - \mathbb{E}[\tilde{H}(T)]] \xrightarrow{\text{law}} X, \tag{32}$$

where $X \sim \mathscr{N}(0, 1)$. Comparing (32) with (29) one notes that the rate of divergence from the trend $\mathbb{E}[\tilde{H}(T)]$ is reduced from $T^{3/2}$ to $\sqrt{\log(T)}T$. As for $\mathbb{E}[\tilde{H}(T)]$, it is important to remark that the overall growth (though not the dominating term which is $4/3T^{3/2}$) depends on the particular form of β. Still assuming $\beta(x) \sim \sqrt{x}$ and letting b be a positive constant, we obtain, for instance, $\mathbb{E}[\tilde{H}(T)] = 4/3T^{3/2} + o(T\sqrt{\log(T)})$ when $\beta(x) = \mathbb{I}_{(0,b]}(x) + x^{1/2}\mathbb{I}_{(b,\infty)}(x)$, and $\mathbb{E}[\tilde{H}(T)] = 4/3T^{3/2} - \log(T)T + o(T\sqrt{\log(T)})$ if $\beta(x) = (1 + x^{1/2})$. Again, comparing these findings with (29) it is apparent that the trend has been reduced from T^2 to $T^{3/2} + o(T^{3/2})$.

(c2) On the other hand, with the beta CRM, we have $K_\rho^{(1)}(x) = 1$ and, consequently, $I_1(T) = \mathbb{E}[\tilde{H}(T)] = 1/2T^2$ whatever the choice of c. Selecting $c(x) \sim \sqrt{x}$ as $x \to \infty$, we obtain $I_2 \sim 16/15T^{5/2}$ and $I_3 \sim d\log(T)T^3$ for some constant $d > 0$. Thus, with $C_0(k, T) = T^{-5/4}$, (19) and (20) are met and Theorem 1 yields

$$\frac{1}{T^{5/4}}\left[\tilde{H}(T) - \frac{1}{2}T^2\right] \xrightarrow{\text{law}} X,$$

where $X \sim \mathscr{N}(0, 16/15)$. Hence, compared with the homogeneous case in (29), the beta CRM does not affect the trend but still decreases the rate of divergence from T^2 to $T^{5/4}$.

(c3) If, instead, we consider the rectangular kernel with $\tau = 1$ combined with an extended gamma CRM such that again $\beta(x) \sim \sqrt{x}$ as $x \to \infty$, it follows that $I_2 \sim$

$4\log(T)$ and $I_3 \to d$ for some constant $d > 0$. Hence, (19) and (20) are satisfied with $C_0(k,T) = (\sqrt{\log(T)})^{-1}$ and, by Theorem 1, we have

$$\frac{1}{\sqrt{\log(T)}}[\tilde{H}(T) - \mathbb{E}[\tilde{H}(T)]] \xrightarrow{\text{law}} X,$$

where $X \sim \mathscr{N}(0,4)$. Hence, we see that the rate of divergence from $\mathbb{E}[\tilde{H}(T)]$ has been reduced with respect to the homogeneous case in (28) decreasing from $T^{1/2}$ to $\sqrt{\log(T)}$. As before, $I_1(T) = \mathbb{E}[\tilde{H}(T)]$ depends on the particular form of β. With $\beta(x) \sim \sqrt{x}$, b being a positive constant, we have $I_1(T) = 4T^{1/2} + o(\sqrt{\log(T)})$ if $\beta(x) = \mathbb{I}_{(0,b]}(x) + x^{1/2}\mathbb{I}_{(b,\infty)}(x)$ and $I_1(T) = 4T^{1/2} - 2\log(T) + o(\sqrt{\log(T)})$ if $\beta(x) = (1 + x^{1/2})$. By comparing these trends with the one in (28) one can appreciate its reduction from T to $T^{1/2} + o(T^{1/2})$.

(c4) Replacing the extended gamma CRM with a beta process we have $I_1(T) = 2T - 1/2$ whatever the choice of c. Moreover, if $c(x) \sim \sqrt{x}$ as $x \to \infty$, we obtain $I_2 \sim 8T^{1/2}$ and $I_3 \sim d\log(T)$ for some $d > 0$. By setting $C_0(k,T) = T^{-1/4}$, (19) and (20) are met and Theorem 1 leads to

$$\frac{1}{T^{1/4}}[\tilde{H}(T) - 2T] \xrightarrow{\text{law}} X,$$

where $X \sim \mathscr{N}(0,8)$. Hence, with respect to (28), the trend is unchanged and the rate of divergence halved.

By means of the previous examples the impact of a nonhomogeneous CRM becomes apparent: a nonhomogeneous CRM allows to reduce both the trend of the cumulative hazard and the rate at which it diverges from its trend. An extended gamma CRM is able to reduce both, whereas a beta CRM affects only the rate of divergence from the trend. Overall, by studying also other examples, not reported here, of functions β and c with the four different kernels considered above, some interesting indications can be drawn. For instance, denote by T^η the rate at which the cumulative hazard based on the homogeneous version of an extended gamma (or beta) CRM diverges from its trend (e.g., $\eta = 3/2$ in the Dykstra–Laud case). Then, by choosing a suitable diverging β (or c) the rate can be tuned at any order in the range $(T^{\eta-1/2}, T^\eta]$. Analogous conclusions can be derived for the trend when using a hazard based on an extended gamma CRM: the trend corresponding to the homogeneous case T^α (e.g., $\alpha = 2$ for the Dykstra–Laud kernel) can be tuned by the choice of β at any rate in the range $(T^{\alpha-1}, T^\alpha]$.

4.2. *Asymptotics for quadratic functionals.* In this section we consider quadratic functionals of the random hazard rate. We derive central limit theorems for the path-second moments and the path-variances of hazard rates with specific kernels and driving CRM. Our results will be mainly based on Theorems 2 and 3. As in the previous section, we first deal with general homogeneous CRM such that $\int_{[1,\infty)} v^4 \rho(dv) < \infty$; this requirement combined with the structure of kernels we consider ensures that (H2) is satisfied. Finally set, as before, $K_\rho^{(i)} = \int_0^\infty s^i \rho(ds)$, for $i = 1, 2, 3, 4$.

(i) *Rectangular kernel.* We start by considering the rectangular kernel and derive CLTs for the path-second moment and for the path-variance of hazard rates. Some simple calculations lead to write, for $T > 2\tau$,

$$k_T^{(1)}(s, x; t, y) = \begin{cases} \frac{st}{T}(x \wedge y + \tau), & x, y < \tau, \\ \frac{st}{T}(x \wedge y + 2\tau - x \vee y), & \tau \le x, y < T - \tau, |x - y| \le 2\tau, \\ \frac{st}{T}[T + \tau - x \vee y], & T - \tau \le x, y < T + \tau, \\ 0, & \text{elsewhere.} \end{cases}$$

Moreover, $k_T^{(2)}(s, x) = sT^{-1}k_T^{(0)}(s, x)$ and for $T > 2\tau$, one has

$$k_T^{(3)}(s, x) = \begin{cases} \frac{sK_\rho^{(1)}}{T}\left[2\tau x + \frac{3}{2}\tau^2\right], & 0 < x < \tau, \\ \frac{sK_\rho^{(1)}}{T}\left[-\frac{1}{2}x^2 + 2\tau(x + \tau)\right], & \tau \le x < 2\tau, \\ \frac{sK_\rho^{(1)}}{T}4\tau^2, & 2\tau \le x < T - \tau, \\ \frac{sK_\rho^{(1)}}{T}[2\tau(T + \tau - x)] & T - \tau \le x < T + \tau, \\ 0, & \text{elsewhere.} \end{cases}$$

In order to apply Theorem 2 let us first consider Condition 1, which allows us to determine the rate function; it turns out that $C_1(k, T) = \sqrt{T}$ since

$$2T\|k_T^{(1)}\|^2_{L^2(\nu^2)} \to \sigma_1^2(k) = \frac{32\tau^3(K_\rho^{(2)})^2}{3}. \tag{33}$$

The verification of conditions 2–6 can be achieved by simple though quite lengthy calculations.

Indeed, letting, for $i = 1, \dots, 4$, d_i be a positive constant, one obtains:

2. $T^2\|k_T^{(1)}\|^4_{L^4(\nu^2)} \sim \frac{d_1}{T} \to 0,$
3. $T^2\|k_T^{(1)} \star_1^1 k_T^{(1)}\|^2_{L^2(\nu^2)} \sim \frac{d_2}{T} \to 0,$
4. $T^2\|k_T^{(1)} \star_2^1 k_T^{(1)}\|^2_{L^2(\nu)} \sim \frac{d_3}{T} \to 0,$
5. $T\|k_T^{(2)} + 2k_T^{(3)}\|^2_{L^2(\nu)} \to \sigma_2^2(k) = 16\tau^2\left[\frac{K_\rho^{(4)}}{4} + 2\tau K_\rho^{(3)}K_\rho^{(1)} + 4\tau^2 K_\rho^{(2)}(K_\rho^{(1)})^2\right],$
6. $T^{3/2}\|k_T^{(2)} + 2k_T^{(3)}\|^3_{L^3(\nu)} \sim \frac{d_4}{T^{1/2}} \to 0.$

Since

$$\frac{1}{T}\int_0^T \mathbb{E}(\tilde{h}(t)^2)\,dt = 2\tau K_\rho^{(2)} + 4\tau^2\big(K_\rho^{(1)}\big)^2 + o(T^{-1/2}), \tag{34}$$

we deduce from Theorem 2 the following asymptotic result, concerning the path-second moment of $\tilde{h}(t)$:

$$T^{1/2}\left\{\frac{1}{T}\int_0^T \tilde{h}(t)^2\,dt - \big(2\tau K_\rho^{(2)} + 4\tau^2\big(K_\rho^{(1)}\big)^2\big)\right\} \xrightarrow{\text{law}} X,$$

where $X \sim \mathscr{N}(0, \sigma_1^2(k) + \sigma_2^2(k))$ with

$$\sigma_1^2(k) + \sigma_2^2(k) = 16\tau^2\left[\frac{K_\rho^{(4)}}{4} + \tau K_\rho^{(3)} K_\rho^{(1)} + \frac{2\tau (K_\rho^{(2)})^2}{3} + \tau^2 K_\rho^{(2)}\big(K_\rho^{(1)}\big)^2\right].$$

Now we concentrate on a CLT involving the path-variance of $\tilde{h}(t)$, that we shall obtain as an application of Theorem 3. In particular, we must verify that Conditions 1, 2 and 3 in the statement of such result are verified, for some appropriate positive constants $\delta(k)$ and $\sigma_3^2(k)$. Indeed, one has that, as $T \to +\infty$,

$$\frac{C_1(k,T)}{(TC_0(k,T))^2} = T^{-1/2} \to 0, \tag{35}$$

$$\frac{2C_1(k,T)}{T^2 C_0(k,T)}\mathbb{E}[\tilde{H}(T)] = \frac{2}{T}\big\{2\tau K_\rho^{(1)} T + o(T)\big\} \to 4\tau K_\rho^{(1)} := \delta(k), \tag{36}$$

and also

$$\|C_1(k,T)\big(k_T^{(2)} + 2k_T^{(3)}\big) - \delta(k) C_0(k,T) k_T^{(0)}\|_{L^2(\nu)}^2 \to 4\tau^2 K_\rho^{(4)}. \tag{37}$$

The fact that $\mathbb{E}[\tilde{H}(T)] = K_\rho^{(1)}\{2T\tau - \frac{1}{2}\tau^2\}$ combined with (34) yields

$$\frac{1}{T}\int_0^T \mathbb{E}\left[\tilde{h}(t) - \frac{\mathbb{E}[\tilde{H}(T)]}{T}\right]^2 dt = 2\tau K_\rho^{(2)} + o(T^{-1/2}).$$

Hence, by using (35)–(37), we deduce from Theorem 3 that

$$\sqrt{T} \times \left\{\frac{1}{T}\int_0^T \left[\tilde{h}(t) - \frac{1}{T}\tilde{H}(T)\right]^2 dt - 2\tau K_\rho^{(2)}\right\} \xrightarrow{\text{law}} X,$$

where $X \sim \mathscr{N}(0, 4\tau^2[\frac{8\tau(K_\rho^{(2)})^2}{3} + K_\rho^{(4)}])$.

(ii) *Ornstein–Uhlenbeck kernel.* Let us now derive the CLT for the path-second moment and the path-variance of hazards based on the Ornstein–Uhlenbeck kernel. For this case we easily obtain

$$k_T^{(1)}(s,x;t,y) = \frac{st}{T} e^{\kappa(x+y)}\big(e^{-2\kappa(x\vee y)} - e^{-2\kappa T}\big)\mathbb{I}_{(0\le x,y\le T)},$$

$$k_T^{(2)}(s,x) = \frac{s^2}{T} e^{2\kappa x}(e^{-2\kappa x} - e^{-2\kappa T})\mathbb{I}_{(0\le x\le T)},$$

$$k_T^{(3)}(s,x) = \frac{sK_\rho^{(1)}}{\kappa T}\big[(e^{-2\kappa T} - e^{-2\kappa x})(e^{\kappa x} - e^{2\kappa x}) + \big(1 - e^{-\kappa(T-x)}\big)^2\big]\mathbb{I}_{(0\le x\le T)},$$

and some tedious algebra allows to derive also $k_T^{(1)} \star_1^1 k_T^{(1)}$ and $k_T^{(1)} \star_2^1 k_T^{(1)}$. Condition 1 in Theorem 2 is verified by choosing $C_1(k,T) = \sqrt{T}$; indeed,

$$2T\|k_T^{(1)}\|^2_{L^2(\nu^2)} \to \sigma_1^2(k) = \frac{2(K_\rho^{(2)})^2}{\kappa}. \tag{38}$$

Standard calculations allow to verify the validity of the other conditions in the statement of Theorem 2. In particular, by letting d_i $(i = 1,\dots,4)$ be a positive constant, one obtains:

2. $T^2\|k_T^{(1)}\|^4_{L^4(\nu^2)} \sim \frac{d_1}{T} \to 0,$
3. $T^2\|k_T^{(1)} \star_1^1 k_T^{(1)}\|^2_{L^2(\nu^2)} \sim \frac{d_2}{T} \to 0,$
4. $T^2\|k_T^{(1)} \star_2^1 k_T^{(1)}\|^2_{L^2(\nu)} \sim \frac{d_3}{T} \to 0,$
5. $T\|k_T^{(2)} + 2k_T^{(3)}\|^2_{L^2(\nu)} \to \sigma_2^2(k) = K_\rho^{(4)} + \frac{8}{\kappa}K_\rho^{(3)}K_\rho^{(1)} + \frac{16}{\kappa^2}K_\rho^{(2)}(K_\rho^{(1)})^2,$
6. $T^{3/2}\|k_T^{(2)} + 2k_T^{(3)}\|^3_{L^3(\nu)} \sim \frac{d_4}{T^{1/2}} \to 0.$

Since, as $T \to +\infty$,

$$\frac{1}{T}\int_0^T \mathbb{E}(\tilde{h}(t)^2)\,dt = K_\rho^{(2)} + \frac{2(K_\rho^{(1)})^2}{\kappa} + o(T^{-1/2}), \tag{39}$$

we deduce from Theorem 2 the following result for the path-second moment:

$$T^{1/2}\left\{\frac{1}{T}\int_0^T \tilde{h}(t)^2\,dt - \left[K_\rho^{(2)} + \frac{2(K_\rho^{(1)})^2}{\kappa}\right]\right\} \xrightarrow{\text{law}} X,$$

where $X \sim \mathscr{N}(0, K_\rho^{(4)} + \frac{8}{\kappa}K_\rho^{(3)}K_\rho^{(1)} + \frac{2(K_\rho^{(2)})^2}{\kappa} + \frac{16}{k^2}K_\rho^{(2)}(K_\rho^{(1)})^2)$. As far as the path-variance is concerned, one verifies easily that the conditions of Theorem 3 are verified, with $\delta(k) = \frac{2^{3/2}}{\sqrt{\kappa}}K_\rho^{(1)}$ and $\sigma_3^2(k) := K_\rho^{(4)}$.

Using (39), it is straightforward to see that

$$\frac{1}{T}\int_0^T \mathbb{E}\left[\tilde{h}(t) - \frac{\mathbb{E}[\tilde{H}(T)]}{T}\right]^2 dt = K_\rho^{(2)} + o(T^{-1/2}).$$

As a consequence, we deduce from Theorem 3 that

$$\sqrt{T} \times \left\{\frac{1}{T}\int_0^T \left[\tilde{h}(t) - \frac{1}{T}\tilde{H}(T)\right]^2 dt - K_\rho^{(2)}\right\} \xrightarrow{\text{law}} X,$$

with $X \sim \mathscr{N}(0, \frac{2(K_\rho^{(2)})^2}{\kappa} + K_\rho^{(4)})$.

Before considering the Dykstra and Laud kernel and the U-shaped kernel, let us make the previous results completely explicit by specifying the background driving CRM. For both the rectangular and the Ornstein–Uhlenbeck kernel the rate function is the same and the CRM affects the variance of the limiting Gaussian random variable for both path-second moment and path-variance of the hazard rate. Take, as before, the generalized gamma CRM with Poisson intensity (6) and denote the Pochhammer symbol by $(a)_n := \Gamma(a+n)/\Gamma(a)$. For this choice we have $K_\rho^{(c)} = [(1-\sigma)_{c-1}](\gamma^{c-\sigma})^{-1}$ for any $c > 0$. For the Ornstein–Uhlenbeck kernel the variance is then given by

$$\sigma_1^2(k) + \sigma_2^2(k) = \frac{(1-\sigma)(16\kappa^{-1}\gamma^{2\sigma} + 2(9-5\sigma)\gamma^\sigma + \kappa(2-\sigma)_2)}{\kappa\gamma^{4-\sigma}} \tag{40}$$

which decreases as κ and γ increase for any given (γ, σ) and (κ, σ), respectively. Moreover, it is maximized by $\sigma = 0$ for low values of κ and γ, whereas, for moderately large values of κ and γ, the maximizing σ increases as κ and γ increase. For instance, if $\kappa = 0.5$ and $\gamma = 2$, the maximizing σ is approximately equal to 0.4 and the overall variance is 7.55. To highlight the incidence of the prior parameters note that with $\kappa = 1$ and $\gamma = 5$, the maximizing σ and the variance are approximately equal to 0.77 and 0.29, respectively. Using the asymptotic variance as a guideline for fixing the prior parameters seems a sensible and straightforward choice since it summarizes in a single expression the various effects of the parameters. Turning to the path-variance, a hazard based on a generalized gamma CRM with Ornstein–Uhlenbeck kernel will have variance given by

$$\sigma_1^2(k) + \sigma_3^2(k) = \frac{(1-\sigma)(2(1-\sigma)\gamma^\sigma + \kappa(2-\sigma)_2)}{\kappa\gamma^{4-\sigma}}, \tag{41}$$

which behaves similarly to (40) but, obviously, leads to smaller values. Considering the same set of parameters as above we have: if $\kappa = 0.5$ and $\gamma = 2$, $\sigma = 0$ maximizes (41) and its value is 0.625; if $\kappa = 1$ and $\gamma = 5$, (41) is maximized by $\sigma \approx 0.1$ leading to a variance of 0.01. Similar considerations hold also for the asymptotic variance of a hazard based on the rectangular kernel combined with a generalized gamma CRM.

Turning attention to quadratic functionals of hazards based on nonhomogeneous CRM, the importance of our Theorem 4 becomes apparent: the verification of the conditions of Theorems 2 and 3 becomes extremely difficult if not impossible. Hence, when it is possible to bound above and below the Poisson intensity of a nonhomogeneous CRM so as to meet the conditions of Theorem 4, we are still able to state that the rate function is $C_1(k, T) = T^{1/2}$ for hazards based on rectangular and Ornstein–Uhlenbeck kernels. Moreover, we can deduce the convergence, along some subsequence $T_{n'}$ of every diverging sequence T_n, of the path-second moment and of the path-variance to a Gaussian random variable with variance taking value in the range $[\sigma_1^2(\bar{k})+\sigma_2^2(\bar{k}),\sigma_1^2(\bar{\bar{k}})+\sigma_2^2(\bar{\bar{k}})]$ and $[\sigma_1^2(\bar{k})+\sigma_3^2(\bar{k}),\sigma_1^2(\bar{\bar{k}})+\sigma_3^2(\bar{\bar{k}})]$, respectively. In order to deduce convergence for every diverging sequence, the structure of the Poisson intensity has to be specified as well. Thus, let us consider again the extended gamma and beta CRMs. As noted in Section 4.1, supposing $\beta(x)=\bar{\beta}$ in (7) and $c(x)=\bar{c}$ in (8), the CRMs become homogeneous and the previous results hold with the same rate functions. Note that, for $a>0$, $K_\rho^{(a)}=\Gamma(a)\bar{\beta}^{-a}$ in the extended gamma case and $K_\rho^{(a)}=\Gamma(a)[(1+\bar{c})_{a-1}]^{-1}$ in the beta case. Hence, with an Ornstein–Uhlenbeck kernel the asymptotic variance of the path-second moment is equal to $(\bar{\beta}^4\kappa^2)^{-1}2(3\kappa^2+9\kappa+8)$ for the former and equal to $[\kappa^2(1+\bar{c})(1+\bar{c})_3]^{-1}2(9\kappa\bar{c}^2+37\kappa\bar{c}+30\kappa+3\kappa^2(1+\bar{c})+8(1+\bar{c})_3)$ for the latter. For the path-variance similar expressions are obtained. If β (or c) are functions bounded by some finite constant M, then we are in the genuinely nonhomogeneous case and, as mentioned above, by Theorem 4 CLTs along subsequences of diverging sequences are granted. To achieve convergence along any sequence, it is enough to suppose that β (or c) are eventually nondecreasing (or nonincreasing), which represents a sensible choice in any application. For instance, considering an extended gamma CRM with nondecreasing β taking values in $[L,M]$ combined with an Ornstein–Uhlenbeck kernel, the path-second moment will converge, along any sequence, to a Gaussian random variable with variance $\sigma_1^2(k)+\sigma_2^2(k)=(M^4\kappa^2)^{-1}2(3\kappa^2+9\kappa+8)$. Analogous considerations hold for the path-variance.

(iii) *Dykstra–Laud and U-shaped kernels.* Our results for quadratic functionals do not apply when choosing the kernel k to be the Dykstra–Laud or U-shaped kernel. Indeed, for both kernels conditions 3, 5 and 6 in Theorem 2 are not met. Moreover, also the additional conditions 1–3 in Theorem 3 are not satisfied. Note that condition 3 represents the most delicate since it involves a contraction. Consider first the Dykstra–Laud kernel. It is easy to see that $k_T^{(1)}(s,x;t,y)=\frac{st}{T}(T-(x\vee y))\mathbb{I}_{(0\le x,y\le T)}$ and that $k_T^{(1)}\star_1^1 k_T^{(1)}(s,x;t,y)=\frac{stK_\rho(T-(x\vee y))}{T^2}[T^2-(x\wedge y)^2-\frac{(T-(x\vee y))^2}{3}]\mathbb{I}_{(0\le x,y\le T)}$. As for condition 1 we obtain with the choice $C_1=T^{-1}$

$$\frac{2}{T^2}\|k_T^{(1)}\|_{L^2(\nu^2)}^2\to\frac{K_\rho^2}{3}.$$

This, however, implies that the quantity in condition 3 converges to a positive constant and the ones in conditions 5 and 6 diverge. In Theorem 3 we obtain that the quantity in condition 1 is equal to 1 and the one in condition 2 diverges. Finally, condition 3 cannot be satisfied since condition 5 in Theorem 2 is violated. For the U-shaped kernel we obtain again $C_1(k,T)=T^{-1}$ and the asymptotic behavior of the various quantities involved in the conditions is the same as the one of the Dykstra and Laud kernel. We have also tried with nonhomogeneous CRM; indeed, it seems possible to obtain $C_1(k,T)=T^{-\eta}$ with any $\eta\in(0,1]$, but the conditions are nonetheless violated.

The fact that our results do not work for the Dykstra–Laud and U-shaped kernels seems to suggest that kernels yielding monotone increasing hazards (at least from some point onward as is the case for the U-shaped kernel) exhibit a too strong growth to be compatible with our conditions. Future research will focus, on one side, on the translation of the conditions into simple and intuitive sufficient ones regarding the behavior of the hazard rate induced by different classes of kernels and, on the other side, to relax the conditions in order to cover models for monotone increasing hazards.

5. From prior to posterior central limit theorems. For the reader interested in Bayesian applications, a natural question is whether our results can be extended in order to study the posterior asymptotic behavior of linear and quadratic functionals of random hazard rates, conditionally on an arbitrarily large number of observations. At a first look, the question seems arduous, since the models we have considered so far are, in general, not conjugate. However, the problem can be completely overcome by using the powerful (and general) theory for posterior models developed in James [22]. Indeed, James proves that the law of the random hazard rate (1), conditionally on $\mathbf{X}$ and on a suitable set of *latent variables* $\mathbf{Y}$, coincides with the distribution of the random object

$$\tilde{h}^{\mathbf{X},\mathbf{Y}}(t) = \int_{\mathbb{X}} k(t,x)\tilde{\mu}^*(dx) + \sum_{i=1}^{k} J_i k(t, Y_i^*), \tag{42}$$

where μ^* is a *nonhomogeneous CRM* with updated intensity measure and the latent variable Y_i^* is the location of a random jump J_i for $i = 1, \dots, k$. It follows that one can render the posterior behavior of any functional of h amenable to asymptotic analysis, by first conditioning on $\mathbf{X}$ and $\mathbf{Y}$, and then by separating the roles of μ^* and of the fixed jumps J_i. In [7], which, in Bayesian terms, represents a logical continuation of the present paper, this idea is exploited and asymptotic results for (42) derived. From a technical point of view, the techniques require an average over the distribution of $\mathbf{Y}|\mathbf{X}$, which can be decomposed into the distribution of $\mathbf{Y}$ given a certain random partition $\mathbf{p}$ and the sample $\mathbf{X}$, and the partition distribution given $\mathbf{X}$. As anticipated in the Introduction, one of the main findings of [7] is that the posterior behavior of linear and quadratic functionals of random hazard rates coincides in most cases with the prior one. This phenomenon is also compared with another asymptotic study of Bayesian nonparametric models, which is customarily referred to as *Bayesian consistency* (see, e.g., [11, 16]).

6. Proofs and further techniques. In this section we collect the proofs of the main results of the paper. As anticipated, we shall make a substantial use of the CLTs, for sequences of single and double Poisson integrals, recently established by [38]. In the next subsection we present some preliminary results concerning double Wiener–Itô integrals, with special attention devoted to weak convergence and central limit theorems. Virtually all of the needed background material, about stochastic integrals of any order with respect to Poisson measures, can be found in [45] and in Chapter 10 of [29]. A different approach, based on Hilbert space techniques, is described in [35]. The reader is also referred to [46] for an updated review of related convergence results.

6.1. *Double integrals and CLTs.* Throughout this section we consider a Poisson CRM $\tilde{N}$ such that (H1) is verified. Recall that $\tilde{N}^c$ is the compensated Poisson measure defined in formulas (12) and (13). For every $f \in L^2_{s,0}(\nu^2)$, we denote by $I_2^{\tilde{N}^c}(f)$ the *double Wiener–Itô integral* of f with respect to $\tilde{N}^c$. The reader is referred to [45] for precise definitions. Here, we shall recall that, if $f \in L^2_{s,0}(\nu^2)$ is a piecewise constant function with support contained in a product set $S \times S \subset (\mathbb{R}^+ \times \mathbb{X})^2$ such that $\nu(S) < +\infty$, then the variable $I_2^{\tilde{N}^c}(f)$ is a genuine ("pathwise") double integral with respect to the restriction to $S \times S$ of the (signed) product measure $\tilde{N}^c(ds, dx)\tilde{N}^c(dt, dy)$. The very nature of f implies that the integration is performed on the intersection between $S \times S$ and the nondiagonal set D_0^2. For a general $f \in L^2_{s,0}(\nu^2)$, $I_2^{\tilde{N}^c}(f)$ is simply the limit in $L^2(\mathbb{P})$ of random variables of the kind $I_2^{\tilde{N}^c}(f_k)$ where each $f_k \in L^2_{s,0}(\nu^2)$ is a piecewise constant function with support in a product set $S_k \times S_k$ with ν^2-finite measure. The following isometric relation is well known: $\forall f_1, f_2 \in L^2_{s,0}(\nu^2)$

$$\begin{aligned} &\mathbb{E}[I_2^{\tilde{N}^c}(f_1) I_2^{\tilde{N}^c}(f_2)] \\ &\quad = 2\int_{\mathbb{R}^+\times\mathbb{X}}\int_{\mathbb{R}^+\times\mathbb{X}} f_1(s,x;t,y) f_2(s,x;t,y)\nu(ds,dx)\nu(dt,dy). \end{aligned} \tag{43}$$

When $f \in L^2_s(\nu^2)$ (hence f does not necessarily vanish on diagonals), we set $I_2^{\tilde{N}^c}(f) = I_2^{\tilde{N}^c}(f\mathbb{I}_{D_0^2})$, and we observe that the isometry property (43) still holds. Indeed, ν is nonatomic, and therefore ν^2 does not charge diagonals [even though $\tilde{N}^c(ds, dx)\tilde{N}^c(dt, dy)$ does]. We also recall the *product formula*

$$\begin{aligned} &\tilde{N}^c(g)\tilde{N}^c(h) \\ &\quad = (g,h)_{L^2(\nu)} + \int_{\mathbb{R}^+\times\mathbb{X}} g(s,x)h(s,x)\tilde{N}^c(ds,dx) + I_2^{\tilde{N}^c}(\widetilde{h\otimes g}), \end{aligned} \tag{44}$$

where $h \otimes g(s, x; t, y) = h(s, x)g(t, y) \in L^2(\nu^2)$ and $(\tilde{\ })$ stands for a symmetrization, which holds for every $f, g \in L^2(\nu)$ such that $g(s, x)h(s, x) \in L^2(\nu)$.

Finally, we state the main results proved in [38]. We consider a sequence of double integrals

$$F_n = I_2^{\tilde{N}^c}(f_n), \qquad n \geq 1, \tag{45}$$

where $f_n \in L^2_{s,0}(\nu^2)$. We will suppose that the following technical assumptions are satisfied: the sequence f_n, $n \geq 1$, in (45) is such that, for every $n \geq 1$,

$$\|f_n\|_{L^2(\nu^2)} > 0 \quad \text{and} \quad f_n \star_2^1 f_n \in L^2(\nu), \tag{1}$$

$$\left\{\int_{\mathbb{R}^+ \times \mathbb{X}} f_n(s, y; \cdot)^4 \nu(ds, dy)\right\}^{1/2} \in L^1(\nu), \tag{2}$$

where we use the notation introduced in (9)–(11), and moreover, as $n \to +\infty$,

$$\int_{\mathbb{R}^+ \times \mathbb{X}} \int_{\mathbb{R}^+ \times \mathbb{X}} f_n(s, y; t, x)^4 \nu(ds, dy)\nu(dt, dx) \to 0. \tag{3}$$

Note that (3) implies, in particular, that $f_n \in L^4(\nu^2)$ for every n. See [38] for a discussion of the role of (1)–(3). In the subsequent sections, we will see how such assumptions restrict the set of the random hazard rates that can be studied by our techniques. The next result is a CLT involving sequences of double integrals.

THEOREM 5 [37]. *Define the sequence* $F_n = I_2^{\tilde{N}^c}(f_n)$ *and* $f_n \in L^2_{s,0}(\nu^2)$, $n \geq 1$, *as in* (45), *and suppose* (1)–(3) *hold. Then*, $f_n \star_1^0 f_n \in L^2(\nu^3)$ *for every* $n \geq 1$, *and moreover*:

1. *if*

$$\begin{aligned} &\|f_n\|^{-2}_{L^2(\nu^2)} \times (f_n \star_1^1 f_n) \to 0 \qquad \text{in } L^2(\nu^2) \quad \text{and} \\ &\|f_n\|^{-2}_{L^2(\nu^2)} \times (f_n \star_2^1 f_n) \to 0 \qquad \text{in } L^2(\nu), \end{aligned} \tag{46}$$

then

$$2^{-1/2}\|f_n\|^{-1}_{L^2(\nu^2)} \times F_n \xrightarrow{\text{law}} X, \tag{47}$$

where $X \sim \mathscr{N}(0, 1)$ *is a standard Gaussian random variable*;

2. *if* $F_n \in L^4(\mathbb{P})$ *for every* n, *then a sufficient condition to have* (46) *is that*

$$\left(2\|f_n\|^2_{L^2(\nu^2)}\right)^{-2}\mathbb{E}(F_n^4) \to 3; \tag{48}$$

3. *if the sequence* $\{(2\|f_n\|^2_{L^2(\nu^2)})^{-2} F_n^4 : n \geq 1\}$ *is uniformly integrable, then conditions* (46), (47) *and* (48) *are equivalent.*

Theorem 5 is proved by using a decoupling technique, known as the *principle of conditioning*, which has been adapted to the framework of CRM by means of the general theory of stable convergence developed in [39]. The next result gives sufficient conditions to have that the law of a random vector, composed of a single and of a double integral, converges weakly to a bivariate Gaussian law. The proof is essentially based on an appropriate version of the *product formulas* for multiple stochastic integrals, proved, for example, in [45].

THEOREM 6 [37]. (A) *Consider a sequence*

$$G_n = \tilde{N}^c(g_n), \qquad n \geq 1,$$

where $g_n \in L^2(\nu) \cap L^3(\nu)$ *and* $\|g_n\|_{L^2(\nu)} > 0$, *and suppose that, as* $n \to +\infty$,

$$\|g_n\|_{L^2(\nu)}^{-3} \int_{\mathbb{R}^+ \times \mathbb{X}} |g_n(s, y)|^3 \nu(ds, dy) \to 0. \tag{49}$$

Then, $\|g_n\|_{L^2(\nu)}^{-1} \times G_n \xrightarrow{\text{law}} X$, *where* $X \sim \mathscr{N}(0, 1)$ *is a centered standard Gaussian random variable*.

(B) *Consider a sequence* $F_n = I_2^{\tilde{N}^c}(f_n)$, $n \geq 1$, *with* $f_n \in L_{s,0}^2(\nu^2)$ *as in* (45), *and a sequence* $G_n = \tilde{N}^c(g_n)$, $n \geq 1$, *as at point* (A). *Suppose moreover that*:

(i) *The sequence* (f_n) *verifies assumptions* (1)–(3), *and satisfies condition* (46).
(ii) *The sequence* (g_n) *satisfies* (49).

Then, *as* $n \to +\infty$,

$$\big(2^{-1/2}\|f_n\|_{L^2(\nu^2)}^{-1} \times F_n, \|g_n\|_{L^2(\nu)}^{-1} \times G_n\big) \xrightarrow{\text{law}} (X, X'), \tag{50}$$

where $X, X' \sim \mathscr{N}(0, 1)$ *are two independent*, *centered standard Gaussian random variables*.

Part (B) of Theorem 6 implies in particular that, whenever conditions (46) and (50) are met, the (componentwise) convergence of $\|f_n\|^{-1} \times F_n$ and $\|g_n\|^{-1} \times G_n$, toward a Gaussian distribution, *implies necessarily* the joint convergence of the vector $(\|f_n\|^{-1}F_n, \|g_n\|^{-1}G_n)$. This conclusion echoes results already established in the framework of Gaussian CRM (see [40]).

Now consider the positive kernel k, which defines $\tilde{h}$ via (1), and suppose (here and for the remainder of the section) that k satisfies assumption (H2). In the next two lemmas we collect some straightforward facts which will be used throughout the sequel.

LEMMA 7. *The two processes* $\tilde{h}(t)$, $t \geq 0$, *and*

$$\tilde{h}_*(t) := \tilde{N}^c((\cdot)k(t, \cdot)) + \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t, x)\nu(ds, dx), \qquad t \geq 0,$$

where

$$\tilde{N}^c((\cdot)k(t, \cdot)) := \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t, x)\tilde{N}^c(ds, dx), \tag{51}$$

have the same law.

PROOF. Use (5) and (14) to compute the two transforms

$$\mathbb{E}\big[e^{i\sum_{j=1}^n \lambda_j \tilde{h}(t_j)}\big] \quad \text{and} \quad \mathbb{E}\big[e^{i\sum_{j=1}^n \lambda_j \tilde{h}_*(t_j)}\big],$$

for every $n \geq 1$, every $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ and every $t_1, \dots, t_n \geq 0$. $\square$

LEMMA 8. *For every* $T > 0$,

$$\int_0^T \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t,x) \tilde{N}^c(ds,dx)\,dt = \tilde{N}^c\big(k_T^{(0)}\big), \tag{52}$$

$$\frac{1}{T}\int_0^T \int_{\mathbb{R}^+ \times \mathbb{X}} s^2 k(t,x)^2 \tilde{N}^c(ds,dx)\,dt = \tilde{N}^c\big(k_T^{(2)}\big), \tag{53}$$

where $k_T^{(0)}$ *and* $k_T^{(2)}$ *are given, respectively, by* (18) *and* (23). *If* $k_T^{(3)} \in L^2(\nu) \cap L^1(\nu)$,

$$\frac{1}{T}\int_0^T \tilde{N}^c((\cdot)k(t,\cdot))\left(\int_{\mathbb{R}^+ \times \mathbb{X}} sk(t,x)\nu(ds,dx)\right) dt = \tilde{N}^c\big(k_T^{(3)}\big). \tag{54}$$

Analogously, for every $T > 0$,

$$\frac{1}{T}\int_0^T I_2^{\tilde{N}^c}\big([(\cdot)k(t,\cdot)] \otimes [(\cdot)k(t,\cdot)]\big)\,dt = I_2^{\tilde{N}^c}\big(k_T^{(1)}\big), \tag{55}$$

where $[(\cdot)k(t,\cdot)] \otimes [(\cdot)k(t,\cdot)](u,x;v,y) := uvk(t,x)k(t,y)$, *and* $k_T^{(1)}$ *is defined according to* (22).

The proof of Lemma 8 is trivial when the map $(t,x) \mapsto k(t,x)$ is piecewise constant; indeed, in this case (52)–(55) follow immediately from the application of a standard Fubini theorem. The general statement is obtained by a density argument; we omit the details here (one can, e.g., mimic the proof of Lemma 13 in [36]).

Finally note that, given two sequences of random variables $\{A_n\}$ and $\{B_n\}$ such that $A_n - B_n \to 0$ in probability, we will sometimes write

$$A_n \overset{\mathbb{P}}{\approx} B_n.$$

6.2. *Proof of Theorem* 1. Use Lemma 7 and relations (51) and (52) to write

$$\begin{aligned}
\tilde{H}(T) &\overset{\text{law}}{=} \int_0^T \tilde{h}_*(t)\,dt \\
&= \int_0^T \tilde{N}^c((\cdot)k(t,\cdot))\,dt + \int_0^T \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t,x)\nu(ds,dx)\,dt \\
&= \int_0^T \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t,x)\tilde{N}^c(ds,dx)\,dt + \int_0^T \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t,x)\nu(ds,dx)\,dt \\
&= \tilde{N}^c\big(k_T^{(0)}\big) + \int_0^T \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t,x)\nu(ds,dx)\,dt,
\end{aligned}$$

which yields, via the relation $\mathbb{E}(\tilde{H}(T)) = \int_0^T \int_{\mathbb{R}^+ \times \mathbb{X}} sk(t,x)\nu(ds,dx)\,dt$,

$$C_0(k,T) \times [\tilde{H}(T) - \mathbb{E}(\tilde{H}(T))] \overset{\text{law}}{=} \tilde{N}^c\big(C_0(k,T) \times k_T^{(0)}\big).$$

Since the isometry property (15) and the assumption (19) yield

$$\mathbb{E}[\tilde{N}^c(C_0(k,T)\times k_T^{(0)})^2] = C_0^2(k,T)\int_{\mathbb{R}^+\times\mathbb{X}}[k_T^{(0)}(s,x)]^2\nu(ds,dx)\to\sigma_0^2(k),$$

we deduce from part (A) of Theorem 6 [in the case $g_n = (C_0(k,T_n)/\sigma_0(k))\times k_{T_n}^{(0)}$, where T_n is any positive sequence diverging to infinity] that, since (20) holds, the CLT (21) must also take place.

6.3. *Proof of Theorem* 2. Use Lemma 7 to write [we adopt once again the notation (51)]

$$\begin{aligned}
&\frac{1}{T}\int_0^T \tilde{h}(t)^2\,dt\\
&\quad\overset{\text{law}}{=}\frac{1}{T}\int_0^T \tilde{N}^c((\cdot)k(t,\cdot))^2\,dt+\frac{1}{T}\int_0^T\left(\int_{\mathbb{R}^+\times\mathbb{X}}sk(t,x)\nu(ds,dx)\right)^2dt\\
&\qquad+\frac{2}{T}\int_0^T\tilde{N}^c((\cdot)k(t,\cdot))\left(\int_{\mathbb{R}^+\times\mathbb{X}}sk(t,x)\nu(ds,dx)\right)dt.
\end{aligned}$$

Now recall that, thanks to (54),

$$\frac{2}{T}\int_0^T\tilde{N}^c((\cdot)k(t,\cdot))\left(\int_{\mathbb{R}^+\times\mathbb{X}}sk(t,x)\nu(ds,dx)\right)dt=\tilde{N}^c(2k_T^{(3)}),$$

so that, by using (26),

$$\begin{aligned}
&C_1(k,T)\times\left\{\frac{1}{T}\int_0^T\tilde{h}(t)^2\,dt-\frac{1}{T}\int_0^T\mathbb{E}[\tilde{h}(t)^2]\,dt\right\}\\
&\quad\overset{\text{law}}{=}C_1(k,T)\times\left\{\frac{1}{T}\int_0^T\tilde{N}^c((\cdot)k(t,\cdot))^2\,dt\right.\\
&\qquad\qquad\left.+\tilde{N}^c(2k_T^{(3)})-\frac{1}{T}\int_0^T\int_{\mathbb{R}^+\times\mathbb{X}}s^2k(t,x)^2\nu(ds,dx)\right\}dt.
\end{aligned}\tag{56}$$

By applying the product formula (44) in the case $g(s,x)=h(s,x)=sk(t,x)$, for every $t\ge 0$ we obtain

$$\begin{aligned}
\tilde{N}^c((\cdot)k(t,\cdot))^2&=\int_{\mathbb{R}^+\times\mathbb{X}}s^2k(t,x)^2\nu(ds,dx)\\
&\quad+\int_{\mathbb{R}^+\times\mathbb{X}}s^2k(t,x)^2\tilde{N}^c(ds,dx)+I_2^{\tilde{N}^c}([(\cdot)k(t,\cdot)]\otimes[(\cdot)k(t,\cdot)]),
\end{aligned}$$

from which we deduce that, thanks to formulas (53) and (55), the expression in (56) is indeed equal to

$$C_1(k,T)\times\{\tilde{N}^c(k_T^{(2)}+2k_T^{(3)})+I_2^{\tilde{N}^c}(k_T^{(1)})\},$$

for every $T > 0$. It follows that Theorem 2 is proved, once it is shown that

$$\big(\tilde{N}^c\big(C_1(k,T) \times \big(k_T^{(2)} + 2k_T^{(3)}\big)\big), I_2^{\tilde{N}^c}\big(C_1(k,T) \times k_T^{(1)}\big)\big) \xrightarrow{\text{law}} (X, X'),$$

where X and X' are independent and such that $X \sim \mathscr{N}(0, \sigma_2^2(k))$ and $X' \sim \mathscr{N}(0, \sigma_1^2(k))$. To this end, we apply part (B) of Theorem 6: according to such a result, it is sufficient to check that, for every positive sequence $T_n \to +\infty$, the two sequences

$$g_n = \frac{C_1(k,T_n)}{\sigma_2(k)}\big(k_{T_n}^{(2)} + 2k_{T_n}^{(3)}\big) \quad \text{and} \quad f_n = \frac{C_1(k,T_n)}{\sigma_1(k)} k_{T_n}^{(1)}, \qquad n \geq 1,$$

satisfy, respectively, condition (49) and conditions (1)–(3) and (46). It is immediately seen that Assumptions 5 and 6 in the statement imply (49), and we are therefore left with the sequence $\{f_n\}$. Conditions (1) and (2) can be checked by standard iterations of the Jensen and Cauchy–Schwarz inequalities (see, e.g., Section 5.1 in [38] for several analogous computations). Finally, (3) is given by Assumption 2 in the statement, whereas Assumptions 3 and 4 give, respectively, the first and the second line in (46). This concludes the proof of Theorem 2.

6.4. *Proof of Theorem* 3. Write first

$$\frac{1}{T}\int_0^T \left[\tilde{h}(t) - \frac{1}{T}\tilde{H}(T)\right]^2 dt = \frac{1}{T}\int_0^T \tilde{h}(t)^2\,dt - \left(\frac{1}{T}\tilde{H}(T)\right)^2, \tag{57}$$

and observe that

$$\begin{aligned} C_1(k,T)\left(\frac{1}{T}\tilde{H}(T)\right)^2 &= \frac{C_1(k,T)}{T^2 C_0(k,T)^2}\{C_0(k,T)[\tilde{H}(T) - \mathbb{E}(\tilde{H}(T))]\}^2 \\ &\quad + \frac{C_1(k,T)}{T^2}\mathbb{E}(\tilde{H}(T))^2 \\ &\quad + 2\frac{C_1(k,T)}{T^2}\mathbb{E}(\tilde{H}(T))[\tilde{H}(T) - \mathbb{E}(\tilde{H}(T))]. \end{aligned} \tag{58}$$

From Assumption 1 in the statement, and since (19) and (20) are in order, we deduce

$$\frac{C_1(k,T)}{T^2 C_0(k,T)^2}\{C_0(k,T)[\tilde{H}(T) - \mathbb{E}(\tilde{H}(T))]\}^2 \xrightarrow{\mathbb{P}} 0. \tag{59}$$

Moreover, Assumption 2 in the statement yields that, as $T \to +\infty$,

$$\begin{aligned} &\frac{2C_1(k,T)}{T^2}\mathbb{E}(\tilde{H}(T))[\tilde{H}(T) - \mathbb{E}(\tilde{H}(T))] \\ &\quad \overset{\mathbb{P}}{\approx} \delta(k) C_0(k,T)[\tilde{H}(T) - \mathbb{E}(\tilde{H}(T))]. \end{aligned} \tag{60}$$

In view of Lemma 7, and by reasoning as in the proofs of Theorems 1 and 2, we infer from relations (57)–(60) that

$$\begin{aligned}
C_1(k,T) \times \left\{ \frac{1}{T}\int_0^T \left[\tilde{h}(t) - \frac{1}{T}\tilde{H}(T)\right]^2 dt - \frac{1}{T}\int_0^T \mathbb{E}[\tilde{h}(t)^2]\,dt + \frac{\mathbb{E}[\tilde{H}(T))]^2}{T^2} \right\} \\
\overset{\text{law}}{=} \tilde{N}^c\bigg(C_1(k,T)\big(k_T^{(2)} + 2k_T^{(3)}\big) - \frac{2C_1(k,T)}{T^2}\mathbb{E}(\tilde{H}(T))k_T^{(0)}\bigg) \\
+ I_2^{\tilde{N}^c}\big(C_1(k,T)k_T^{(1)}\big) \\
\overset{\mathbb{P}}{\approx} \tilde{N}^c\big(C_1(k,T)\big(k_T^{(2)} + 2k_T^{(3)}\big) - \delta(k)C_0(k,T)k_T^{(0)}\big) + I_2^{\tilde{N}^c}\big(C_1(k,T)k_T^{(1)}\big).
\end{aligned}$$

The conclusion is deduced from Assumption 3 in the statement, by applying Theorem 6 in the case

$$\begin{aligned}
g_n &= \frac{C_1(k,T_n)(k_{T_n}^{(2)} + 2k_{T_n}^{(3)}) - \delta(k)C_0(k,T_n)k_{T_n}^{(0)}}{\sigma_3(k)}, \\
f_n &= \frac{C_1(k,T_n)}{\sigma_1(k)}k_{T_n}^{(1)}, \qquad n \geq 1,
\end{aligned}$$

where $T_n \to +\infty$.

6.5. *Proof of Theorem* 4. To prove Part (A), observe that the assumptions imply the existence of two constants $0 < D_1 < D_2 < +\infty$, such that, for T sufficiently large,

$$D_1 < C_0^2(k,T) \times \int_{\mathbb{R}^+\times\mathbb{X}} \big[k_T^{(0)}(s,x)\big]^2 \nu(ds,dx) < D_2.$$

Standard arguments yield therefore that, for every sequence $T_n \to +\infty$, there exists a subsequence $T_{n'}$ such that, as $n' \to +\infty$,

$$C_0^2(k,T_{n'}) \times \int_{\mathbb{R}^+\times\mathbb{X}} \big[k_{T_{n'}}^{(0)}(s,x)\big]^2 \nu(ds,dx) \to \sigma^2(k) > 0,$$

where $\sigma^2(k)$ is some well-chosen positive constant. Moreover,

$$\begin{aligned}
C_0^3(k,T_{n'}) &\times \int_{\mathbb{R}^+\times\mathbb{X}} \big[k_{T_{n'}}^{(0)}(s,x)\big]^3 \nu(ds,dx) \\
&\leq C_0^3(k,T_{n'}) \int_{\mathbb{R}^+} \big[\overline{\overline{k}}_{T_{n'}}^{(0)}(s,x)\big]^3 \nu(ds,dx) \\
&\sim C_0^3(\overline{\overline{k}},T_{n'}) \int_{\mathbb{R}^+} \big[\overline{\overline{k}}_{T_{n'}}^{(0)}(s,x)\big]^3 \nu(ds,dx) \to 0.
\end{aligned}$$

The proofs of parts (B) and (C) are based on analogous computations, and are omitted.

7. Conclusions and future work. (I) Future research will focus on the generalization of our asymptotic results to general multiplicative intensity models [1], which include a wide variety of popular models such as Cox proportional hazards regression models, multiple decrement models, birth and death processes and nonhomogeneous Poisson processes. To fix ideas consider the Cox proportional hazards regression model, in which Z_i is an m-dimensional vector of covariates recorded for the ith individual and θ is a m-dimensional vector of unknown regression coefficients. Then the proportional hazards model is specified in terms of the hazard function relationship as

$$h_i(t) = h_0(t) \exp(\theta' Z_i),$$

where h_0 represents the so-called baseline hazard function. A Bayesian treatment leads to considering h_0 and θ to be random and, hence, by choosing $\tilde{h}_0$ to be a mixture as in (1) and π to be a prior for $\tilde{\theta}$, one obtains a semiparametric random hazard rate function for the ith individual of the form

$$\tilde{h}_i(t) = \exp(\tilde{\theta}' Z_i) \int_{\mathbb{X}} k(t, x) \tilde{\mu}(dx). \tag{61}$$

Bayesian analysis of the Cox model within this setup has been pursued in [20, 21, 34]. Since (1) still represents the basic building block of (61) and, indeed, also of other multiplicative intensity models, we aim at extending our results to random objects such as (61) and expect to obtain CLTs for which the limiting random variable is a suitable mixture of Gaussian distributions.

(II) The techniques exploited in Section 6, for deriving the main results of this paper, can be further generalized. As already mentioned, they are indeed based on a very general decoupling criterion, known as the *principle of conditioning*. As shown in [37–39], this principle can be applied to a wide class of stochastic integrals with respect to completely random measures, including multiple Wiener–Itô integrals of any order $n > 2$. In particular, we expect that the results of the present paper can be suitably extended to accommodate the asymptotic analysis of nonlinear and nonquadratic functionals, such as, for example, path-moments of order greater than 2. Note that results of this type are already available in the Gaussian case. See, for example, [40].

(III) By suitably tailoring the general results proved in [37–39], one could study the asymptotic behavior of more general linear and quadratic functionals associated with processes of the type (1), where the parameter t lives in a general space (e.g., the plane). Some examples of such processes are presented in [20, 49]. In such a framework, a crucial point is the choice of an appropriate definition of *path-variance*. Note that a similar analysis has been already performed for quadratic functionals of bivariate Gaussian processes (such as the Brownian sheet or the Kiefer process) in [9], where these results have been applied to the asymptotic analysis of independence test statistics.

Acknowledgment. We are grateful to an anonymous referee for a careful reading of the manuscript and for insightful suggestions.

REFERENCES

[1] AALEN, O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6** 701–726. MR491547

[2] BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63** 167–241. MR1841412

[3] BRIX, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. in Appl. Probab.* **31** 929–953. MR1747450

[4] CIFARELLI, D. M. and MELILLI, E. (2000). Some new results for Dirichlet priors. *Ann. Statist.* **28** 1390–1413. MR1805789

[5] CIFARELLI, D. M. and REGAZZINI, E. (1990). Distribution functions of means of a Dirichlet process. *Ann. Statist.* **18** 429–442. MR1041402

[6] DE BLASI, P. and HJORT, N. L. (2007). Bayesian survival analysis in proportional hazard models with logistic relative risk. *Scand. J. Statist.* **34** 229–257. MR2325252

[7] DE BLASI, P., PECCATI, G. and PRÜNSTER, I. (2007). Asymptotics for posterior hazards. Technical report. Submitted.

[8] DALEY, D. J. and VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes*. Springer, New York. MR950166

[9] DEHEUVELS, P., PECCATI, G. and YOR, M. (2006). On quadratic functionals of the Brownian sheet and related processes. *Stochastic Process. Appl.* **116** 493–538. MR2199561

[10] DOKSUM, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.* **2** 183–201. MR0373081

[11] DRĂGHICI, L. and RAMAMOORTHI, R. V. (2003). Consistency of Dykstra–Laud priors. *Sankhyā* **65** 464–481. MR2028910

[12] DYKSTRA, R. L. and LAUD, P. (1981). A Bayesian nonparametric approach to reliability. *Ann. Statist.* **9** 356–367. MR606619

[13] EPIFANI, I., LIJOI, A. and PRÜNSTER, I. (2003). Exponential functionals and means of neutral-to-the-right priors. *Biometrika* **90** 791–808. MR2024758

[14] FERGUSON, T. S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2** 615–629. MR0438568

[15] FERGUSON, T. S. and PHADIA, E. G. (1979). Bayesian nonparametric estimation based on censored data. *Ann. Statist.* **7** 163–186. MR515691

[16] GHOSH, J. K. and RAMAMOORTHI, R. V. (2003). *Bayesian Nonparametrics*. Springer, New York. MR1992245

[17] HJORT, N. L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.* **18** 1259–1294. MR1062708

[18] HO, M.-W. (2006). A Bayes method for a monotone hazard rate via **S**-paths. *Ann. Statist.* **34** 820–836. MR2283394

[19] ISHWARAN, H. and JAMES, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* **96** 161–173. MR1952729

[20] ISHWARAN, H. and JAMES, L. F. (2004). Computational methods for multiplicative intensity models using weighted gamma processes: Proportional hazards, marked point processes, and panel count data. *J. Amer. Statist. Assoc.* **99** 175–190. MR2054297

[21] JAMES, L. F. (2003). Bayesian calculus for gamma processes with applications to semiparametric intensity models. *Sankhyā* **65** 179–206. MR2016784

[22] JAMES, L. F. (2005). Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages. *Ann. Statist.* **33** 1771–1799. MR2166562

[23] JAMES, L. F. (2006). Poisson calculus for spatial neutral to the right processes. *Ann. Statist.* **34** 416–440. MR2275248

[24] KALBFLEISCH, J. D. (1978). Non-parametric Bayesian analysis of survival time data. *J. Roy. Statist. Soc. Ser. B* **40** 214–221. MR517442

[25] KIM, Y. (1999). Nonparametric Bayesian estimators for counting processes. *Ann. Statist.* **27** 562–588. MR1714717

[26] KIM, Y. and LEE, J. (2003). Bayesian analysis of proportional hazard models. *Ann. Statist.* **31** 493–511. MR1983539

[27] KINGMAN, J. F. C. (1967). Completely random measures. *Pacific J. Math.* **21** 59–78. MR0210185

[28] KINGMAN, J. F. C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. Oxford Univ. Press, New York. MR1207584

[29] KWAPIEŃ, S. and WOYCZYŃSKI, W. A. (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Boston, MA. MR1167198

[30] LIJOI, A., MENA, R. H. and PRÜNSTER, I. (2005). Hierarchical mixture modeling with normalized inverse-Gaussian priors. *J. Amer. Statist. Assoc.* **100** 1278–1291. MR2236441

[31] LO, A. Y. (1982). Bayesian nonparametric statistical inference for Poisson point processes. *Z. Wahrsch. Verw. Gebiete* **59** 55–66. MR643788

[32] LO, A. Y. and WENG, C.-S. (1989). On a class of Bayesian nonparametric estimates. II. Hazard rate estimates. *Ann. Inst. Statist. Math.* **41** 227–245. MR1006487

[33] NIETO-BARAJAS, L. E. and WALKER, S. G. (2004). Bayesian nonparametric survival analysis via Lévy driven Markov processes. *Statist. Sinica* **14** 1127–1146. MR2126344

[34] NIETO-BARAJAS, L. E. and WALKER, S. G. (2005). A semi-parametric Bayesian analysis of survival data based on Lévy-driven processes. *Lifetime Data Anal.* **11** 529–543. MR2213503

[35] NUALART, D. and VIVES, J. (1990). Anticipative calculus for the Poisson process based on the Fock space. In *Séminaire de Probabilités XXIV*, 1988/89. *Lecture Notes in Mathematics* **1426** 154–165. Springer, Berlin. MR1071538

[36] PECCATI, G. (2001). On the convergence of multiple random integrals. *Studia Sci. Math. Hungar.* **37** 429–470. MR1874695

[37] PECCATI, G. and TAQQU, M. S. (2007). Limit theorems for multiple stochastic integrals. Preprint. Available at http://geocities.com/giovannipeccati/.

[38] PECCATI, G. and TAQQU, M. S. (2007). Central limit theorems for double Poisson integrals. *Bernoulli*. To appear.

[39] PECCATI, G. and TAQQU, M. S. (2007). Stable convergence of generalized L^2 stochastic integrals and the principle of conditioning. *Electron. J. Probab.* **12** no. 15, 447–480 (electronic). MR2299924

[40] PECCATI, G. and TUDOR, C. A. (2005). Gaussian limits for vector-valued multiple stochastic integrals. In *Séminaire de Probabilités XXXVIII*. *Lecture Notes in Mathematics* **1857** 247–262. Springer, Berlin. MR2126978

[41] REGAZZINI, E., GUGLIELMI, A. and DI NUNNO, G. (2002). Theory and numerical analysis for exact distributions of functionals of a Dirichlet process. *Ann. Statist.* **30** 1376–1411. MR1936323

[42] REGAZZINI, E., LIJOI, A. and PRÜNSTER, I. (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31** 560–585. MR1983542

[43] ROTA, G.-C. and WALLSTROM, T. C. (1997). Stochastic integrals: A combinatorial approach. *Ann. Probab.* **25** 1257–1283. MR1457619

[44] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. (Translated from the Japanese.) MR1739520

[45] SURGAILIS, D. (1984). On multiple Poisson stochastic integrals and associated Markov semigroups. *Probab. Math. Statist.* **3** 217–239. MR764148

[46] SURGAILIS, D. (2003). Non-CLTs: U-statistics, multinomial formula and approximations of multiple Itô–Wiener integrals. In *Theory and Applications of Long-Range Dependence* (P. Doukhan, G. Oppenheim and M.S. Taqqu, eds.) 129–142. Birkhäuser, Boston, MA. MR1956047

[47] WALKER, S. and DAMIEN, P. (1998). A full Bayesian non-parametric analysis involving a neutral to the right process. *Scand. J. Statist.* **25** 669–680. MR1666808

[48] WALKER, S. and MULIERE, P. (1997). Beta-Stacy processes and a generalization of the Pólya-urn scheme. *Ann. Statist.* **25** 1762–1780. MR1463574

[49] WOLPERT, R. L. and ICKSTADT, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika* **85** 251–267. MR1649114

[50] WOLPERT, R. L. and TAQQU, R. L. (2005). Fractional Ornstein–Uhlenbeck Lévy processes and the Telecom process. *Signal Process.* **85** 1523–1545.

LABORATOIRE DE STATISTIQUE
THÉORIQUE ET APPLIQUÉE
UNIVERSITÉ PARIS VI
175 RUE DU CHEVALERET
75013 PARIS
FRANCE
E-MAIL: giovanni.peccati@gmail.com

DIPARTIMENTO DI STATISTICA
E MATEMATICA APPLICATA
UNIVERSITÀ DEGLI STUDI DI TORINO
PIAZZA ARBARELLO 8
10122 TORINO
ITALY
E-MAIL: igor@econ.unito.it