



HAL
open science

Stein's method for diffusive limit of queueing processes

Eustache Besançon, Laurent Decreusefond, Pascal Moyal

► **To cite this version:**

Eustache Besançon, Laurent Decreusefond, Pascal Moyal. Stein's method for diffusive limit of queueing processes. Queueing Systems, In press. hal-01784139v2

HAL Id: hal-01784139

<https://hal.science/hal-01784139v2>

Submitted on 13 Jun 2018 (v2), last revised 2 May 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

STEIN'S METHOD FOR DIFFUSIVE LIMITS OF MARKOV PROCESSES

EUSTACHE BESANÇON, LAURENT DECREUSEFOND, AND PASCAL MOYAL

ABSTRACT. Donsker Theorem is perhaps the most famous invariance principle result for Markov processes. It states that when properly normalized, a random walk behaves asymptotically like a Brownian motion. This approach can be extended to general Markov processes whose driving parameters are taken to their limits, which can lead to insightful results in contexts like large distributed systems or queueing networks. The purpose of this paper is to assess the rate of convergence in these so-called diffusion approximations. To this end, we extend the functional Stein method introduced for the Brownian approximation of Poisson processes to two simple examples of queueing models. By doing so, we complete the recent applications of Stein's method in the queueing context, with results concerning the whole trajectory of the considered process, instead of its stationary distribution.

1. MOTIVATIONS

It is usually said that the Central Limit Theorem (CLT for short) indicates that the rate of convergence in the Law of Large Numbers is of the order of $1/\sqrt{n}$ because one can informally write

$$\frac{1}{n} \sum_{j=1}^n X_j \simeq \mathbf{E}[X_1] + \frac{1}{\sqrt{n}} \mathcal{N}(0, 1).$$

Going further, one can then investigate the rate of convergence in the CLT itself. For this, one needs to define a suitable notion of distance between laws of random variables (r.v.'s for short). In the present context, we will be led to use an alternative distance to the classical Prohorov one, inducing the same topology whenever the state-space E of the considered r.v.'s is a separable metric space. Specifically, denote by $\text{Lip}_1(E)$, the set of 1-Lipschitz continuous functions from E to the real line \mathbb{R} . The so-called Kolmogorov-Rubinstein distance (sometimes called Wasserstein distance) between two probability measures \mathbf{P} and \mathbf{Q} on E is then defined by

$$d_{\text{KR}}(\mathbf{P}, \mathbf{Q}) = \sup_{F \in \text{Lip}_1(E)} \left(\int_E F \, d\mathbf{P} - \int_E F \, d\mathbf{Q} \right).$$

In Theorem 11.3.3 of [17], it is shown that

$$\left(d_{\text{KR}}(\mathbf{P}, \mathbf{Q}_n) \xrightarrow{n \rightarrow \infty} 0 \right) \iff \left(\mathbf{Q}_n \text{ converges in dist. to } \mathbf{P} \right).$$

This formulation is particularly well suited to be estimated via Stein's method (SM for short).

The SM was first introduced in an article by Stein [29] to quantify the rate of convergence in the Central Limit Theorem and was soon extended to the Poisson distribution by Chen [10]. In its first step, the method involves characterizing the target distribution with a functional operator A such that $\mathbf{E}_{\mathbf{Q}}[AF] = 0$ for any F in a sufficiently large set of test functions \mathcal{F} , if and only if $\mathbf{Q} = \mathbf{P}$. Barbour [2] then introduced the generator interpretation that made possible the extension of Stein's

method to many other probability distributions: in many cases this functional operator A can be viewed as the infinitesimal generator of a Markovian semi-group $(P_t, t \geq 0)$ whose stationary measure is \mathbf{P} . This means that we can write

$$(1) \quad d_{\mathcal{F}}(\mathbf{P}, \mathbf{Q}) := \sup_{F \in \mathcal{F}} \left| \mathbf{E}_{\mathbf{P}} [F] - \mathbf{E}_{\mathbf{Q}} [F] \right| = \sup_{F \in \mathcal{F}} \left| \mathbf{E}_{\mathbf{Q}} \left[\int_0^{\infty} AP_t F \, dt \right] \right|.$$

If we choose \mathcal{F} to be Lip_1 , we have an interesting representation of the Kolmogorov-Rubinstein between \mathbf{P} and \mathbf{Q} . The function

$$x \in E \mapsto \int_0^{\infty} AP_t F(x) \, dt$$

is one possible expression of the solution of the so-called Stein equation. If \mathbf{P} is the Gaussian distribution on \mathbb{R} , then $(P_t, t \geq 0)$ is known as the Ornstein-Uhlenbeck semi-group whose regularizing properties induce that the solution of the Stein equation has bounded first and second order derivatives. This observation is the first step of the numerous papers on the SM (see [3] and references therein).

A very important breakthrough was made by Nourdin and Peccati [21] who showed that alternatively, the right-hand-side of (1) could be transformed and amenable to further simplifications, by using integration by parts in the Malliavin calculus sense. This was the starting point of a bunch of articles with a wide area of applications: Berry-Esseen theorem, iterated-logarithm theorem (see [20] and references therein), limit theorems on manifolds, Poisson approximation [23], etc. As a result of these almost fifty years of intense activity, a huge number of Gaussian or Poisson convergence results have been quantified. Closer to the class of models we have in mind, let us mention the fruitful recent applications of the SM, to assess the rate of convergence of the stationary distributions of various processes involved in queueing: Erlang-A and Erlang-C systems in [7]; a system with reneging and phase-type service time distributions (in which case the target distribution is the stationary distribution of a piecewise OU process) in [8].

The literature is much more restricted when it comes to cases where the limiting distribution is that of a whole stochastic process, like a Brownian motion or a (marked) Poisson point process. The first work in that direction is due to Barbour [2] which established the first quantified version of Donsker Theorem, resorting to ideas closely related to Malliavin calculus. In [11], a different technique was used to estimate the convergence rate of the normalized Poisson process to the Brownian motion. The paper of Shih [27] extends the original approach of the SM in abstract Wiener spaces. Besides the technical points which are evidently more involved, the main difference between convergence to random variables and convergence to random processes is that for the latter, we generally have a large choice of functional spaces of reference. For instance, a Brownian motion can be seen either as a square integrable process, as a continuous process, as an α -Hölder continuous process for any $\alpha < 1/2$, or even as an element of a fractional Sobolev spaces as defined below. Changing the topology modifies the admissible test functions: The evaluation of the trajectory at time t_0 is Lipschitz continuous on a Hölder space but it is not defined on the space of square integrable functions. Moreover, as already seen in [11], the convergence rate may also depend on the chosen space.

So far, the trajectorial version of Stein's method has been applied to estimate the convergence rate of explicit processes towards the Brownian motion or Poisson point processes [16]. This does not represent all kinds of situations where we know that a sequence of processes converges to a diffusion process. Here, we have in mind the vast literature on diffusion approximations, allowing to efficiently simulate an asymptotic version of the process under consideration, or assess the order of the fluctuations around its fluid limit or its mean field, along the various applications.

The most basic example is that of the M/M/1 queue with initial condition nx_0 , arrival rate λn and service rate μn . It is well known (see e.g. Section 5.7 in [25]) that if L_n^\dagger denotes the process which counts the customers in the system, then

$$(2) \quad Z_n^\dagger = \frac{\sqrt{n}}{\lambda + \mu} \left(\frac{L_n^\dagger}{n} - \bar{L}^\dagger \right) \Longrightarrow B,$$

on the time interval $[0, x_0(\mu - \lambda)]$, where " \Longrightarrow " denotes the weak convergence of processes, B is a standard Brownian motion and \bar{L}^\dagger is the solution of the equation

$$\bar{L}^\dagger(t) = x_0 + \lambda t - \mu \int_0^t \mathbf{1}_{\{\bar{L}^\dagger(s^-) > 0\}} ds, \quad t \geq 0.$$

The convergence holds in distribution over \mathbb{D} , the Skorohod space of continuous-on-the-right-with-left-limits functions. The principle of the proof is to show that the sequence of processes $(Z_n^\dagger : n \geq 1)$ is tight in the convenient topology and that the finite dimensional distributions of Z_n^\dagger converge to that of the Brownian motion. One approach is to view L_n^\dagger as the solution of a stochastic differential equation driven by a finite number of independent Poisson processes on the real line:

$$(3) \quad L_n^\dagger(t) = x_0 + N_{\lambda n}(t) - \int_0^t \mathbf{1}_{\{L_n^\dagger(s^-) > 0\}} dN_{\mu n}(s),$$

where for any $\alpha > 0$, N_α denote a Poisson process on \mathbb{R}^+ of intensity α . This yields an implicit definition of L_n^\dagger which, using martingale convergence, is sufficient to prove the tightness of $(Z_n^\dagger : n \geq 1)$. However, due to the reflection term in (3) the process L_n^\dagger is not well suited to a straightforward development of the SM. But it is easily seen that until the first time τ_n^\dagger when it reaches 0, L_n^\dagger is the difference of two independent Poisson processes. We can thus proceed in two steps: deducing from a classical large deviation result, the existence of a strictly positive horizon T such that the probability of $\{\tau_n^\dagger \leq T\}$ decays exponentially, and on the complement, adapting the results in [11] on the convergence rate of (linear combinations of) conveniently normalized independent Poisson processes to a Brownian motion. This an interesting and original extension of the SM.

Another canonical example of a diffusion limit arising in queueing is the rescaled M/M/ ∞ queue, whose limiting process is an Ornstein-Uhlenbeck process (see [4] and Section 6.6 in [25], or [14] regarding non-exponential service times), for which we do not know *a priori* a characterizing operator as in Eqn. (1). To circumvent this difficulty we first apply an integral transformation to the process counting the number of busy servers, so that we are reduced to prove the convergence to a time-changed Brownian motion. Then, via a representation by a Marked Poisson process and an interpolation, it is sufficient to estimate the convergence rate in a finite dimensional functional space.

The two models thus involve very different techniques, therefore a generalization of our results can be envisioned only on a case by case basis. Although it is frustrating, this fact is not so surprising, as it is reminiscent of most diffusion approximation results themselves which, within a well establish general procedure (tightness and convergence of fidi distributions, martingale convergence), often also resort in detail to ad-hoc arguments.

The paper is organized as follows. In Section 2, we explain the functional framework and introduce the Malliavin calculus for Brownian motion and marked Poisson point process. In Section 3, we show that the distance we aim to compute can be split into three parts, each one we handle differently. In particular, in subsection 3.3, we develop our approach of the Stein method. It is an extension to the functional setting of [23, Theorem 3.1]. In Section 4, we apply the previous results to the M/M/1 and to the M/M/ ∞ queue in Section 5.

2. PRELIMINARIES

2.1. Functional spaces. We need to introduce several spaces of functions. Throughout the whole paper, we fix a time horizon $T > 0$.

Definition 2.1. *The Skorohod space $\mathbb{D}([0, T])$ is the space of right continuous with left limits (rcll) functions from $[0, T]$ into \mathbb{R} . It is usually equipped with the distance*

$$(4) \quad d_{\mathbb{D}}(f, g) = \inf_{\phi \in \text{Hom}_T} \left(\max(\|Id - \phi\|_{L^\infty([0, T])}, \|f - g \circ \phi\|_{L^\infty([0, T])}) \right)$$

where Hom_T is the space of increasing homeomorphisms from $[0, T]$ into itself.

It contains \mathcal{C} , the space of continuous functions on $[0, T]$, as well as \mathcal{E} , the set of step-wise functions. In \mathcal{C} , it is interesting to focus on the Hölder continuous functions: $f \in \text{Hol}(\eta)$ whenever

$$\|f\|_{\text{Hol}(\eta)} = \sup_{s \neq t \in [0, T]} \frac{|f(t) - f(s)|}{|t - s|^\eta} < \infty.$$

As in [13, 19], we consider the fractional Sobolev spaces $W_{\eta, p}$ defined for $\eta \in (0, 1]$ and $p \geq 1$ as the closure of \mathcal{C}^1 functions with respect to the norm

$$\|f\|_{\eta, p}^p = \int_0^T |f(t)|^p dt + \iint_{[0, T]^2} \frac{|f(t) - f(s)|^p}{|t - s|^{1+p\eta}} dt ds.$$

For $\eta = 1$, $W_{1, p}$ is the completion of \mathcal{C}^1 for the norm:

$$\|f\|_{1, p}^p = \int_0^T |f(t)|^p dt + \int_0^T |f'(t)|^p dt.$$

They are known to be Banach spaces and to satisfy the Sobolev embeddings [1, 18]:

$$W_{\eta, p} \subset \text{Hol}(\eta - 1/p) \text{ for } \eta - 1/p > 0$$

and

$$W_{\eta, p} \subset W_{\alpha, q} \text{ for } 1 \geq \eta \geq \alpha \text{ and } \eta - 1/p \geq \alpha - 1/q.$$

As a consequence, since $W_{1, p}$ is separable (see [6]), so does $W_{\eta, p}$. We need to compute the $W_{\eta, p}$ norm of primitive of step functions.

Lemma 2.2. *Let $0 \leq s_1 < s_2 \leq T$ and consider*

$$h_{s_1, s_2}(t) = \int_0^t \mathbf{1}_{[s_1, s_2]}(r) dr.$$

There exists $c > 0$ such that for any s_1, s_2 , we have

$$(5) \quad \|h_{s_1, s_2}\|_{W_{\eta, p}} \leq c |s_2 - s_1|^{1-\eta}.$$

Proof. Remark that for any $s, t \in [0, T]$,

$$|h_{s_1, s_2}(t) - h_{s_1, s_2}(s)| \leq |t - s| \wedge (s_2 - s_1).$$

The result then follows from the definition of the $W_{\eta, p}$ norm. \square

We also need to introduce the Besov-Liouville spaces of fractional derivatives. For $f \in L^1([0, T]; dt)$, (denoted by L^1 for short) the left and right fractional integrals of f are defined by :

$$(I_{0+}^\alpha f)(x) = \frac{1}{\Gamma(\alpha)} \int_0^x f(t)(x-t)^{\alpha-1} dt, \quad x \geq 0,$$

$$(I_{T-}^\alpha f)(x) = \frac{1}{\Gamma(\alpha)} \int_x^T f(t)(t-x)^{\alpha-1} dt, \quad x \leq 1,$$

where $\alpha > 0$ and $I_{0+}^0 = I_{T-}^0 = \text{Id}$. For any $\alpha \geq 0$, $p, q \geq 1$, any $f \in L^p$ and $g \in L^q$ where $p^{-1} + q^{-1} \leq \alpha$, we have :

$$(6) \quad \int_0^T f(s)(I_{0+}^\alpha g)(s) ds = \int_0^T (I_{T-}^\alpha f)(s)g(s) ds.$$

The Besov-Liouville space $I_{0+}^\alpha(L^p) := \mathcal{I}_{\alpha,p}^+$ is usually equipped with the norm :

$$(7) \quad \|I_{0+}^\alpha f\|_{\mathcal{I}_{\alpha,p}^+} = \|f\|_{L^p}.$$

Analogously, the Besov-Liouville space $I_{T-}^\alpha(L^p) := \mathcal{I}_{\alpha,p}^-$ is usually equipped with the norm :

$$\|I_{T-}^\alpha f\|_{\mathcal{I}_{\alpha,p}^-} = \|f\|_{L^p}.$$

These spaces are particularly interesting in view of their relationships with the spaces of Hölder continuous functions.

Theorem 2.3 (Sobolev embeddings, [18, 26]). *We have the following embedding properties.*

- (1) If $0 < \alpha < 1$, $1 < p < 1/\alpha$, then I_{0+}^α is a bounded operator from L^p into L^q with $q = p(1 - \alpha p)^{-1}$.
- (2) For any $0 < \alpha < 1$ and any $p \geq 1$, $\mathcal{I}_{\alpha,p}^+$ is continuously embedded in $\text{Hol}(\alpha - 1/p)$ provided that $\alpha - 1/p > 0$.
- (3) For any $\alpha' \geq \alpha$ and p, p' such that $\alpha' - 1/p' \geq \alpha - 1/p$, $\mathcal{I}_{\alpha',p'} \subset \mathcal{I}_{\alpha,p}$.
- (4) For $1 \geq \alpha > \eta > \zeta > 0$, the following embeddings are continuous (even compact)

$$W_{\alpha,p} \subset \mathcal{I}_{\eta,p}^+ \subset W_{\zeta,p}.$$

It may be useful to keep in mind the following diagram where all arrows represent continuous embeddings. For any $\eta > 0$, any $\epsilon \in (0, \eta)$,

$$\begin{array}{ccccccc} W_{\eta,1/(\eta-\epsilon)} & \longrightarrow & \text{Hol}_0(\epsilon) & \longrightarrow & \mathcal{C}_0 & \longrightarrow & \mathcal{C} & \longrightarrow & \mathbb{D} \\ & & & & & & \downarrow & & \uparrow \\ & & & & & & W_{\eta,1/(\eta+\epsilon)} & \longleftarrow & \mathcal{E} & \longleftarrow & \mathcal{E}_0 \end{array}$$

where \mathbb{D}_0 (respectively \mathcal{C}_0 , Hol_0 , \mathcal{E}_0) represents the elements of \mathbb{D} (respectively \mathcal{C} , Hol , \mathcal{E}) which are null at time 0.

2.2. Wiener space. Since we want to compare some measure with the distribution of the Brownian motion, sometimes called the Wiener measure, we need to construct precisely the functional framework. We refer to [22, 24, 30] for details about Malliavin calculus in the Gaussian setting.

Let $B = (B(t), t \in [0, T])$ be a standard one-dimensional Brownian motion. Since it has Hölder continuous sample-paths of any order less than 1/2, we can say that almost-surely, B belongs to $\mathcal{I}_{\eta,p}^+$ for any

$$(\eta, p) \in \Lambda = \{(\eta, p) \in \mathbb{R}^+ \times \mathbb{R}^+, 0 < \eta - 1/p < 1/2\}.$$

We denote by $\mathbf{P}_{\eta,p}$, the distribution of B over $\mathcal{I}_{\eta,p}^+$. The spaces $\mathcal{I}_{\eta,p}^+$ are Banach spaces, for which there exists the notion of Fréchet derivative. For $F : \mathcal{I}_{\eta,p}^+ \rightarrow \mathbb{R}$, it is differentiable whenever

$$(8) \quad \lim_{\epsilon \rightarrow 0} \epsilon^{-1} (F(x + \epsilon h) - F(x))$$

exists for any $h \in \mathcal{I}_{\eta,p}^+$ and defines an element of $(\mathcal{I}_{\eta,p}^+)^*$,

$$\lim_{\epsilon \rightarrow 0} \epsilon^{-1} (F(x + \epsilon h) - F(x)) = \langle DF(x), h \rangle_{(\mathcal{I}_{\eta,p}^+)^*, \mathcal{I}_{\eta,p}^+}.$$

In particular, as in finite dimension, Fréchet differentiability implies continuity. In the present context, the functions we are going to consider are random variables,

meaning that they are defined up to a negligible set, so that no hypothesis of continuity can be enforced. Moreover, as Eqn. (8) suggests, if $F = G$ almost-surely, we must be sure that $F(\cdot + h) = G(\cdot + h)$ almost-surely for any $h \in W_{\eta,p}$, i.e. the push-forward measure of $\mathbf{P}_{\eta,p}$ by the map $\tau_h : x \mapsto x + h$ is absolutely continuous with respect to $\mathbf{P}_{\eta,p}$. For this property to hold, the Cameron-Martin theorem says that we must restrict the perturbation h to belong to $\mathcal{I}_{1,2}^+$.

Theorem 2.4 (Cameron-Martin). *For any $h \in \mathcal{I}_{1,2}^+$, for any bounded functional $F : W_{\eta,p} \rightarrow \mathbb{R}$*

$$(9) \quad \mathbf{E}[F(B+h)] = \mathbf{E} \left[F(B) \exp \left(\int_0^T \dot{h}(s) dB(s) - \frac{1}{2} \|h\|_{\mathcal{I}_{1,2}^+}^2 \right) \right],$$

where \dot{h} is the time derivative of $h \in \mathcal{I}_{1,2}^+$, so that \dot{h} belongs to $L^2([0,T])$ and the stochastic integral has to be taken in the Itô sense.

Otherwise stated, Eqn. (9) means that the distribution of $B+h$ is absolutely continuous with respect to $\mathbf{P}_{\eta,p}$ and that its Radon-Nykodim derivative is given by the exponential factor of the RHS of (9).

Because of this theorem, the space $\mathcal{I}_{1,2}^+$ plays a crucial role in the stochastic calculus of variations. We have the following scheme

$$\begin{array}{ccc} W_{\eta,p}^* & \xrightarrow{i_{\eta,p}^*} & (\mathcal{I}_{1,2}^+)^* \\ & & \downarrow \simeq \\ L^2 & \xrightarrow{I_{0+}^1} & \mathcal{I}_{1,2}^+ \xrightarrow{i_{\eta,p}} W_{\eta,p} \end{array}$$

The map $i_{\eta,p}$ is the embedding from $\mathcal{I}_{1,2}^+$ into $W_{\eta,p}$. The pivotal space, i.e. the Hilbert space identified to itself, is, in this context, the space $\mathcal{I}_{1,2}^+$ and not L^2 as it often happens. This means that $i_{\eta,p}^*$ is the adjoint of $i_{\eta,p}$ after this identification.

We can now introduce the concept of Gross-Sobolev or weak derivative.

Definition 2.5. *A function $F : W_{\eta,p} \rightarrow \mathbb{R}$ is said to be cylindrical if it is of the form*

$$(10) \quad F = f(\delta_B h_1, \dots, \delta_B h_k)$$

where f belongs to the Schwartz space on \mathbb{R}^k , h_1, \dots, h_k belong to $\mathcal{I}_{1,2}^+$ and $\delta_B h$ is the Itô integral of \dot{h} :

$$\delta_B h = \int_0^1 \dot{h}(s) dB(s).$$

Remark that if \dot{u} belongs to L^2 , then

$$\begin{aligned} \int_0^s \dot{u}(s) d(B+h)(s) &= \int_0^s \dot{u}(s) dB(s) + \int_0^s \dot{u}(s) \dot{h}(s) ds \\ &= \int_0^s \dot{u}(s) dB(s) + \langle \dot{u}, \dot{h} \rangle_{L^2} = \int_0^s \dot{u}(s) dB(s) + \langle u, h \rangle_{\mathcal{I}_{1,2}^+}. \end{aligned}$$

Hence, if F is cylindrical

$$\left. \frac{d}{d\varepsilon} F(B + \varepsilon h) \right|_{\varepsilon=0} = \sum_{j=1}^k \partial_j f(\delta_B h_1, \dots, \delta_B h_k) \langle h_j, h \rangle_{\mathcal{I}_{1,2}^+}$$

This motivates the following definition

Definition 2.6. For F as in (10), let ∇F be the element of $L^2(W_{\eta,p}; \mathcal{I}_{1,2}^+)$ defined by

$$\nabla F = \sum_{j=1}^k \partial_j f(\delta_B h_1, \dots, \delta_B h_k) h_j$$

and let $\nabla^{(2)} F$ be the element of $L^2(W_{\eta,p}; \mathcal{I}_{1,2}^+ \otimes \mathcal{I}_{1,2}^+)$

$$\nabla^{(2)} F = \sum_{j,l=1}^k \partial_{jl}^{(2)} f(\delta_B h_1, \dots, \delta_B h_k) h_j \otimes h_l.$$

Consider the norm

$$\|F\|_{2,2}^2 = \|F\|_{L^2}^2 + \mathbf{E} \left[\|\nabla F\|_{\mathcal{I}_{1,2}^+}^2 \right] + \mathbf{E} \left[\|\nabla^{(2)} F\|_{\mathcal{I}_{1,2}^+ \otimes \mathcal{I}_{1,2}^+}^2 \right],$$

where

$$\|\nabla F\|_{\mathcal{I}_{1,2}^+}^2 = \int_0^1 \left(\sum_{j=1}^k \partial_j f(\delta_B h_1, \dots, \delta_B h_k) \dot{h}_j(s) \right)^2 ds$$

and

$$\|\nabla^{(2)} F\|_{(\mathcal{I}_{1,2}^+)^{\otimes 2}}^2 = \int_0^1 \int_0^1 \left(\sum_{j,l=1}^k \partial_{jl}^{(2)} f(\delta_B h_1, \dots, \delta_B h_k) \dot{h}_j(s) \dot{h}_l(r) \right)^2 ds dr.$$

The set $\mathbb{D}_{2,2}$ is the completion of the set of cylindrical functions with respect to the norm $\|\cdot\|_{2,2}$.

Remark 1. Note that if h belongs to $\mathcal{I}_{2,2}^\pm = (I_{0+}^1 \circ I_{1-}^1)(L^2) \subset \mathcal{I}_{1,2}^+$ then

$$\nabla f(\delta_B h) = f'(\delta_B h) h$$

belongs to $L^2(W_{\eta,p}; \mathcal{I}_{2,2}^\pm)$. This means that for such a functional, its gradient is more regular, in the sense that it belongs to a smaller space, than for ordinary elements of $\mathbb{D}_{2,2}$.

Since we identified $\mathcal{I}_{1,2}^+$ with its dual, the space $\mathcal{I}_{2,2}^\pm$ is in duality with L^2 : For $h \in \mathcal{I}_{2,2}^\pm$, there exists $\ddot{h} \in L^2$ such that $h = I_{0+}^1(I_{1-}^1(\ddot{h}))$. Hence for $k \in \mathcal{I}_{1,2}^+$, we have

$$\begin{aligned} \left| \langle h, k \rangle_{\mathcal{I}_{1,2}^+} \right| &= \left| \int_0^1 I_{1-}^1(\ddot{h})(s) k(s) ds \right| = \left| \int_0^1 \ddot{h}(s) I_{0+}^1(k)(s) ds \right| \\ &= \left| \int_0^1 \ddot{h}(s) k(s) ds \right| \leq \|\ddot{h}\|_{L^2} \|k\|_{L^2}. \end{aligned}$$

Since $\mathcal{I}_{1,2}^+$ is dense in L^2 , we can extend this duality pairing to $h \in \mathcal{I}_{2,2}^\pm$ and $k \in L^2$.

This leads to the following definition.

Definition 2.7. A function $F : W_{\eta,p} \rightarrow \mathbb{R}$ is said to belong to the class $\Sigma_{\eta,p}$ whenever it belongs to $\text{Lip}_1(W_{\eta,p})$, belongs to $\mathbb{D}_{2,2}$ and satisfies

$$(11) \quad \left| \left\langle \nabla^{(2)} F(x) - \nabla^{(2)} F(x+g), h \otimes k \right\rangle_{\mathcal{I}_{1,2}^+} \right| \leq \|g\|_{W_{\eta,p}} \|h\|_{L^2} \|k\|_{L^2},$$

for any $x \in W_{\eta,p}$, $g \in \mathcal{I}_{1,2}^+$, $h, k \in L^2$. This means that $\nabla^{(2)} F$ is an element of the space $\text{Lip}_1(W_{\eta,p}; (\mathcal{I}_{2,2}^\pm)^{\otimes 2})$.

If $F : W_{\eta,p} \rightarrow \mathbb{R}$ is thrice differentiable in the direction of $\mathcal{I}_{1,2}^+$ and such that

$$\sup_{x \in W_{\eta,p}} \|\nabla^{(3)} F\|_{(\mathcal{I}_{2,2}^\pm)^{\otimes 3}} < \infty$$

then by the fundamental theorem of calculus

$$\begin{aligned} & \left| \left\langle \nabla^{(2)} F(x) - \nabla^{(2)} F(x+g), h \otimes k \right\rangle_{\mathcal{I}_{1,2}^+} \right| \\ & \leq \|\nabla^{(3)} F\|_{L^\infty(W_{\eta,p}; (\mathcal{I}_{2,2}^\pm)^{\otimes 3})} \|g\|_{L^2} \|h\|_{L^2} \|k\|_{L^2}. \end{aligned}$$

Since $W_{\eta,p}$ is continuously embedded in L^2 ,

$$\|\nabla^{(3)} F\|_{L^\infty(W_{\eta,p}; (\mathcal{I}_{2,2}^\pm)^{\otimes 3})}^{-1} F \in \Sigma_{\eta,p}.$$

Our main results below will be more easily expressed for test functions in the following set,

$$(12) \quad \mathcal{L}_{\eta,p} = \{\text{bounded functions of } \text{Lip}_1(\mathbb{D}) \text{ whose restriction to } W_{\eta,p} \text{ belongs to } \Sigma_{\eta,p}\}.$$

Let us also define the following distances and norm: for any two processes U and V in the convenient spaces,

$$(13) \quad d_{\Sigma_{\eta,p}}(U, V) = \sup_{F \in \Sigma_{\eta,p}} |\mathbf{E}[F(U)] - \mathbf{E}[F(V)]|;$$

$$(14) \quad d_{\mathcal{L}_{\eta,p}}(U, V) = \sup_{F \in \mathcal{L}_{\eta,p}} |\mathbf{E}[F(U)] - \mathbf{E}[F(V)]|;$$

$$(15) \quad \|V - U\|_{\infty, T} = \sup_{t \in [0, T]} |V(t) - U(t)| \quad \text{a.s..}$$

2.3. Poisson point process. We now introduce the minimum framework to get an integration by parts for Poisson point processes. For details, we refer to [15, 24] Let E be a complete and separable metric space equipped with a σ -finite measure ν . Let \mathfrak{N}_E be the space of locally finite configurations on E , i.e. the set of at most denumerable subsets of E with no accumulation point. Such a set ϕ can be described as a set or as a sum of atomic measures:

$$\phi \simeq \sum_{x \in \phi} \varepsilon_x,$$

where ε_x is the Dirac measure at x , so that for any $\psi : E \rightarrow \mathbb{R}$,

$$\int_E \psi \, d\phi = \sum_{x \in \phi} \psi(x).$$

For ν a σ -finite measure on E , a Poisson point process of control measure ν is an \mathfrak{N}_E -valued random variable, say N_ν , such that for any $\psi : E \rightarrow \mathbb{R}$, with compact support,

$$\mathbf{E} \left[\exp \left(- \sum_{x \in N_\nu} \psi(x) \right) \right] = \exp \left(- \int_E 1 - e^{-\psi(x)} \, d\nu(x) \right).$$

The multivariate Campbell-Mecke formula states that for any integer $k \geq 1$, for any non-negative $F : E^k \times \mathfrak{N}_E$,

$$(16) \quad \mathbf{E} \left[\sum_{x_1, \dots, x_k \in N_\nu^\#} F(x_1, \dots, x_k, N_\nu) \right] = \int_{E^k} \mathbf{E} \left[F \left(x_1, \dots, x_k, N_\nu + \sum_{j=1}^k \varepsilon_{x_j} \right) \right] \otimes_{j=1}^k d\nu(x_j),$$

where the sum in the left-hand-side runs through the k -uples of distinct points of the configuration N_ν . We say that $F : E \rightarrow \mathbb{R}$ belongs to $\text{dom } D$ whenever

$$\mathbf{E} \left[\int_E \left(F(N_\nu + \varepsilon_x) - F(N_\nu) \right)^2 d\nu(x) \right] < \infty$$

and we set, for any $x \in E$,

$$D_x F(N_\nu) = F(N_\nu + \varepsilon_x) - F(N_\nu - \varepsilon_x),$$

where $N_\nu - \varepsilon_x$ is to be understood as N_ν whenever $x \notin N_\nu$. Let

$$\text{dom } \delta_\nu = \left\{ u : \mathfrak{N}_E \times E \rightarrow \mathbb{R}, \mathbf{E} \left[\int_E |u(N_\nu, x)|^2 d\nu(x) \right] < \infty \right\}.$$

Then, for $F \in \text{dom } D$ and $u \in \text{dom } \delta_\nu$, the multivariate Campbell-Mecke formula entails that

$$(17) \quad \mathbf{E} [F \delta_\nu u] = \mathbf{E} \left[\int_E DF(x) u(N_\nu, x) d\nu(x) \right],$$

where

$$\delta_\nu u = \int_E u(N_\nu - \varepsilon_x, x) dN_\nu(x) - \int_E u(N_\nu, x) d\nu(x).$$

Note that if u is deterministic,

$$(18) \quad \delta_\nu u = \int_E u(x) (dN_\nu(x) - d\nu(x)) \text{ and } D_x \delta_\nu u = u(x).$$

Moreover,

$$(19) \quad \mathbf{E} [\delta_\nu u] = 0 \text{ and } \mathbf{E} \left[(\delta_\nu u)^2 \right] = \int_E |u(x)|^2 d\nu(x).$$

3. DISTANCES BETWEEN PROBABILITY DISTRIBUTIONS

3.1. Distances on functional spaces. Two measures are comparable only if they are supported on the same space. For real or multivariate random variables, their distribution is canonically supported either by \mathbb{N} , \mathbb{R} or \mathbb{R}^n , etc. When dealing with functional spaces, a given process can naturally belong to several metric spaces. The sample-paths of continuous time Markov chains are piecewise constant, thus (see [11]) belong to $W_{\eta,p}$ for any (η, p) such that $\eta - 1/p < 0$ and also to \mathbb{D} . On the other hand, trajectories of diffusion processes belong to $W_{\eta,p}$ for $\eta - 1/p < 1/2$. This means that two factors may contribute to the distance between the distribution of a CTMC and that of a diffusion process: the difference between the dynamics and the gap of regularity; the latter being in some sense of lesser importance for the probabilist. In an effort to see the importance of each of these terms, we consider an intermediate process which has at least the regularity of the diffusion and a stochastic behavior similar to that of the stochastic process under study.

For $f \in \mathbb{D}$, we consider its affine interpolation on $[0, T]$ of mesh T/n :

$$(20) \quad \pi_n f : [0, T] \longrightarrow \mathbb{R}$$

$$t \longmapsto \sum_{i=0}^{n-1} \frac{n}{T} \left(f \left(\frac{(i+1)T}{n} \right) - f \left(\frac{iT}{n} \right) \right) I_{0+}^1 \left(\mathbf{1}_{\left[\frac{iT}{n}, \frac{(i+1)T}{n} \right)} \right) (t).$$

Remark that a stochastic process and its affine interpolation have similar dynamics since they coincide at each point of the subdivision, whose mesh tends to 0. In these conditions, it is meaningful to evaluate the distance between the distribution of some stochastic process L_n and the distribution of its affine interpolation, the distance between $\pi_n L_n$ and $\pi_n B$ where B is a standard Brownian motion, and then the distance between $\pi_n B$ and B . In view of what we said earlier, these distances are evaluated for the topology of the fractional Sobolev spaces $W_{\eta,p}$ for $\eta - 1/p < 0$,

$\eta - 1/p < 1$ and $\eta - 1/2 < 0$ respectively. In the end, by the triangular inequality we get the distance between the distribution of L_n and the Gaussian measure in the smallest spaces $W_{\eta,p}$ for $\eta - 1/p < 0$ but in passing we obtain some insights on the different factors which lead to the discrepancy. The distance between L_n and $\pi_n L_n$ and the distance between $\pi_n B$ and B are due to the gap of regularity between the sample-paths, whereas the difference between the laws of $\pi_n L_n$ and $\pi_n B$ is due to the dissimilarity of their stochastic behavior.

3.2. Distance between sample-paths. In what follows, \mathbb{N} is the set of positive integers. As mentioned above, we need to estimate the distance between the distribution of L_n and of $\pi_n L_n$ and the distance between the laws of $\pi_n B$ and of B .

Regarding the latter, since $\pi_n B$ and B are defined on the same probability space, we can resort to a more precise result, Proposition 13.20 in [19], claiming that for some c ,

$$\mathbf{E} [\|\pi_n B - B\|_{W_{\eta,p}}] \leq c n^{-(1/2-\eta)}, \quad n \in \mathbb{N},$$

for any $\eta < 1/2$. Then, recalling (13) we immediately get

$$(21) \quad \mathbf{E} [d_{\Sigma_{\eta,p}}(\pi_n B, B)] \leq c n^{-(1/2-\eta)}, \quad n \in \mathbb{N}.$$

We can also estimate the distance between the sample-paths of Birth-and-death processes and their interpolation. Specifically,

Lemma 3.1. *Let $n \in \mathbb{N}$, and let X be a \mathbb{N} -valued Markov jump process on $[0, T]$ of infinitesimal generator \mathcal{A} . Suppose that there exists two constants $J \in \mathbb{N}$ and $\alpha > 0$ such that*

- *the amplitude of jumps of X is bounded by $J > 0$, i.e. for all $i, j \in \mathbb{N}$, $\mathcal{A}(i, j) = 0$ whenever $|j - i| > J$;*
- *the intensities of jumps of X are bounded by $n\alpha$, i.e. for all $i, j \in \mathbb{N}$, $i \neq j$, $\mathcal{A}(i, j) \leq n\alpha$.*

Then,

$$\mathbf{E} [\|X_n - \pi_n X_n\|_{\infty, T}] \leq 2J \frac{\log n}{\log \log n}.$$

Proof. Fix $n \in \mathbb{N}$. For any $t \in [0, T]$ we have that

$$\begin{aligned} & |X_n(t) - \pi_n X_n(t)| \\ &= \left| X_n(t) - X_n\left(\frac{iT}{n}\right) - \frac{n}{T} \left(t - \frac{iT}{n}\right) \left(X_n\left(\frac{(i+1)T}{n}\right) - X_n\left(\frac{iT}{n}\right) \right) \right| \\ &\leq 2 \sup_{t \in \left[\frac{iT}{n}; \frac{(i+1)T}{n}\right]} \left| X_n(t) - X_n\left(\frac{iT}{n}\right) \right|, \end{aligned}$$

so that

$$(22) \quad \mathbf{E} [\|X_n - \pi_n X_n\|_{\infty, T}] \leq 2\mathbf{E} \left[\max_{i \in [0, n-1]} \sup_{t \in \left[\frac{iT}{n}; \frac{(i+1)T}{n}\right]} \left| X_n(t) - X_n\left(\frac{iT}{n}\right) \right| \right].$$

But for any i and any $t \in \left[\frac{iT}{n}; \frac{(i+1)T}{n}\right]$ we have that

$$\left| X_n(t) - X_n\left(\frac{iT}{n}\right) \right| \leq J (A_n^i + D_n^i),$$

where A_n^i and D_n^i denote respectively the number of up and down jumps of the process X_n within the interval $\left[\frac{iT}{n}; \frac{(i+1)T}{n}\right]$. In turn, by assumption $A_n^i + D_n^i$ is

stochastically dominated by a Poisson r.v., say P^i , of parameter $\alpha n \frac{T}{n} = \alpha T$. All in all, we obtain with (22) that

$$\mathbf{E} [\| X_n, \pi_n X_n \|_{\infty, T}] \leq 2J \mathbf{E} \left[\max_{i \in [0, n-1]} P^i \right],$$

and we conclude using Proposition A.1. \square

3.3. Functional Stein method. For $\xi = (\xi_k, k = 1, \dots, n)$ a finite sequence of positive real numbers, consider $Y = (Y_k, k = 1, \dots, n)$ a family of independent centered Gaussian random variables such that $\text{var}(Y_k) = \xi_k^2$ and

$$B_\xi := \sum_{j=1}^n Y_j h_j^n,$$

where for $0 \leq i \leq n-1$,

$$h_i^n := \sqrt{\frac{n}{T}} I_{0^+}^1 (\mathbf{1}_{[\frac{iT}{n}, \frac{(i+1)T}{n}]}).$$

Set

$$\begin{aligned} T_{\eta, p}^n : \mathbb{R}^n &\longrightarrow W_{\eta, p} \\ (y_1, \dots, y_n) &\longmapsto \sum_{j=1}^n y_j h_j^n. \end{aligned}$$

On \mathbb{R}^n , put μ_ξ^n the Gaussian measure of density:

$$(y_1, \dots, y_n) \longmapsto \frac{1}{(2\pi)^{n/2} \prod_{j=1}^n \xi_j} \exp \left(-\frac{1}{2} \sum_{j=1}^n \frac{y_j^2}{\xi_j^2} \right).$$

For any $k = 1, \dots, n$, consider the \mathbb{R} -valued process, damped Ornstein-Uhlenbeck process:

$$X^{\xi_k}(x, t) = e^{-t}x + \xi_k \sqrt{2} \int_0^t e^{-(t-s)} dB(s),$$

where B is an ordinary Brownian motion. X^{ξ_k} can alternatively be described as the solution of the stochastic differential equation

$$dX(t) = -X_t dt + \xi_k \sqrt{2} dB(t), \quad X(0) = x$$

The Itô formula easily entails that if we set

$$P_t^{\xi_k} f(x) = \mathbf{E} [f(X^{\xi_k}(x, t))] \quad \text{and} \quad L^{\xi_k} f(x) = -x f'(x) + \xi_k^2 f''(x),$$

then, for $f \in L^1(\mu_{\xi_k}^1)$,

$$(23) \quad \frac{d}{dt} P_t^{\xi_k} f(x) = L^{\xi_k} P_t^{\xi_k} f(x) = P_t^{\xi_k} L^{\xi_k} f(x) \quad \text{and} \quad (P_t^{\xi_k} f)'(x) = e^{-t} P_t^{\xi_k} f'(x).$$

Moreover, the distribution of $X^{\xi_k}(x, t)$ is Gaussian with mean $e^{-t}x$ and variance $\xi_k^2(1 - e^{-2t})$, hence

$$(24) \quad P_t^{\xi_k} f(x) = \int_{\mathbb{R}} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\mu_{\xi_k}^1(y).$$

It is then straightforward that

$$P_t^{\xi_k} f(x) \xrightarrow{t \rightarrow \infty} \int_{\mathbb{R}} f d\mu_{\xi_k}^1.$$

Combine (23) and (24) to obtain

$$(25) \quad \int_{\mathbb{R}} f d\mu_{\xi_k}^1 - f(x) = \int_0^\infty L^{\xi_k} P_t^{\xi_k} f(x) dt.$$

We now transfer this construction onto $W_{\eta,p}$. Let $\mathbf{P}_{\eta,p}^n$ be the push-forward of μ_ξ^n by the map $T_{\eta,p}^n$. Remark that $\mathbf{P}_{\eta,p}^n$ is supported by $W^n = \text{span}(h_j^n, j = 1, \dots, n)$ and that for any $(\eta, p) \in \Lambda$,

$$W^n \subset \mathcal{I}_{1,2}^+ \subset W_{\eta,p}.$$

We denote by $\mathcal{I}_{1,2}^n$, the space W^n equipped with the scalar product of $\mathcal{I}_{1,2}^+$ and by $W_{\eta,p}^n$, the space W^n with the norm induced by $W_{\eta,p}$. The space W^n is finite dimensional so that the distinction between the norms may seem spurious but it is still mandatory to keep track of the underlying infinite dimensional setting. For $y \in \mathbb{R}^n$, \mathbb{X} is the $W_{\eta,p}^n$ -valued process defined by

$$\mathbb{X}(T_{\eta,p}^n y, t) = \sum_{j=1}^n X^{\xi_j}(y_j, t) h_j^n.$$

For $y \in \mathbb{R}^n$, let

$$\mathbb{P}_t F(T_{\eta,p}^n y) = \mathbf{E} [F(\mathbb{X}(T_{\eta,p}^n y, t))].$$

By tensorization of the previous construction (or more directly using the general theory of abstract Wiener spaces [9, 12, 27]), we see that $\mathbf{P}_{\eta,p}^n$ is the stationary and invariant measure of the Markov process \mathbb{X} , whose generator is given by

$$\begin{aligned} \mathbb{L}F(T_{\eta,p}^n y) &= - \sum_{j=1}^n y_j \langle h_j^n, \nabla F(T_{\eta,p}^n y) \rangle_{W_{\eta,p}^n, (W_{\eta,p}^n)^*} \\ &\quad + \sum_{j=1}^n \xi_j^2 \langle h_j^n \otimes h_j^n, \nabla^{(2)} F(T_{\eta,p}^n y) \rangle_{(\mathcal{I}_{1,2}^+)^n}. \end{aligned}$$

This means in particular that (25) holds true in the new form:

$$(26) \quad \int_{\mathbb{R}} F d\mathbf{P}_{\eta,p}^n - F(T_{\eta,p}^n y) = \int_0^\infty \mathbb{L}\mathbb{P}_t^\xi F(T_{\eta,p}^n y) dt.$$

We can now state and prove the functional Stein's theorem which is the cornerstone of the following. In spirit, it is the multidimensional version of Theorem 3.1 of [23]. For $(u_j, j = 1, \dots, n) \in L^2(E, \mathbb{R}^n)$, set

$$u = \sum_{j=1}^n u_j \otimes h_j^n \text{ and } \delta_\nu u = \sum_{j=1}^n \delta_\nu u_j h_j^n.$$

Theorem 3.2. *Assume that $(u_k, k = 1, \dots, n)$ is an orthogonal family of elements of $L^2(\nu)$. For any $k \in \{1, \dots, n\}$, let*

$$(27) \quad \xi_k^2 = \int_E u_k(x)^2 d\nu(x).$$

Consider $Y = (Y_k, k = 1, \dots, n)$ a family of independent centered Gaussian random variables such that $\text{var}(Y_k) = \xi_k^2$ and let

$$B_\xi = \sum_{j=1}^n Y_j h_j^n$$

For any $F \in \Sigma_{\eta,p}$,

$$(28) \quad \begin{aligned} &|\mathbf{E}[F(B_\xi)] - \mathbf{E}[F(\delta_\nu u)]| \\ &\leq \frac{1}{2} n^{-3/2+\eta} \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \int_E |u_j(x) u_k(x)| |u_l(x)| d\nu(x). \end{aligned}$$

Remark 2. *If we compare Theorem 3.2 with Theorem 3.1 of [23], we retrieve a third order moment. It is not surprising that we have crossed moments since we need to control how the correlations between the different coordinates of δu do vanish as n tends to infinity. We already know that each of the components tend to a Gaussian distribution, the only point at stake is then to show that they become more and more independent as the size of the vector increases.*

At last, remark that there is no term involving the second order moments, this is due to the hypothesis of orthogonality.

Proof of Theorem 3.2. According to (26), we have

$$(29) \quad \mathbf{E} [F(B_\xi)] - \mathbf{E} [F(\delta_\nu u)] = \mathbf{E} \left[\int_0^\infty \mathbb{L}\mathbb{P}_t F(\delta_\nu u) dt \right].$$

According to the integration by parts formula and to the fundamental theorem of calculus, we get

$$\begin{aligned} & \sum_{j=1}^n \mathbf{E} \left[\delta_\nu u_j \langle h_j^n, \nabla \mathbb{P}_t F(\delta_\nu u) \rangle_{\mathcal{I}_{1,2}^+} \right] \\ &= \sum_{j=1}^n \mathbf{E} \left[\int_E u_j(x) \langle h_j^n, \nabla \mathbb{P}_t F(\delta_\nu u + u(x)) - \nabla \mathbb{P}_t F(\delta_\nu u) \rangle_{\mathcal{I}_{1,2}^+} d\nu(x) \right] \\ &= \sum_{j,k=1}^n \mathbf{E} \left[\int_E \int_0^1 u_j(x) u_k(x) \langle h_j^n \otimes h_k^n, \nabla^{(2)} \mathbb{P}_t F(\delta_\nu u + r u(x)) \rangle_{\mathcal{I}_{1,2}^+} dr d\nu(x) \right]. \end{aligned}$$

Since the u_k 's are orthogonal, in view of Eqn. (27),

$$\begin{aligned} & \mathbf{E} \left[\sum_{j,k=1}^n \int_E \int_0^1 u_j(x) u_k(x) \langle h_j^n \otimes h_k^n, \nabla^{(2)} \mathbb{P}_t F(\delta_\nu u) \rangle_{\mathcal{I}_{1,2}^+} dr d\nu(x) \right] \\ &= \mathbf{E} \left[\sum_{j=1}^n \xi_j^2 \langle h_j^n \otimes h_j^n, \nabla^{(2)} \mathbb{P}_t F(\delta_\nu u) \rangle_{\mathcal{I}_{1,2}^+} \right]. \end{aligned}$$

Hence

$$\begin{aligned} & \mathbf{E} [\mathbb{L}\mathbb{P}_t F(\delta_\nu u)] \\ &= \sum_{j,k=1}^n \int_E \int_0^1 \mathbf{E} \left[\langle h_j^n \otimes h_k^n, \nabla^{(2)} \mathbb{P}_t F(\delta_\nu u + r u(x)) - \nabla^{(2)} \mathbb{P}_t F(\delta_\nu u) \rangle_{\mathcal{I}_{1,2}^+} \right] \\ & \quad \times u_j(x) u_k(x) dr d\nu(x). \end{aligned}$$

Recall that $\|h_j^n\|_{L^2} \leq n^{-1/2}$ and note that

$$\|u(x)\|_{W_{\eta,p}} \leq \sum_{l=1}^n |u_l(x)| \|h_l^n\|_{W_{\eta,p}} = n^{-(1/2-\eta)} \sum_{l=1}^n |u_l(x)|.$$

Since F belongs to $\Sigma_{\eta,p}$

$$(30) \quad |\mathbf{E} [\mathbb{L}\mathbb{P}_t F(\delta_\nu u)]| \leq n^{-3/2+\eta} e^{-2t} \sum_{j,k,l=1}^n \int_0^1 |u_j(x) u_k(x)| |u_l(x)| d\nu(x).$$

Plug (30) into (29) yields (28). \square

4. THE M/M/1 QUEUE

The M/M/1 queue consists in a single server with infinite queue, where the service times are independently and identically distributed, according to an exponential distribution of parameter μ . The customers arrive at the time epochs of a Poisson process of intensity $\lambda > 0$. Let $L^\dagger(t)$ denote the number of customers in the system (including the one in service, if any) at time $t \geq 0$. The process $(L^\dagger(t), t \geq 0)$ counting the number of customers in the system is clearly a birth and death process, that is ergodic if and only if $\lambda/\mu < 1$. If the initial size of the system is $x \in \mathbb{N}$, then L^\dagger obeys the SDE

$$L^\dagger(t) = x + N_\lambda(t) - \int_0^t \mathbf{1}_{\{L(s^-) > 0\}} N_\mu(ds),$$

for two independent Poisson processes N_λ and N_μ . This process is rescaled by accelerating time by an arbitrarily large factor n , while multiplying the number of customers in the initial state by the same factor, and then dividing the number of customers in the system at any time by n : with obvious notation, for all $t \geq 0$ we obtain

$$\overline{L}_n^\dagger(t) = x + \frac{N_{n\lambda}(t)}{n} - \frac{N_{n\mu}(t)}{n} + \frac{1}{n} \int_0^t \mathbf{1}_{\{L_n^\dagger(ns^-) = 0\}} dN_{n\mu}(s).$$

It is a well established fact (see e.g. Proposition 5.16 in [25]) that the sequence $(\overline{L}_n^\dagger : n \geq 1)$ converges in probability and uniformly over compact sets, to a deterministic process

$$\overline{L}^\dagger : t \mapsto \overline{L}^\dagger(t) = (x + \lambda t - \mu t)^+,$$

and that

$$\begin{aligned} Z_n^\dagger(t) &:= \frac{\sqrt{n}}{\sqrt{\lambda + \mu}} \left(\overline{L}_n^\dagger(t) - \overline{L}^\dagger(t) \right) \\ &= \frac{\sqrt{n}}{\sqrt{\lambda + \mu}} \left[\left(\frac{N_{n\lambda}(t)}{n} - \lambda t \right) - \left(\frac{N_{n\mu}(t)}{n} - \mu t \right) \right] \\ &\quad + \frac{\sqrt{n}}{\sqrt{\lambda + \mu}} \left(\frac{1}{n} \int_0^t \mathbf{1}_{\{\overline{L}_n^\dagger(s^-) = 0\}} dN_{n\mu}(s) - \int_0^t \mathbf{1}_{\{\overline{L}^\dagger(s) = 0\}} ds \right) \end{aligned}$$

converges in distribution on \mathbb{D} to the standard Brownian motion B .

We aim at controlling the speed of the latter convergence. For that purpose, we bound for any fixed n and any horizon T , the $\mathcal{L}_{\eta,p}$ -distance between these processes, defined by (14). We have the following result,

Theorem 4.1. *Suppose that $\lambda < \mu$ and let $T \leq \frac{x}{\mu - \lambda}$. Then, for all $n \in \mathbb{N}$ we have that*

$$d_{\mathcal{L}_{\eta,p}}(Z_n^\dagger, B) \leq \frac{c \log n}{\log \log n \sqrt{n}},$$

where B is a standard Brownian motion.

Proof. Fix $T \leq \frac{x}{\mu - \lambda}$. Then for all $n \in \mathbb{N}$, recalling (13) and (20), by the very definition of the set $\mathcal{L}_{\eta,p}$ in (12) we have that

$$(31) \quad d_{\mathcal{L}_{\eta,p}}(Z_n^\dagger, B) \leq d_{\mathcal{L}_{\eta,p}}(Z_n^\dagger, \pi_n Z_n^\dagger) + d_{\Sigma_{\eta,p}}(\pi_n Z_n^\dagger, \pi_n B) + d_{\Sigma_{\eta,p}}(\pi_n B, B).$$

First observe that the function \overline{L}^\dagger is affine hence $\pi_n \overline{L}^\dagger$ on $[0, T]$. Moreover, the operator π_n is linear and the elements of $\mathcal{L}_{\eta,p}$ are 1-Lipschitz-continuous, thus we

have that for all n ,

$$(32) \quad \begin{aligned} d_{\mathcal{L}_{\eta,p}}(Z_n^\dagger, \pi_n Z_n^\dagger) &\leq \mathbf{E} [\| Z_n^\dagger - \pi_n Z_n^\dagger \|_{\infty, T}] \leq \frac{1}{\sqrt{n(\lambda + \mu)}} \mathbf{E} [\| L_n^\dagger - \pi_n L_n^\dagger \|_{\infty, T}] \\ &\leq \frac{c \log n}{\log \log n \sqrt{n}}, \end{aligned}$$

where the last inequality follows from applying Lemma 3.1 to the Markov processes $(L_n^\dagger : n \geq 1)$ for $J \equiv 1$ and $\alpha \equiv \lambda \vee \mu$. Now, for any $n \in \mathbb{N}$, if we let $\tau_0^n = \inf\{t > 0, L_n^\dagger(t) = 0\}$, for any $F \in \Sigma_{\eta,p}$ we have that

$$(33) \quad \begin{aligned} \mathbf{E} [|F(\pi_n Z_n^\dagger) - F(\pi_n B)|] \\ = \mathbf{E} [|F(\pi_n Z_n^\dagger) - F(\pi_n B)| \mathbf{1}_{\{T < \tau_0^n\}}] + \mathbf{E} [|F(\pi_n Z_n^\dagger) - F(\pi_n B)| \mathbf{1}_{\{T \geq \tau_0^n\}}]. \end{aligned}$$

We first prove that for some $c > 0$,

$$(34) \quad \mathbf{E} [|F(\pi_n Z_n^\dagger) - F(\pi_n B)| \mathbf{1}_{\{T < \tau_0^n\}}] \leq \frac{c}{\sqrt{n}}, \quad n \in \mathbb{N}.$$

Fix $n \in \mathbb{N}$. On the event $\{T \leq \tau_0^n\}$, for any $t \in [0, T)$ we have that

$$\begin{aligned} Z_n^\dagger(t) &= \frac{1}{\sqrt{\lambda + \mu}} \left(\sqrt{\lambda} \left(\frac{N_{n\lambda}(t)}{\sqrt{\lambda n}} - \sqrt{\lambda n t} \right) - \sqrt{\mu} \left(\frac{N_{n\mu}(t)}{\sqrt{\mu n}} - \sqrt{\mu n t} \right) \right) \\ &=: \frac{1}{\sqrt{\lambda + \mu}} \left(Z_{\lambda,n}^\dagger(t) - Z_{\mu,n}^\dagger(t) \right). \end{aligned}$$

Consider $N_{n(\lambda+\mu)}^\dagger$ the marked Poisson point process of control measure

$$d\nu_n^\dagger(s, r) = n(\lambda + \mu) ds \otimes \left(\frac{\lambda}{\lambda + \mu} \varepsilon_1(dr) + \frac{\mu}{\lambda + \mu} \varepsilon_{-1}(dr) \right).$$

That is to say, $N_{n(\lambda+\mu)}^\dagger$ is constructed as an ordinary Poisson process on the positive real line with intensity $n(\lambda + \mu)$ and each atom is assigned a mark $+1$ or -1 , independent of the atom location and of the other marks, with respective probability $\lambda(\lambda + \mu)^{-1}$ and $\mu(\lambda + \mu)^{-1}$. By the thinning property of Poisson processes, the point process which counts the atom $N_{n(\lambda+\mu)}^\dagger$ with mark $+1$ (respectively -1) is a Poisson process with intensity $n\lambda$ (respectively $n\mu$). For any $t \in [0, T]$, let

$$\begin{aligned} v_t &: [0, T] \times \{-1, 1\} \longrightarrow \mathbb{R} \\ (s, r) &\longmapsto \frac{1}{\sqrt{n(\lambda + \mu)}} r \mathbf{1}_{[0,t)}(s). \end{aligned}$$

Then, we have

$$Z_n^\dagger(t) \stackrel{\text{dist}}{=} \delta_{\nu_n^\dagger}(v_t)$$

and

$$\pi_n Z_{\lambda,n}^\dagger \stackrel{\text{dist}}{=} \sum_{i=0}^{n-1} \frac{n}{T} \delta_{\nu_n^\dagger}(v_{(i+1)T/n} - v_{iT/n}) I_{0+}^1 \left(\mathbf{1}_{\left[\frac{iT}{n}, \frac{(i+1)T}{n}\right)} \right).$$

For $i = 1, \dots, n$, let

$$\begin{aligned} u_i^\dagger(s, r) &= \sqrt{\frac{n}{T}} \left(v_{(i+1)T/n}(s, r) - v_{iT/n}(s, r) \right) \\ &= \frac{1}{\sqrt{T(\lambda + \mu)}} r \mathbf{1}_{[iT/n, (i+1)T/n)}(s), \end{aligned}$$

so that

$$\pi_n Z_{\lambda,n}^\dagger \stackrel{\text{dist}}{=} \sum_{i=0}^{n-1} \delta_{\nu_n^\dagger}(u_i^\dagger) h_i^n.$$

It is clear that

$$\int_{[0,T] \times \{-1,1\}} u_i^\dagger(s,r) u_j^\dagger(s,r) d\nu_n^\dagger(s,r) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Moreover,

$$\sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \int_E |u_j^\dagger u_k^\dagger u_l^\dagger| d\nu_n^\dagger = \frac{1}{T^{3/2} \sqrt{\lambda + \mu}} \sum_{i=1}^n \int_{(i-1)T/n}^{iT/n} n ds = \frac{n}{\sqrt{T(\lambda + \mu)}}.$$

Let $\xi^\dagger = (1, k = 1, \dots, n)$. Since $\pi_n B = B_{\xi^\dagger}$, as a consequence of Theorem 3.2, we obtain (34).

Regarding the second term on the right-hand side of (33), fix $n \in \mathbb{N}$ and observe that F is in particular bounded, so there exists a constant c such that

$$\mathbf{E} \left[|F(\pi_n Z_n^\dagger) - F(\pi_n B)| \mathbf{1}_{\{T > \tau_0^n\}} \right] \leq c \mathbf{P} [T > \tau_0^n].$$

But $\mathbf{P} [T > \tau_0^n]$ tends to 0 with exponential speed from Theorem 11.9 of [28]: if $\rho < 1$, for any $x > 0$ and any $y < 0$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{P} \left[\tau_0^n \leq \frac{x}{\lambda - \mu} + y \right] = -f(y),$$

where f is strictly positive on $(0, \infty)$. This shows that

$$\mathbf{E} \left[|F(Z_n^\dagger) - F(B)| \mathbf{1}_{\{T > \tau_0^n\}} \right] \leq ce^{-n}$$

for some d , which, together (34) in (33), shows that for some constant c ,

$$d_{\Sigma_{n,p}}(\pi_n Z_n^\dagger, \pi_n B) \leq \frac{c}{\sqrt{n}}, n \in \mathbb{N}.$$

This, together with (32) and (21) in (31), concludes the proof. \square

5. THE M/M/ ∞ QUEUE

We consider an M/M/ ∞ queue: a potentially unlimited number of servers attend customers that enter the system following a Poisson process of intensity λ , requesting service times that are exponentially distributed of parameter μ (where $\lambda, \mu > 0$).

If $L^\sharp(t)$ denotes the number of customers in the system at time t , L^\sharp is an ergodic Markov process which obeys the SDE

$$L^\sharp(t) = x_0 + N_\lambda(t) - \sum_{i=1}^{+\infty} \int_0^t \mathbf{1}_{\{L^\sharp(s^-) \geq i\}} N_\mu^i(ds), \quad t \geq 0,$$

where N_λ is a Poisson process of intensity λ , the N_μ^i 's are independent Poisson processes of intensity μ , and x_0 is the initial number of customers at time 0. For simplicity, we assume throughout this Section that the system is initially empty, i.e.

$$(35) \quad x_0 = 0.$$

The process L^\sharp is then rescaled by accelerating time by a factor $n \in \mathbb{N}$, and dividing the size of the system by n . From (35), the corresponding n -th rescaled process is thus defined by

$$\overline{L}_n^\sharp : t \mapsto \frac{N_{\lambda n}(t)}{n} - \frac{1}{n} \sum_{i=1}^{+\infty} \int_0^t \mathbf{1}_{\{\overline{L}_n^\sharp(s^-) > \frac{i}{n}\}} N_\mu^i(ds).$$

It is a well known fact (see e.g. [4] or Theorem 6.14 in [25]) that the sequence of processes $(\overline{L}_n^\sharp, n \geq 0)$ converges in L^1 and uniformly over compact sets to the deterministic function

$$(36) \quad L^\sharp : t \mapsto \overline{L}^\sharp(t) = \rho - \rho e^{-\mu t},$$

where $\rho = \lambda/\mu$. Moreover, if we define for all n the process

$$(37) \quad Z_n^\sharp : t \mapsto Z_n^\sharp(t) = \sqrt{n} \left(\overline{L}_n^\sharp(t) - \overline{L}^\sharp(t) \right),$$

the sequence $(Z_n^\sharp : n \geq 0)$ converges in distribution to the process X^\sharp defined by

$$(38) \quad X^\sharp : t \mapsto X^\sharp(t) = X^\sharp(0)e^{-\mu t} + \int_0^t e^{-\mu(t-s)} \sqrt{h(s)} dB(s),$$

where $h(t) = \lambda(2 - e^{-\mu t})$ for all $t \geq 0$ and B is the standard Brownian motion. The following result is shown in [25] (eq. (6.23) in Chapter 6),

Proposition 5.1. *The sequence of processes $(Y_n^\sharp : n \geq 0)$ defined by*

$$(39) \quad t \mapsto Y_n^\sharp(t) = Z_n^\sharp(t) - Z_n^\sharp(0) + \mu \int_0^t Z_n^\sharp(s) ds$$

converges in distribution to $B \circ \gamma$, where

$$(40) \quad \gamma(t) = 2\lambda t - \rho(1 - e^{-\mu t}) \text{ for all } t \geq 0.$$

This section is devoted to assessing the rate of convergence in Proposition 5.1, on any fixed time interval $[0, T]$, where $T > 0$ is fixed throughout.

5.1. An integral transformation. The process X^\sharp defined in (38) is an Ornstein-Uhlenbeck process which can be analyzed by introducing a one to one mapping between the space of rcll functions and $\mathbb{R} \times \mathbb{D}_0$.

Proposition 5.2. *The mapping*

$$\Theta : \begin{cases} \mathbb{D}([0, T]) & \longrightarrow \mathbb{R} \times \mathbb{D}_0([0, T]) \\ f & \longmapsto \left(f(0), f(\cdot) - f(0) + \tau \int_0^\cdot f(s) ds \right) \end{cases}$$

is linear, continuous and one to one.

Proof. Let us fix $\eta \in \mathbb{D}_0([0, T])$ and consider the following integral equation of unknown function z ,

$$z(t) - z(0) = -\tau \int_0^t z(s) ds + \eta(t).$$

We clearly have for all $t \geq 0$,

$$z(t) = z(0)e^{-\tau t} + \eta(t) - \tau \int_0^t e^{-\tau(t-s)} \eta(s) ds,$$

hence Θ is bijective and for all $(x, \eta) \in \mathbb{R} \times \mathbb{D}_0([0, T])$,

$$(41) \quad \Theta^{-1}(x, \eta) = \left(t \mapsto xe^{-\tau t} + \eta(t) - \tau \int_0^t e^{-\tau(t-s)} \eta(s) ds \right).$$

Linearity and continuity are straightforward. \square

Lemma 5.3. *On the subset of $\{0\} \times \Theta(\mathcal{I}_{\alpha, 2}^+)$ whose image by Θ^{-1} is in $\mathcal{I}_{\alpha, 2}^+$, Θ^{-1} is Lipschitz continuous.*

Proof. For all η and $\omega \in \Theta(\mathcal{I}_{\alpha,2}^+)$ and all $t \in [0, 1]$, we have that

$$\Theta^{-1}(0, \eta) - \Theta^{-1}(0, \omega) = \eta(t) - \omega(t) - \tau \int_0^t e^{-\tau(t-s)}(\eta(s) - \omega(s)) \, ds.$$

Hence

$$\|\Theta^{-1}(0, \eta) - \Theta^{-1}(0, \omega)\|_{\alpha,2} < t \|\eta - \omega\|_{\alpha,2} + \tau \|\eta - \omega\|_{\alpha,2} \sqrt{\int_0^t e^{-2\tau(t-s)} \, ds},$$

so that for some positive number k ,

$$\|\Theta^{-1}(x, \eta) - \Theta^{-1}(y, \omega)\|_{\alpha,2} < k \|\eta - \omega\|_{\alpha,2}.$$

The proof is thus complete. \square

As they are either piecewise linear or piecewise constant L^\sharp, Z_n^\sharp and Y_n^\sharp all belong to $\mathcal{I}_{\alpha,2}^+$. And since the $\mathcal{I}_{\alpha,2}^+$ are closed sets, X^\sharp and $B \circ \gamma$ also belong to $\mathcal{I}_{\alpha,2}^+$. As the following corollary demonstrates, the Lipschitz property of Θ^{-1} enables to reduce the calculation of the rate of convergence of Z_n^\sharp towards X^\sharp to that of Y_n^\sharp towards $B \circ \gamma$,

Corollary 5.4. *Almost surely, for some positive constant c we have*

$$d_{\Sigma_{n,p}}(\pi_n Z_n^\sharp, X^\sharp) \leq c d_{\Sigma_{n,p}}(\pi_n Y_n^\sharp, B \circ \gamma).$$

Proof. From Proposition 5.1, Y_n^\sharp converges in distribution to $B \circ \gamma$. From the Lipschitz continuity of Θ and the Continuous Mapping Theorem, setting $\tau = \mu$ this entails the convergence in distribution

$$\Theta(Z_n^\sharp) = (0, Y_n^\sharp) \Rightarrow (0, B \circ \gamma)$$

However, expression (6.34) in [21] shows that for all t , $\Theta(X^\sharp(t)) = (0, B \circ \gamma(t))$ which, together with the linearity of Θ and of the operator π_n for any n , concludes the proof. \square

5.2. Alternative representation. As a consequence of a Corollary 5.4, the rate of convergence of $(G_n : n \geq 0)$ towards X^\sharp is bounded by that of $(Y_n : n \geq 0)$ towards $B \circ \gamma$, which we now address.

For that purpose, and following Section 7.2 of [25], an alternative representation of L^\sharp in terms of a marked Poisson process will prove useful. A point (x, z) represents a customer arriving at time x and requiring a service of duration z . Let $N_{\lambda,\mu}$ be a Poisson process on $\mathbb{R}^+ \times \mathbb{R}^+$ of control measure $\lambda \, dx \otimes \mu e^{-\mu z} \, dz$. At any time $t \geq 0$, the number of busy servers at t equals the number of points located in the shaded trapeze bounded by the axes of equation $x = 0$ and $x = t$, and above the line $z = t - x$: in other words

$$L^\sharp(t) = \int_{C_t} dN_{\lambda,\mu}(x, z),$$

where

$$(42) \quad C_t = \{(x, z), 0 \leq x \leq t, z \geq t - x\}.$$

Fix a positive integer n throughout this section. After scaling, for all $t \geq 0$ we get that

$$\overline{L_n^\sharp}(t) = \frac{1}{n} N_{\lambda n, \mu}(C_t).$$

Let us denote for all (x, z) in the positive orthant by

$$d\nu_n^\sharp(x, z) := \lambda n \, dx \otimes \mu e^{-\mu z} \, dz,$$

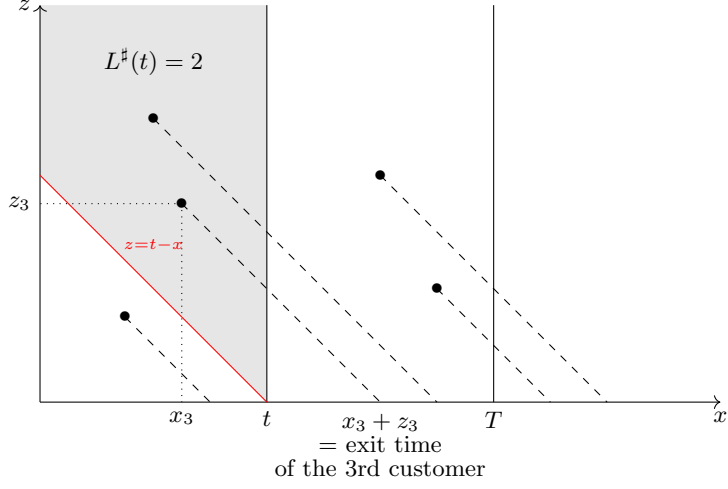


FIGURE 1. Representation of the $M/M/\infty$ queue

the control measure of $N_{n\lambda,\mu}$. As readily follows from (36), the fluid limit $\bar{L}^\#$ can be written for all $t \geq 0$ as

$$\bar{L}^\#(t) = \frac{1}{n} \int \mathbf{1}_{C_t}(x, z) \, d\nu_n^\#(x, z),$$

in a way that

$$(43) \quad Z_n^\#(t) = \frac{1}{\sqrt{n}} \int \mathbf{1}_{C_t} (dN_{\lambda n, \mu} - d\nu_n^\#),$$

for C_t defined by (42). We deduce that for all $t \geq 0$,

$$(44) \quad \begin{aligned} Y_n^\#(t) &= \frac{1}{\sqrt{n}} \int \mathbf{1}_{C_t} (dN_{\lambda n, \mu} - d\nu_n^\#) \\ &\quad + \mu \int_0^t \frac{1}{\sqrt{n}} \int \mathbf{1}_{C_s} (dN_{\lambda n, \mu} - d\nu_n^\#) \, du \\ &= \frac{1}{\sqrt{n}} \delta_{\lambda n, \mu}(\mathbf{1}_{C_t}) + \mu \int_0^t \frac{1}{\sqrt{n}} \delta_{\lambda n, \mu}(\mathbf{1}_{C_u}) \, du, \end{aligned}$$

where $\delta_{\lambda n, \mu}$ is the compensated integral with respect to the Poisson process $N_{\lambda n, \mu}$, see (18).

5.3. Reduction to the finite dimension. Fix $n \in \mathbb{N}$ and recall (20). It follows from (44) that

$$\begin{aligned} \pi_n Y_n^\# &= \sum_{i=0}^{n-1} \frac{n}{T} \left(Y_n^\# \left(\frac{(i+1)T}{n} \right) - Y_n^\# \left(\frac{iT}{n} \right) \right) I_{0^+}^1 \left(\mathbf{1}_{[\frac{iT}{n}, \frac{(i+1)T}{n}]} \right) \\ &= \sum_{i=0}^{n-1} \left(\delta_{\lambda n, \mu} \left(\mathbf{1}_{C_{\frac{(i+1)T}{n}}} - \mathbf{1}_{C_{\frac{iT}{n}}} \right) + \mu \int_{\frac{iT}{n}}^{\frac{(i+1)T}{n}} \delta_{\lambda n, \mu}(\mathbf{1}_{C_u}) \, du \right) h_i^n \\ &= \sum_{i=0}^{n-1} \delta_{\lambda n, \mu}(u_i^\#) h_i^n, \end{aligned}$$

where for all i and all $(x, z) \in \mathbb{R}^2$,

$$(45) \quad u_i^\sharp(x, z) = \frac{1}{\sqrt{T}} \left(\mathbf{1}_{C_{\frac{(i+1)T}{n}}}(x, z) - \mathbf{1}_{C_{\frac{iT}{n}}}(x, z) + \mu \int_{\frac{iT}{n}}^{\frac{(i+1)T}{n}} \mathbf{1}_{C_u}(x, z) \, du \right).$$

Let us denote for any $i = 0, \dots, n-1$,

$$\xi_i^\sharp := \sqrt{\gamma \left(\frac{(i+1)T}{n} \right) - \gamma \left(\frac{iT}{n} \right)}.$$

The following result is proven in appendix B,

Proposition 5.5. *For any n , the family $(u_i^\sharp, i = 1, \dots, n)$ has the following properties:*

- (i) *It is orthogonal in $L^2(u_n^\sharp)$;*
- (ii) *For some constant c independent of n ,*

$$\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \int_E |u_i^\sharp u_j^\sharp u_k^\sharp| \, d\nu_n^\sharp \leq nc.$$

- (iii) *For any $i \in \{1, \dots, n\}$,*

$$\int \int u_i^\sharp u_i^\sharp \, d\nu_n^\sharp = \frac{n}{T} (\xi_i^\sharp)^2.$$

Notice that for a large enough n , for all $t \geq 0$,

$$\frac{n}{t} (\xi_i^\sharp)^2 \xrightarrow{i/n \rightarrow t} \gamma'(t) \text{ and for a fixed } i, \frac{n}{t} (\xi_i^\sharp)^2 \xrightarrow{n \rightarrow \infty} \gamma'(0).$$

We thus have the following result,

Proposition 5.6. *For some c , for all positive integer n , the respective interpolations of Y_n^\sharp and $B \circ \gamma$ satisfy*

$$d_{\Sigma_{\eta,p}}(\pi_n Y_n^\sharp, \pi_n(B \circ \gamma)) \leq \frac{c}{\sqrt{n}}.$$

Proof. Recall that $\pi_n(B \circ \gamma)$ can be represented as

$$\pi_n(B \circ \gamma) \stackrel{\text{dist}}{=} \sum_{j=1}^n Y_j^\sharp h_j^n = B_{\xi^\sharp},$$

where $(Y_k^\sharp, k = 1, \dots, n)$ is a family of independent centered Gaussian random variables such that $\text{var}(Y_k^\sharp) = (\xi_k^\sharp)^2$ for all k . From assertion (i) of Proposition 5.5, we can apply Theorem 3.2 : for any $F \in \Sigma_{\eta,p}$,

$$|\mathbf{E}[F(B_{\xi^\sharp})] - \mathbf{E}[F(\delta_{\lambda n, \mu} u^\sharp)]| \leq \frac{1}{2} n^{-3/2+\eta} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \int_E |u_i^\sharp u_j^\sharp| |u_k^\sharp| \, d\nu_n^\sharp.$$

Assertions (ii) and (iii) of Proposition 5.5 allow us to conclude. \square

5.4. Speed of convergence. We can now state our main result of this section,

Theorem 5.7. *On the interval $[0, T]$, let $(G_n : n \geq 1)$ be the sequence of processes defined for all n by (37) and X^\sharp be the process defined by (38). Then there exists a constant $c > 0$ such that for all n ,*

$$d_{\mathcal{L}_{\eta,p}}(Z_n^\sharp, X^\sharp) \leq \frac{c \log n}{\log \log n \sqrt{n}}.$$

Proof. For any $n \in \mathbb{N}$ we have that

$$\begin{aligned}
(46) \quad d_{\mathcal{L}_{\eta,p}}(Z_n^\sharp, X^\sharp) &\leq d_{\mathcal{L}_{\eta,p}}(Z_n^\sharp, \pi_n Z_n^\sharp) + d_{\Sigma_{\eta,p}}(\pi_n Z_n^\sharp, X^\sharp) \\
&\leq d_{\mathcal{L}_{\eta,p}}(Z_n^\sharp, \pi_n Z_n^\sharp) + c d_{\Sigma_{\eta,p}}(\pi_n Y_n^\sharp, B \circ \gamma) \\
&\leq d_{\mathcal{L}_{\eta,p}}(Z_n^\sharp, \pi_n Z_n^\sharp) + c d_{\Sigma_{\eta,p}}(\pi_n Y_n^\sharp, \pi_n B \circ \gamma) + c d_{\Sigma_{\eta,p}}(\pi_n B \circ \gamma, B \circ \gamma),
\end{aligned}$$

where we applied Corollary 5.4 in the second inequality.

First define the stopping times

$$\tau_n^\sharp = \inf \{t \geq 0 : N_{n\lambda}(t) \geq 2\lambda n T\}, \quad n \in \mathbb{N}.$$

Then, as all functions of $\mathcal{L}_{\eta,p}$ are bounded and Lipschitz continuous we obtain that for all n ,

$$\begin{aligned}
(47) \quad d_{\mathcal{L}_{\eta,p}}(Z_n^\sharp, \pi_n Z_n^\sharp) &\leq \sup_{F \in \mathcal{L}_{\eta,p}} \mathbf{E} \left[|F(Z_n^\sharp) - F(\pi_n Z_n^\sharp)| \mathbf{1}_{\{T < \tau_n^\sharp\}} \right] + c \mathbf{P}[T \geq \tau_n^\sharp] \\
&\leq \mathbf{E} \left[d_{\mathbb{D}}(Z_n^\sharp, \pi_n Z_n^\sharp) \mathbf{1}_{\{T < \tau_n^\sharp\}} \right] + c \mathbf{P}[T \geq \tau_n^\sharp] \\
&\leq \mathbf{E} \left[\|Z_n^\sharp(\cdot \wedge \tau_n^\sharp) - \pi_n(Z_n^\sharp(\cdot \wedge \tau_n^\sharp))\|_{\infty, T} \mathbf{1}_{\{T < \tau_n^\sharp\}} \right] + c \mathbf{P}[T \geq \tau_n^\sharp].
\end{aligned}$$

On the one hand, from Tchebychev inequality, we have for all n ,

$$(48) \quad \mathbf{P}[T \geq \tau_n^\sharp] = \mathbf{P}[N_{n\lambda}(T) \geq 2\lambda n T] \leq \frac{\text{Var}(N_{n\lambda}(T))}{(\lambda n T)^2} \leq \frac{c}{n}.$$

Also, for any n on $\{T < \tau_n^\sharp\}$ we have that

$$L_n^\sharp(t \wedge \tau_n^\sharp) \leq N_{n\lambda}(t) \leq 2\lambda n T,$$

therefore the Markov process $L_n^\sharp(\cdot \wedge \tau_n^\sharp)$ satisfies to the Assumptions of Lemma 3.1 for $J \equiv 1$ and $\alpha \equiv \lambda \vee (\mu T)$. Thus we obtain as in (32) that for any n ,

$$\begin{aligned}
(49) \quad \mathbf{E} \left[\|Z_n^\sharp(\cdot \wedge \tau_n^\sharp) - \pi_n(Z_n^\sharp(\cdot \wedge \tau_n^\sharp))\|_{\infty, T} \mathbf{1}_{\{T < \tau_n^\sharp\}} \right] &\leq \frac{1}{\sqrt{n}} \mathbf{E} [\|L_n^\sharp - \pi_n L_n^\sharp\|_{\infty, T}] + \sqrt{n} \|\bar{L}^\sharp - \pi_n \bar{L}^\sharp\|_{\infty, T} \\
&\leq \frac{c \log n}{\log \log n \sqrt{n}}.
\end{aligned}$$

Recalling (36), we use the fact that

$$\begin{aligned}
\sqrt{n} \|\bar{L}^\sharp - \pi_n \bar{L}^\sharp\|_{\infty, T} &\leq 2\sqrt{n} \max_{i \in [0, n-1]} \sup_{t \in [\frac{iT}{n}; \frac{(i+1)T}{n}]} \left| e^{-\mu t} - e^{-\mu \frac{iT}{n}} \right| \\
&\leq 2\sqrt{n} \left(e^{-\frac{\mu}{n}} - 1 \right) \leq \frac{c}{\sqrt{n}}.
\end{aligned}$$

Finally, gathering (49) with (48) in (47) entails that for all n ,

$$d_{\mathcal{L}_{\eta,p}}(Z_n^\sharp, \pi_n Z_n^\sharp) \leq \frac{c \log n}{\sqrt{n}}$$

which, together with with Proposition 5.6 and (21) in (46), concludes the proof. \square

APPENDIX A. MOMENT BOUND FOR POISSON VARIABLES

By following closely Chapter 2 in [5], we show hereafter a moment bound for the maximum of n Poisson variables. (Notice that, contrary to Exercise 2.18 in [5] we do not assume here that the Poisson variables are independent.)

Proposition A.1. *Let $n \in \mathbb{N}$ and let $X_i, i = 1, \dots, n$ be Poisson random variables of parameter ν . Then for some c depending only on ν we have that*

$$(50) \quad \mathbf{E} \left[\max_{i=1, \dots, n} X_i \right] \leq c \frac{\log n}{\log \log n}.$$

Proof. Denote for all i , $Z_i = X_i - \nu$, and by Ψ_{Z_i} the moment generating function of Z_i . By Jensen's inequality and the monotonicity of $\exp(\cdot)$ we get that

$$\exp \left(u \mathbf{E} \left[\max_{i=1, \dots, n} Z_i \right] \right) \leq \mathbf{E} \left[\max_{i=1, \dots, n} \exp(uZ_i) \right] \leq \sum_{i=1}^n \mathbf{E} [\exp(uZ_1)] \leq n \exp(\Psi_{Z_1}(u)).$$

After a quick algebra, this readily implies that

$$\mathbf{E} \left[\max_{i=1, \dots, n} Z_i \right] \leq \inf_{u \in \mathbb{R}} \left(\frac{\log n + \nu(e^u - u - 1)}{u} \right) = \frac{\log n + \nu \left(e^{\frac{a}{W(a)}} - 1 - W(a) - 1 \right)}{1 + W(a)},$$

where W is the so-called Lambert function, solving the equation $W(x)e^{W(x)} = x$ over $[-1/e, \infty]$, and $a = \frac{\log(n/e^\nu)}{e^\nu}$. This entails in turn that

$$\mathbf{E} \left[\max_{i=1, \dots, n} X_i \right] \leq \nu e \frac{a}{W(a)} - \nu + \nu = \frac{\log(n/e^\nu)}{W(\log(n/e^\nu)/e^\nu)}.$$

We conclude by observing that $W(z) \geq \log(z) - \log \log(z)$ for all $z > e$. Therefore there exists $c > 0$ such that for $n \geq \exp(e^{\nu+1} + \nu)$,

$$\mathbf{E} \left[\max_{i=1, \dots, n} X_i \right] \leq \frac{\log(n/e^\nu)}{\log(\log(n/e^\nu)/e^\nu) - \log \log(\log(n/e^\nu)/e^\nu)} \leq c \frac{\log n}{\log \log n},$$

which completes the proof. \square

APPENDIX B. PROOF OF PROPOSITION 5.5

Without loss of generality we set $T = 1$. Fix n throughout this section, and denote for all $i \in [1, n]$ and $(x, z) \in \mathbb{R}^2$,

$$\alpha_i(x, z) = \mathbf{1}_{C_{\frac{i}{n}}}(x, z), \quad \beta_i(x, z) = \int_{\frac{i}{n}}^{\frac{i+1}{n}} \mathbf{1}_{C_u}(x, z) \, du.$$

Proof of (i). Recall (45), and fix two indexes $1 \leq i < j \leq n$. We have that

$$(51) \quad \begin{aligned} \int \int u_i^\# u_j^\# \, d\nu_n^\# &= \int \int (\alpha_{i+1} - \alpha_i)(\alpha_{j+1} - \alpha_j) \, d\nu_n^\# \\ &+ \mu \int \int \beta_i(\alpha_{j+1} - \alpha_j) \, d\nu_n^\# + \mu \int \int \beta_j(\alpha_{i+1} - \alpha_i) \, d\nu_n^\# + \mu^2 \int \int \beta_j \beta_i \, d\nu_n^\# \\ &=: I_1 + I_2 + I_3 + I_4, \end{aligned}$$

where straightforward computations show that

$$\begin{aligned}
I_1 &= \lambda n \left(2e^{-\mu \frac{i-i}{n}} - e^{-\mu \frac{i-i-1}{n}} - e^{-\mu \frac{i-i+1}{n}} \right); \\
I_2 &= \frac{\lambda n}{\mu} \left(2e^{-\mu \frac{i-i}{n}} - e^{-\mu \frac{i-i-1}{n}} - e^{-\mu \frac{i-i+1}{n}} \right) - \lambda \left(e^{-\mu \frac{i+1}{n}} - e^{-\mu \frac{i}{n}} \right); \\
I_3 &= \frac{\lambda n}{\mu} \left(-2e^{-\mu \frac{i-i}{n}} + e^{-\mu \frac{i-i-1}{n}} + e^{-\mu \frac{i-i+1}{n}} \right); \\
I_4 &= \frac{\lambda n}{\mu} \left(-2e^{-\mu \frac{i-i}{n}} + e^{-\mu \frac{i-i-1}{n}} + e^{-\mu \frac{i-i+1}{n}} \right) + \lambda \left(e^{-\mu \frac{i+1}{n}} - e^{-\mu \frac{i}{n}} \right).
\end{aligned}$$

Adding up the above in (51) yields the result. \square

Proof of (ii). For all $1 \leq i, j, k \leq n$ we write

$$\begin{aligned}
(52) \quad I_{i,j,k} &:= \int_{\mathbb{R}^2} |u_i^\# u_j^\# u_k^\#| \, d\nu_n^\# \leq \int |(\alpha_{i+1} - \alpha_i)(\alpha_{j+1} - \alpha_j)(\alpha_{k+1} - \alpha_k)| \, d\nu_n^\# \\
&+ \int |(\alpha_{i+1} - \alpha_i)(\alpha_{j+1} - \alpha_j)\mu\beta_k| \, d\nu_n^\# + \int |(\alpha_{j+1} - \alpha_j)(\alpha_{k+1} - \alpha_k)\mu\beta_i| \, d\nu_n^\# \\
&+ \int |(\alpha_{i+1} - \alpha_i)(\alpha_{k+1} - \alpha_k)\mu\beta_j| \, d\nu_n^\# + \int |(\alpha_{i+1} - \alpha_i)\mu^2\beta_j\beta_k| \, d\nu_n^\# \\
&+ \int |(\alpha_{j+1} - \alpha_j)\mu^2\beta_i\beta_k| \, d\nu_n^\# + \int |(\alpha_{k+1} - \alpha_k)\mu^2\beta_i\beta_j| \, d\nu_n^\# \\
&+ \int |\mu^3\beta_i\beta_j\beta_k| \, d\nu_n^\# =: \sum_{l=1}^8 I_{i,j,k}^l.
\end{aligned}$$

It can be easily retrieved that

$$\begin{aligned}
I_{i,i,i}^1 &= n \left(\frac{\lambda}{n} - \frac{\lambda}{\mu} \left(1 - e^{-\frac{\mu}{n}} \right) \left(1 - e^{\mu \frac{i+1}{n}} \right) \right) \leq \frac{\lambda}{\mu}; \\
I_{i,j,k}^1 &= 0, \quad 1 \leq i < j < k \leq n; \\
I_{i,i,k}^1 &= \frac{\lambda n}{\mu} \left(e^{\mu \frac{i+1}{n}} - e^{\mu \frac{i}{n}} \right) \left(e^{-\mu \frac{k}{n}} - e^{-\mu \frac{k+1}{n}} \right) \leq \frac{\lambda}{\mu n}, \quad i = j < k,
\end{aligned}$$

and the other cases can be treated similarly. Also, simple computations show that if $i < j$,

$$\mu \int |(\alpha_{i+1} - \alpha_i)(\alpha_{j+1} - \alpha_j)\beta_k| \, d\nu_n^\# \leq \lambda \left(e^{\mu \frac{i+1}{n}} - e^{\mu \frac{i}{n}} \right) \left(e^{-\mu \frac{j}{n}} - e^{-\mu \frac{j+1}{n}} \right) \leq \frac{\lambda}{n^2},$$

whereas if $i = j$, the above integral is upper bounded by

$$2\lambda \left(2 + e^{-\mu \frac{i+1}{n}} - e^{-\mu \frac{i}{n}} - 2e^{-\frac{\mu}{n}} \right) \leq \frac{2\lambda\mu}{n}.$$

It readily follows that in all cases, $I_{i,j,k}^2, I_{i,j,k}^3$ and $I_{i,j,k}^4$ are less than cn^{-1} for some constant c . Reasoning similarly, we also obtain that for all i, j, k ,

$$\mu^2 \int |(\alpha_{i+1} - \alpha_i)\mu^2\beta_j\beta_k| \, d\nu_n^\# \leq \frac{\mu^2}{n} \left(\frac{\lambda}{n} - \frac{\lambda}{\mu} \left(1 - e^{-\frac{\mu}{n}} \right) \left(1 - e^{\mu \frac{i+1}{n}} \right) \right) \leq \frac{\lambda}{\mu n^2},$$

so that in all cases the $I_{i,j,k}^5, I_{i,j,k}^6$ and $I_{i,j,k}^7$'s are less than cn^{-2} for some c . Finally, observing that for all u, v, w ,

$$\int \int \mathbf{1}_{C_u} \mathbf{1}_{C_v} \mathbf{1}_{C_w} \lambda \mu e^{-\mu y} \, dx \, dy = \frac{\lambda}{\mu} \left(e^{-\mu(\max(u,v,w) - \min(u,v,w))} - e^{-\mu \max(u,v,w)} \right)$$

we can similarly bound $I_{i,j,k}^8$ by a cn^{-2} for all i, j, k . To summarize, all the $I_{i,j,k}$'s are less than cn^{-2} for some c , except for the $I_{i,i,i}^1$'s, $i = 1, \dots, n$, which are bounded

by a constant but are only n in number, and all terms where one index appears twice, which are less than cn^{-1} for some c , but are only n^2 in number. Hence (ii). \square

Proof of (iii). We have for all $1 \leq i \leq n$,

$$(53) \quad \int \int u_i^\# u_i^\# d\nu_n^\# = \int \int \alpha_{i+1} d\nu_n^\# + \int \int \alpha_i d\nu_n^\# - 2 \int \int \alpha_{i+1} \alpha_i d\nu_n^\# \\ + 2\mu \int \int \beta_i \alpha_{i+1} d\nu_n^\# - 2\mu \int \int \beta_i \alpha_i d\nu_n^\# + \mu^2 \int \int \beta_i \beta_i d\nu_n^\# \\ = J_1 + J_2 + J_3 + J_4 + J_5 + J_6,$$

where straightforward calculations show that

$$J_1 = \frac{\lambda n}{\mu} \left(1 - e^{-\mu \frac{i+1}{n}}\right); \quad J_2 = \frac{\lambda n}{\mu} \left(1 - e^{-\mu \frac{i}{n}}\right); \\ J_3 = -2 \frac{\lambda n}{\mu} \left(e^{-\frac{\mu}{n}} - e^{-\mu \frac{i+1}{n}}\right); \quad J_4 = 2 \frac{\lambda n}{\mu} (1 - e^{-\frac{\mu}{n}}) - 2\lambda e^{-\mu \frac{i+1}{n}}; \\ J_5 = -2 \frac{\lambda n}{\mu} (1 - e^{-\frac{\mu}{n}}) - 2 \frac{\lambda n}{\mu} (e^{-\mu \frac{i+1}{n}} - e^{-\mu \frac{i}{n}}); \\ J_6 = \lambda \left(2 + 2e^{-\mu \frac{i+1}{n}} + \frac{2n}{\mu} (e^{-\mu \frac{i+1}{n}} - e^{-\mu \frac{i}{n}} + e^{-\frac{\mu}{n}} - 1)\right).$$

Recalling (40), adding up the J_j 's, $j = 1, \dots, 6$, concludes the proof. \square

REFERENCES

- [1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Academic Press, 2003.
- [2] A. D. Barbour, *Stein's method for diffusion approximations*, Probability Theory and Related Fields **84** (1990), no. 3, 297–322.
- [3] A. D. Barbour and L. H. Y. Chen, *An introduction to Stein's method*, Lecture Notes Series, vol. 4, National University of Singapore, 2005.
- [4] A. A. Borovkov, *Limit laws for queueing processes in multichannel systems*, Sibirsk Mat **8** (1967), no. 5, 983–1004.
- [5] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities*, Oxford University Press, 2013.
- [6] H. Brézis, *Analyse fonctionnelle*, Masson, 1987.
- [7] A. Braverman, J.G. Dai, and J. Feng, *Stein's method for steady-state diffusion approximations: an introduction through the Erlang-A and Erlang-C models*, Stochastic Systems **6** (2017), no. 2, 301–366.
- [8] A. Braverman and J.G. Dai, *Stein's method for steady-state diffusion approximations of M/PH/N+M systems*, Annals of Applied Probability **27** (2017), no. 1, 550–581.
- [9] S. Chatterjee and E. Meckes, *Multivariate normal approximation using exchangeable pairs*, ALEA. Latin American Journal of Probability and Mathematical Statistics **4** (2008), 257–283.
- [10] L. H. Y. Chen, *On the convergence of Poisson binomial to Poisson distributions*, The Annals of Probability **2** (1974), no. 1, 178–180.
- [11] L. Coutin and L. Decreusefond, *Stein's method for Brownian approximations*, Communications on Stochastic Analysis **7** (2013), no. 3, 349–372.
- [12] L. Coutin and L. Decreusefond, *Convergence rate in the Rough Donsker theorem*, arXiv:1707.01269 [math] (2017), arXiv: 1707.01269.
- [13] L. Decreusefond, *Stochastic integration with respect to Volterra processes*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **41** (2005), no. 2, 123–149.
- [14] L. Decreusefond and P. Moyal, *A Functional Central Limit Theorem for the M/GI/ ∞ queue*, Annals of Applied Probability **18** (2008), no. 6, 2156–2178.
- [15] L. Decreusefond and P. Moyal, *Stochastic modeling and analysis of telecom networks*, ISTE Ltd and John Wiley & Sons Inc, 2012.
- [16] L. Decreusefond, M. Schülte, and C. Thäle, *Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry*, Annals of probability **44** (2016), no. 3, 2147–2197.
- [17] R. M. Dudley, *Real analysis and probability*, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002.

- [18] D. Feyel and A. de la Pradelle, *On Fractional Brownian Processes*, Potential Analysis **10** (1999), no. 3, 273–288.
- [19] P. K. Friz and N. B. Victoir, *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*, Cambridge University Press, 2010.
- [20] I. Nourdin and G. Peccati, *Normal Approximations with Malliavin Calculus: From Stein’s Method to Universality*, Cambridge University Press, 2012.
- [21] I. Nourdin and G. Peccati, *Stein’s method on Wiener chaos*, Probability Theory and Related Fields **145** (2009), no. 1-2, 75–118.
- [22] D. Nualart, *The Malliavin Calculus and Related Topics*, vol. 17, SpringerVerlag, 1995.
- [23] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet, *Stein’s method and Normal approximation of Poisson functionals*, The Annals of Probability **38** (2010), no. 2, 443–478.
- [24] N. Privault, *Stochastic analysis in discrete and continuous settings with normal martingales*, Lecture Notes in Mathematics, vol. 1982, Springer-Verlag, Berlin, 2009.
- [25] P. Robert, *Stochastic networks and queues*, Applications of Mathematics, vol. 52, Springer-Verlag, Berlin, 2003.
- [26] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives: theory and applications*, Gordon and Breach Science Publishers, Switzerland ; Philadelphia, Pa., USA, 1993.
- [27] H.-H. Shih, *On Stein’s method for infinite-dimensional Gaussian approximation in abstract Wiener spaces*, Journal of Functional Analysis **261** (2011), no. 5, 1236–1283.
- [28] A. Shwartz and A. Weiss, *Large Deviations For Performance Analysis: Queues, Communication and Computing*, Chapman and Hall/CRC, London ; New York, 1995.
- [29] C. Stein, *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, The Regents of the University of California, 1972.
- [30] A. S. Üstünel, *Analysis on Wiener Space and Applications*, arXiv:1003.1649 **12** (2010), no. 1, 85–90.

LTCI, TELECOM PARISTECH, UNIVERSITÉ PARIS-SACLAY, 75013 PARIS, FRANCE
E-mail address: `eustache.besancon@mines-telecom.fr`

LTCI, TELECOM PARISTECH, UNIVERSITÉ PARIS-SACLAY, 75013 PARIS, FRANCE
E-mail address: `laurent.decreusefond@mines-telecom.fr`

LMAC, UTC - 57 AVENUE DE LANDSHUT, 60203 COMPIÈGNE, FRANCE
E-mail address: `pascal.moyal@utc.fr`