On mean-field $GI/GI/1$ queueing model:
existence, uniqueness, convergence *

A.Yu. Veretennikov†

July 3, 2018

Abstract

A mean-field extension of the queueing system $GI/GI/1$ is considered. The process is constructed as a Markov solution of a martingale problem. Uniqueness in distribution is also established.

1 Introduction

Mean-field approach in the theory of queueing systems allows to take into consideration large interacting ensembles of queues by using the idea of replacing these interactions by a suitable “mean field”. In particular, this approach showed rather fruitful in systems with countable state spaces. In this work we propose a method of constructing a more general extension of the system $GI/GI/1$ – or, more precisely, $GI/GI/1/\infty$ – under certain restrictions on intensities of arrivals and service, which intensities may both depend on the state as well as on the marginal distribution of the process. Existence and weak uniqueness is discussed on the basis of compactness of measures, Skorokhod’s unique probability space Lemma, total variation metric and a Skorokhod–Girsanov’s density of measures theorem for jump processes. The basis for the study in the sections 2.1 and 2.2 is a technique similar to the one developed in a recent preprint on McKean-Vlasov equations joint with Yulia Mishura [13]. Note

*This study has been funded by the Russian Academic Excellence Project ‘5-100’ (section 2.2) and by the RFBR grant 17-01-00633_a (section 2.1).
†School of Mathematics, University of Leeds, Leeds, LS2 9JT, UK; email: a.veretennikov@leeds.ac.uk & National Research University Higher School of Economics, Moscow, Russian Federation, & Institute for Information Transmission Problems, Moscow, Russian Federation
that close results on “propagation of chaos” for various – yet, quite different from ours – mean-field SDE models under various assumptions were earlier established in [4], [11], [14], et al. It is likely that the established results may be useful in the area of mathematical theory of reliability, see [6]. Some earlier results on mean-field queueing models could be found, in particular, in [1], [3], [9], see also further references therein. However, both models and especially methods in the present paper are different.

The paper consists of Introduction, Main section and Appendix. The Main section consists of two subsections related to the first two topics shown in the title, with one theorem in each and with the proof of this theorem; the third on convergence is postponed to further publications (if it is not allowed to change the title). The Appendix contains the statement of Skorokhod’s Lemma about an equivalence of weak convergence of a sequence of processes to a convergence in probability of processes with the same distributions on a unique probability space.

2 Main section

The state space of the process under consideration is the union

$$\mathcal{X} := (0, x) \cup \bigcup_{n=1}^{\infty}(n, x, y), \quad x, y \geq 0.$$  

The meaning of $n$ here is the number of “customers” in the system; the value $x$ stands for the elapsed time from the last arrival, while $y$ signifies the elapsed time of the current service. There is only one server which works without breaks (if there is at least one customer in the system) and it is always in a working state. All newly arrived customers stand in a queue of the infinite capacity, and for simplicity only we assume the FIFO discipline of service. It is assumed that at any time $t$ at any state $X = (n, x, y)$ (or $X = (0, x)$ for $n = 0$) there are intensities of service $\Lambda^{-}[t, X_t, \mu_t]$ and arrivals $\Lambda^{+}[t, X_t, \mu_t]$, where $\mu_t$ is the distribution of the random variable $X_t$ itself. Note that occasionally we will be using notation $(0, x, y)$ where $y$ is a “false” variable, i.e., we identify all such triples with any $y$ with a couple $(0, x)$. The process is piecewise–linear Markov (PLMP, see [5]), which simply means that between any jumps the continuous components $-(x, y)$ if $n > 0$, or just $x$ if $n = 0$ – grow linearly with rate 1, while the discrete component $n$ remains unchanged.
The assumptions:

(A1) There are Borel measurable, non-negative and bounded functions $\lambda^+(t, X, Y)$ and $\lambda^-(t, X, Y)$.

(A2) 

$$\Lambda^\pm [t, X, \mu] = \int \lambda^\pm (t, X, Y) \mu(dY)$$

(NB: Automatically, both $\Lambda^\pm$ are Borel functions of $(t, X)$.)

(A3) The functions $\lambda^\pm (t, X, Y)$ are continuous in all variables.

(A4) The functions $\lambda^\pm (t, X, Y)$ are uniformly bounded away from zero except for $\lambda^-(t, (0, x), (n, y, y')) = 0$, for any $x, y, y' \geq 0$ (no jump down from a state with zero customers).

Let us emphasize that neither Lipschitz nor any other regularity of the intensities $\lambda^\pm$ is assumed, except for continuity in (A3).

Note that both intensities $\Lambda^\pm$ may include additional (non-negative) terms not depending on the measure, say, $\lambda^\pm_0 (t, X)$; in particular, this may be helpful so as to justify the assumption (A4), as the terms $\lambda^\pm_0 (t, X)$ can be easily assumed bounded away from zero uniformly and independently of $N$.

For $X \in \mathcal{X}$ let us denote

$$X^+ := (n + 1, 0, y), \quad \text{for } X = (n, x, y),$$

$$X^- := (n - 1, x, 0), \quad \text{for } X = (n, x, y), \quad n \geq 1.$$

Naturally, $X^-$ is not defined for $X = (0, x)$.

2.1 Existence

The initial value $X_0$ of the process may be distributed, which distribution is denoted by $\mu_0$ (in particular, $\mu_0$ may be a delta-measure concentrated at one point).

**Theorem 1** Let the assumptions (A1)–(A3) be satisfied. Then for any initial distribution $\mu_0$ on $\mathcal{X}$, on some probability space there exists a Markov process $(X_t, t \geq 0)$
with marginal distributions $\mu_t$ and intensities $\Lambda[t, X_t, \mu_t], H[t, X_t, \mu_t]$; in other words, such that for any bounded continuous function $g(X)$ with bounded continuous derivatives in $(x, y)$, the expression

$$M_t := g(X_t) - g(X_0) - \int_0^t L(s, X_s, \mu_s)g(X_s) \, ds$$

(1)

is a martingale, where for $X = (n, x, y), X' = (n', x', y')$, $n \geq 0, t \geq 0$,

$$L(t, X', \mu)g(X) := \Lambda^+[t, X', \mu](g(X^+) - g(X))$$

$$+ 1(n > 0)\Lambda^-[t, X', \mu](g(X^-) - g(X))$$

$$+ \frac{\partial}{\partial x}g(n, x, y) + 1(n > 0)\frac{\partial}{\partial y}g(n, x, y).$$

Moreover, for any given measure-valued function $(\mu_s, s \geq 0)$ in $L(s, X_s, \mu_s)$, the martingale problem (see [8]) has a weakly unique solution.

Note that the generator of the Markov process is, of course, $L(t, X, \mu)$; different variables $X$ and $X'$ are needed only for the convenience of the proof. Equivalently, Dynkin’s identity holds true for any function $g(X)$ from the same class,

$$\mathbb{E}_{0, X_0}g(X_t) = g(X_0) + \mathbb{E}_{0, X_0}\int_0^t L(s, X_s, \mu_s)g(X_s) \, ds.$$  

(2)

Moreover, equivalently, for any $0 \leq t_1 < t_2 \ldots < t_{m+1}$, and for any Borel bounded functions $\phi_k(X), X \in \mathcal{X}$,

$$\mathbb{E}_{0, X_0}\left(g(X_{t_{m+1}}) - g(X_{t_m}) - \int_{t_m}^{t_{m+1}} L(s, X_s, \mu_s)g(X_s) \, ds\right) \prod_{k=1}^m \phi_k(X_{t_k}) = 0.$$  

(3)

The latter formula may be called another version of Dynkin’s identity and will be the basis for establishing existence. With a bit of abuse of the standard terminology, (3) may also be called a martingale problem. Note, however, that weak uniqueness (= uniqueness in distribution) in this Theorem given $(\mu_s, s \geq 0)$ does not mean a total uniqueness in distribution of the process under construction because there is no claim of uniqueness of $(\mu_s, s \geq 0)$, not even talking about a distribution in the space of trajectories.
Proof of Theorem 1. For any $n \geq 1$ consider a process $(X^n_t)$, with initial data $X^n_0, \delta = X_0$ and intensities of jumps up and down, respectively,

$$
\Lambda^+ [t, X^n_{t-1/n}, \mu^n_{t-1/n}], \quad \Lambda^- [t, X^n_{t-1/n}, \mu^n_{t-1/n}],
$$

where $X^n_t$ with $t < 0$ is understood as $X^n_0$, and similarly for $\mu^n_t$. The process $(X^n_t)$ for each $n$ are constructed by induction successfully on the intervals $[0, 1/n], [1/n, 2/n], \text{etc}$. Due to the boundedness assumption on both intensities, there is no blow up and the processes for any $n$ are defined for any $t \geq 0$ as càdlàg pure jump processes. Moreover, for any $t$ probability of jump exactly at time $t$ for any $X^n_t$ equals zero.

The processes $(X^n_t, t \geq 0)$ for $n \geq 1$ being constructed, let us introduce on some probability space independent equivalent processes $(\xi^n_t, t \geq 0)$; let $\mathbb{E}'$ stand in all cases for the integration with respect to the third variable, e.g.,

$$
\mathbb{E}' \lambda^\pm(t, X^n_t, \xi^n_t) := \int \lambda^\pm(t, X^n_t, Y) \mu^n_t(dY).
$$

It can be checked that the assumptions of the Lemma 1 from the Appendix are satisfied. Hence, on some new probability space there are equivalent – and, hence, Markov with the same generators – processes $(\tilde{X}^n_t, \tilde{\xi}^n_t)$, and a limiting pair $(\tilde{X}_t, \tilde{\xi}_t)$ such that for some subsequence $(\tilde{X}^n_{t'}, \tilde{\xi}^n_{t'}) \xrightarrow{\mathbb{P}} (\tilde{X}_t, \tilde{\xi}_t)$, $n' \to \infty$, for each $t$. It follows due to the boundedness of all intensities that the limiting process $(\tilde{X}_t, \tilde{\xi}_t)$ is also stochastically continuous. More than that, with probability one the pair $(\tilde{X}_t, \tilde{\xi}_t)$ is a pure jump process with a finite number of jumps on any bounded interval. Moreover, the property $\lim_{h \downarrow 0} \sup_n \sup_{t,s \leq T, |t-s| \leq h} \mathbb{P}(|\tilde{X}_n^t - \tilde{X}_n^s| > \epsilon) = 0$ implies that for any $\epsilon > 0$ there is a following convergence in probability,

$$
\tilde{X}^n_{t-1/n'} \xrightarrow{\mathbb{P}} \tilde{X}_t, \quad n' \to \infty.
$$

The analogue of Dynkin’s formula (3) for the pair $(\tilde{X}^n_{t'}, \tilde{\xi}^n_{t'})$ reads,

$$
\mathbb{E}_{0, \tilde{X}^n_0} \left[ \left( g(\tilde{X}^n_{t_{m+1}}) - g(\tilde{X}^n_{t_m}) - \int_{t_m}^{t_{m+1}} \mathbb{E}' L(s, \tilde{X}^n_{s-1/n'}, \tilde{\xi}^n_{s-1/n'}) g(\tilde{X}^n_s) ds \right) \times \prod_{k=1}^{m} \phi_k(\tilde{X}^n_{t_k}) \right] = 0,
$$

(4)
which formula follows straightforwardly from the “complete expectation” arguments (cf., for example, [16]).

By continuity of \(\lambda\) and \(h\), and due to the stochastic continuity of the processes \(\tilde{X}\) and \(\tilde{\xi}\), and since all integrand expressions are bounded, and by virtue of Lebesgue’s bounded convergence Theorem, we obtain from (1) in the limit with continuous bounded functions \((\phi_k)\),

\[
E_{0,X_0} \left( g(\tilde{X}_{t_{m+1}}) - g(\tilde{X}_t) - \int_{t_m}^{t_{m+1}} \mathbb{E}'L(s, \tilde{X}_s, \tilde{\xi}_s) g(\tilde{X}_s) \, ds \right) \prod_{k=1}^{m} \phi_k(\tilde{X}_{t_k}) = 0. \tag{5}
\]

Since the distribution of the random variable \(\tilde{\xi}_t\) is the same as the one of \(\tilde{X}_t\) – let us denote it by \(\tilde{\mu}_t\) – then (5) can be equivalently written as

\[
E_{0,X_0} \left( g(\tilde{X}_{t_{m+1}}) - g(\tilde{X}_t) - \int_{t_m}^{t_{m+1}} Lg(s, \tilde{X}_s, \tilde{\mu}_s) \, ds \right) \prod_{k=1}^{m} \phi_k(\tilde{X}_{t_k}) = 0. \tag{6}
\]

Due to the properties of measures on \(\mathbb{R}^d\), the formula (6) holds true for any Borel bounded functions \((\phi_k)\), too. Due to [7], solution of the “martingale problem” (6) – or, more precisely, of the martingale problem

\[
M_t := g(\tilde{X}_{t_{m+1}}) - g(\tilde{X}_t) - \int_{t_m}^{t_{m+1}} Lg(s, \tilde{X}_s, \tilde{\mu}_s) \, ds \quad \text{is a martingale}, \tag{7}
\]

with a given family of marginal measures \((\tilde{\mu}_s, s \geq 0)\) is unique. Hence, according to [10], or [8, Theorem 4.4.2] the limiting process \(\tilde{X}\) is Markov. The form of its generator with the required intensities \(\Lambda^\pm\) follows from (6). This finishes the proof of the Theorem [11].

\section{2.2 Weak uniqueness}

Emphasize that we will use essentially boundedness of all intensities and the condition that they are (uniformly) bounded away from zero. While it is clear that the boundedness from above may be relaxed for the purpose of establishing existence – e.g., under Lyapunov type conditions, or under a linear growth, or otherwise, – and that boundedness away from zero is not required for the existence at all, yet for the uniqueness both boundedness from above and from below seems essential (although also could be, apparently, slightly relaxed). On the other hand, continuity of the intensities in this section is not necessary and they are not assumed.
Theorem 2 Let the assumptions (A1)–(A2) and (A4) be satisfied. Then, for any fixed distribution \( \mathcal{L}(X_0) \), there exists no more than one distribution of the process \((X_t, t \geq 0)\) with required intensities \( \Lambda[t, x, \mu_t] \) and \( H[t, x, \mu_t] \).

Recall that no Lipschitz assumptions on the intensities are assumed. In the calculus the total variation metric will be used.

In the proof we will re-denote all intensities by \( \lambda^\pm \) and \( \Lambda^\pm \) with \( \pm \) signs; those with “plus” will correspond to jumps up (arrivals) while those with “minus” – to jumps down (service). Let \( \bar{\Lambda}[t, X, \mu] = \Lambda^+[t, X, \mu] + \Lambda^-[t, X, \mu] \).

Proof of Theorem 2 is based on Skorokhod–Girsanov’s change of measure formula for jump processes (see, e.g., [12]). Suppose there are two solutions, \((X^1_t, \mu^1_t)\) and \((X^2_t, \mu^2_t)\). Denote by \( \Omega_n \) the event that the trajectory \( X \) has precisely \( n \) jumps on \([0, T]\). Recall – see, e.g., [12], [15] – that on the interval of time \([0, T]\) the density of one distribution with respect to the other – we denote them by \( \mathbb{P}^{\mu_i}, i = 1, 2 \) – on a typical trajectory \( \omega = (t^\pm_1, \ldots, t^\pm_n) \) with overall \( n \geq 0 \) jumps up \( (t^+_i) \) or down \( (t^-_j) \) reads,

\[
\rho_T := \frac{d\mathbb{P}^n}{d\mathbb{P}^1}(\omega) |_{\Omega_n} = \prod_{i=1}^{n} \frac{\Lambda^+[t^+_i, X_{t^+_i}, \mu^1_{t^+_i}]}{\Lambda^-[t^+_i, X_{t^+_i}, \mu^1_{t^+_i}]} \exp \left( - \int_0^T (\bar{\Lambda}[t, X_t, \mu^2_t] - \bar{\Lambda}[t, X_t, \mu^1_t]) \, dt \right),
\]

where \( X = (X_s, 0 \leq s \leq T) \) and \( (t^\pm_i) \) are the moments of jumps of the trajectory \( X \), up or down, respectively; we keep the same sign at \( \Lambda \), too, i.e., \( \Lambda^+[t^+, \ldots] \) or, respectively, \( \Lambda^-[t^-, \ldots] \). The usual convention \( \prod_{i=1}^0 \ldots = 1 \) is assumed. Note that, of course, the number of jumps \( n \) is random – i.e., it is a function of the trajectory – but in any case it is almost surely finite due to the boundedness of the intensities.
Note also that the expression $\rho_T$ above is a probability density. We have,

$$
\mathbb{E}^{\mu_1} \prod_{i=1}^{n} \frac{\Lambda^\pm[t^\pm_i, X_{t_i}, \mu^2_{t_i}]}{\Lambda^\pm[t^\pm_i, X_{t_i}, \mu^1_{t_i}]} \exp \left( -\int_0^T (\bar{\Lambda}[t, X_t, \mu^2_{t}] - \bar{\Lambda}[t, X_t, \mu^1_{t}]) \, dt \right) 
$$

$$
= \sum_{n=0}^{\infty} \mathbb{E}^{\mu_1}(\Omega_n) \prod_{i=1}^{n} \frac{\Lambda^\pm[t^\pm_i, X_{t_i}, \mu^2_{t_i}]}{\Lambda^\pm[t^\pm_i, X_{t_i}, \mu^1_{t_i}]} \exp \left( -\int_0^T (\bar{\Lambda}[t, X_t, \mu^2_{t}] - \bar{\Lambda}[t, X_t, \mu^1_{t}]) \, dt \right) 
$$

$$
= \sum_{n=0}^{\infty} \mathbb{E}^{\mu_2} \int \cdots \int \prod_{i=1}^{n} \Lambda^\pm[t^\pm_i, X_{t_i}, \mu^2_{t_i}] \exp \left( -\int_0^T \bar{\Lambda}[t, X_t, \mu^2_{t}] \, dt \right) \prod_{i=1}^{n} dt_i 
$$

$$
= \sum_{n=0}^{\infty} \mathbb{P}^{\mu_2}(\Omega_n) = 1. 
$$

Note that given the initial state $X_0$, the value without expectation $\mathbb{E}^{\mu_2}$ here equals, actually,

$$
\sum_{n=0}^{\infty} \int \cdots \int \prod_{i=1}^{n} \Lambda^\pm[t^\pm_i, X_{t_i}, \mu^2_{t_i}] \exp \left( -\int_0^T \bar{\Lambda}[t, X_t, \mu^2_{t}] \, dt \right) \prod_{i=1}^{n} dt_i,
$$

which itself equals identically one, while expectation $\mathbb{E}^{\mu_2}$ relates to integration of each term over $X_0$ if it is random. It is worthwhile to recall that the rule of the evolution of the trajectory $X$ between the moments of jumps $t_i$ is deterministic and linear with rate one for the continuous components, and the discrete component does not change between any two consequent jumps.

Now, we want to estimate the distance in total variation between two probability measures in the space of trajectories, $\mu^1_{[0,T]}$ and $\mu^2_{[0,T]}$ and then to use the inequality that the distance between the marginals of any two measures does not exceed the distance of the measures themselves,

$$
\varphi_T := \|\mu^1_{T} - \mu^2_{T}\|_{TV} \leq \|\mu^1_{[0,T]} - \mu^2_{[0,T]}\|_{TV} = 2 - 2\mathbb{E}^{\mu_1} (\rho_T \wedge 1) =: \psi_T.
$$

Now, the idea is to estimate the right hand side in the last term via $\varphi$ and, hence, to show that, at least, for small values of $T > 0$ this value equals zero. If this is realized, then the claim that $\varphi_t = 0$ for $t \leq T$, $t \leq 2T$, etc., and, eventually, for all $t \geq 0$ would follow by induction. In fact, we will be able to estimate the right hand side via another expression with $\psi_T$ itself. Note, by the way, that although normally marginal distributions of any process may not determine the distribution
in the space of trajectories, in our case with intensities it is, of course, the case which follows from [7], as mentioned already in the proof of the Theorem 1.

The first goal is to find a suitable lower bound for the value \( \mathbb{E}^{\mu_1}(\rho_T \wedge 1) \) from below. Let us split it as follows:

\[
\mathbb{E}^{\mu_1}(\rho_T \wedge 1) = \sum_{n=0}^{\infty} \mathbb{E}^{\mu_1}(\Omega_n)(\rho_T \wedge 1).
\]

Further, we have for \( n = 0 \),

\[
\mathbb{E}^{\mu_1}(\Omega_0)(\rho_T \wedge 1) = \mathbb{E}^{\mu_1}(\Omega_0) \exp\left(-\int_0^T (\bar{\Lambda}[t, X_t, \mu^2_t] - \bar{\Lambda}[t, X_t, \mu^1_t])\, dt\right) \wedge 1 \\
\geq \exp\left(-\int_0^T \|\lambda\||\mu^2_t - \mu^1_t|_{TV}\, dt\right)\mathbb{E}^{\mu_1}(\Omega_0) \\
\geq \exp(-\|\lambda\| TV_T)\mathbb{E}^{\mu_1}(\Omega_0).
\]

All norms like \( \|\lambda\| \) are sup-norms (except for the total variation norm, which is always shown explicitly). We used the fact that \( |\bar{\Lambda}[t, X, \mu]| \leq \|\lambda\| \), and that

\[
|\bar{\Lambda}[t, X_t, \mu^2_t] - \bar{\Lambda}[t, X_t, \mu^1_t]| \leq \|\lambda\|\|\mu^2_{[0,t]} - \mu^1_{[0,t]}|_{TV} \leq \|\lambda\|\|\mu^1_{[0,T]} - \mu^2_{[0,T]}|_{TV}, \ 0 \leq t \leq T.
\]

Similarly for \( n \geq 1 \), with a notation \( \tilde{\Lambda}^\pm[t^\pm, \ldots] := \ln \Lambda^\pm[t^\pm, \ldots] \),

\[
\mathbb{E}^{\mu_1}(\Omega_n)(\rho_T \wedge 1) = \mathbb{E}^{\mu_1}(\Omega_n) \left\{ \prod_{i=1}^{n} \frac{\Lambda^\pm[t^\pm_i, X_{t_i}, \mu^2_{t_i}]}{\Lambda^\pm[t^\pm_i, X_{t_i}, \mu^1_{t_i}]} \times \right.
\left. \exp\left(-\int_0^T (\bar{\Lambda}[t, X_t, \mu^2_t] - \bar{\Lambda}[t, X_t, \mu^1_t])\, dt\right) \right\} \wedge 1 \\
\geq \mathbb{E}^{\mu_1}(\Omega_n) \exp\left(-\sum_{i=1}^{n} |\bar{\Lambda}[t^\pm_i, X_{t_i}, \mu^2_{t_i}] - \bar{\Lambda}[t^\pm_i, X_{t_i}, \mu^1_{t_i}]| \right) \\
\times \exp\left(-\int_0^T |\bar{\Lambda}[t, X_t, \mu^2_t] - \bar{\Lambda}[t, X_t, \mu^1_t]|\, dt\right).
\]
Minimum with 1 here was dropped after all multipliers were estimated from below by the values less than one. Further, since the derivative of \( \ln x \) is bounded on any interval \( 0 < a \leq x \leq b \), say, by a constant \( K \), we have with \( a = \inf \lambda(\ldots) =: \lambda \) and \( b = \|\lambda\| \),

\[
|\tilde{\Lambda}^\pm[t_i^\pm, X_t, \mu_{t_i}^2] - \tilde{\Lambda}[t_i^\pm, X_t, \mu_{t_i}^1]| \leq K|\Lambda^\pm[t_i^\pm, X_t, \mu_{t_i}^2] - \Lambda^\pm[t_i^\pm, X_t, \mu_{t_i}^1]|
\]

\[
\leq K\|\Lambda\|\|\mu_{t_i}^2 - \mu_{t_i}^1\|_{TV}.
\]

Hence,

\[
\mathbb{E}^{\mu_1} 1(\Omega_n) (\rho_T \wedge 1)
\]

\[
\geq \mathbb{E}^{\mu_1} 1(\Omega_n) \exp(-\sum_{i=1}^n K\|\Lambda\|\|\mu_{t_i}^2 - \mu_{t_i}^1\|_{TV})
\]

\[
\times \exp(-\int_0^T \|\lambda\|\|\mu_{t_i}^2 - \mu_{t_i}^1\|_{TV} dt).
\]

(Here by definition \( t_0 = 0 \).) Here \( \lambda > 0 \) stands for the infimum of all intensities, which infimum is positive by the assumption. Thus, using the bound \( 1 - \exp(-a) \leq a \) and estimates \( \mathbb{E}^{\mu_1} 1(\Omega_0) \leq \exp(-\lambda T) \) and \( \mathbb{E}^{\mu_1} 1(\Omega_n) \leq \frac{(\|\lambda\| T)^n}{n!} \exp(-\lambda T), n \geq 1, \) we
\[
\frac{1}{2} \psi_T = 1 - \sum_n \mathbb{E}^{\lambda^1_1}(\Omega_n) (\rho_T \wedge 1)
\]

\[
\leq (1 - \exp(-\|\lambda\| T\psi_T)) \mathbb{E}^{\lambda^1_1}(\Omega_0)
\]

\[
+ \sum_{n=1}^\infty \mathbb{E}^{\lambda^1_1}(\Omega_n) \left(1 - \exp\left(-K \sum_{i=1}^n \|\Lambda\| \|\mu^2_t - \mu^1_t\|_{TV}\right) \times \exp\left(-\int_0^T \|\lambda\| \|\mu^2_t - \mu^1_t\|_{TV} dt\right)\right)
\]

\[
\leq \|\lambda\| T\psi_T \mathbb{E}^{\mu^1_1}(\Omega_0)
\]

\[
+ \sum_{n=1}^\infty \mathbb{E}^{\mu^1_1}(\Omega_n) (1 - \exp (-nK\|\lambda\| \psi_T - T\|\lambda\| \psi_T))
\]

\[
\leq \psi_T \exp(-\lambda T) \left(\|\lambda\| T + \sum_{n=1}^\infty (nK\|\Lambda\| + T\|\lambda\|) \frac{\|\lambda\| T^n}{n!}\right)
\]

\[
= T\psi_T \exp(-\lambda T) \left(\|\lambda\| + \sum_{n=0}^\infty ((n + 1)K\|\Lambda\| + T\|\lambda\|) \frac{\|\lambda\|^{n+1} T^n}{(n + 1)!}\right).
\]

The series in the right hand side here converges and does not exceed some constant, say, \(C > 0\), if \(T \leq 1\). Hence, overall, we obtain,

\[
0 \leq \frac{1}{2} \psi_T \leq C T\psi_T, \quad T \leq 1.
\]

This implies that

\[
\psi_T = 0, \quad T < (2C)^{-1} \wedge 1,
\]

and, therefore, also

\[
\phi_T = 0, \quad T < (2C)^{-1} \wedge 1,
\]

as required. In other words, we have shown that the two marginal measures \(\mu^1_t\) and \(\mu^2_t\) coincide for all \(t < (2C)^{-1} \wedge 1\).

Further, note the constant \(C\) in this calculus does not depend on the initial distribution of the process. Hence, using the Markov property of the process and
repeating the same arguments on \([T, 2T]\), \([2T, 3T]\), etc., by induction we conclude that
\[ \psi_t = 0, \quad t \geq 0, \]
and, therefore, also
\[ \varphi_t = 0, \quad t \geq 0, \]
as required. So, the two measures \(\mu^1\) and \(\mu^2\) on the space of trajectories are equal. The Theorem 2 is proved.

3 Appendix

The following celebrated Lemma is stated for the convenience of the reader.

**Lemma 1 (Skorokhod [15, Ch.1, §6])** Let \(\xi^n_t\) \((t \geq 0, n = 0, 1, \ldots)\) be some \(d\)-dimensional stochastic processes defined on some probability space and let for any \(T > 0, \epsilon > 0\) the following hold true:

\[
\lim_{c \to \infty} \sup_n \sup_{t \leq T} \mathbb{P}(|\xi^n_t| > c) = 0,
\]

\[
\lim_{h \downarrow 0} \sup_n \sup_{t,s \leq T; |t-s| \leq h} \mathbb{P}(|\xi^n_t - \xi^n_s| > \epsilon) = 0.
\]

Then there exists a subsequence \(n' \to \infty\) and a new probability can be constructed with processes \(\tilde{\xi}^{n'}_t, t \geq 0\) and \(\tilde{\xi}_t, t \geq 0\), such that all finite-dimensional distributions of \(\tilde{\xi}^{n'}\) coincide with those of \(\xi^n\) and such that for any \(\epsilon > 0\) and all \(t \geq 0\),

\[
\mathbb{P}(|\tilde{\xi}^{n'}_t - \tilde{\xi}_t| > \epsilon) \to 0, \quad n' \to \infty.
\]

Acknowledgements

The techniques used in this paper were stimulated by the methods developed in a long-term joint work on formally quite different McKean-Vlasov SDE equations with Yu. Mishura, as well as in fruitful discussions of the author on the same subject with D. Siska, and L. Szpruch. S. Pirogov, A. Rybko, and G. Zverkina helped to find some (quite a few) technicalities to be corrected in the earlier versions of the text. The author is sincerely thankful to all these colleagues. The deepest gratitude is to Professor Alexander Dmitrievich Solovyev (06.09.1927 – 06.04.2001) who was the author’s supervisor at BSc and MSc programmes at University.
References

[1] Baccelli, F., Rybko, A.N., Shlosman, S.B.: Queueing networks with mobile servers: the mean-field approach, Problems Inform. Transmission, 52(2), 178-199 (2016)

[2] Billingsley, P.: Convergence of Probability Measures. New York, NY: John Wiley & Sons (1999)

[3] Borovkov, K.A.: Propagation of chaos for queueing networks, Theory of Probability and its Applications, 42:3, 385-394 (1998)

[4] Dawson, D.A.: Critical dynamics and fluctuations for a mean-field model of cooperative behavior. Journal of Statistical Physics. 31(1), 29-85 (1983)

[5] Gnedenko, B.V., Kovalenko, I.N.: Introduction to Queueing Theory, 2nd ed., Birkhäuser, Boston et al. (1989)

[6] Gnedenko, B.V., Belyaev, Yu.K., Solovyev, A.D.: Mathematical Methods in Reliability Theory, Academic Press, New York (1969)

[7] Davis, M.H.A.: Piecewise–Deterministic Markov Processes: A General Class of Non–Diffusion Stochastic Models. J. Royal Stat. Soc. Ser. B (Methodological). 46(3), 353–388 (1984)

[8] Ethier, S.N., Kurtz, T.G.: Markov processes: Characterisation and Convergence. Wiley, Hoboken, N.J. (2005)

[9] Karpelevich, F.I., Rybko, A.N.: Asymptotic Behavior of the Thermodynamical Limit for Symmetric Closed Queueing Networks, Problems Inform. Transmission, 36(2), 154-179 (2000)

[10] Krylov, N.V.: On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes. Mathematics of the USSR - Izvestija, 7(3), 691-709 (1973)

[11] Léonard, C.: Une loi des grands nombres pour des systèmes de diffusions avec interactions et à coefficients non bornés. AIHP, sec. B. 22(2), 237-262 (1986)

[12] Liptser, R.Sh., Shiryaev, A.N.: Theory of martingales, Kluwer Academic Publishers Group, Dordrecht (1989)
[13] Mishura, Yu.S., Veretennikov, A.Yu.: Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations, https://arxiv.org/abs/1603.02212

[14] Oelschläger, K.: A martingale approach to the law of large numbers for weakly interacting stochastic processes. Ann. Prob. 12(2), 458-479 (1984)

[15] Skorokhod, A.V.: Studies in the theory of random processes. Addison-Wesley, Reading, Mass. (1965)

[16] Veretennikov, A.Yu., Zverkina, G.A.: Simple Proof of Dynkins Formula for Single-Server Systems and Polynomial Convergence Rates, Markov Processes Relat. Fields. 20, 479-504 (2014)