

On a toy model of interacting neurons

Nicolas Fournier^a and Eva Löcherbach^b

^aLaboratoire de Probabilités et Modèles aléatoires, UMR 7599, UPMC, Case 188, 4 place Jussieu,
F-75252 Paris Cedex 5, France. E-mail: nicolas.fournier@upmc.fr

^bCNRS UMR 8088, Département de Mathématiques, Université de Cergy-Pontoise, 2 avenue Adolphe Chauvin,
95302 Cergy-Pontoise Cedex, France. E-mail: eva.loecherbach@u-cergy.fr

Received 13 October 2014; revised 3 May 2015; accepted 22 July 2015

Abstract. We continue the study of a stochastic system of interacting neurons introduced in De Masi, Galves, Löcherbach and Presutti (*J. Stat. Phys.* **158** (2015) 866–902). The system consists of N neurons, each spiking randomly with rate depending on its membrane potential. At its spiking time, the neuron potential is reset to 0 and all other neurons receive an additional amount $1/N$ of potential. Moreover, electrical synapses induce a deterministic drift of the system towards its average potential. We prove propagation of chaos of the system, as $N \rightarrow \infty$, to a limit nonlinear jumping stochastic differential equation. We consequently improve on the results of (*J. Stat. Phys.* **158** (2015) 866–902), since (i) we remove the compact support condition on the initial datum, (ii) we get a rate of convergence in $1/\sqrt{N}$. Finally, we study the limit equation: we describe the shape of its time-marginals, we prove the existence of a unique nontrivial invariant distribution, we show that the trivial invariant distribution is not attractive, and in a special case, we establish the convergence to equilibrium.

Résumé. Cet article continue l'étude du système stochastique de neurones en interaction introduit par De Masi, Galves, Löcherbach et Presutti (*J. Stat. Phys.* **158** (2015) 866–902). Le système est composé de N neurones. Chaque neurone décharge un potentiel d'action à des instants aléatoires, à un taux qui dépend de son potentiel de membrane. Ce potentiel est alors remis à 0, et tous les autres neurones reçoivent une charge supplémentaire de $1/N$. De plus, des synapses électriques induisent une dérive déterministe qui attire le système vers sa valeur moyenne. Nous établissons la propriété de propagation du chaos lorsque $N \rightarrow \infty$, vers la solution d'une équation différentielle stochastique non-linéaire à sauts. Nous améliorons les résultats obtenus dans (*J. Stat. Phys.* **158** (2015) 866–902) puisque (i) nous levons la condition de support compact imposée aux données initiales, (ii) nous obtenons une vitesse de convergence en $1/\sqrt{N}$. Enfin, nous proposons une étude de l'équation limite : nous décrivons la forme de ses lois marginales (en temps), nous démontrons l'existence d'une unique loi invariante non-triviale et montrons que la mesure invariante triviale n'est pas attractive. Enfin, nous obtenons la convergence vers l'équilibre dans un cas particulier.

MSC: 60K35; 60G55; 60F17

Keywords: Piecewise deterministic Markov processes; Mean-field interaction; Biological neural nets; Interacting particle systems; Nonlinear stochastic differential equations

1. Introduction and main results

1.1. The model

We consider, for each $N \geq 1$, a family of i.i.d. Poisson measures $(\mathbf{N}^i(ds, dz))_{i=1, \dots, N}$ on $\mathbb{R}_+ \times \mathbb{R}_+$ having intensity measure $ds dz$, as well as a family $(X_0^{N,i})_{i=1, \dots, N}$ of \mathbb{R}_+ -valued random variables independent of the Poisson measures. The object of this paper is to study the Markov process $X_t^N = (X_t^{N,1}, \dots, X_t^{N,N})$ taking values in \mathbb{R}_+^N and

solving, for $i = 1, \dots, N$, for $t \geq 0$,

$$\begin{aligned}
 X_t^{N,i} &= X_0^{N,i} - \lambda \int_0^t (X_s^{N,i} - \bar{X}_s^N) ds - \int_0^t \int_0^\infty X_{s-}^{N,i} \mathbf{1}_{\{z \leq f(X_{s-}^{N,i})\}} \mathbf{N}^i(ds, dz) \\
 &\quad + \frac{1}{N} \sum_{j \neq i} \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f(X_{s-}^{N,j})\}} \mathbf{N}^j(ds, dz),
 \end{aligned}
 \tag{1}$$

where $\bar{X}_t^N = N^{-1} \sum_{j=1}^N X_t^{N,j}$. The coefficients of this system are $\lambda \geq 0$ and a function $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$ satisfying (at least) the following assumption.

Assumption 1. f is nondecreasing, $f(0) = 0$, $f(x) > 0$ for all $x > 0$, $\lim_{\infty} f = \infty$ and $f \in C^1(\mathbb{R}_+)$.

Proposition 2. Grant Assumption 1, let $\lambda \geq 0$ and $N \geq 1$ be fixed. Then there exists a unique càdlàg adapted strong solution $(X_t^N)_{t \geq 0}$ to (1) taking values in \mathbb{R}_+^N .

1.2. Formal description and goals

This paper continues a study started in De Masi, Galves, Löcherbach and Presutti [6]. The particle system (1) is the model of interacting neurons considered in [6], inspired by a work of Galves and Löcherbach [13]. The system is made of N neurons. Each $X_t^{N,i}$ models the membrane potential at time t of the i th neuron. Interactions between neurons are due to two types of synapses, chemical and electrical synapses. Chemical synapses are characterized through *spiking* of the neurons, i.e. a fast trans-membrane current. Spiking occurs randomly following a point process of rate $f(x)$ for a neuron of which the membrane potential equals x . At its spiking time, the membrane potential of the neuron is reset to a resting potential which we choose to be equal to 0. At the same time, the action of the chemical synapses induces an increase of the membrane potentials of the other neurons: they receive an additional amount $1/N$ of potential. Our model does not take into account a refractory period. Electrical synapses, which are due to gap junctions, work constantly over time and tend to synchronize the membrane potentials of the neurons. Such electrical synapses are typical for systems requiring fast responses to stimuli, often found in animals. They induce a constant drift of the system towards the average membrane potential of the system, at speed λ .

It is well known that neuronal interactions can exhibit very complicated interaction structures. Our model only takes into account the average effect of the interactions. We are thus working with a toy model where interactions are of *mean-field* type.

Concerning f , we think of functions of the type $f(x) = (x/x_0)^\xi$ with ξ large and some *soft threshold* x_0 . In this case, for x the membrane potential of the neuron, spiking occurs at very low rate if $x < x_0$ and with very high rate if $x > x_0$. Note that in the biological literature it is often assumed that spiking occurs when the membrane potential reaches a fixed threshold x_0 , which would correspond to $f(x) = \infty \mathbf{1}_{[x_0, \infty)}$. However, a well-defined fixed threshold does not seem to exist in *in vivo* neurons, see e.g. Jahn, Berg, Hounsgaard and Ditlevsen [18] who propose a statistical study showing that a point process model in which the jump intensity depends on the membrane potential is well adapted. We therefore propose a smooth firing rate depending on the membrane potential of the form $f(x) = cx^\xi$ with ξ quite large and $c > 0$.

We are interested in the evolution of a large system of neurons, i.e. in the limit $N \rightarrow \infty$. We prove a weak law of large numbers for the empirical measure of the system (*propagation of chaos*): we show that the empirical distribution of the system becomes deterministic as $N \rightarrow \infty$ and tends to the law of a limit process $(Y_t)_{t \geq 0}$ which solves a *nonlinear* jumping SDE.

Such a result has already been achieved in [6] in the case of a *compact support*, i.e. when the initial conditions $X_0^{N,i}$ are uniformly bounded. It is then possible to control the evolution of the support of the law of the process over time. Consequently, the propagation of chaos can be shown for any locally Lipschitz continuous function f , exactly as if it was globally Lipschitz continuous and bounded.

The case where the initial conditions are not compactly supported is more delicate, at least when f is not globally Lipschitz continuous. Our results work under quite weak moment conditions on the initial datum, for quite general

functions f . We obtain a rate of convergence in $1/\sqrt{N}$, as one expects. These results should remain true when adding a diffusive component to the dynamics of individual neurons, at the cost of a higher level of technicality.

Finally, we propose a short study of the limit equation. We describe the shape of its time-marginals, we prove the existence of a unique nontrivial invariant distribution, we show that the trivial invariant distribution is not attractive, and when $\lambda = 0$, we establish the convergence to equilibrium for a class of initial conditions.

Let us mention that all the results and proofs below have been elaborated thinking of the case where $f(x) = x^\xi$ with $\xi \geq 2$, which thus satisfies all the conditions of the paper.

1.3. References

Using a mean-field approach in order to describe the typical behavior of a neuron within a large population of similarly behaving neurons from a macroscopic point of view is by now classical in neuromathematics. A lot of effort has been spent by the neuromathematical community focussing on the study of leaky integrate-and-fire models and their mean-field limits; in these models the membrane potential of a neuron is described by a (finite or infinite dimensional) diffusion process, and spiking occurs when reaching a threshold.

Recent interesting papers using a fixed threshold are those of Delarue, Inglis, Rubenthaler and Tanré [4] and [5], see also Cáceres, Carrillo and Perthame [3]. Here, the membrane potential is described by a one dimensional diffusion process. The existence of a fixed threshold may lead to severe mathematical problems related to a possible blow-up of the limit nonlinear equation. Such a blow-up appears when a macroscopically large proportion of neurons spike at the same time. Avalanches and synchronization are phenomena which are related to such a blow-up.

Notice that on the contrary, our model does not include a diffusive part in the evolution of each neuron's membrane potential. As indicated above, we expect our result on propagation of chaos to remain true when adding such a diffusive component. But of course none of the above mentioned phenomena such as blow-up or avalanches appear in our model since spiking occurs at a smooth rate which is finite all over the state space. As we have already mentioned, this choice of modeling is motivated by biological considerations. Consequently, we do not have to face the same difficulties. The problems we have to deal with are linked to the jump part of the equation, more specifically to the fact that the spiking rate is not globally Lipschitz.

Recently, Inglis and Talay [16] have proposed a model of the integrate and fire type where neurons do spike when hitting a fixed threshold but where the effect of a spike is not instantaneously transmitted to the other neurons. As a consequence, their model does not present blow-up phenomena neither.

For an excellent overview of the mean-field approach in integrate-and-fire models with a strong modeling point of view, we refer the reader to Faugeras, Touboul and Cessac [8]. They specifically deal with the case where several big populations of neurons interact through their neural efficacies which are chosen to be i.i.d. Gaussian random variables. In particular, they also deal with negative synaptic weights. Actually the extension of our model to the multi-population case, including also inhibitory synapses, seems to be quite straightforward and is part of a future work.

Finally, in a recent article, Luçon and Stannat [20] consider a population of mean field interacting diffusions which are attached to spatial positions and evolve within a random environment. Here, the spiking is encoded within the diffusion model (as in the Fitzhugh–Nagumo model). The interaction strength between two neurons depends on their spatial distance and may show singularities. Moreover, the coefficients of the underlying diffusion are of polynomial growth and therefore not globally Lipschitz neither. However, the techniques and results obtained in this article are clearly far from the considerations we are interested in the present paper.

To summarize, our aim is not to build a model which describes the full neurophysiological reality, but to study a simple model describing some basic biological features and to concentrate on the randomness hidden behind the spike times. It is inspired by integrate-and-fire models, but spiking occurs randomly, with state-dependent intensity and the system does not contain any other source of randomness. Let us finally mention that this model can also be interpreted in terms of an associated nonlinear Hawkes process including a variable memory structure. We refer to Hansen, Reynaud-Bouret and Rivoirard [14] for an interesting statistical study of the neuronal interaction graph using Hawkes processes.

From the purely probabilistic point of view, propagation of chaos is a popular topic since the seminal works of Kac [19], McKean [21,22] and Sznitman [25,26]. Generally, one tries to prove that the time-evolution of a particle, interacting with a large number of other particles, can be approximated by a *nonlinear* process. By *nonlinear*, in the

sense of McKean, we mean that the law of the process itself is involved in its dynamics. There exist essentially two kinds of proofs.

- The first one, based on *coupling* and often used in [26], provides a (often optimal) rate of convergence but works mainly when all the parameters of the model are globally Lipschitz continuous. However, it sometimes happens that the non-Lipschitz terms are not really a problem, when they have, roughly, the good sign: see Malrieu [23], who studies some McKean–Vlasov equation with, roughly, a convex interaction potential. More recently, it has been shown in Bolley, Cañizo and Carrillo [1] that the coupling method can also apply to the case of nonglobally Lipschitz parameters, under some very restrictive exponential moment conditions. They also get a (almost optimal) rate of convergence. This idea has also been exploited for the Boltzmann equation in [12].
- The second method, elaborated in [25] when studying the Boltzmann equation, is based on tightness/consistency/uniqueness of the nonlinear process. It applies much more generally (it requires only some a priori bounds and some continuity of the parameters), but does not provide any rate of convergence.

In the present paper we make use of the two methods and investigate to which extent they can be applied. Roughly, the tightness/consistency/uniqueness works very well, under some very light assumptions on f and on the initial conditions. But the most important point of the paper is that, still for quite a general class of functions f (as x^ξ with $\xi \geq 2$), we show that the coupling method also works, without imposing some exponential moment conditions. This is very specific to the model under study, relies on quite fine computations, and on the use of an ad hoc distance. As previously mentioned we get an optimal rate of convergence.

Finding an ad hoc distance is a classical strategy to prove uniqueness of the solution or to study its large time behavior, in all fields of differential equations. It is a good approach, in the sense that it often allows for many developments, such as stability and convergence of approximate models. But each model requires its own study and the *good* distance often looks mysterious. The *distance* may or may not depend on the precise parameters of the model. Let us quote a few papers. For example, Tanaka [27,28] discovered, using a specific nonlinear jumping SDE, that the Wasserstein distance with quadratic cost between two solutions of the homogeneous Boltzmann equation for Maxwell molecules is decreasing, providing the first uniqueness result for the Boltzmann equation in the physically reasonable case without cutoff. Bolley, Guillin and Malrieu [2] were able to precisely study, using a nonlinear Brownian SDE, the large-time behavior of solutions to a Vlasov–Fokker–Planck equation by introducing an ad hoc modification of the Wasserstein distance depending on the parameters of the equation. They also quantified, with similar tools, the convergence of some particle systems. In [9,11], we introduced an ad hoc distance to prove uniqueness of some infinite stochastic interacting particle systems undergoing coalescence. Here also, the distance was depending on the interaction kernel.

However, the study proposed in the present paper, and in particular the proof of the uniqueness of the limit equation, is situated in a completely different mathematical framework compared to the above mentioned papers, and the specific choice of an ad hoc distance that we propose is a new feature.

1.4. The limit equation

Assume that the $X_0^{N,i}$ are i.i.d. with common law g_0 on \mathbb{R}_+ . Simple considerations show that the solution $(X_t^N)_{t \geq 0}$ should behave, for N large, as N independent copies of the solution to the following *nonlinear*, in the sense of McKean, SDE. Let Y_0 be a g_0 -distributed random variable, independent of a Poisson measure $\mathbf{N}(ds, dz)$ on $\mathbb{R}_+ \times \mathbb{R}_+$ having intensity measure $ds dz$. An \mathbb{R}_+ -valued càdlàg adapted process $(Y_t)_{t \geq 0}$ is said to solve the nonlinear SDE if

$$Y_t = Y_0 - \lambda \int_0^t (Y_s - \mathbb{E}[Y_s]) ds - \int_0^t \int_0^\infty Y_{s-} \mathbf{1}_{\{z \leq f(Y_{s-})\}} \mathbf{N}(ds, dz) + \int_0^t \mathbb{E}[f(Y_s)] ds. \tag{2}$$

For PDE specialists, let us mention that for $(Y_t)_{t \geq 0}$ a solution to (2), $g(t) = \mathcal{L}(Y_t)$ solves the following nonlinear PDE in weak form: for any $\phi \in C_b^1(\mathbb{R}_+)$, the set of C^1 -functions on $[0, \infty)$ such that ϕ and ϕ' are bounded, for any $t \geq 0$,

$$\int_0^\infty \phi(x) g(t, dx) = \int_0^\infty \phi(x) g(0, dx) + \int_0^t \int_0^\infty ([\phi(0) - \phi(x)] f(x) + \phi'(x) [a_s - \lambda x]) g(s, dx) ds,$$

where $a_t = \int_0^\infty [f(x) + \lambda x]g(t, dx)$. Setting also $p_t = \int_0^\infty f(x)g(t, x) dx$, the strong equation writes

$$\forall t > 0, \forall x > 0, \quad \partial_t g(t, x) = [\lambda x - a_t] \partial_x g(t, x) + [\lambda - f(x)]g(t, x) \quad \text{and} \quad \forall t > 0, \quad g(t, 0) = p_t/a_t,$$

with $g(0, x)$ a given probability density on $[0, \infty)$.

The nonlinear SDE (2) is not clearly well posed, unless one assumes e.g. that f is globally Lipschitz-continuous and bounded. Under Assumption 1, we are generally only able to check the weak existence, that is existence of a filtered probability space on which there is a Poisson measure \mathbf{N} and a càdlàg adapted process $(Y_t)_{t \geq 0}$ such that (2) holds true for all $t \geq 0$.

Assumption 3. $f \in C^2(\mathbb{R}_+)$ is convex increasing and $\sup_{x \geq 1} [f'(x)/f(x) + f''(x)/f'(x)] < \infty$.

Theorem 4. Grant Assumption 1 and suppose that $\lambda \geq 0$.

(i) Assume only that $\mathbb{E}[Y_0] < \infty$. Then there is weak existence of a solution $(Y_t)_{t \geq 0}$ to (2) satisfying $\int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty$ for all $t \geq 0$.

(ii) Assume now that the law of Y_0 is compactly supported. Then there exists a path-wise unique solution $(Y_t)_{t \geq 0}$ to (2) such that there is a deterministic locally bounded function $A : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that a.s., $\sup_{[0, \infty)} (Y_t/A(t)) < \infty$.

(iii) Grant now Assumption 3 and assume that $\mathbb{E}[f(Y_0)] < \infty$. Then there is a path-wise unique solution to (2) satisfying $\sup_{[0, t]} \mathbb{E}[f(Y_s)] < \infty$ for all $t \geq 0$.

Let us mention that Assumption 3 can be slightly relaxed: if for example $f = f_1 + f_2$ with f_1 satisfying Assumptions 1 and 3 and $f_2(x) = \int_0^x \psi(y) dy$ with $\psi \in C_c^1([0, \infty))$ nonnegative, then Theorem 4(iii) still holds true. In fact, what we really need is that the conclusions of Lemma 18 below are satisfied.

Let us comment on the results of Theorem 4. Point (i) is not hard: it is checked by compactness and is actually a consequence of Theorem 5(i)–(ii) below. The only noticeable point is that the condition $\mathbb{E}[Y_0] < \infty$ is sufficient to guarantee that indeed, $\int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty$ for all $t \geq 0$, which is sufficient to handle a proof by compactness. Point (ii) is not very complicated and has already been proven in [6]. The only difficult point is to check that if Y_0 is bounded, then Y_t is a priori bounded for all t . Once this is seen, the function f can be considered as if it was bounded and globally Lipschitz continuous. Finally, (iii) is much more delicate and goes clearly beyond the results of [6]. Indeed, when computing the time derivative of $\mathbb{E}[|X_t - Y_t|]$, for X and Y two solutions to (2), some nonlinear terms appear: there is no hope to conclude uniqueness by the Gronwall lemma. One possibility is to use the famous $x|\log x|$ extension of the Gronwall lemma, but this requires to have some bounds for something like $\sup_{[0, T]} \mathbb{E}[\exp(f(Y_t))] < \infty$, see Bolley, Cañizo and Carrillo [1] or [12] for such considerations, but this is not very satisfying, since it requires the strong condition that $\mathbb{E}[\exp(f(Y_0))] < \infty$. We thus search for a more convenient “distance.” We first observe that when time-differentiating $\mathbb{E}[|f(X_t) - f(Y_t)|]$, the contribution of the most unpleasant term (the Poisson integral) is nonpositive: it gives exactly $-\mathbb{E}[|f(X_t) - f(Y_t)|^2]$, which is a very good point. However, the other terms cause problems for small values of X, Y , if f vanishes too fast at 0 (e.g. $f(x) = x^\xi$ with $\xi \geq 2$). To overcome this difficulty, it actually suffices to work with $\mathbb{E}[|H(X_t) - H(Y_t)|]$, where $H(x) \simeq f(x) + x \wedge 1$. Of course, it is more convenient to use a smooth version of $x \wedge 1$, so that we will work with $H(x) = \arctan x + f(x)$. It is likely that we could also use $H(x) = \ell(x) + f(x)$, with any smooth increasing function $\ell(x)$ behaving like ax near 0 (for some $a > 0$) and tending to some constant $b > 0$ as $x \rightarrow \infty$.

1.5. Propagation of chaos

We start with a general weak result. The set $\mathbb{D}(\mathbb{R}_+)$ of càdlàg functions on \mathbb{R}_+ is endowed with the topology of the Skorokhod convergence on compact time intervals, see Jacod and Shiryaev [17].

Theorem 5. Grant Assumption 1 and suppose that $\lambda \geq 0$. Consider a probability distribution g_0 on \mathbb{R}_+ such that $\int_0^\infty y g_0(dy) < \infty$. For each $N \geq 1$, consider the unique solution $(X_t^N)_{t \geq 0}$ to (1) starting from some i.i.d. g_0 -distributed initial conditions $X_0^{N,i}$.

(i) The sequence of processes $(X_t^{N,1})_{t \geq 0}$ is tight in $\mathbb{D}(\mathbb{R}_+)$.

- (ii) The sequence of empirical measures $\mu_N = N^{-1} \sum_{i=1}^N \delta_{(X_t^{N,i})_{t \geq 0}}$ is tight in $\mathcal{P}(\mathbb{D}(\mathbb{R}_+))$.
- (iii) Any limit point μ of μ_N a.s. belongs to $\mathcal{S} := \{\mathcal{L}((Y_t)_{t \geq 0}) : (Y_t)_{t \geq 0} \text{ solution to (2) with } \mathcal{L}(Y_0) = g_0 \text{ and satisfying } \int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty \text{ for all } t \geq 0\}$.
- (iv) If moreover (a) g_0 is compactly supported or (b) $\int_0^\infty f(y)g_0(dy) < \infty$ and f satisfies Assumption 3, then μ_N goes in probability to $\mathcal{L}((Y_t)_{t \geq 0})$, where $(Y_t)_{t \geq 0}$ is the unique solution to (2).

Points (i), (ii) and (iii) are not very difficult. The fact that f is not globally Lipschitz continuous is not really a problem when working by compactness. And of course, point (iv), which is usually called propagation of chaos, is a consequence of points (ii) and (iii) and of the uniqueness results of Theorem 4. Again, the above theorem has already been proven in the case of compact support in [6]; but the techniques employed in [6] cannot be used in the general case where g_0 is not compactly supported. Under a few additional conditions, we get a quantified version of the above convergence, at least concerning the time marginals.

Assumption 6. There is a constant C such that for all $x, y \in \mathbb{R}_+$, $f(x + y) \leq C(1 + f(x) + f(y))$.

Theorem 7. Grant Assumptions 1, 3 and 6 and suppose that $\lambda \geq 0$ and that $\int_0^\infty f^2(y)g_0(dy) < \infty$. Consider, for each $N \geq 1$, the unique solution $(X_t^N)_{t \geq 0}$ to (1) starting from some i.i.d. g_0 -distributed initial conditions $X_0^{N,i}$. Consider also the unique solution $(Y_t^{N,1})_{t \geq 0}$ to (2) starting from $Y_0 = X_0^{N,1}$ and driven by the Poisson random measure $\mathbf{N}^1(ds, dz)$. The law of $(Y_t^{N,1})_{t \geq 0}$ does not depend on N , and we denote by $g(t) := \mathcal{L}(Y_t^{N,1})$. Introduce $H(x) = f(x) + \arctan(x)$. Then for all $T > 0$, there is a constant C_T , depending only on T, λ, f and $\int_0^\infty f^2(y)g_0(dy)$ such that

$$\sup_{[0, T]} (\mathbb{E}[|X_t^{N,1} - Y_t^{N,1}|] + \mathbb{E}[|H(X_t^{N,1}) - H(Y_t^{N,1})|]) \leq \frac{C_T}{\sqrt{N}}.$$

Assume furthermore that $\int_0^\infty y^{2+\varepsilon} g_0(dy) < \infty$ for some $\varepsilon > 0$. Then for all $T > 0$, there is a constant C_T , depending only on $T, \lambda, f, \varepsilon$ and $\int_0^\infty [f^2(y) + y^{2+\varepsilon}]g_0(dy)$ such that

$$\sup_{[0, T]} \mathbb{E} \left[\mathcal{W}_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{X_t^{N,i}}, g(t) \right) \right] \leq \frac{C_T}{\sqrt{N}}.$$

The Monge–Kantorovich–Wasserstein distance between two probability measures μ and ν on \mathbb{R}_+ with finite expectations is defined by $\mathcal{W}_1(\mu, \nu) = \inf\{\mathbb{E}[|U - V|], \mathcal{L}(U) = \mu \text{ and } \mathcal{L}(V) = \nu\}$. The moment condition $\int_0^\infty f^2(y)g_0(dy) < \infty$ is very reasonable: somewhere in the proof, we will have to study the convergence of $N^{-1} \sum_{j=1}^N f(Y_t^{N,j})$ to $\mathbb{E}[f(Y_t^{N,1})]$. If we want a rate of convergence of order $1/\sqrt{N}$, such an assumption is needed.

1.6. Large time behavior of the limit process

First, we study the possible invariant measures.

Theorem 8. Grant Assumption 1 and let $\lambda \geq 0$. Then the nonlinear equation (2) has exactly two invariant probability measures supported in \mathbb{R}_+ . The first one is δ_0 . The second one is of the form $g(dx) = g(x) dx$, with $g : [0, \infty) \mapsto [0, \infty)$ defined as follows.

- (i) If $\lambda > 0$, then

$$g(x) = \frac{p}{p + \lambda m - \lambda x} \exp\left(-\int_0^x \frac{f(y)}{p + \lambda(m - y)} dy\right) \mathbf{1}_{\{0 \leq x < m + p/\lambda\}},$$

where $p > 0$ and $m > 0$ are uniquely determined by the constraints $\int_0^\infty g(dx) = 1$, $\int_0^\infty xg(dx) = m$. Furthermore, we have $\int_0^\infty f(x)g(dx) = p$ and $m + p/\lambda > 1$.

(ii) If $\lambda = 0$, then

$$g(x) = \exp\left(-\frac{1}{p} \int_0^x f(y) dy\right),$$

where $p > 0$ is uniquely determined by the constraint $\int_0^\infty g(x) dx = 1$. Furthermore, it holds that $\int_0^\infty f(x) \times g(x) dx = p$.

Starting from a (reasonable) nontrivial initial condition, it is likely that Y_t goes in law to g as $t \rightarrow \infty$. When $\lambda = 0$, we can prove such a result under a few assumptions.

Proposition 9. Grant Assumptions 1 and 3 and assume that $\lambda = 0$. Suppose moreover that the solution $(Y_t)_{t \geq 0}$ to (2) starts from $Y_0 \sim g_0(x) dx$ where $g_0 \in C_b^1([0, \infty))$ satisfies $g_0(0) = 1$, $\int_0^\infty f^2(x)g_0(x) dx < \infty$ and $\int_0^\infty |g_0'(x)| dx < \infty$. Denote by $g(t)$ the law of Y_t and write g for the invariant probability measure defined in Theorem 8(ii). Then we have $\lim_{t \rightarrow \infty} \|g(t) - g\|_{TV} = 0$, where $\|\cdot\|_{TV}$ denotes the total variation distance. If furthermore there are $c > 0$ and $\xi \geq 1$ such that $f(x) \geq cx^\xi$ for all $x \in [0, 1]$, then we have the estimate $\|g(t) - g\|_{TV} \leq C(1+t)^{-1/\xi}$.

Our proof, which is probably not optimal, relies on the use of the *strong version* of the PDE satisfied by $g(t)$. The regularity conditions on g_0 , as well as the *structure* condition $g_0(0) = 1$, will imply that $g(t)$ has a sufficiently regular density. As can be seen in the next subsection (see also (20) in Section 7.3), if $g_0(0) \neq 1$, then $g(t, y)$ will be discontinuous for all $t > 0$ (it will have one jump at some value $y_t \in (0, \infty)$ depending on t).

When $\lambda > 0$, the situation is more intricate and we have not been able to prove the convergence to equilibrium. One reason is that the nondegenerate invariant probability measure is more complicated, compactly supported and possibly not continuous at the right extremity of its support. In any case, the computation handled to treat the case $\lambda = 0$ does not extend. A natural approach would be to show first that $\lim_{t \rightarrow \infty} \mathbb{E}[\lambda Y_t + f(Y_t)]$ exists. However, $\mathbb{E}[\lambda Y_t + f(Y_t)]$ does not solve a closed equation, and we did not succeed. The only result we are able to prove is that Y_t cannot go in law to the invariant measure δ_0 . Our proof, which was as usual elaborated in the case where $f(x) = x^\xi$, actually extends to the following situation.

Assumption 10.

- (i) It holds that $\limsup_{x \rightarrow \infty} [f'(x)/f(x)] < 1$.
- (ii) There are $\xi \geq 1$, $\zeta \geq \xi - 1$ and some constants $0 < c < C$ such that $cx^\xi \leq f(x) \leq C(x^{\xi-1} + x^\zeta)$.

Proposition 11. Let $\lambda \geq 0$ and grant Assumptions 1 and 3. Assume that $\mathbb{P}(Y_0 = 0) < 1$ and that $\mathbb{E}[f^2(Y_0)] < \infty$ and consider the unique solution $(Y_t)_{t \geq 0}$ to (2). If $\lambda > 0$, grant moreover Assumption 10 and suppose that $\mathbb{E}[Y_0^{\zeta+1}] < \infty$. Then Y_t does not go to 0 in law as $t \rightarrow \infty$.

1.7. Shape of the time-marginals of the nonlinear SDE

The next theorem shows that random spiking creates density near 0, even if the system starts from a singular initial condition (see also Theorem 2 of [6] in the case of a smooth initial condition).

Theorem 12. Let $\lambda \geq 0$, grant Assumptions 1 and 3 and suppose that $\mathbb{E}[f^2(Y_0)] < \infty$ and that $\mathbb{P}(Y_0 = 0) < 1$. Consider the unique solution $(Y_t)_{t \geq 0}$ to (2), set $p_t = \mathbb{E}[f(Y_t)]$, $a_t = \lambda \mathbb{E}[Y_t] + \mathbb{E}[f(Y_t)]$ and denote by $g(t)$ the law of Y_t . The functions $t \mapsto a_t$ and $t \mapsto p_t$ are continuous and positive on $[0, \infty)$. Introduce the deterministic flow as follows: for $x \in [0, \infty)$ and $0 \leq s \leq t$,

$$g_{s,t}(x) = e^{-\lambda(t-s)}x + \int_s^t e^{-\lambda(t-u)}a_u du.$$

For $t > 0$ fixed and $y \in [0, \varphi_{0,t}(0)]$, let $\beta_t(y) \in [0, t]$ be uniquely determined by $\varphi_{\beta_t(y),t}(0) = y$. For $y \geq \varphi_{0,t}(0)$, let $\gamma_t(y) = (y - \varphi_{0,t}(0))e^{\lambda t}$, which satisfies $\varphi_{0,t}(\gamma_t(y)) = y$. It holds that for any $t > 0$,

$$(t, dy) = \frac{p_{\beta_t(y)}}{a_{\beta_t(y)}} \exp\left(\int_{\beta_t(y)}^t (\lambda - f(\varphi_{\beta_t(y),s}(0))) ds\right) \mathbf{1}_{\{y \in [0, \varphi_{0,t}(0)]\}} dy + \exp\left(-\int_0^t f(\varphi_{0,s}(\gamma_t(y))) ds\right) \mathbf{1}_{\{y \in [\varphi_{0,t}(0), \infty)\}} (g_0 \circ \gamma_t^{-1})(dy).$$

In particular, since $\beta_t(0) = t$, the density of $g(t)$ at 0 is given by $g(t, 0) = p_t/a_t$.

1.8. Plan of the paper

Section 2 consists of collecting some useful a priori bounds for the particle system and the limit process. In Section 3, we check the path-wise uniqueness of the limit process. Section 4 is devoted to the proof of Theorem 4, the propagation of chaos without rate of convergence. Section 5 shows the quantified propagation of chaos result. In Section 6, we investigate the possible invariant distributions of the limit process. The shape of the time-marginals is studied in Section 7, in which we also prove the nonextinction result (Proposition 11) and the trend to equilibrium when $\lambda = 0$ (Proposition 9).

1.9. Constants

In the whole paper, C stands for a (large) finite constant and c stands for a (small) positive constant. Their values may change from line to line. They are allowed to depend only on f, λ and g_0 , any other dependence will be indicated in subscript. For example, C_T is a finite constant depending only on f, λ, g_0 and T .

2. A priori bounds

The aim of this section is to establish some fundamental bounds for the particle system and for the limit process. Before that, we proceed to some elementary considerations.

Remark 13. Grant Assumptions 1 and 3 and additionally Assumption 6 for point (iv).

- (i) There is $c > 0$ such that $f(x) \geq cx$ for all $x \geq 1$.
- (ii) For all $A > 0$, there is $C_A > 0$ such that for all $x \geq 0$, $f(x + A) \leq C_A(1 + f(x))$.
- (iii) There is $C > 0$ such that $f(x) \leq C \exp(Cx)$ for all $x \geq 0$.
- (iv) For all $A > 0$, there is $C_A > 0$, $f(A(x + y)) \leq C_A(1 + f(x) + f(y))$ for all $x, y \geq 0$.

Proof. Point (i) is obvious since f is convex increasing and since $f(0) = 0$. Point (iii) is easily checked using that f is increasing as well as point (ii) with $A = 1$. Point (iv) is an immediate consequence of Assumption 6. To check (ii), we will prove that there is $a > 0$ such that $f(x + a) \leq 2f(x) + 1$ for all $x \geq 0$, which clearly suffices. By Assumption 3, there is $B > 0$ such that $f'(x) \leq B(1 + f(x))$ for all $x \geq 0$. Fix $a := 1/(2B)$ and write $f(x + a) = f(x) + \int_x^{x+a} f'(y) dy \leq f(x) + aB(1 + \sup_{[x, x+a]} f) = f(x) + aB(1 + f(x + a)) = f(x) + (1 + f(x + a))/2$, whence $f(x + a) \leq 2f(x) + 1$ as desired. \square

We now study the limit equation.

Proposition 14. Grant Assumption 1, suppose that $\lambda \geq 0$ and that $\mathbb{E}[Y_0] < \infty$. There is a constant $C > 0$ depending only on λ, f and $\mathbb{E}[Y_0]$ such that a solution $(Y_t)_{t \geq 0}$ to (2) a priori satisfies

$$a.s., \text{ for all } t \geq 0, \quad Y_t \leq Y_0 + C(1 + t), \tag{3}$$

$$\text{for all } t \geq 0, \quad \int_0^t \mathbb{E}[Y_s f(Y_s)] ds \leq 2\mathbb{E}[Y_0] + 2f(2)t. \tag{4}$$

Proof. Taking expectations in (2), we see that

$$\mathbb{E}[Y_t] = \mathbb{E}[Y_0] + \int_0^t \mathbb{E}[f(Y_s)(1 - Y_s)] ds \leq \mathbb{E}[Y_0] + f(2)t - \frac{1}{2} \int_0^t \mathbb{E}[Y_s f(Y_s)] ds,$$

because $f(x)(1 - x) = -xf(x)/2 + f(x)(1 - x/2) \leq -xf(x)/2 + f(2)$ for $x \geq 0$, since f is nonnegative and nondecreasing. Inequality (4) then follows from the fact that $\mathbb{E}[Y_t] \geq 0$. Recalling the nonlinear SDE (2), we see that $Y_t \leq Y_0 + \int_0^t \mathbb{E}[\lambda Y_s + f(Y_s)] ds$ for all $t \geq 0$ a.s. But there exists a constant C , depending on f and λ , such that $\lambda y + f(y) \leq C(1 + yf(y))$ for all $y \geq 0$: it suffices to use that f is positive and nondecreasing. Consequently, $\mathbb{E}[\lambda Y_s + f(Y_s)] \leq C(1 + \mathbb{E}[Y_s f(Y_s)])$ for all $s \geq 0$ and (3) follows from (4). \square

We now turn to the particle system.

Proposition 15. *Grant Assumption 1 and suppose that $\lambda \geq 0$. Any solution $(X_t^N)_{t \geq 0}$ to (1) a.s. satisfies that for all $t \geq 0$, all $i = 1, \dots, N$,*

$$X_t^{N,i} \leq X_0^{N,i} + (4\lambda t + 4)(\bar{X}_0^N + Z_t^N), \quad (5)$$

$$\frac{1}{N} \sum_{j=1}^N \int_0^t \int_0^\infty (1 + X_{s-}^{N,j}) \mathbf{1}_{\{z \leq f(X_{s-}^{N,j})\}} \mathbf{N}^j(ds, dz) \leq 3\bar{X}_0^N + 4Z_t^N, \quad (6)$$

where $Z_t^N := N^{-1} \sum_{j=1}^N \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f(2)\}} \mathbf{N}^j(ds, dz)$. Furthermore, it holds that for any $T \geq 0$,

$$\mathbb{P}\left(\forall i = 1, \dots, N, \sup_{[0, T]} X_t^{N,i} \leq X_0^{N,i} + (4\lambda T + 4)(\bar{X}_0^N + 2f(2)T)\right) \geq 1 - e^{-NTf(2)(3-e)}. \quad (7)$$

Proof. We start with the following observation: taking the (empirical) mean of (1), we find

$$\bar{X}_t^N = \bar{X}_0^N + \frac{1}{N} \sum_{i=1}^N \int_0^t \int_0^\infty \left(\frac{N-1}{N} - X_{s-}^{N,i}\right) \mathbf{1}_{\{z \leq f(X_{s-}^{N,i})\}} \mathbf{N}^i(ds, dz) \quad (8)$$

which implies, since $\bar{X}_t^N \geq 0$ and $(N-1)/N \leq 1$, that

$$\frac{1}{N} \sum_{i=1}^N \int_0^t \int_0^\infty (X_{s-}^{N,i} - 1) \mathbf{1}_{\{z \leq f(X_{s-}^{N,i})\}} \mathbf{N}^i(ds, dz) \leq \bar{X}_0^N.$$

Using that $x - 1 \geq (x + 1)/3 - (4/3)\mathbf{1}_{\{x \leq 2\}}$ for all $x \geq 0$ and that f is nondecreasing, we deduce that

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \int_0^t \int_0^\infty (1 + X_{s-}^{N,i}) \mathbf{1}_{\{z \leq f(X_{s-}^{N,i})\}} \mathbf{N}^i(ds, dz) \\ & \leq 3\bar{X}_0^N + \frac{4}{N} \sum_{i=1}^N \int_0^t \int_0^\infty \mathbf{1}_{\{X_{s-}^{N,i} \leq 2\}} \mathbf{1}_{\{z \leq f(X_{s-}^{N,i})\}} \mathbf{N}^i(ds, dz). \end{aligned}$$

Since f is nondecreasing, (6) follows. Recalling (8) and using (6), we realize that

$$\bar{X}_t^N \leq \bar{X}_0^N + \frac{1}{N} \sum_{i=1}^N \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f(X_{s-}^{N,i})\}} \mathbf{N}^i(ds, dz) \leq 4\bar{X}_0^N + 4Z_t^N.$$

Now, for all $1 \leq i \leq N$, starting from (1),

$$\begin{aligned} X_t^{N,i} &\leq X_0^{N,i} + \lambda \int_0^t \bar{X}_s^N ds + \frac{1}{N} \sum_{j=1}^N \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f(X_s^{N,j})\}} \mathbf{N}^j(ds, dz) \\ &\leq X_0^{N,i} + \lambda \int_0^t (4\bar{X}_0^N + 4Z_s^N) ds + 3\bar{X}_0^N + 4Z_t^N. \end{aligned}$$

Hence (5) follows from the fact that Z_t^N is a.s. a nondecreasing function of time. Finally, the deviation estimate (7) simply relies on (5) and the inequality

$$\mathbb{P}(Z_T^N \geq 2f(2)T) \leq e^{-2f(2)NT} \mathbb{E}[e^{NZ_T^N}] = e^{-NTf(2)(3-e)}, \quad (9)$$

which uses that Z_T^N is the empirical mean of N i.i.d. Poisson($f(2)T$)-random variables. \square

The above estimate is largely sufficient to give the

Proof of Proposition 2. Suppose first that f is bounded. Then using only that f is measurable (and nonnegative), we can apply Theorem 9.1 in Chapter IV of Ikeda and Watanabe [15]: there is a path-wise unique solution $(X_t^N)_{t \geq 0}$ to (1) defined on $[0, \infty)$.

For a general f satisfying Assumption 1 and a fixed truncation level $K > 0$, we consider the unique solution $(X_t^{N,K})_{t \geq 0}$ to (1) with f replaced by $f \wedge K$ and we introduce $\tau_K = \inf\{t \geq 0 : |X_t^{N,K}| \geq K\}$. By path-wise uniqueness, it holds that $X_t^{N,K} = X_t^{N,K+1}$ for all $K \in \mathbb{N}$ and all $t \in [0, \tau_K]$ and that $\tau_K < \tau_{K+1}$ for all $K \in \mathbb{N}$, almost surely. Setting $\tau = \sup_K \tau_K$, it is not hard to conclude that there is a path-wise unique solution $(X_t^N)_{t \in [0, \tau)}$ to (1) defined on $[0, \tau)$ and that $\limsup_{t \rightarrow \tau} |X_t^N| = \infty$ on the event $\{\tau < \infty\}$.

Recall now (5): a.s., $X_t^{N,i} \leq X_0^{N,i} + C(1+t)(\bar{X}_0^N + Z_t^N)$ for all $i = 1, \dots, N$, all $t \geq 0$. Observe also that obviously, $\sup_{[0, T]} Z_t^N < \infty$ a.s. for all $T > 0$. Hence $\tau = \infty$ a.s., which completes the proof. \square

3. Path-wise uniqueness for the nonlinear SDE

Let us first consider the case with compact support.

Proposition 16. *Suppose Assumption 1 and that $\lambda \geq 0$. Path-wise uniqueness holds true for the nonlinear SDE (2), in the class of processes $(Y_t)_{t \geq 0}$ such that there is a deterministic locally bounded function $A : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that a.s., $\sup_{t \geq 0} (Y_t/A(t)) < \infty$.*

Note that the above condition is a priori satisfied for g_0 compactly supported thanks to (3).

Proof of Proposition 16. Consider two solutions $(Y_t)_{t \geq 0}$ and $(X_t)_{t \geq 0}$ to (2), driven by the same Poisson measure \mathbf{N} and with $Y_0 = X_0$. A very rough computation shows that there is a constant C , depending only on λ , such that

$$\begin{aligned} \mathbb{E}[|X_t - Y_t|] &\leq C \int_0^t \mathbb{E}[|X_s - Y_s|(1 + f(X_s) + f(Y_s)) + |f(X_s) - f(Y_s)|(1 + X_s + Y_s)] ds \\ &\quad + C \int_0^t (|\mathbb{E}[X_s] - \mathbb{E}[Y_s]| + |\mathbb{E}[f(X_s)] - \mathbb{E}[f(Y_s)]|) ds. \end{aligned} \quad (10)$$

But we know by assumption that a.s., $\max\{Y_t, X_t\} \leq A(t)$ for some deterministic locally bounded function A . Since f is C^1 on $[0, \infty)$, it is Lipschitz continuous and bounded on compacts. We thus easily check that for all T , there is a constant C_T such that for all $t \in [0, T]$,

$$\mathbb{E}[|X_t - Y_t|] \leq C_T \int_0^t \mathbb{E}[|X_s - Y_s|] ds.$$

Finally, we know by assumption that the function $t \mapsto \mathbb{E}[|X_t - Y_t|]$ is locally bounded. We thus may apply the Gronwall Lemma and deduce that $\mathbb{E}[|X_t - Y_t|] = 0$ for all $t \geq 0$ as desired. \square

Proposition 17. *Let $\lambda \geq 0$ and grant Assumptions 1 and 3. Path-wise uniqueness holds true for the nonlinear SDE (2) in the class of processes $(Y_t)_{t \geq 0}$ such that $\sup_{[0, T]} \mathbb{E}[f(Y_t)] < \infty$ for all $T \geq 0$.*

More generally, for any pair of solutions $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ to (2), satisfying $\sup_{[0, T]} (\mathbb{E}[f(X_t)] + \mathbb{E}[f(Y_t)]) < \infty$ for all $T \geq 0$, driven by the same Poisson measure but with possibly different initial conditions, it holds that for all $T \geq 0$,

$$\sup_{[0, T]} \mathbb{E}[|H(X_t) - H(Y_t)|] \leq C_T \mathbb{E}[|H(X_0) - H(Y_0)|], \tag{11}$$

where $H(x) = f(x) + \arctan x$.

Here again, (3) (and Remark 13(ii)) shows that the condition is a priori satisfied if $\mathbb{E}[f(Y_0)] < \infty$. As already mentioned, a proof based on $\mathbb{E}[|X_t - Y_t|]$ does not seem to work: one finds an inequality like (10) (even with a finer computation using the Itô formula), from which it seems difficult to conclude.

The rest of the section is devoted to the proof of Proposition 17.

Lemma 18. *Grant Assumptions 1 and 3 and let $H(x) = f(x) + \arctan(x)$. There is a constant C such that for all $x, y \in \mathbb{R}_+$, we have*

- (0) $|H''(x)| \leq CH'(x)$,
- (i) $x + H'(x) \leq C(1 + f(x))$,
- (ii) $|x - y| + |H'(x) - H'(y)| + |f(x) - f(y)| \leq C|H(x) - H(y)|$,
- (iii) $-\text{sg}(x - y)(xH'(x) - yH'(y)) \leq C|H(x) - H(y)|$,
- (iv) $-(f(x) \wedge f(y))|H(x) - H(y)| + |f(x) - f(y)|(H(x) \wedge H(y) - |H(x) - H(y)|) \leq C|H(x) - H(y)|$.

Proof. First, $|H''(x)| \leq |\arctan''(x)| + f''(x) \leq C + f''(x)$. If $x \leq 1$, we deduce that $|H''(x)| \leq C \leq CH'(x)$, while if $x \geq 1$, we recall that $f''(x) \leq Cf'(x)$, whence $|H''(x)| \leq C(1 + f'(x)) \leq Cf'(x) \leq CH'(x)$ as desired.

We next check (i). We have $x + H'(x) \leq x + f'(x) + 1$. If $x \leq 1$, we just write $x + H'(x) \leq C \leq C(1 + f(x))$. If now $x \geq 1$, since $f'(x) \leq Cf(x)$ by Assumption 3, we find that $x + H'(x) \leq 2x + Cf(x) \leq Cf(x)$ by Remark 13(i).

In order to prove (ii), it is sufficient to check that $1 + |H''(x)| + f'(x) \leq CH'(x)$ for all $x \geq 0$. First, $1 \leq CH'(x)$ for all $x \geq 0$, because $H'(x) \geq f'(1) > 0$ if $x \geq 1$, while $H'(x) \geq \arctan'(x) \geq 1/2$ if $x \leq 1$. Next, $f'(x) \leq H'(x)$ is obvious. Finally, $|H''(x)| \leq CH'(x)$ has already been seen.

Concerning point (iii),

$$-\text{sg}(x - y)(xH'(x) - yH'(y)) = -\text{sg}(x - y)(xf'(x) - yf'(y)) - \text{sg}(x - y) \left(\frac{x}{1 + x^2} - \frac{y}{1 + y^2} \right).$$

The first term on the RHS is negative, because $xf'(x)$ is nondecreasing. The second one can be roughly bounded by $C|x - y|$ which in turn is bounded by $C|H(x) - H(y)|$ due to point (ii).

Finally, we rewrite the LHS of point (iv) as

$$-(f(x) \vee f(y))|H(x) - H(y)| + (H(x) \wedge H(y))|f(x) - f(y)|.$$

Using that $|f(x) - f(y)| \leq |H(x) - H(y)|$ because $H(x) = f(x) + \arctan(x)$ with both f and \arctan nondecreasing, that $f(x) \vee f(y) \geq f(x)$ and $H(x) \wedge H(y) \leq H(x)$, we get an upper-bound in

$$-f(x)|H(x) - H(y)| + H(x)|H(x) - H(y)| = (\arctan x)|H(x) - H(y)| \leq \pi|H(x) - H(y)|/2.$$

This completes the proof. \square

Proof of Proposition 17. Let thus $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ be two solutions of (2) driven by the same Poisson measure and satisfying $\sup_{[0, T]} \mathbb{E}[f(X_t) + f(Y_t)] < \infty$ for all $T > 0$. We apply the Itô formula for jump processes and take

expectations to compute $\mathbb{E}[|H(X_t) - H(Y_t)|]$. Actually, one has to first consider a regularized version of the absolute value and then to pass to the limit, but this causes no difficulty. See the proof of Theorem 7 where such a regularization procedure is handled. We find, using that H is increasing (whence $\text{sg}(H(x) - H(y)) = \text{sg}(x - y)$) and that $H(0) = 0$,

$$\begin{aligned} \mathbb{E}[|H(X_t) - H(Y_t)|] &= \mathbb{E}[|H(X_0) - H(Y_0)|] - \lambda \int_0^t \mathbb{E}[\text{sg}(X_s - Y_s)(H'(X_s)X_s - H'(Y_s)Y_s)] ds \\ &\quad + \lambda \int_0^t \mathbb{E}[\text{sg}(X_s - Y_s)(H'(X_s)\mathbb{E}[X_s] - H'(Y_s)\mathbb{E}[Y_s])] ds \\ &\quad + \int_0^t \mathbb{E}[-(f(X_s) \wedge f(Y_s))|H(X_s) - H(Y_s)| \\ &\quad + |f(X_s) - f(Y_s)|(H(X_s) \wedge H(Y_s) - |H(X_s) - H(Y_s)|)] ds \\ &\quad + \int_0^t \mathbb{E}[\text{sg}(X_s - Y_s)(H'(X_s)\mathbb{E}[f(X_s)] - H'(Y_s)\mathbb{E}[f(Y_s)])] ds \\ &=: \mathbb{E}[|H(X_0) - H(Y_0)|] + I_t + J_t + K_t + L_t. \end{aligned}$$

First, it follows from Lemma 18(iii)–(iv) that $I_t + K_t \leq C \int_0^t \mathbb{E}[|H(X_s) - H(Y_s)|] ds$. Next we write

$$\begin{aligned} J_t + L_t &\leq \int_0^t \mathbb{E}[|H'(X_s) - H'(Y_s)|](\lambda\mathbb{E}[X_s] + \mathbb{E}[f(X_s)]) ds \\ &\quad + \int_0^t \mathbb{E}[H'(Y_s)](\lambda|\mathbb{E}[X_s - Y_s]| + |\mathbb{E}[f(X_s) - f(Y_s)]|) ds. \end{aligned}$$

Using Lemma 18(i)–(ii), we thus find

$$J_t + L_t \leq C \int_0^t \mathbb{E}[|H(X_s) - H(Y_s)|]\mathbb{E}[1 + f(X_s) + f(Y_s)] ds.$$

Since $\sup_{t \in [0, T]} \mathbb{E}[f(X_t) + f(Y_t)] < \infty$ by assumption, we conclude that for all $T \geq 0$, there is a constant C_T such that for all $t \in [0, T]$,

$$\mathbb{E}[|H(X_t) - H(Y_t)|] \leq \mathbb{E}[|H(X_0) - H(Y_0)|] + C_T \int_0^t \mathbb{E}[|H(X_s) - H(Y_s)|] ds.$$

We know by assumption that $t \mapsto \mathbb{E}[|H(X_t) - H(Y_t)|]$ is locally bounded. Hence (11) follows from the Gronwall Lemma. Path-wise uniqueness is immediately deduced by injectivity of H . \square

4. Propagation of chaos without rate

In this section, we prove Theorem 5 and conclude the proof of Theorem 4. We start with tightness.

Proof of Theorem 5(i)–(ii). First, it is well known that point (ii) follows from point (i) and the exchangeability of the system, see Sznitman [26, Proposition 2.2(ii)]. We thus only prove (i). We consider a probability distribution g_0 on \mathbb{R}_+ such that $\int_0^\infty x g_0(dx) < \infty$ and, for each $N \geq 1$, the unique solution $(X_t^N)_{t \geq 0}$ to (1) starting from some i.i.d. g_0 -distributed initial conditions $X_0^{N,i}$. To show that the family $((X_t^{N,1})_{t \geq 0})_{N \geq 1}$ is tight $\mathbb{D}(\mathbb{R}_+)$, we use the criterion of Aldous, see Jacod and Shiryaev [17, Theorem 4.5, page 356]. It is sufficient to prove that

- (a) for all $T > 0$, all $\varepsilon > 0$, $\lim_{\delta \downarrow 0} \limsup_{N \rightarrow \infty} \sup_{(S, S') \in A_{\delta, T}} \mathbb{P}(|X_{S'}^{N,1} - X_S^{N,1}| > \varepsilon) = 0$, where $A_{\delta, T}$ is the set of all pairs of stopping times (S, S') such that $0 \leq S \leq S' \leq S + \delta \leq T$ a.s.,
- (b) for all $T > 0$, $\lim_{K \uparrow \infty} \sup_N \mathbb{P}(\sup_{t \in [0, T]} X_t^{N,1} \geq K) = 0$.

Point (b) follows from (5): we know that $\sup_{[0,T]} X_t^{N,1} \leq X_0^{N,1} + (4\lambda T + 4)(\bar{X}_0^N + Z_T^N)$, with Z_T^N the mean of N i.i.d. $\text{Poisson}(Tf(2))$ -distributed random variables. Hence, setting $m = \int_0^\infty xg_0(dx)$,

$$\sup_N \mathbb{E} \left[\sup_{[0,T]} X_t^{N,1} \right] \leq m + (4\lambda T + 4)(m + f(2)T) < \infty.$$

To check (a), we will use several times that for $0 \leq t \leq T$, all $N \geq 1$, all $j = 1, \dots, N$,

$$\int_0^t \mathbb{E}[X_s^{N,j} f(X_s^{N,j})] ds \leq 3m + 4f(2)t \leq C_T. \tag{12}$$

Indeed, take expectations in (6), use exchangeability and recall that $\mathbb{E}[\bar{X}_0^N] = m$ and $\mathbb{E}[Z_t^N] = tf(2)$.

We next consider $(S, S') \in A_{\delta,T}$ and write

$$\begin{aligned} |X_{S'}^{N,1} - X_S^{N,1}| &\leq \int_S^{S'} \int_0^\infty X_u^{N,1} \mathbf{1}_{\{z \leq f(X_u^{N,1})\}} \mathbf{N}^1(du, dz) + \frac{1}{N} \sum_{j=2}^N \int_S^{S'} \int_0^\infty \mathbf{1}_{\{z \leq f(X_u^{N,j})\}} \mathbf{N}^j(du, dz) \\ &\quad + \lambda \int_S^{S'} X_u^{N,1} du + \lambda \int_S^{S'} \bar{X}_u^N du \\ &=: I_{S,S'} + J_{S,S'} + K_{S,S'} + L_{S,S'}. \end{aligned}$$

We first note that $I_{S,S'} > 0$ implies that $\tilde{I}_{S,S'} := \int_S^{S'} \int_0^\infty \mathbf{1}_{\{z \leq f(X_u^{N,1})\}} \mathbf{N}^1(du, dz) \geq 1$, whence

$$\mathbb{P}(I_{S,S'} > 0) \leq \mathbb{P}(\tilde{I}_{S,S'} \geq 1) \leq \mathbb{E}[I_{S,S'}] \leq \mathbb{E} \left[\int_S^{S+\delta} f(X_u^{N,1}) du \right].$$

But for all $A > 0$, $f(x) \leq f(A) + A^{-1}xf(x)$ because f is nondecreasing. Hence, by (12),

$$\mathbb{P}(I_{S,S'} > 0) \leq \delta f(A) + \frac{1}{A} \int_0^T \mathbb{E}[X_u^{N,1} f(X_u^{N,1})] du \leq \delta f(A) + \frac{C_T}{A}.$$

Choosing $A = f^{-1}(\delta^{-1/2})$ (recall that $\lim_{\infty} f = \infty$ and consider a generalized notion of inverse function if necessary), we end with

$$\mathbb{P}(I_{S,S'} > 0) \leq \delta^{1/2} + \frac{C_T}{f^{-1}(\delta^{-1/2})}.$$

We proceed similarly to check that

$$\mathbb{E}[J_{S,S'}] \leq \frac{1}{N} \sum_{j=2}^N \mathbb{E} \left[\int_S^{S+\delta} f(X_u^{N,j}) du \right] \leq \delta^{1/2} + \frac{C_T}{f^{-1}(\delta^{-1/2})}.$$

Next, we write, for any $A > 0$, using that $x \leq A + xf(x)/f(A)$ and then (12),

$$\mathbb{E}[K_{S,S'}] \leq \lambda \mathbb{E} \left[\int_S^{S+\delta} X_u^{N,1} du \right] \leq \lambda \delta A + \frac{\lambda}{f(A)} \int_0^T \mathbb{E}[X_u^{N,1} f(X_u^{N,1})] du \leq \lambda \delta A + \frac{\lambda C_T}{f(A)}.$$

We choose $A = \delta^{-1/2}$ and get

$$\mathbb{E}[K_{S,S'}] \leq \lambda \delta^{1/2} + \frac{\lambda C_T}{f(\delta^{-1/2})}.$$

The same arguments show that

$$\mathbb{E}[L_{S,S'}] \leq \lambda \delta^{1/2} + \frac{\lambda C_T}{f(\delta^{-1/2})}.$$

We can now conclude that for $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}(|X_{S'}^{N,1} - X_S^{N,1}| > \varepsilon) &\leq \mathbb{P}(I_{S,S'} > 0) + \mathbb{P}(J_{S,S'} > \varepsilon/4) + \mathbb{P}(K_{S,S'} > \varepsilon/4) + \mathbb{P}(L_{S,S'} > \varepsilon/4) \\ &\leq \delta^{1/2} + \frac{C_T}{f^{-1}(\delta^{-1/2})} + \frac{4}{\varepsilon} \left((1 + 2\lambda)\delta^{1/2} + \frac{C_T}{f^{-1}(\delta^{-1/2})} + \frac{2\lambda C_T}{f(\delta^{-1/2})} \right). \end{aligned}$$

This last quantity does not depend on $N \geq 1$ nor on $(S, S') \in A_{\delta,T}$ and tends to 0 as $\delta \rightarrow 0$. This completes the proof. \square

We now turn to the consistency result.

Proof of Theorem 5(iii). We consider a probability distribution g_0 on \mathbb{R}_+ such that $\int_0^\infty x g_0(dx) < \infty$ and, for each $N \geq 1$, the unique solution $(X_t^N)_{t \geq 0}$ to (1) starting from some i.i.d. g_0 -distributed initial conditions $X_0^{N,i}$. We introduce $\mu_N = N^{-1} \sum_{i=1}^N \delta_{(X_t^{N,i})_{t \geq 0}}$, which is a $\mathcal{P}(\mathbb{D}(\mathbb{R}_+))$ -valued random variable. By Theorem 5(ii), this sequence is tight. We thus consider a (not relabeled) subsequence μ_N going in law to some $\mathcal{P}(\mathbb{D}(\mathbb{R}_+))$ -valued random variable μ . We want to show that μ a.s. belongs to $\mathcal{S} := \{\mathcal{L}((Y_t)_{t \geq 0}) : (Y_t)_{t \geq 0} \text{ solution to (2) with } \mathcal{L}(Y_0) = g_0 \text{ and satisfying } \int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty \text{ for all } t \geq 0\}$.

Step 1. For $t \geq 0$, we introduce $\pi_t : \mathbb{D}(\mathbb{R}_+) \mapsto \mathbb{R}_+$ defined by $\pi_t(\gamma) = \gamma_t$. We claim that $Q \in \mathcal{P}(\mathbb{D}(\mathbb{R}_+))$ belongs to \mathcal{S} if the following conditions are satisfied:

- (a) $Q \circ \pi_0^{-1} = g_0$;
- (b) for all $t \geq 0$, $\int_{\mathbb{D}(\mathbb{R}_+)} \int_0^t \gamma_s f(\gamma_s) ds Q(d\gamma) < \infty$;
- (c) for any $0 \leq s_1 < \dots < s_k < s < t$, any $\varphi_1, \dots, \varphi_k \in C_b(\mathbb{R}_+)$, any $\varphi \in C_b^2(\mathbb{R}_+)$,

$$\begin{aligned} F(Q) &:= \int_{\mathcal{D}(\mathbb{R}_+)} \int_{\mathcal{D}(\mathbb{R}_+)} Q(d\gamma) Q(d\tilde{\gamma}) \varphi_1(\gamma_{s_1}) \cdots \varphi_k(\gamma_{s_k}) \\ &\quad \times \left[\varphi(\gamma_t) - \varphi(\gamma_s) - \int_s^t f(\gamma_u) (\varphi(0) - \varphi(\gamma_u)) du - \int_s^t \varphi'(\gamma_u) [f(\tilde{\gamma}_u) + \lambda(\tilde{\gamma}_u - \gamma_u)] du \right] = 0. \end{aligned}$$

Indeed, consider a process $(Y_t)_{t \geq 0}$ of which the law Q satisfies the above three points. By (a), $\mathcal{L}(Y_0) = g_0$. By (b), $\int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty$ for all $t \geq 0$. Finally, (c) implies that for all $\varphi \in C_b^2(\mathbb{R}_+)$,

$$\varphi(Y_t) - \int_0^t [\varphi(0) - \varphi(Y_s)] f(Y_s) ds - \int_0^t \varphi'(Y_s) [\mathbb{E}[f(Y_s)] + \lambda(\mathbb{E}[Y_s] - Y_s)] ds$$

is a martingale. By Jacod and Shiryaev [17, Theorem II.2.42, page 86], this implies that Y is a semimartingale with characteristics (B, C, ν) given by

$$B_t = \int_0^t [\lambda(\mathbb{E}[Y_s] - Y_s) + \mathbb{E}[f(Y_s)] + Y_s f(Y_s)] ds, \quad C_t = 0, \quad \nu(ds, dx) = f(Y_{s-}) ds \delta_{-Y_{s-}}(dx).$$

We have chosen the truncation function $h(x) = x$ (i.e. *no truncation*) since Y possesses only large jumps. Finally, [17, Theorem III.2.26, page 157] implies that there is a Poisson measure $\mathbf{N}(ds, dz)$ on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $ds dz$ such that Y solves (2).

Step 2. Here we check that for any $t \geq 0$, a.s., $\mu(\{\gamma : \Delta\gamma(t) \neq 0\}) = 0$. We assume by contradiction that there exists $t > 0$ such that $\mu(\{\gamma : \Delta\gamma(t) \neq 0\}) > 0$ with positive probability. Hence there are $a, b > 0$ such that the event $E := \{\mu(\{\gamma : |\Delta\gamma(t)| > a\}) > b\}$ has a positive probability. For every $\varepsilon > 0$, we have $E \subset \{\mu(\mathcal{B}_a^\varepsilon) > b\}$, where $\mathcal{B}_a^\varepsilon :=$

$\{\gamma : \sup_{s \in (t-\varepsilon, t+\varepsilon)} |\Delta\gamma(s)| > a\}$, which is an open subset of $\mathbb{D}(\mathbb{R}_+)$. Thus $\mathcal{P}_{a,b}^\varepsilon := \{Q \in \mathcal{P}(\mathbb{D}(\mathbb{R}_+)) : Q(\mathcal{B}_a^\varepsilon) > b\}$ is an open subset of $\mathcal{P}(\mathbb{D}(\mathbb{R}_+))$. The Portmanteau theorem implies then that for any $\varepsilon > 0$,

$$\liminf_{N \rightarrow \infty} \mathbb{P}(\mu_N \in \mathcal{P}_{a,b}^\varepsilon) \geq \mathbb{P}(\mu \in \mathcal{P}_{a,b}^\varepsilon) \geq \mathbb{P}(E) > 0.$$

But, for all $N > 1/a$ (so that for each $i = 1, \dots, N$, the only jumps of $X^{N,i}$ that may exceed a are those produced by the Poisson measure \mathbf{N}^i),

$$\{\mu_N \in \mathcal{P}_{a,b}^\varepsilon\} \subset \left\{ \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\left\{ \int_{t-\varepsilon}^{t+\varepsilon} \mathbf{1}_{\{z \leq f(X_u^{N,i})\}} \mathbf{N}^i(du, dz) \geq 1 \right\}} \geq b \right\},$$

whence, using exchangeability,

$$\mathbb{P}(\mu_N \in \mathcal{P}_{a,b}^\varepsilon) \leq \frac{1}{bN} \sum_{i=1}^N \mathbb{E} \left(\int_{t-\varepsilon}^{t+\varepsilon} \mathbf{1}_{\{z \leq f(X_u^{N,i})\}} \mathbf{N}^i(du, dz) \right) = \frac{1}{b} \int_{t-\varepsilon}^{t+\varepsilon} \mathbb{E}[f(X_u^{N,1})] du.$$

We now observe that for any $A > 0$, $\int_{t-\varepsilon}^{t+\varepsilon} \mathbb{E}[f(X_u^{N,1})] du \leq 2\varepsilon f(A) + A^{-1} \int_{t-\varepsilon}^{t+\varepsilon} \mathbb{E}[X_u^{N,1} f(X_u^{N,1})] du \leq 2\varepsilon f(A) + CA^{-1}$ by (12). Choosing $A = f^{-1}(\varepsilon^{-1/2})$, we conclude that $\int_{t-\varepsilon}^{t+\varepsilon} \mathbb{E}[f(X_u^{N,1})] du \leq C(\sqrt{\varepsilon} + 1/f^{-1}(\varepsilon^{-1/2}))$, which does not depend on N and tends to 0 as $\varepsilon \rightarrow 0$. We thus have the contradiction

$$0 < \mathbb{P}(E) \leq \liminf_{\varepsilon \rightarrow 0} \liminf_{N \rightarrow \infty} \mathbb{P}(\mu_N \in \mathcal{P}_{a,b}^\varepsilon) = 0.$$

Step 3. Our limit μ a.s. satisfies (a), because $\mu \circ \pi_0^{-1}$ is the limit in law of $\mu^N \circ \pi_0^{-1} = N^{-1} \sum_{i=1}^N \delta_{X_0^{N,i}}$, which goes to g_0 because the $X_0^{N,i}$ are i.i.d. with common law g_0 . It also a.s. satisfies (b) since for all $t \geq 0$ and $K > 0$, using the Fatou lemma and (12),

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{D}(\mathbb{R}_+)} \int_0^t [(\gamma_s f(\gamma_s)) \wedge K] ds \mu(d\gamma) \right] &\leq \liminf_N \mathbb{E} \left[\int_{\mathbb{D}(\mathbb{R}_+)} \int_0^t [(\gamma_s f(\gamma_s)) \wedge K] ds \mu_N(d\gamma) \right] \\ &= \liminf_N N^{-1} \sum_{i=1}^N \int_0^t \mathbb{E}[X_s^{N,i} f(X_s^{N,i})] ds < \infty. \end{aligned}$$

The conclusion follows by letting $K \rightarrow \infty$.

Step 4. It remains to check that μ a.s. satisfies (c). We thus consider $F : \mathbb{D}(\mathbb{R}_+) \mapsto \mathbb{R}$ as in (c).

Step 4.1. Here we prove that $\lim_N \mathbb{E}[|F(\mu_N)|] = 0$. We have

$$\begin{aligned} F(\mu_N) &= \frac{1}{N} \sum_{i=1}^N \varphi_1(X_{s_1}^{N,i}) \cdots \varphi_k(X_{s_k}^{N,i}) \\ &\quad \times \left[\varphi(X_t^{N,i}) - \varphi(X_s^{N,i}) - \int_s^t f(X_u^{N,i}) [\varphi(0) - \varphi(X_u^{N,i})] du - \lambda \int_s^t \varphi'(X_u^{N,i}) (\bar{X}_u^N - X_u^{N,i}) du \right. \\ &\quad \left. - \int_s^t \varphi'(X_u^{N,i}) \frac{1}{N} \sum_{j=1}^N f(X_u^{N,j}) du \right]. \end{aligned}$$

But recalling (1) and using the Itô formula for jump processes,

$$\begin{aligned} \varphi(X_t^{N,i}) &= \varphi(X_0^{N,i}) + \int_0^t \int_0^\infty [\varphi(0) - \varphi(X_{u-}^{N,i})] \mathbf{1}_{\{z \leq f(X_{u-}^{N,i})\}} \mathbf{N}^i(du, dz) + \lambda \int_0^t \varphi'(X_u^{N,i}) (\bar{X}_u^N - X_u^{N,i}) du \\ &\quad + \sum_{j \neq i} \int_0^t \int_0^\infty \left(\varphi\left(X_{u-}^{N,i} + \frac{1}{N}\right) - \varphi(X_{u-}^{N,i}) \right) \mathbf{1}_{\{z \leq f(X_{u-}^{N,j})\}} \mathbf{N}^j(du, dz). \end{aligned}$$

Consequently, using the notation $\tilde{\mathbf{N}}^i(du, dz) = \mathbf{N}^i(du, dz) - du dz$ and setting

$$M_t^{N,i} := \int_0^t \int_0^\infty [\varphi(0) - \varphi(X_{u-}^{N,i})] \mathbf{1}_{\{z \leq f(X_{u-}^{N,i})\}} \tilde{\mathbf{N}}^i(du, dz),$$

$$\Delta_t^{N,i} := \sum_{j \neq i} \int_0^t \int_0^\infty \left(\varphi\left(X_{u-}^{N,i} + \frac{1}{N}\right) - \varphi(X_{u-}^{N,i}) \right) \mathbf{1}_{\{z \leq f(X_{u-}^{N,j})\}} \mathbf{N}^j(du, dz) - \int_0^t \varphi'(X_u^{N,i}) \frac{1}{N} \sum_{j=1}^N f(X_u^{N,j}) du,$$

we see that

$$F(\mu_N) = \frac{1}{N} \sum_{i=1}^N \varphi_1(X_{s_1}^{N,i}) \cdots \varphi_k(X_{s_k}^{N,i}) [(M_t^{N,i} - M_s^{N,i}) + (\Delta_t^{N,i} - \Delta_s^{N,i})].$$

Since the Poisson measures \mathbf{N}^i are i.i.d., the martingales $M^{N,i}$ are orthogonal. Using exchangeability and the boundedness of the φ_k , we thus find that

$$\mathbb{E}[|F(\mu_N)|] \leq C_F \frac{1}{\sqrt{N}} \mathbb{E}[(M_t^{N,1} - M_s^{N,1})^2]^{1/2} + C_F \mathbb{E}[|\Delta_t^{N,1}| + |\Delta_s^{N,1}|]. \tag{13}$$

First, since φ is bounded and using (12) (recall that $f(x) \leq f(1) + xf(x)$),

$$\mathbb{E}[(M_t^{N,1} - M_s^{N,1})^2] = \int_s^t \mathbb{E}[(\varphi(0) - \varphi(X_u^{N,1}))^2 f(X_u^{N,1})] du \leq C_F \int_0^t \mathbb{E}[f(X_u^{N,1})] du \leq C_F.$$

Next,

$$\begin{aligned} |\Delta_t^{N,1}| &\leq \int_0^t \int_0^\infty \left| \varphi\left(X_{u-}^{N,1} + \frac{1}{N}\right) - \varphi(X_{u-}^{N,1}) \right| \mathbf{1}_{\{z \leq f(X_{u-}^{N,1})\}} \mathbf{N}^1(du, dz) \\ &\quad + \left| \sum_{j=1}^N \int_0^t \int_0^\infty \left(\varphi\left(X_{u-}^{N,1} + \frac{1}{N}\right) - \varphi(X_{u-}^{N,1}) \right) \mathbf{1}_{\{z \leq f(X_{u-}^{N,j})\}} \tilde{\mathbf{N}}^j(du, dz) \right| \\ &\quad + \sum_{j=1}^N \int_0^t \left| \varphi\left(X_u^{N,1} + \frac{1}{N}\right) - \varphi(X_u^{N,1}) - \frac{1}{N} \varphi'(X_u^{N,1}) \right| f(X_u^{N,j}) du \\ &=: I_t^N + J_t^N + K_t^N. \end{aligned}$$

Using that φ' is bounded and (12), we find

$$\mathbb{E}[I_t^N] \leq \frac{C_F}{N} \int_0^t \mathbb{E}[f(X_u^{N,1})] du \leq \frac{C_F}{N}.$$

Moreover, since φ'' is bounded and by (12) again,

$$\mathbb{E}[K_t^N] \leq \frac{C_F}{N^2} \sum_{j=1}^N \int_0^t \mathbb{E}[f(X_u^{N,j})] du \leq \frac{C_F}{N}.$$

Finally, using the independence of the Poisson measures \mathbf{N}^j , that φ' is bounded and (12),

$$\mathbb{E}[(J_t^N)^2] = \sum_{j=1}^N \int_0^t \mathbb{E} \left[\left(\varphi\left(X_{u-}^{N,1} + \frac{1}{N}\right) - \varphi(X_{u-}^{N,1}) \right)^2 f(X_u^{N,j}) \right] du \leq \frac{C_F}{N^2} \sum_{j=1}^N \int_0^t \mathbb{E}[f(X_u^{N,j})] du \leq \frac{C_F}{N}.$$

All this implies that $\mathbb{E}[|\Delta_t^{N,1}|] \leq C_F/\sqrt{N}$ whence, coming back to (13), $\mathbb{E}[|F(\mu_N)|] \leq C_F/\sqrt{N}$.

Step 4.2. Clearly, F is continuous at any point $Q \in \mathcal{P}(\mathbb{D}(\mathbb{R}_+))$ such that $Q(\gamma : \Delta\gamma(s_1) = \dots = \Delta\gamma(s_k) = \Delta\gamma(s) = \Delta\gamma(t) = 0) = 1$ and such that $\int_{\mathbb{D}(\mathbb{R}_+)} \int_0^t [\gamma_u + f(\gamma_u)] du Q(d\gamma) < \infty$. Our limit point μ a.s. satisfies these two conditions by Steps 2 and 3 (because $x + f(x) \leq C(1 + xf(x))$). Since μ is the limit in law of μ_N and since F is a.s. continuous at μ , we thus deduce that for any $K > 0$, $\mathbb{E}[|F(\mu)| \wedge K] = \lim_N \mathbb{E}[|F(\mu_N)| \wedge K]$. Consequently, $\mathbb{E}[|F(\mu)| \wedge K] \leq \limsup_N \mathbb{E}[|F(\mu_N)|]$ for all $K > 0$. Using Step 4.1, we deduce that $\mathbb{E}[|F(\mu)| \wedge K] = 0$ for any $K > 0$. By the monotone convergence theorem, we conclude that $\mathbb{E}[F(\mu)] = 0$, whence $F(\mu) = 0$ a.s. \square

We can finally study the well posedness of the nonlinear SDE.

Proof of Theorem 4. Point (i) (weak existence assuming only Assumption 1 and that $\mathbb{E}[Y_0] < \infty$) follows from Theorem 5(ii)–(iii): we have built at least one weak solution, passing to the limit in the particle system, and we have seen that this solution satisfies that $\int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty$ for all $t \geq 0$.

For point (ii) (strong well posedness under Assumption 1 when $g_0 = \mathcal{L}(Y_0)$ is compactly supported), we only have to check that the solution built in point (i) satisfies that there is a deterministic locally bounded function $A : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that a.s., $Y_t \leq A(t)$ for all $t \geq 0$. This will conclude the proof, since such a weak existence result, together with the path-wise uniqueness proven in Proposition 16, will imply the strong well posedness. We thus assume that $\text{Supp } g_0 \subset [0, K]$ and set $A(t) := K + \int_0^t (\lambda \mathbb{E}[Y_s] + \mathbb{E}[f(Y_s)]) ds$, which is clearly locally bounded since $\int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty$ for all $t \geq 0$. Then it is obvious, recalling (2), that a.s., for all $t \geq 0$, $Y_t \leq A(t)$.

To check point (iii) (strong well posedness under Assumptions 1 and 3 when $\mathbb{E}[f(Y_0)] < \infty$), it suffices to prove that the solution built in point (i) satisfies $\sup_{[0,t]} \mathbb{E}[f(Y_s)] < \infty$ for all $t \geq 0$. Again, this weak existence, together with the strong uniqueness of Proposition 17, will complete the proof. Put $C(t) := \int_0^t (\lambda \mathbb{E}[Y_s] + \mathbb{E}[f(Y_s)]) ds$, which is again locally bounded, and observe from (2), that a.s., for all $t \geq 0$, $Y_t \leq Y_0 + C(t)$. Since $\mathbb{E}[f(Y_0)] < \infty$, we immediately conclude, using Remark 13(ii), that $\sup_{[0,t]} \mathbb{E}[f(Y_s)] < \infty$ for all $t \geq 0$, as desired. \square

Finally, we can give the

Proof of Theorem 5(iv). First grant Assumption 1 and assume that g_0 is compactly supported. We have seen in Theorem 5(ii)–(iii) that μ_N is tight and that any limit point μ a.s. belongs to $\mathcal{S} = \{\mathcal{L}((Y_t)_{t \geq 0}) : (Y_t)_{t \geq 0} \text{ solution to (2) with } \mathcal{L}(Y_0) = g_0 \text{ and satisfying } \int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty \text{ for all } t \geq 0\}$. But arguing as in the proof of Theorem 4(ii), we see that $\mathcal{S} = \mathcal{S}'$, where $\mathcal{S}' = \{\mathcal{L}((Y_t)_{t \geq 0}) : (Y_t)_{t \geq 0} \text{ solution to (2) with } \mathcal{L}(Y_0) = g_0 \text{ and such that a.s., for all } t \geq 0, Y_t \leq A(t) \text{ for some deterministic locally bounded function } A\}$. As seen in Theorem 4(ii), \mathcal{S}' is reduced to one point. The conclusion follows: μ_N goes in probability, as $N \rightarrow \infty$, to the unique element of \mathcal{S}' .

Next grant Assumptions 1 and 3 and assume that $\int_0^\infty f(y)g_0(dy) < \infty$. We have seen in Theorem 5(ii)–(iii) that μ_N is tight and that any limit point μ a.s. belongs to $\mathcal{S} = \{\mathcal{L}((Y_t)_{t \geq 0}) : (Y_t)_{t \geq 0} \text{ solution to (2) satisfying } \int_0^t \mathbb{E}[Y_s f(Y_s)] ds < \infty \text{ for all } t \geq 0\}$. But arguing as in the proof of Theorem 4(ii), we see that $\mathcal{S} = \mathcal{S}''$, where $\mathcal{S}'' = \{\mathcal{L}((Y_t)_{t \geq 0}) : (Y_t)_{t \geq 0} \text{ solution to (2) with } \mathcal{L}(Y_0) = g_0 \text{ and such that } \sup_{[0,t]} \mathbb{E}[f(Y_s)] < \infty \text{ for all } t \geq 0\}$. As seen in Theorem 4(iii), \mathcal{S}'' is reduced to one point. The conclusion follows. \square

5. Quantified propagation of chaos

The aim of this section is to prove Theorem 7. We thus impose Assumptions 1, 3 and 6 and we fix an initial distribution g_0 such that $\int_0^\infty f^2(x)g_0(dx) < \infty$. We consider an i.i.d. family $X_0^{N,i}$ of g_0 -distributed random variables, an i.i.d. family of Poisson measures $\mathbf{N}^i(ds, dz)$ on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity measure $ds dz$, we denote, for each $N \geq 1$, by $(X_t^N)_{t \geq 0} = (X_t^{N,1}, \dots, X_t^{N,N})_{t \geq 0}$ the solution to (1). Finally, we denote by $(Y_t^{N,i})_{t \geq 0}$, for every $N \geq 1$, every $i = 1, \dots, N$, the path-wise unique (thanks to Theorem 4(iii)) solution to (2) starting from $X_0^{N,i}$ and driven by the Poisson measure \mathbf{N}^i . Obviously, for every $N \geq 1$, the processes $(Y_t^{N,i})_{t \geq 0}$, $i = 1, \dots, N$, are i.i.d.

To prove Theorem 7, we will essentially mimic the path-wise uniqueness proof of Theorem 4 to control $\sup_{[0,T]} \mathbb{E}[|H(X_t^{N,1}) - H(Y_t^{N,1})|]$ by C_T/\sqrt{N} . But there are a number of technical difficulties. First, we need to work on $[0, \tau_N^T]$, for some well-chosen stopping time τ_N^T that is asymptotically greater than T . Next, we will rather

study $\mathbb{E}[(N^{-1} + (H(X_t^{N,1}) - H(Y_t^{N,1}))^2)^{1/2}]$: this changes nothing to the result, but allows for a more rigorous proof (we apply the Itô formula to a true C^2 function) and allows for the control of a second derivative, see Lemma 20(i), that would explode without the additional N^{-1} term. We start with some more moment estimates.

Lemma 19.

(i) For all $T > 0$, there is C_T depending only on T, λ, g_0 and f such that

$$\mathbb{E}\left[\sup_{[0,T]} f^2(Y_t^{N,1})\right] \leq C_T \quad \text{and} \quad \sup_N \mathbb{E}\left[\sup_{[0,T]} f^2(X_t^{N,1})\right] \leq C_T.$$

(ii) For all $T \geq 1$, we can find a constant $R_T > 0$ such that the stopping time

$$\tau_N^T := \inf\left\{t \geq 0 : N^{-1} \sum_{i=1}^N (f(X_t^{N,i}) + f(Y_t^{N,i})) \geq R_T\right\}$$

satisfies, for some constants $C > 0$ (and C_T) depending only on λ, g_0 and f (and T).

$$\mathbb{P}(\tau_N^T \leq T) \leq \frac{C}{N} \quad \text{and} \quad \mathbb{E}\left[\sup_{[0,T]} (1 + f(X_t^{N,1}) + f(Y_t^{N,1})) \mathbf{1}_{\{\tau_N^T \leq T\}}\right] \leq \frac{C_T}{\sqrt{N}}.$$

Proof. Recalling (3), it a.s. holds that for all $t \geq 0, Y_t^{N,1} \leq X_0^{N,1} + C(1 + t)$. Using Remark 13(ii) and that $\mathbb{E}[f^2(X_0^{N,1})] = \int_0^\infty f^2(x)g_0(dx) < \infty$, we immediately deduce that $\mathbb{E}[\sup_{[0,T]} f^2(Y_t^{N,1})] \leq C_T$.

Next, (5) tells us that a.s., for all $t \geq 0, X_t^{N,1} \leq X_0^{N,1} + C(1 + T)(\bar{X}_0^N + Z_T^N)$. By Remark 13(iv),

$$\sup_{[0,T]} f^2(X_t^{N,1}) \leq C_T(1 + f^2(X_0^{N,1}) + f^2(\bar{X}_0^N) + f^2(Z_T^N)).$$

But f^2 being convex, $f^2(\bar{X}_0^N) \leq N^{-1} \sum_{i=1}^N f^2(X_0^{N,i})$. Consequently, $\mathbb{E}[\sup_{[0,T]} f^2(X_t^{N,1})] \leq C_T(1 + \int_0^\infty f^2(x) \times g_0(dx) + \mathbb{E}[f^2(Z_T^N)])$. To end the proof of (i), it suffices to recall that Z_T^N is the mean of N i.i.d. Poisson($Tf(2)$)-distributed random variables: since $f(x) \leq Ce^{Cx}$ by Remark 13(iii), a simple computation shows that indeed, $\sup_N \mathbb{E}[f^2(Z_T^N)] < \infty$.

Using again (3) and (5), we see that a.s., for all $t \in [0, T]$, all $i = 1, \dots, N, X_t^{N,i} \leq X_0^{N,i} + C(1 + T)(\bar{X}_0^N + Z_T^N)$ and $Y_t^{N,i} \leq X_0^{N,i} + C(1 + t)$. Consequently, using Remark 13(iv) and the convexity of f (whence $f(\bar{X}_0^N) \leq N^{-1} \sum_{i=1}^N f(X_0^{N,i})$),

$$\sup_{[0,T]} \frac{1}{N} \sum_{i=1}^N (f(X_t^{N,i}) + f(Y_t^{N,i})) \leq C_T \left(1 + f(Z_T^N) + \frac{1}{N} \sum_{i=1}^N f(X_0^{N,i})\right).$$

The bounds $\mathbb{P}(Z_T^N \geq 2f(2)T) \leq \exp(-NTf(2)(3 - e))$, see (9), and

$$\mathbb{P}\left(\frac{1}{N} \sum_{i=1}^N f(X_0^{N,i}) \geq \int_0^\infty f(x)g_0(dx) + 1\right) \leq \frac{\text{Var}(f(X_0^{N,1}))}{N} \leq \frac{C}{N}$$

imply that, with the choice $R_T = C_T(1 + f(2f(2)T) + \int_0^\infty f(x)g_0(dx) + 1)$,

$$\mathbb{P}(\tau_N^T \leq T) \leq \exp(-NTf(2)(3 - e)) + C/N \leq C/N$$

as desired. The last inequality immediately follows, using (i) and the Cauchy–Schwarz inequality. □

We carry on with a technical lemma similar to Lemma 18.

Lemma 20. Grant Assumptions 1 and 3 and recall that $H(x) = f(x) + \arctan x$. Define, for $N \geq 1$, $a_N(x, y) := [N^{-1} + (H(x) - H(y))^2]^{1/2}$.

- (i) It holds that $|\partial_x a_N(x, y)| \leq H'(x)$ and $|\partial_{xx} a_N(x, y)| \leq |H''(x)| + \sqrt{N}(H'(x))^2$.
- (ii) We have $|\partial_x a_N(x, y) + \partial_y a_N(x, y)| \leq |H'(x) - H'(y)|$.
- (iii) There is $C > 0$ such that $-[x\partial_x a_N(x, y) + y\partial_y a_N(x, y)] \leq C a_N(x, y)$.
- (iv) Finally, there is $C > 0$ such that

$$\begin{aligned} \Delta_N(x, y) &:= (f(x) \wedge f(y)) [a_N(0, 0) - a_N(x, y)] + (f(x) - f(y))_+ [a_N(0, y) - a_N(x, y)] \\ &\quad + (f(y) - f(x))_+ [a_N(x, 0) - a_N(x, y)] \\ &\leq \frac{f(x) \wedge f(y)}{\sqrt{N}} + C a_N(x, y). \end{aligned}$$

Proof. Points (i) and (ii) follow from direct computations. For (iii), using the expression of H ,

$$\begin{aligned} -[x\partial_x a_N(x, y) + y\partial_y a_N(x, y)] &= \frac{-(H(x) - H(y))}{[N^{-1} + (H(x) - H(y))^2]^{1/2}} [xH'(x) - yH'(y)] \\ &= \frac{-(H(x) - H(y))}{[N^{-1} + (H(x) - H(y))^2]^{1/2}} [xf'(x) - yf'(y)] \\ &\quad + \frac{-(H(x) - H(y))}{[N^{-1} + (H(x) - H(y))^2]^{1/2}} \left[\frac{x}{1+x^2} - \frac{y}{1+y^2} \right]. \end{aligned}$$

The first term on the RHS is nonpositive, because both $H(x)$ and $xf'(x)$ are nondecreasing. The second one is roughly bounded by $|x/(1+x^2) - y/(1+y^2)| \leq |x - y|$ which is bounded, recalling Lemma 18(ii), by $C|H(x) - H(y)| \leq C a_N(x, y)$. To prove (iv), we first observe, since a_N is symmetric and f is nondecreasing, that

$$\Delta_N(x, y) = \frac{f(x) \wedge f(y)}{\sqrt{N}} - (f(x) \vee f(y)) a_N(x, y) + |f(x) - f(y)| a_N(0, x \wedge y).$$

Noting that $|f(x) - f(y)| \leq |H(x) - H(y)| \leq a_N(x, y)$, we deduce that

$$\Delta_N(x, y) \leq \frac{f(x) \wedge f(y)}{\sqrt{N}} + a_N(x, y) (a_N(0, x \wedge y) - f(x) \vee f(y)).$$

The conclusion follows, since $a_N(0, x \wedge y) - f(x) \vee f(y) \leq N^{-1/2} + H(x) \wedge H(y) - f(x) \vee f(y)$, which is obviously bounded by $1 + \pi/2$. □

We are now ready to give the

Proof of Theorem 7. We fix $T > 0$ and define R_T and τ_N^T as in Lemma 19(ii). In the whole proof, we work on the time interval $[0, T]$. Recall that a_N and Δ_N were defined in Lemma 20.

Step 1. This is the main step of the proof. We show that there is a constant C_T such that for all $N \geq 1$, $\sup_{[0, T]} \mathbb{E}[a_N(X_{t \wedge \tau_N^T}^{N,1}, Y_{t \wedge \tau_N^T}^{N,1})] \leq C_T N^{-1/2}$. Applying the Itô formula for jump processes, we find that

$$\mathbb{E}[a_N(X_{t \wedge \tau_N^T}^{N,1}, Y_{t \wedge \tau_N^T}^{N,1})] = N^{-1/2} + I + J + \lambda K + \lambda L,$$

where

$$I = \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \Delta_N(X_s^{N,1}, Y_s^{N,1}) ds \right],$$

$$\begin{aligned}
 J &= \sum_{j=2}^N \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} f(X_s^{N,j}) \left[a_N \left(X_s^{N,1} + \frac{1}{N}, Y_s^{N,1} \right) - a_N(X_s^{N,1}, Y_s^{N,1}) \right] ds \right] \\
 &\quad + \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_y a_N(X_s^{N,1}, Y_s^{N,1}) \mathbb{E}[f(Y_s^{N,1})] ds \right], \\
 K &= -\mathbb{E} \left[\int_0^{t \wedge \tau_N^T} (\partial_x a_N(X_s^{N,1}, Y_s^{N,1}) X_s^{N,1} + \partial_y a_N(X_s^{N,1}, Y_s^{N,1}) Y_s^{N,1}) ds \right], \\
 L &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} (\partial_x a_N(X_s^{N,1}, Y_s^{N,1}) \bar{X}_s^N + \partial_y a_N(X_s^{N,1}, Y_s^{N,1}) \mathbb{E}[Y_s^{N,1}]) ds \right].
 \end{aligned}$$

By Lemma 20(iv) and Lemma 19(i),

$$I \leq \frac{1}{\sqrt{N}} \int_0^t \mathbb{E}[f(Y_s^{N,1})] ds + C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} a_N(X_s^{N,1}, Y_s^{N,1}) ds \right] \leq \frac{C_T}{\sqrt{N}} + C \int_0^t \mathbb{E}[a_N(X_{s \wedge \tau_N^T}^{N,1}, Y_{s \wedge \tau_N^T}^{N,1})] ds.$$

Lemma 20(iii) implies that

$$K \leq C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} a_N(X_s^{N,1}, Y_s^{N,1}) ds \right] \leq C \int_0^t \mathbb{E}[a_N(X_{s \wedge \tau_N^T}^{N,1}, Y_{s \wedge \tau_N^T}^{N,1})] ds.$$

We next write $L = L_1 + L_2 + L_3$, with

$$\begin{aligned}
 L_1 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_x a_N(X_s^{N,1}, Y_s^{N,1}) [\bar{X}_s^N - \bar{Y}_s^N] ds \right], \\
 L_2 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} [\partial_x a_N(X_s^{N,1}, Y_s^{N,1}) + \partial_y a_N(X_s^{N,1}, Y_s^{N,1})] \bar{Y}_s^N ds \right], \\
 L_3 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_y a_N(X_s^{N,1}, Y_s^{N,1}) (\mathbb{E}[Y_s^{N,1}] - \bar{Y}_s^N) ds \right].
 \end{aligned}$$

Using the Cauchy–Schwarz inequality, Lemma 20(i) and the fact that the $Y_s^{N,i}$ are i.i.d.,

$$L_3 \leq \frac{1}{\sqrt{N}} \int_0^t \mathbb{E}[H'(Y_s^{N,1})^2]^{1/2} (\text{Var } Y_s^{N,1})^{1/2} ds \leq \frac{C_T}{\sqrt{N}}.$$

We used Lemma 18(i), which tells us that $x + H'(x) \leq C(1 + f(x))$, whence $\sup_{[0,T]} \mathbb{E}[H'(Y_s^{N,1})^2] \leq C_T$ and $\sup_{[0,T]} \text{Var } Y_s^{N,1} \leq C_T$ by Lemma 19(i). Next, Lemmas 20(ii) and 18(ii) tell us that $|\partial_x a_N(x, y) + \partial_y a_N(x, y)| \leq |H'(x) - H'(y)| \leq C|H(x) - H(y)| \leq Ca_N(x, y)$. Consequently,

$$L_2 \leq C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} |\bar{Y}_s^N| a_N(X_s^{N,1}, Y_s^{N,1}) ds \right] \leq C_T \int_0^t \mathbb{E}[a_N(X_{s \wedge \tau_N^T}^{N,1}, Y_{s \wedge \tau_N^T}^{N,1})] ds.$$

We used that, by definition of τ_N^T and since $x \leq C(1 + f(x))$ (see Lemma 18(i)), $|\bar{Y}_s^N| \leq C(1 + N^{-1} \sum_{i=1}^N f(X_s^{N,i})) \leq C(1 + R_T)$ for all $s \in [0, \tau_N^T]$ a.s. Finally, using that τ_N^T does not break the exchangeability and Lemma 20(i), we write

$$\begin{aligned}
 L_1 &= \frac{1}{N} \sum_{j=1}^N \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_x a_N(X_s^{N,1}, Y_s^{N,1}) [X_s^{N,j} - Y_s^{N,j}] ds \right] \\
 &= \frac{1}{N} \sum_{j=1}^N \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_x a_N(X_s^{N,j}, Y_s^{N,j}) [X_s^{N,1} - Y_s^{N,1}] ds \right]
 \end{aligned}$$

$$\begin{aligned} &\leq \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \left(\frac{1}{N} \sum_{j=1}^N H'(X_s^{N,j}) \right) |X_s^{N,1} - Y_s^{N,1}| ds \right] \\ &\leq C_T \int_0^t \mathbb{E}[a_N(X_{s \wedge \tau_N^T}^{N,1}, Y_{s \wedge \tau_N^T}^{N,1})] ds. \end{aligned}$$

The last inequality uses that, by definition of τ_N^T and since $H'(x) \leq C(1 + f(x))$ (see Lemma 18(i)), $|N^{-1} \times \sum_{j=1}^N H'(X_s^{N,j})| \leq C_T$ for all $s \in [0, \tau_N^T]$ a.s. It also uses that $|x - y| \leq C|H(x) - H(y)| \leq Ca_N(x, y)$ by Lemma 18(ii).

We finally write $J = J_1 + J_2 + J_3 + J_4$, where, using again exchangeability,

$$\begin{aligned} J_1 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} f(X_s^{N,2}) \left((N-1) \left[a_N \left(X_s^{N,1} + \frac{1}{N}, Y_s^{N,1} \right) - a_N(X_s^{N,1}, Y_s^{N,1}) \right] - \partial_x a_N(X_s^{N,1}, Y_s^{N,1}) \right) ds \right], \\ J_2 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} f(X_s^{N,2}) (\partial_x a_N(X_s^{N,1}, Y_s^{N,1}) + \partial_y a_N(X_s^{N,1}, Y_s^{N,1})) ds \right], \\ J_3 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_y a_N(X_s^{N,1}, Y_s^{N,1}) [f(Y_s^{N,2}) - f(X_s^{N,2})] ds \right], \\ J_4 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_y a_N(X_s^{N,1}, Y_s^{N,1}) [\mathbb{E}[f(Y_s^{N,2})] - f(Y_s^{N,2})] ds \right]. \end{aligned}$$

We start with J_1 . Using Lemma 20(i),

$$\begin{aligned} &|(N-1)[a_N(x + 1/N, y) - a_N(x, y)] - \partial_x a_N(x, y)| \\ &\leq |a_N(x + 1/N, y) - a_N(x, y)| + |N[a_N(x + 1/N, y) - a_N(x, y)] - \partial_x a_N(x, y)| \\ &\leq N^{-1} \sup_{z \in [x, x+1/N]} [|\partial_x a_N(z, y)| + |\partial_{xx} a_N(z, y)|] \\ &\leq N^{-1} \sup_{z \in [x, x+1/N]} [H'(z) + |H''(z)| + \sqrt{N}(H'(z))^2] \\ &\leq CN^{-1/2}(1 + f^2(x)). \end{aligned}$$

The last inequality uses that $|H''(x)| \leq CH'(x)$ (see Lemma 18), the fact that $H'(x) \leq C(1 + f(x))$ (see Lemma 18(i)) and that $\sup_{[x, x+1/N]} f(z) \leq C(1 + f(x))$ (see Remark 13(iv)). Consequently,

$$J_1 \leq \frac{C}{\sqrt{N}} \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} f(X_s^{N,2})(1 + f^2(X_s^{N,1})) ds \right].$$

By Lemmas 20(ii) and 18(ii), $|\partial_x a_N(x, y) + \partial_y a_N(x, y)| \leq |H'(x) - H'(y)| \leq C|H(x) - H(y)| \leq Ca_N(x, y)$. Hence

$$J_2 \leq C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} f(X_s^{N,2}) a_N(X_s^{N,1}, Y_s^{N,1}) \right].$$

Lemmas 20(i) and 18(i) imply that $|\partial_y a_N(x, y)| \leq H'(y) \leq C(1 + f(y))$ and we obviously have $|f(x) - f(y)| \leq |H(x) - H(y)| \leq a_N(x, y)$. It follows that

$$J_3 \leq C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} (1 + f(Y_s^{N,1})) a_N(X_s^{N,2}, Y_s^{N,2}) \right] = C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} (1 + f(Y_s^{N,2})) a_N(X_s^{N,1}, Y_s^{N,1}) \right].$$

We have checked that

$$J_1 + J_2 + J_3 \leq C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} (1 + f(X_s^{N,2}) + f(Y_s^{N,2})) \left(a_N(X_s^{N,1}, Y_s^{N,1}) + \frac{1 + f^2(X_s^{N,1})}{\sqrt{N}} \right) ds \right].$$

Using exchangeability and then the definition of τ_N^T , we thus can write

$$\begin{aligned} J_1 + J_2 + J_3 &\leq C \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \left(\frac{1}{N} \sum_{j=1}^N (1 + f(X_s^{N,j}) + f(Y_s^{N,j})) \right) \left(a_N(X_s^{N,1}, Y_s^{N,1}) + \frac{1 + f^2(X_s^{N,1})}{\sqrt{N}} \right) ds \right] \\ &\leq C(1 + R_T) \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \left(a_N(X_s^{N,1}, Y_s^{N,1}) + \frac{1 + f^2(X_s^{N,1})}{\sqrt{N}} \right) ds \right] \\ &\leq C_T \int_0^t \mathbb{E} [a_N(X_{s \wedge \tau_N^T}^{N,1}, Y_{s \wedge \tau_N^T}^{N,1})] ds + \frac{C_T}{\sqrt{N}}. \end{aligned}$$

The last inequality uses that $\sup_N \sup_{[0,T]} \mathbb{E}[f^2(X_t^{N,1})] < \infty$ by Lemma 19(i). Finally, using again exchangeability, that $|\partial_y a_N(x, y)| \leq C(1 + f(y))$, the Cauchy–Schwarz inequality and that the $Y_s^{N,i}$ are i.i.d.,

$$\begin{aligned} J_4 &= \mathbb{E} \left[\int_0^{t \wedge \tau_N^T} \partial_y a_N(X_s^{N,1}, Y_s^{N,1}) \left(\mathbb{E}[f(Y_s^{N,2})] - \frac{1}{N-1} \sum_{j=2}^N f(Y_s^{N,j}) \right) ds \right] \\ &\leq C \int_0^t \mathbb{E} [(1 + f(Y_s^{N,1}))^2]^{1/2} \frac{[\text{Var } f(Y_s^{N,1})]^{1/2}}{\sqrt{N-1}} ds. \end{aligned}$$

Again, we conclude that $J_4 \leq C_T N^{-1/2}$ since $\sup_{[0,T]} \mathbb{E}[f^2(Y_t^{N,1})] < C_T$, as shown in Lemma 19.

All in all, we have checked that $\mathbb{E}[a_N(X_{t \wedge \tau_N^T}^{N,1}, Y_{t \wedge \tau_N^T}^{N,1})] \leq C_T N^{-1/2} + C_T \int_0^t \mathbb{E}[a_N(X_{s \wedge \tau_N^T}^{N,1}, Y_{s \wedge \tau_N^T}^{N,1})] ds$. We conclude the step using the Gronwall lemma.

Step 2. It is not hard to complete the proof. First, gathering Step 1 (recall that $|H(x) - H(y)| \leq a_N(x, y)$) and Lemma 19(ii) (recall that $H(x) \leq \pi/2 + f(x)$), we find, for all $t \in [0, T]$,

$$\mathbb{E}[|H(X_t^{N,1}) - H(Y_t^{N,1})|] \leq \mathbb{E}[|H(X_{t \wedge \tau_N^T}^{N,1}) - H(Y_{t \wedge \tau_N^T}^{N,1})|] + \mathbb{E}[(H(X_t^{N,1}) + H(Y_t^{N,1})) \mathbf{1}_{\{t \wedge \tau_N^T < t\}}] \leq \frac{C_T}{\sqrt{N}}.$$

Moreover, $|x - y| \leq C|H(x) - H(y)|$ by Lemma 18(ii), whence $\sup_{[0,T]} \mathbb{E}[|X_t^{N,1} - Y_t^{N,1}|] \leq C_T N^{-1/2}$.

We next assume additionally that $\int_0^\infty y^{2+\varepsilon} g_0(dy) < \infty$ for some $\varepsilon > 0$. Recalling (3), this obviously implies that $\sup_{[0,T]} \mathbb{E}[(Y_t^{N,1})^{2+\varepsilon}] \leq C_T$. Since the $Y_t^{N,i}$ are i.i.d. \mathbb{R} -valued random variables, it is well known, see e.g. [10, Theorem 1 with $d = 1, p = 1, q = 2 + \varepsilon$], that

$$\mathbb{E} \left[\mathcal{W}_1 \left(N^{-1} \sum_{i=1}^N \delta_{Y_t^{N,i}}, \mathcal{L}(Y_t^{N,1}) \right) \right] \leq \frac{C \mathbb{E}[(Y_t^{N,1})^{2+\varepsilon}]^{1/(2+\varepsilon)}}{\sqrt{N}} \leq \frac{C_T}{\sqrt{N}}.$$

But it follows from exchangeability that

$$\mathbb{E} \left[\mathcal{W}_1 \left(N^{-1} \sum_{i=1}^N \delta_{X_t^{N,i}}, N^{-1} \sum_{i=1}^N \delta_{Y_t^{N,i}} \right) \right] \leq \frac{1}{N} \sum_{i=1}^N \mathbb{E}[|X_t^{N,i} - Y_t^{N,i}|] = \mathbb{E}[|X_t^{N,1} - Y_t^{N,1}|] \leq \frac{C_T}{\sqrt{N}}.$$

Using the triangular inequality for \mathcal{W}_1 , we conclude that for all $t \in [0, T]$,

$$\mathbb{E} \left[\mathcal{W}_1 \left(N^{-1} \sum_{i=1}^N \delta_{X_t^{N,i}}, \mathcal{L}(Y_t^{N,1}) \right) \right] \leq \frac{C_T}{\sqrt{N}}$$

as desired. □

6. Invariant distributions

Here we prove Theorem 7. We thus only impose Assumption 1. We start with the following remark.

Proposition 21. *Let \mathbf{N} be a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $ds dz$, let $\lambda \geq 0$ and $a \geq 0$.*

(i) *The \mathbb{R}_+ -valued SDE*

$$Z_t = Z_0 - \int_0^t \int_0^\infty Z_s - \mathbf{1}_{\{z \leq f(Z_{s-})\}} \mathbf{N}(ds, dz) + \int_0^t (a - \lambda Z_s) ds \tag{14}$$

has a path-wise unique solution for every nonnegative initial condition Z_0 .

(ii) *Furthermore, (14) has a unique invariant probability measure g_a . It is given by $g_0 = \delta_0$ if $a = 0$ and by $g_a(dx) = g_a(x) dx$ if $a > 0$, where (with the convention that $a/\lambda = \infty$ if $\lambda = 0$),*

$$g_a(x) = \frac{p_a}{a - \lambda x} \exp\left(-\int_0^x \frac{f(y)}{a - \lambda y} dy\right) \mathbf{1}_{\{0 \leq y < a/\lambda\}},$$

where $p_a > 0$ is such that $\int_0^\infty g_a(x) dx = 1$. It automatically holds that $\int_0^\infty f(x)g_a(dx) = p_a$.

Proof. Point (i) is straightforward. All the coefficients being locally Lipschitz-continuous, we have *local* strong existence and uniqueness, i.e. strong existence and uniqueness on $[0, \tau)$, where $\tau = \inf\{t \geq 0 : Z_t = \infty\}$. But having a look at (14), we see that a.s., for all $t \geq 0$, $Z_t \leq Z_0 + at$. Hence $\tau = \infty$ a.s.

Point (ii) is straightforward if $a = 0$. Indeed, δ_0 is clearly an invariant distribution. It is unique, because for any initial condition, Z_t tends a.s. to 0 as $t \rightarrow \infty$. Indeed, if $\lambda > 0$, then $0 \leq Z_t \leq e^{-\lambda t} Z_0$. If now $\lambda = 0$, then $Z_t = Z_0 \mathbf{1}_{\{t < \tau_0\}}$, where τ_0 follows an exponential distribution with parameter $f(Z_0)$ (conditionally on Z_0).

We next prove (ii) when $a > 0$. We first claim that the homogeneous Markov process Z has exactly one invariant probability distribution which is supported in $[0, a/\lambda]$ (or $[0, \infty)$ if $\lambda = 0$). This follows from the classical theory of Markov processes, since 0 is a positive Harris recurrent state of Z . Indeed, let $\tau_0 = \inf\{t \geq 0 : Z_t = 0\}$. Then for any initial condition $z > 0$, $\mathbb{E}_z(\tau_0) < \infty$. This can be easily checked, using e.g. that starting from $z > 0$, $Z_t = e^{-\lambda t} z + (1 - e^{-\lambda t})a/\lambda \geq \min\{z, a/\lambda\}$ for all $t \in [0, \tau_0)$, so that Z jumps to zero with a rate bounded from below by $\min\{f(z), f(a/\lambda)\} > 0$. As a consequence, the successive jump times of Z to 0 induce a regeneration scheme, and Z is positive Harris recurrent implying the uniqueness of the invariant probability measure. Moreover, it is clear that $Z_t \leq a/\lambda$ for every $t \geq \tau_0$, which implies that the support of the invariant probability is included in $[0, a/\lambda]$.

It thus only remains to check that g_a is indeed an invariant probability measure for (14). The computations below include the case where $\lambda = 0$. It suffices to prove that for all $\phi \in C_b^1(\mathbb{R}_+)$,

$$\int_0^\infty [\phi(0) - \phi(x)] f(x) g_a(dx) + \int_0^\infty \phi'(x)(a - \lambda x) g_a(dx) = 0. \tag{15}$$

Indeed, the infinitesimal generator associated to the SDE (14) is given by $\mathcal{L}_a \phi(x) = [\phi(0) - \phi(x)] f(x) + \phi'(x)(a - \lambda x)$. First, a direct computation shows that

$$\int_0^\infty f(x) g_a(dx) = p_a \int_0^{a/\lambda} \frac{f(x)}{a - \lambda x} \exp\left(-\int_0^x \frac{f(y)}{a - \lambda y} dy\right) dx = -p_a \left[\exp\left(-\int_0^x \frac{f(y)}{a - \lambda y} dy\right) \right]_{x=0}^{x=a/\lambda} = p_a.$$

The last equality uses that $f(a/\lambda) > 0$. Hence, (15) reduces to

$$\int_0^{a/\lambda} \phi(x) f(x) g_a(x) dx - \int_0^{a/\lambda} \phi'(x)(a - \lambda x) g_a(x) dx = \phi(0) p_a.$$

Proceeding to an integration by parts in the second integral and using that $f(x)g_a(x) + [(a - \lambda x)g_a(x)]' = 0$ for all $x \in (0, a/\lambda)$, we see that (15) reduces to

$$-[\phi(x)(a - \lambda x)g_a(x)]_{x=0}^{x=a/\lambda} = \phi(0) p_a.$$

This is easily checked, since $ag_a(0) = p_a$ and since $\lim_{x \uparrow a/\lambda} (a - \lambda x)g_a(x) = 0$. \square

We now study for which values of a an invariant measure of (14) is an invariant measure of (2).

Lemma 22. *Adopt the notation of Proposition 21. When $a = 0$, we define $p_0 = 0 = \int_0^\infty f(x)g_0(dx)$. We also introduce, for $a \geq 0$, $m_a := \int_0^\infty xg_a(dx)$. The equation $a = \lambda m_a + p_a$ has the solution $a = 0$ and exactly one positive solution a_* . Furthermore, it holds that $a_* > \lambda$.*

Proof. The proof below works whenever $\lambda > 0$ or $\lambda = 0$. Evidently, $a = 0$ solves $a = \lambda m_a + p_a$. Let now $a > 0$. Since $\int_0^\infty g_a(dx) = 1$, we have

$$\frac{1}{p_a} = \int_0^{a/\lambda} \frac{1}{a - \lambda x} \exp\left(-\int_0^x \frac{f(y)}{a - \lambda y} dy\right) dx =: \Gamma_1(a).$$

Next,

$$m_a = p_a \int_0^{a/\lambda} \frac{x}{a - \lambda x} \exp\left(-\int_0^x \frac{f(y)}{a - \lambda y} dy\right) dx =: p_a \Gamma_2(a).$$

Hence a solves $a = \lambda m_a + p_a$ if and only if $a/p_a - \lambda m_a/p_a = 1$, i.e. $a\Gamma_1(a) - \lambda\Gamma_2(a) = 1$, i.e.

$$\Gamma(a) := a\Gamma_1(a) - \lambda\Gamma_2(a) = \int_0^{a/\lambda} \exp\left(-\int_0^x \frac{f(y)}{a - \lambda y} dy\right) dx = 1.$$

But Γ is continuous and strictly increasing, $\Gamma(0) = 0$ and $\Gamma(\infty) = \infty$, so that the equation $\Gamma(a) = 1$ has exactly one solution a_* . Finally, we obviously have $\Gamma(\lambda) < 1$, so that $a_* > \lambda$. \square

We are now able to give the

Proof of Theorem 8. Consider an invariant probability measure g , supported by \mathbb{R}_+ , for the nonlinear SDE (2). Let $Y_0 \sim g$ and consider $(Y_t)_{t \geq 0}$ solution to (2). Then for all $t \geq 0$, $Y_t \sim g$, so that $\mathbb{E}[Y_t] = m$ and $\mathbb{E}[f(Y_t)] = p$, where $m = \int_0^\infty xg(dx)$ and $p = \int_0^\infty f(x)g(dx)$. Consequently, $(Y_t)_{t \geq 0}$ solves (14) with $a = p + \lambda m$. Since $(Y_t)_{t \geq 0}$ is stationary, we deduce from Proposition 21 that $g = g_a$. But of course we have the constraint that $a = p_a + \lambda m_a$, whence $a = 0$ or $a = a_*$ by Lemma 22. Hence either $g = \delta_0$ or $g = g_{a_*}$.

Consider now $Y_0 \sim g$, with $g = \delta_0$ or $g = g_{a_*}$. Then the solution $(Y_t)_{t \geq 0}$ to (14) (with $a = 0$ or $a = a_*$) is stationary by Proposition 21. Since furthermore $\mathbb{E}[\lambda Y_t + f(Y_t)] = \int_0^\infty [\lambda x + f(x)]g(dx) = \lambda m_a + p_a = a$ by Lemma 22 since $a = 0$ or $a = a_*$, we conclude that $(Y_t)_{t \geq 0}$ also solves (2). Consequently, g is an invariant measure for (2).

We thus have checked that (2) has exactly two invariant probability distributions, which are δ_0 and g_{a_*} . Finally g_{a_*} is indeed the probability measure g defined in the statement (where $p = p_{a_*}$ and $m = m_{a_*}$) and we have that $m + p/\lambda = a_*/\lambda > 1$. \square

7. Shape of the time-marginals and large-time behavior

The aim of this section is to prove Theorem 12 and Propositions 11 and 9. We thus consider $\lambda \geq 0$, grant Assumptions 1 and 3 and suppose that $\mathbb{E}[f^2(Y_0)] < \infty$ and $\mathbb{P}(Y_0 = 0) < 1$. We consider the unique solution $(Y_t)_{t \geq 0}$ to (2), we set $p_t = \mathbb{E}[f(Y_t)]$, $m_t = \mathbb{E}[Y_t]$, $a_t = \lambda m_t + p_t$ and denote by $g(t)$ the law of Y_t . We also recall that for $x \in [0, \infty)$ and $0 \leq s < t$, $\varphi_{s,t}(x) = e^{-\lambda(t-s)}x + \int_s^t e^{-\lambda(t-u)}a_u du$. We also introduce

$$\kappa_{s,t}(x) = \exp\left(-\int_s^t f(\varphi_{s,u}(x)) du\right). \tag{16}$$

Notice that φ satisfies the flow property: one can directly check that for all $0 \leq r \leq s \leq t$, all $x \in [0, \infty)$, $\varphi_{r,t}(y) = \varphi_{s,t}(\varphi_{r,s}(y))$.

7.1. Time-marginals

Let us first proceed to a few technical considerations.

Lemma 23. *Under the above conditions,*

- (i) $t \mapsto m_t, t \mapsto p_t$ and $t \mapsto a_t$ are locally Lipschitz continuous on $[0, \infty)$,
- (ii) for all $t > 0, \lim_{h \downarrow 0} h^{-1} \mathbb{E}[1 - \kappa_{t-h,t}(Y_{t-h})] = p_t$.

Proof. Using Remark 13(ii) and (3), we observe that $t \mapsto \mathbb{E}[f^2(Y_t)]$ is locally bounded. We now prove (i). By the Itô formula, we have $m_t = m_0 + \int_0^t \mathbb{E}[(1 - Y_s) f(Y_s)] ds$ and $p_t = p_0 + \int_0^t \mathbb{E}[f'(Y_s)(p_s + \lambda(m_s - Y_s)) - f^2(Y_s)] ds$. But under Assumptions 1 and 3, there is $C > 0$ such that $x + f'(x) \leq C(1 + f(x))$. We easily conclude that $s \mapsto \mathbb{E}[(1 - Y_s) f(Y_s)]$ and $s \mapsto \mathbb{E}[f'(Y_s)(p_s + \lambda(m_s - Y_s)) - f^2(Y_s)]$ are locally bounded. The conclusion follows.

We next fix $t > 0$ and prove (ii). We write $|p_t - h^{-1} \mathbb{E}[1 - \kappa_{t-h,t}(Y_{t-h})]| \leq \Delta_h^1 + \Delta_h^2 + \Delta_h^3$, where

$$\begin{aligned} \Delta_h^1 &:= |p_t - p_{t-h}|, \\ \Delta_h^2 &:= h^{-1} \left| \mathbb{E} \left[hf(Y_{t-h}) - \int_{t-h}^t f(\varphi_{t-h,u}(Y_{t-h})) du \right] \right|, \\ \Delta_h^3 &:= h^{-1} \left| \mathbb{E} \left[\int_{t-h}^t f(\varphi_{t-h,u}(Y_{t-h})) du - (1 - \kappa_{t-h,t}(Y_{t-h})) \right] \right|. \end{aligned}$$

First, we have $\lim_{h \downarrow 0} \Delta_h^1 = 0$ by point (i). Next, we see that for $h \in (0, t \wedge 1]$, for $u \in [t - h, t]$ and for $x \geq 0$, it holds that $\varphi_{t-h,u}(x) \leq x + C_t h$ and $|x - \varphi_{t-h,u}(x)| \leq C_t(1 + x)h$, for some constant C_t . Hence $|f(x) - f(\varphi_{t-h,u}(x))| \leq (\sup_{[0, x+C_t h]} f') \times C_t(1 + x)h$. Using Assumption 3 and Remark 13(ii), we get that $|f(x) - f(\varphi_{t-h,u}(x))| \leq C_t(1 + f(x))(1 + x)h$. All this implies that $\Delta_h^2 \leq C_t h \mathbb{E}[(1 + Y_{t-h})(1 + f(Y_{t-h}))] \leq C_t h$ (because, as already seen, $s \mapsto \mathbb{E}[Y_s f(Y_s)]$ is locally bounded), which tends to 0 as $h \downarrow 0$. Finally, since $|y - (1 - \exp(-y))| \leq y^2$ for all $y \geq 0$,

$$\Delta_h^3 \leq h^{-1} \mathbb{E} \left[\left(\int_{t-h}^t f(\varphi_{t-h,u}(Y_{t-h})) du \right)^2 \right] \leq \mathbb{E} \left[\int_{t-h}^t f^2(\varphi_{t-h,u}(Y_{t-h})) du \right].$$

As previously, we use Remark 13(ii) to get $f^2(\varphi_{t-h,u}(x)) \leq f^2(x + C_t h) \leq C_t(1 + f^2(x))$ (if $h \in (0, t \wedge 1]$), whence $\Delta_h^3 \leq C_t h \mathbb{E}[1 + f^2(Y_{t-h})] \leq C_t h$ (since $s \mapsto \mathbb{E}[f^2(Y_s)]$ is locally bounded), which tends to 0 as $h \downarrow 0$. This completes the proof. □

We now introduce, for $t \geq 0, \tau_t := \sup\{s \in [0, t] : \Delta Y_s \neq 0\}$, the last jump instant before t . We adopt the convention that $\sup \emptyset = 0$: if there is no jump during $[0, t]$, we set $\tau_t = 0$.

Lemma 24. *Under the above conditions,*

- (i) a.s., for all $t \geq 0, Y_t = \varphi_{0,t}(Y_0) \mathbf{1}_{\{\tau_t=0\}} + \varphi_{\tau_t,t}(0) \mathbf{1}_{\{\tau_t>0\}}$,
- (ii) for all $t > 0, \mathbb{P}(\tau_t = 0 | Y_0) = \kappa_{0,t}(Y_0)$,
- (iii) for all $t > 0, \mathbb{P}(Y_t = 0) < 1$.

Proof. From (2), we have $Y_r = Y_0 \mathbf{1}_{\{\tau_t=0\}} + \int_{\tau_t}^r (a_s - \lambda Y_s) ds$ for all $t \geq 0$ and all $r \in [\tau_t, t]$. Solving this ODE, we find $Y_t = e^{-\lambda t} Y_0 \mathbf{1}_{\{\tau_t=0\}} + \int_{\tau_t}^t e^{-\lambda(t-s)} a_s ds$, which proves point (i). But $\tau_t = 0$ implies that $\tau_s = 0$ for all $s \in [0, t]$, whence $Y_s = \varphi_{0,s}(Y_0)$ on $[0, t]$. As a consequence, $\{\tau_t = 0\} = \{\int_0^t \int_0^\infty \mathbf{1}_{\{z \leq f(\varphi_{0,s}(Y_0))\}} \mathbf{N}(ds, dz) = 0\}$, so that $\mathbb{P}(\tau_t = 0 | Y_0) = \exp(-\int_0^t f(\varphi_{0,s}(Y_0)) ds)$, as claimed in point (ii). Using that $Y_t \geq Y_0 e^{-\lambda t}$ on $\{\tau_t = 0\}$ and point (ii), we see that

$$\mathbb{P}(Y_t > 0) \geq \mathbb{P}(Y_0 > 0, \tau_t = 0) = \mathbb{E}[\kappa_{0,t}(Y_0) \mathbf{1}_{\{Y_0>0\}}] > 0$$

since $Y_0 > 0$ occurs with positive probability. This proves (iii). □

The law of τ_t is absolutely continuous on $(0, t]$, as shown in the next proposition. This smoothness property will allow us to show that jumps indeed create a density for Y_t .

Proposition 25. *Under the above conditions, for all $t > 0$, the law of τ_t is given by $h_t(ds) = \mathbb{E}[\kappa_{0,t}(Y_0)]\delta_0(ds) + p_s\kappa_{s,t}(0)\mathbf{1}_{\{0 < s < t\}} ds$.*

Proof. First, $\mathbb{P}(\tau_t = 0) = \mathbb{E}[\kappa_{0,t}(Y_0)]$ as desired by Lemma 24(ii). We next introduce the filtration $\mathcal{F}_s = \sigma(\{Y_0, \mathbf{N}([0, r] \times A) : r \in [0, s], A \in \mathcal{B}([0, \infty))\})$ and the process $J_s = \sum_{r \in [0, s]} \mathbf{1}_{\{\Delta Y_r \neq 0\}} = \int_0^s \int_0^\infty \mathbf{1}_{\{z \leq f(Y_{r-})\}} \mathbf{N}(dr, dz)$ which counts the number of jumps of Y . We consider $0 < s - h < s < t$ and observe that $\{\tau_t \in (s - h, s]\} = \{J_s > J_{s-h}\} \cap \{J_t = J_s\}$. The event $\{J_s > J_{s-h}\}$ is \mathcal{F}_s -measurable. When Y does not jump during $(s, t]$, $Y_{r-} = Y_r = \varphi_{s,r}(Y_s)$ for all $r \in (s, t]$: this follows from Lemma 24(i), from the semi-group property of the flow φ , and from the fact that $\tau_r = \tau_s = \tau_t$ when Y does not jump during $(s, t]$. Consequently, $\{J_t = J_s\} = \{\int_s^t \int_0^\infty \mathbf{1}_{\{z \leq f(\varphi_{s,r}(Y_s))\}} \mathbf{N}(dr, dz) = 0\}$, whence

$$\mathbb{P}(J_t = J_s \mid \mathcal{F}_s) = \mathbb{P}\left(\int_s^t \int_0^\infty \mathbf{1}_{\{z \leq f(\varphi_{s,r}(Y_s))\}} \mathbf{N}(dr, dz) = 0 \mid \mathcal{F}_s\right) = \kappa_{s,t}(Y_s).$$

We conclude that $\mathbb{P}(\tau_t \in (s - h, s]) = \mathbb{E}[\kappa_{s,t}(Y_s)\mathbf{1}_{\{J_s > J_{s-h}\}}]$. On the event $\{J_s > J_{s-h}\}$, the process Y jumps (at least once) to 0 during $(s - h, s]$, so that $Y_s \in [0, \varphi_{s-h,s}(0)]$ by Lemma 24(i). Hence,

$$\left|\mathbb{P}(\tau_t \in (s - h, s]) - \mathbb{E}[\kappa_{s,t}(0)\mathbf{1}_{\{J_s > J_{s-h}\}}]\right| \leq \sup_{x \in [0, \varphi_{s-h,s}(0)]} |\kappa_{s,t}(x) - \kappa_{s,t}(0)| \times \mathbb{E}[J_s - J_{s-h}].$$

Using that $\mathbb{E}[J_s - J_{s-h}] = \int_{s-h}^s p_r dr \leq Ch$ (by Lemma 23(i)), that $\varphi_{s-h,s}(0) \rightarrow 0$ as $h \rightarrow 0$, and the (obvious) continuity of $x \mapsto \kappa_{s,t}(x)$, we conclude that

$$\limsup_{h \rightarrow 0} \frac{1}{h} \left|\mathbb{P}(\tau_t \in (s - h, s]) - \kappa_{s,t}(0)\mathbb{P}(J_s > J_{s-h})\right| = 0. \quad (17)$$

Next, arguing exactly as in Lemma 24(ii), we get $\mathbb{P}(J_s > J_{s-h}) = 1 - \mathbb{E}[\kappa_{s-h,s}(Y_{s-h})]$. Hence, we deduce from Lemma 23(ii) that

$$\lim_{h \rightarrow 0} \frac{1}{h} \mathbb{P}(J_s > J_{s-h}) = \mathbb{E}[f(Y_s)] = p_s. \quad (18)$$

Gathering (17) and (18), we deduce that indeed, the density of the law of τ_t at point $s \in (0, t)$ exists and equals $p_s\kappa_{s,t}(0)$. \square

We are now able to give the

Proof of Theorem 12. We have already seen that $t \mapsto a_t$ and $t \mapsto p_t$ are continuous (by Lemma 23(i)) and positive (by Lemma 24(iii)). We now fix $t > 0$. By Lemma 24(i), $Y_t = \varphi_{0,t}(Y_0)\mathbf{1}_{\{\tau_t=0\}} + \varphi_{\tau_t,t}(0)\mathbf{1}_{\{\tau_t>0\}}$. Hence for any bounded measurable $\phi : [0, \infty) \mapsto \mathbb{R}$,

$$\mathbb{E}[\phi(Y_t)] = \mathbb{E}[\phi(\varphi_{\tau_t,t}(0)\mathbf{1}_{\{\tau_t>0\}})] + \mathbb{E}[\phi(\varphi_{0,t}(Y_0)\mathbf{1}_{\{\tau_t=0\}})] =: A(\phi) + B(\phi).$$

Clearly, $\varphi_{\tau_t,t}(0) < \varphi_{0,t}(0)$ when $\tau_t > 0$ and $\varphi_{0,t}(Y_0) \geq \varphi_{0,t}(0)$. Using Proposition 25, we can write

$$A(\phi) = \int_0^t \phi(\varphi_{s,t}(0)) p_s \kappa_{s,t}(0) ds.$$

Recall that for $y \in [0, \varphi_{0,t}(0)]$, $\beta_t(y) \in [0, t]$ is uniquely defined by $\varphi_{\beta_t(y),t}(0) = y$. The change of variables $s \mapsto y = \varphi_{s,t}(0)$, for which $dy = -e^{-\lambda(t-s)} a_s ds$ and $s = \beta_t(y)$, gives us

$$A(\phi) = \int_0^{\varphi_{0,t}(0)} \phi(y) \frac{p_{\beta_t(y)}}{a_{\beta_t(y)}} \kappa_{\beta_t(y),t}(0) e^{\lambda(t-\beta_t(y))} dy.$$

Next recall that $\gamma_t(y) = (y - \varphi_{0,t}(0))e^{\lambda t}$ for $y \geq \varphi_{0,t}(0)$. Using Lemma 24(ii) and the change of variables $x \mapsto y = \varphi_{0,t}(x)$, for which $\gamma_t(y) = x$, we find

$$B(\phi) = \mathbb{E}[\phi(\varphi_{0,t}(Y_0))\kappa_{0,t}(Y_0)] = \int_0^\infty \phi(\varphi_{0,t}(x))\kappa_{0,t}(x)g_0(dx) = \int_{\varphi_{0,t}(0)}^\infty \phi(y)\kappa_{0,t}(\gamma_t(y))(g_0 \circ \gamma_t^{-1})(dy).$$

We have proved that

$$\mathbb{E}[\phi(Y_t)] = \int_0^\infty \phi(y) \left[\frac{p_{\beta_t}(y)}{a_{\beta_t}(y)} \kappa_{\beta_t(y),t}(0) e^{\lambda(t-\beta_t(y))} \mathbf{1}_{\{y < \varphi_{0,t}(0)\}} dy + \kappa_{0,t}(\gamma_t(y)) \mathbf{1}_{\{y \geq \varphi_{0,t}(0)\}} (g_0 \circ \gamma_t^{-1})(dy) \right].$$

Replacing κ by its expression, one finds the formula claimed in the statement. □

7.2. Nonextinction

We first consider the easy case where $\lambda = 0$.

Proof of Proposition 11 when $\lambda = 0$. Using Theorem 12, we see that the law of Y_t has a density bounded by 1, on $[0, \varphi_{0,t}(0))$ (because for all $x \in [0, \varphi_{0,t}(0))$, there is $s \in (0, t]$ such that $\varphi_{s,t}(0) = x$). For all $t > 1$, $\varphi_{0,t}(0) = \int_0^t a_s ds > \int_0^1 a_s ds =: \alpha > 0$ by Lemma 24(iii). Hence, for $\varepsilon := \min\{\alpha, 1/2\}$, it holds that $\Pr(Y_t \leq \varepsilon) \leq \varepsilon \leq 1/2$ for all $t > 1$. Consequently, Y_t cannot tend to 0 in probability as $t \rightarrow \infty$. □

To study the case where $\lambda > 0$, we need the following lemma.

Lemma 26. *Grant Assumptions 1, 3, 10(i) and suppose that $\lambda \geq 0$. Assume that $\mathbb{E}[f(Y_0)] < \infty$ and consider the unique solution $(Y_t)_{t \geq 0}$ to (2) built in Theorem 4(iii). There is $q > 1$ such that $\sup_{t \geq t_0} \mathbb{E}[f^q(Y_t)] < \infty$ for all $t_0 > 0$.*

Proof. By Assumption 10(i), there exists $q > 1$ such that $q \limsup_{x \rightarrow \infty} [f'(x)/f(x)] < 1$. We now divide the proof into two steps.

Step 1. Here we assume that Y_0 has a bounded support, so that Y_t is uniformly bounded on each compact time interval by (3). This ensures that all the computations below are licit. Applying the Itô formula for jump processes and taking expectations, we easily check that

$$\frac{d}{dt} \mathbb{E}[f^q(Y_t)] = -\lambda \mathbb{E}[(f^q)'(Y_t)(Y_t - \mathbb{E}[Y_t])] - \mathbb{E}[f^{q+1}(Y_t)] + \mathbb{E}[(f^q)'(Y_t)] \mathbb{E}[f(Y_t)].$$

Now, as shown e.g. in [24], it holds that $\mathbb{E}[\varphi_1(X)]\mathbb{E}[\varphi_2(X)] \leq \mathbb{E}[\varphi_1(X)\varphi_2(X)]$ for any $[0, \infty)$ -valued random variable X and any pair of nondecreasing functions $\varphi_1, \varphi_2 : [0, \infty) \mapsto \mathbb{R}$. Using that f^q is convex (since $q > 1$ and by Assumption 3), we deduce that $\mathbb{E}[(f^q)'(Y_t)(Y_t - \mathbb{E}[Y_t])] \geq 0$ and that $\mathbb{E}[(f^q)'(Y_t)] \mathbb{E}[f(Y_t)] \leq \mathbb{E}[(f^q)'(Y_t)f(Y_t)]$. We thus have

$$\frac{d}{dt} \mathbb{E}[f^q(Y_t)] \leq -\mathbb{E}[f^{q+1}(Y_t) - (f^q)'(Y_t)f(Y_t)].$$

Recalling now that $q \limsup_{x \rightarrow \infty} [f'(x)/f(x)] < 1$, we can find two constants $c > 0$ and $C \geq 0$ such that $f^{q+1}(x) - (f^q)'(x)f(x) \geq cf^{q+1}(x) - C$ for all $x \geq 0$. Indeed, find $x_0 > 0$ such that $a := \sup_{x \geq x_0} [f'(x)/f(x)] < 1/q$, observe that $f^{q+1}(x) - (f^q)'(x)f(x) = f^{q+1}(x) - qf'(x)f^q(x) \geq (1 - aq)f^{q+1}(x)$ for $x \geq x_0$, and conclude by setting $c = (1 - aq) > 0$ and $C = \sup_{[0, x_0]} (f^q)'(x)f(x)$. We deduce that

$$\frac{d}{dt} \mathbb{E}[f^q(Y_t)] \leq C - c\mathbb{E}[f^{q+1}(Y_t)] \leq C - c\mathbb{E}[f^q(Y_t)]^{1+1/q}.$$

The conclusion classically follows: there is a constant K , not depending on $\mathbb{E}[f^q(Y_0)]$ such that $\mathbb{E}[f^q(Y_t)] \leq K(1 + t^{-q})$.

Step 2. We next only assume that $\mathbb{E}[f(Y_0)] < \infty$. We introduce $Y_0^A = \min\{Y_0, A\}$ and the unique solution $(Y_t^A)_{t \geq 0}$ to (2) starting from Y_0^A . By Step 1, we know that for all $t \geq 0$, uniformly in A , $\mathbb{E}[f^q(Y_t^A)] \leq K(1 + t^{-q})$. But we also know by Proposition 17 that for each $t \geq 0$, Y_t^A goes in law to Y_t as $A \rightarrow \infty$ (we have to verify that $\mathbb{E}[|H(Y_0^A) - H(Y_0)|] \rightarrow 0$, which is not difficult by dominated convergence since $\mathbb{E}[H(Y_0)] \leq \pi/2 + \mathbb{E}[f(Y_0)] < \infty$ by assumption). The conclusion follows. \square

Proof of Proposition 11 when $\lambda > 0$. We work by contradiction and assume that Y_t goes in law (and thus in probability) to 0 as $t \rightarrow \infty$. By Lemma 26, we know that there is $q > 1$ such that $\sup_{t \geq 1} \mathbb{E}[f^q(Y_t)] < \infty$. This implies that $\sup_{t \geq 1} \mathbb{E}[Y_t^q] < \infty$ by Remark 13(i). Consequently, Y_t and $f(Y_t)$ are uniformly integrable (for $t \geq 1$), so that the Lebesgue theorem tells us, since Y_t goes in probability to 0, that $a_t = \lambda \mathbb{E}[Y_t] + E[f(Y_t)]$ tends to 0.

We use Lemma 24(i) to write $Y_t = e^{-\lambda t} Y_0 \mathbf{1}_{\{\tau_t=0\}} + \varphi_{\tau_t,t}(0)$. First, there is $t_0 > 0$ such that

$$\text{for all } t \geq t_0, \quad \varphi_{\tau_t,t}(0) \leq \varphi_{0,t}(0) \leq 1/2. \quad (19)$$

Indeed, we consider $t_1 > 0$ such that $a_t \leq \lambda/4$ for all $t \geq t_1$. Then we see that $\varphi_{\tau_t,t}(0) \leq \varphi_{0,t}(0) \leq e^{-\lambda t} \int_0^{t_1} e^{\lambda u} a_u du + (\lambda/4) \int_{t_1}^t e^{-\lambda(t-u)} du \leq C e^{-\lambda t} + 1/4$, whence the conclusion.

Taking expectations in (2), we see that $(d/dt)\mathbb{E}[Y_t] = \mathbb{E}[(1 - Y_t)f(Y_t)]$. But for $t \geq t_0$ and on the event $\{\tau_t > 0\}$, we have $Y_t = \varphi_{\tau_t,t}(0) \leq 1/2$ and thus $(1 - Y_t)f(Y_t) \geq 0$. Next on the event $\{\tau_t = 0\}$, we have $Y_t = \varphi_{0,t}(Y_0)$. Consequently,

$$\frac{d}{dt} \mathbb{E}[Y_t] \geq \mathbb{E}[(1 - \varphi_{0,t}(Y_0))f(\varphi_{0,t}(Y_0))\mathbf{1}_{\{\tau_t=0\}}].$$

Hence, by Lemma 24(ii),

$$\frac{d}{dt} \mathbb{E}[Y_t] \geq \mathbb{E}[(1 - \varphi_{0,t}(Y_0))f(\varphi_{0,t}(Y_0))\kappa_{0,t}(Y_0)] \geq I_t - J_t,$$

where

$$I_t := \mathbb{E}[(1 - \varphi_{0,t}(Y_0))f(\varphi_{0,t}(Y_0))\kappa_{0,t}(Y_0)\mathbf{1}_{\{\varphi_{0,t}(Y_0) < 3/4\}}],$$

$$J_t := \mathbb{E}[(\varphi_{0,t}(Y_0) - 1)f(\varphi_{0,t}(Y_0))\kappa_{0,t}(Y_0)\mathbf{1}_{\{\varphi_{0,t}(Y_0) > 1\}}].$$

We will now prove that there is $t_2 > t_0$ such that $I_t > J_t$ for all $t \geq t_2$. This will end the proof, since then $(d/dt)\mathbb{E}[Y_t] > 0$ for all $t \geq t_2$, so that $\mathbb{E}[Y_t]$, and thus a fortiori a_t , cannot go to 0. To prove that I_t is eventually greater than J_t , we will check that, with $\xi \geq 1$ defined in Assumption 10(ii),

$$(a) \quad \liminf_{t \rightarrow \infty} \frac{e^{\lambda \xi t}}{\kappa_{0,t}(0)} I_t > 0, \quad (b) \quad \limsup_{t \rightarrow \infty} \frac{e^{\lambda \xi t}}{\kappa_{0,t}(0)} J_t = 0.$$

Let us first prove (b). It holds that $\kappa_{0,t}(Y_0)/\kappa_{0,t}(0) \leq 1$. Next, recalling that $\varphi_{0,t}(Y_0) = e^{-\lambda t} Y_0 + \varphi_{0,t}(0) \leq e^{-\lambda t} Y_0 + 1/2$ for $t \geq t_0$ by (19), we see that $\mathbf{1}_{\{\varphi_{0,t}(Y_0) > 1\}} \leq \mathbf{1}_{\{e^{-\lambda t} Y_0 > 1/2\}}$. On the set $e^{-\lambda t} Y_0 > 1/2$, it holds that $\varphi_{0,t}(Y_0) \leq e^{-\lambda t} Y_0 + 1/2 \leq 2e^{-\lambda t} Y_0$. Finally, by Assumption 10(ii), we write that $e^{\lambda \xi t} (\varphi_{0,t}(Y_0) - 1)f(\varphi_{0,t}(Y_0)) \leq C e^{\lambda \xi t} [(2e^{-\lambda t} Y_0)^\xi + (2e^{-\lambda t} Y_0)^{\zeta+1}] \leq C [Y_0^\xi + Y_0^{\zeta+1}] \leq C(1 + Y_0^{\zeta+1})$ because $\zeta \geq \xi - 1$ by assumption. All in all, we have checked that for $t \geq t_0$,

$$\frac{e^{\lambda \xi t}}{\kappa_{0,t}(0)} J_t \leq C \mathbb{E}[(1 + Y_0^{\zeta+1})\mathbf{1}_{\{Y_0 > e^{\lambda t}/2\}}].$$

Since $\mathbb{E}[Y_0^{\zeta+1}] < \infty$ by assumption, the monotone convergence theorem shows the validity of (b).

We finally prove (a). We have $\varphi_{0,t}(Y_0) \geq e^{-\lambda t} Y_0$. Hence by Assumption 10(ii), we may write, on the event $\{\varphi_{0,t}(Y_0) < 3/4\}$, that $(1 - \varphi_{0,t}(Y_0))f(\varphi_{0,t}(Y_0)) \geq c e^{-\lambda \xi t} Y_0^\xi$. We next recall that, as previously, for $t \geq t_0$, $\varphi_{0,t}(Y_0) \leq$

$e^{-\lambda t} Y_0 + 1/2$. Consequently, $\mathbf{1}_{\{\varphi_{0,t}(Y_0) < 3/4\}} \geq \mathbf{1}_{\{e^{-\lambda t} Y_0 < 1/4\}}$. Finally,

$$\begin{aligned} \frac{\kappa_{0,t}(Y_0)}{\kappa_{0,t}(0)} &= \exp\left(-\int_0^t [f(\varphi_{0,s}(Y_0)) - f(\varphi_{0,s}(0))] ds\right) \\ &\geq \exp\left(-\int_0^t e^{-\lambda s} Y_0 \times \left(\sup_{[0, \varphi_{0,s}(Y_0)]} f'\right) ds\right), \end{aligned}$$

because $\varphi_{0,s}(Y_0) = e^{-\lambda s} Y_0 + \varphi_{0,s}(0)$. But $\sup_{s \geq 0} \varphi_{0,s}(Y_0) \leq Y_0 + C$: this follows from the expression of φ and the fact that $t \mapsto a_t$ is a bounded function, since it is locally bounded and tends to 0. Hence, using Assumptions 1, 3 and Remark 13(ii), $\sup_{[0, \varphi_{0,s}(Y_0)]} f' \leq C(1 + f(Y_0))$, and we end with

$$\frac{\kappa_{0,t}(Y_0)}{\kappa_{0,t}(0)} \geq \exp(-CY_0(1 + f(Y_0))).$$

All in all, we see that for $t \geq t_0$,

$$\frac{e^{\lambda \xi t}}{\kappa_{0,t}(0)} I_t \geq c \mathbb{E}[Y_0^\xi \exp(-CY_0(1 + f(Y_0))) \mathbf{1}_{\{Y_0 < e^{\lambda t}/4\}}].$$

This last quantity tends, by monotone convergence, to $c \mathbb{E}[Y_0^\xi \exp(-CY_0(1 + f(Y_0)))] > 0$. We have checked (a). \square

7.3. Trend to equilibrium when $\lambda = 0$

This final part is dedicated to the proof of Proposition 9. We thus work under all the assumptions above and suppose furthermore that $\lambda = 0$ and that the initial condition g_0 has a density $g_0 \in C_b^1([0, \infty))$ satisfying $g_0(0) = 1$ and $\int_0^\infty |g'_0(y)| dy < \infty$.

Since $\lambda = 0$, we simply have $\varphi_{s,t}(x) = x + A_t - A_s$, where $A_t = \int_0^t a_s ds$. For $t \geq 0$ and $y \in [0, A_t]$, $\beta_t(y) \in [0, t]$ is defined by $A_t - A_{\beta_t(y)} = y$. And for $t \geq 0$ and $y \geq A_t$, $\gamma_t(y) = y - A_t$. We thus know from Theorem 12 that $g(t)$ has a density on $[0, \infty)$ given by

$$g(t, y) = \kappa_{\beta_t(y), t}(0) \mathbf{1}_{\{y < A_t\}} + \kappa_{0,t}(y - A_t) g_0(y - A_t) \mathbf{1}_{\{y \geq A_t\}}. \tag{20}$$

Observe that $g(t, A_t -) = g(t, A_t +) = \kappa_{0,t}(0)$. A little study, using our assumptions on g_0 and that the map $t \mapsto a_t$ is continuous and positive, shows that $g(t, y)$ is continuous on $[0, \infty)^2$, of class C^1 on $\{(t, y) \in [0, \infty)^2 : y \neq A_t\}$ and that $\sup_{[0, T]} (\|\partial_y g(t, \cdot)\|_{L^\infty(\mathbb{R})} + \|\partial_y g(t, \cdot)\|_{L^1(\mathbb{R})}) < \infty$ for all $T > 0$.

Let $\phi \in C_b^1([0, \infty))$. Applying the Itô formula to compute $\phi(Y_t)$, taking expectations and differentiating the obtained formula, we find that

$$\begin{aligned} \frac{d}{dt} \int_0^\infty \phi(x) g(t, x) dx &= \int_0^\infty [\phi(0) - \phi(x) + p_t \phi'(x)] g(t, x) dx \\ &= - \int_0^\infty \phi(x) [f(x) g(t, x) + p_t \partial_x g(t, x)] dx. \end{aligned}$$

The second equality follows from an integration by parts, which is licit because for t fixed, $g(t, \cdot)$ is continuous, piece-wise C^1 and $\int_0^\infty |\partial_y g(t, y)| dy < \infty$. The boundary term disappears since $g(t, 0) = 1$.

We introduce now $g(x) = \exp(-p^{-1} \int_0^x f(y) dy)$ as in Theorem 8(ii), which solves $p \partial_x g(x) + f(x) g(x) = 0$. Hence, for any $\phi \in C_b^1$,

$$\begin{aligned} \frac{d}{dt} \int_0^\infty \phi(x) (g(t, x) - g(x)) dx &= - \int_0^\infty \phi(x) f(x) (g(t, x) - g(x)) dx - p_t \int_0^\infty \phi(x) (\partial_x (g(t, x) - g(x))) dx \\ &\quad + (p - p_t) \int_0^\infty \phi(x) \partial_x g(x) dx. \end{aligned}$$

We thus can apply Lemma 27 below with $a(t, x) = g(t, x) - g(x)$ and $b(t, x) = -f(x)(g(t, x) - g(x)) - p_t \partial_x (g(t, x) - g(x)) + (p - p_t) \partial_x g(x)$, which both belong to $L_{\text{loc}}^\infty([0, T], L^1([0, \infty))$ because $(1 + f(x))g(x) + |\partial_x g(x)|$ is integrable, because p_t is locally bounded, because $\int_0^\infty f(x)g(t, x) dx = p_t$ and because $\sup_{[0, T]} \|\partial_x g(t, \cdot)\|_{L^1(\mathbb{R})} < \infty$ for all $T > 0$, to deduce that $t \mapsto \int_0^\infty |g(t, x) - g(x)| dx$ is continuous and satisfies, for a.e. $t \geq 0$,

$$\begin{aligned} \frac{d}{dt} \int_0^\infty |g(t, x) - g(x)| dx &= - \int_0^\infty \text{sg}(g(t, x) - g(x)) f(x) (g(t, x) - g(x)) dx \\ &\quad - p_t \int_0^\infty \text{sg}(g(t, x) - g(x)) \partial_x (g(t, x) - g(x)) dx \\ &\quad + (p - p_t) \int_0^\infty \text{sg}(g(t, x) - g(x)) \partial_x g(x) dx. \end{aligned}$$

The second term on the RHS equals $-p_t \int_0^\infty \partial_x [g(t, x) - g(x)] dx = p_t |g(t, 0) - g(0)| = 0$. Using that g is decreasing, we easily check that the last term is bounded by

$$|p_t - p| \int_0^\infty |\partial_x g(x)| dx = |p - p_t| g(0) = |p - p_t| = \left| \int_0^\infty f(x) (g(t, x) - g(x)) dx \right|.$$

Thus, for a.e. $t \geq 0$,

$$\begin{aligned} \frac{d}{dt} \|g(t) - g\|_{L^1} &\leq - \int_0^\infty f(x) |g(t, x) - g(x)| dx + \left| \int_0^\infty f(x) (g(t, x) - g(x)) dx \right| \\ &= -2 \min \left\{ \int_0^\infty f(x) (g(t, x) - g(x))_+ dx, \int_0^\infty f(x) (g(t, x) - g(x))_- dx \right\}. \end{aligned}$$

But by (20), $g(t, x) \leq 1$ for all $t \geq 1$, all $x \in [0, A_1]$ and $A_1 = \int_0^1 a_s ds > 0$. Consequently, for $t \geq 1$, all $\varepsilon \in [0, A_1]$, since f is nondecreasing,

$$\int_0^\infty f(x) (g(t, x) - g(x))_+ dx \geq f(\varepsilon) \int_\varepsilon^\infty (g(t, x) - g(x))_+ dx \geq f(\varepsilon) \left[\int_0^\infty (g(t, x) - g(x))_+ dx - \varepsilon \right].$$

But $g(t)$ and g being two probability density functions, $\int_0^\infty (g(t, x) - g(x))_+ dx = \|g(t) - g\|_{L^1}/2$. Thus for all $t \geq 1$, all $\varepsilon \in [0, A_1]$,

$$\int_0^\infty f(x) (g(t, x) - g(x))_+ dx \geq f(\varepsilon) [\|g(t) - g\|_{L^1}/2 - \varepsilon].$$

Since $g(x) \leq 1$ for all $x \geq 0$, a similar estimate holds true for $\int_0^\infty f(x) (g(t, x) - g(x))_- dx$. All in all, we have proved that for a.e. $t \geq 1$, all $\varepsilon \in [0, A_1]$,

$$\frac{d}{dt} \|g(t) - g\|_{L^1} \leq -f(\varepsilon) [\|g(t) - g\|_{L^1} - 2\varepsilon].$$

Choosing $\varepsilon = \min\{A_1, \|g(t) - g\|_{L^1}/4\}$ and introducing the function $\Phi(x) = xf(A_1 \wedge (x/4))/2$, this implies, still for a.e. $t \geq 1$, that

$$\frac{d}{dt} \|g(t) - g\|_{L^1} \leq -\Phi(\|g(t) - g\|_{L^1}).$$

Since Φ is nonnegative increasing on $(0, \infty)$ and vanishes only at 0, we easily conclude, using that $t \mapsto \|g(t) - g\|_{L^1}$ is continuous, that indeed, $\lim_{t \rightarrow \infty} \|g(t) - g\|_{L^1} = 0$. Recalling that $2\|g(t) - g\|_{\text{TV}} = \|g(t) - g\|_{L^1}$ implies the result.

If now $f(x) \geq cx^\xi$ on $[0, 1]$ for some $\xi \geq 1$, there clearly exists another constant $c > 0$ such that $\Phi(x) \geq cx^{\xi+1}$ for all $x \in [0, 2]$ (recall that f is nondecreasing). But $\|g(t) - g\|_{L^1}$ always belongs to $[0, 2]$. We thus have, for a.e.

$t \geq 1$, $(d/dt)\|g(t) - g\|_{L^1} \leq -c\|g(t) - g\|_{L^1}^{\xi+1}$. It is then not hard to deduce that there is a constant C such that $\|g(t) - g\|_{L^1} \leq C(1+t)^{-1/\xi}$ for all $t \geq 0$.

It remains to check the following lemma, which is well-known folklore since the seminal work of DiPerna and Lions [7]. We unfortunately found no precise reference.

Lemma 27. *Let $a, b : [0, \infty) \times \mathbb{R} \mapsto \mathbb{R}$ belong to $L^\infty_{\text{loc}}([0, \infty), L^1(\mathbb{R}))$, that is $\sup_{[0, T]} \int_{\mathbb{R}} (|a(t, x)| + |b(t, x)|) dx < \infty$ for all $T > 0$. Assume that for all $\phi \in C^1_b(\mathbb{R})$, all $t \geq 0$,*

$$\frac{d}{dt} \int_{\mathbb{R}} \phi(y)a(t, y) dy = \int_{\mathbb{R}} \phi(y)b(t, y) dy. \tag{21}$$

Then $t \mapsto \int_{\mathbb{R}} |a(t, x)| dx$ is continuous and for a.e. $t \geq 0$,

$$\frac{d}{dt} \int_{\mathbb{R}} |a(t, x)| dx = \int_{\mathbb{R}} \text{sg}(a(t, x))b(t, x) dx,$$

where $\text{sg}(u) = \mathbf{1}_{\{u>0\}} - \mathbf{1}_{\{u<0\}}$.

Proof. We introduce $\rho_\varepsilon = (2\pi\varepsilon)^{-1/2} \exp(-x^2/(2\varepsilon))$ for $\varepsilon > 0$ and $x \in \mathbb{R}$ and define $a_\varepsilon(t, x) = [a(t, \cdot) \star \rho_\varepsilon](x) = \int_{\mathbb{R}} a(t, y)\rho_\varepsilon(x - y) dy$ and $b_\varepsilon(t, x) = [b(t, \cdot) \star \rho_\varepsilon](x)$. It is well known that for all $t \geq 0$, $\lim_{\varepsilon \rightarrow 0} (\|a(t, \cdot) - a_\varepsilon(t, \cdot)\|_{L^1} + \|b(t, \cdot) - b_\varepsilon(t, \cdot)\|_{L^1}) = 0$. It is also clear that for all $t \geq 0$, all $\varepsilon > 0$, $\|a_\varepsilon(t, \cdot)\|_{L^1} \leq \|a(t, \cdot)\|_{L^1}$ and $\|b_\varepsilon(t, \cdot)\|_{L^1} \leq \|b(t, \cdot)\|_{L^1}$.

Step 1. Applying (21) with $\phi(y) = \rho_\varepsilon(x - y)$, we find that for all $t \geq 0$, all $x \in \mathbb{R}$, $\partial_t a_\varepsilon(t, x) = b_\varepsilon(t, x)$. Hence for any $\psi \in C^1_b(\mathbb{R})$, $\partial_t \psi(a_\varepsilon(t, x)) = \psi'(a_\varepsilon(t, x))b_\varepsilon(t, x)$. We conclude that for all $t \geq 0$,

$$\int_{\mathbb{R}} \psi(a_\varepsilon(t, x)) dx = \int_{\mathbb{R}} \psi(a_\varepsilon(0, x)) dx + \int_0^t \int_{\mathbb{R}} \psi'(a_\varepsilon(s, x))b_\varepsilon(s, x) dx ds. \tag{22}$$

Step 2. We now pass to the limit as $\varepsilon \rightarrow 0$ in (22) to deduce that for all $t \geq 0$, all $\psi \in C^2_b(\mathbb{R})$,

$$\int_{\mathbb{R}} \psi(a(t, x)) dx = \int_{\mathbb{R}} \psi(a(0, x)) dx + \int_0^t \int_{\mathbb{R}} \psi'(a(s, x))b(s, x) dx ds. \tag{23}$$

First, we clearly have that $\lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}} \psi(a_\varepsilon(t, x)) dx = \int_{\mathbb{R}} \psi(a(t, x)) dx$ because ψ is globally Lipschitz continuous and because $\lim_{\varepsilon \rightarrow 0} \|a(t, \cdot) - a_\varepsilon(t, \cdot)\|_{L^1} = 0$. The first term on the RHS is of course treated similarly. We next introduce $\Delta_\varepsilon = |\int_0^t \int_{\mathbb{R}} \psi'(a_\varepsilon(s, x))b_\varepsilon(s, x) dx ds - \int_0^t \int_{\mathbb{R}} \psi'(a(s, x))b(s, x) dx ds|$ and write

$$\Delta_\varepsilon \leq \|\psi'\|_\infty \int_0^t \int_{\mathbb{R}} |b_\varepsilon(s, x) - b(s, x)| dx ds + \int_0^t \int_{\mathbb{R}} |b(s, x)| |\psi'(a_\varepsilon(s, x)) - \psi'(a(s, x))| dx ds = I_\varepsilon + J_\varepsilon.$$

First, $\lim_{\varepsilon \rightarrow 0} I_\varepsilon = 0$ by dominated convergence, because $\int_{\mathbb{R}} |b_\varepsilon(s, x) - b(s, x)| dx$ is bounded on $[0, t]$ and tends to 0 for each $s \geq 0$. Next, we observe that for all (large) $K > 0$,

$$J_\varepsilon \leq 2\|\psi'\|_\infty \int_0^t \int_{\mathbb{R}} |b(s, x)| \mathbf{1}_{\{|b(s, x)|>K\}} dx ds + \|\psi''\|_\infty K \int_0^t \int_{\mathbb{R}} |a_\varepsilon(s, x) - a(s, x)| dx ds.$$

The second term tends to 0 as $\varepsilon \rightarrow 0$, for the same reasons as for I_ε . We thus conclude that $\limsup_{\varepsilon \rightarrow 0} \Delta_\varepsilon \leq 2\|\psi'\|_\infty \int_0^t \int_{\mathbb{R}} |b(s, x)| \mathbf{1}_{\{|b(s, x)|>K\}} dx ds$ for all $K > 0$. But this last quantity tends to 0 as $K \rightarrow \infty$, so that finally, $\lim_{\varepsilon \rightarrow 0} \Delta_\varepsilon = 0$ and (23) is verified.

Step 3. Here we verify, and this will conclude the proof, that

$$\int_{\mathbb{R}} |a(t, x)| dx = \int_{\mathbb{R}} |a(0, x)| dx + \int_0^t \int_{\mathbb{R}} \text{sg}(a(s, x))b(s, x) dx ds. \tag{24}$$

We consider a sequence of even smooth nonnegative functions $\psi_n \in C_b^2(\mathbb{R})$, such that $\psi_n(u)$ increases to $|u|$, for each $u \in \mathbb{R}$, as $n \rightarrow \infty$, such that $\psi_n'(u)$ tends to $\text{sg}(u)$ for each $u \in \mathbb{R}$ and such that $\sup_n \|\psi_n'\|_\infty \leq 2$. The choice $\psi_n(u) = \sqrt{x^2 + 1/n} - \sqrt{1/n}$ is possible. By Step 2, we find, for all $t \geq 0$, all $n \geq 1$,

$$\int_{\mathbb{R}} \psi_n(a(t, x)) dx = \int_{\mathbb{R}} \psi_n(a(0, x)) dx + \int_0^t \int_{\mathbb{R}} \psi_n'(a(s, x)) b(s, x) dx ds.$$

By monotone convergence, we have $\lim_n \int_{\mathbb{R}} \psi_n(a(t, x)) dx = \int_{\mathbb{R}} |a(t, x)| dx$ and $\lim_n \int_{\mathbb{R}} \psi_n(a(0, x)) dx = \int_{\mathbb{R}} |a(0, x)| dx$. It also holds true that $\lim_n \int_0^t \int_{\mathbb{R}} \psi_n'(a(s, x)) b(s, x) dx ds = \int_0^t \int_{\mathbb{R}} \text{sg}(a(s, x)) b(s, x) dx ds$, by dominated convergence. Indeed, we know that $\psi_n'(a(s, x)) \rightarrow \text{sg}(a(s, x))$ for each s, x , and $|\psi_n'(a(s, x)) b(s, x)| \leq 2|b(s, x)|$, which is integrable on $[0, t] \times \mathbb{R}$ by assumption. \square

Acknowledgements

This paper is a continuation of [6], and EL thanks Marzio Cassandro, Anna De Masi and Errico Presutti for their collaborative participation and the many discussions we had in L'Aquila and Rome, especially concerning the invariant distribution of the limit process. EL also thanks Carl Graham for many discussions at the beginning of the work. This article was produced as part of the project Labex MME-DII (ANR11-LBX-0023-01) and as part of the activities of FAPESP Research, Dissemination and Innovation Center for Neuromathematics (Grant 2013/07699-0, S. Paulo Research Foundation). The authors warmly thank the referees for their suggestions that allowed them to improve subsequently the presentation of the paper.

References

- [1] F. Bolley, J. A. Cañizo and J. A. Carrillo. Stochastic mean-field limit: Non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.* **21** (11) (2011) 2179–2210. [MR2860672](#)
- [2] F. Bolley, A. Guillin and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov–Fokker–Planck equation. *ESAIM Math. Model. Numer. Anal.* **44** (2010) 867–884. [MR2731396](#)
- [3] M. J. Cáceres, J. A. Carrillo and B. Perthame. Analysis of nonlinear noisy integrate & fire neuron models: Blow-up and steady states. *J. Math. Neurosci.* **1** (2011) Article 7. [MR2853216](#)
- [4] F. Delarue, J. Inglis, S. Rubenthaler and E. Tanré. Global solvability of a networked integrate-and-fire model of McKean–Vlasov type. *Ann. Appl. Probab.* **25** (2015) 2096–2133. [MR3349003](#)
- [5] F. Delarue, J. Inglis, S. Rubenthaler and E. Tanré. Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Process. Appl.* **125** (2015) 2451–2492. [MR3322871](#)
- [6] A. De Masi, A. Galves, E. Löcherbach and E. Presutti. Hydrodynamic limit for interacting neurons. *J. Stat. Phys.* **158** (2015) 866–902. [MR3311484](#)
- [7] R. J. DiPerna and P. L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** (1989) 511–548. [MR1022305](#)
- [8] O. Faugeras, J. Touboul and B. Cessac. A constructive mean-field analysis of multi-population neural networks with random synaptic weights and stochastic inputs. *Front. Comput. Neurosci.* **3** (2009) 1–28.
- [9] N. Fournier. On some stochastic coalescents. *Probab. Theory Related Fields* **136** (2006) 509–523. [MR2257133](#)
- [10] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** (2015) 707–738. [MR3383341](#)
- [11] N. Fournier and E. Löcherbach. Stochastic coalescence with homogeneous-like interaction rates. *Stochastic Process. Appl.* **119** (2009) 45–73. [MR2485019](#)
- [12] N. Fournier and S. Mischler. Rate of convergence of the Nanbu particle system for hard potentials. *Ann. Probab.* To appear, 2016.
- [13] A. Galves and E. Löcherbach. Infinite systems of interacting chains with memory of variable length – a stochastic model for biological neural nets. *J. Stat. Phys.* **151** (5) (2013) 896–921. [MR3055382](#)
- [14] N. Hansen, P. Reynaud-Bouret and V. Rivoirard. Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21** (2015) 83–143. [MR3322314](#)
- [15] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam, 1989. [MR1011252](#)
- [16] J. Inglis and D. Talay. Mean-field limit of a stochastic particle systems smoothly interacting through threshold hitting-times and applications to neural networks with dendritic component. *SIAM J. Math. Anal.* **47** (5) (2015) 3884–3916. [MR3411723](#)
- [17] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, 2nd edition. Springer, Berlin, 2003. [MR1943877](#)
- [18] P. Jahn, R. Berg, J. Hounsgaard and S. Ditlevsen. Motoneuron membrane potentials follow a time inhomogeneous jump diffusion process. *J. Comput. Neurosci.* **31** (3) (2011) 563–579. [MR2864731](#)

- [19] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. III (Berkeley and Los Angeles, 1956)* 171–197. University of California Press, Berkeley. [MR0084985](#)
- [20] E. Luçon and W. Stannat. Mean field limit for disordered diffusions with singular interactions. *Ann. Appl. Probab.* **24** (5) (2014) 1946–1993. [MR3226169](#)
- [21] H. P. McKean. Speed of approach to equilibrium for Kac’s caricature of a Maxwellian gas. *Arch. Ration. Mech. Anal.* **21** (1966) 347–367. [MR0214112](#)
- [22] H. P. McKean. Propagation of chaos for a class of nonlinear parabolic equations. In *Stochastic Differential Equations* 41–57. Washington, *Lecture Series in Differential Equations (Session 7, Catholic University)*. Air Force Office Sci. Res., Arlington, VA, 1967. [MR0233437](#)
- [23] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE’s. *Stochastic Process. Appl.* **95** (2001) 109–132. [MR1847094](#)
- [24] K. D. Schmidt. On inequalities for moments and the covariance of monotone functions. *Insurance Math. Econom.* **55** (2014) 91–95. [MR3179802](#)
- [25] A.-S. Sznitman. Equations de type de Boltzmann, spatialement homogènes. *Z. Wahrsch. Verw. Gebiete* **66** (1984) 559–592. [MR0753814](#)
- [26] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX – 1989* 165–251. *Lecture Notes in Math.* **1464**. Springer, Berlin, 1991. [MR1108185](#)
- [27] H. Tanaka. On the uniqueness of Markov process associated with the Boltzmann equation of Maxwellian molecules. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)* 409–425. Wiley, New York, 1978. [MR0536022](#)
- [28] H. Tanaka. Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete* **46** (1978/79) 67–105. [MR0512334](#)