The Annals of Applied Probability
2021, Vol. 31, No. 5, 2350–2375
https://doi.org/10.1214/20-AAP1649

BRANCHING DIFFUSION REPRESENTATION FOR NONLINEAR CAUCHY PROBLEMS AND MONTE CARLO APPROXIMATION

BY PIERRE HENRY-LABORDÈRE[1] AND NIZAR TOUZI[2]

[1]*Société Générale, Global Market Quantitative Research, pierre.henry-labordere@sgcib.com*
[2]*Ecole Polytechnique Paris, Centre de Mathématiques Appliquées, nizar.touzi@polytechnique.edu*

We provide probabilistic representations of the solution of some semilinear hyperbolic and high-order PDEs based on branching diffusions. These representations pave the way for an approximation of the solution by the standard Monte Carlo method, whose error estimate is controlled by the standard central limit theorem, thus partly bypassing the curse of dimensionality. We illustrate the numerical implications in the context of some popular PDEs in physics such as nonlinear Klein–Gordon equation, a simplified scalar version of the Yang–Mills equation, a fourth-order nonlinear beam equation and the Gross–Pitaevskii PDE as an example of nonlinear Schrödinger equations.

1. Introduction. Similar to the intimate connection between the heat equation and the Brownian motion, linear (second-order) parabolic partial differential equations are connected to stochastic processes. More precisely, the Feynman–Kac formula states that a linear parabolic PDE with infinitesimal generator $\mathcal{L} := b \cdot \partial_x + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^{\mathrm{T}}\partial^2_{xx^{\mathrm{T}}}]$, and terminal condition $f_1(\cdot)$ at time T, can be written as a conditional expectation of $f_1(X_T)$ involving the Itô process X associated with the generator $\mathcal{L}$. This connection allows to devise numerically approximations of the solution of such PDEs by probabilistic (Monte Carlo) methods, which represents a clear advantage in high-dimensional problems as the error estimate induced by the central limit theorem is dimension-free.

An important focus was put on the extension to nonlinear (second-order) parabolic PDEs, see, for example, [12] for a quick review of existing methods.

– A first attempt was achieved by exploiting the stochastic representation by means of backward stochastic differential equations, see Bally and Pagès [3], Bouchard and Touzi [6], Zhang [22], and the extension to the fully nonlinear setting by Fahim, Touzi and Warin [11]. The induced numerical method involves repeated computations of conditional expectations, resulting in a serious dimension dependence of the corresponding numerical methods. In fact, as highlighted in [11], this method can be viewed as part of the traditional finite-elements algorithm.

– For special types of nonlinearities, one may also use the representations available in the literature on Galton–Watson branching diffusions and their extension to superprocesses. While Galton–Watson branching diffusions are connected to the so-called KPP (Kolmogorov–Petrovskii–Piskunov) equations [20], superprocesses induce a connection with a more general class of nonlinear parabolic and elliptic PDEs, with nonlinearity depending only on the unknown function. See, for example, Le Gall [18]. These connections have potential numerical implications but still require to develop some approximation methods for superprocesses (which, up to our knowledge, are not available in the current literature). We also refer to the branching process interpretation of Le Jan and Sznitman [19] for the incompressible Navier–Stokes equations in $\mathbb{R}^3$ in Fourier space, which has not been followed by any numerical implications.

Received January 2018; revised October 2020.
MSC2020 subject classifications. 35A99, 35C15, 60J85, 65C05.
Key words and phrases. Duhamel formula, nonlinear initial value partial differential equations, branching processes.

Inspired by the extension of KPP equations by Rasulov, Raimov and Mascagni [21] and Henry–Labordère [13] suggested a numerical method based on an extension of the McKean [20] branching diffusion representation of the so-called KPP equation to a class of semilinear second order parabolic PDEs with power nonlinearity in the value function u.

The resulting algorithm is purely probabilistic. In particular its convergence is controlled by the central limit theorem, with rate of convergence independent of the dimension of the underlying state, thus proving a high potential for high-dimensional problems.[1] The validity of this method in the path-dependent case, and for further analytic nonlinearities which are of the power type in the triple $(u, \partial_x u, \partial^2_{xx^{\mathrm{T}}} u)$ is analyzed in [13–16]. We also refer to Agarwal and Claisse [1] for the extension to elliptic semilinear PDEs, and Bouchard, Tan, Warin and Zou [5] for Lipschitz nonlinearity in the pair $(v, \partial_x v)$. A critical ingredient for the extension is the use of Galton–Watson trees weighted by some Malliavin automatic differentiation weights.

Our objective in this paper is to show that the above branching diffusion approach extends to more general Cauchy problems, including hyperbolic and higher order PDEs. Such an extension may seem to be due to the intimate connection between parabolic second order PDEs and diffusions generators. However, probabilistic representations of some specific examples of hyperbolic PDEs did appear in the previous literature. The first such relevant work traces back to Kac [17] in the context of the one-dimensional telegrapher equation:

$$\partial_{tt} u - c\partial_{xx} u + (2\beta)\partial_t u = 0, \qquad u(0,x) \equiv f_1(x), \qquad \partial_t u(0,x) \equiv f_2(x),$$

where β and c are constant parameters, and the boundary data f_1, f_2 are some given functions. Observing that we may express $u(t,x) = \frac{1}{2}[u_0 + u_1](t,x)$, where the pair (u_0, u_1) solves the coupled system of first order PDEs:

$$\partial_t u_j + (-1)^j c\partial_x u_j - \beta(u_{1-j} - u_j) = 0,$$
$$u_j(0,x) = f_1(x) - \frac{(-1)^j}{c}\int_0^x f_2(y)\,dy, \quad j = 0, 1.$$

A branching mechanism representation of u_0 and u_1 is obtained by following McKean's representation for (interacting) KPP equations (with zero diffusion).

We next mention the work of Dalang, Mueller and Tribe [8] who introduced an alternative stochastic representation for a class of *linear* Cauchy problems $Lu = Fu$, with a potential F, including the linear wave equation. Their starting point is the well-known representation

$$u(t,x) = w(t,x) + \int_0^t \int_{\mathbb{R}^d} V(t-r, x-y) u(t-r, x-y) S(r, dy)\,dr, \tag{1.1}$$

where w is the solution with zero potential, and S is a fundamental solution, restricted to be representable by a signed measure with $\sup_t |S(t, \mathbb{R}^d)| < \infty$. By substituting formally $u(t-r, x-y)$ on the right-hand side by the last expression of u, one formally obtains a candidate representation for u as $\sum_{m\geq 0} H_m(t,x)$, where $H_0 = w$, and $H_m = \int_0^t \int_{\mathbb{R}^d} V(t - r, x - y) H_{m-1}(t-r, x-y) S(r, dy)\,dr$, $m \geq 1$. Finally, by convenient normalization of the kernel S, the last expression induces a probabilistic representation.

Subsequently, Bakhtin and Mueller [2] considered the one-dimensional nonlinear wave equation

$$\partial_t^2 u - \partial_x^2 u = \sum_{j\geq 0} a_j u^j \quad \text{on } \mathbb{R}_+ \times \mathbb{R}, \qquad u(0,\cdot) = f_1, \qquad \partial_t u(0,\cdot) = f_2 \quad \text{on } \mathbb{R},$$

[1] Notice however that this solves only partly the curse of dimensionality as the variance of the error may still depend on the dimension.

and obtained a probabilistic representation by means of stochastic cascades, which mimics exactly the McKean branching diffusion representation of the KPP equation [20], similar to [13, 15]. Our main contribution in this paper is to show that such a representation holds for a wider class of Cauchy problems, in arbitrary dimension, and with analytic nonlinearity in $(u, \partial_x u)$.

Our starting point is that the representations of [8] and [2] are closely related to the McKean [20] representation of KPP equations, and the corresponding extensions in [13–15]. This in fact opens the door to a much wider extension reported in Section 2 in the context of linear Cauchy problems with constant coefficients, and in Section 3 in the context of nonlinear Cauchy problems with constant coefficients, and analytic nonlinearity in the value function u. The crucial tool for our extension is the so-called Duhamel formula which expresses the solution of such an equation as the convolution (i.e., integration) of the boundary conditions with respect to a family of fundamental solutions. This is in contrast with (1.1) which uses the single fundamental solution S, and requires that the solution w of the zero-potential equation be given. Then, the probabilistic representation appears naturally after convenient normalization of the fundamental solutions.

It is also remarkable that the Malliavin automatic differentiation technique, exploited in [14, 16] in order to address semilinear parabolic second order PDEs, extends naturally to the context of general Cauchy problems by introducing the space gradient of the fundamental solution. This observation is the key-ingredient for our extension in Section 4 to a general class of semilinear Cauchy problems with analytic nonlinearity in $(u, \partial_x u)$.

The performance of the Monte Carlo numerical method induced by our representation is illustrated in Section 5 on some relevant examples from mathematical physics. We start with two examples of semilinear wave equations: the Klein–Gordon equation in dimensions 1,2,3, which has a power nonlinearity in u, and a simplified version of the Yang–Mills equation (in dimension 1), which contains a nonlinearity in the space gradient. We observe that due to some restricting conditions which will be detailed in Sections 2, 3 and 4, our Monte Carlo approximation method does not apply to the multi-dimensional wave equation with space gradient nonlinearity. We next report some numerical experiments in the context of the nonlinear one-dimensional Beam equation which contains two derivatives in time and four space derivatives. We finally consider the Gross–Pitaevskii equation in dimensions 1, 2 and 3, as an example of nonlinear Schrödinger equations. All of the numerical results reveal an excellent performance of our Monte Carlo approximation method.

We finally emphasize that, throughout the paper, we consider the Cauchy problem on $\mathbb{R}_+ \times \mathbb{R}^d$, thus ignoring the important case of restricted domain $\mathcal{D} \subset \mathbb{R}^d$ for the space variable. We mention that Chatterjee [7] proved that the function $u(t,x) = \mathbb{E}_x[f(tX + \sqrt{\tau}Z, B_\tau)]$, with independent r.v. X, Z with standard Cauchy and normal distributions, respectively, and $\tau := \inf\{t > 0 : B_t \notin \mathcal{D}\}$ is the first exit time of an independent Brownian motion from the domain $\mathcal{D}$, solves the wave equation on $\mathbb{R}_+ \times \mathcal{D}$. However, this does not provide a representation for the wave equation on a restricted space domain as the determination of f from given $f_1(x) := u(0,x)$ and $f_2(x) := \partial_t u(0,x)$ is not transparent.

2. Probabilistic representation for linear Cauchy problems.

2.1. *Nonhomogeneous Cauchy problem.* For a smooth function $\phi : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}$, we denote $\partial_t^0 \phi = D^0 \phi = \phi$, and

$$\partial_t^j \phi := \frac{\partial^j \phi}{\partial t^j} \quad \text{and} \quad D^\alpha \phi := \frac{\partial^{|\alpha|} \phi}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}, \quad \text{for all } j \geq 1, \alpha \in \mathbb{N}^d.$$

Given two integers $N, M \geq 1$, we denote $\mathbb{N}_M^d := \{\alpha \in \mathbb{N}^d : |\alpha| \leq M\}$, and we consider some scalar parameters $(a_j)_{1\leq j\leq N}, (b_\alpha)_{\alpha\in\mathbb{N}_M^d} \subset \mathbb{R}$ with

$$a_N = 1 \quad \text{and} \quad \{\alpha \in \mathbb{N}_M^d : |\alpha| = M \text{ and } b_\alpha \neq 0\} \neq \varnothing.$$

Throughout this paper, we consider nonlinear partial differential equations defined by means of the following linear Cauchy problem

$$\sum_{n=1}^{N} a_n \partial_t^n u - \sum_{\alpha\in\mathbb{N}_M^d} b_\alpha D^\alpha u = F \quad \text{on } \mathbb{R}_+ \times \mathbb{R}^d, \tag{2.1}$$

$$\partial_t^{n-1} u(0,\cdot) = p_n f_n \quad \text{on } \mathbb{R}^d, n = 1, \dots, N, \tag{2.2}$$

where the boundary data and the source term satisfy the following conditions:

$$f_n : \mathbb{R}^d \longrightarrow \mathbb{R}, \quad 1 \leq n \leq N \quad \text{and} \quad F : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R} \quad \text{are bounded continuous,} \tag{2.3}$$

and $p_1, \dots, p_N$ are scalar parameters in the simplex:

$$p_n > 0 \quad \text{for all } n = 1, \dots, N, \text{ and } p_1 + \cdots + p_N = 1.$$

Clearly, one can reduce to the case $p_n = N^{-1}$, for all $n = 1, \dots, N$, however we keep introducing such a finite probability measure so as to highlight that this arbitrariness may be useful from the numerical viewpoint.

2.2. *Duhamel's formula for linear Cauchy problems.* We next recall the general solution of nonhomogeneous Cauchy problems with constant coefficients. We first introduce the $\mathbb{C}^N$-valued function with components $\hat{g} := (\hat{g}_1, \dots, \hat{g}_N)$:

$$\hat{g}(t,\xi) := (2\pi)^{-\frac{d}{2}} e^{tB(\xi)^{\mathrm{T}}} \mathrm{e}_1, \quad t \geq 0, \xi \in \mathbb{R}^d, \tag{2.4}$$

where $(\mathrm{e}_1, \dots, \mathrm{e}_N)$ is the canonical basis of $\mathbb{R}^N$,

$$B(\xi) := \left(\begin{array}{c|c} 0 & \\ \vdots & I_{N-1} \\ 0 & \\ \hline b(\xi) & -a_1 \quad \cdots \quad -a_{N-1} \end{array}\right) \quad \text{and} \quad b(\xi) := \sum_{\alpha\in\mathbb{N}_M^d} i^{|\alpha|} b_\alpha \xi^\alpha, \quad \xi \in \mathbb{R}^d, \tag{2.5}$$

with $\xi^\alpha := \xi_1^{\alpha_1} \dots \xi_d^{\alpha_d}$, for all multi-index $\alpha \in \mathbb{N}^d$. The polynomial function $b : \mathbb{R}^d \longrightarrow \mathbb{R}$ is called *the symbol* of the Cauchy problem (2.1).

As standard, we denote by $\mathcal{S}$ the Schwartz space of rapidly decreasing functions on $\mathbb{R}^d$, and by $\mathcal{S}'$ the corresponding dual space of tempered distributions. We recall that this space contains all (distributions represented by) polynomially growing functions. However, notice that in dimension $d = 1$, by taking $M = 1$, $N = 2$, $b_0 = 0$, $b_1 = 0$, and $b_2 = -1$, we obtain $\hat{g}(t,\xi) = e^{t\xi^2}$. In particular, $\hat{g}(t,\cdot) \notin \mathcal{S}'$ for all $t > 0$ (this corresponds the Heat equation with the wrong sign for the diffusion coefficient). For this reason, the following assumption is needed. Unfortunately, we could not find in the literature any sufficient conditions on the coefficients of the PDE.

ASSUMPTION 2.1. There exists $T \in (0, \infty]$ such that $\hat{g}(t,\cdot) \in \mathcal{S}'$ for all $t \in [0, T)$.

Under this assumption, we may introduce the so-called Green functions as the inverse Fourier transform with respect to the space variable:

$$g(t,\cdot) := \mathfrak{F}^{-1} \hat{g}(t,\cdot), \quad t \in [0, T), \tag{2.6}$$

in the distribution sense, that is, $\langle \mathfrak{F}^{-1}\hat{g}(t,\cdot),\varphi\rangle = \langle \hat{g}(t,\cdot),\mathfrak{F}^{-1}\varphi\rangle$ for all $\varphi \in \mathcal{S}$, $t \in [0,T)$, where

$$\mathfrak{F}^{-1}\varphi(x) := (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{i\xi\cdot x}\varphi(\xi)\,d\xi \quad \text{for all } \varphi \in \mathcal{S}.$$

ASSUMPTION 2.2. For all $n = 1,\dots,N$:

(i) $(t,x) \longmapsto (g_n(t,\cdot) * \phi)(x)$ is continuous on $[0,T)\times\mathbb{R}^d$, for all bounded continuous function ϕ on $\mathbb{R}^d$;

(ii) $g_n(t,\cdot)$ may be represented by a signed measure $g_n(t,dx) = g_n^+(t,dx) - g_n^-(t,dx)$, with total variation measure $|g_n| := g_n^+ + g_n^-$ absolutely continuous with respect to some probability measure μ_n; the corresponding densities γ_n, γ_n^+ and γ_n^-, defined by

$$g_n^+(t,dx) = \gamma_n^+(t,x)\mu_n(t,dx), \qquad g_n^-(t,dx) = \gamma_n^-(t,x)\mu_n(t,dx), \qquad \gamma_n := \gamma_n^+ - \gamma_n^-,$$

satisfy $|\gamma_n(t,\cdot)|_\infty < \infty$, $\gamma_n(t,\mathbb{R}^d) < \infty$, and $\gamma_N(\cdot,\mathbb{R}^d) \in \mathbb{L}^1([0,t])$ for all $t \in [0,T)$.

Again, we could not find any sufficient conditions on the coefficients of the PDE which guarantee the validity of the last requirements. Notice that there is no hope for item (i) to hold without any further conditions. For instance, Example 2.9 below illustrates a situation where $g_1(t,\cdot) \notin \mathbb{L}^1(\mathbb{R})$, even in one space dimension. Then, the constant function $\varphi := \mathbf{1}$ is so that $g_n(t,\cdot) * \mathbf{1} = \infty$.

REMARK 2.3. As $g_n(t,\cdot) \in \mathbb{L}^1(\mathbb{R}^d)$ by Assumption 2.2, we may choose $\mu_n(t,dz) = \|g_n(t,\cdot)\|^{-1}_{\mathbb{L}^1(\mathbb{R}^d)}|g_n|(t,dz)$ and $\gamma_n(t,z) = \operatorname{sgn}(g_n)(t,z)\|g_n(t,\cdot)\|_{\mathbb{L}^1(\mathbb{R}^d)}$, where $\operatorname{sgn}(\alpha) := \mathbf{1}_{\{\alpha\ge 0\}} - \mathbf{1}_{\{\alpha<0\}}$, and we may have a simplified statement of Assumption 2.2 in terms of this particular choice of dominating measures. However by introducing more general dominating measures μ_n, we would like to highlight that this additional degree of freedom may have important numerical implications of the subsequent representation results of the paper.

PROPOSITION 2.4. *Let F, $(f_n)_{1\le n\le N}$ be as in (2.3). Then, under Assumptions 2.1 and 2.2, the Cauchy problem (2.1)–(2.2) has a unique solution in $C_b^0([0,T)\times\mathbb{R}^d,\mathbb{R})$ given by*

$$u(t,x) := \sum_{n=1}^N p_n\big(f_n * g_n(t,\cdot)\big)(x) + \int_0^t \big(F(t-s,\cdot) * g_N(s,\cdot)\big)(x)\,ds, \tag{2.7}$$
$$t \in [0,T),\ x \in \mathbb{R}^d.$$

Of course the last result is well known, however we emphasize that it is usually stated under different assumptions on g_n, f_n and F. Namely, one may typically relax Assumption 2.2 and assume that f and F have bounded support so as to guarantee that the convolutions involved in the representation (2.7) are well defined and satisfy the property $\mathfrak{F}(f_n * g_n(t,\cdot)) = (2\pi)^{\frac{d}{2}}\mathfrak{F}(f_n)\mathfrak{F}(g_n(t,\cdot))$ for all $t \in [0,T)$, $n = 1,\dots,N$. Our conditions in Proposition 2.4 are suitable for the subsequent use in this paper both for the nonlinear extension and, more importantly, for the numerical implications in terms of Monte Carlo approximation.

For the convenience of the reader, we report the proof of Proposition 2.4.

PROOF. First, the conditions on the densities γ_n contained in Assumption 2.2 guarantee that the function u defined in (2.7) is bounded and continuous. Then the distribution represented by u is in $\mathcal{S}'$, and we may define the corresponding Fourier transform in the space

variable $\hat{u}(t,\cdot) := \mathfrak{F}(u(t,\cdot))$ in the distribution sense. By standard calculation using the properties of the Fourier transform, we see that

$$\hat{u}(t,\xi) = \sum_{n=1}^{N} p_n \hat{f}_n(\xi)\hat{g}_n(t,\xi) + \int_0^t \hat{F}(t-s,\xi)\hat{g}_N(s,\xi)\,ds, \quad t \geq 0, \xi \in \mathbb{R}^d, \tag{2.8}$$

where $\hat{f} := \mathfrak{F}(f)$ and $\hat{F} := \mathfrak{F}(F(t,\cdot))$. By the definition of $\hat{g}$ in (2.4), we see that, for every fixed $\xi \in \mathbb{R}^d$, the function $\hat{u}(\cdot,\xi)$ is the unique solution of the ODE

$$\sum_{n=1}^{N} a_n \partial_t^n \hat{u} - b(\xi)\hat{u} = \hat{F}(\cdot,\xi) \quad \text{on } \mathbb{R}_+, \qquad \partial_t^{n-1}\hat{u}(0,\xi) = p_n \hat{f}_n(\xi), \quad n = 1,\dots,N, \tag{2.9}$$

which can be written equivalently in terms of the function $\hat{v} := (\hat{u}, \partial_t\hat{u}, \dots, \partial_t^{N-1}\hat{u})^\top$:

$$\partial_t \hat{v} = B\hat{v} + \hat{F}\mathrm{e}_N, \quad \text{on } \mathbb{R}_+ \quad \text{and} \quad \hat{v}(0,\cdot) = \sum_{n=1}^{N} p_n \hat{f}_n \mathrm{e}_n.$$

Obviously, the last ODE has a unique solution with closed form obtained by the variation of the constant method $\hat{v}(t,\xi) := e^{tB(\xi)}\hat{v}(0,\xi) + \int_0^t e^{sB(\xi)}\hat{F}(t-s)e_N\,ds$, whose first entry induces the solution $\hat{u}$ introduced in (2.8). To conclude the proof, it suffices to observe that any solution $\tilde{u} \in C_b^0([0,T)\times\mathbb{R}^d)$ of the Cauchy problem (2.1)–(2.2) has a well-defined Fourier transform in the distribution sense satisfying the ODE (2.9) for all fixed $\xi \in \mathbb{R}^d$. □

REMARK 2.5. For later use, we provide some details about the diagonalization of the matrix $B(\xi)$, needed for the computation of the Green functions g as the inverse Fourier transform of the associated $\hat{g}$. Direct examination reveals that the eigenvalues of the matrix $B(\xi)$ are the roots of the corresponding characteristic polynomial $b(\xi) - \sum_{n=1}^{N} a_n\lambda^n$. Assume that $B(\xi)$ has N distinct (simple) eigenvalues $(\lambda_j(\xi))_{1\leq n\leq N} \in \mathbb{C}^N$. Then, denoting by $\mathrm{diag}[\lambda]$ the diagonal matrix with $\mathrm{diag}[\lambda]_{j,j} = \lambda_j$, it follows that

$$B(\xi) = P(\xi)\,\mathrm{diag}\big[\lambda(\xi)\big]P(\xi)^{-1}, \quad \text{where } P(\xi)_{j,\ell} := \lambda_\ell(\xi)^j, 1 \leq j,\ell \leq N.$$

The matrix $P(\xi)$ is the so-called Vandermonde matrix whose inverse is given by

$$\big\{P(\xi)^{-1}\big\}_{j,n} = \frac{\Lambda_n(\xi)}{\lambda_j(\xi)\prod_{\ell\neq j}(\lambda_\ell - \lambda_j)(\xi)}, \quad j,n = 1,\dots,N,$$

where

$$\Lambda_N(\xi) = 1 \quad \text{and} \quad \Lambda_n(\xi) := (-1)^{n-1} \sum_{\substack{1\leq \ell_1 \leq \cdots \leq \ell_{N-n} \leq N \\ \ell_1,\dots,\ell_{N-n}\neq n}} (\lambda_{\ell_1}\cdots\lambda_{\ell_{N-n}})(\xi) \quad \text{for } n < N.$$

Therefore, by the definition of $\hat{g}_n$ in (2.4), we have for $(r,\xi) \in \mathbb{R}_+\times\mathbb{R}^d$:

$$\hat{g}_n(r,\xi) = (2\pi)^{-\frac{d}{2}}\Lambda_n(\xi)\sum_{j=1}^{N}\frac{e^{r\lambda_j(\xi)}}{\prod_{\ell\neq j}(\lambda_\ell - \lambda_j)(\xi)}, \quad n = 1,\dots,N. \tag{2.10}$$

2.3. *Probabilistic representation.* Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space supporting two random variables τ and I, with

$$\tau \text{ and } I \text{ independent}, \ \mathbb{P}[\tau \in dt] = \rho(t)\mathbf{1}_{\{t\geq 0\}}\,dt \text{ and } \mathbb{P}[I = n] = p_n, n = 1,\dots,N, \tag{2.11}$$

for some $C^0(\mathbb{R}_+,\mathbb{R})$ density function $\rho > 0$ on $(0,\infty)$. We shall denote $\bar{\rho}(t) := \int_t^\infty \rho(s)\,ds$.

Recall the densities γ_n and the dominating probability measures $\mu_n(t,\cdot)$, $n=1,\dots,N$, introduced in Assumption 2.2. For all $t\geq 0$, we introduce the random variables

$$X_t^n := x + Z_t^n \text{ independent of } (I,\tau), \text{ with } \mathbb{P}\big[Z_t^n \in dz\big] = \mu_n(t,dz), n=1,\dots,N. \tag{2.12}$$

The following representation is a rewriting of Proposition 2.4 in terms of the last notation.

PROPOSITION 2.6. *Let $(f_n)_{1\leq n\leq N}$ and F be as in (2.3). Then, under Assumptions 2.1 and 2.2, the unique $C_b^0([0,T]\times\mathbb{R}^d,\mathbb{R})$ solution of the Cauchy problem (2.1)–(2.2) is*:

$$u(t,x) = \mathbb{E}\bigg[\mathbf{1}_{\{\tau\geq t\}}\frac{\gamma_I(t,Z_t^I)}{\bar\rho(t)}f_I\big(X_t^I\big) + \mathbf{1}_{\{\tau<t\}}\frac{\gamma_N(\tau,Z_\tau^N)}{\rho(\tau)}F\big(t-\tau,X_\tau^N\big)\bigg],\quad t<T, x\in\mathbb{R}^d.$$

2.4. *Examples*. We conclude this section by examining some examples. In particular Example 2.9 does not satisfy Assumption 2.2.

EXAMPLE 2.7 (Heat equation). Let $N=1$, $M=2$, $b_\alpha=0$ whenever $|\alpha|\leq 1$, and $b_{1,1}=b_{2,2}=1$, $b_{1,2}=b_{2,1}=0$. Then $B(\xi)=b(\xi)=-|\xi|^2$, $\xi\in\mathbb{R}^d$, implying that Assumption 2.3 is valid with $T=\infty$, and

$$g_1(r,z) := (2\pi)^{-d}\int e^{-|\xi|^2 r+i\xi z}\,d\xi = (4\pi r)^{-d/2}e^{-\frac{|z|^2}{4r}},\quad (r,z)\in\mathbb{R}_+\times\mathbb{R}^d,$$

is the heat kernel which is positive, has unit total mass, and satisfies Assumption 2.2. Then, choosing $\gamma_1(t,\cdot)=1$ leads to $Z_t^1=\sqrt{2t}Z$ with $Z\in N(0,I_d)$ a d-Gaussian r.v.

EXAMPLE 2.8 (Airy equation). Let $d=1$, $N=1$, $M=3$, $b_0=b_1=b_2=0$ and $b_3=1$. Then $B(\xi)=b(\xi)=-i\xi^3$ is scalar valued, and

$$g_1(r,z) := (2\pi)^{-1}\int e^{i\xi z-r\xi^3}\,d\xi = (3r)^{-1/3}\mathrm{Ai}\big((3r)^{-1/3}z\big),\quad (r,z)\in\mathbb{R}_+\times\mathbb{R},$$

where we introduced the Airy function $\mathrm{Ai}(x) := (\pi)^{-d}\int_0^\infty \cos\big(-\frac{\xi^3}{3}+x\xi\big)\,d\xi$, $x\in\mathbb{R}$. Notice that $\int|g_1|(t,dz)=\infty$, so that Assumption 2.2 fails in this example.

EXAMPLE 2.9 (Wave equation). Let $N=2$, with $a_1=0$, $M=2$, $b_\alpha=0$ whenever $|\alpha|\leq 1$, and $b_{1,1}=b_{2,2}=1$, $b_{1,2}=b_{2,1}=0$. Then

$$B(\xi)=\begin{pmatrix}0 & 1\\ -|\xi|^2 & 0\end{pmatrix},\qquad e^{B(\xi)r}=(2\pi)^{-\frac{d}{2}}\begin{pmatrix}\cos\big(r|\xi|\big) & |\xi|^{-1}\sin\big(r|\xi|\big)\\ -|\xi|\sin\big(r|\xi|\big) & \cos\big(r|\xi|\big)\end{pmatrix},\quad \xi\in\mathbb{R}^d,$$

and, for $(r,z)\in\mathbb{R}_+\times\mathbb{R}^d$,

$$g_2(r,dz) = (2\pi)^{-d}\int\frac{\sin\big(r|\xi|\big)}{|\xi|}e^{i\xi\cdot z}\,d\xi,$$

$$g_1(r,dz) = (2\pi)^{-d}\int\cos\big(r|\xi|\big)e^{i\xi\cdot z}\,d\xi = \partial_r g_2(r,dz).$$

Direct calculations, using Kirchhoff's formula (in) see, for example, [10]), provide

$$g_2(r,dz) = \begin{cases}\dfrac{1}{2}1_{\{|z|<r\}}\,dz & \text{for } d=1,\\ \dfrac{1}{2\pi}\big(r^2-|z|^2\big)^{-\frac{1}{2}}1_{\{|z|<r\}}\,dz & \text{for } d=2,\\ \dfrac{\sigma_r(dz)}{4\pi r} & \text{for } d=3,\end{cases}$$

where $\sigma_r(dz)$ denotes the surface area on $\partial B(0,r)$. We directly compute that $\|g_2(t,\cdot)\|_{\mathbb{L}^1(\mathbb{R}^d)} = t$. Then, choosing $\gamma_2(t,\cdot) = t$ leads to $Z_t^2 = tZ$ where:

- in dimension $d = 1$, Z has a uniform distribution on $[-1,1]$,
- in dimension $d = 2$, the law of Z is defined by the density $\frac{1}{2\pi}\frac{1}{\sqrt{1-z^2}}1_{|z|<1}$,
- in dimension $d = 3$, the law of Z is $\frac{1}{4\pi}\mu_{S^2}(dz)$, where μ_{S^2} denotes the volume measure on the unit sphere.

EXAMPLE 2.10 (Beam equation). Let $d = 1$, $N = 2$, with $a_1 = 0$, $M = 4$, $b_1 = b_2 = b_3 = 0$ and $b_4 = 1$ corresponding to the fourth-order PDE $\partial_t^2 u + D^4 u = 0$. Then,

$$g_1(r,z) = \partial_r g_2(r,z) \quad \text{and} \quad g_2(r,z) = \sqrt{r}G\left(\frac{z}{\sqrt{r}}\right) \quad \text{where}$$

$$G(0) = (2\pi)^{-\frac{1}{2}}, \tag{2.13}$$

$$2G'(x) = \operatorname{Im}\left(\int_0^{\frac{x}{\sqrt{2\pi}}} e^{i\pi t^2/2}\,dt\right) - \operatorname{Re}\left(\int_0^{\frac{x}{\sqrt{2\pi}}} e^{i\pi t^2/2}\,dt\right), \quad x \in \mathbb{R}.$$

Note that $\int_{\mathbb{R}} |G(x)|\,dx < \infty$ as $G(x) \sim_{|x|\to\infty} (\frac{2}{\pi})^{\frac{1}{2}} x^{-2}(\cos(\frac{x^2}{4}) - \sin(\frac{x^2}{4}))$. Then, we may choose $\gamma_2(t,z) = t\|G\|_{\mathbb{L}^1(\mathbb{R}^d)}\operatorname{sgn}G(z)$, thus inducing $Z_t^2 = \sqrt{t}Z$ with Z distributed according to $\|G\|^{-1}_{\mathbb{L}^1(\mathbb{R}^d)}|G(z)|$.

All of the results of the present paper extend to the case of Cauchy problems for complex-valued functions, with coefficients $(a_n)_{1\le n<N}$ and $(b_\alpha)_{\alpha\in\mathbb{N}_M^d}$ in $\mathbb{C}$, thus allowing to include, for instance, the Schrödinger equation (see Example 2.11) and its semilinear extension as the Gross–Pitaevskii PDE (see Section 5.4). This extension follows by simply applying the methodology described throughout the paper separately to the real part and the imaginary part of the representation.

EXAMPLE 2.11 (Schrödinger equation and analytical continuation). Consider the Schrödinger equation of a free particle (with source term)

$$i\,\partial_t u = -\frac{1}{2}\Delta u + F, \quad u(0,\cdot) = f_1. \tag{2.14}$$

The corresponding Duhamel formula is

$$u(t,x) = \int_{\mathbb{R}^d} f_1(y) g_1(t, x-y)\,dy - i\int_0^t \int_{\mathbb{R}^d} g_1(s, x-y) F(s,y)\,ds\,dy,$$

with $g_1(t,x) = (2\pi i t)^{\frac{-d}{2}} e^{-\frac{x^2}{2it}}$. Note the coefficient $-i$ in front of F as $a_1 = i$ here. By setting $y - x = \sqrt{it}z$, this can be written as

$$u(t,x) = \int_{-\infty e^{-\frac{i\pi}{4}}}^{\infty e^{-\frac{i\pi}{4}}} f_1\big(x + e^{\frac{i\pi}{4}}\sqrt{t}z\big)e^{-\frac{z^2}{2}}\,dz - i\int_0^t \int_{-\infty e^{\frac{-i\pi}{4}}}^{\infty e^{\frac{-i\pi}{4}}} F\big(s, x + e^{\frac{i\pi}{4}}\sqrt{s}z\big)e^{-\frac{z^2}{2}}\,dz\,ds.$$

By assuming that $|f_1(x + e^{\frac{i\pi}{4}+i\theta}\sqrt{t}R)|e^{-\frac{R^2}{2}\cos(2\theta)}$ and $|F(s, x + e^{\frac{i\pi}{4}+i\theta}\sqrt{s}R)|e^{-\frac{R^2}{2}\cos(2\theta)}$ goes to *zero* as $R \to \infty$ when $\theta \in [-\frac{\pi}{4}, 0]^d \cup [\frac{3\pi}{4}, \pi]^d$, the integration over z in $[-\infty e^{-\frac{i\pi}{4}}, \infty e^{-\frac{i\pi}{4}}]^d$ can be deformed into $[-\infty, +\infty]^d$ by analytical continuation and we obtain for some standard Gaussian r.v. Z on $\mathbb{R}^d$:

$$u(t,x) = \mathbb{E}\big[f_1\big(x + e^{\frac{i\pi}{4}}\sqrt{t}Z\big)\big] - i\int_0^t \mathbb{E}\big[F\big(s, x + e^{\frac{i\pi}{4}}\sqrt{s}Z\big)\big]\,ds.$$

3. A first class of semilinear Cauchy problems. In this section, we consider the semilinear Cauchy problem:

$$\sum_{n=1}^{N} a_n \partial_t^n u - \sum_{\alpha \in \mathbb{N}_M^d} b_\alpha D^\alpha u = \sum_{j \geq 0} q_j c_j u^j \quad \text{on } \mathbb{R}_+ \times \mathbb{R}^d, \tag{3.1}$$

$$\partial_t^{n-1} u(0, \cdot) = p_n f_n \quad \text{on } \mathbb{R}^d, n = 1, \ldots, N, \tag{3.2}$$

where the nonlinearity is defined by means of an atomic probability measure $(q_j)_{j \geq 0}$, with $q_j \geq 0$ and $\sum_{j \geq 0} q_j = 1$, together with the functions

$$c_j : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}, \quad j \geq 0, \text{ bounded continuous, for all } j \geq 0. \tag{3.3}$$

The power series on the right-hand side of (3.1) is well defined provided that

$$H_1(s) := \sum_{j \geq 0} q_j \|c_j\|_\infty s^j \tag{3.4}$$

has a strictly positive radius of convergence, so that it is well defined at least in some neighborhood of the origin.

3.1. *The branching mechanism.* A particle of generation $\nu \in \mathbb{N}$ is a multi-integer $k := (0, k_1, \ldots, k_\nu) \in \mathbf{K}_\nu := \{0\} \times \mathbb{N}^\nu$. With this notation, the ancestor is denoted by 0, the children of the ancestor by $(0, 1), (0, 2), \ldots$, the children of $(0, 1)$ are $(0, 1, 1), (0, 1, 2), \ldots$, and so on.

We set by convention $\mathbf{K}_0 := \{0\}$, and we denote by $\mathbf{K} := \bigcup_{\nu \geq 0} \mathbf{K}_\nu$ the collection of all particles. For $\nu \geq 1$, and a particle $k := (0, k_1, \ldots, k_\nu) \in \mathbf{K}_\nu$, we denote by $k_- := (0, k_1, \ldots, k_{\nu-1}) \in \mathbf{K}_{\nu-1}$ its parent particle.

We next introduce independent families of random variables $(\tau_k, I_k^0, J_k)_{k \in \mathbf{K}}$:

- I_k^0 and τ_k are i.i.d. copies of the random variables I and τ, respectively, as introduced in (2.11);
- J_k are i.i.d. random variables with $\mathbb{P}[J_k = j] = q_j$ for all $j \geq 0$.

The time occurrences of the branching events, prior to t, are recorded through the sequence $(T_k^t)_k$ defined for all $t \geq 0$ by

$$T_{0-}^t := 0 \quad \text{and} \quad T_k^t := t \wedge (T_{k_-}^t + \tau_k), \quad \text{for all } k \in \mathbf{K}.$$

With these notation, each particle k lives on the time interval $[T_{k_-}^t, T_k^t]$. The branching mechanism is then the following:

- Start from particle 0.
- At the first branching time $T_0^t = t \wedge \tau_0$, particle zero dies out; if $T_0^t < t$, it generates J_0 descendants labelled $(0, 1), \ldots, (0, J_0)$.
- Each descendant particle k undergoes the same behavior, independently of its peer-particles: it dies out at the branching time T_k^t, and generates J_k descendants labelled $(k, 1), \ldots, (k, J_k)$, whenever $T_k^t < t$.
- We denote by $\mathcal{K}_s^t$ the collection of all living particles at time s, and by $\overline{\mathcal{K}}_t := \bigcup_{s \leq t} \mathcal{K}_s^t$ the collection of all particles which have been living prior to time t. For simplicity we set $\mathcal{K}_t := \mathcal{K}_t^t$.
- We finally denote for all particles $k \in \overline{\mathcal{K}}_t$

$$I_k^t := I_k^0 \mathbf{1}_{k \in \mathcal{K}_t} + N \mathbf{1}_{k \in \overline{\mathcal{K}}_t \setminus \mathcal{K}_t}. \tag{3.5}$$

3.2. *Probabilistic representation.* Given the branching mechanism defined in the previous subsection, we introduce the corresponding branching process:

$$X_0^0 = x \quad \text{and} \quad X^k_{T^t_k} := X^{k_-}_{T^t_{k_-}} + Z^k_{T^t_k} \quad \text{for all } k \in \overline{\mathcal{K}}_t, \tag{3.6}$$

where $(Z^k_{T^t_k})_k$ are independent with distribution, conditional on the type I^t_k and the lifetime $T^t_k - T^t_{k_-}$ of the particle k, given by

$$\mathbb{P}\big[Z^k_{T^t_k} \in dz | I^t_k, \Delta T^t_k\big] = \mu_{I^t_k}(\Delta T^t_k, dz) \quad \text{with } \Delta T^t_k := T^t_k - T^t_{k_-}. \tag{3.7}$$

We now introduce, for all $t \geq 0$, and $x \in \mathbb{R}^d$, the random variable

$$\xi_{t,x} := \prod_{k \in \mathcal{K}_t} \frac{\gamma_{I^t_k}(\Delta T^t_k, Z^k_{T^t_k})}{\bar{\rho}(\Delta T^t_k)} f_{I^t_k}(X^k_t) \prod_{k \in \overline{\mathcal{K}}_t \setminus \mathcal{K}_t} \frac{\gamma_N(\Delta T^t_k, Z^k_{T^t_k})}{\rho(\Delta T^t_k)} c_{J_k}\big(t - T^t_k, X^k_{T^t_k}\big). \tag{3.8}$$

Recall the power series H_1 defined in (3.4).

ASSUMPTION 3.1. The power series H_1 has a radius of convergence $R_1 \in (0, \infty]$, and:

(i) $r_1 := \sup_{1 \leq n \leq N} \|f_n\|_\infty \|\gamma_n\|_\infty < R_1$;
(ii) there are constants $T > 0$ and $s_1 > r_1$ such that $\int_{r_1}^{s_1} H_1(s)^{-1}\, ds = T \|\gamma_N\|_\infty$.

We observe that in most of the practical situations, only a finite number of coefficients c_j are nonzero, inducing $R_1 = +\infty$. In this case (i) is a slight reinforcement of Assumption 2.2. Finally, (ii) is a standard condition for Galton–Watson branching mechanisms.

THEOREM 3.2. *Let $f \in C^0_b(\mathbb{R}^d, \mathbb{R}^n)$ and $c_j \in C^0_b(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R})$. Then, under Assumptions 2.1, 2.2 and 3.1, we have $\xi_{t,x} \in \mathbb{L}^1$ and $u(t,x) := \mathbb{E}[\xi_{t,x}]$ is a solution of the semilinear Cauchy problem (3.1)–(3.2).*

PROOF. 1. We first verify the $\mathbb{L}^1$-integrability of $\xi_{t,x}$. From the expression of $\xi_{t,x}$ in (3.8), we see that

$$|\xi_{t,x}| \leq \chi_t := \prod_{k \in \mathcal{K}_t} \frac{r_1}{\bar{\rho}(\Delta T^t_k)} \prod_{k \in \overline{\mathcal{K}}_t \setminus \mathcal{K}_t} \frac{\overline{c}_{J_k} \|\gamma_N\|_\infty}{\rho(\Delta T^t_k)}. \tag{3.9}$$

Our objective is now to prove that $t \longmapsto \omega(t) := \mathbb{E}\chi_t < \infty$. By standard arguments, this function is related to the ODE

$$\partial_t w = \|\gamma_N\|_\infty H_1(w) \quad \text{and} \quad w(0) = r_1. \tag{3.10}$$

Let us first verify that this ODE has a nonexploding solution under Assumption 3.1(ii). As the power series function H_1 has radius R_1, the condition $r_1 < R_1$ is necessary to find a (finite) solution of the last ODE. Next, for an arbitrary $L > 0$, notice that H_1 is Lipschitz on $[-L, L]$, so that the last ODE has a unique nonnegative solution w up to T_L, where

$$\lim_{t \to T_L} w(t) = L \quad \text{whenever } T_L < \infty,$$

and the nonnegativity of w follows from the fact that $w(0) > 0$ and all derivatives $H_1^{(i)}(0) \geq 0$ for all $i \geq 0$. Then, it follows from direct integration of the ODE that

$$\|\gamma_N\|_\infty T_L = \int_0^{T_L} \frac{\partial_s w}{H_1(w)}(s)\, ds = \int_{r_1}^{L} \frac{ds}{H_1(s)} \quad \text{as long as } T_L < \infty.$$

This proves that w is a nonexploding solution on $[0,T]$ if and only if $\int_{r_1}^{s_1} \frac{ds}{H_1(s)} = \|\gamma_N\|_\infty T$ for some constant $s_1 > r_1$. Under this condition, we obtain by direct integration of (3.10)

$$\begin{aligned} w(t) &= r_1 + \|\gamma_N\|_\infty \sum_{j\geq 0} q_j \bar{c}_j \int_0^t w(s)^j \, ds \\ &= \mathbb{E}\Big[\frac{r_1 \mathbf{1}_{\{\tau_0 \geq t\}}}{\bar{\rho}(t)} + \mathbf{1}_{\{\tau_0 < t\}} \frac{\|\gamma_N\|_\infty \|c_{J_0}\|_\infty}{\rho(\tau_0)} w(\tau_0)^{J_0} \Big] \\ &= \mathbb{E}\Big[\frac{r_1 \mathbf{1}_{\{\tau_0 \geq t\}}}{\bar{\rho}(t)} + \mathbf{1}_{\{\tau_0 < t\}} \frac{\|\gamma_N\|_\infty \|c_{J_0}\|_\infty}{\rho(\tau_0)} \\ &\quad \times \prod_{k \in \mathcal{K}_t^1} \Big(\frac{r_1 \mathbf{1}_{\{T_k^t \geq t\}}}{\bar{\rho}(\tau_k)} + \mathbf{1}_{\{T_k^t < t\}} \frac{\|\gamma_N\|_\infty \|c_{J_0}\|_\infty}{\rho(\tau_k)} w(T_k^t)^{J_k} \Big) \Big], \end{aligned}$$

where $\mathcal{K}_t^n$ denotes the particles generated after n branching events prior to t and the last two equalities follow from the definition of the branching mechanism together with the tower property. Iterating up to the nth generation, this provides

$$\begin{aligned} v(t) &= \mathbb{E}\chi_t^n \quad \text{where} \\ \chi_t^n &:= \prod_{k \in \mathcal{K}_t^n} \frac{r_1}{\bar{\rho}(\Delta T_k^t)} \prod_{k \in \bigcup_{j<n} \mathcal{K}_t^j} \frac{\|\gamma_N\|_\infty \|c_{J_0}\|_\infty}{\rho(\Delta T_k^t)} \\ &\quad + \prod_{k \in \bigcup_{j \leq n} \mathcal{K}_t^j} \frac{\|\gamma_N\|_\infty \|c_{J_0}\|_\infty}{\rho(\Delta T_k^t)} \prod_{k \in \mathcal{K}_t^{n+1}} w(T_{k-}^t)^{J_{k-}}. \end{aligned}$$

As $w \geq 0$, it follows that $\chi_t^n \geq 0$, and we deduce from Fatou's lemma that $\infty > w(t) = \liminf_{n\to\infty} \mathbb{E}\chi_t^n \geq \mathbb{E}\chi_t$, implying the required integrability of the bound χ_t in (3.9).

2. We next prove that $v(t,x) := \mathbb{E}[\xi_{t,x}]$ is a solution of the nonlinear Cauchy problem (3.1). To see this, observe that

$$\xi_{t,x} = \mathbf{1}_{\{\tau_0 \geq t\}} \frac{\gamma_{I_0^t}(t, Z_t^0)}{\bar{\rho}(t)} f_{I_0^t}(X_t^0) + \mathbf{1}_{\{\tau_0 < t\}} \frac{\gamma_N(\tau_0, Z_{\tau_0}^0)}{\rho(\tau_0)} c_{J_0}(t-\tau_0, X_{\tau_0}^0) \prod_{j=1}^{J_0} \xi_{t-\tau_0, X_{\tau_0}^0}^{(j)},$$

where $\xi_{t-\tau_0, X_{\tau_0}^0}^{(j)}$ have the same distribution, conditional on $(\tau_0, X_{\tau_0}^0)$. Taking expectations, and using the tower property, it follows from the definition of the function v that

$$\begin{aligned} v(t,x) &= \mathbb{E}\Big[\mathbf{1}_{\{\tau_0 \geq t\}} \frac{\gamma_{I_0^t}(t, Z_t^0)}{\bar{\rho}(t)} f_{I_0^t}(X_t^0) + \mathbf{1}_{\{\tau_0 < t\}} \frac{\gamma_N(\tau_0, Z_{\tau_0}^0)}{\rho(\tau_0)} c_{J_0}(t-\tau_0, X_{\tau_0}^0) v(t-\tau_0, X_{\tau_0}^0)^{J_0} \Big] \\ &= \sum_{n=1}^N \int p_n\big(f_n * g_n(t,\cdot)\big)(x) + \int_0^t \big(F(t-s,\cdot) * g_N(s,\cdot)\big)(x) \, ds, \end{aligned}$$

where $F := \sum_{j\geq 0} q_j c_j v^j$. Since v is bounded, it follows from Proposition 2.4 that v is a $C_b^0([0,T] \times \mathbb{R}^d, \mathbb{R})$ solution of the nonlinear Cauchy problem (3.1)–(3.2). □

REMARK 3.3 (Finite propagation speed). From the simulation of Z_t^2 in the case of the wave equation, we deduce directly the finite propagation speed property: If $f_1 = f_2 = 0$ and $c_0 = 0$ on a ball $B(x_0, t_0)$ of center x_0 and radius t_0, then $u = 0$ is a solution of the PDE within the cone $K(x_0, t_0) := \{(x,t) : 0 \leq t \leq t_0 \text{ and } |x - x_0| \leq t_0 - t\}$.

Motivated by the numerical implications of Section 3.3, we now provide some sufficient conditions for $\xi_{t,x}$ to have a finite pth moment. In particular, for $p = 2$, this guarantees that the error estimate of the Monte Carlo approximation induced by the representation of Theorem 3.2 is characterized by the standard central limit theorem, independently of the dimension d of the space variable x. Our first requirement is on the power series:

$$H_p(s) := \sum_{j\geq 0} (q_j \|c_j\|_\infty)^p s^j. \tag{3.11}$$

ASSUMPTION 3.4. The power series H_p has a radius of convergence $R_p \in (0, \infty]$, and:

(i) $r_p := \max_{1\leq n\leq N} \|f_n\|_\infty^p \|\gamma_n \bar{\rho}^{\frac{1-p}{p}}\|_\infty^p < R_p$, and $\alpha_p := \|\gamma_N \rho^{\frac{1-p}{p}}\|_\infty^p < \infty$;
(ii) there are constants $T > 0$ and $s_p > r_p$ such that $\int_{r_p}^{s_p} H_p(s)^{-1}\, ds = \alpha_p T$.

THEOREM 3.5. *Let $f \in C_b^0(\mathbb{R}^d, \mathbb{R}^n)$ and $c_j \in C_b^0(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R})$. Then, under Assumptions 2.1, 2.2 and 3.4, we have $\xi_{t,x} \in \mathbb{L}^p$ for all $t \in [0, T]$.*

PROOF. Similar to the calculation in the previous proof, we have

$$|\xi_{t,x}|^p \leq \prod_{k\in\mathcal{K}_t} \frac{r_p}{\bar{\rho}(\Delta T_k^t)} \prod_{k\in\overline{\mathcal{K}}_t\setminus\mathcal{K}_t} \frac{\alpha_p \|c_{J_k}\|_\infty^p}{\rho(\Delta T_k^t)}.$$

The right-hand side is now related to the ODE

$$\partial_t w = \alpha_p H_p(w) \quad \text{and} \quad w(0) = r_p.$$

The integrability of $|\xi_{t,x}|^p$ can now be verified by following the same line of argument as in Step 1 of the proof of Theorem 3.2. □

3.3. *Numerical implications.* Assuming that uniqueness holds for the nonlinear Cauchy problem (3.1)–(3.2), the representation of Theorem 3.2 opens the door to a Monte Carlo approximation for the solution u. This is easily obtained by producing independent copies $(\xi_{t,x}^s)_{s\geq 1}$ copies of the r.v. $\xi_{t,x}$, so that

$$\hat{u}_S(t,x) := \frac{1}{S}\sum_{s=1}^S \xi_{t,x}^s \longrightarrow u(t,x), \quad \text{a.s. as } S \to \infty,$$

by the law of large numbers.

Moreover, under Assumption 3.4 with $p = 2$, we have $\mathbb{E}[\xi_{t,x}^2] < \infty$ by Theorem 3.5, and it follows from the central limit theorem that

$$\sqrt{S}(\hat{u}_S - u)(t,x) \longrightarrow \mathcal{N}\big(0, \mathbb{V}\mathrm{ar}[\xi_{t,x}]\big), \quad \text{in distribution.}$$

This shows that the error estimate of our Monte Carlo approximation is of the order $\sqrt{S}$, and is therefore dimension-free. Again, notice however that this solves only partly the curse of dimensionality, as the variance of the error is driven by $\mathbb{V}\mathrm{ar}[\xi_{t,x}]$ which may still be exploding in terms of the dimension parameter. Despite this, we emphasize that the present method is the first attempt in the literature to address the approximation of initial value problems in high dimension.

Finally, we discuss the complexity of our numerical method in Section 5.1.2 when we detail our algorithm.

4. Further nonlinear Cauchy problems. In this section, we consider the following semilinear Cauchy problem with polynomial nonlinearity in the pair (u, Du):

$$\sum_{n=1}^{N} a_n \partial_t^n u - \sum_{\alpha \in \mathbb{N}_M^d} b_\alpha D^\alpha u = \sum_{j \geq 0} q_j c_{j,0} u^{\ell_{j,0}} \prod_{h=1}^{H} (c_{j,h} \cdot Du)^{\ell_{j,h}} \quad \text{on } \mathbb{R}_+ \times \mathbb{R}^d, \tag{4.1}$$

$$\partial_t^{n-1} u(0, \cdot) = p_n f_n \quad \text{on } \mathbb{R}^d, n = 1, \dots, N, \tag{4.2}$$

where $(\ell_j)_{j \geq 0} \subset \mathbb{N}^{1+H}$ is a sequence of vector integers $\ell_j = (\ell_{j,0}, \dots, \ell_{j,H})$, and the nonlinearity is defined by means of an atomic probability measure $(q_j)_{j \geq 0}$, with $q_j > 0$ and $\sum_{j \geq 0} q_j = 1$, together with the functions

$$c_{j,0} \in C_b^0(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}) \quad \text{and} \quad c_{j,h} \in C_b^0(\mathbb{R}_+ \times \mathbb{R}^d, B_1(\mathbb{R}^d)), \quad j \geq 0, 1 \leq h \leq H. \tag{4.3}$$

Here, $B_1(\mathbb{R}^d)$ denotes the unit ball in $\mathbb{R}^d$ in the sense of the Euclidean norm. Notice that the above Cauchy problem covers the particular case of (3.1)–(3.2) by setting $\ell_{j,h} = 0$ for all $h = 1, \dots, H$.

By standard Fourier transform theory, the partial gradients of the Green functions $g = (g_n)_{1 \leq n \leq N}$, in the distribution sense, are given by

$$Dg_n(t, \cdot) = i \mathfrak{F}^{-1}\big(\xi \mapsto \xi \hat{g}_n(t, \xi)\big), \quad t \geq 0, n = 1, \dots, N.$$

Similar to the probability measures μ_n introduced in Assumption 2.2, we now assume that the distributions Dg_n can be represented by signed measures.

ASSUMPTION 4.1.

(i) For all $n = 1, \dots, N$ and $m = 1, \dots, d$, the distribution $\partial_{x_m} g_n(t, \cdot)$ may be represented by a signed measure with total variation $|\partial_{x_m} g_n(t, \cdot)|$.

(ii) The measure $|Dg_n(t, \cdot)| := \sum_m |\partial_{x_m} g_n(t, \cdot)|$ is absolutely continuous w.r.t. some probability measure $\mu_n^1(t, \cdot)$ so that we may define the density vector $\gamma_n^1 = (\gamma_{n,1}^1, \dots, \gamma_{n,d}^1)$ by

$$Dg_n(t, dx) = \gamma_n^1(t, x) \mu_n^1(t, dx), \quad t \geq 0, x \in \mathbb{R}^d.$$

Notice that although $\gamma_n^1(t, \cdot)$ is defined $d\mu_n^1(t, \cdot)$-a.s. this will be sufficient for our needs. Our starting point is the following "automatic differentiation property", which follows by direct differentiation of the Duhamel formula of Proposition 2.4.

PROPOSITION 4.2. *In addition to the conditions of Proposition 2.4, let Assumption 4.1 hold true. Then, the solution u of the linear Cauchy problem (2.1)–(2.2) is differentiable with respect to the space variable with*

$$Du(t, x) = \mathbb{E}\Big[\mathbf{1}_{\{\tau \geq t\}} \frac{\gamma_I^1(t, Z_t^{1,I})}{\bar{\rho}(t)} f_I\big(X_t^{1,I}\big) + \mathbf{1}_{\{\tau < t\}} \frac{\gamma_N^1(\tau, Z_\tau^{1,N})}{\rho(\tau)} F\big(t - \tau, X_\tau^{1,N}\big)\Big],$$

where (τ, I) are as in (2.11), and $X_\tau^{1,n} = x + Z_\tau^{1,n}$, with $Z_t^{1,n}$ distributed as $\mu_n^1(t, \cdot)$ independent of (I, τ).

EXAMPLE 4.3. Let us illustrate the last result on our main examples.

(i) Heat equation: we directly compute that $Dg_1(t, dx) = -\frac{x}{2t} g_1(t, dx)$.

(ii) Wave equation, $d = 1$: we have $Dg_2(t, dx) = \frac{1}{2}(\delta(x + t) - \delta(x - t))\, dx$, but $Dg_1 = D^2 g_2$ cannot be represented as a signed measure, thus violating Assumption 4.1. However, we may still handle the one-dimensional wave equation by reducing to the case $f_1 = 0$, see Section 5.2.

(iii) Wave equation, $d > 1$: Assumption 4.1 is not satisfied as Dg_1 and Dg_2 involve first-order derivative of the delta function supported on the light cone $\{(x,t) : t^2 - x^2 = 0\}$.

(iv) Beam equation: we have $Dg_2(t, dx) = G'(\frac{x}{\sqrt{t}})\, dx$.

In order to introduce the probabilistic representation of the solution of (4.1)–(4.2), we consider the branching mechanism defined in Section 3.1, where we modify the definition of the independent i.i.d. random variables $(J_k)_{k\in\mathbf{K}}$ and we introduce the types of particles $(\theta_k^t)_{k\in\mathbf{K}}$ as follows:

- $\mathbb{P}[J_k = \ell_j] = q_j$ for all $j \geq 0$, and we denote $\bar{J}_{k,-1} := 0$, $\bar{J}_{k,h} := \bar{J}_{k,h-1} + J_{k,h}$, $h = 0, \dots, H$;
- an arbitrary particle $k \in \overline{\mathcal{K}}_t \setminus \mathcal{K}_t$ branches at time T_k into $\bar{J}_{k,H}$ particles; the hth block of descendant particles are labelled

$$(k, j), \quad j = \bar{J}_{k,h-1} + 1, \dots, \bar{J}_{k,h}, \quad h = 0, \dots, H; \tag{4.4}$$

- we assign to the hth block of new particles (4.4) the types:

$$\theta^t_{(0)} := 0 \quad \text{and} \quad \theta^t_{(k,j)} := h \quad \text{for } k \in \overline{\mathcal{K}}_t \setminus \mathcal{K}_t,\ j = \bar{J}_{k,h-1} + 1, \dots, \bar{J}_{k,h}, h = 0, \dots, H. \tag{4.5}$$

Finally, we introduce a branching process which differs slightly from (3.6)–(3.7). Let

$$\hat{X}^0_0 = x \quad \text{and} \quad \hat{X}^k_{T^t_k} := \hat{X}^{k-}_{T^t_{k-}} + \hat{Z}^k_{T^t_k} \quad \text{for all } k \in \overline{\mathcal{K}}_t, \tag{4.6}$$

where $(Z^k_{T^t_k})_k$ are independent with distribution, conditional on (I^t_k, θ^t_k) and the lifetime $\Delta T^t_k := T^t_k - T^t_{k-}$ of the particle k, given by

$$\mathbb{P}\big[\hat{Z}^k_{T^t_k} \in dz | I^t_k, \theta^t_k, \Delta T^t_k\big] = \mathbf{1}_{\{\theta^t_k = 0\}} \mu_{I^t_k}(\Delta T^t_k, dz) + \mathbf{1}_{\{\theta^t_k \neq 0\}} \mu^1_{I^t_k}(\Delta T^t_k, dz). \tag{4.7}$$

The main goal of this section is to provide a representation of the solution of the Cauchy problem (4.1)–(4.2) by means of the random variable

$$\hat{\xi}_{t,x} := \prod_{k\in\mathcal{K}_t} \frac{\mathcal{W}_k}{\bar{\rho}(\Delta T^t_k)} \big[f_{I^t_k}(\hat{X}^k_t) - \mathbf{1}_{\{\theta^t_k \neq 0\}} f_{I^t_k}(\hat{X}^k_{T_{k-}})\big] \prod_{k\in\overline{\mathcal{K}}_t\setminus\mathcal{K}_t} \frac{\mathcal{W}_k}{\rho(\Delta T^t_k)} c_{J_k,0}\big(t - T^t_k, \hat{X}^k_{T^t_k}\big), \tag{4.8}$$

where the random weights $\mathcal{W}_k$ are given by

$$\mathcal{W}_k := \mathbf{1}_{\{\theta^t_k = 0\}} \gamma_{I^t_k}\big(\Delta T^t_k, \hat{Z}^k_{T^t_k}\big) + \mathbf{1}_{\{\theta^t_k \neq 0\}} c_{J_k,\theta^t_k}\big(t - T^t_{k-}, \hat{X}^{k-}_{T^t_{k-}}\big) \cdot \gamma^1_{I^t_k}\big(\Delta T^t_k, \hat{Z}^k_{T^t_k}\big), \quad k \in \overline{\mathcal{K}}_t.$$

Here $\mathcal{W}^k$ accounts for the differentiation which corresponds to introducing the corresponding weight γ^1 as in Proposition 4.2 in case of a nonzero type particle. We also emphasize the presence of an additional correction in the first product of (4.8) over the particles $k \in \mathcal{K}_t$, consisting in substracting $\mathbf{1}_{\{\theta^t_k \neq 0\}} f_{I^t_k}(\hat{X}^k_{T_{k-}})$, compare with (3.8) where no such correction is needed. Similar to [14], this correction is necessary in order to control for the potential explosion due to the automatic differentiation weight in the case where the particle has a nonzero type, and achieved by requiring the addition Lipschitz condition on the boundary data f, see Theorem 4.4 below.

We denote

$$\begin{aligned} \hat{r}_p := \sup_{\substack{0\leq t\leq T\\ 1\leq n\leq N}} \Big\{ \|f_n\|^p_\infty \int |\gamma_n(t,z)|^{p-1} |g_n|(t, dz) \Big\} \\ \vee \Big\{ \|\nabla f_n\|^p_\infty \int |z|^p |\gamma^1_n(t,z)|^{p-1} |Dg_n|(t, dz) \Big\} \end{aligned} \tag{4.9}$$

$$/\bar\rho(t)^{p-1},$$

$$\hat\alpha_p := \sup_{0\le t\le T} \frac{\{\int |\gamma_N(t,z)|^{p-1}|g_N(t,dz)|\} \vee \{\int |\gamma_N^1(t,z)|^{p-1}|Dg_N(t,dz)|\}}{\rho(t)^{p-1}}. \tag{4.10}$$

We shall provide in Proposition 4.8 below some sufficient conditions which guarantee that the coefficients $\hat\alpha_p$ and $\hat r_p$ are finite. The following result provides a probabilistic representation of the solution of the semilinear Cauchy (4.1) under a condition involving the above $\hat r_p$ and $\hat\alpha_p$. Recall the power series H_p introduced in (3.11).

THEOREM 4.4. *Let Assumptions 2.1, 2.2 and 4.1 hold true, and assume that f is bounded and Lipschitz. Let $p>1$, and ρ be a positive density function with support on $(0,\infty)$ such that the constants $\hat r_p$ and $\hat\alpha_p$ defined in (4.9)–(4.10) satisfy*

$$\hat r_p < R_p, \qquad \hat\alpha_p<\infty \quad \text{and} \quad \int_{\hat r_p}^{\hat s_p} \frac{ds}{H_p(s)} = \hat\alpha_p T \quad \text{for some } \hat s_p > \hat r_p \text{ and } T>0. \tag{4.11}$$

Then, $\hat\xi_{t,x}\in\mathbb{L}^p$ for $t\in[0,T]$, and $u(t,x):=\mathbb{E}[\hat\xi_{t,x}]$ is a bounded continuous solution of the semilinear Cauchy problem (4.1).

PROOF. Similar to the proof of Theorem 3.2, we directly estimate that

$$|\hat\xi_{t,x}|^p \le \prod_{k\in\mathcal{K}_t} \frac{1}{\bar\rho(\Delta T_k^t)^p}\big(\mathbf{1}_{\theta_k^t=0}\|f_{I_k^t}\|_\infty^p|\gamma_{I_k^t}(\Delta T_k^t,\hat Z_{T_k^t}^k)|^p + \mathbf{1}_{\theta_k^t\neq0}\|\nabla f_{I_k^t}\|_\infty^p|Z_{T_k^t}^k|^p|\mathcal{W}_k|^p\big)$$
$$\times \prod_{k\in\overline{\mathcal{K}}_t\setminus\mathcal{K}_t}\left|\frac{\mathcal{W}_k}{\rho(\Delta T_k^t)}\right|^p \|c_{J_k,0}\|_\infty^p.$$

By the independence of the ΔT_k^t's and the $Z_{T_k^t}^k$'s, this provides

$$\mathbb{E}|\hat\xi_{t,x}|^p \le \mathbb{E}\prod_{k\in\mathcal{K}_t}\frac{\hat r_p}{\bar\rho(\Delta T_k^t)}\prod_{k\in\overline{\mathcal{K}}_t\setminus\mathcal{K}_t}\frac{\hat\alpha_p\|c_{J_k,0}\|_\infty^p}{\rho(\Delta T_k^t)}.$$

The required result follows by the same line of argument as in the proof of Theorem 3.2. □

In the rest of this section, we provide sufficient conditions for $\hat r_p$ and $\hat\alpha_p$ to be finite, as required in (4.11). In preparation for this, we need some estimates on the Green functions g_n, $n=1,\dots,N$. Recall that $\{\alpha\in\mathbb{N}_M^d : |\alpha|=M \text{ and } b_\alpha\neq0\}\neq\varnothing$, and define *the principal symbol*

$$b_M(\xi) := i^M\sum_{|\alpha|=M} b_\alpha\xi^\alpha = e^{i\pi\eta_M(\xi)}|b_M(\xi)|$$
$$\text{where } \eta_M(\xi) := \frac{1+M-\operatorname{sg}\{b_M(\xi)\}}{2};\ \xi\in\mathbb{R}^d.$$

LEMMA 4.5. *For all $\xi\in\mathbb{R}^d$, the matrix $B(\varepsilon^{-1}\xi)$ has N simple eigenvalues $\lambda_n(\varepsilon^{-1}\xi)$, $n=1,\dots,N$, for sufficiently small $\varepsilon>0$, with asymptotics*

$$\lim_{\varepsilon\searrow0}\varepsilon^{\frac{1}{\sigma}}\lambda_n(\varepsilon^{-1}\xi) = \lambda_n^0(\xi) := |b_M(\xi)|^{\frac{1}{N}}e^{\frac{i\pi}{N}(\eta_M(\xi)+2n)}, \quad \text{where } \sigma:=\frac{N}{M}.$$

PROOF. Recall from Remark 2.5 that the spectrum of the matrix $B(\varepsilon^{-1}\xi)$ consists of the solutions of the characteristic polynomial $\sum_{n=1}^{N} a_n\lambda^n = b(\varepsilon^{-1}\xi)$.

As $\{\alpha \in \mathbb{N}_M^d : |\alpha| = M \text{ and } b_\alpha \neq 0\} \neq \varnothing$, we see that $\varepsilon^M b(\varepsilon^{-1}\xi) \longrightarrow b_M(\xi)$. Then, denoting by λ_ε an arbitrary solution of the last characteristic polynomial, it follows that $\varepsilon^M \sum_{n=1}^{N} a_n\lambda_\varepsilon^n \longrightarrow b_M(\xi)$, and we deduce that $(\lambda_\varepsilon)_{\varepsilon>0}$ has no finite accumulation point. Together with the normalization $a_N = 1$, this in turn implies that $\sum_{n=1}^{N} a_n\lambda_\varepsilon^n \sim \lambda_\varepsilon^N$ as $\varepsilon \searrow 0$, and therefore

$$\lim_{\varepsilon\searrow 0} \varepsilon^M \lambda_\varepsilon^N = b_M(\xi) = |b_M(\xi)| e^{i\pi\eta_M(\xi)} := \lambda_N^0(\xi)^N,$$

which can be written equivalently as $\lim_{\varepsilon\searrow 0}(\varepsilon^{\frac{M}{N}} \frac{\lambda_\varepsilon}{\lambda_N^0(\xi)})^N = 1$. Hence, the limiting spectrum of the matrix $B(\varepsilon^{-1}\xi)$ consists of N simple eigenvalues expressed in terms of the unit roots:

$$\lim_{\varepsilon\searrow 0} \varepsilon^{\frac{M}{N}} \lambda_n(\varepsilon^{-1}\xi) = \lambda_N^0(\xi) e^{\frac{2ni\pi}{N}} = \lambda_n^0(\xi), \quad n = 1, \dots, N. \qquad \square$$

The last result shows that the asymptotics of $B(\varepsilon^{-1}\xi)$ are in the context of Remark 2.5. Then, it follows that the space Fourier transforms of the Green functions are given by (2.10). Then, by convenient scaling, we expect the corresponding limits

$$\hat{g}_n^0 := (2\pi)^{-\frac{d}{2}} \Lambda_n^0 \sum_{j=1}^{N} \frac{e^{\lambda_j^0}}{\prod_{\ell\neq j}(\lambda_\ell^0 - \lambda_j^0)}, \quad n = 1, \dots, N, \tag{4.12}$$

with

$$\Lambda_N^0(\xi) := 1, \qquad \Lambda_n^0(\xi) := (-1)^{n-1} \sum_{\substack{1\le \ell_1 \le \cdots \le \ell_{N-n} \le N \\ \ell_1, \dots, \ell_{N-n} \neq n}} (\lambda_{\ell_1}^0 \cdots \lambda_{\ell_{N-n}}^0)(\xi) \quad \text{for } n < N.$$

In order to justify this convergence, we need the following additional condition.

ASSUMPTION 4.6. For all $\varphi \in \mathcal{S}$, the family $\{t^{1-n}\hat{g}_n(t, t^{-\sigma}\cdot)\varphi, t \in (0, \varepsilon]\}$ is uniformly integrable in $\mathbb{R}^d$, for some $\varepsilon > 0$.

We finally introduce the corresponding scaled Green functions

$$g_n^\sigma(t, \cdot) := g_n(t, \cdot) \circ (t^\sigma X)^{-1}, \quad t \in [0, T], n = 1, \dots, N,$$

where X the canonical map on $\mathbb{R}^d$, that is, $X(x) = x$ for all $x \in \mathbb{R}^d$. In particular, in the case where g_n can be represented by a function, $g_n^\sigma(t, x) = t^{\sigma d} g_n(t, t^\sigma x)$.

ASSUMPTION 4.7.

(i) The functions $t \longmapsto \int |z|^p |\gamma_n^1(t, z)|^p \mu_n^1(t, dz)$, $n = 1, \dots, N$ and $t \longmapsto \int |\gamma_N^1(t, z)|^p \mu_N^1(t, dz)$ are continuous on $(0, T]$;

(ii) $\gamma_n^1(t, t^\sigma z) = O(1)$ near the origin $t = 0, n = 1, \dots, N$;

(iii) The following families are uniformly integrable for some $\varepsilon > 0$:

$$\{z \longmapsto |z|^p |\gamma_n^1(t, t^\sigma z)|^{p-1} |Dg_n^\sigma|(t, dz)\}_{t\le\varepsilon}, \quad 1 \le n \le N \quad \text{and}$$

$$\{z \longmapsto |\gamma_N^1(t, t^\sigma z)|^{p-1} |Dg_N^\sigma|(t, dz)\}_{t\le\varepsilon}.$$

PROPOSITION 4.8. *Let Assumptions 2.1, 2.2, 4.1, 4.6 and 4.7 hold true. Assume further that the density $\rho \in C^0(\mathbb{R}_+)$, strictly positive on $(0, \infty)$, and $\lim_{t\searrow 0} t^{N-1+\sigma p}\rho(t)^{1-p} < \infty$. Then, for any bounded Lipschitz function f, we have $\hat{r}_p + \hat{\alpha}_p < \infty$.*

To prepare the proof of this result, we isolate the following pointwise convergence.

LEMMA 4.9. *The space Fourier Green functions $\hat{g}$ satisfy the short time asymptotics*

$$t^{1-n}\hat{g}_n(t,t^{-\sigma}\xi)\longrightarrow \hat{g}_n^0(\xi)\quad \text{as } t\searrow 0, \text{for all } \xi\in\mathbb{R}^d.$$

If in addition Assumption 4.6 holds true, then the last convergence holds in $\mathcal{S}'$, and the short time asymptotics of the scaled Green functions are given by

$$t^{1-n}g_n^\sigma(t,\cdot)\longrightarrow g_n^0:=\mathfrak{F}^{-1}\hat{g}_n^0\quad \text{as } t\searrow 0, \text{in } \mathcal{S}', n=1,\dots,N.$$

PROOF. First, the pointwise convergence of $\hat{G}_n(t,\cdot):=t^{1-n}\hat{g}_n(t,t^{-\sigma}\cdot)$ towards $\hat{g}_n^0$, as $t\searrow 0$, follows from direct application of Lemma 4.5 together with the observations reported in Remark 2.5. The uniform integrability condition of Assumption 4.6 guarantees that $\langle\hat{G}_n(t,\cdot),\varphi\rangle\longrightarrow\langle\hat{g}_n^0(t,\cdot),\varphi\rangle$ for all $\varphi\in\mathcal{S}$, that is, $\hat{G}_n(t,\cdot)\longrightarrow\hat{g}_n^0$ as $t\searrow 0$ in $\mathcal{S}'$. This in turn implies the convergence of the corresponding Fourier inverse $\mathfrak{F}^{-1}\hat{G}_n(t,\cdot)$ towards $g_n^0:=\mathfrak{F}^{-1}\hat{g}_n^0$ as $t\searrow 0$ in $\mathcal{S}'$. It remains to relate the distribution $\mathfrak{F}^{-1}\hat{G}_n(t,\cdot)$ to the Green function g_n. To see this, we use the properties of the Fourier transform in $\mathcal{S}'$ as defined by means of arbitrary test functions $\varphi\in\mathcal{S}$ as follows:

$$\langle\mathfrak{F}^{-1}\hat{G}_n,\varphi\rangle=\langle\hat{G}_n,\mathfrak{F}^{-1}\varphi\rangle=t^{\sigma d}\langle\hat{g}_n(t,\cdot),(\mathfrak{F}^{-1}\varphi)(t^\sigma\cdot)\rangle=t^{\sigma d}\langle g_n(t,\cdot),\mathfrak{F}((\mathfrak{F}^{-1}\varphi)(t^\sigma\cdot))\rangle.$$

We finally observe by direct calculation that $\mathfrak{F}((\mathfrak{F}^{-1}\varphi)(\lambda\cdot))=\lambda^{-d}\varphi(\lambda^{-1}\cdot)$ for all $\varphi\in\mathcal{S}$ and all constant $\lambda\in\mathbb{R}$. Then

$$\langle\mathfrak{F}^{-1}\hat{G}_n,\varphi\rangle=\langle g_n(t,\cdot),\varphi(t^{-\sigma}\cdot)\rangle=\langle g_n^\sigma(t,\cdot),\varphi\rangle.$$

By the arbitrariness of $\varphi\in\mathcal{S}$, this provides that $\mathfrak{F}^{-1}\hat{G}_n(t,\cdot)=g_n^\sigma(t,\cdot)$, thus completing the proof. □

PROOF OF PROPOSITION 4.11. By Assumption 4.1(i), together with the fact that $\bar{\rho}(T)\le\rho\le 1$ on $[0,T]$, we only need to justify that

$$\sup_{\substack{0\le t\le T\\1\le n\le N}}\int|z|^p|\gamma_n^1(t,z)|^{p-1}|\partial_x g_n|(t,dz))$$
$$+\sup_{0\le t\le T}\rho(t)^{1-p}\int|\gamma_N^1(t,z)|^{p-1}|\partial_x g_N|(t,dz))<\infty.$$

Assumption 4.7(i) ensures that the functions inside the last suprema are continuous on $(0,\infty)$. Then, in order to prove that $\hat{r}_p$ and $\hat{\alpha}_p$ are finite, it suffices to verify that the functions inside the last supremum are bounded near $t=0$.

To see this, we first observe that the convergence of the scaled Green functions in $\mathcal{S}'$ in Lemma 4.9 implies that

$$t^{1-n+\sigma|\alpha|}D^\alpha g_n^\sigma(t,\cdot)\longrightarrow D^\alpha g_n^0\quad \text{as } t\searrow 0, \text{in } \mathcal{S}', \text{ for all } \alpha\in\mathbb{N}^d, n=1,\dots,N. \tag{4.13}$$

Then, we compute by a direct change of variables that

$$\begin{aligned}\int|z|^p|\gamma_n^1(t,z)|^{p-1}|\partial_x g_n|(t,dz)&=\int|t^\sigma z|^p|\gamma_n^1(t,t^\sigma z)|^{p-1}|\partial_x g_n^\sigma|(t,dz)\\&=t^{n-1+\sigma p}\int|z|^p t^{1-n}|\gamma_n^1(t,t^\sigma z)|^{p-1}|\partial_x g_n^\sigma|(t,dz)\\&=\mathrm{O}(t^{n-1+\sigma p})\quad\text{near the origin},\end{aligned}$$

by Assumption 4.7(ii)–(iii), together with (4.13). This implies that the limit is zero as $n \geq 1$ and $p > 1$. Similarly,

$$\begin{aligned}
&\rho(t)^{1-p} \int |\gamma_N^1(t,z)|^{p-1} |\partial_x g_N|(t,dz) \\
&\quad = t^{N-1+\sigma p} \rho(t)^{1-p} \int t^{1-N} |\gamma_N^1(t,t^\sigma z)|^{p-1} |\partial_x g_N^\sigma|(t,dz) \\
&\quad = \mathrm{O}\big(t^{N-1+\sigma p} \rho(t)^{1-p}\big) \quad \text{near the origin,}
\end{aligned}$$

again by Assumption 4.7(ii)–(iii). □

5. Numerical examples.

5.1. *Wave semilinear PDE.* We consider the nonlinear Klein–Gordon wave equation in $\mathbb{R}^d$ for $1 \leq d \leq 3$:

$$(\partial_{tt}^2 - \Delta)u + u^3 + u^2 = 0, \qquad u(0,x) = f_1(x), \qquad \partial_t u(0,x) = f_2(x). \tag{5.1}$$

To the best of our knowledge, the current literature only considers approximation of the solution by deterministic numerical schemes, see, for example, [9]. Due to the curse of dimensionality, mainly $d = 1$ and $d = 2$ have been considered. To illustrate the efficiency of our algorithm, we solve this equation in $d = 1, 2$ and 3. In our numerical experiments, we take the initial conditions

$$f_1(x) := -\frac{12}{9 + 2(\sum_{i=1}^d x_i)^2} \quad \text{and} \quad f_2(x) := -\frac{48\sqrt{d+1}(\sum_{i=1}^d x_i)}{(2(\sum_{i=1}^d x_i)^2 + 9)^2},$$

for which the explicit solution is $u(t,x) = -\frac{12}{9+2(\sqrt{d+1}t - \sum_{i=1}^d x_i)^2}$.

We choose $\rho(t) = \beta e^{-\beta t}$ and $\bar{\rho}(t) = e^{-\beta t}$. For convenience, we set $u(t,x) := U(t,x) + f_1(x)$, and we compute that U satisfies the nonhomogeneous nonlinear wave PDE

$$\begin{aligned}
&(\partial_{tt}^2 - \Delta)U + U^3 + (3f_1 + 1)U^2 + (3f_1^2 + 2f_1)U + (f_1^3 + f_1^2 - \Delta f_1) = 0, \\
&U(0,x) = 0, \qquad \partial_t U(0,x) = f_2(x).
\end{aligned}$$

Note that as $U(0,x) = 0$, we do not need to simulate our branching particles according to the distribution g_1 but only g_2. This was our motivation for introducing the function U.

We directly compute that $r_p < (\frac{1+d}{6})^p t^p$ and $\alpha_p < t^p \beta^{1-p}$. Furthermore, $|(3f_1 + 1)|_\infty = 1$, $|(3f_1^2 + 2f_1)|_\infty = 8/3$ and $|(f_1^3 + f_1^2 - \Delta f_1)|_\infty < 1$. This implies that the assumptions in Theorem 3.5 are satisfied.

The Monte Carlo approximation of $U(t, X_0)$ (and therefore $u(t, X_0) = U(t, X_0) + f_1(X_0)$) can be described by the following meta-algorithm.

5.1.1. *Meta-algorithm.*

1. Start at $t_0 = 0$ at the position X_0 and initialize a weight $\mathcal{W} := 1$.
2. Simulate an exponential r.v. τ with (arbitrarily) constant intensity β (i.e., default time) and simulate the particle at the new position $X_\tau = X_0 + Z_\tau$ at τ. More precisely, we draw uniform variables $(U_i)_{i=1,2}$ on $[0,1]$ and set

$$\begin{aligned}
&Z_\tau^1 = \tau(2U_1 - 1), \quad d = 1, \\
&Z_\tau^1 = \sqrt{1 - U_1^2}\cos(2\pi U_2)\tau, \qquad Z_\tau^2 = \sqrt{1 - U_1^2}\sin(2\pi U_2)\tau, \quad d = 2, \\
&Z_\tau^1 = \cos(2\pi U_1)\cos(2\pi U_2)\tau, \qquad Z_\tau^2 = \sin(2\pi U_1)\cos(2\pi U_2)\tau, \\
&Z_\tau^3 = \sin(2\pi U_2)\tau, \quad d = 3.
\end{aligned}$$

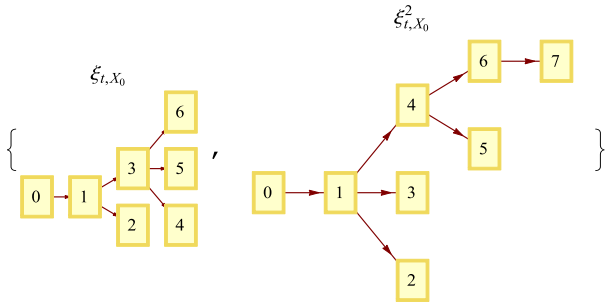

FIG. 1. *Galton–Watson trees associated to* $\xi^{(1)}_{t,X_0}$ *and* $\xi^{(2)}_{t,X_0}$.

At $t_1 := \tau < t$, the particle dies and we create 0, 1, 2 or 3 descendants with probability $p := 1/4$. We then multiply the weight $\mathcal{W}$ by the mass τ and according to the number of descendants, we update again the weight $\mathcal{W}$ by

$$\mathcal{W} := \begin{cases} \mathcal{W} \times (-p^{-1})(f_1^3 + f_1^2 - \Delta f_1)(X_\tau)\beta^{-1}e^{\beta(t_1-t_0)} & \text{if 0 descendant,} \\ \mathcal{W} \times (-p^{-1})(3f_1^2 + 2f_1)(X_\tau)\beta^{-1}e^{\beta(t_1-t_0)} & \text{if 1 descendant,} \\ \mathcal{W} \times (-p^{-1})(3f_1 + 1)(X_\tau)\beta^{-1}e^{\beta(t_1-t_0)} & \text{if 2 descendants,} \\ \mathcal{W} \times (-p^{-1})\beta^{-1}e^{\beta(t_1-t_0)} & \text{if 3 descendants,} \end{cases}$$

where 0 descendant means that the particle dies.

3. For each descendant, we apply independently Steps 2 and 3 until the default time—say τ_n—is greater than the maturity t. In this case, we multiply $\mathcal{W}$ by

$$\mathcal{W} := \mathcal{W} \times e^{\beta(t-\tau_{n-1})}.$$

4. Finally, for all particles alive at time t (with locations $(X_t^k)_{k\in\mathcal{K}_t}$), compute

$$\mathcal{W} \prod_{k\in\mathcal{K}_t} f_2(X_t^k),$$

and average the result using M Monte Carlo paths.

5.1.2. *Complexity.* Our algorithm boils down to the simulation of uniform random variables in order to generate the Z's, the exponential r.v. τ_i and the number of descendants at each τ_i. As our Monte Carlo estimator is square-integrable, the rate of convergence is given by the central limit theorem. In our numerical experiment, we compute the standard deviation of our Monte Carlo estimator in order to check the convergence. We should emphasize again that we do need to fine-tune some numerical parameters and no approximations are needed.

We emphasize again that the present method is the first attempt in the literature to address the problem of numerical approximation of the solution of a general initial value problem. Therefore, we have no benchmark for the comparison of our numerical results, and we restrict the following section to the illustration of the efficiency of our approximation method.

5.1.3. *Numerics.* In Figure 1, we have plotted two examples of Galton–Watson trees corresponding to the functionals

$$\xi^{(1)}_{t,X_0} := \frac{1}{\beta^2 p^2}(3f_1(X_{\tau_1}) + 1)e^{\beta\tau_1}e^{\beta(\tau_3-\tau_1)}e^{3\beta(t-\tau_3)}e^{\beta(t-\tau_1)}f_2(X_t^6)f_2(X_t^5)f_2(X_t^4)f_2(X_t^2),$$

$$\begin{aligned}\xi^{(2)}_{t,X_0} := {} & -\frac{1}{\beta^3 p^3}(3f_1(X_{\tau_4}) + 1)(3f_1^2 + 2f_1)(X_t^6) \\ & \times e^{\beta\tau_1}e^{\beta(\tau_4-\tau_1)}e^{\beta(\tau_6-\tau_4)}f_2(X_t^7)f_2(X_t^5)f_2(X_t^3)f_2(X_t^2) \\ & \times e^{\beta(t-\tau_6)}e^{\beta(t-\tau_4)}e^{2\beta(t-\tau_1)}.\end{aligned}$$

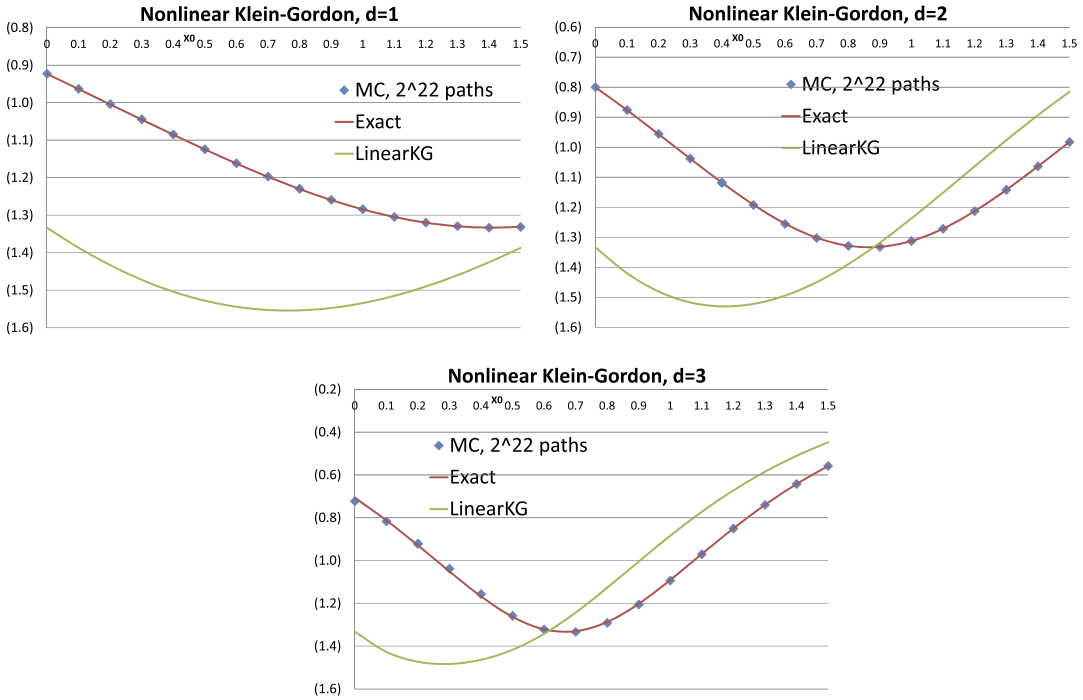

FIG. 2. *Numerical solutions of the nonlinear Klein–Gordon PDE* (5.1) *for* $d = 1, 2, 3$ *as a function of* x_0 (*all the coordinates in* $\mathbb{R}^d$ *are equal to* x_0) *and* $t = 1$.

In the present very simple examples, in order to alleviate the figure, we have used simpler notation to label the branching particles than those in Section 3.1.

Note that by construction, the result is independent of β when M is large. There is an optimal choice of β that minimizes the variance of our Monte Carlo estimator. We have chosen $\beta = 1$ in our numerical experiments and $M = 2^{22}$ for which the standard deviation of our estimator is less than 0.01. Below, we have plotted our numerical result for $u(t, X_0)$ again our analytical solution for $d = 1, 2, 3$ as a function of $x_0 \in [0, 1.5]$ (all the coordinates in $\mathbb{R}^d$ are equal to x_0) and $t = 1$, see Figure 2. We obtain a perfect match. In order to see that our numerical solutions captures perfectly the additional nonlinearity $u^3 + u^2$, we have also shown for completeness the (analytical) solution of the linear wave equation (denoted "LinearKG"):

$$(\partial_{tt} - \Delta)u = 0, \qquad u(0, x) = f_1(x), \qquad \partial_t u(0, x) = f_2(x). \tag{5.2}$$

This also highlights that the nonlinear PDE differs from the linear PDE that is simpler to solve.

5.2. *Yang–Mills PDE: A toy model.* Here we consider the semilinear wave equation in $\mathbb{R}$:

$$\partial_{tt} u - \Delta u + u^3 + u\partial_x u = 0, \qquad u(0, x) = f_1(x), \qquad \partial_t u(0, x) = f_2(x). \tag{5.3}$$

We take the initial conditions

$$f_1(x) = -\frac{1}{1 - x}, \qquad f_2(x) = \frac{1}{(1 - x)^2},$$

for which the explicit solution is $u(t, x) = -\frac{1}{1+t-x}$. Assumption (2.2) is satisfied for g_2 only for $d = 1$, this is why we restrict to this case. For $d > 1$, ∇g_2 involves a derivative of delta

function supported on the lightcone. Notice that the singularity of f_1 and f_2 are not seen by our numerical algorithm because of our choice of the initial position x_0 and the finite speed property of the wave equation: The singularity $x = 1$ is not attainable.

This example can be interpreted as a scalar version of the Yang–Mills hyperbolic system. Indeed, the Lagrangian associated to the Yang–Mills theory on $\mathbb{R}^{1+d}$ is

$$\int F^a_{\mu\nu} F^{\mu\nu a}\, d^4x, \qquad F^{\mu\nu a} := \eta^{\mu\alpha}\eta^{\nu\beta} F^a_{\alpha\beta}$$

with $\eta_{00} = -1$, $\eta_{ii} = 1$, $i = 1, \dots d$ and 0 otherwise. Here we use the convention of implicit summation for repeated indices (a, μ, ν). The curvature is $F^a_{\mu\nu} := \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - f^a_{bc} A^b_\mu A^c_\nu$. The Euler–Lagrange equations (written in the gauge frame $\partial_\mu A^{\mu,a} = 0$) give the system of hyperbolic PDE

$$\partial_\mu F^{\mu\nu a} + f^a_{bc} A^b_\mu F^{\mu\nu c} = 0.$$

Note that the initial boundary conditions at $t = 0$, $A^a_\nu(t = 0, x)$ and $\partial_t A^a_\nu(t = 0, x)$ need to satisfy a constraint condition in order to preserve the gauge condition for all t. Using the expression for the curvature $F^a_{\mu\nu}$, we get that $A^{\mu a} := \eta^{\mu\alpha} A^a_\alpha$ is solution of a system of hyperbolic PDE of the form

$$\Box A^{\nu a} - f^a_{bc} A^{\mu b} \partial_\mu A^{\nu c} + f^a_{\mathrm{bc}} A^b_\mu \big(\partial^\mu A^{\nu c} - \partial^\nu A^{\mu c} - f^c_{de} A^{\mu d} A^{\nu e}\big) = 0$$

with $\Box := \partial_{tt} - \Delta$. This can be schematically written as

$$\Box A - f A \partial A - f^2 A^3 = 0$$

hence our PDE (5.3).

Note that our branching Monte Carlo algorithm can be easily adapted to solve a system of semilinear PDEs. All we need to do is to index particles with a type (μa) corresponding to a coordinate of the solution, see an example with the complex Gross–Pitaevskii equation in Section 5.4.

As in the previous section, we set $u(t, x) = f_1(x) + U(t, x)$ for which U satisfies the nonhomogeneous hyperbolic PDE

$$\partial_{tt} U - \Delta U + U^3 + 3 f_1 U^2 + U \sum_{i=1}^d \partial_{x_i} U + \left(3 f_1^2(x) + \sum_i^d \partial_{x_i} f_1\right) U + f_1 \sum_i^d \partial_{x_i} U$$
$$+ \left(f_1^3 + \sum_i^d f_1 \partial_{x_i} f_1 - \Delta f_1\right) = 0, \quad U(0, x) = 0, \partial_t u(0, x) = f_2(x).$$

The Monte Carlo approximation of $U(t, X_0)$ (and therefore $u(t, X_0) = U(t, X_0) + f_1(X_0)$) can be described by the following meta-algorithm:

1. Start with a type (0) particle at $t_0 = 0$ at position X_0, and initialize a weight $\mathcal{W} := 1$;
2. Simulate an exponential r.v. τ with (arbitrarily) constant intensity β, and simulate the particle at the new position $X_\tau = X_0 + Z_\tau$ at τ. More precisely, we draw a uniform variable (U) on $[0, 1]$ and set $Z_\tau = \tau(2U - 1)$;
3. At $t_1 := \tau < t$, the particle dies and we create 0, 1, 2 or 3 descendants with probability $p := 1/6$; in case of 1 descendant, the type assigned is (0) or (1); similarly, in case of 2 descendants, their types can be both (0), both (1), or one of each type; in case of 3 descendants, they are all of type (0); we then multiply the weight $\mathcal{W}$ by the mass τ (see Remark 2.3) if the

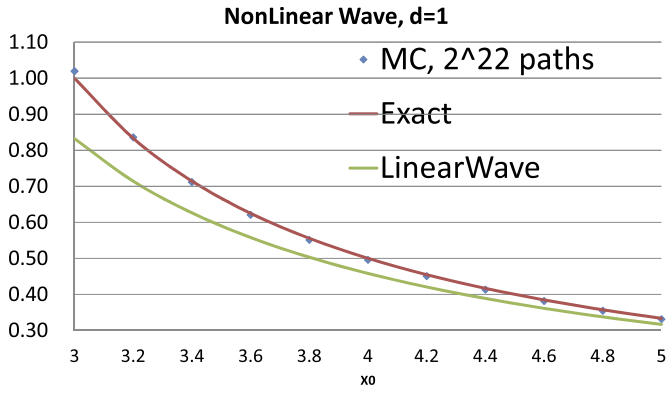

FIG. 3. *Numerical solution of the* 1*-d semilinear wave PDE* (5.3) *as a function of* x_0 *and* $t = 1$.

type of the particle is (0) and according to the number of descendants and type, we update also the weight $\mathcal{W}$ by

$$\mathcal{W} := \begin{cases} \mathcal{W} - p^{-1}\big(f_1^3 + f_1\partial_x f_1 - \Delta f_1\big)(X_\tau)\beta^{-1}e^{\beta(t_1-t_0)}, & \text{0 descendant,} \\ \mathcal{W} \times -p^{-1}\big(3f_1^2(x) + \partial_x f_1\big)(X_\tau)\beta^{-1}e^{\beta(t_1-t_0)}, & \text{1 descendant, type (0),} \\ \mathcal{W} \times -p^{-1} f_1(X_\tau)\beta^{-1}e^{\beta(t_1-t_0)}, & \text{1 descendant, type (1),} \\ \mathcal{W} \times (-3/p) f_1(X_\tau)\beta^{-1}e^{\beta(t_1-t_0)}, & \text{2 descendants, type (0),} \\ \mathcal{W} \times -p^{-1}\beta^{-1}e^{\beta(t_1-t_0)}, & \text{2 descendants, type (0) and (1),} \\ \mathcal{W} \times -p^{-1}\beta^{-1}e^{\beta(t_1-t_0)}, & \text{3 descendants,} \end{cases}$$

where 0 descendant means that the particle dies. The particle of type (1) is then simulated on the light cone, meaning that $Z_\tau = \tau$ or $Z_\tau = -\tau$ with probability $1/2$. In the last case, the weight is multiplied by -1.

4. For each particle, we apply independently Steps 2 and 3 until the default time—say τ_n—is greater than the maturity t. In this case, we multiply $\mathcal{W}$ by

$$\mathcal{W} := \mathcal{W} \times e^{\beta(t-\tau_{n-1})}.$$

5. Finally, compute for all particles in $\mathcal{K}_t$

$$\mathcal{W} \prod_{k\in\mathcal{K}_t} \big(f_2(X_t^k)1_{\text{type=0}} - f_2(X_t^k - \Delta X_t^k)1_{\text{type=1}}\big),$$

where $\Delta X_t^k = \pm(t - \tau_{n-1})$ with probability $1/2$ and average the result using M Monte Carlo paths.

Below, we have plotted our numerical result for $u(t, X_0)$ again our analytical solution for $d = 1$ as a function of $x_0 \in [3, 5]$ and $t = 1$, see Figure 3. We obtain a perfect match. For completeness, we have also shown the (analytical) solution of the linear wave equation (denoted "Linearwave")

$$(\partial_{tt} - \Delta)u = 0, \qquad u(0, x) = f_1(x), \qquad \partial_t u(0, x) = f_2(x) \tag{5.4}$$

to show that our numerical solutions capture perfectly the additional nonlinearity $u^3 + u\partial_x u$. Note that as $X_t \in [x_0 - t, x_0 + t]$, the singularity of f_2 at $x = 1$ is not relevant.

5.3. *Nonlinear beam PDE.* We consider the nonlinear beam equation:

$$\partial_t^2 u + \partial_x^4 u + u^2 + h(t, x) = 0, \qquad u(0, x) = \tanh(x), \qquad \partial_t u(0, x) = \cosh(x)^{-2}, \tag{5.5}$$

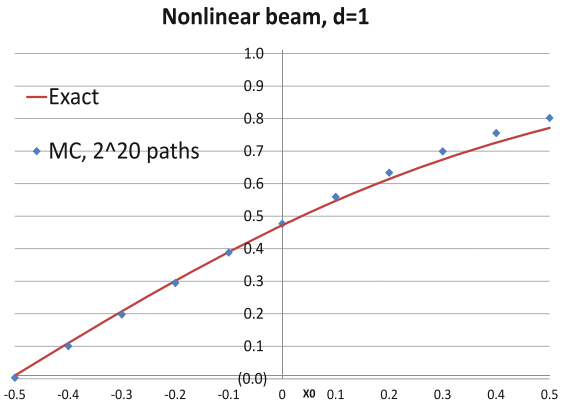

FIG. 4. *Numerical solution of nonlinear beam PDE (5.5) for $d = 1$ as a function of x_0 and $t = 0.5$.*

for which the explicit solution is $u(t,x) = \tanh(x+t)$ for a suitable choice of h. Here we follow the same discussion as in Section 5.1. Below, we have plotted our numerical result for $u(t, X_0)$ again our exact solution for $d = 1$ as a function of $x_0 \in [-0.5, 0.5]$ and $t = 0.5$, see Figure 4. Here we obtain a small error due to the fact that the one-dimensional density $\frac{|G(z)|}{\int_{\mathbb{R}} |G(z)|\,dz}$ with G given by (2.13) has been computed on an interval $[-10, 10]$ and stored for computational purpose.

5.4. *Gross–Pitaevskii PDE.* The Gross–Pitaevskii PDE reads

$$i\,\partial_t u(t,x) = -\frac{1}{2}\Delta u(t,x) + h\big|u(t,x)\big|^2 u(t,x), \quad x \in \mathbb{R}^d \tag{5.6}$$

with h a constant. This equation describes a Bose–Einstein condensate at zero or very low temperature. This has been recently solved using a time-splitting spectral method [4]. This deterministic method suffers from the curse of dimensionality and requires suitable mesh size controls. Below, we present our Monte Carlo algorithm. We emphasize again that our algorithm is partially immune to the dimension, as the error is controlled by the standard central limit theorem (with variance possibly dimension-dependent), so that convergence to the true solution is guaranteed provided that the standard deviation of our Monte Carlo estimate converges to zero.

We set $h := -1$ and $f_1(x) := \frac{\sqrt{d}}{\cosh(\sum_{i=1}^d x_i)}$ for which the explicit solution is

$$u(t,x) = e^{\frac{idt}{2}} \frac{\sqrt{d}}{\cosh(\sum_{i=1}^d x_i)}.$$

PDE (5.6) can be written as a two-dimensional PDE system with polynomial nonlinearity:

$$i\,\partial_t u(t,x) = -\frac{1}{2}\Delta u + guu^*u,$$

$$-i\,\partial_t u^*(t,x) = -\frac{1}{2}\Delta u^* + guu^*u^*.$$

From Example 2.11, $Z_\tau = e^{\frac{i\pi}{4}}\sqrt{\tau}Z \in \mathbb{C}^d$ and $Z_\tau^* = e^{-\frac{i\pi}{4}}\sqrt{\tau}Z \in \mathbb{C}^d$ with $Z \in N(0,d)$.

Meta-algorithm.

1. Start at $t_0 = 0$ at the position X_0 with a particle of type 0 and initialize a *complex* weight $\mathcal{W} := 1$.

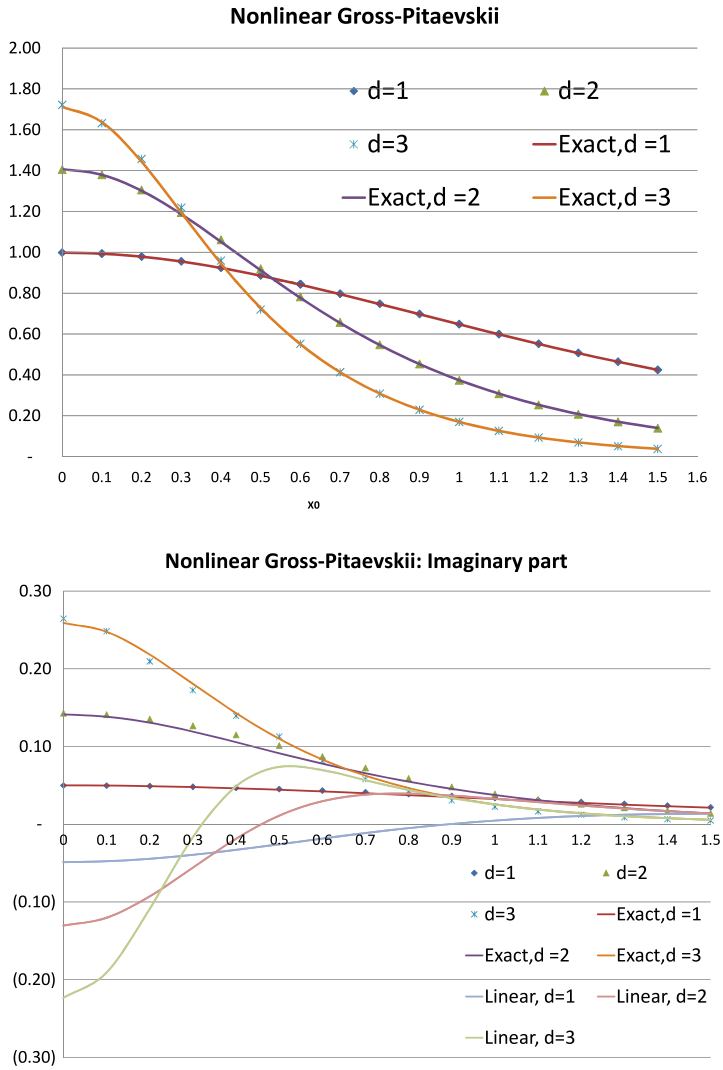

FIG. 5. *Numerical solutions for* $\mathrm{Re}(u(t,x_0))$ *(left) and* $\mathrm{Im}(u(t,x_0))$ *(right) of the Gross–Pitaevskii PDE* (5.6) *for* $d=1,2,3$ *as a function of* x_0 *(all the coordinates in* $\mathbb{R}^d$ *are equal to* x_0*) and* $t=0.1$.

2. Simulate an exponential r.v. τ with (arbitrarily) constant intensity β and simulate the particle at the new (complex) position $X_\tau = X_0 + e^{\frac{i\pi}{4}}\sqrt{\tau}Z$ (resp. $X_\tau = X_0 + e^{-\frac{i\pi}{4}}\sqrt{\tau}Z$) at τ if the particle is of type 0 (resp. of type 1) with Z a d-dimensional standard (real) Gaussian variable.

At $t_1 := \tau < t$, the particle dies and we create 2 descendants of type 0 and one of type 1 (resp. 2 descendants of type 1 and one of type 0) if the particle is of type 0 (resp. type 1). We then multiply the weight $\mathcal{W}$ by

$$\mathcal{W} := \begin{cases} \mathcal{W} \times (i)\beta^{-1} e^{\beta(t_1 - t_0)} & \text{type 0,} \\ \mathcal{W} \times (-i)\beta^{-1} e^{\beta(t_1 - t_0)} & \text{type 1.} \end{cases}$$

3. For each descendant, we apply independently Steps 2 and 3 until the default time—say τ_n—is greater than the maturity t. In this case, we multiply $\mathcal{W}$ by

$$\mathcal{W} := \mathcal{W} \times e^{\beta(t-\tau_{n-1})}.$$

4. Finally, for all particles alive at time t (with locations $(X_t^k)_{k\in\mathcal{K}_t}$), compute

$$\mathcal{W} \prod_{k\in\mathcal{K}_t:\text{Type}=0} f_1(X_t^k) \prod_{k\in\mathcal{K}_t:\text{Type}=1} f_1(X_t^k)^*,$$

and average the result using M Monte Carlo paths.

We have chosen $\beta = 1$ in our numerical experiments and $M = 2^{22}$ for which the standard deviation of our estimator is less than 0.01. Below, we have plotted our numerical result for $\text{Re}(u(t, x_0))$ and $\text{Im}(u(t, x_0))$ again our analytical solution for $d = 1, 2, 3$ as a function of $x_0 \in [0, 1.5]$ (all the coordinates in $\mathbb{R}^d$ are equal to x_0) and $t = 0.1$, see Figure 5. We obtain a perfect match. For completeness, we have also shown the (numerical) solution of the linear Schrödinger equation (i.e., $g := 0$), denoted "Linear".

Funding. This work benefits from the financial support of the ERC Advanced Grant 321111, and the Chairs *Financial Risk* and *Finance and Sustainable Development*.

REFERENCES

[1] AGARWAL, A. and CLAISSE, J. (2020). Branching diffusion representation of semi-linear elliptic PDEs and estimation using Monte Carlo method. *Stochastic Process. Appl.* **130** 5006–5036. MR4108480 https://doi.org/10.1016/j.spa.2020.02.009

[2] BAKHTIN, Y. and MUELLER, C. (2010). Solutions of semilinear wave equation via stochastic cascades. *Commun. Stoch. Anal.* **4** 425–431. MR2677199 https://doi.org/10.31390/cosa.4.3.07

[3] BALLY, V. and PAGÈS, G. (2003). Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic Process. Appl.* **106** 1–40. MR1983041

[4] BAO, W., JAKSCH, D. and MARKOWICH, P. A. (2003). Numerical solution of the Gross–Pitaevskii equation for Bose–Einstein condensation. *J. Comput. Phys.* **187** 318–342. MR1977789 https://doi.org/10.1016/S0021-9991(03)00102-5

[5] BOUCHARD, B., TAN, X., WARIN, X. and ZOU, Y. (2017). Numerical approximation of BSDEs using local polynomial drivers and branching processes. *Monte Carlo Methods Appl.* **23** 241–263. MR3745458 https://doi.org/10.1515/mcma-2017-0116

[6] BOUCHARD, B. and TOUZI, N. (2004). Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.* **111** 175–206. MR2056536 https://doi.org/10.1016/j.spa.2004.01.001

[7] CHATTERJEE, S. Stochastic solutions of the wave equation. Available at arXiv:1306.2382.

[8] DALANG, R. C., MUELLER, C. and TRIBE, R. (2008). A Feynman–Kac-type formula for the deterministic and stochastic wave equations and other P.D.E.'s. *Trans. Amer. Math. Soc.* **360** 4681–4703. MR2403701 https://doi.org/10.1090/S0002-9947-08-04351-1

[9] DEHGHAN, M. and SHOKRI, A. (2009). Numerical solution of the nonlinear Klein–Gordon equation using radial basis functions. *J. Comput. Appl. Math.* **230** 400–410. MR2532333 https://doi.org/10.1016/j.cam.2008.12.011

[10] EVANS, L. C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Amer. Math. Soc., Providence, RI. MR2597943 https://doi.org/10.1090/gsm/019

[11] FAHIM, A., TOUZI, N. and WARIN, X. (2011). A probabilistic numerical method for fully nonlinear parabolic PDEs. *Ann. Appl. Probab.* **21** 1322–1364. MR2857450 https://doi.org/10.1214/10-AAP723

[12] GUYON, J. and HENRY-LABORDÈRE, P. (2014). *Nonlinear Option Pricing*. *Chapman & Hall/CRC Financial Mathematics Series*. CRC Press, Boca Raton, FL. MR3155635

[13] HENRY-LABORDÈRE, P. (2012). Counterparty Risk Valuation: A Marked Branching Diffusion Approach. Risk magazine.

[14] HENRY-LABORDÈRE, P., OUDJANE, N., TAN, X., TOUZI, N. and WARIN, X. (2019). Branching diffusion representation of semilinear PDEs and Monte Carlo approximation. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 184–210. MR3901645 https://doi.org/10.1214/17-aihp880

[15] HENRY-LABORDÈRE, P., TAN, X. and TOUZI, N. (2014). A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Process. Appl.* **124** 1112–1140. MR3138609 https://doi.org/10.1016/j.spa.2013.10.005

[16] HENRY-LABORDÈRE, P., TAN, X. and TOUZI, N. (2017). Unbiased simulation of stochastic differential equations. *Ann. Appl. Probab.* **27** 3305–3341. MR3737926 https://doi.org/10.1214/17-AAP1281

[17] KAC, M. (1974). A stochastic model related to the telegrapher's equation. *Rocky Mountain J. Math.* **4** 497–509. MR0510166 https://doi.org/10.1216/RMJ-1974-4-3-497

[18] LE GALL, J.-F. (1999). *Spatial Branching Processes, Random Snakes and Partial Differential Equations. Lectures in Mathematics ETH Zürich*. Birkhäuser, Basel. MR1714707 https://doi.org/10.1007/978-3-0348-8683-3

[19] LE JAN, Y. and SZNITMAN, A. S. (1997). Stochastic cascades and 3-dimensional Navier–Stokes equations. *Probab. Theory Related Fields* **109** 343–366. MR1481125 https://doi.org/10.1007/s004400050135

[20] MCKEAN, H. P. (1975). Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piskunov. *Comm. Pure Appl. Math.* **28** 323–331. MR0400428 https://doi.org/10.1002/cpa.3160280302

[21] RASULOV, A., RAIMOVA, G. and MASCAGNI, M. (2010). Monte Carlo solution of Cauchy problem for a nonlinear parabolic equation. *Math. Comput. Simulation* **80** 1118–1123. MR2610073 https://doi.org/10.1016/j.matcom.2009.12.009

[22] ZHANG, J. A. (2004). Numerical scheme for backward stochastic differential equations: Approximation by step processes. *Ann. Appl. Probab.* **14** 459–488.